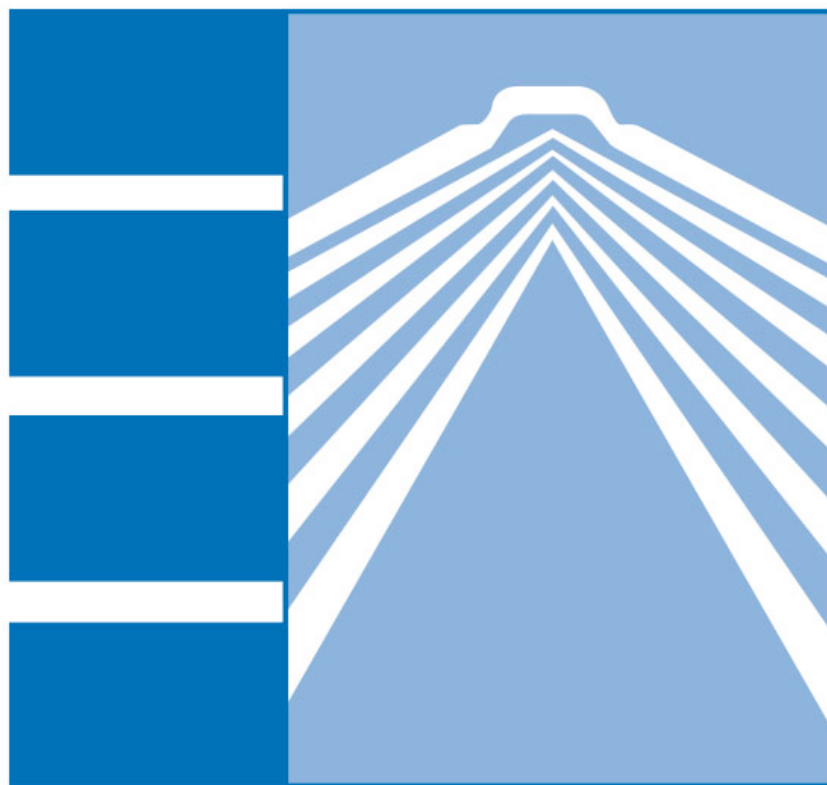
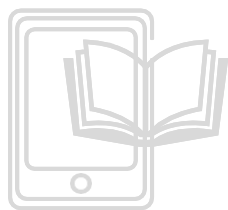


Curso de GENÉTICA MOLECULAR E INGENIERÍA GENÉTICA



Marta Izquierdo Rojo

PIRÁMIDE



BASE LEGAL PARA DIGITALIZACIÓN DE LIBROS CON DERECHOS DE AUTOR

El Sistema Integrado de Bibliotecas de la Universidad Central del Ecuador, digitalizará su acervo bibliográfico siempre y cuando sea para fines educativos y de investigación. No se permite la reproducción y distribución para la comercialización directa e indirecta del mismo.

La digitalización del material bibliográfico se lo realiza de acuerdo al Código Orgánico De La Economía Social de los Conocimientos, Creatividad e Innovación Art. 212 numeral 9, Literal ii “(...) Una biblioteca o archivo podrá, además, realizar los siguientes actos (...) La reproducción electrónica y comunicación pública de obras de su colección para ser consultadas gratuita y simultáneamente hasta por un número razonable de usuarios, sólo en terminales de redes de la respectiva institución o para usuarios de esa institución bajo su control, en condiciones que garanticen que no se puedan hacer copias electrónicas de esas reproducciones” y literal vii “La reproducción, adaptación, traducción, transformación, arreglo, distribución y comunicación de una obra protegida por derechos de autor o una prestación protegida por derechos conexos, en uno o más formatos accesibles para el uso exclusivo de personas con discapacidad”.

Este material se considera un producto intelectual a favor de su autor; por tanto, la titularidad de sus derechos se encuentra protegida por el Código Orgánico De La Economía Social de los Conocimientos, Creatividad e Innovación. La violación a dichos derechos constituye un delito que será responsabilidad del usuario.

Este libro está disponible físicamente en:

BIBLIOTECA		PISO	
		ESTANTE	
		BANDEJA	
CONVENIO INTERINSTITUCIONAL		UNIVERSIDAD	

Curso de
GENÉTICA MOLECULAR
E INGENIERÍA GENÉTICA

MARTA IZQUIERDO ROJO

CATEDRÁTICA DE BIOLOGÍA MOLECULAR
DE LA UNIVERSIDAD AUTÓNOMA DE MADRID

Curso de
**GENÉTICA MOLECULAR
E INGENIERÍA GENÉTICA**

EDICIONES PIRÁMIDE

Edición en versión digital

Está prohibida la reproducción total o parcial de este libro electrónico, su transmisión, su descarga, su descompilación, su tratamiento informático, su almacenamiento o introducción en cualquier sistema de repositorio y recuperación, en cualquier forma o por cualquier medio, ya sea electrónico, mecánico, conocido o por inventar, sin el permiso expreso escrito de los titulares del copyright.

© Marta Izquierdo Rojo, 2014

© Primera edición electrónica publicada por Ediciones Pirámide (Grupo Anaya, S. A.), 2014

Para cualquier información pueden dirigirse a piramide_legal@anaya.es

Juan Ignacio Luca de Tena, 15. 28027 Madrid

Teléfono: 91 393 89 89

www.edicionespiramide.es

ISBN digital: 978-84-368-3124-5

Índice

Prólogo	11
Abreviaturas	13
1. Conceptos básicos sobre el gen, el genoma y su expresión	17
1.1. La información genética	17
1.2. Estructura de los ácidos nucleicos	19
1.2.1. Estructura del ADN	21
1.2.2. Estructura y función de los ARN	21
1.3. Curvas de reasociación: grados de repetición en las secuencias	24
1.4. Niveles de empaquetamiento del ADN	28
1.5. Expresión de la información genética y su regulación	31
1.5.1. Procesamiento de los mensajeros	35
1.5.2. Traducción	35
1.6. Variaciones en la cantidad de ADN y número de genes a lo largo de la escala evolutiva	39
1.7. Genética y epigenética en relación con la variabilidad entre células, individuos y especies	42
1.8. Distribución de tareas y porcentajes en el genoma humano: metodología empleada en su determinación	45
Lecturas complementarias	48
2. Recombinación	49
2.1. Recombinación entre cromosomas homólogos	49
2.1.1. La recombinación homóloga en procariotas	50
2.1.2. La recombinación homóloga en eucariotas	57
2.2. Recombinación entre sitios concretos	61
2.2.1. La recombinación específica de sitio del fago λ	61
2.2.2. Otras recombinaciones específicas de sitio	63
2.3. Transposición	69
2.3.1. Transposición vía ADN	69
2.3.2. Transposición vía ARN: retrotransposones	77
2.3.3. Efectos celulares y control de la transposición	79
Lecturas complementarias	83

3. Enzimas y técnicas de utilización general en manipulación genética.....	85
3.1. Degradación y síntesis de ácidos nucleicos <i>in vitro</i>	85
3.1.1. Degradación química	85
3.1.2. Degradación enzimática	88
3.1.3. Síntesis de ácidos nucleicos.....	90
3.2. Enzimas de restricción	96
3.2.1. Enzimas de restricción de tipo II.....	98
3.2.2. Enzimas de restricción de otros tipos	102
3.3. Genomas tratados con enzimas de restricción	103
3.3.1. Polimorfismos de fragmentos de restricción	107
3.3.2. La huella genética individual.....	109
3.4. Hibridación de ácidos nucleicos en soporte sólido ordenado en matrices.	112
3.5. Secuenciación del ADN: distintos métodos y aproximaciones.....	114
3.5.1. Método químico.....	114
3.5.2. Método enzimático.....	114
3.5.3. Secuenciación masiva o de nueva generación	116
Lecturas complementarias	122
 4. Métodos de amplificación de secuencias: ADN recombinante y reacción en cadena de la polimerasa	123
4.1. Elementos del ADN recombinante	124
4.2. Elementos de la clonación molecular en bacterias	124
4.2.1. Plásmidos como vectores.....	124
4.2.2. El bacteriófago λ como vector.....	127
4.2.3. Otros vectores para la obtención de ADN recombinante en bacterias.....	131
4.3. Tipos de clonación molecular y aplicaciones: genotecas	137
4.3.1. Elaboración de una genoteca de ADN.....	137
4.3.2. Elaboración de una genoteca de ADNc	138
4.3.3. Identificación del recombinante deseado de una genoteca	142
4.3.4. Aplicaciones de las genotecas de ADN: paseo cromosómico	142
4.3.5. Elección del vector más idóneo para el tipo de clonación	146
4.3.6. Expresión de los productos clonados en bacterias.....	146
4.4. Amplificación por la reacción en cadena de la polimerasa (PCR)	151
4.4.1. Variaciones concretas de la técnica básica	153
4.5. Algunas aplicaciones de la PCR	161
 5. Mutaciones, remodelación de genomas y cáncer.....	167
5.1. Tipos de mutaciones	167
5.2. Mutaciones espontáneas: tipos y frecuencia.....	172
5.3. Mutaciones provocadas: agentes mutagénicos	175
5.4. Mutagénesis al azar y selección específica.....	177

5.5. Mutagénesis dirigida	178
5.5.1. Estrategia de provocar deleciones en secuencias clonadas	178
5.5.2. Estrategia de mutaciones puntuales en secuencias clonadas vía M13	180
5.5.3. Obtención de mutaciones puntuales por PCR.....	180
5.6. Mutagénesis <i>in vivo</i>	181
5.7. Mutagénesis controlada a nivel de genomas completos: bacterias sintéticas.....	187
5.8. Mutaciones y cáncer	188
Lecturas complementarias	191
6. Transferencia genética en levaduras e insectos	193
6.1. Clonación en levaduras: <i>Saccharomyces cerevisiae</i>	193
6.2. Métodos de transformación y selección en levaduras	194
6.3. Tipos de vectores en levaduras.....	195
6.3.1. Vectores de integración cromosómica.....	196
6.3.2. Vectores de replicación autónoma	198
6.4. Traducción y modificaciones posteriores como glicosilación y secreción de productos	204
6.5. Otras levaduras de interés	205
6.6. Transferencia génica en insectos como sistema de expresión alternativo al de levaduras: los baculovirus como vectores	206
7. Transferencia génica a células de mamíferos. Terapia génica	209
7.1. Tipos y métodos de transfección.....	209
7.2. Métodos de selección y genes señal para transformaciones estables y transitorias.....	211
7.2.1. Genes de selección estable	212
7.2.2. Genes señal para transferencias estables.....	213
7.2.3. Selección transitoria	214
7.3. ADN desnudo y vectores	214
7.3.1. SV40 y vectores mixtos derivados	215
7.3.3. Retrovirus y lentivirus	218
7.3.4. Adenovirus	222
7.3.5. Adenoasociados	224
7.3.6. Vectores derivados del <i>Herpes</i>	226
7.3.7. El virus de la vacuna	227
7.3.8. Cromosomas artificiales de mamíferos.....	230
7.4. Terapia génica	231
7.4.1. Terapia génica de enfermedades monogénicas.....	233
7.4.2. Terapia génica de enfermedades multigénicas: cáncer	237
7.4.3. Terapia génica por cirugía genómica	241
7.4.4. Terapia génica, consideraciones finales	241
Lecturas complementarias	242

8. Individuos transgénicos y clónicos, clonación terapéutica	243
8.1. Métodos de obtención de animales transgénicos	243
8.1.1. Microinyección de ADN a huevos fecundados.....	244
8.1.2. Obtención de transgénicos por introducción de ADN a embrio- blastos del blastocisto.....	248
8.2. Expresión del transgén y aplicaciones de animales transgénicos.....	251
8.3. Individuos clónicos por trasplante de núcleos.....	255
8.4. Animales clónico-transgénicos de interés farmacológico y biomédico	261
8.5. Clonación terapéutica	264
Lecturas complementarias	267
9. Transferencia génica a plantas	269
9.1. Vectores.....	270
9.1.1. Plásmidos	270
9.1.2. Virus como vectores	273
9.2. Métodos de cultivo.....	278
9.3. Métodos de transferencia del ADN a la planta	279
9.4. Expresión del transgén.....	282
9.5. Plantas transgénicas en la agricultura	283
Lecturas complementarias	296
Índice de materias	297

Prólogo

Este libro se basa en otro anterior de título *Ingeniería genética y transferencia génica*. Tanto aquél como éste abordan los principios de manipulación genética, clonaje y transferencia de genes en distintos sistemas: bacteriano, de levaduras, animal y vegetal. Describen asimismo las técnicas de producción de animales transgénicos y clónicos, analizando las aplicaciones y problemas que de todo ello se pueden derivar. También analizan la posibilidad de utilizar plantas transgénicas como fuente de antibióticos, vacunas o plásticos biodegradables. En este *Curso de Genética Molecular e Ingeniería Genética* se revisan y actualizan los conceptos, técnicas y aplicaciones de la manipulación genética y la transferencia de este material entre organismos dispares. Las novedades son numerosas de principio a fin. Asimismo, la posibilidad de obtener el texto con las ilustraciones a todo color en la versión digital permite clarificar y amenizar las ilustraciones, complejas en algunos casos. El conocimiento de que aproximadamente el 80% del genoma está implicado de modo directo o indirecto en la regulación de un 3% que representan los genes que se traducen a proteínas es novedad; como el 17% restante se puede asignar a restos de retrovirus, transposones o secuencias cuya función se está investigando (Piwi, por ejemplo), no queda mucho sitio para el ADN «basura» que, hasta no hace mucho, representaba la mayor parte de los genomas de mamíferos. Un capítulo dedicado a la recombinación es también novedad; este tema tan básico e interesante no siempre está presente como tal en libros generales de biología molecular e ingeniería gené-

tica y, sin embargo, toda la manipulación genética, transposición o la integración de virus es pura recombinación. La mutagénesis controlada a nivel de genomas completos con la producción de bacterias totalmente sintéticas nos ha llegado sin previo aviso. La terapia génica está recogiendo sus primeros éxitos como opción terapéutica en algunas enfermedades monogénicas y la cirugía genómica nos llama a la puerta, muy decidida a entrar. Los animales transgénicos, clónicos y clónico-transgénicos avanzan a buen ritmo hacia las granjas especializadas para la producción de carne, leche modificada u órganos para trasplantes. La terapia celular y la clonación terapéutica son la gran promesa de la medicina regenerativa, aunque aún quedan algunos problemas por resolver. La transferencia génica en plantas es responsable de los cultivos transgénicos que tanto debate ha generado y aún genera, pero las plantas transgénicas con maduración controlada de flores y frutos, semillas híbridas a gran escala o resistentes a herbicidas están entre nosotros para quedarse porque, aunque no resuelvan todos los problemas de la alimentación humana, pueden paliar alguno de ellos.

Es mi deseo que este libro represente una ayuda y un estímulo para el estudiante, el estudioso y el curioso.

Quisiera expresar mi más profunda gratitud a **José Antonio Pérez** por muchas de las ilustraciones de este libro. Siempre es un placer dar las gracias a los colegas que han revisado generosamente algunos de los capítulos de esta obra aportando numerosas y acertadas sugerencias. Estoy espe-

cialmente agradecida a **Fernando Carrasco** por todo lo referente a secuenciación y a **Antonio Jiménez** por la exhaustiva y magnífica revisión del capítulo de levaduras. Finalmente estoy en deuda permanente con **Juan Antonio García Álvarez** que no sólo contribuye desinteresadamente en la docencia de la UAM todos los años, sino que ha revisado y aportado mejoras sustanciales en el capítulo de transferencia génica a plantas. Mi agra-

decimiento a la Universidad Autónoma de Madrid, donde ejerzo la docencia desde hace muchos años, a mis colegas del Departamento de Biología Molecular y a los alumnos, a quienes dedico este libro.

Madrid, febrero de 2014.

MARTA IZQUIERDO ROJO

Abreviaturas

ADA	Enzima deaminasa de adenosinas.
ADN	Ácido desoxirribonucleico.
ADNasa	Desoxirribonucleasa.
ADNpol	Polimerasas de ADN.
AMP	Adenina monofosfato.
Amp	Ampicilina.
ARN	Ácido ribonucleico.
ARNasa	Ribonucleasa.
ARNi	Ácido ribonucleico de interferencia.
ARNm	Ácido ribonucleico mensajero.
ARNnc	Ácido ribonucleico no codificante.
ARN-pi	ARN de interferencia asociado a la proteína nuclear Piwi.
miARN	Micro-ARN.
ARNr	Ácido ribonucleico ribosomal.
ARNt	Ácido ribonucleico de transferencia.
BrdU	Bromodesoxiuridina.
CAB	Cromosoma artificial de bacteria.
CAL	Cromosoma artificial de levadura.
CAM	Cromosoma artificial de mamífero.
CaMV	Versión anglosajona del virus del mosaico de la coliflor, <i>Cauliflower mosaic virus</i> .
cat	Gen correspondiente a la enzima cloranfenicol acetiltransferasa.
cis	Actuación sobre la misma unidad física de ADN o cromosoma.
CEN	Centrómero.
CRISPR	Del inglés: <i>clustered regularly interspaced short palindromic repeat</i> . Región cromosómica rica en elementos repetidos.
CSC	Células troncales tumorales, del inglés <i>cancer stem cells</i> .
DHF	Dihidrofolato.
dhfr	Gen que codifica la enzima dehidrofolato reductasa.
DIG	Digoxigenina; hapteno esteroideo.
EDTA	Etilen-di-amin-tetra-ácido-acético, agente quelante de iones divalentes.
EI	Elemento de iniciación del ARN (levaduras).
EPO	Eritropoyetina.
FACS	Del inglés: <i>fluorescence-activated cell sorting</i> .
gfp	Gen que codifica la proteína verde fluorescente; del inglés, <i>green fluorescent protein</i> .
GGTA1	Gen que codifica la enzima α 1, 3-galactosil-transferasa (α 1,3GT).
GMP	Guanina monofosfato.

<i>gus</i>	Gen correspondiente a la enzima β -glucuronidasa.
HAT	Hipoxantina, aminopterina y timina.
<i>hc</i>	Gen correspondiente a la hormona de crecimiento.
Hfr F ⁺	Alta frecuencia de recombinación, del inglés: <i>high frequency of recombination</i> , portador del factor F de fertilidad.
<i>hgro</i>	Gen que codifica la quinasa de higromicina.
HUVEC	Células endoteliales de cordón umbilical humanas.
<i>IRES</i>	Del inglés, <i>internal ribosome entry site</i> , lugar de entrada de los ribosomas.
iPS	Células troncales pluripotentes inducidas, del inglés: <i>induced pluripotent stems</i> .
<i>i-tetrVP16</i>	Gen fusión del represor de la tetraciclina con el activador <i>VP16</i> .
<i>tkHs</i>	Gen correspondiente a la enzima quinasa de timidinas del virus <i>Herpes simplex</i> .
HXGPRT	Hipoxantina-guanina-fosforribosil transferasa.
Ig	Inmunoglobulina.
IMP	Inosina monofosfato.
IRES	Secuencia de iniciación interna para ribosomas.
kb	1.000 bases para sistemas monocatenarios (ARN) y 1.000 pares de bases para sistemas bicatenarios (ADN).
kD	Kilodaltons, unidad de tamaño de proteínas equivalente a 1.000 Daltons.
KO	Del inglés: <i>knock out</i> , dejar fuera de combate, eliminar.
<i>lac</i>	Gen correspondiente a la enzima β -galactosidasa.
LDM	Leucodistrofia metacromática.
<i>leu</i>	Gen que codifica por la enzima β -isopriomalato dehidrogenasa, indispensable en la biosíntesis de leucina.
LHh	Lactoferrina humana.
<i>luc</i>	Gen que codifica por la luciferasa.
mgb	Megabase (1.000 kb).
MLV	γ -retrovirus, del inglés: <i>morine leukaemia virus</i> .
MSC	Células troncales mesenquimales humanas.
MTX	Metotrexato.
ng	Nanogramo, 10 ⁻⁹ gramos.
<i>neo</i>	Gen que codifica la enzima neomicina-acetiltransferasa. Confiere resistencia a los antibióticos: neomicina, kanamicina y G418.
ORF	Del inglés: <i>Open Reading Frame</i> , pauta de lectura abierta sin codones de terminación.
<i>paf</i>	Gen correspondiente a la proteína anticongelante de la faneca.
PAM	Plásmido artificial de mamíferos.
pb	Pares de bases.
pBR	Plásmido de Bolívar y Rodríguez.
pCL	Plásmido centromérico de levaduras.
PCR	Reacción en cadena de la polimerasa (versión anglosajona).
PEG	Polietilen-glicol.
pg	Picogramo, 10 ⁻¹² gramos.
PIT	Región posterior a la iniciación de la transcripción.
PM	Proteína de movimiento.
pRL	Plásmido de replicación en levaduras.
PTF	Polimorfismo en el tamaño de fragmentos de restricción.
<i>puro</i>	Gen que codifica la enzima puromicín-acetiltransferasa, confiere resistencia al antibiótico puromicina.
pVV	Promotor del virus de la vacuna.
RCP	Reacción en cadena de la polimerasa (versión en castellano).
S	Unidad Svedberg para medir el coeficiente de sedimentación de una partícula sometida a centrifugación. Las unidades S no son aditivas.

SA	Secuencia activadora en levaduras.
SCID	Síndrome de la inmunodeficiencia combinada severa.
SD	Secuencia Shine y Dalgarno de iniciación de la transcripción en bacterias.
Sf	<i>Spodoptera frugiperda</i> .
SIDA	Síndrome de inmunodeficiencia adquirida.
SIE	Secuencia de iniciación de empalme.
SIT	Secuencia de iniciación de la transcripción.
SRA	Secuencia de replicación autónoma.
ssb	Proteínas de unión a ADN de hebra sencilla, del inglés: <i>single stranded binding</i> .
STE	Secuencia de terminación de empalme.
SUR	Secuencia de unión al ribosoma en bacterias.
SV40	Virus de los simios 40.
TALEN	Del inglés: <i>transcription activator-like effector nucleases</i> .
<i>tat</i>	Gen tetraciclina-acetiltransferasa.
TEL	Telómero.
THF	Tetrahidrofolato.
<i>tk</i>	Gen correspondiente a la enzima quinasa de timidinas.
TMN, TMF	Timidina monofosfato.
TMV	Virus del mosaico del tabaco, versión anglosajona.
Tn	Transposón.
<i>trans</i>	Actuación sobre una unidad física distinta de ADN o cromosoma.
TRI	Terminal repetido invertido.
TRL	Terminal repetido largo.
TRP	Triptófano.
UAS	Del inglés: <i>upstream activating sequences</i> (secuencias de activación corriente arriba).
URA	Orotidín 5 difosfato decarboxilasa, precursor de uridín monofosfato.
VAA	Vector adenoasociado.
VHS	Virus <i>Herpes simplex</i> .
VMP	Virus del mosaico del pepino.
VP16	Proteína activadora del virus <i>Herpes simplex</i> .
VPNAc	Virus de la poliedrosis nuclear de <i>Autographa californica</i> .
VTMR	Virus del tumor de mama de ratón.
WAS	Síndrome de Wiskott-Aldrich.
XMP	Xantina monofosfato.
XGPRT	Xantina-guanina-fosforribosil transferasa.
YAC	Del inglés: <i>yeast artificial chromosome</i> (cromosoma artificial de levadura).

1

Conceptos básicos sobre el gen, el genoma y su expresión

1.1. LA INFORMACIÓN GENÉTICA

El soporte físico de la información genética es la molécula de ácido desoxirribonucleico (ADN). Así, el ADN almacena y transmite a sus descendientes las instrucciones precisas para producir nuevos seres a su imagen y semejanza. Toda célula recibe una dotación génica (de material hereditario) de su progenitora (o progenitores), y con ella debe ser capaz de producir un gran número de proteínas y realizar numerosas funciones. El **genoma** es el conjunto de todo el material genético de un individuo o de una especie. Debe permitir que un ser humano se desarrolle, crezca y se maneje, física e intelectualmente, durante las distintas etapas de una vida. El genoma de una especie es como el producto final de un texto que generaciones de individuos han contribuido a corregir, modificar y optimizar para transmitir todos y cada uno de los procesos en él descritos. Durante mucho tiempo se pensó, erróneamente, que la naturaleza del material genético era tan compleja como la de sus productos, y hasta 1944 se aceptó la naturaleza proteica del material genético. Sólo las proteínas parecían tener la suficiente versatilidad para dar origen al abanico de distintas funciones y moléculas observables en la célula. Hasta los experimentos pioneros de Frederick Griffith en 1928, los espectaculares de Oswald Avery, Colin McLeod y Madyn McCarty y su sustancia transformante en 1944 y, finalmente, los de Alfred Hershey y Martha Chase en 1952 no queda totalmente demostrado que el ADN es realmente el material genético. Pronto se puso de manifiesto

que la versatilidad no radicaba en la naturaleza de la molécula (que es monótona y sólo tiene cuatro bases), sino en la información que de ella emanaba al tratar como una unidad cada conjunto de tres pares de bases.

Considerando que la esencia de la vida es realmente la experiencia acumulada a través de generaciones, la biología trata fundamentalmente de descubrir cómo los seres vivos son capaces de transmitir ese saber adquirido de una forma perdurable, estable y a la vez flexible. Mirando una simple bacteria en un medio de crecimiento adecuado, se puede ver cómo asimila los nutrientes del entorno, se alarga y se divide en dos. Las dos bacterias hijas hacen exactamente lo mismo. Ocasionalmente, la réplica sale un poco distinta y transmite la diferencia a sus descendientes. La perpetuación viene definida por el proceso de la replicación. En la replicación, el ADN paterno se copia de modo semiconservativo, originando dos moléculas hijas, cada una constituida por una hebra paterna y la otra recién sintetizada. Tanto si la célula tiene un cromosoma como si tiene muchos, todo el ADN se replicará antes de la división celular. Si el genoma es pequeño y circular, como el de procariotes, tendrá únicamente un origen de replicación, y si, por el contrario, el genoma es muy grande y hay varios cromosomas, tendrá muchos orígenes de replicación simultáneos (figura 1.1).

El mecanismo de la replicación es enormemente complejo y en él participan numerosas enzimas. De entre ellas, la ADN polimerasa es capaz de sintetizar directamente una nueva banda de

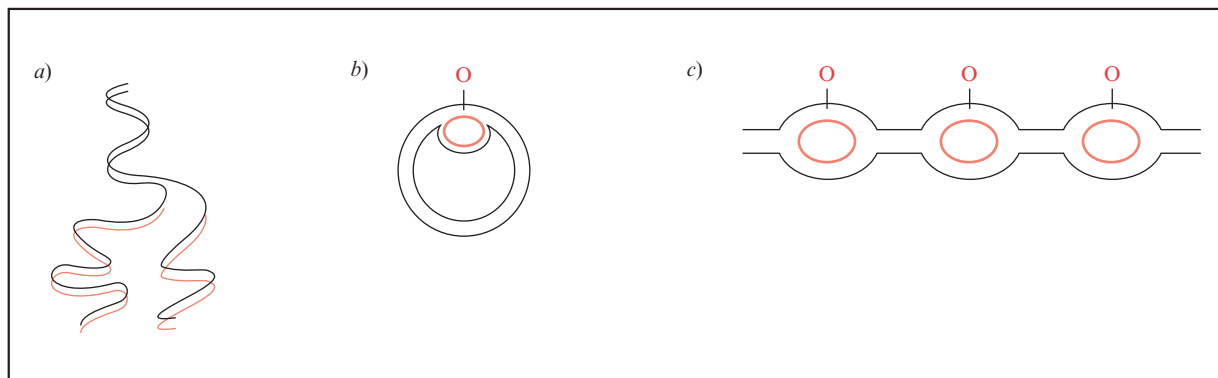


Figura 1.1. Replicación del ADN. a) Replicación del ADN de modo semiconservativo. Cada molécula hija llevará una hebra parental y la otra de nueva síntesis. b) Replicación de un plásmido con un único origen de replicación. c) Replicación de un ADN eucariótico con muchos orígenes de replicación.

ADN a partir de un molde o patrón, pero hay muchas otras enzimas que, a su vez, son esenciales para una correcta replicación. Tanto procariotas como eucariotas poseen varias ADN polimerasas, pero sólo una de ellas realiza la función de copia; las demás están implicadas en funciones de reparación o reemplazamiento de fragmentos cortos en el ADN.

Las modificaciones de la información que se escapan a los mecanismos de reparación constituyen las mutaciones. Una mutación puede representar, por ejemplo, el cambio de una base (mutaciones puntuales) o la eliminación de un fragmento más o menos grande de ADN (deleciones). En todos los casos se producirá una alteración estable y hereditaria en los lugares afectados. Todos los organismos sufren un cierto número de mutaciones espontáneas como resultado de sucesos incontrolables e inevitables. Las mutaciones pueden, no obstante, inducirse con agentes que modifican determinadas bases en el ADN o producen rupturas múltiples en la cadena doble.

La definición de **gen** ha variado a lo largo de los años; actualmente el gen se considera una unidad de información estable y hereditaria suficiente y necesaria para realizar al menos una función. Una definición moderna sería la de *un fragmento de ADN portador de la unidad mínima de significación genética, que se transcribe a ARN*.

La mayoría de los genes eucarióticos que llegan a traducirse poseen dos tipos de secuencias: los intrones, que se transcriben pero son eliminados antes de llegar al citoplasma, y los exones, que formarán parte de las proteínas. Los genes de bacterias carecen de intrones mientras que los genes de organismos más evolucionados poseen un número variable de intrones, y la definición de gen abarcaría tanto a exones como a intrones de la misma unidad de significación genética. Una definición más amplia que la de gen es la de elemento funcional, que engloba también el ADN implicado en regulación y en adquisición de estructuras relajadas o compactas. **El elemento funcional define un fragmento discreto de ADN que codifica por un producto concreto** (como una proteína o un ARN no codificante) **o muestra una configuración estructural reproducible y característica** (como la distendida, que permite la unión a proteínas, o la muy plegada y condensada de heterocromatina).

La transcripción es el proceso por el cual se obtiene una molécula de ARN a partir de un fragmento de ADN y **la genómica estudia todos los genes que se transcriben en una célula, tejido u organismo**. De manera análoga, la proteómica estudia las proteínas totales de una célula, tejido u organismo. **El exoma** se refiere al conjunto de exones de un genoma.

1.2. ESTRUCTURA DE LOS ÁCIDOS NUCLEICOS

La información genética está contenida en la propia molécula de ADN. El código o la clave radica en el orden en el que se presentan las cuatro bases: adenina, citosina, guanina y timina. La célula interpreta el código y elabora los productos necesarios para su ciclo vital. La información se transcribe en la célula a ácido ribonucleico mensajero de una sola hebra (ARNm), cuya lectura se va a traducir en proteínas concretas, que serán los peones encargados de ejecutar las órdenes. Las proteínas, a su vez, pueden regular la expresión del ADN, con lo que se completa el ciclo. Así pues, la información genética pasa ma-

yoritariamente de una célula a otra a través del ADN (ácido desoxirribonucleico, de hebra doble) y del núcleo al citoplasma a través de los ARNm (ácidos ribonucleicos mensajeros, de una sola hebra).

Los ácidos nucleicos son polinucleótidos cuyas unidades, los nucleótidos, constan de una base nitrogenada, una pentosa y un grupo fosfato. El tipo de pentosa determina que el ácido nucleico sea ADN o ARN. En el ADN, la pentosa es la 2-desoxirribosa, mientras que en el ARN es la ribosa (figura 1.2). La ausencia de un grupo hidroxilo en el carbono 2' de la pentosa da al ADN una gran resistencia a la hidrólisis por álcali, contrariamente al ARN, que a pH alcalino y 37 °C se rompe en nucleótidos en pocos minutos.

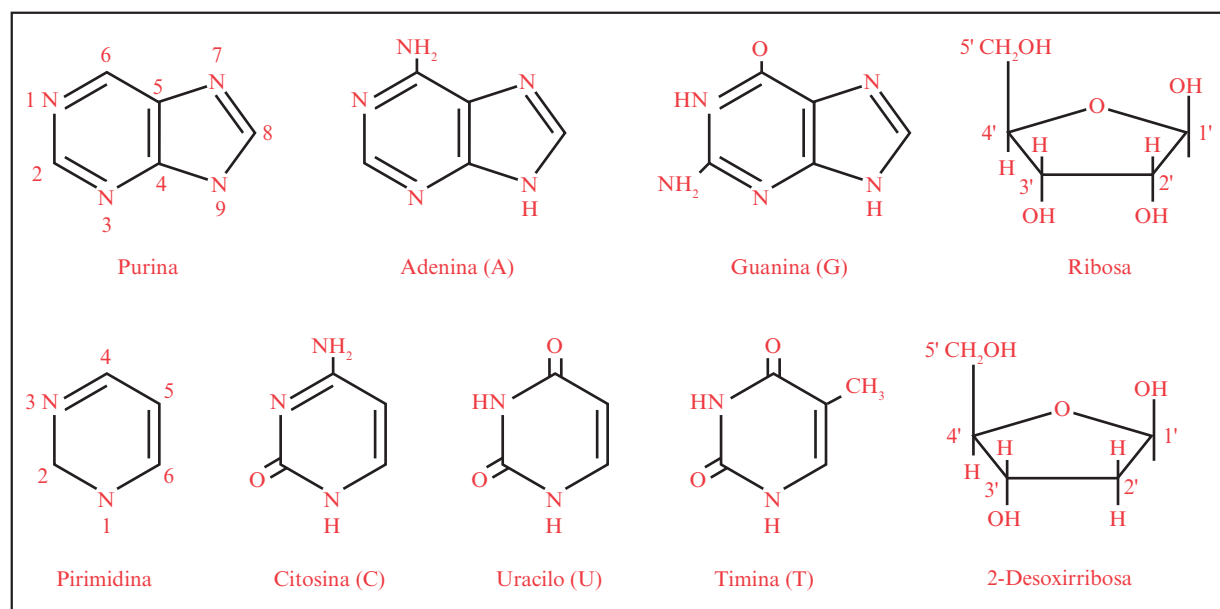


Figura 1.2. Componentes de los ácidos nucleicos, composición química y numeración.

Las bases nitrogenadas corresponden a dos categorías: purinas y pirimidinas. Las dos purinas adenina (A) y guanina (G) están presentes tanto en el ADN como en el ARN. En cuanto a las pirimidinas, el ADN posee citosina (C) y timina (T), mientras que en el ARN la timina está sustituida

por uracilo (U). El uracilo difiere de la timina en un grupo metilo en el carbono 5 (figura 1.2). Cabría preguntarse por qué el ADN tiene una base distinta al ARN; la respuesta puede tener mucho que ver con la función que presta el ADN como molécula depositaria de la información genética.

La citosina se desamina espontáneamente a uracilo con cierta frecuencia, y si en el ADN hubiera uracilo, sería prácticamente imposible discernir entre los originados por desaminación de la citosina y los propios de la molécula del ADN. Esto conllevaría una gran tasa de error difícilmente compatible con la función primordial del ADN de transmitir fielmente la información genética de generación en generación. Si el ADN no posee normalmente uracilo, todos los uracilos que se encuentran en el ADN deben provenir de la desaminación de la citosina y deben ser eliminados. De ello se encarga la enzima uracil N-glicosilasa que se pasea por el ADN y se para en cada timina y en cada uracilo; si reconoce el grupo —CH_3 de la timina pasa de largo; pero en su ausencia corta la base y los sistemas de reparación celulares colocan la complementaria a la de la cadena opuesta.

Sin embargo, muchas citosinas en el ADN se encuentran metiladas y la desaminación oxidativa de esta base así modificada origina timina. Timinas apareadas con guaninas también son reconocidas y eliminadas por un sistema mucho más lento e ineficiente que el anteriormente descrito.

Las bases son estructuras planas que, cuando se unen a la posición 1' de las pentosas por un enlace N-glicosílico, constituyen los nucleósidos. Cuando un grupo fosfato se une a un nucleósido se forma el nucleótido correspondiente. El lugar más común para la esterificación es el grupo hidroxilo del carbono 5' de la ribosa o la desoxirribosa. En esta posición se pueden añadir de uno a tres grupos fosfato creando los mono, di o trinucleótidos; por ejemplo: adenina monofosfato (AMP), adenina difosfato (ADP) o adenina trifosfato (ATP) (figura 1.3).

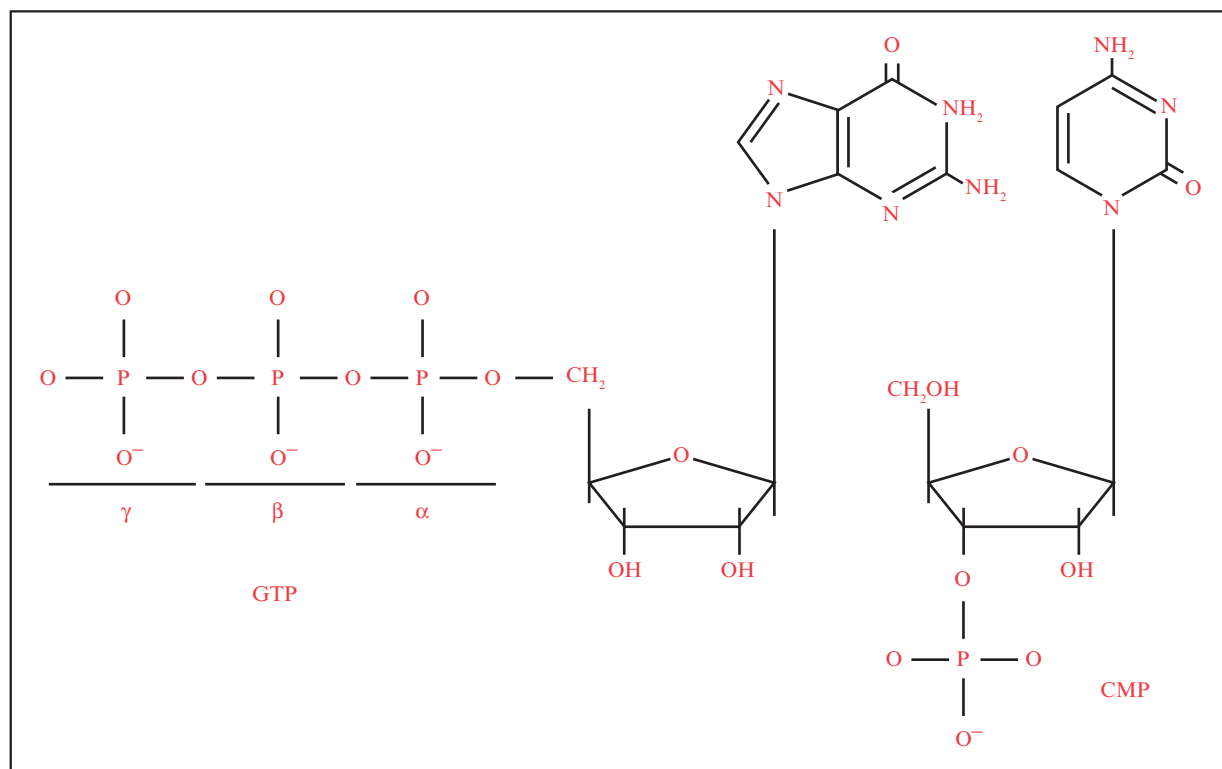


Figura 1.3. A la izquierda, un trifosfato 5' de guanosina donde se muestran los fosfatos α (alfa), β (beta) y γ (gamma); a la derecha, un monofosfato 3' de citosina.

El primer grupo fosfato (α) situado en la posición 5' está unido por un enlace de tipo éster, pero los otros dos grupos fosfato (β y γ) lo están por enlaces fosfo-anhídrido altamente energéticos que liberan una gran cantidad de energía al hidrolizarse. El grupo fosfato puede estar unido también a la posición 3' de la pentosa, como muestra la figura 1.3.

1.2.1. Estructura del ADN

El ADN es una doble hélice en la que las bases se encuentran en el interior de la molécula (figura

1.4) y los grupos fosfato en el exterior. La adenina siempre se aparea con la timina, y la guanina con la citosina. La unión entre las bases complementarias se realiza a través de puentes de hidrógeno: dos entre timinas y adeninas, y tres entre guaninas y citosinas.

La síntesis del ADN se produce siempre en la dirección 5' fosfato $\rightarrow$ 3' OH, ya que sólo existen ADN polimerasas capaces de añadir nucleótidos a extremos 3'OH libres. Las dos hebras que forman la hélice se definen por sus extremos: 5' fosfato y 3'OH. Estas hebras tienen orientaciones opuestas: una va en dirección 5' $\rightarrow$ 3' y su complementaria en la 3' $\rightarrow$ 5' (figura 1.5). La ruptura de los puentes de hidrógeno por calor, álcali o com-

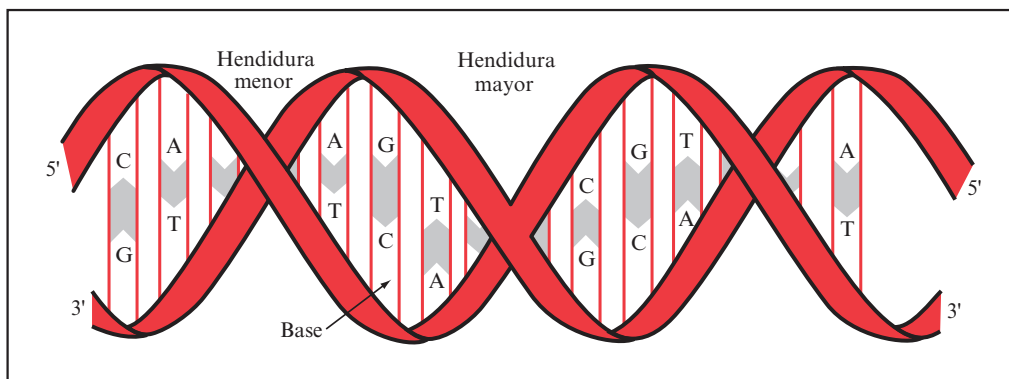


Figura 1.4. La molécula de ADN mostrando la hebra doble, el apareamiento de las bases y los extremos 5' fosfato y 3'OH.

puestos químicos produce la separación física de las dos hebras del ADN y se denomina desnaturalización. La desnaturalización por calor es total a los 90 °C, y por álcali, a pH superiores a 11,3. En ambos casos el proceso es reversible, y al desaparecer el agente desnaturalizante la molécula renaturaliza, esto es, vuelve a adquirir la estructura doble helicoidal original.

1.2.2. Estructura y función de los ARN

Las diferencias estructurales entre ARN y ADN radican fundamentalmente en: 1) el azúcar

de los ARN es ribosa; 2) la base uracilo sustituye a la timina en todos los ARN (figura 1.2), y 3) los ARN son en su mayoría de hebra sencilla, aunque pueden adquirir estructuras secundarias con diversos grados de plegamientos y complementariedad. Está generalmente aceptado que el ARN precedió evolutivamente al ADN.

En la célula hay muchos tipos de ARN, y el tamaño de los mismos es una característica que suele emplearse para diferenciarlos. Así hablamos de los ARN largos nucleares precursores de los ARN mensajeros (ARNm), que son los verdaderos transportadores de la información contenida en muchos genes del núcleo al citoplasma, donde

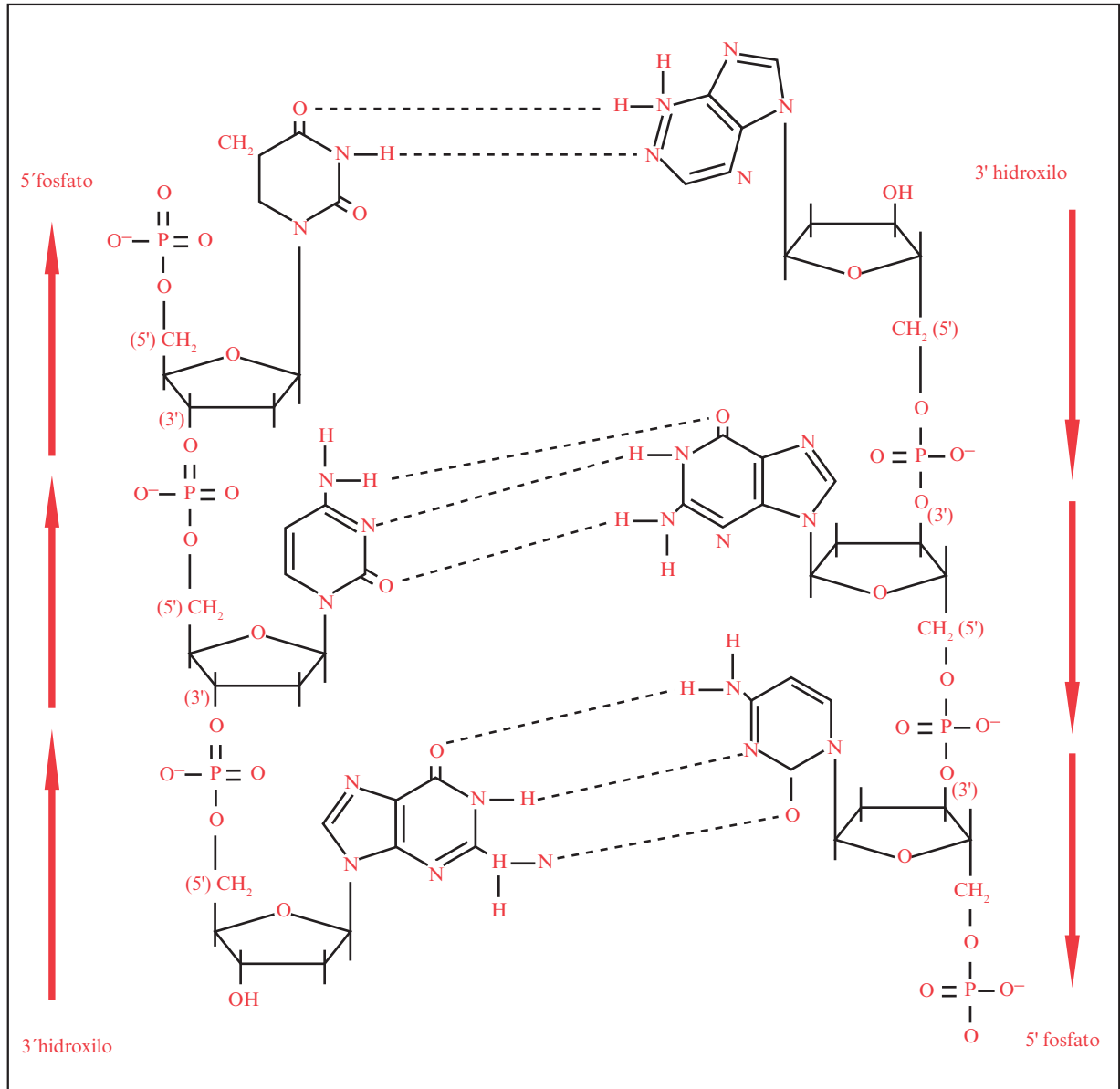


Figura 1.5. Carácter antiparalelo de la doble hélice y apareamiento de bases. Las dos hebras irían en la dirección 5' fosfato → 3'OH (dirección de la síntesis), pero los extremos superior e inferior de cada hebra serán opuestos: la hebra izquierda comenzará con un fosfato en la posición 5' y terminará con un grupo 3'OH en la desoxirribosa, mientras que la hebra derecha comenzará con un 3'OH en el azúcar y terminará con un fosfato en la posición 5'.

se traducirán a proteínas. La maduración de los ARN largos nucleares y la transformación en

mensajeros implica una serie de cortes y empalmes (que denominaremos ajuste) y una reducción

considerable del tamaño de los mismos. La mayor parte de los ARNm poseen una región de poliadenina en el extremo 3' que se considera la cola de poliA. Esta cola se añade después de la transcripción, pero antes de su andadura del núcleo al citoplasma. Los ARNm eucarióticos poseen además una modificación en el extremo 5' consistente en una guanina conectada al primer nucleótido del mensajero por una unión 5'-5' fosfato y metilada (caperuza). Esta modificación se realiza al inicio de la transcripción. Dentro de los ARN de pequeño tamaño encontramos 4 clases: 1) los ARN pequeños nucleares; 2) los ARN pequeños nucleolares que se encuentran en el nucleolo; 3) los micro-ARN, y 4) los ARN de transferencia (ARNt). Sus funciones se describen seguidamente.

El ARN participa activamente en la transmisión de la información, la regulación y el control de los procesos biológicos. De entre sus numerosas y variadas funciones celulares podemos destacar:

- a) Trasladar la información a distintas regiones celulares (ARNm, ARNt).
- b) Servir de soporte estructural de orgánulos subcelulares, como los ribosomas. Los ribosomas de mamíferos contienen 4 moléculas de ARN ribosomal (ARNr) y 82 proteínas. La subunidad pequeña (40S, siendo S una unidad de sedimentación Svedberg) alberga un ARNr de 18S (1.874 bases), mientras que la subunidad grande (60S) alberga 3 ARNr: uno de 28S (4.718 bases), otro de 7,8S (160 bases) y un tercero de 5S (120 bases). En los ribosomas, los ARN constituyen la mayor parte de su masa y su columna vertebral. La mayoría de las distintas proteínas se asientan y tienen algún tipo de contacto con un ARNr.
- c) En ocasiones, la molécula de ARN puede tener actividad enzimática, como es el caso de las ribozimas, moléculas de ARN capaces de producir un corte enzimático entre distintas regiones de su propia molécula o a un ARN complementario parcialmente.
- d) Control a nivel de transcripción o traducción, como los ARN no codificantes (ARNnc). Algunos de éstos constituyen un nuevo tipo de ARN pequeños nucleares que se asocian con promotores (PASR del inglés: *promotor associated short RNA*) y con terminadores de la transcripción (TASR del inglés: *termini-associated short RNA*). Su función concreta sin embargo no está aún esclarecida.

Además tenemos los ARN de interferencia (ARNi) y los micro-ARN (miARN). Estos pequeños ARN de hebra doble pueden tener procedencia exógena o endógena (ARNi) o ser parte del propio genoma (miARN). Su modo de acción es semejante: se asocian al complejo enzimático RISC (del inglés: *RNAi silencing complex*) que desnaturaliza y se queda con una de las dos hebras del ARN, la de menor estabilidad termodinámica en el extremo 5' (rica en A/T, frente a G/C). A continuación, busca ARNm complementarios y produce un corte en mitad de la región de homología (figura 1.6). Otras ARNasas celulares se encargan de romper totalmente la molécula diana. Como el complejo RISC-ARNi es reutilizable, en poco tiempo se puede eliminar muy eficientemente cualquier mensajero (silenciamiento). Cuando la homología entre el miARN y el ARNm no es total y reside en un extremo no traducible de la molécula (figura 1.6), puede producirse un bloqueo inicial en la traducción que puede ser seguido, bien por una rápida deadenilación del mensajero, o por la eliminación de la cola de poli-A, lo que en ambos casos desencadena su eliminación. El ARNi también puede ejercer su acción silenciadora a nivel de ADN; en este caso hibridará con el ARN naciente de forma análoga al citoplasmático, eliminándolo a la vez que induce metilaciones de genes diana a través de las enzimas metil-transferasas de algunas histonas (figura 1.6). En todos los casos la acción es de carácter negativo, es decir, eliminando un mensajero concreto o silenciando un gen por metilación. Los efectos positivos que pudieran derivarse de su acción serían indirectos, como al eliminar un represor aumentaría la expresión de sus genes diana.

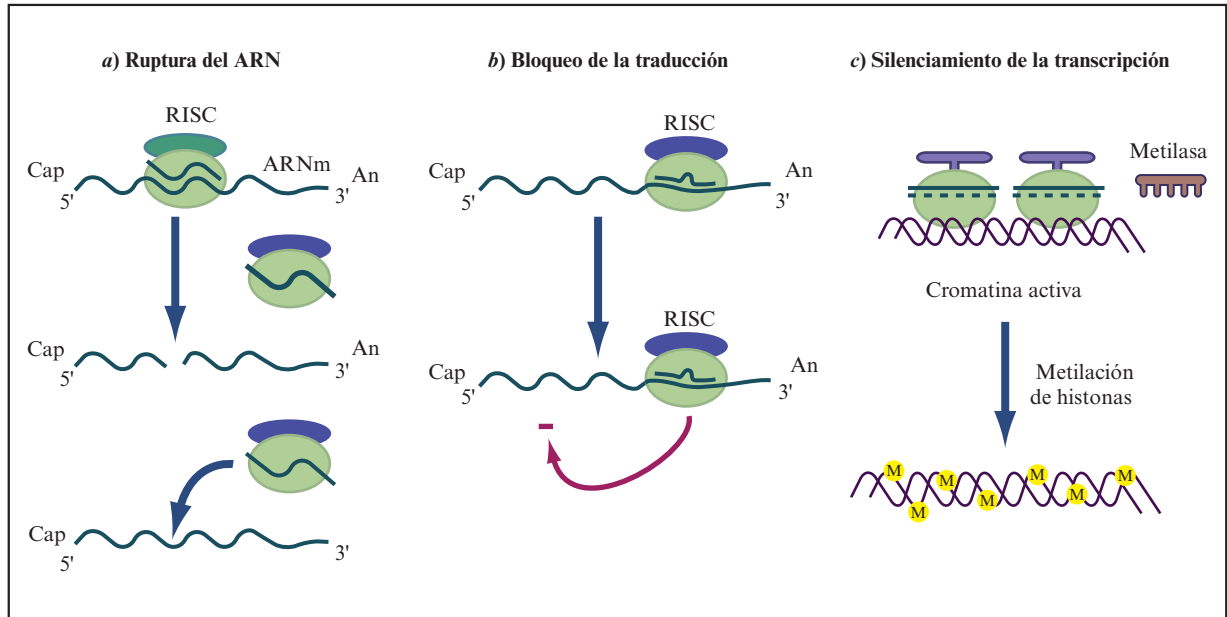


Figura 1.6. ARNi y miARN. Mecanismos de acción de los ARNi y miARN cuando: a) se actúa a nivel de mensajero y la homología es total; b) se actúa en las regiones no traducibles del ARNm y la homología es parcial, y c) actuación a nivel de represión genómica.

1.3. CURVAS DE REASOCIACIÓN: GRADOS DE REPETICIÓN EN LAS SECUENCIAS

La estabilidad de la estructura de hebra doble y helicoidal de los ácidos nucleicos se debe al apilamiento de las bases en el centro de la molécula y los puentes de hidrógeno que las mantienen apareadas y en proximidad física. Al elevarse la temperatura, los puentes de hidrógeno se rompen y la molécula pasa de su estado ordenado (nativo) a otro de enrollamiento desordenado y hebra simple (desnaturalizado). La desnaturalización del ADN puede detectarse por el aumento de absorción de luz ultravioleta, a una longitud de onda de 260 nm. Los anillos heterocíclicos de las bases del ADN absorben más cantidad de luz cuando no están apilados ni forman puentes de hidrógeno en el interior de la doble hélice. Este fenómeno se denomina hipercromicidad y se manifiesta de forma que el ADN desnaturalizado absorbe un 37% más de luz, a 260 nm, que en su forma helicoidal o apareada.

Si una solución de ADN de doble hebra tiene una absorbancia de $A_{260} = 1$, la misma solución desnaturalizada tendrá $A_{260} = 1,37$. Esta relación define al ADN de hebra doble como hipocrómico, mientras que el de hebra sencilla se considera hipercrómico (la absorbancia de las bases libres equivaldría a $A_{260} = 1,6$).

La temperatura a la cual la mitad de las moléculas han pasado al estado desnaturalizado se denomina temperatura de fusión (T_m) (figura 1.7) por analogía con el proceso físico de fusión que representa el cambio del estado sólido a líquido de la materia por acción del calor. La temperatura de fusión (T_m , del inglés *melting temperature*) varía de unos organismos a otros, de acuerdo con su porcentaje de G + C. Cuanto mayor es éste, mayor será la temperatura de fusión. Apareamientos erróneos ocasionales entre bases no complementarias también originarán un descenso de la T_m con respecto a una hélice perfectamente apareada.

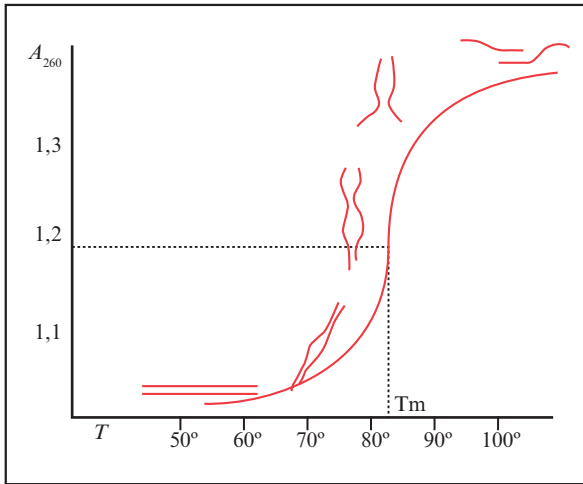


Figura 1.7. Desnaturalización del ADN y curva de fusión. A medida que se aumenta la temperatura (T), aumenta la absorbancia a 260 nm (A_{260}) y el ADN adquiere las distintas configuraciones moleculares que se representan. T_m es la temperatura de fusión, a la cual la mitad de las moléculas están ya desapareadas.

La desnaturalización por calor puede además romper los enlaces fosfodiéster, por lo que en muchas ocasiones se utilizan agentes desnaturalizantes alternativos, entre otros, pH elevado y agua destilada. Un pH superior a 11 elimina los puentes de hidrógeno entre las bases al cambiar la carga de muchos de los grupos que intervienen en estas uniones no covalentes. El agua destilada impide la neutralización de los grupos fosfato de la molécula por sales de sodio y magnesio; al ser los grupos fosfato de marcado carácter negativo, su repulsión causa la separación física de las hebras y, por tanto, su desnaturalización.

La molécula de ADN, aun en estado nativo, mantiene un equilibrio dinámico en el que pequeñas regiones, particularmente si son ricas en A/T, se desnaturalizan y renaturalizan espontáneamente, constituyendo las denominadas palpitaciones o respiración del ADN.

Compuestos químicos como la urea o la formamida son agentes desnaturalizantes porque interfieren reaccionando directamente con las bases e impidiendo un apareamiento correcto entre ellas

a temperaturas a las que normalmente estarían apareadas. La T_m se reduce proporcionalmente a la concentración del agente desnaturalizante añadido. El formaldehído es otro agente desnaturalizante cuya acción, a diferencia de los casos anteriores, es irreversible porque establece uniones covalentes con los grupos NH_2 de las bases e impide el apareamiento de éstas. Este compuesto es capaz de reaccionar con el ADN a temperatura ambiente. El hecho de que las moléculas de formaldehído puedan reaccionar con estos grupos a temperatura ambiente implica que las bases están continuamente apareándose y desapareándose en un equilibrio dinámico de formación y ruptura de puentes de hidrógeno. Este fenómeno, anteriormente descrito como palpitaciones del ADN, se cree que puede facilitar el acceso de numerosas proteínas reguladoras al interior de la molécula.

El proceso de recuperación de la estructura nativa, una vez que el agente desnaturalizante ha desaparecido, se denomina renaturalización o reasociación, proceso muy valioso para evaluar la relación genética entre diversos organismos, mostrándonos las similitudes, identidades o divergencias entre ADN de distintas procedencias.

Se trata de un proceso al azar en el que las moléculas desnaturalizadas no renaturalizan con sus complementarias originales, sino con cualquier molécula presente con la que formen un híbrido estable.

La renaturalización del ADN tiene lugar en dos pasos:

1. Dos secuencias complementarias colisionan al azar y se aparean formando una hélice corta.
2. El apareamiento se extiende a todo lo largo de las hebras a modo de cremallera.

El ADN renaturalizado recupera todas las propiedades físicas y químicas del ADN original. En ocasiones, el proceso de renaturalización se denomina hibridación, debido a que las moléculas desnaturalizadas no renaturalizan con su hebra original, sino con cualquier otra con la que tengan complementariedad.

Hibridación de ácidos nucleicos en líquido: curvas de reasociación y grados de repetición de las secuencias

Cuando el ADN de una especie se purifica, se corta en pequeños fragmentos, se desnaturaliza (100 °C) y se pone en condiciones de sal y temperatura que le permiten renaturalizar (72 °C, PBS: tampón fosfato pH 7), el proceso nos proporciona información valiosa sobre el tamaño y niveles de repetitividad de sus secuencias. La representación gráfica en coordenadas cartesianas suele considerarse: en ordenadas, el porcentaje de ADN que ya ha encontrado su pareja y, por tanto, está en forma de doble hebra; en abscisas se representa el valor Cot , que es el producto de la concentración inicial de ADN (en moles) por el tiempo de incubación. Los distintos puntos en la gráfica son el resultado de determinar el porcentaje de ADN en forma de doble hebra a distintos tiempos (en segundos). El porcentaje de ADN de doble hebra cada tiempo se estima por la propiedad de ser retenido en una columna de fosfato de calcio cristalino (HAP = hidroxipatita). La cinética de la curva resultante es de pseudoprimer orden y constituye la curva de reasociación característica de la especie. El valor $Cot_{1/2}$ representa el tiempo que tardan la mitad de las moléculas de un organismo en encontrar su pareja y nos indica el tamaño del genoma de la especie, siempre que no haya secuencias repetidas (figura 1.8, *a*).

La complejidad del genoma de una especie viene definida por su tamaño y sus niveles de repetitividad. Los genomas más sencillos son aquellos con un tamaño pequeño y carentes de secuencias repetidas, como los virus o las bacterias. En ellos, la comparación de sus curvas de reasociación ($Cot_{1/2}$) es una medición de sus tamaños relativos. Sin embargo, a medida que subimos en la escala evolutiva, las curvas de reasociación se complican conteniendo varios puntos de inflexión. Cada punto de inflexión de la curva nos indica un cambio en el nivel de repetitividad; así la curva de reasociación de un mamífero poseerá típicamente tres curvas superpuestas (figura 1.8 *b*, *c*, *d*) y tres niveles de repetitividad. La curva más rápida en reasociar estará formada por secuencias

muy repetidas que representan aproximadamente el 20% del genoma, el 30% del genoma estará formado por secuencias de repetitividad intermedia y el 40% restante representa la reasociación de secuencias únicas con una repetitividad media inferior a 10 copias por genoma haploide. En estos experimentos no se suele alcanzar un 100% de renaturalización y los resultados son aproximados. Las secuencias muy repetidas pueden estar presentes varios millones de veces por genoma haploide, mientras que las de repetitividad intermedia lo harán alrededor de unas 500 veces.

Si bien, como hemos mencionado, el valor $Cot_{1/2}$ en especies sin ADN repetido es proporcional al tamaño de su genoma, en especies con varios niveles de repetitividad es el componente único el que más se aproxima al tamaño real del genoma.

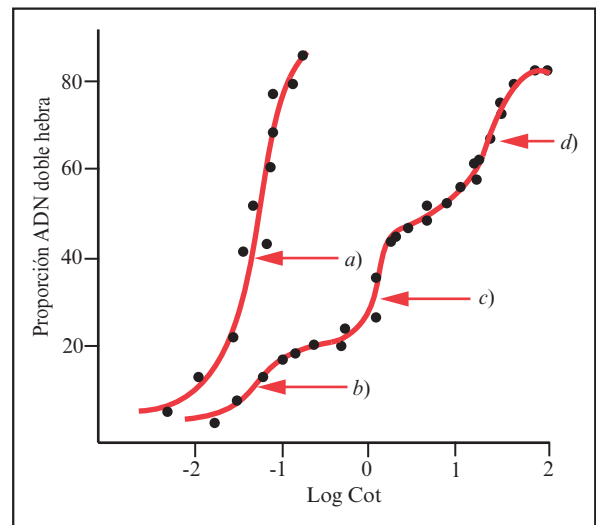


Figura 1.8. Curvas de reasociación de dos organismos con distinta complejidad genómica. El valor $Cot_{1/2}$ de un organismo sin ADN repetido representa su tamaño (flecha *a*). En organismos más evolucionados y típicamente mamíferos observamos curvas complejas con tres niveles de repetitividad y nuevos $Cot_{1/2}$ (flechas *b*, *c*, *d*) dentro de cada curva.

En curvas de reasociación con varios componentes, **el componente de reasociación más rápido**

representa secuencias muy repetidas (figura 1.8, *b*). Estas secuencias se llaman también ADN satélite por formar una pequeña banda diferenciada de la principal en gradientes de cloruro de cesio. Esto implica que su porcentaje de G/C es ligeramente diferente al de la mayoría del ADN genómico. Una parte de este ADN parece tener funciones estructurales como el mantenimiento de la estructura muy condensada del ADN de centrómeros y telómeros en las células de mamíferos (heterocromatina). El origen de estas secuencias se relaciona bien con fenómenos de replicación a saltos, es decir, salto de la ADN polimerasa durante el proceso de replicación, bien con amplificaciones laterales repentinas que originan un gran número de copias en tándem o bien con recombinaciones desiguales. Asimismo, pueden representar restos de retrovirus o transposones que no han podido ser eliminados eficazmente por la célula. El tamaño de la unidad repetida y el grado de repetitividad son muy variables en los distintos genomas. Un ejemplo de secuencias altamente repetidas son las denominadas «Alu», llamadas así por poseer una diana para esta enzima de restricción. Las secuencias Alu están emparentadas con elementos móviles o transposones, son específicas de primates y se encuentran unas 750.000 veces por genoma haploide en chimpancés y en el hombre. La unidad repetida tiene 280 pares de bases (pb), que a su vez contiene un dímero de 120 pb y está flanqueada por repeticiones cortas directas de 6-18 pb. Uno de los dímeros tiene una secuencia interna de 32 nucleótidos no presente en la otra. Además en el genoma humano encontramos secuencias largas muy repetidas que se transcriben, como son las LINE (del inglés: *long interspersed elements*), las SINE (del inglés: *short interspersed elements*) que pueden tener sólo 10 pb y estar presentes millones de veces por genoma haploide, así como terminales repetidos de transposones o elementos retrovirales no funcionales.

Las secuencias de repetitividad intermedia pueden representar genes idénticos presentes unos cientos de veces por genoma, como los genes que codifican por las histonas o el ARN ribosomal. También pueden representar familias de genes, es decir, genes parecidos pero no idénticos con fun-

ciones similares, como la familia de las globinas, las actinas, las inmunoglobulinas, kinasas, etc. (figura 1.8, *c*). Probablemente las familias de genes procedan de genes únicos que se han duplicado varias veces y han evolucionado separadamente durante miles o millones de años. Los miembros pueden estar juntos o dispersos y los intrones divergen mucho más que los exones. La separación de las copias suele ocurrir por translocación a otro cromosoma o por transposición. En el caso de las globinas se ha calculado una divergencia del 0,12% por millón de años transcurridos. Duplicaciones y divergencia son también la base de los pseudogenes. Estos últimos son genes inactivos o degenerados que mantienen una cierta semejanza con el gen original. Por ejemplo, el gen de la β -globina se duplica hace 55 millones de años, su segundo gen permanece funcional 22 millones de años más y se convierte en pseudogén desde hace 33 millones de años. Los pseudogenes acumulan mutaciones en lugares importantes frente a lugares silenciosos (como la tercera base del codón). Los pseudogenes también pueden tener su origen en una retrotransposición de un ARNm funcional que se ha insertado al azar en el genoma (capítulo 2). En este caso, el pseudogén es muy parecido al ARNm, pero carece de promotor e intrones, en los flancos lleva dos terminales repetidos directos y a veces incluye en el extremo 3' una cola de poliA. Al carecer de promotor, en principio no se puede transcribir a no ser que se integre cerca de una región promotora. En el genoma humano se han encontrado 11.224 pseudogenes de los cuales 863 se transcriben.

A su vez, las familias de genes crecen o se contraen en número de miembros por sobrecrecimientos desiguales durante la recombinación homóloga. Estas recombinaciones desacopladas entre dos familias de genes provocan una gran variabilidad en el tamaño final y número de miembros a lo largo de la escala evolutiva.

Las secuencias únicas son genes o regiones presentes entre 1 y 10 veces por genoma haploide (figura 1.8, *d*). En el genoma humano incluye 20.687 genes que codifican proteínas y cuyos exones representan el 3% del genoma (realmente, según las últimas estimaciones, el 2,94%). Cada uno

de estos genes tiene un número variable de exones e intrones y una media de 6 posibilidades de cortes y empalmes (ajustes) alternativos que en la mayoría de los casos dan lugar aproximadamente a 4 proteínas ligeramente distintas. En total, estos genes, con sus promotores, exones, intrones y regiones no traducibles, arrastran al 40% del genoma en las hibridaciones.

En general, los virus, bacterias y levaduras carecen de ADN repetido y, por tanto, dan curvas de reasociación sin puntos de inflexión. El cangrejo, por otra parte, carece de ADN moderadamente repetido, por lo que sus curvas de reasociación sólo tendrán dos componentes. Las plantas en general duplican su ADN con mucha facilidad, por lo que muchas especies carecen de ADN único o representa una proporción muy baja del total, sus curvas de reasociación tendrán valores de $Cot_{1/2}$ más pequeños que los del componente más lento de mamíferos.

De modo similar a las curvas de reasociación Cot , existen las Rot , en las que se pueden determinar los niveles de repetitividad de los mensajeros de una línea celular o tejido, así como una estimación del número de mensajeros distintos que posee la línea celular o tejido. Los experimentos son similares, pero la hibridación se realiza entre una pequeña parte de un ADN copia del ARNm marcado radiactivamente, con una gran cantidad de ARNm sin marcar que constituye el valor Ro . El valor $Rot_{1/2}$ de un solo mensajero servirá de patrón para comparar con el valor obtenido en mezclas complejas de mensajeros.

1.4. NIVELES DE EMPAQUETAMIENTO DEL ADN

El ADN nunca está totalmente desnudo. En procariotas el ADN está asociado a proteínas estructurales (HU, IHF, H1 y P) y reguladoras formando el nucleoide. En eucariotas el ADN interacciona con una gran variedad de proteínas, se espiraliza y condensa para formar la cromatina y los cromosomas.

Los cromosomas constituyen el orden superior de empaquetamiento del ADN (figura 1.9) y se pueden ver al microscopio óptico. La unidad estructural

inmediatamente por debajo del cromosoma es la fibra de cromatina de 30 nm, visible al microscopio electrónico. Se compone de una superhélice de 6-7 nucleosomas por vuelta. El ADN está enrollado alrededor de un octámero de proteínas denominadas histonas formando una unidad denominada nucleosoma. Cada nucleosoma es como un disco formado por 8 proteínas básicas (ricas en lisina y arginina) muy conservadas a lo largo de la escala evolutiva y que constituyen el corazón del nucleosoma; además, hay una histona distinta en la parte exterior del nucleosoma. Hay cinco tipos de histonas: H1, H2A, H2B, H3 y H4 (con variaciones, por ejemplo, en espermatozoides y en cromatina activa). La histona H1 es esencial en el empaquetamiento de estructuras de orden superior, ya que forma parte tanto del nucleosoma (localizada externamente) como del espacio entre nucleosomas que constituye la región de enlace. Es la histona de mayor tamaño. El corazón del nucleosoma estaría formado por: 2 H2A, 2 H2B, 2 H3 y 2 H4. La doble hélice de ADN se enrolla a la parte exterior del octámero, ocupando 145 pares de bases (pb) por nucleosoma; la región de enlace comprende unos 60 pb, pero puede variar. El conjunto de un nucleosoma y su región de enlace constituye el mononucleosoma. Cada uno de éstos abarca aproximadamente 200 pb.

Las histonas y el ADN contactan cada 10 pares de bases a través de la hendidura menor del ADN. Estudios de cristalografía de alta resolución con rayos X han revelado que, dentro del corazón del nucleosoma, cada histona está compuesta por tres módulos helicoidales y dos protuberancias o colas amino-terminales. Las colas de cada histona son largas y no muestran ninguna estructura secundaria, pero contienen los residuos que sufren las acetilaciones, fosforilaciones y metilaciones asociadas a la decondensación y, por tanto, a la actividad transcripcional de la cromatina. Las colas de las histonas H3 y H2B pasan a través de las hendiduras menores del ADN extendiéndose hacia el exterior del nucleosoma, y una de las colas protuberantes de la histona H4 llega a hacer numerosos contactos con el dímero H2A-H2B del nucleosoma vecino.

Durante la recombinación, la replicación y la transcripción, la estructura de la cromatina cambia y las histonas colaboran temporalmente con proteínas heterogéneas de diversa índole que facilitan su

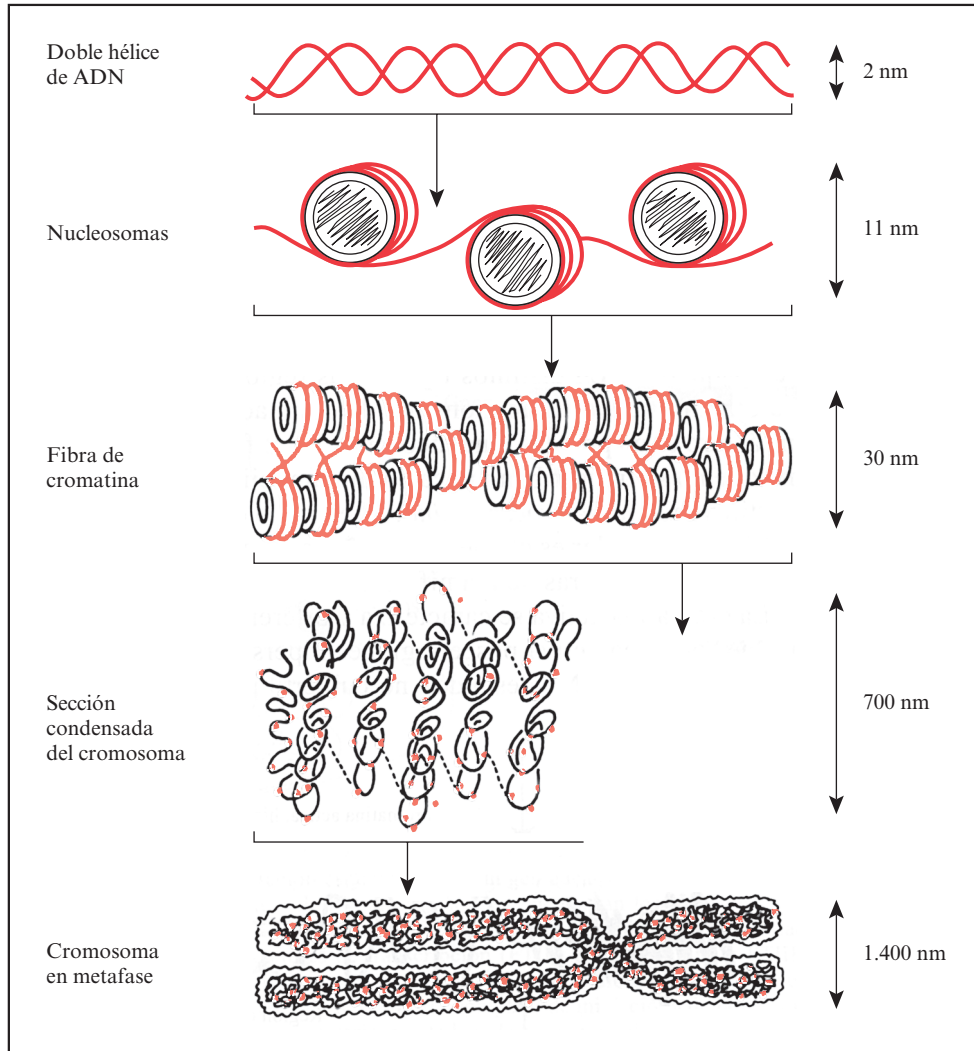


Figura 1.9. Esquema de los distintos niveles de empaquetamiento del ADN hasta formar el cromosoma. La fibra de 2 nm (nanómetros) de ADN envuelve un octámero de histonas para constituir el nucleosoma. Los nucleosomas forman filamentos de 11 nm que se empaquetan para formar la fibra de cromatina de 30 nm. El siguiente nivel de empaquetamiento es ya la fibra condensada de cromosoma (700 nm) y el propio cromosoma (1.400 nm).

propia modificación estructural. Los factores de transcripción facilitan cambios estructurales en las histonas que permiten la activación génica. Un aumento en la acetilación de las colas de las histonas se asocia con transcripción activa, mientras que una hipacetilación implica represión. Se han encontrado acetiltransferasas y deacetilasas que forman

complejos con ciertos factores de transcripción con el fin de remodelar la región apropiada en la cromatina. Así, hay factores de transcripción que inducen la metilación de la lisina que ocupa la posición 4 a partir del terminal amino de la cadena polipeptídica de la histona H3 (Lys4-H3me) y la acetilación de la lisina que ocupa la 9.^a posición (Lys9-H3ac,

H3K9ac) formándose una cromatina activa y abierta que suele estar asociada con islas CpG no metiladas. Por el contrario, una metilación de la misma lisina en la posición 9 de la histona H3 suele implicar condensación e inactividad (H3k9me3).

Se pueden distinguir hasta siete configuraciones de la cromatina con distintos niveles de empaquetamiento. Las formas menos condensadas constituyen la eucromatina y funcionalmente se corresponden con cromatina activa, que es la forma asociada a la accesibilidad de factores de transcrip-

ción al ADN y expresión de genes concretos; por el contrario, las formas más condensadas y empaquetadas se asocian con la cromatina inactiva. El paso de una configuración a otra es un proceso dinámico y puede ser frecuente. Las regiones de cromatina que no se transcriben nunca en ningún tejido están aún más condensadas, constituyendo la heterocromatina. El ejemplo de configuración heterocromática más espectacular es el que acompaña la inactivación por metilación (impronta) de uno de los cromosomas X en las hembras de mamíferos.

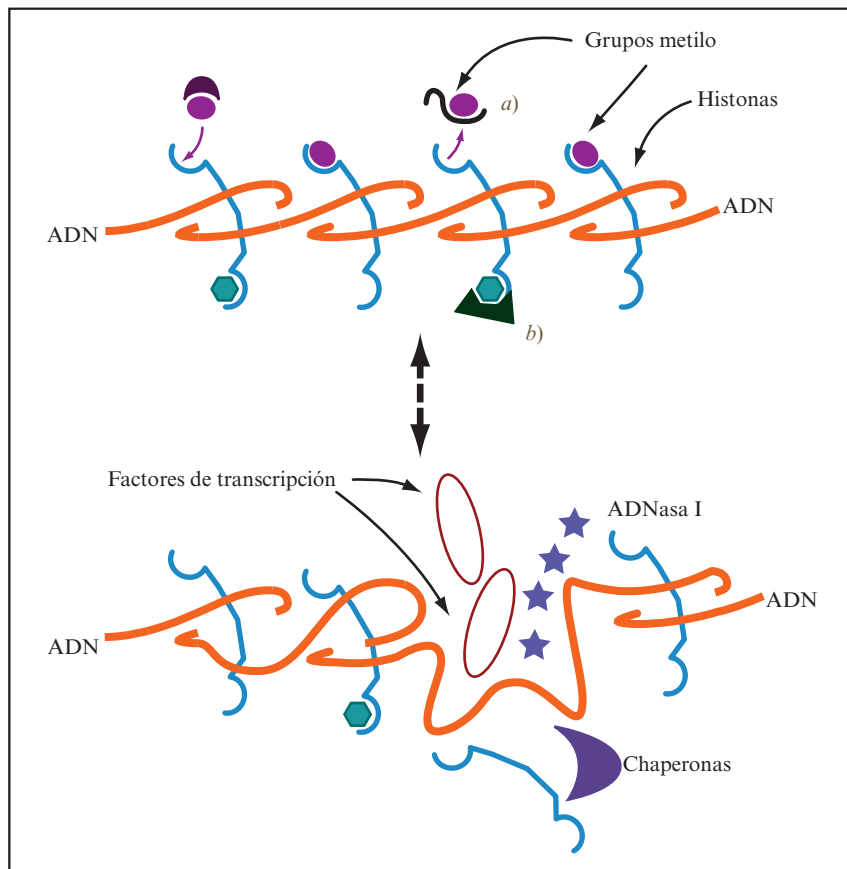


Figura 1.10. Representación esquemática de la cromatina inactiva (arriba) y activa (abajo). Enzimas que modifican la cromatina, como Set1, añaden grupos metilo en lugares específicos de la histona. Otras proteínas pueden eliminar las modificaciones químicas añadidas con antelación *a)*. Hay proteínas que se unen a las modificaciones químicas añadidas *b)*. La cromatina activa está menos condensada que la inactiva, es accesible a factores de transcripción y es hipersensible a la acción de la enzima ADNasa I que rompe el ADN en estas regiones. Además, las chaperonas de las histonas pueden desplazar y reemplazar las histonas en el ADN. Todo este escenario controla la accesibilidad de los factores de transcripción a la cromatina activa y el ADN.

La cromatina activa se caracteriza y diferencia de la inactiva en su estructura y en su hipersensibilidad a la ADNasa I. Respecto a la estructura, en levaduras, la mosca *Drosophila* y humanos se ha encontrado que la cromatina activa posee nucleosomas de unión inestable que permiten fácil acceso a la maquinaria reguladora. Estos nucleosomas inestables poseen dos variantes de las histonas: al variante H3.3 en vez de la H3 y la H2A.Z en vez de la H2A.

1.5. EXPRESIÓN DE LA INFORMACIÓN GENÉTICA Y SU REGULACIÓN

En organismos complejos, sólo un número limitado de genes se expresa en un momento dado en un tipo celular concreto, definiendo así su función. La regulación de la expresión génica es, pues, fundamental para establecer y mantener la función de una célula. La activación de la transcripción en eucariotas está mediada por una estructura abierta de la cromatina que permite la unión al ADN de factores de transcripción y ARN-polimerasas. Se estima que aproximadamente un 15% del genoma humano está en forma de cromatina activa en un momento dado. Existe un pequeño número de factores de transcripción de alta afinidad que, desplazando los nucleosomas, se unen al ADN creando una región de cromatina activa sensible a la ADNasa I. Esta configuración facilita la unión de otros factores de menor afinidad que estabilizan la estructura adquirida y permiten la iniciación de la transcripción por el complejo de iniciación, pero en ningún caso la unión de diferentes factores de transcripción entre sí es al azar.

Los promotores son, por definición, los lugares precisos donde se une la ARN polimerasa y sus co-factores, para formar el complejo de pre-iniciación y comenzar la transcripción. Los alrededores de este lugar, unos 50 pb alrededor del lugar preciso del inicio de la transcripción, también se incluyen en la región promotora. El proceso de iniciación de la transcripción se estabiliza o se reprime por la unión de numerosos y variados factores de transcripción a lugares concretos en el ADN. Los factores son proteínas que contribuyen a que se

inicie la transcripción y pueden actuar por unión directa al ADN o por unión a otros factores de transcripción. Las histonas, por otra parte, también sufren modificaciones químicas que contribuyen a las asociaciones-disociaciones con el ADN y con otras proteínas (figura 1.11).

A nivel de ADN, podemos encontrar alguna o varias de las siguientes secuencias cercanas al lugar donde se inicia la síntesis del ARN y que son regiones cortas (típicamente de 5-15 nucleótidos) a las que se unen factores de transcripción en un momento dado:

- a) La secuencia de iniciación de la transcripción (SIT) ha sido consensuada en: Pi-PiANW_iPiPi, donde Pi = pirimidina, A = adenina, N = cualquier nucleótido y W = adenina o timina. La A también representa la posición del primer nucleótido que se transcribe a ARN (+1).
- b) Una región rica en AT situada a unos 30 nucleótidos antes del lugar de inicio de la transcripción. A esta región, cuya secuencia consenso es TATAAA, se le ha denominado la caja TATA.
- c) Una región posterior a la iniciación de la transcripción (PIT) de secuencia consenso PuGW_iPiV, donde Pu = purina, G = guanina y V = A o C
- d) Dos elementos que se sitúan uno anterior (ATA) y otro posterior (PTA) a la caja TATA y que se suelen unir al factor de transcripción IIB (TFIIB).
- e) Elemento MTE (del inglés, *motif ten element*), que puede sustituir a PIT y a la caja TATA o actuar en sinergia con ellos y con SIT.
- f) Islas CpG, siendo éstas, regiones con un contenido en G+C superior al 55% y un tamaño mínimo de 500 pb (en mamíferos suele tener entre 1 y 2 kilobases).

El tratamiento con la enzima ADNasa I impedirá la degradación del ADN en toda aquella región promotora que esté físicamente cubierta por los numerosos factores de transcripción presentes en un momento dado. Así pues, las regiones sen-

sibles a la ADNasa I suelen ser más amplias que las regiones promotoras; éstas suelen estar protegidas tanto por factores que activan la transcripción como por los que la reprimen.

En organismos superiores la región promotora puede ser de dos tipos:

1. *Estrecha*, que en al menos el 10-25% de los casos posee una clara caja TATA y es muy común en el control de genes específicos de tejidos concretos. A la caja TATA se une específicamente una proteína que forma parte del complejo de preiniciación y ayuda en la elección del lugar exacto de iniciación. Este tipo de promotores cortos o estrechos suelen estar inactivos por defecto, la activación la desencadenan factores de transcripción específicos que, en una primera fase, facilitan la remodelación de la cromatina, lo que a su vez facilita el acceso de nuevo factores de transcripción a más lugares en el ADN.
2. *Ancha*, con presencia de islas CpG en muchos casos y sin cajas TATA, más abundante en organismos a partir de vertebrados y ubicuas en muchos tipos de células. Son características de genes que corresponden a proteínas de mantenimiento celular de expresión continuada y estable. Típicamente los promotores largos o anchos poseen varios lugares de iniciación (SIT) y el alto contenido en CG de estas regiones contribuye a la emisión de señales cromatínicas que ayudan a mantener la forma abierta y facilitan la iniciación.

Cada tipo de promotor se asocia con modificaciones concretas en la cromatina y factores de transcripción favoritos (figura 1.11), pero no se descarta la existencia de una gran variedad de modalidades intermedias.

Muchos promotores anchos solapan parcialmente con islas CpG que pueden o no contener citosinas metiladas dependiendo o no de la actividad del promotor. Islas con citosinas metiladas se asocian a silencio transcripcional, pero el grado de metilación/acetilación de algunas histonas y la

estructura de la cromatina constituye un nivel de control de la transcripción superior a la metilación de islas CpG. Así, algunas veces podemos encontrar transcripción activa en islas CpG con citosinas metiladas cuando las histonas mantienen su metilación y acetilación características de la cromatina activa, por ejemplo, metilación en la Lys4-H3 y acetilación en las Lys9-H3. En otros casos, como en las larvas de *Drosophila*, se ha establecido una correlación entre la presencia de la hormona ecdisona y la acetilación de la lisina en la posición 23 de la H3 que provoca la expresión de los genes localizados en las regiones 74EF y 75B de los cromosomas politénicos (gigantes) de las glándulas salivares de las larvas. La respuesta se puede visualizar al microscopio, ya que estas regiones se abultan y forman las llamadas «linternas chinas». En este caso concreto, la activación del gen no implica acetilación de la lisina 9 de la H3. Uno de los productos de estos genes es un nuevo factor de transcripción que activará muchos otros *loci* responsables del paso de larva a pupa en el desarrollo de la mosca. En general, la metilación de la lisina 9 o la 27 de la H3 se asocia con represión de la transcripción, mientras que la metilación en la lisina 4, la 36 o la 79 de la H3 se asocia con activación de la transcripción.

Los *factores de transcripción* son muy numerosos y variados, en humanos, por ejemplo, hay más de 1.400 factores de transcripción con la capacidad de reconocer y unirse a una secuencia específica del ADN. Se pueden dividir en «generales» y «específicos». Los generales se asocian al complejo de pre-iniciación en los lugares cercanos al lugar de iniciación de la transcripción y se encuentran en la mayoría de los genes que se expresan, mientras que los específicos suelen actuar más alejados de la SIT y su presencia varía en los distintos tipos celulares o tejidos. Sin embargo, puede haber sustituciones de unos factores de transcripción generales por otros del mismo tipo, cuyo cambio conlleva especificidad. Así en el cerebro muchos genes específicos de tejido se controlan con promotores anchos sin cajas TATA.

Otra característica de los factores de transcripción es que son multifuncionales y se pueden unir a distintos locus genómicos con patrones y

combinaciones diversas para especificar la activación o represión de genes. La mayoría de los complejos de factores de transcripción actúan positivamente facilitando la iniciación de la transcripción, pero algunos actúan negativamente a modo de represores de la iniciación. Hay factores que estimulan la transcripción de un gen e inhiben la de otro, dependiendo de con quien se asocian; los hay específicos de la célula, del tejido o del estadio de desarrollo. En humanos se ha encontrado que la combinación de distintos factores es determinante y depende del contexto y de la situación celular. Diferentes combinaciones se unen a dianas distintas, así el patrón de asociación de factores de transcripción no es igual en las regiones proximales al gen (promotores) y distales (lejanas al gen); pero muchos factores de transcripción pueden unirse simultáneamente a promotores y a potenciadores distantes. Hay además combinaciones de factores de transcripción que distinguen específicamente entre el alelo paterno y el materno. Se confirma en humanos que muchos represores se autorregulan manteniendo niveles de producción estables.

Cuando los factores de transcripción abandonan una región promotora, en muchos casos, al menos en humanos, se produce una metilación pasiva del ADN libre.

En el genoma humano se han observado 3 tipos de interacciones entre los factores de transcripción y el ADN: 1) factores que se unen con fuerza en regiones reguladoras dentro de la región promotora; 2) factores que se unen fuertemente sólo en regiones alejadas de los promotores, y 3) factores que se unen con fuerza en cualquier región proximal o distal a los promotores. En la región de control de la β -hemoglobina, por ejemplo, encontramos una región de hipersensibilidad a la ADNasa I de unos 110 pb, ocupada por docenas de factores de transcripción. Dependiendo del tipo celular, podemos encontrar entre 5 y 8 lugares de unión a factores de transcripción en estas regiones hipersensibles a la ADNasa I. Dentro de estas regiones encontraremos otras que son resistentes por estar protegidas por los propios factores de transcripción. También existen factores que predominantemente se

encuentran asociados a la cromatina en su forma compacta. En un momento dado, aproximadamente el 8% del genoma humano está interactuando con factores de transcripción de diversa índole. Por otro lado, grupos de genes pueden transcribirse bajo una señal de regulación común compartiendo elementos específicos de regulación. Ejemplos de ello son los elementos de respuesta al choque térmico o a las hormonas esteroideas, que son reconocidos por factores específicos que modulan la expresión de grupos de genes. De entre los factores de transcripción que se unen físicamente al ADN, son particularmente conocidos los denominados dedos. Las proteínas dedo poseen varias moléculas de zinc, por lo que se denominan dedos de zinc. Deben su nombre a la forma tridimensional que tienen las protuberancias de aminoácidos que salen de las regiones de unión a zinc, muy semejantes a dedos.

Las ARN polimerasas son realmente las máquinas ejecutoras, pero la regulación propiamente dicha está en manos de los factores de transcripción y las proteínas reguladoras. En eucariotas hay 3 ARN polimerasas (ARNpol): la ARNpol I, que transcribe el ARN ribosomal; la ARNpol III, que transcribe algunos ARN de pequeño tamaño como el ARNt, y la ARNpolimerasa II, que es la enzima que en humanos sintetiza la mayor parte del ARNmensajero, ARN no codificante y micro-ARN, a partir de un molde de ADN. El corazón de la enzima posee 12 subunidades, pero es incapaz de reconocer específicamente un promotor; para ello, requiere una serie de factores de transcripción y un conjunto de proteínas reguladoras. El complejo de iniciación está constituido por la ARN polimerasa y los factores que se unen a ella en el momento de la iniciación. El sentido de la transcripción no parece que esté predeterminado y en principio la ARN polimerasa podría tirar para cualquiera de los dos lados; sin embargo, parece cada vez más probable que bucles que unen el final de un gen con el principio del mismo son los que determinan la dirección de la síntesis. La finalización de la transcripción viene determinada en la mayoría de los mensajeros por varias señales: una región rica en GC, la formación de híbridos ADN-ARN y la presencia de la señal de

adenilación. Una región rica en GC provoca la ralentización de la ARN polimerasa, la formación de híbridos entre el ARNm recién sintetizado y el ADN puede inducir a la ARNpol a proseguir en sentido contrario al canónico, creándose un ARN de doble banda que facilitará la terminación.

Cada célula de un tejido concreto posee del orden de 200.000 regiones específicas hipersensi-

bles a la ADNasa I, número muy superior al de promotores. En el genoma humano también se han encontrado unos 2 millones de potenciadores sin diana adscrita por el momento y se estima que cada gen puede tener una media de 10 elementos reguladores. Cada tipo celular es, pues, el resultado de una serie de combinaciones y permutaciones de los elementos reguladores de sus genes, que le confieren características únicas.

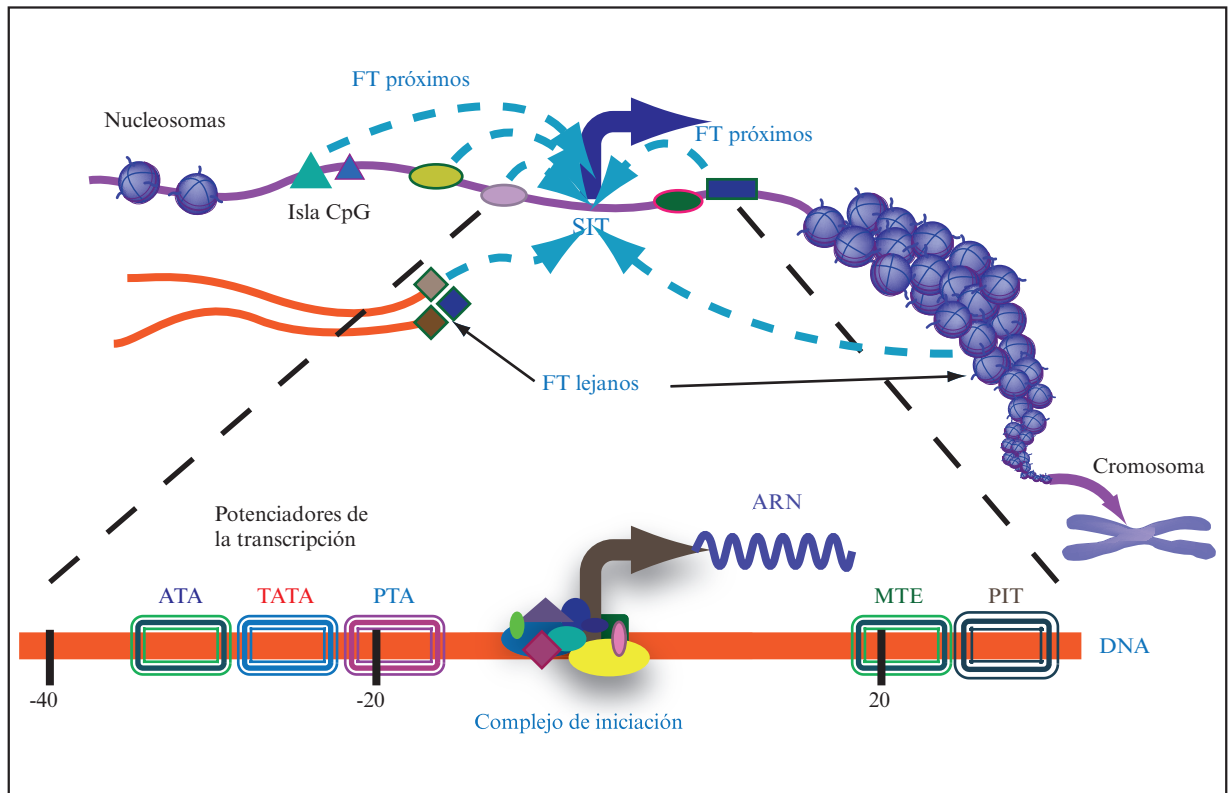


Figura 1.11. Visión esquemática de los elementos funcionales que intervienen en la expresión de un gen humano. En la parte superior el desenrollamiento de la cromatina muestra las regiones hipersensibles a la ADNasa I. Un aumento de esta región nos muestra en el flanco 5', y a una distancia variable del lugar de inicio de la transcripción, los elementos *potenciadores* de la transcripción; más cerca del inicio de la transcripción están los *promotores*, lugares de entrada de la ARN polimerasa. A continuación, los factores de transcripción que se unen a las regiones reguladoras cercanas a los promotores formando el *complejo de iniciación* que transcribe los intrones y exones que posteriormente han de ser procesados. Se muestran: la región que corresponde con una isla CpG en el ADN, la secuencia de iniciación de la transcripción (SIT), las regiones con secuencias potenciadoras de la transcripción posteriores a la iniciación de la transcripción (PIT, MTE), factores de transcripción (FT) próximos y lejanos, nucleosomas, varias regiones anteriores a la iniciación de la transcripción que son potenciadores (ATA, TATA, PTA).

La regulación de cada gen concreto en el espacio y en el tiempo parece ser de tipo espacial, es decir, formada por una red de lazos cromosomales en la que potenciadores y promotores, distantes físicamente entre sí, interaccionan y comunican su información reguladora. En el genoma humano se han identificado cientos de interacciones de largo alcance que siempre incluyen alguna región de iniciación de la transcripción. Más aún, se ha estimado que cada región de iniciación de la transcripción interacciona con una media de 4 elementos cromosomales distantes, lo que indica la existencia de una red compleja de conexiones cromatínicas reguladoras. Además, la regulación que se ejerce es de tipo dinámico cambiando en el espacio y en el tiempo y con cada tipo celular. La regulación combinada proximal (regiones promotoras) y distal (lazos) parece ser la característica que más nos diferencia a los humanos del resto de los organismos. La estructura conectiva reticular y la actuación de factores de transcripción a distancia es mucho mayor en humanos que en otros organismos.

1.5.1. Procesamiento de los mensajeros

La región en el ADN correspondiente al mensajero propiamente dicho se denomina región codificante y determina el orden lineal de los aminoácidos en la proteína a la que dará lugar. Esta región puede estar interrumpida varias veces por secuencias ajenas a la proteína codificada. Las partes que se expresan se denominan exones, y las que interrumpen la continuidad del gen, intrones. El flanco 3' del mensajero contiene las señales para la terminación de la transcripción y los procesos de maduración. En la mayoría de los genes eucarióticos, al terminar la transcripción, hay en el ARN una secuencia de poliadenilación (AAUAAA) que determina la ruptura y adición de la cola de poli-adenina a unos 10-30 nucleótidos de distancia hacia el extremo 3' desde la secuencia señal. Además, puede haber en esta región numerosos sitios de unión a proteínas reguladoras como las secuencias de unión a los miARN.

Los intrones se eliminan por un proceso complejo de corte y empalme (ajuste) que tiene lugar

en el núcleo, inmediatamente después de la transcripción cuando aún el ARN naciente está asociado a la cromatina. Hay la sospecha de que la cromatina puede estar involucrada en la regulación del complejo sistema de ajuste que sufre el ARN-pre-mensajero, antes de su transporte al citoplasma. Las secuencias consenso que en humanos determinan la eliminación de un intrón son: AG↓GU(A o G)AGU en el flanco 5' y NCAG↑G en el 3'. El número de intrones es variable, pudiendo ser de 0 a 50. El tamaño del intrón también es variable, encontrándose en su mayoría entre 75 y 2.000 nucleótidos. Hay mensajeros que utilizan distintos exones en distintos tejidos, los hay que utilizan distintos lugares de iniciación o distintos lugares de terminación. En el gen de la miosina humana, por ejemplo, tenemos dos lugares de iniciación de la transcripción que son del tipo estrecho con presencia de cajas TATA y una región no traducible (figura 1.12). La miosina embrionaria se transcribe a partir de la segunda caja TATA empezando en la región no traducible 2 (NT2) seguida por los exones: 2, 3, 5, 6, 7, 8 y 9. La miosina adulta, por el contrario, utiliza la caja TATA1, la región no traducible 1 y el exón E1, el primer intrón resulta ser la caja TATA2 y NT2, pero en este caso, del exón 1 pasamos directamente al 4, ausente en la miosina embrionaria, para seguir por los exones 5, 6, 7, 8 y 9. Bastante más complejo es el caso del gen *Dscam* (del inglés: *Down syndrome cell adhesion molecule*) de *Drosophila*, cuyo producto es homólogo a la molécula de adhesión celular descrita originalmente en pacientes con síndrome de Down en humanos. *Dscam* codifica por un receptor de membrana con 3 regiones bien diferenciadas: extracelular, transmembrana e intracelular, que facilita conexiones precisas entre distintos tipos de neuronas. Las partes variables son extracelulares, cada receptor recibe una señal del exterior que se traduce en un movimiento del axón para formar una determinada conexión nerviosa; todo esto en los conos oculares de la mosca. Hay una gran diversidad de receptores y se requiere que cada uno de ellos responda a un nivel distinto de una misma señal. El gen *Dscam* posee 115 exones: 20 fijos y 95 variables, que dan lugar a 38.016 receptores distintos de guía de axones (figura 1.12, b). En humanos el gen *Dscam* no

tiene la misma función ni ajuste que en *Drosophila*, pero los genes de las inmunoglobulinas operan con mecanismos semejantes.

Cada gen humano se estima que tiene 6 posibilidades de ajustes alternativos y suele generar

una media de 4 proteínas distintas. Esta versatilidad ofrece posibilidades de que un mismo gen pueda realizar funciones ligeramente distintas según las necesidades de la célula o según el tipo de células. El orden de los exones no se altera.

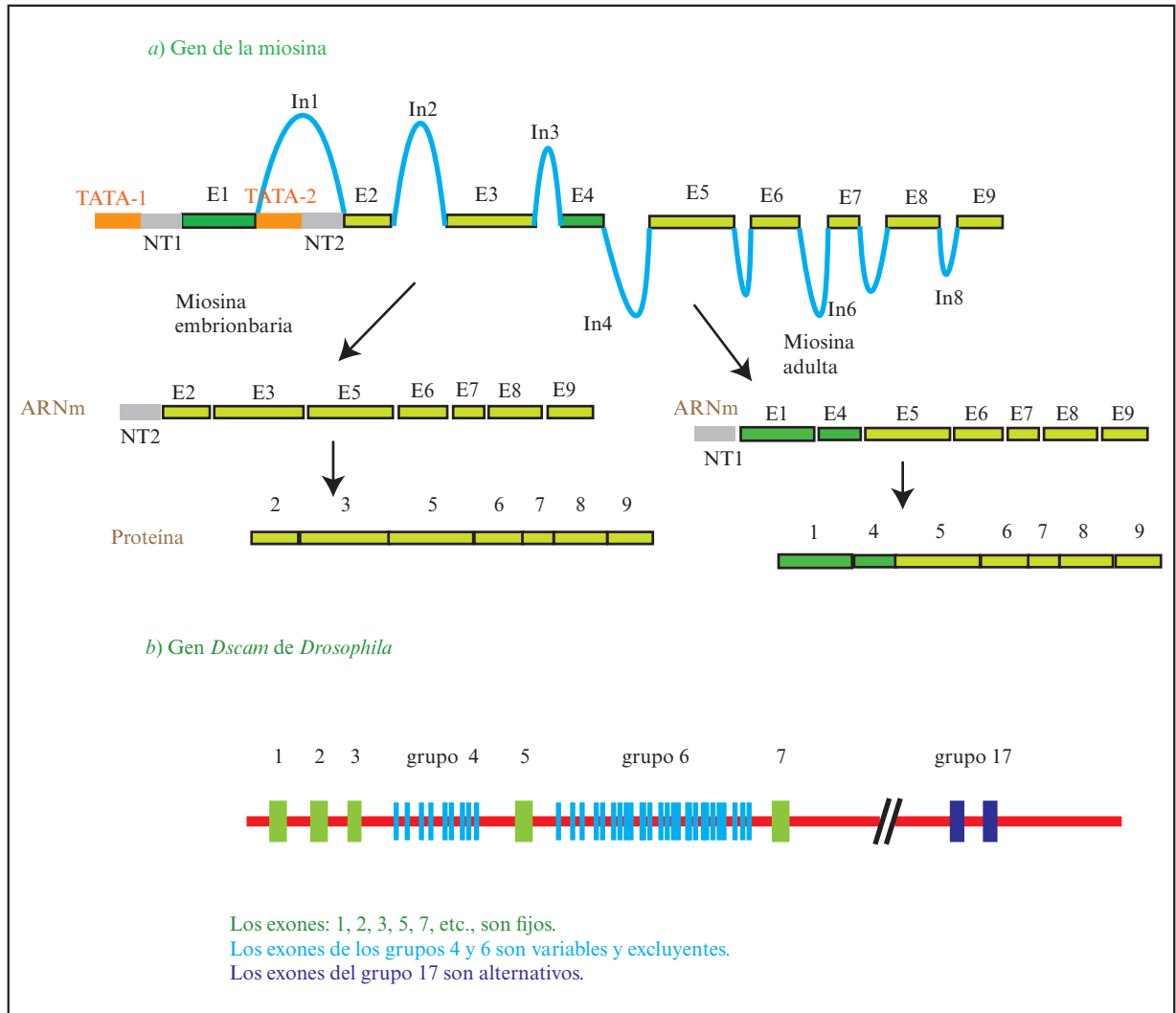


Figura 1.12. Unidades de transcripción complejas: *a)* Utilización de distintos exones en la miosina embrionaria y miosina adulta. La miosina embrionaria utiliza la TATA2 para iniciar la transcripción en la región no traducible 2 (NT2) y consta de los exones 2, 3, 5, 6, 7, 8 y 9; la miosina adulta utiliza la TATA1 y la transcripción se inicia en la región no traducible 1 (NT1), posee los exones 1, 4, 5, 6, 7, 8 y 9. *b)* El gen *Dscam* de *Drosophila* que tiene 115 exones: los exones 1, 2, 3, 5, 7, etc., son fijos; los exones del grupo 4 y del grupo 6 (48 exones) son variables y excluyentes; el grupo 17 tiene 2 exones alternativos.

1.5.2. Traducción

La traducción de los mensajeros se realiza en los polirribosomas. Los sistemas de traducción de procariotas y eucariotas son, hasta cierto punto, intercambiables, en el sentido de que mensajeros eucarióticos, aunque con dificultades, se pueden traducir en procariotas, y viceversa. Las diferencias mayores están en la formación de los complejos de iniciación y reconocimiento, así como en la frecuencia de uso de los distintos codones del código genético.

En el proceso de iniciación intervienen la subunidad pequeña del ribosoma, el ARNm y los factores de iniciación. En bacterias hay tres posibles codones de iniciación: AUG, GUG y UUG. En eucariotas siempre se utiliza el codón AUG para iniciar la síntesis, y, aunque el ARN_t iniciador (ARN_{t_i}) es distinto al de elongación, la metionina inicial no está formilada (ARN_{t_f}) como en procariotas. En organismos superiores hay al menos nueve factores de iniciación en comparación con los tres de procariotas.

El código genético está degenerado en el sentido de que distintos tripletes (especialmente en la tercera base) codifican el mismo aminoácido (tabla 1.1), pero distintos organismos usan preferentemente siempre los mismos tripletes de entre todos los comunes para un aminoácido. Esto se traduce en una abundancia o escasez de los ARN_t más usados frente a los poco frecuentes. Un ARN que posea varios tripletes de poco uso para un sistema determinado se traducirá con dificultad.

Hay tres codones que no corresponden a ningún aminoácido y que carecen del ARN_t correspondiente: son los codones de terminación de la traducción. Estos codones tienen nombres dados en su descubrimiento: UAG = ámbar, UAA = ocre y UGA = ópalo.

Durante el proceso de traducción también hay regulación de la expresión génica. Por ejemplo, la iniciación puede ser obstruida por proteínas que se unen a los lugares de reconocimiento del ribosoma por el mensajero, impidiendo así la formación del complejo de iniciación. El caso de la tubulina es de este tipo. La tubulina es la proteína fundamental de los microtúbulos y su producción está controla-

TABLA 1.1

Degeneración del código genético donde se pone de manifiesto cómo varios codones diferentes corresponden a un mismo aminoácido

Aminoácido	Tripletes del código
Alanina	GCU, GCC, GCA y GCG
Arginina	CGU, CGC, CGA, CGG, AGA y AGG
Asparagina	GAU y GAC
Glicina	GGU, GGC, GGA y GGG
Leucina	CUU, CUC, CUA, CUG, UUA y UUG
Prolina	CCU, CCC, CCA y CCG

da por las existencias celulares de la proteína. Cuando los niveles celulares son elevados, la tubulina libre interfiere con uno de los primeros estadios de la traducción, uniéndose directamente a la cadena naciente. El resultado es la degradación del mensajero evitando así la producción de más tubulina. Más aún, la tubulina recién sintetizada puede sufrir una serie de modificaciones reversibles y dinámicas que pueden condicionar la función de los microtúbulos. Los microtúbulos son polímeros de α - y β -tubulina dinámicamente ensamblados y presentes en todas las células eucarióticas. Participan en numerosos procesos celulares incluyendo: mitosis, movilidad celular, transporte intracelular, secreción, forma celular y polarización. Así, aunque la estructura de los microtúbulos se conserva en distintos tipos celulares y organismos, su función es variada y las modificaciones posteriores a la traducción son selectivas e importantes para su función posterior. Tomaremos como ejemplo los microtúbulos de las neuronas. Las neuronas son células polarizables y excitables con múltiples extensiones, una de las cuales se especializa en axón y las otras en dendritas (figura 1.13). En los microtúbulos de las neuronas encontramos diversas modificaciones posteriores a la traducción de sus componentes: 1) La eliminación de la tirosina del extremo carboxi-terminal de la α -tubulina (detirosinación) y la subsiguiente eliminación adicional del penúltimo glutamato generando la forma

$\Delta 2$ -tubulina. En las neuronas, por ejemplo, esta modificación tiene una función estabilizadora, es irreversible y no ocurre en las β -tubulinas. 2) La adición de varios residuos de glutamato (glutamilación) o de glicina (glicilación) en las colas de los extremos carboxi-terminal de ambas tubulinas α y β . La modificación es reversible, pero no se conocen las enzimas que llevan a cabo la retirada de los residuos. En el caso de las neuronas encontramos muchas tubulinas glutamiladas y ninguna glicilada. 3) La acetilación en la lisina 40 localizada en muchos casos en el espacio interno del microtúbulo y cuya función en neuronas se relaciona con la

memoria a corto y largo plazo. Se tiene la idea de que, particularmente en neuronas, las modificaciones selectivas de los microtúbulos son las que producen la sintonía fina de la regulación de los procesos que controlan la diferenciación neuronal, la guía del axón, el transporte neuronal, la sinapsis y la memoria a corto y largo plazo. No serán iguales las modificaciones observadas en los microtúbulos de las dendritas a las del axón de una misma neurona; ni las de la parte media del axón coincidirán totalmente con las de la parte apical de crecimiento, generando patrones microtubulares específicos asociados a funciones concretas (figura 1.13).

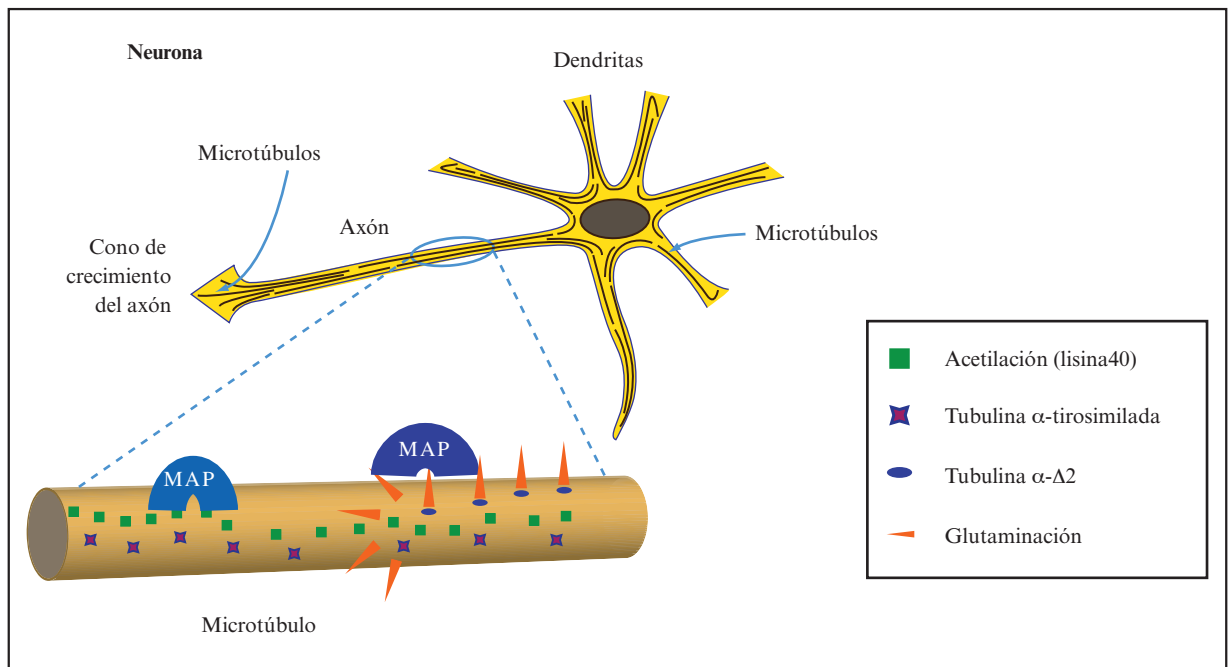


Figura 1.13. Representación esquemática de los tipos de modificaciones posteriores a la traducción que podemos encontrar en los microtúbulos de un axón de una neurona. Estas modificaciones serán ligeramente distintas a las que encontraríamos, por ejemplo, en las dendritas de la misma neurona, o en el cono terminal de crecimiento del axón, creando así un patrón subcelular específico.

Como ya se ha comentado con anterioridad (1.2.2 y 1.5), los ARN no codificantes (ARNnc), fundamentalmente el ARN de interferencia y los micro-ARN, regulan de manera muy eficiente la

expresión de un gen a nivel de traducción y ellos a su vez están regulados a nivel de ADN por factores de transcripción. El ARNi degradará de modo rápido y eficaz cualquier mensajero com-

plementario a su secuencia y el micro-ARN impedirá la traducción cuando la complementariedad con su diana no sea perfecta y ésta se encuentre en la región 3' no traducible del mensajero (figura 1.6). Las ventajas de una regulación por ARNi o miARN en vez de factores de transcripción es su rapidez de producción, ya que al no ser proteínas no necesitan traducirse para actuar, además su destrucción cuando ya no se necesiten es barata, ya que son pequeños. Por la misma razón estos ARN pueden evolucionar rápidamente. Las interacciones de los pequeños ARNnc con sus dianas por hibridación de sus ácidos nucleicos son más sencillas que las de los ADN y los factores de transcripción que son ADN/proteína. Factores de transcripción y miARN pueden cooperar en sus genes diana, ya que éstos pueden poseer en el ADN varios elementos en *cis* de unión a factores de transcripción y más de una región de reconocimiento para miRNA que disminuirá drásticamente los niveles de la proteína. Unos y otros sólo requieren entre 6-8 nucleótidos específicos, como región de reconocimiento.

1.6. VARIACIONES EN LA CANTIDAD DE ADN Y NÚMERO DE GENES A LO LARGO DE LA ESCALA EVOLUTIVA

Para contestar a la pregunta de si representa la cantidad de ADN de los organismos (genotipo) su complejidad morfológica (fenotipo), podemos analizar la cantidad de ADN de algunas especies que han servido como modelo a lo largo de la escala evolutiva. Hoy día están secuenciados los genomas de muchas especies y conocemos sus tamaños con gran precisión. A través de la base de datos norteamericana NCBI (National Center for Biotechnology Information) podemos acceder a los valores sobre tamaños de genoma, número de genes estimado y número de proteínas de una gran variedad de organismos. Los datos según esta fuente de información son bastante más elevados que los de otras fuentes (ENCODE por ejemplo), pero internamente coherentes. La bacteria *Escherichia coli* tiene un tamaño de genoma de 5×10^6 pares de bases (pb); la levadura *Saccha-*

romyces cerevisiae tiene 12×10^6 pb, la mosca del vinagre *Drosophila melanogaster* $1,5 \times 10^8$ pb, el gallo *Gallus gallus* 1×10^9 pb, la vaca *Bos taurus* $2,6 \times 10^9$ pb y el hombre *Homo sapiens* 3×10^9 pb. Hasta aquí podríamos admitir que una mayor complejidad morfológica va asociada a un mayor tamaño de genoma, aunque la diferencia entre estas 3 últimas especies parece demasiado pequeña para los enormes cambios fenotípicos que se observan. Efectivamente, un análisis algo más exhaustivo nos muestra algunas desviaciones (tabla 1.2). Dentro del reino vegetal, la pequeña planta con flores, crucífera, de la familia de las mostazas y modelo por excelencia *Arabidopsis thaliana* tiene un tamaño de genoma de 1×10^8 pb, ligeramente inferior al de *Drosophila* y la planta del arroz *Oryza sativa* ligeramente superior, con $3,8 \times 10^8$ pb. El maíz, con 2×10^9 pb, supera en tamaño al gallo y se acerca mucho al de la vaca y al del hombre. El hombre se ve ligeramente superado en tamaño de genoma por su pariente vecino el chimpancé *Pan troglodytes* con $3,1 \times 10^9$ pb. Así pues, la cantidad de ADN de un organismo no necesariamente correlaciona con su nivel de complejidad dentro de la escala evolutiva (figura 1.14).

Otro parámetro interesante de comparar es la cantidad de genes y proteínas de los organismos con el grado de complejidad morfológica y funcional de los mismos. Volviendo a los datos del NCBI, el número estimado de genes, así como el de proteínas, de algunos organismos a lo largo de la escala animal nos indica que tampoco este parámetro es un fiel reflejo del grado de complejidad que observamos en los distintos organismos (tabla 1.1). La bacteria *E. coli* posee 5.000 genes estimados a partir del ADN secuenciado teniendo en cuenta principalmente las pautas abiertas de lectura y ausencia de codones de terminación; el número de proteínas estimado para este organismo es de 4.800. El porcentaje del genoma implicado en la codificación de estos genes es del 100%. En la levadura *S. cerevisiae* aproximadamente el 70% de su genoma codifica por 5.500 genes y 4.300 proteínas. En la rama vegetal el número de genes de las dos especies estudiadas es similar al número de genes estimado para el hombre; sin embargo, el porcentaje del genoma implicado en su codificación es

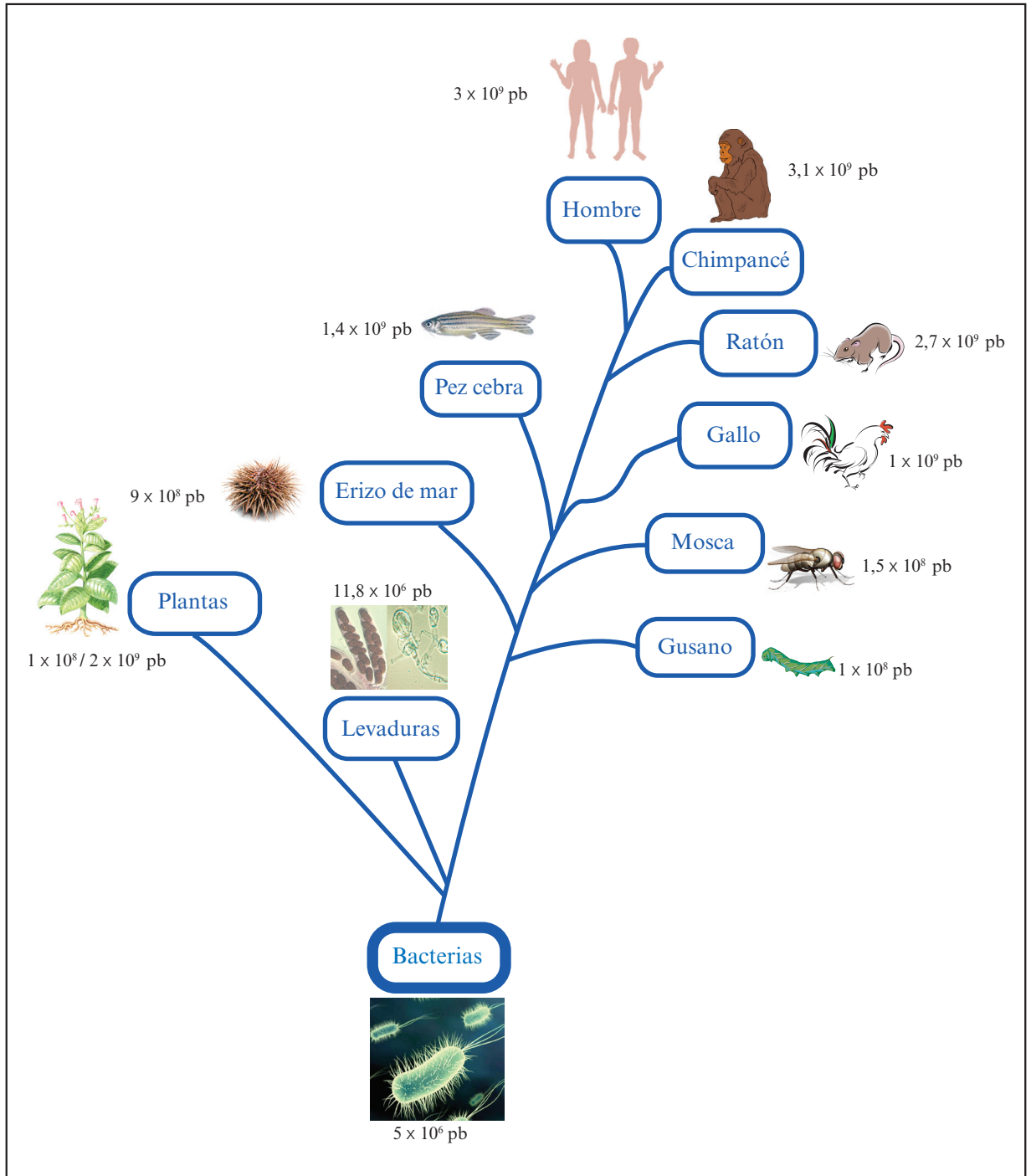


Figura 1.14. Tamaños de los genomas de diversas especies (en pares de bases: pb).

bastante superior (50 y 13% frente a un 2,9%). En la rama animal el número estimado de genes se mantiene entre 15.000 de la mosca del vinagre y 36.300 del ratón. El hombre, con 35.000 genes¹, está cercano al ratón y al chimpancé (30.200 genes) pero los errores de cálculo sólo nos permiten asegurar que todos ellos tienen un número de genes muy similar y un porcentaje del genoma implicado muy parecido. El chimpancé (*Pan troglodytes*)

tiene muchas más regiones duplicadas en su genoma que el hombre, lo que, con el tiempo, probablemente ha originado nuevos genes.

En cuanto al porcentaje genómico implicado en la codificación a proteínas, la tendencia es que a mayor complejidad morfológico-funcional menor porcentaje del genoma implicado en la expresión de genes traducibles a proteínas, aunque hay algunas excepciones (tabla 1.2).

TABLA 1.2

Cantidad de ADN en algunos organismos modelo y número estimado de genes y proteínas, porcentaje aproximado del genoma que representan los genes

Organismo	Cantidad de ADN pb	Genes/ Proteínas	% del genoma*
<i>Escherichia coli</i> (bacteria)	5×10^6	5.000/4.800	100%
<i>Saccharomyces cerevisiae</i> (levadura)	$11,8 \times 10^6$	5.500/4.300	70%
<i>Arabidopsis thaliana</i> (planta con flor)	1×10^8	33.600/35.400	50%
<i>Oriza sativa</i> (arroz)	$3,8 \times 10^8$	34.500/28.500	13%
<i>Zea mays</i> (maíz)	2×10^9	—	—
<i>Caenorhabditis elegans</i> (gusano)	1×10^8	21.200/23.900	32%
<i>Strongylocentrotus purpuratus</i> (erizo de mar)	9×10^8	23.200/22.700	4%
<i>Drosophila melanogaster</i> (mosca del vinagre)	$1,5 \times 10^8$	15.500/24.100	16%
<i>Danio rerio</i> (pez cebra)	$1,4 \times 10^9$	28.700/27.400	3%
<i>Gallus gallus</i> (gallo)	1×10^9	17.500/16.800	2,6%
<i>Mus musculus</i> (ratón)	$2,7 \times 10^9$	36.300/29.000	2%
<i>Bos Taurus</i> (vaca)	$2,6 \times 10^9$	27.100/22.000	1,5%
<i>Pan troglodytes</i> (chimpancé)	$3,1 \times 10^9$	30.200/32.500	1,4%
<i>Homo sapiens</i> (hombre)	3×10^9	35.000/32.000	2,9%

* El porcentaje del genoma que codifica por genes es aproximado y se ha calculado considerando un tamaño medio de un gen de 1.500 nucleótidos, excepto en humanos donde se han utilizado los datos reales según ENCODE.

¹ El número de genes asignados a la especie humana varía según la fuente elegida. Así, de 35.000 según la base de datos norteamericana NCBI (National Center for Biotechnology Information) a 20.687 según ENCODE (Encyclopedia of DNA elements).

En conclusión: ni la cantidad de ADN, ni el número de genes, ni el número de proteínas celulares correlacionan con la complejidad morfológico-funcional que observamos en los distintos

organismos. La adquisición progresiva de funciones más evolucionadas y complejas está relacionada con las variaciones en la regulación de la expresión génica. Hoy sabemos que el 80% del genoma humano, por ejemplo, está implicado en la estricta regulación del 2,9% que constituyen los exones de los genes que codifican por proteínas.

1.7. GENÉTICA Y EPIGENÉTICA EN RELACIÓN CON LA VARIABILIDAD ENTRE CÉLULAS, INDIVIDUOS Y ESPECIES

Se podría decir que la genética propone y la epigenética dispone. En sentido muy amplio, la epigenética estudia las modificaciones internas que diferencian dos células con idéntica dotación génica. El término «epigenética» lo acuñó Conrad Waddington en 1942 para describir las manifestaciones externas de la actividad genética durante el desarrollo. Aunque el término precede a la moderna biología molecular, hoy sigue teniendo mucha actualidad para describir las modificaciones que, sin cambiar la secuencia del ADN, afectan a la metilación del ADN o a la metilación, acetilación, fosforilación y ubiquitinación de las histonas. Otro factor imprescindible de evaluar en la creación de variabilidad es la transcripción y procesamiento de los ARN, incluyendo la función reguladora de los micro-ARN, así como los cortes y empalmes (ajuste) diferenciales que crean distintas proteínas a partir de un mismo ARN.

Todas las células de un organismo reciben una dotación génica idéntica y, sin embargo, la morfología y función de los distintos órganos es muy dispar. La identidad y función de una célula está gobernada por su patrón de expresión génica durante el desarrollo y en su destino final. No sabemos cómo se establecen y mantienen con precisión los distintos patrones de expresión/represión característicos de cada tejido, pero ellos son la base fundamental de las diferencias entre los distintos órganos de un ser humano. Como ya hemos indicado con anterioridad, la secuenciación de las regiones susceptibles al tratamiento con la

enzima ADNasa I (que degrada las zonas con cromatina activa y accesible a los factores de transcripción, figura 1.10), de 120 tipos de células humanas ha revelado la existencia de 2,9 millones de lugares en el genoma donde factores de transcripción se unen al ADN. De ellos, 3.700 son comunes para todos los tipos de células y aproximadamente un tercio se encontraron sólo en un tipo celular. Así pues, cada tipo celular usa una combinatoria de factores reguladores especial y específica. Esta riqueza de posibilidades ayuda a explicar cómo es posible que un puñado de genes que no llega al 3% del genoma pueda crear, mantener y dirigir todas las acciones de un ser humano.

Los mismos fenómenos que actúan a nivel de individuo para crear los distintos tejidos pueden crear diferencias a nivel de individuos de la misma especie.

Cada vez es más evidente que el genoma es extremadamente conservador y que las diferencias, a nivel de ADN, entre dos individuos de la misma especie son mínimas. Las mayores diferencias genéticas en las especies con dos sexos están entre ellos. En la especie humana las mujeres tienen dos cromosomas X, que tienen un tamaño considerable (150×10^6 pb), mientras que los varones poseen un cromosoma Y en vez del segundo cromosoma X, de tamaño mucho más pequeño (50×10^6 pb).

Si bien las diferencias entre distintas especies tienen un fuerte componente genético basado en duplicaciones, deleciones, reordenamientos y mutaciones que se han ido sucediendo en el tiempo durante millones de años, las diferencias entre individuos de la misma especie dependen además de otros factores como el de la reproducción sexual, que baraja, reordena e intercambia genes de individuos distintos, así como de la regulación de la transcripción durante el desarrollo, que incluye metilación del ADN y procesamiento de los ARN.

El erizo de mar comparte unos 7.000 genes con el ser humano y un gran número de factores de transcripción son intercambiables entre organismos separados hace millones de años. Un ejemplo de conservación de factores de transcrip-

ción es el caso de los genes *Pax* que controlan el desarrollo del ojo en vertebrados y mamíferos. El equivalente del gen *Pax6* de ratón en *Drosophila* es el gen *eyeless* que desencadena la formación del ojo. La forzada expresión de *eyeless* en lugares ectópicos del animal, como el ala, la pata o la antena, origina la formación de un ojo de manera ectópica en el ala, pata o antena. De forma similar, la expresión del gen *Pax6* de ratón en los halterios (alas rudimentarias truncadas) y ala de la mosca origina una mosca con 6 ojos: los 2 normales en su posición correspondiente por la expresión de *eyeless*, 2 en la base de los halterios y 2 en la base de las alas (figura 1.15.a). Tanto *Pax6* como *eyeless* son genes «maestros» que codifican por sendos factores de transcripción que desencadenan una cascada de activaciones y represiones génicas encadenadas que terminan con la producción del ojo. Si bien *Pax6*, como hemos visto, produce ojos (no funcionales) en la mosca, la inversa no se cumple; tampoco *Pax6* es capaz de rescatar mutantes *eyeless*, que en homocigosis es letal.

La diferencia genética entre el chimpancé y el hombre es del 1,4%. Dos individuos de la misma especie comparten el 99,8% de sus genes y dos hermanas el 99,98%. Todo ello nos confirma que es la regulación de la expresión génica, y no los genes, lo que en último término crea la variabilidad entre individuos. El caso más claro lo presentan los gemelos univitelinos que reciben la misma dotación genética y, por tanto, las diferencias fenotípicas que se observan entre ellos han de atribuirse a la diferente expresión del genoma y a factores ambientales.

Si tomamos como ejemplo dos individuos de la especie humana, podremos asignar una parte no despreciable de la variabilidad al componente genético. La reproducción sexual y la recombinación homóloga durante la formación de los gametos aseguran intercambio forzoso de genomas distintos.

Existen dos circunstancias especialmente predisuestas a crear variabilidad en los organismos: 1) cambios genéticos o epigenéticos durante el desarrollo y 2) variaciones epigenéticas durante la transcripción y procesamiento de los ARN.

Cambios genéticos durante el desarrollo. El paso de una célula huevo a un organismo adulto se realiza a través de un conjunto de interacciones moleculares y celulares que desencadenan un programa coordinado de activación y represión de genes específicamente regulado en el espacio y en el tiempo. Pequeñas variaciones en el desarrollo de este programa provocará variantes fenotípicas estables. Estas pequeñas variaciones pueden tener un origen genético, como las mutaciones y reordenaciones, o epigenético, como las metilaciones de genes o histonas.

Las mutaciones son alteraciones de diversa índole en el ADN. Estas alteraciones pueden ser tan pequeñas como el cambio de una base por otra durante la replicación del ADN. El efecto de las mutaciones puede variar de nulo a letal, pasando por una variedad de estados intermedios que pueden afectar a un producto o desencadenar una serie de eventos que amplifiquen el efecto final. Así pequeñas mutaciones o translocaciones no tendrán impacto en el porcentaje de variabilidad de un genoma, pero pueden desencadenar un cambio fenotípico reconocible. La segmentación en *Drosophila* se produce por una expresión en cascada de los genes homeóticos (implicados en formar diversas partes del cuerpo) seguida de una serie de interacciones entre proteínas reguladoras que van activando sucesivamente los genes estructurales responsables de las distintas partes del cuerpo. Por ejemplo, el locus homeótico *bithorax* dirige la formación del tórax del animal. Una mutación en la unidad transcripcional *ultrabithorax* (*Ubx*) puede tener efectos espectaculares como provocar una mosca con dos pares extra de alas (figura 1.15.b) o *antennapedia* (*Ant*) que desencadena el desarrollo de una pata en el lugar de la antena (figura 1.15.c). En el primer caso se trata de la falta de expresión del gen *Ubx*, lo que provoca la transformación homeótica de la región torácica 3 de la mosca en la región inmediatamente anterior, es decir, la torácica 2, donde se encuentran las alas. La mutación de *antennapedia* que presentamos va asociada a una translocación provocada por rayos X, por lo que no hay pérdida o ganancia de ADN, simplemente una relocalización de secuencias reguladoras y estructurales.

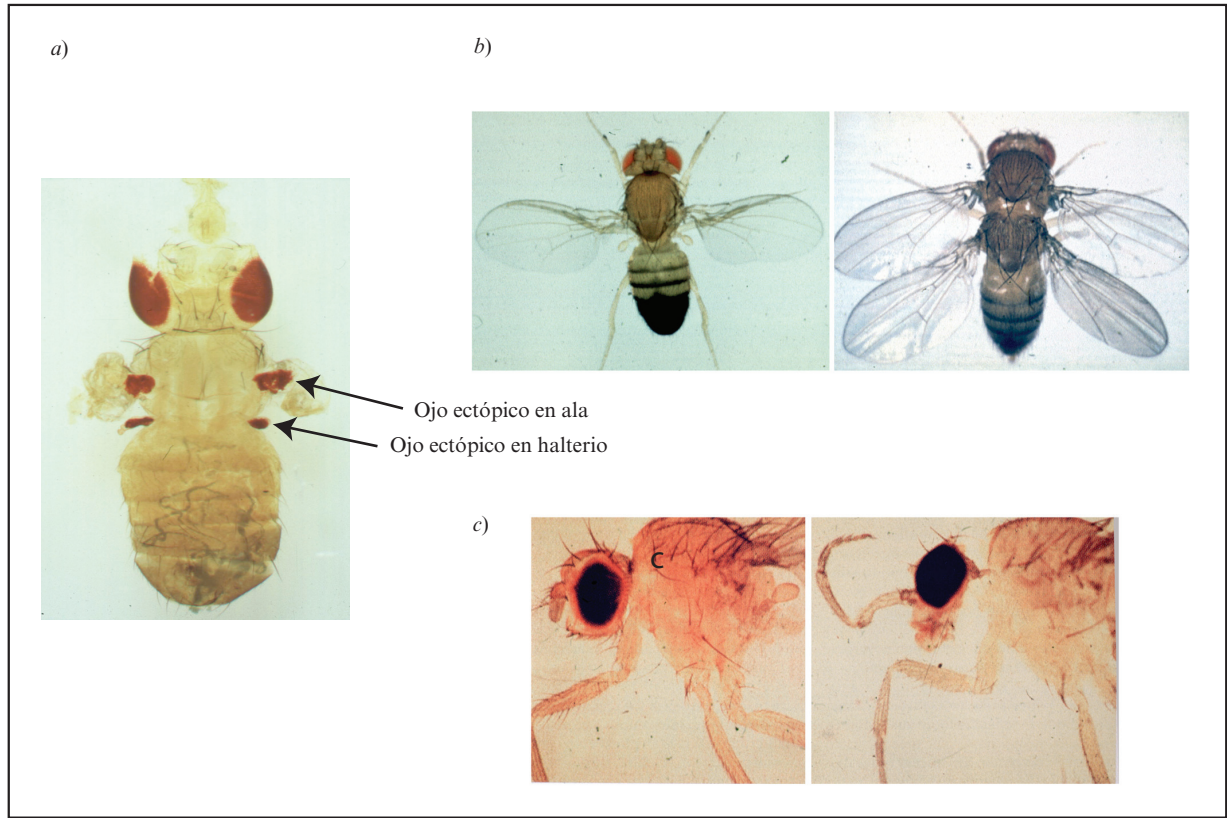


Figura 1.15. Efecto del intercambio de genes y mutaciones homeóticas en *Drosophila melanogaster*. a) La expresión del gen maestro *Pax6* de ratón, en la base de los halterios y alas, origina una mosca con 6 ojos. b) Una mutación en el complejo transcripcional *ultrabitorax* (*Ubx*) puede provocar una mosca con 4 alas. c) Una translocación que implique el locus *antennapedia* (*Ant*) puede dar lugar a una mosca con una pata en vez de antena. Fotos cortesía de Ginés Morata.

Cambios epigenéticos durante el desarrollo.

Durante las primeras fases del desarrollo embrionario de mamíferos, el genoma de las distintas células sufrirá una serie de desmetilaciones y metilaciones *de novo* sucesivas que durante la diferenciación a los tejidos adquirirá un perfil específico estable. Perfil re-programable durante la gametogénesis. En mamíferos, la metilación de citosinas en las regiones ricas en el dinucleótido CpG que preceden a muchos genes causan su silenciamiento. Estas regiones se llaman islas CpG y se encuentran dentro de la región promotora de una gran variedad de genes. Cuando una isla CpG no está metilada, la cromatina adquiere una for-

ma abierta que facilita la unión del ADN a la ARN polimerasa o factores de transcripción. La metilación desencadena un cambio en la estructura de la cromatina que impide el acceso de moléculas reguladoras o iniciadoras de la transcripción, por lo que el gen no se expresará. Pequeños cambios en los patrones de metilación de genes concretos provocarán el mismo efecto que una mutación, pero sin alteración del ADN.

Transcripción y procesamiento de los ARN.

Hoy sabemos que en las células de los organismos se transcribe un porcentaje sorprendentemente alto del genoma a ARN. En algunas de ellas, el 75% del genoma se transcribe en un momento

dado. En el hombre se transcribe el 62% del genoma y sólo el 5,5% representa exones. Se han identificado 62.403 lugares donde se inicia la transcripción en el genoma humano y muchos de los transcritos no funcionan ni como ARN mensajero (ARNm), ni ARN de transferencia (ARNt), ni ARN ribosomal (ARNr), son elementos reguladores de la expresión génica y son los ARN no codificantes (ARNnc). Estos ARNnc controlan muchos procesos biológicos, como proliferación, diferenciación y desarrollo, donde regulan la expresión génica a varios niveles incluyendo la modificación de la cromatina, transcripción, modificaciones en el ARN, cortes y empalmes en el ARN (ajuste), estabilidad del ARN y traducción. Los miARNs cooperan con factores de transcripción y se cree que en organismos superiores controlan la expresión génica de aproximadamente el 30% de los genes que codifican proteínas y el 1% de genes de todo tipo. En el ratón muchos transcritos se originan a partir de promotores alternativos, pueden sufrir ajustes variables o usar regiones de poliadenilación distintas. No sólo un gran porcentaje del genoma de mamíferos se transcribe, sino que además es frecuente que se realice de las dos hebras del ADN; en el ratón el 62% del genoma transcribe las dos hebras. Esto genera ARN de doble banda que es el sustrato para que se desencadene la respuesta mediada por ARNi y miARN. Los miARN desarrollan un papel especialmente relevante durante el desarrollo. En ratón y *Drosophila* los niveles de expresión de los genes de las dianas de los miARNs son mucho más bajos en los tejidos diferenciados que en el embrión, sugiriendo su implicación en los procesos de especialización tisular. Además, cada miARN puede regular un gran número de ARNm, y un solo ARNm puede ser diana de varios miARN. Pequeñas variaciones en la función de algún ARNi o miARN contribuirá en la variabilidad del organismo sin prácticamente contribuir a la divergencia genómica.

El mecanismo de ajuste extiende considerablemente la capacidad codificadora del ADN. El 40% de los genes de *Drosophila* y el 74% de los genes humanos se procesan por ajustes alternati-

vos dando lugar a productos ligeramente distintos en la mayoría de los casos, pero muy distintos ocasionalmente. Ya hemos descrito el caso de la miosina embrionaria y adulta o el caso extremo del gen *Dscam* (figura 1.12, *a*), donde el gen es el mismo en todas las células (conos del ojo), pero la combinatoria de los transcritos crea una enorme variabilidad que da lugar a 38.016 receptores de membrana ligeramente distintos. También en *Drosophila*, el sexo macho o hembra del animal, va asociado a distintos cortes y empalmes alternativos secuenciales. El gen *sxl* (del inglés, *sex lethal*) de la mosca posee 4 exones y 3 intrones, pero el exón 3 lleva un codón de terminación de cadena (figura 1.16). El gen se expresa siempre que la razón entre cromosomas X y juegos autosómicos sea 0,5 o 1 (en machos $XY + 2A \rightarrow 1/2 = 0,5$ y en hembras $XX + 2A \rightarrow 2/2 = 1$); la no expresión del gen es letal. El ajuste alternativo es el siguiente: en machos se ajusta el exón 1 con el 3, con lo que da un producto truncado de *sxl*; en hembras se ajusta el exón 1 con el 2 y con el 4 dando como producto la proteína Tra (del inglés, *transformer*) que es un regulador del ajuste de una cascada de proteínas que ejecutarán que ese organismo sea hembra.

1.8. DISTRIBUCIÓN DE TAREAS Y PORCENTAJES EN EL GENOMA HUMANO: METODOLOGÍA EMPLEADA EN SU DETERMINACIÓN

Si con los conocimientos actuales tratamos de asignar tareas concretas a porcentajes del genoma humano, nos encontramos con que el 3% son genes que se traducen a proteínas de todo tipo: estructurales, reguladoras, enzimas, y un 80% lo constituyen regiones implicadas directa o indirectamente en la regulación del 3% anterior, esto es: promotores (70.300), potenciadores (400.000), ARNsnc, dedos de zinc, terminadores, islas CpG etcétera. Del 17% restante, se podría aún asignar una parte a la adquisición y mantenimiento de estructuras y formas: centrómeros, telómeros, heterocromatina, etc., y el resto serían pseudogenes, restos de retrovirus, transposones inactivos, se-

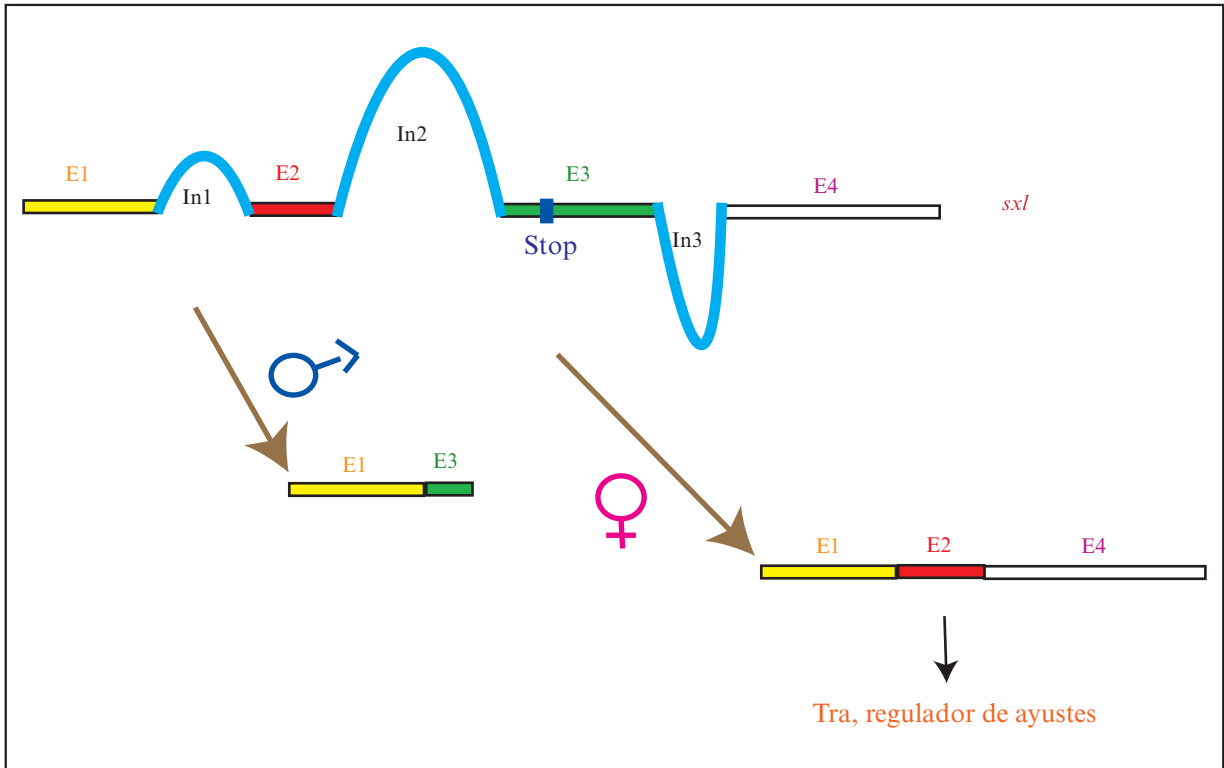


Figura 1.16. Ajuste alternativo en el gen *sxl* de *Drosophila*. Los machos ajustan el exón 1 con el 3 y las hembras el 1, 2 y 4. Sólo las hembras producen Tra que es un regulador del corte y empalme de muchos otros ARN.

cuencias altamente repetidas sin función conocida, o simplemente secuencias de función aún desconocida. Así, la mayor parte del ADN de cada célula está implicada, en funciones de control génico. La regulación de cada gen se vislumbra como un proceso altamente sofisticado que implica regiones próximas y lejanas al gen regulado, en una combinatoria única y dinámica que varía en el espacio y en el tiempo. No parece que el genoma tenga mucho ADN «basura», como se creía hasta hace no mucho tiempo.

Todos estos datos relativos a los porcentajes del genoma implicados en la síntesis de ARN, la síntesis de proteínas, la regulación de la expresión génica en los distintos tipos celulares o la adquisición de estructuras se han conseguido utilizando las siguientes técnicas:

1. El porcentaje de ADN que se transcribe a ARN se ha determinado por varios métodos: 1) el aislamiento y la secuenciación del ARN nuclear recién sintetizado; 2) la captura del extremo 5' metilado (con caperuza) y secuenciación de una pequeña parte de este extremo para confirmar los lugares de la iniciación de la transcripción (método CAGE); 3) la captura simultánea de ARN con caperuza y extremos 3' de poliA, lo que implica un ARN maduro y listo para viajar al citoplasma. Una vez aislados, se secuencian una parte de los dos extremos para confirmar la identidad del ARN (método ARN-PET).
2. Los lugares de control de la expresión génica en el ADN por unión de proteínas re-

guladoras, principalmente factores de transcripción, se han podido identificar utilizando también 3 métodos: 1) Inmunoprecipitación de la cromatina seguido de secuenciación (método ChIP-seq). Este método consiste en provocar una unión covalente entre la cromatina, el ADN y las proteínas que en un momento dado están unidas al ADN; a continuación se añade un anticuerpo específico contra un factor de transcripción concreto o una proteína de unión a la cromatina provocando su inmunoprecipitación. Seguidamente se secuencia el ADN arrastrado en la inmunoprecipitación. 2) Tratamiento con ADNasa I (figura 1.17), que produce cortes en regiones de cromatina activa, es decir, regiones en las que factores de transcripción sustituyen a la cromatina; este tratamiento se continúa con una secuenciación de las re-

giones hipersensibles a la ADNasa I (método ADNasa-seq). 3) Tratamiento con formaldehído y secuenciación (método FAIRE-seq). El tratamiento con formaldehído provoca uniones covalentes entre el ADN y las proteínas unidas a él. Además, en las regiones con nucleosomas y cromatina la proporción de ADN es pequeña en comparación con la de proteínas, mientras que en las regiones donde la cromatina está sustituida por factores de transcripción, la proporción de ADN es superior a la de proteínas. Una extracción fenólica subsecuente al tratamiento con formaldehído arrastrará las proteínas libres y las uniones fuertes ADN-cromatina a la fase fenólica, mientras que el ADN desnudo o con pocas proteínas unidas se quedará en la fase acuosa, de donde el ADN es rescatado y secuenciado.

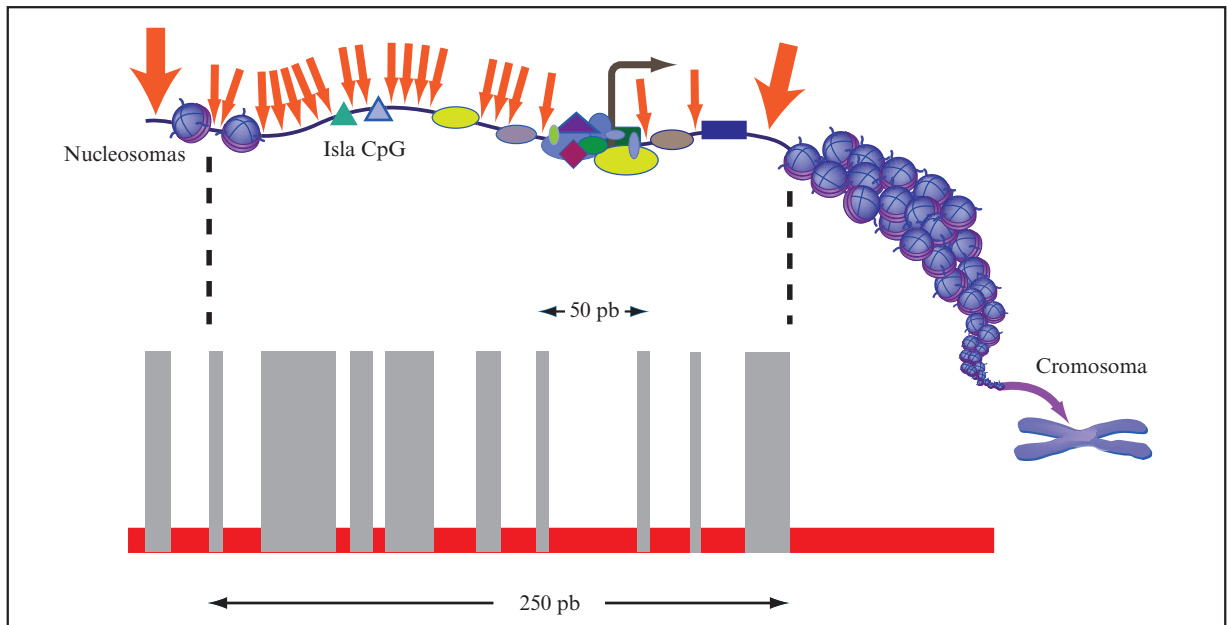


Figura 1.17. Tratamiento con ADNasa I. Con un primer tratamiento a núcleos intactos obtendríamos fragmentos de unos 250 pb en regiones de cromatina abierta (flechas anchas). Estos fragmentos pueden ser secuenciados o vueltos a tratar con mayores concentraciones de ADNasa I (flechas estrechas) para delimitar los espacios protegidos y resistentes al ataque, debido a la presencia de factores de transcripción. En la parte inferior, el análisis generalizado y la huella protectora (barras) debida a la presencia de los distintos factores de transcripción y ARN polimerasas.

3. Las regiones con dinucleótidos CpG no metilados se asocian con ADN activo, mientras que si éstos están metilados implica silenciamiento génico. Para determinar el grado de metilación de regiones del ADN se emplea el método RRBS (del inglés: *reduced representation bisulfite sequencing*) que consiste en un primer tratamiento del ADN total y genómico con una enzima de restricción que tenga CpG en la diana de reconocimiento (Bgl II, por ejemplo) y selección de una fracción de fragmentos relativamente pequeños (500-600 pb) que estarán, de manera natural, enriquecidos en islas CpG. El tratamiento con bisulfito de una parte de esta fracción de ADN desnaturalizada convertirá las

citiosinas no metiladas en uracilo, que ahora ya no apareará con la guanina, sino con adenina; la amplificación por la PCR (del inglés: *polymerase chain reaction*), la secuenciación posterior y comparación con la secuencia no tratada con bisulfito, nos dirá el estatus de metilación que tenía el ADN secuenciado porque el par C/G no metilado se convertirá en A/T, mientras que el metilado permanece como tal.

4. Regiones que se traducen a proteínas. La presencia de islas CpG y regiones con pautas de lectura abiertas sin codones de terminación anuncian durante la secuenciación la presencia de una región que codifica una proteína.

LECTURAS COMPLEMENTARIAS

Biología molecular e ingeniería genética (2.^a edición) de Ángel Herráez. Elsevier España. 2012.

Cálculo en biología molecular y biotecnología (edición en español de la 2.^a edición en inglés *Calculations for molecular biology and biotechnology: a guide to mathematics in the laboratory*). Frank H. Stephenson, traducción de Jorge Lloberas Cavero y Annabel Valledor Fernández. Elsevier España. 2012.

Chromatin proteins: key responders to stress. Karen T. Smith, Jerry L. Workman. *PLOS Biology*, 10 e1001371. 2012.

Encode explained revisión de Josep R. Ecker sobre 7 artículos publicados en la revista *Nature*, vol. 489, 2012.

Pseudogenes. Yusuf Tutar en *Comparative and functional genomes*. Artículo de revisión. Hindawi Publishing Corporation. Doi:10.1155/2012/424526. 2012.

The human encyclopaedia revisión de Brenda Maher sobre 7 artículos publicados en la revista *Nature*, vol. 489, 2012.

Tubulin post-translational modifications: encoding functions on the neuronal microtubule cytoskeleton, revisión de Carsten Janke y Matthias Kneussel, *Trends in neurosciences*, 33, 362-372. 2010.

La recombinación es el intercambio de material genético entre dos moléculas de ADN de procedencia dispar. El proceso de recombinación y separación de los cromosomas homólogos de forma independiente está íntimamente ligado a la evolución hasta el punto de que es difícil concebir ésta sin intercambio de material genético. Dentro de un mismo cromosoma, en ausencia de recombinación podría ser complicada la separación de alteraciones favorables de las deletéreas y así probar el posible efecto beneficioso de las primeras. La recombinación, en suma, está lanzando continuamente nuevas proposiciones de dentro a fuera, después, por selección natural, se quedarán las beneficiosas y se eliminarán las desfavorables.

Podemos distinguir tres tipos de recombinación:

- **Recombinación homóloga.** En organismos superiores coincide con la primera profase meiótica que tiene lugar durante la formación de los gametos masculinos (espermatogénesis) y femeninos (oogénesis). La homología entre las secuencias que participan es imprescindible. La meiosis reduce la dotación cromosómica de las células diploides a la mitad (haploides) en preparación para la fecundación. Esta reducción ocurre cuando los cromosomas homólogos se dirigen a polos opuestos de la célula durante la primera división meiótica después de haber intercambiado material genético en la recombinación. Cada cromosoma homólogo, recombinado o no, irá a un gameto diferente.

La probabilidad de que dos regiones cromosómicas (*loci*) recombinen dependerá de la distancia entre ellas, siendo directamente proporcional. Si dos *loci* están muy próximos en el cromosoma, es difícil que se produzca un sobrecruzamiento entre ambos y se considera que están *ligados*. Si, por el contrario, los *loci* están físicamente muy separados dentro del mismo cromosoma, la probabilidad de que ocurra una recombinación en posiciones intermedias es muy alta y, por tanto, de que se hereden juntos muy baja.

- **La recombinación específica de sitio** tiene lugar entre secuencias específicas; la homología, de existir, es secundaria. Tiene lugar, por ejemplo, entre un lugar concreto en el genoma del fago λ y un lugar específico de la bacteria o un sitio concreto de un virus adenoasociado y el cromosoma 19 humano, es un mecanismo de integración del genoma viral en el de la célula infectada.
- **Recombinación al azar.** Unión de dos moléculas de ADN distintas. El ejemplo más característico es la transposición: movimiento de unas secuencias de un lugar a otro dentro del mismo genoma.

2.1. RECOMBINACIÓN ENTRE CROMOSOMAS HOMÓLOGOS

Es el intercambio de ADN realizado por enzimas que utilizan como sustrato dos moléculas de

ADN de doble hebra, homólogos. En mamíferos, aunque la frecuencia de recombinación debería ser constante a lo largo de toda la región de homología, hay desviaciones dependiendo de diversos efectos locales y globales. Por ejemplo, la recombinación está ausente de las regiones de heterocromatina muy condensada, y en el ser humano, la frecuencia global de recombinación es el doble en la ovogénesis que en la espermatogénesis; en *Drosophila*, la frecuencia de recombinación en machos es prácticamente nula. Una frecuencia de recombinación normal sería de 1 o 2 entrecruzamientos por cada par de cromosomas homólogos; parece existir una cierta interferencia que impide la presencia de un entrecruzamiento en las proximidades de otro. Se denomina *quiasma* al lugar en el que las hebras de ADN se entrecruzan y se inicia el intercambio. En mamíferos la recombinación implica ruptura y unión de dos cadenas de ADN de banda doble, pertenecientes a dos cromátidas no hermanas de dos cromosomas homólogos.

2.1.1. La recombinación homóloga en procariotas

El intercambio de secuencias entre moléculas de ADN homólogas se describe por primera vez en bacterias (*Escherichia coli*), pero la recombinación en este organismo no es enteramente equivalente a la que tiene lugar entre cromosomas de organismos superiores durante la meiosis. Las bacterias son organismos haploides y para que tenga lugar una recombinación homóloga es imprescindible la presencia de dos moléculas de ADN de doble hebra cada una; así, en estos organismos, el intercambio de material genético se realiza fundamentalmente en dos circunstancias: 1) después de la conjugación, es decir, de la transferencia de ADN de una bacteria a otra, 2) inmediatamente después de la replicación y durante la reparación de huecos en el ADN creados por fallos en la replicación. La recombinación también puede tener lugar entre genomas de bacteriófagos (el fago λ por ejemplo) después de su replicación en el seno de la bacteria infectada. En todos los

casos han de generarse extremos 3' libres y regiones de hebra sencilla.

La conjugación es un proceso bacteriano de transferencia de material genético mediado por los factores de fertilidad F. Estos factores son realmente episomas, es decir, replicones autónomos que pueden, bien mantenerse en el citoplasma, o integrarse en el cromosoma de la bacteria (figura 2.1). La cepa de *E. coli* que posee el factor se denomina Hfr F^+ (alta frecuencia de recombinación, del inglés, *high frequency of recombination*), y la cepa receptora que carece del factor, F^- . Cuando el factor F está integrado en el cromosoma, puede producirse una ruptura en una de las dos hebras del factor en una región específica denominada: ori T (origen de la transferencia) situada cerca de un extremo del episoma. La ruptura origina un extremo 5' fosfato que va a liderar la transferencia llevándose enganchado el cromosoma bacteriano. El ADN que se traslada de una célula a otra es de hebra sencilla, durante la transferencia se produce la duplicación de la hebra de ADN y, ya en la célula receptora, el ADN transferido crea un diploide parcial (merozigoto) que será un magnífico sustrato para la recombinación homóloga. El ADN que permanece en la célula donante también pasa a hebra doble. Los genes necesarios para la movilización están en el plásmido (*tra*, de transferencia). La transmisión se realiza a través de los denominados pelos sexuales o *pili* producidos por las cepas Hfr F^+ . Los canales de conjugación son observables al microscopio electrónico. Las cepas receptoras F^- se convierten en F^+ si reciben una copia completa del episoma, lo cual no ocurre con una frecuencia muy elevada, ya que una parte importante de F va a la cola de la transferencia y los *pili* no son muy resistentes, con lo que la interrupción del proceso ocurre a menudo. Los episomas incapaces de integrarse en el cromosoma bacteriano son, por definición, los plásmidos propiamente dichos. La región de ADN de la célula bacteriana receptora, que es sustituida por la homóloga donante durante la recombinación, no sobrevivirá y será eliminada por la célula.

Inmediatamente después de la replicación del ADN en bacterias y antes de que se divida la célula bacteriana, tenemos también dos moléculas

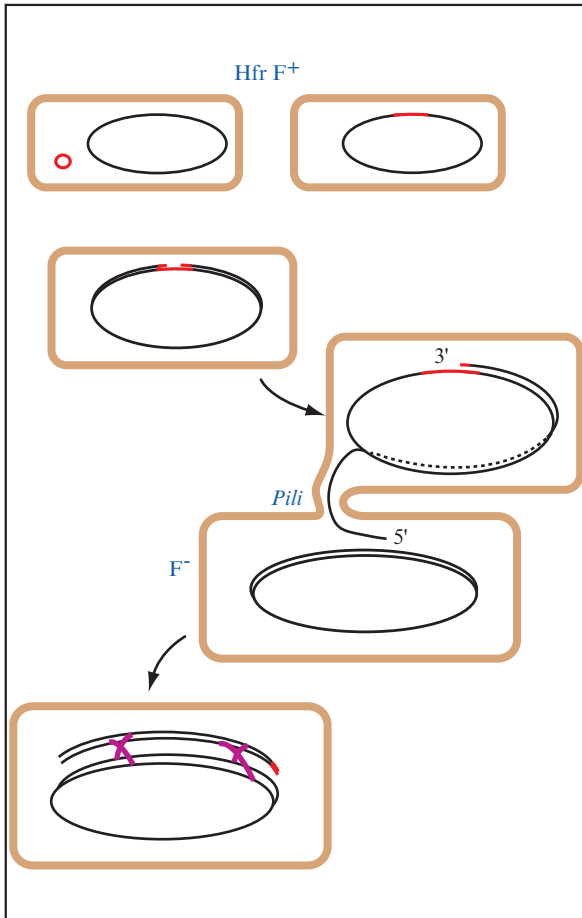


Figura 2.1. Conjugación mediada por los factores de fertilidad F. En la parte superior: formas libre e integrada de un episoma que en este caso es el factor de fertilidad F. Más abajo, ruptura de una de las dos hebras del factor (en color) y transferencia del episoma movilizándolo al cromosoma de la bacteria en forma de hebra sencilla, al llegar a la bacteria receptora se convertirá en ADN de doble banda y tendrá lugar la recombinación homóloga con el cromosoma de la bacteria.

de ADN de dos hebras cada una, sustrato de la recombinación.

Existen dos modelos para explicar la recombinación homóloga en bacterias: el modelo de Holliday y el de Meselson-Radding. En ambos casos participan dos moléculas de ADN de distinta pro-

cedencia, pero homólogas (que pueden aparear). **El modelo de Holliday** (Robin Holliday, 1964) postula que la recombinación se inicia con la ruptura de dos de las cuatro hebras de ADN de hebra sencilla que participan. El corte ocurre aproximadamente a la misma altura y las hebras implicadas poseen la misma polaridad (figura 2.2). Seguidamente cada hebra rota deja a su otra mitad, cruza y se aparea con su complementaria del ADN homólogo, desplazando a su equivalente. El intercambio de hebras es recíproco y se van generando moléculas híbridas (heterodúplexes) a medida que el quiasma avanza en una dirección. El desplazamiento puede ocurrir en cualquiera de las dos direcciones posibles durante una distancia muy variable. El paso siguiente es una segunda ruptura doble con dos opciones: 1) que participen las mismas hebras que han sufrido el primer corte o 2) que participen las dos hebras que aún permanecen intactas. En el primer caso volvemos al estado inicial devolviendo a su sitio las hebras que habían intercambiado pareja y como resultado tendremos dos heterodúplexes, que no recombinantes, ya que sólo una de las dos hebras ha participado en el intercambio. Por el contrario, si las rupturas afectan a las hebras aún intactas, se repetirán los eventos ya explicados para las dos hebras intercambiadas y tendremos un segundo quiasma que, de ir en la misma dirección que el primero, irá generando un intercambio recíproco de las dos hebras de ADN de un cromosoma con las de su homólogo, hasta el final del cromosoma. El resultado es una recombinación propiamente dicha.

En algunos plásmidos y adenovirus, fotografiados al microscopio electrónico en el momento de la recombinación homóloga, se aprecia una estructura de forma cruciforme denominada la enrucijada de Holliday. La estructura representa el resultado de una apertura en las hebras que están intercambiando material genético, seguido de una rotación. La representación de esta estructura en el plano, y dependiendo de que las segundas rupturas ocurran en horizontal o en vertical, nos permite visualizar el resultado previsto cuando se implican las mismas hebras de la ruptura inicial o las que permanecían intactas (figura 2.3). En el pri-

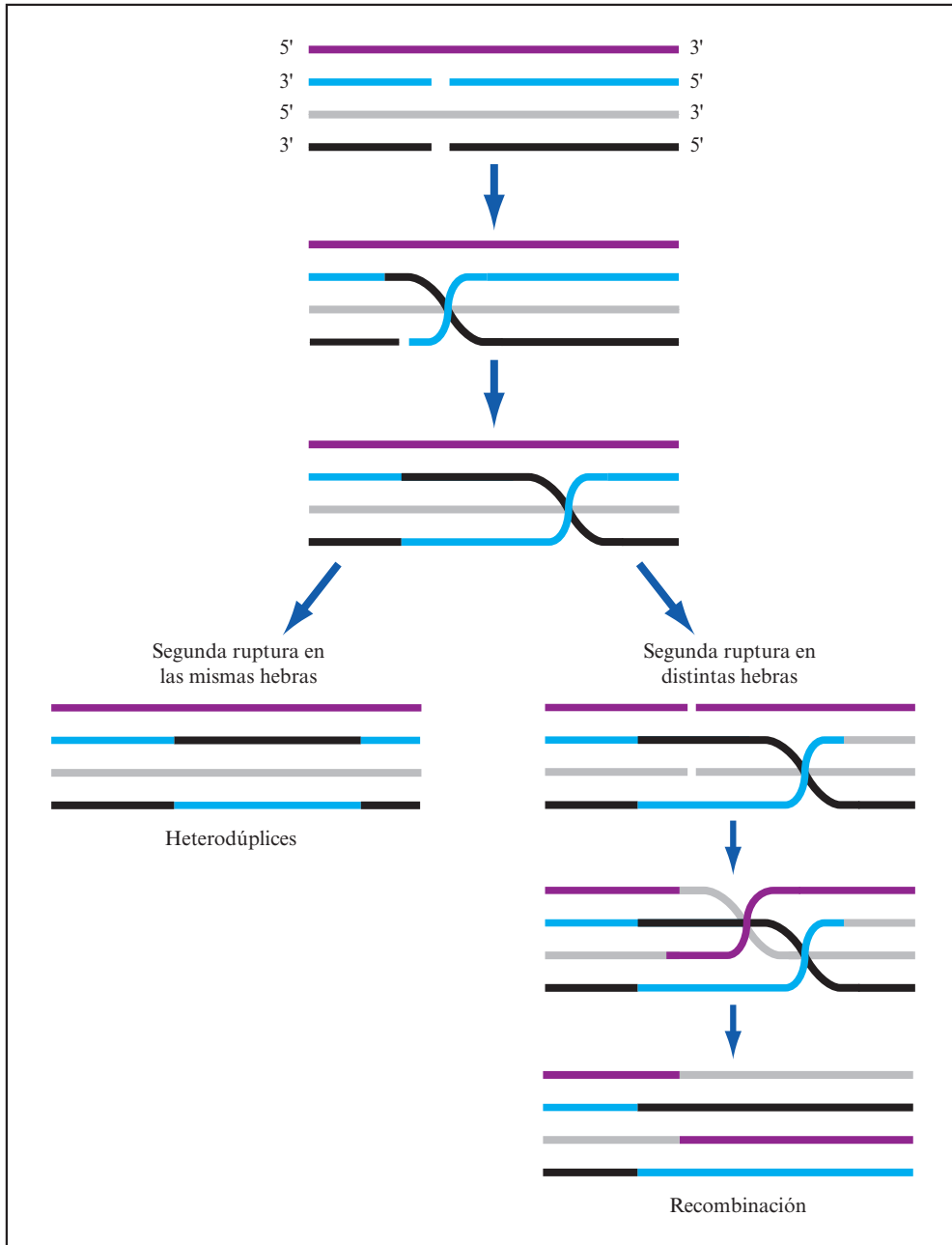


Figura 2.2. Modelo de Holliday de recombinación en procariotas. Proceso de iniciación mediado por rupturas en dos hebras procedentes de cromosomas distintos manteniendo la misma polaridad ($3' \rightarrow 5'$). La segunda ruptura puede dar lugar a heterodúplexes (lado izquierdo, abajo) o recombinantes (lado derecho, abajo), según que tenga lugar en las mismas hebras o no que la primera ruptura, respectivamente. Se muestra en el lado derecho la migración, hasta el final del cromosoma, de los dos quiasmas generados con las rupturas en las cuatro hebras.

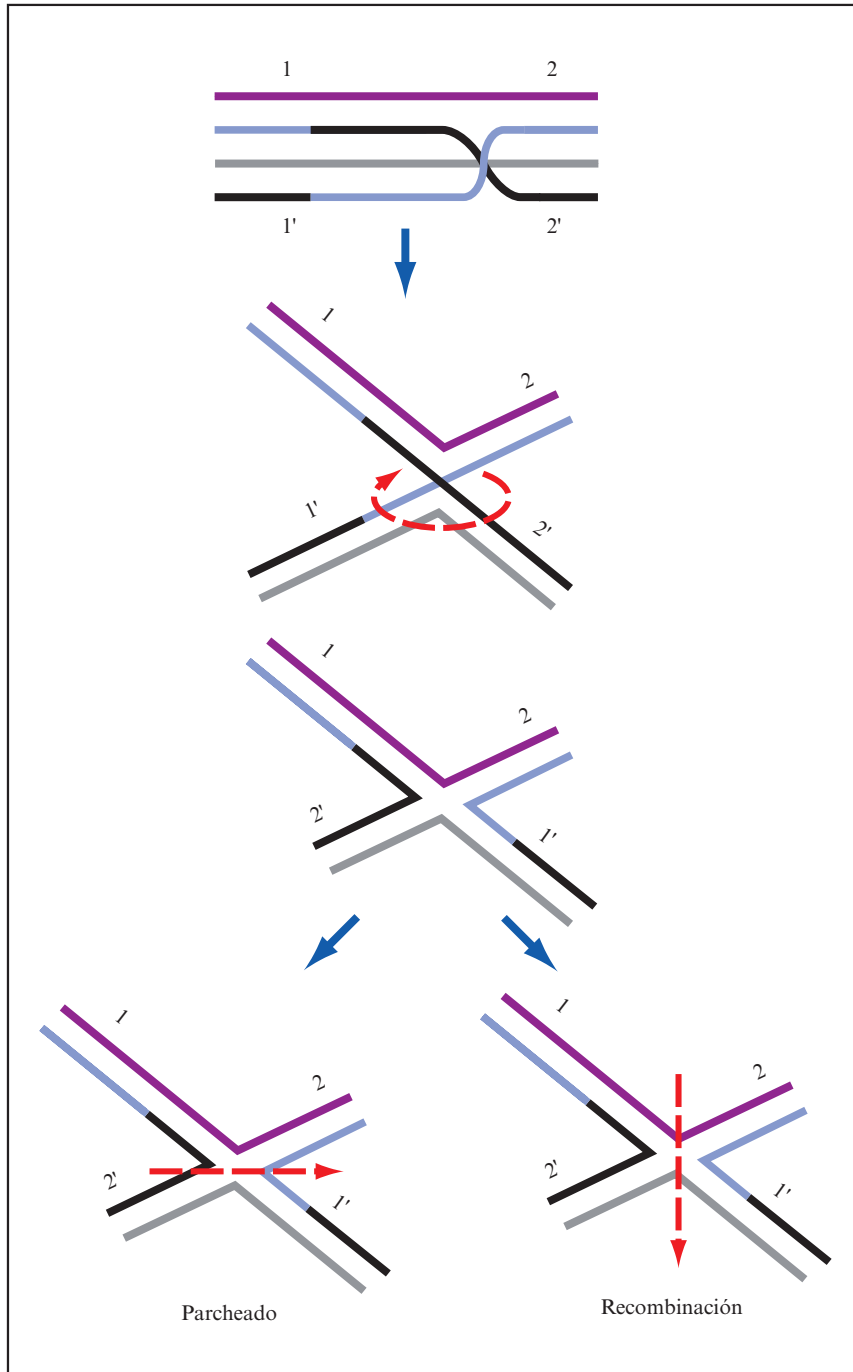


Figura 2.3. Esquema en el plano de la encrucijada de Holliday. La ruptura en horizontal generaría heterodúplexes (parcheo), mientras que los cortes en vertical darían lugar a recombinantes.

mer caso se liberarían heterodúplices, y en el segundo, recombinantes.

El modelo de Meselson-Radding supera uno de los puntos más débiles del modelo de Holliday que es el doble corte inicial a la misma altura en hebras de ADN homólogas. Este modelo postula una sola mella inicial en el ADN: el extremo 3'OH generado es enganchado por la ADNpol que realiza síntesis de ADN, mientras que el extremo 5'P libre invade y ataca la cadena contraria no cortada (figura 2.4a). Se forma un bucle al desplazar la cadena autóctona que se elimina seguidamente al formarse el quiasma y ligarse la hebra invasora. En este punto el modelo prosigue, en coincidencia con la propuesta de Holliday, con la formación de la encrucijada y resolución con un segundo doble corte en horizontal (parchado) o vertical (recombinación). Además, las bacterias utilizan la recombinación como mecanismo de reparación en caso de huecos en una de las hebras, rescatando el material genético de la hebra complementaria intacta (figura 2.4b). Este mecanismo de recombinación/reparación se acerca al modelo de Meselson-Radding más que al de Holliday, ya que requiere síntesis de ADN. Una vez que se ha formado el quiasma y avanza, ya no hay distinción entre los distintos modelos. Las dos opciones de resolución son comunes, según que la segunda ruptura afecte a las mismas hebras que la primera o no.

En el proceso de recombinación de procariontes intervienen una serie de enzimas sin cuya colaboración no sería posible llevar a cabo esta compleja maquinaria de intercambio de material genético. En bacterias se han descrito entre 10 y 20 *loci* cuyas mutaciones generan individuos deficientes en recombinación (rec^-). Participan enzimas generales como polimerasas de ADN (ADNpol), ligasas de ADN, topoisomerasas y proteínas de unión a ADN de hebra simple (ssb, del inglés *single stranded binding*), así como otras más específicas de la recombinación como las Rec A, Rec BCD o Rec G, cuyas letras implican el orden en el que se descubrieron, no siempre coincidente con el orden de actuación en el proceso de recombinación. También intervienen enzimas que se descubrieron en estudios de reparación del ADN como las Ruv AB y Ruv C.

Las ADNpol son enzimas que sintetizan una hebra de ADN nueva complementaria a otra que funciona de molde. Necesitan un extremo 3'OH libre donde agarrarse y extienden este extremo añadiendo nucleótidos uno a uno y complementarios a la hebra que funciona como molde. En *E. coli* se han descrito 5 ADNpol de las cuales la pol I es la que mayoritariamente participa en los procesos de recombinación y reparación de huecos en el ADN. Las ligasas unen extremos 3'OH con 5'P (fosfato). Las topoisomerasas catalizan cambios en la topología del ADN provocando rupturas transitorias en una o las dos hebras del ADN, pasando la hebra/s por el hueco y re-ligando las hebras. Las topoisomerasas de tipo I realizan una ruptura transitoria en una de las dos hebras del ADN, mientras que las topoisomerasas de tipo II actúan introduciendo dos mellas transitorias en las dos hebras del ADN. Las ssbs tienen gran afinidad por ADN de hebra sencilla impidiendo que éste aparezca con su complementario para formar una doble hélice. Actúan de forma cooperativa de modo que la unión de nuevos monómeros se favorece.

En la recombinación, la primera que actúa es Rec BCD y su misión parece ser la de crear regiones de ADN de hebra sencilla con extremos libres 3'. Este complejo enzimático de varias subunidades (B, C y D) se introduce en el ADN allí donde haya una doble ruptura y se desplaza a una velocidad de unas 1.000 pb/segundo en la dirección 5' → 3' hasta encontrar una secuencia *chi* (del inglés *cross-over hot spot instigador*: 5'GCTGGTGG3'), donde se detiene unos 10 segundos. Las secuencias *chi* favorecen la recombinación, por lo que se consideran puntos calientes recombinacionales, se encuentran por puro azar cada 5 a 10 kb pero no son imprescindibles. Rec BCD tiene varias actividades: ATPasa que le permite desplazarse, helicasa con la que desenrolla el ADN, exonucleasa V con la que puede ir eliminando nucleótidos a partir de un extremo 3' libre del ADN y endonucleasa con la que puede cortar una hebra del ADN en cualquier punto interno. Así Rec BCD se engancha a una doble mella en el ADN y se desplaza desenrollando y desapareando el ADN con su actividad helicasa; con su actividad

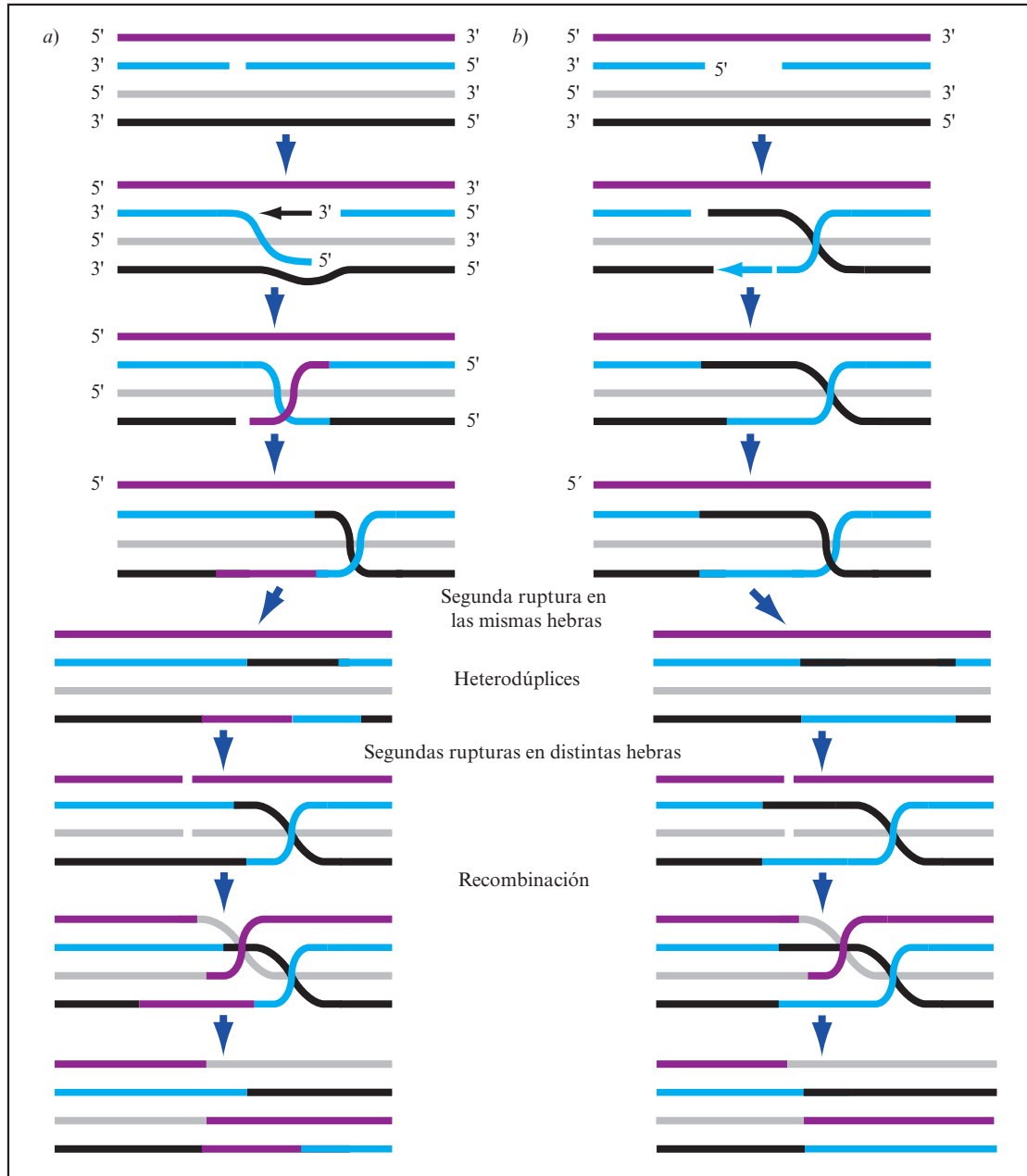


Figura 2.4. Modelo de Meselson-Radding de recombinación en procariotas. *a)* Proceso de iniciación mediante una sola mecha en una de las hebras del ADN. Síntesis y desplazamiento de la hebra cortada, invasión y primera ruptura de la segunda hebra, rotación y encrucijada de Holliday con las segundas rupturas que puede dar lugar a heterodúplexes o recombinantes (abajo), según que tenga lugar en las mismas hebras o no que la primera ruptura, respectivamente. *b)* Reparación/recombinación de un hueco generado durante la replicación del ADN, a partir de la formación del primer quiasma el proceso es como *a)*.

exonucleasa degrada parcialmente el ADN de hebra sencilla con el extremo 3'. Al llegar a una secuencia *chi*, se detiene y rompe la hebra superior del ADN a una distancia entre 4 y 6 bases de *chi*. Rec BCD reconoce la hebra superior de la secuencia *chi* desapareada y libera la subunidad D (nu-

cleasa), Rec BC sigue avanzando más despacio como helicasa (figura 2.5).

La presencia del extremo 3' libre permite e incentiva la llegada de Rec A que se une al extremo. Rec A facilita la invasión de una molécula de hebra simple a otra de doble hebra con desplazamiento

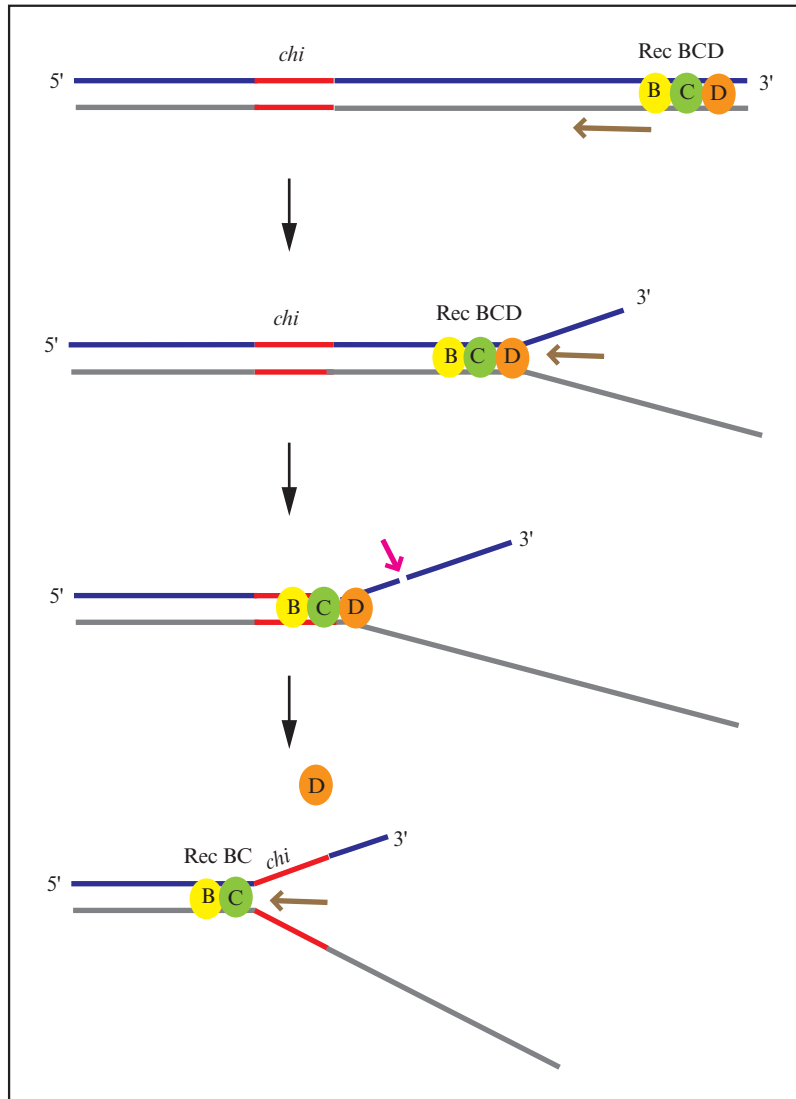


Figura 2.5. Reacciones coordinadas de las diversas actividades de Rec BCD en un sustrato de ADN cercano a una secuencia *chi*. Rec BCD se acerca a una secuencia *chi* degradando parcialmente el extremo 3'. Al llegar, realiza un corte endonucleolítico, pierde Rec D y continúa sólo con la actividad helicasa.

de su idéntica. Seguidamente interviene el sistema Ruv (enzimas de reparación de los daños provocados por luz ultravioleta) que actúa y resuelve la encrucijada de Holliday. Ruv A, parte del complejo Ruv AB, reconoce los quiasmas y cada una de sus 4 subunidades se une a una hebra del ADN. Ruv B es un exámero que forma una especie de anillo alrededor de la molécula de ADN, es en realidad una helicasa con actividad ATPasa que facilita el desplazamiento del punto de intersección de los cromosomas (figura 2.6). La presencia del complejo Ruv AB permite al quiasma viajar a una velocidad de 10-20 pb por segundo y desplazar a las moléculas de Rec A que pudieran quedar en el ADN. Otra helicasa, Rec G, puede suplir las actividades de Ruv AB como se observa en mutantes de este último complejo, pero mutantes en ambos Ruv AB y Rec G son incapaces de recombinar.

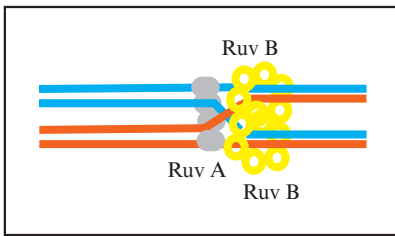


Figura 2.6. Actuación de Ruv A y Ruv B durante la movilidad del quiasma en la recombinación.

La endonucleasa Ruv C reconoce específicamente encrucijadas de tipo Holliday y produce los cortes que resuelven el proceso, pero no es la única enzima que lo puede hacer. La enzima UvrD de *E. coli* es una helicasa implicada en la reparación de apareamientos incorrectos, que también puede resolver encrucijadas de Holliday; asimismo, la endonucleasa I de T7 o la resolvasa del bacteriófago SPP1 que infecta *Bacillus subtilis*, son eficientes resolviendo este tipo de encrucijadas.

2.1.2. La recombinación homóloga en eucariotas

La recombinación homóloga en eucariotas es muy importante durante la meiosis y gametogéne-

sis, pero también ocurre en células somáticas durante la mitosis como mecanismo de reparación cuando uno de los cromosomas homólogos está dañado; en estas circunstancias el ADN reparador procederá de la cromátida hermana.

Durante la meiosis, la recombinación homóloga es responsable del intercambio de material genético entre los cromosomas paternos y maternos promoviendo la diversidad genética de la progenie y facilitando que en un mismo cromosoma pueda haber genes paternos y maternos que se transmitirán ligados a la descendencia. También facilita una correcta segregación de los cromosomas que han recombinado.

En eucariotas la iniciación del proceso de reparación/recombinación ocurre cuando existe una ruptura en cada una de las dos hebras de una misma cromátida. Esta doble mella es imprescindible para la relajación y desenrollamiento de las hebras que participarán en el intercambio. A la ruptura le sigue una digestión de los extremos 5' que genera extremos 3' protuberantes. Uno de estos extremos invadirá y desplazará su idéntica en la cromátida hermana permitiendo síntesis de ADN utilizando la complementaria a la hebra invasora como molde. Se forma un lazo en forma de una D y se pueden seguir dos rutas: una de estabilización del lazo D con la formación de quiasmas, y otra de desplazamiento con reparación pero no intercambio (figura 2.7). En el primer caso, el segundo extremo 3'OH creado inicia la síntesis de ADN utilizando como molde la hebra desplazada por la invasora, a la vez que la prolongación de la invasora se engancha al extremo 5' resultado de la digestión original. Se forman dos quiasmas que viajarán en la misma o diferente dirección y que se resolverán como encrucijadas de Holliday. Así, en cada caso, puede haber dos nuevas posibilidades: que el segundo corte ocurra en las mismas hebras donde ocurrió el primer corte generando heterodúplexes o que el segundo corte ocurra en hebras distintas generando recombinantes por el corte y unión de estas hebras con desplazamiento hasta el final de los cromosomas. Siempre que al menos uno de los dos quiasmas genere corte y unión de las dos hebras no implicadas en el primer corte, se gene-

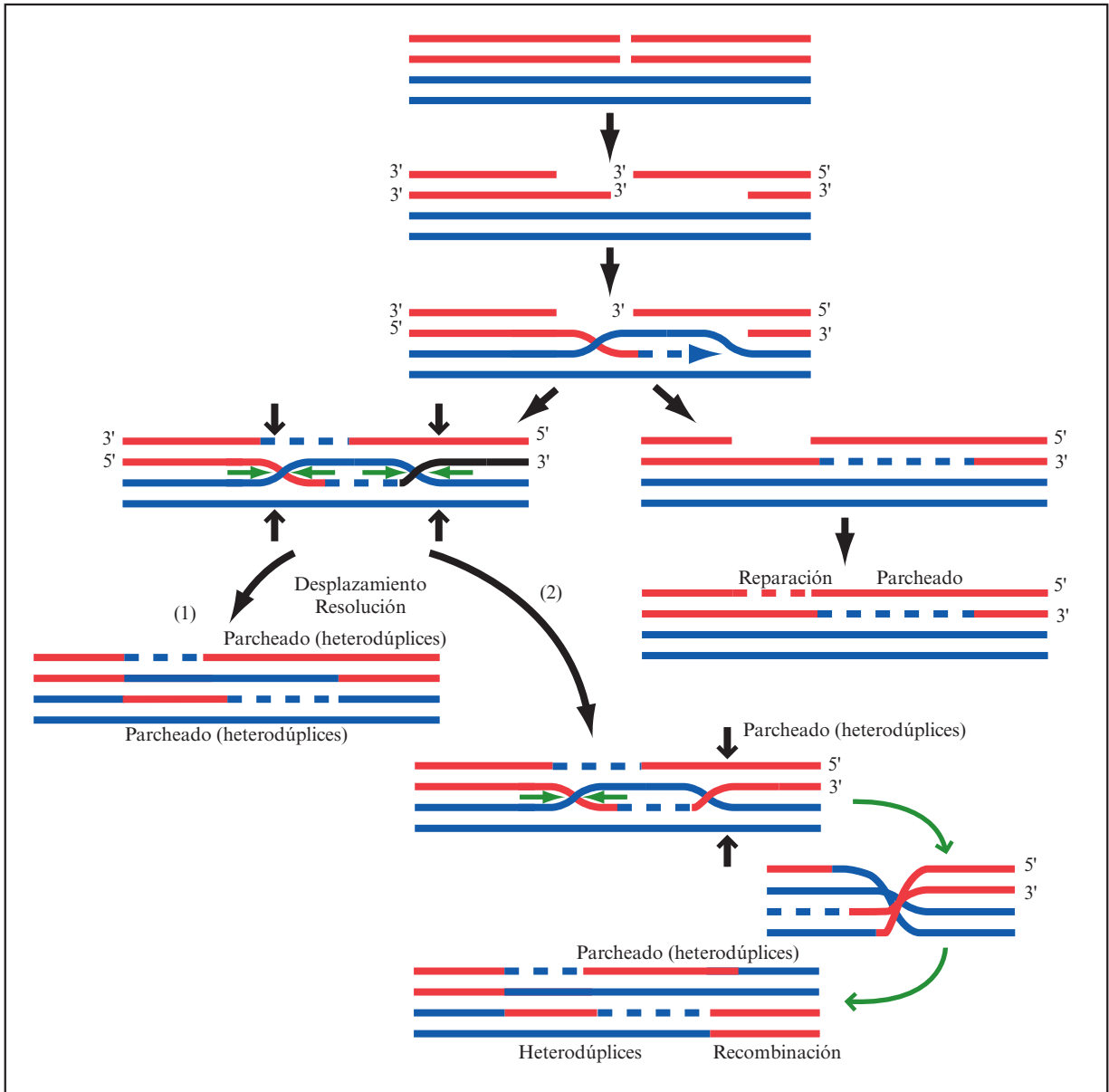


Figura 2.7. Posibles rutas en la recombinação homóloga de eucariotas. Ruptura en las dos hebras del ADN y digestión de los extremos 5' agrandando el hueco. Invasión de uno de los extremos 3'OH y formación del lazo D. Por el ramal izquierdo formación de dos quiasmas, desplazamiento y resolución con cortes en las mismas hebras en las que ocurrieron los primeros cortes (1) o en las otras hebras (2) (las flechas indican posibilidades de corte). En el caso (1) se generan heterodúplexes, en el (2) se asume que en el quiasma de la izquierda el segundo corte mediado por la encrucijada de Holliday ocurre en las mismas hebras en las que ocurrió el primero, pero en el quiasma de la derecha el corte y unión tiene lugar en las hebras que no habían participado en las primeras mellas (las flechas indican los lugares de corte), produciéndose una verdadera recombinação. Por el ramal derecho, síntesis y unión con ligasa que genera reparación y un parcheado.

rá un recombinante. Alternativamente, la hebra invasora se puede desplazar y acabar enganchándose al extremo 5' resultado de la digestión original, al tiempo que síntesis de ADN termina de rellenar el hueco creado en la hebra superior. Así tendremos reparación y parcheado, pero no quiasmas ni recombinación.

El parcheado son regiones de heterodúplexes que, al no aparearse correctamente, pueden requerir la acción de las enzimas de reparación que a su vez pueden elegir como molde cualquiera de las dos hebras, por lo que pueden producir recombinaciones en pequeñas regiones del heterodúplex o incluso conversión génica, es decir, regiones en las que uno de los alelos se ha convertido en el homólogo (figura 2.8).

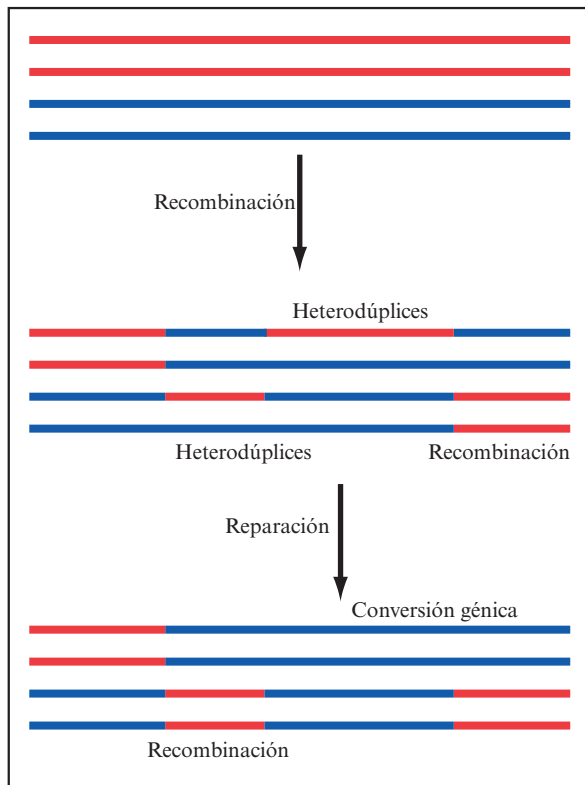


Figura 2.8. Posibles maneras de reparar un parcheado. Como resultado de la reparación se ha producido una nueva recombinación y una conversión génica.

Proteínas implicadas en el proceso de recombinación en eucariotas

El factor inicial y desencadenante de la recombinación es una doble ruptura, una en cada hebra de una de las dos cromátidas del ADN. Si bien estas rupturas pueden aparecer esporádicamente al azar durante la replicación en la mitosis, durante la meiosis son imprescindibles para iniciar una recombinación que es obligatoria en el proceso de gametogénesis y que no se puede confiar al azar, ya que la meiosis no proseguiría en ausencia de recombinación. Así pues, en meiosis la proteína Spo 11 autolesiona su propio ADN con rupturas en las dos hebras de una de las cromátidas, acción que desencadenará la recombinación. La ruptura se produce en regiones de cromatina accesible aparentemente al azar aunque es improbable que los extremos 3' sean guanina. A continuación actúa el complejo MRX (Mre11-Rad50-Xrs2) que degrada los extremos 5' y origina extremos protuberantes 3' de unos 100 nucleótidos protegidos por la proteína de replicación A (RPA, del inglés *replication protein A*) que alisa el ADN eliminando estructuras secundarias. Seguidamente interviene Rad 51 que se asocia con uno de los extremos 3' y facilita la invasión de la cromátida hermana desplazando a su idéntica. Toda una familia de proteínas asociadas a Rad 51, como Rad 52 y Rad 55-Rad 57, ayudan en esta misión. Realizada la tarea, Rad 54 y otras proteínas desplazan a Rad 51 del escenario dando paso a enzimas de síntesis de ADN. El papel que tiene la mayoría de estas enzimas se descubrió obteniendo mutantes y viendo la función bloqueada durante la meiosis de *Saccharomyces cerevisiae*. Rad 51 en levaduras es equivalente a Rec A en bacterias y Dmcl sería la equivalente en *Drosophila melanogaster* y otros organismos superiores, pero Dmcl actúa exclusivamente en meiosis, no en reparación. En la meiosis de eucariotas también se han identificado varias resolvasas de encrucijadas tipo Holliday, como Mus 81 y Yen 1 presentes en levaduras y mamíferos.

Además de todas las enzimas descritas en este apartado, son también importantes, aunque no son específicas únicamente de la recombinación,

las topoisomerasas, enzimas encargadas de cambiar la topología del ADN mediante rupturas transitorias de una o las dos hebras del ADN. Creado el hueco, pasan la/s hebra/s intactas a través de dicho hueco y vuelven a unir las moléculas cortadas. Así relajan las superestructuras cromatínicas. Las de tipo I provocan una mella en una de las hebras del ADN permitiendo una rotación o desenrollamiento sobre la fija para seguidamente volver a unir los extremos cortados; las de tipo II cortan en las dos hebras permitiendo el paso de otro ADN de doble hebra. Las topoisomerasas son imprescindibles para desenrollar el ADN y permitir que se lleven a cabo funciones tan importantes como la replicación, transcripción o recombinación.

Citología de la recombinación en la meiosis eucariótica

Durante las mitosis, pero muy especialmente en la primera profase meiótica, los cromosomas se asocian de dos en dos cada uno con su homólogo, lo que implica un reconocimiento y distinción de las cromátidas hermanas frente al resto. Esta asociación culmina en la meiosis con la aparición de una estructura muy prominente denominada el complejo sinaptonémico, que media en la conexión íntima de los homólogos en el estadio de paquitene. El complejo sinaptonémico es una estructura

eminentemente proteica parecida a una cremallera con dos elementos axiales laterales donde se localizan los ADN de las cromátidas y un elemento central rico en proteínas y filamentos transversales (figura 2.9). Los elementos axiales de los complejos sinaptonémicos de mamíferos (ratón y mono) están formados por cohexinas, unas proteínas involucradas en la separación de las cromátidas y ADN. Respecto al ADN, los complejos sinaptonémicos parecen estar enriquecidos en secuencias repetidas SINE (apartado 1.3, 1.^{er} capítulo), por lo que quizá estas secuencias sean como puntos de anclaje de la estructura sinaptonémica. También hay evidencia de que el ADN en estos organismos se une al elemento axial correspondiente a través de la proteína SYCP3, una proteína que, junto a SYCP1 y SYCP2, son constituyentes del propio elemento lateral. El elemento central es fundamentalmente proteico con filamentos transversales, pero sin evidencia clara de ADN asociado. El grupo de proteínas ZMM está directamente implicado en la formación y mantenimiento del elemento central. El grupo incluye al menos 7 proteínas distintas: Zip1 (del inglés *zip*, cremallera) que es un componente integral y mayoritario del elemento central, Zip2, Zip3 y Zip4 que facilitan interacciones entre proteínas; Mer3, Msh4 y Msh 5 que se cree facilitan el proceso de la recombinación, pero no se conocen los detalles de su modo de acción. En contra de las apariencias, el inicio de la recom-

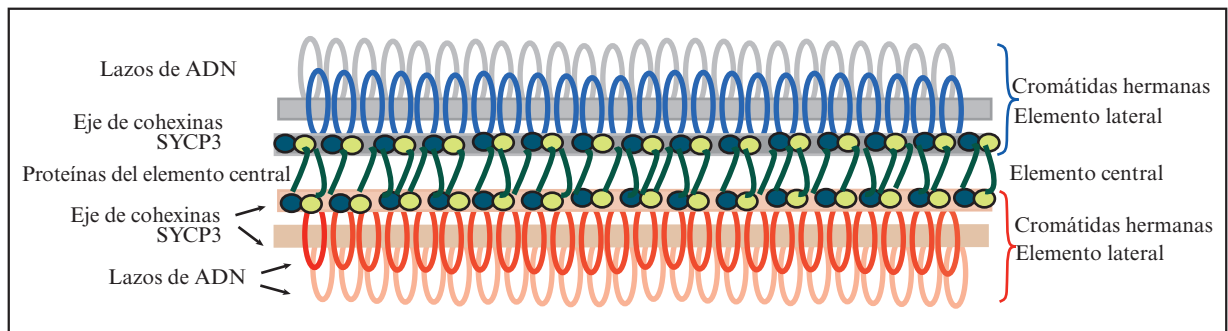


Figura 2.9. Estructura esquemática del complejo sinaptonémico que recuerda una cremallera. Se muestran los dos elementos laterales donde se encuentran proteínas y el ADN que va a recombinar. El elemento central está formado por numerosas proteínas que se supone facilitan el intercambio del ADN implicado, sin embargo, es difícil detectar ADN como parte del elemento central del complejo sinaptonémico.

binación con las rupturas del ADN en las dos hebras de una misma cromátida precede normalmente a la formación del complejo sinaptonémico. Sin embargo, en algunos mutantes de la enzima Spo11 de organismos como *Drosophila melanogaster* o *Caenorhabditis elegans* no hay rupturas iniciales ni recombinación y, sin embargo, se forma el complejo sinaptonémico. Por el contrario, puede haber recombinación sin formación del complejo como lo demuestran mutantes en las proteínas que forman dicho complejo sinaptonémico.

Aunque parece obvio que el complejo sinaptonémico debe tener un papel importante en la recombinación homóloga, su función concreta aún no está totalmente esclarecida; el punto más problemático es la ausencia de ADN asociado al elemento central.

2.2. RECOMBINACIÓN ENTRE SITIOS CONCRETOS

Hay un tipo de intercambio de secuencias que siempre ocurre entre dos lugares concretos. Normalmente estos sitios específicos pueden tener entre 14 y 50 pb de longitud y en ocasiones son idénticos, pero la homología, de existir, tiene un papel secundario. Las enzimas que intervienen en el proceso se denominan genéricamente recombinasas y se conocen más de 100 distintas. Los modelos clásicos de recombinación específica de sitio los constituye la integración de un genoma viral en un lugar concreto del ADN hospedador. Los virus que infectan bacterias se denominan bacteriófagos (o, más coloquialmente, fagos) y el bacteriófago mejor estudiado es el fago lambda (λ). Este fago se integra en un lugar concreto de la bacteria por recombinación específica de sitio y también se escinde del cromosoma bacteriano por este mecanismo. El fago P1 se circulariza al llegar a la bacteria mediante recombinación específica de sitio, y los parvovirus se integran en un lugar concreto del cromosoma 19 humano por este tipo de recombinación. También las levaduras tienen un plásmido que cataliza la inversión de parte de sus genes por este procedimiento. Veremos en detalle algunos de estos ejemplos.

2.2.1. La recombinación específica de sitio del fago λ

Los bacteriófagos son parásitos bacterianos que sólo pueden dividirse y proliferar en el interior de la bacteria. Las funciones básicas y de supervivencia de un fago son: la protección de su ácido nucleico del medio, la introducción en el interior de una bacteria, la conversión de la bacteria infectada en un sistema de producción de partículas fágicas y, por último, la liberación al medio de los fagos completos. Esta serie de sucesos constituyen el ciclo lítico, denominándose virulentos a los fagos que sólo son capaces de llevar a cabo este ciclo. Los fagos atemperados, por el contrario, tienen como alternativa la integración en el genoma del huésped y permanecer en lisogenia, es decir, como profago, duplicándose con el cromosoma del huésped y expresando sólo el gen represor, imprescindible para impedir que se inicie el ciclo lítico. El fago atemperado λ , tal como se aísla de la partícula, es una molécula lineal de doble banda de 48,5 kb (kilobases). Se conoce su secuencia totalmente y en los extremos de la cadena posee un corto fragmento de 12 bases. Estos extremos son complementarios y permiten la circularización del profago antes de la integración en el genoma de la bacteria. Estas secuencias protuberantes constituyen los extremos cohesivos «cos».

En este apartado nos interesa el mecanismo por el cual el fago λ se integra en el cromosoma bacteriano y se escinde de él si salta del ciclo lisogénico al lítico. Son reacciones entre dos lugares concretos en las que, más que un intercambio de ADN propiamente dicho, lo que ocurre es una adición o escisión de un ADN foráneo en el cromosoma de la bacteria. Pero antes de entrar en los detalles del proceso vamos a analizar algunos aspectos de la vida del bacteriófago. Cuando un fago λ infecta a una bacteria *E. coli*, lo primero que hace es circularizarse mediante los extremos cohesivos *cos* y así evitar el ataque de exonucleasas que pudieran dañar la integridad de su genoma. Seguidamente hay una decisión que tomar: ciclo lítico con la producción masiva de partículas fágicas, muerte de la bacteria y liberación de la progenie *versus* integración en un lugar concreto

del cromosoma bacteriano y funcionar como una callada parte del genoma parasitado (profago). La decisión de seguir el ciclo lisogénico o el lítico depende de un complejo equilibrio entre una serie de productos génicos del fago y del hospedador. Uno de los puntos clave parece ser la competición entre un factor represor del fago denominado CI y un anti-represor denominado Cro, por controlar los dos operadores que flanquean el represor (figura 2.10). Un operador en procariotas es una región reguladora de la expresión génica donde hay un sitio de unión para la ARNpol y otro para la unión del represor. En el caso del fago λ estas dos regiones se solapan, con lo que la entrada de la ARNpol al operador asegura una transcripción y traducción de los genes contiguos y un impedimento estérico para el acceso del represor al operador; por el contrario, si el represor es el primero en ganar acceso al operador, la ARNpol no podrá entrar y la represión de los genes contiguos estará

asegurada. En λ hay dos operadores/promotores a la derecha e izquierda del represor: P_D/O_D y P_I/O_I . Hay también un promotor P_M que controla la expresión del represor y que es estimulado por el propio represor. Así durante la lisogenia sólo se expresa el represor. Pero, volviendo a la elección de la vía lítica o la lisogénica, los niveles del producto CII tienen una influencia crítica en la elección. CII es un factor de transcripción que actúa sobre el promotor temprano P_E (del inglés: *early*) y facilita la transcripción de un ARN de interferencia anti-*cro* (gen que codifica por Cro) y que sin embargo tiene la misma orientación y transcribe el represor. Así pues, si los niveles de CII son altos, el represor gana el acceso a los operadores y establece la lisogenia. Si, por el contrario, los niveles de CII son bajos, Cro se sitúa en los operadores y se inicia la cascada de eventos que culminan en el ciclo lítico. El efecto de CIII es secundario, ya que ayuda a proteger a CII impi-

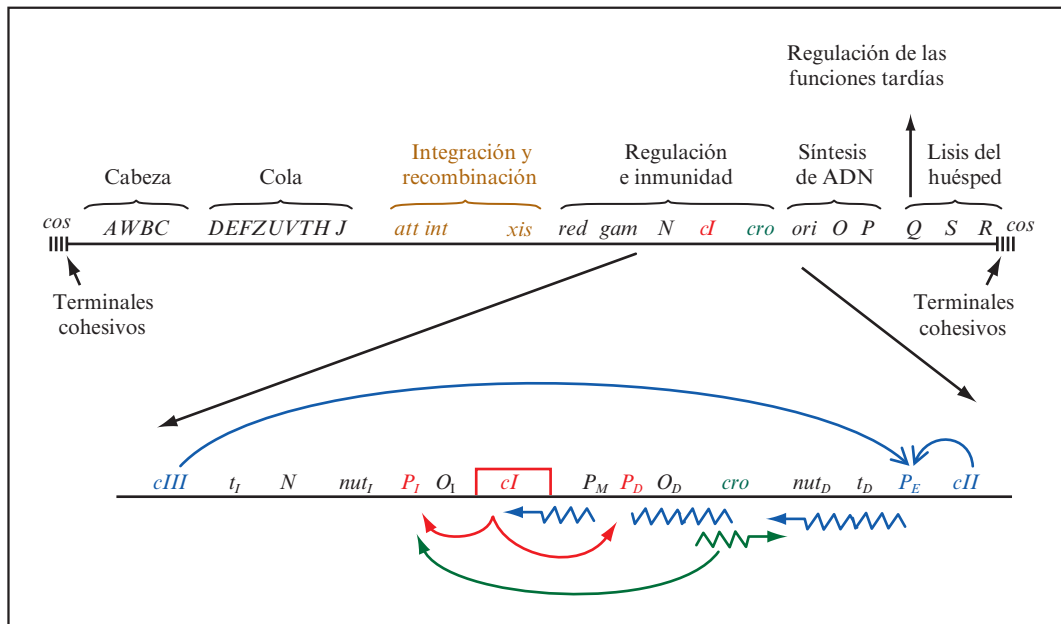


Figura 2.10. Control del establecimiento de la lisis o lisogenia en el fago λ . Visión esquemática del genoma del fago con los genes relativos a las distintas funciones. Ampliada la parte relativa a la lucha entre el represor *cI* y el anti-represor *cro* por el control de los promotores P_I y P_D . También CII y CIII sobre P_E para decidir el establecimiento de la lisogenia. La ausencia de CII provoca la inducción del ciclo lítico. Los genes, con excepciones, se escriben en minúscula y cursiva; las proteínas correspondientes en mayúscula.

diendo su degradación. La influencia de la célula hospedadora en los niveles de CII es también crítica, ya que *E. coli* posee unas proteasas que degradan CII y, por tanto, en su presencia se induce el ciclo lítico. El problema es que estas proteasas se activan por crecimiento en medios de cultivo ricos. Así pues, *lambda* lisará células que crecen muy bien y lisogenizará cultivos pobres.

Si la decisión es a favor de la lisogenia, la recombinasa que facilitará la integración es la integrasa Int (gen correspondiente: *int*), que es una nucleasa y topoisomerasa. La integración es una recombinación entre el sitio en el fago *att P* (del inglés: *attachment phage*) y el sitio en la bacteria *att B*. La eliminación de *att B* del genoma de la bacteria provoca una disminución en la eficiencia de integración de unas 1.000 veces y en este caso el lugar de integración pasa a ser al azar. Tanto *att P* como *att B* tienen 3 componentes denominados respectivamente POP' y BOB', donde la secuencia O es común para ambos sitios y es donde ocurre físicamente la recombinación (figura 2.11). El tamaño de *att P* es de 240 pb mientras que *att B* sólo ocupa 23 pb, y de éstas la región central O representa 15 pb. Además de Int, la integración requiere la presencia de una proteína de la bacteria denominada: factor de integración del hospedador (FIH). Después de la integración se generan dos nuevos sitios denominados BOP' (*att* izquierda) y POB' (*att* derecha) en los extremos del profago que serán los sustratos de la recombinasa Xis encargada de la escisión del profago y salto al ciclo lítico. Aunque el proceso de la recombinación podría ser reversible, la presencia o ausencia de Xis determina la direccionalidad del proceso. Xis es un producto del gen *xis* del fago, que no estará presente a la llegada del fago a la bacteria.

La ruptura de los ADN de la bacteria y del fago ocurre de modo secuencial semejante al descrito en el apartado 2.1.1 para la recombinación procariótica. Primero se producen mellas en una de las dos hebras de cada recombinante y a una distancia de 7 pb se produce el segundo corte en las otras dos hebras que no habían intervenido en el primer corte (figura 2.11). La integrasa es una proteína monomérica con capacidad de corte y

unión (endonucleasa y ligasa). La primera reacción de corte genera dos extremos 3'-fosfato y otros dos 5'-OH; los extremos 3'P se unen covalentemente cada uno a una tirosina de un monómero enzimático, dejando libres los extremos 5'OH que atacan los 3' P-tirosina del otro lado produciéndose el primer quiasma que migra 7 pb y se produce el segundo corte en las hebras que no habían participado en el primero (figura 2.11). Las integrasas están relacionadas con las topoisomerasas, pero la diferencia fundamental entre ambas es que las primeras cortan y unen moléculas de ADN de distinta procedencia, mientras que las topoisomerasas rompen y unen las mismas moléculas de ADN.

La proteína FIH tiene 20 kD (kilodaltons), dos subunidades distintas que proceden de los genes *himA* y *himD* y la capacidad de enrollar ADN en la superficie del intasoma, complejo multiproteico donde tiene lugar la recombinación descrita. El intasoma se asocia primero con *attP*, Int y FIH; así configurado atrapa *attB* y las reacciones comienzan.

La escisión del fago λ es un proceso de recombinación muy similar al descrito para la integración aunque con algunas diferencias, como que además de Int se necesita Xis y se prescinde de FIH. En este caso, Int sólo reconocerá *att D* y *att I* si Xis está presente.

2.2.2. Otras recombinaciones específicas de sitio

1. El sistema Cre-lox del fago P1

Un sistema más sencillo que el descrito para el fago λ es el del fago P1. La recombinasa Cre codificada por el fago cataliza la recombinación entre dos secuencias idénticas que limitan los extremos del bacteriófago. La longitud de estos sitios específicos es de 34 pb y se denominan regiones *loxP*. Cada *loxP* consta de dos regiones de 13 pb en orientación invertida y una región espaciadora de 8 pb asimétrica. Con esta secuencia, se pueden formar troncos monocatenarios apareados en paralelo, que terminarán en los lazos de 8 pb (figura 2.12). Estructura esta última que reconocerá

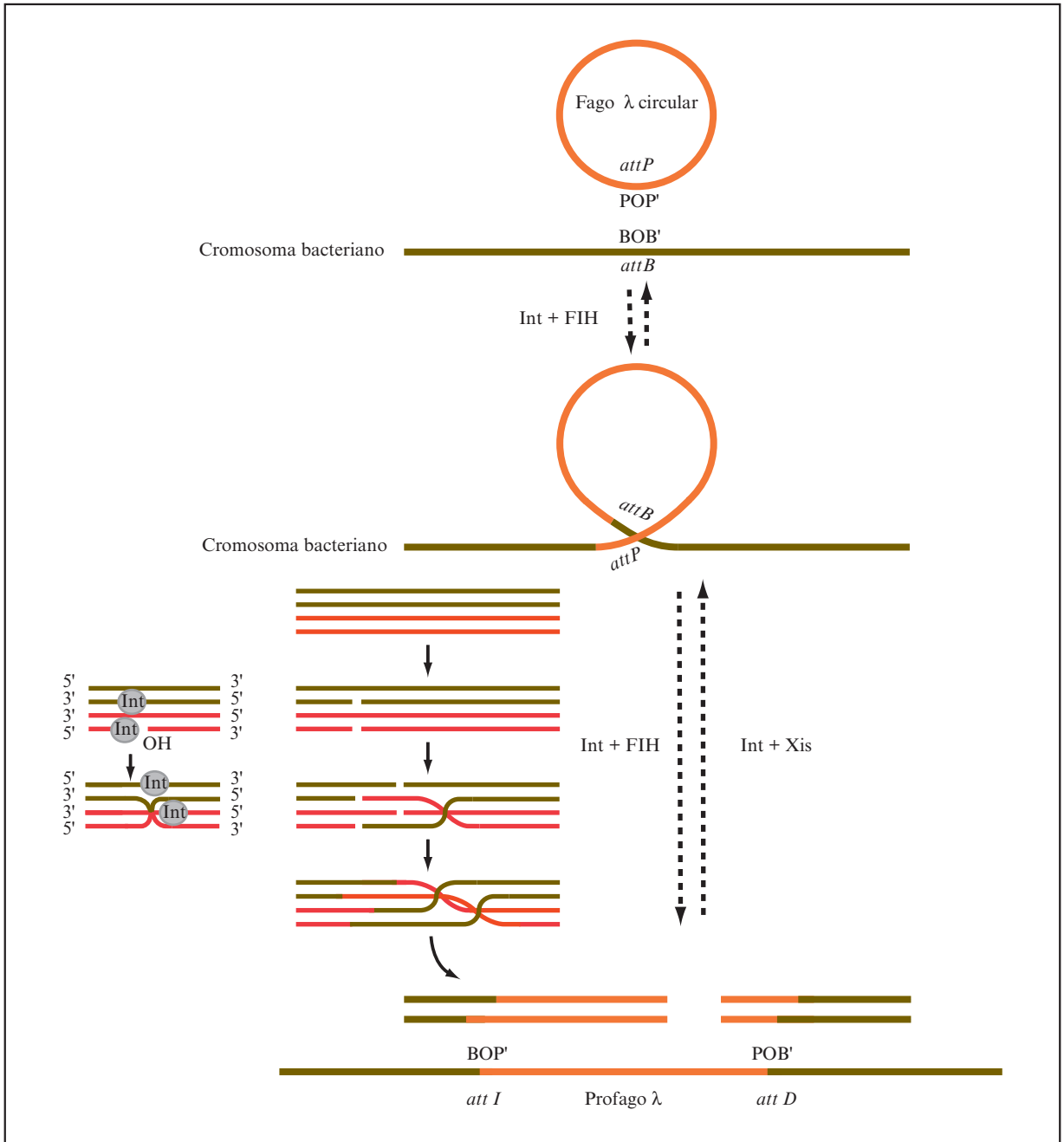


Figura 2.11. Recombinación específica de sitio del fago λ. Visión esquemática de la integración y escisión del genoma fágico en el cromosoma bacteriano a través de *att P* (POP') y *att B* (BOB'). La escisión ocurre entre *att I* (BOP') y *att D* (POB') y requiere la presencia de Xis además de Int y FIH. En la parte izquierda de la figura se representa el mecanismo molecular de esta recombinación con el detalle de las primeras rupturas mediadas por la integrasa (Int).

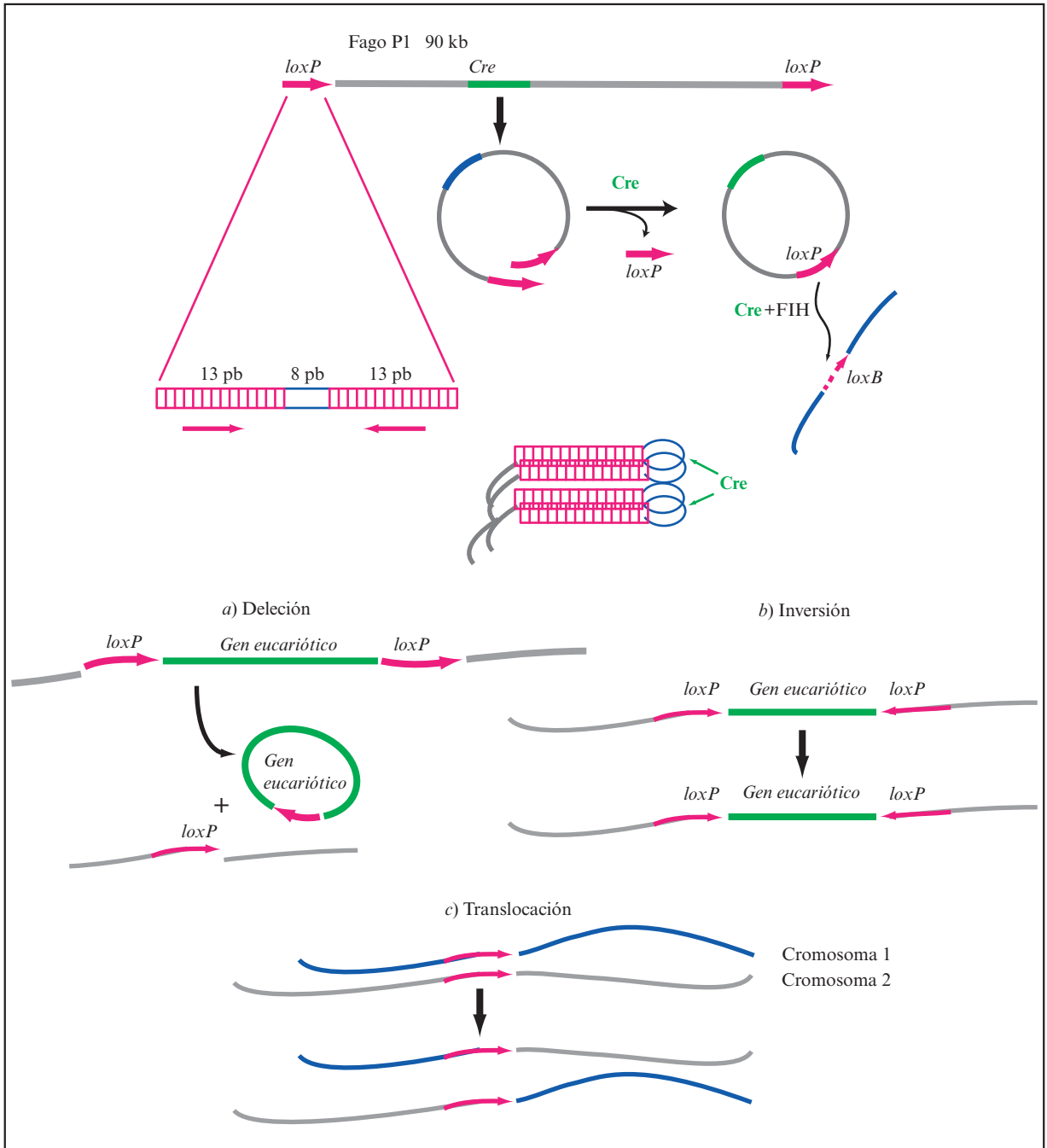


Figura 2.12. Recombinación específica de sitio del fago P1. Adquisición de la forma circular del fago por recombinación específica de sitio mediada por Cre. También integración en el cromosoma bacteriano mediada por Cre y FIH. Adaptación a genomas eucarióticos para provocar: deleciones, inversiones o translocaciones.

Cre para cortar y unir las moléculas recombinantes. Cre es suficiente y necesaria para llevar a cabo la recombinación que en el fago tiene como función convertir la forma lineal del bacteriófago en una forma circular que sea resistente a las exonucleasas bacterianas. La recombinasa posee 343 aminoácidos y dos regiones que se definen por su proximidad a los extremos amino o carboxilo. Sólo la región próxima al terminal carboxilo tiene actividad de corte y unión. La región asimétrica de 8 pb determina la orientación del *loxP* y el tipo de corte: escisión si los dos *loxP* están en la misma orientación e inversión si están en orientaciones opuestas.

El fago P1, una vez circularizado, puede integrarse en el sitio *loxB* de la bacteria, proceso que también se lleva a cabo por una recombinación específica entre los sitios *loxP* y *loxB*. Esta recombinación no es muy eficiente en el caso de este fago y para que se lleve a cabo es necesaria la presencia del factor de integración del huésped (FIH), además de Cre.

Debido a su simplicidad, el sistema Cre-*lox* se ha adaptado para su uso en multitud de células eucarióticas, entre ellas células de muchos mamíferos incluido el hombre, para llevar a cabo recombinaciones específicas de sitio. Se pueden inducir deleciones, inversiones o translocaciones según orientemos los *loxP*. Tendremos una deleción con los dos *loxP* en la misma orientación; una inversión con los dos *loxP* en orientaciones opuestas, y una translocación con los *loxP* en distintos cromosomas (figura 2.12).

2. El sistema ϕ C31 de *Streptomyces*

El sistema Cre-*lox* tiene el inconveniente de que es potencialmente reversible, aunque sea termodinámicamente improbable que ocurra. El sistema del fago atemperado ϕ C31 que infecta *Streptomyces* no es reversible, por lo que cada vez está siendo más utilizado en organismos superiores (*Drosophila*, el pez cebra, células de ratón y humanas). La integrasa de ϕ C31 facilita la recombinación entre el *attP* del fago (39 pb) y el de la bacteria *attB* (34 pb) sin necesidad de cofactores de la célula hospedadora. Además, esta recombinación

es estable y no revierte, ya que no se ha encontrado hasta el momento un cofactor que junto a la recombinasa reconozca las secuencias híbridas *attD* (lado derecho) y *attI* (lado izquierdo) creadas con la integración. En sus aplicaciones eucarióticas, la recombinación suele ser más eficiente si se introduce el *attP* en el genoma eucariota y se facilita *attB* en un plásmido portador de la secuencia a integrar. También se han obtenido recombinaciones sin añadir al genoma eucariótico un sitio *attP*, debido a que por azar hay lugares pseudo-*attP* en la mayoría de los genomas grandes (se han detectado 27.924 en el genoma humano). Si lo que buscamos es la escisión de un gen concreto, pondremos *attP* y *attB* en los extremos del gen a eliminar (figura 2.13). Esta metodología también nos permite acercar un gen a un promotor eliminando las secuencias intermedias.

3. El sistema FLP-FRT de *Saccharomyces cerevisiae*

Algunas cepas de *S. cerevisiae* poseen un plásmido endógeno de 2 μ m de longitud que no confiere a la célula ninguna característica especial. Se le denomina simplemente círculo de 2 μ m. Es estable y se comporta como un minicromosoma circular. Posee la estructura típica de la cromatina, es nuclear y su replicación parece estar bajo el mismo control genético que el resto de los cromosomas. Su presencia no es esencial para la levadura; las cepas que lo poseen se denominan *cir*⁺, y las que carecen de él *cir*^o. El círculo está dividido en dos partes que se conectan a través de dos repeticiones invertidas de 599 pb que constituyen el sustrato de la recombinasa FLP. El sistema FLP-FRT (del inglés: *flippase-flippase recognition target*) es responsable de las inversiones que ocurren en este plásmido cuya función no está clara (figura 2.14). La secuencia FRT consta de dos regiones invertidas alrededor de una secuencia asimétrica central al estilo de *loxP*. La recombinación entre dos sitios FRT causa el intercambio de las secuencias colindantes, pero preserva el sitio FRT, como ocurría en el caso de *loxP*, por lo que también aquí teóricamente el proceso podría ser reversible. Sus aplicaciones en eucariotas son prácticamente

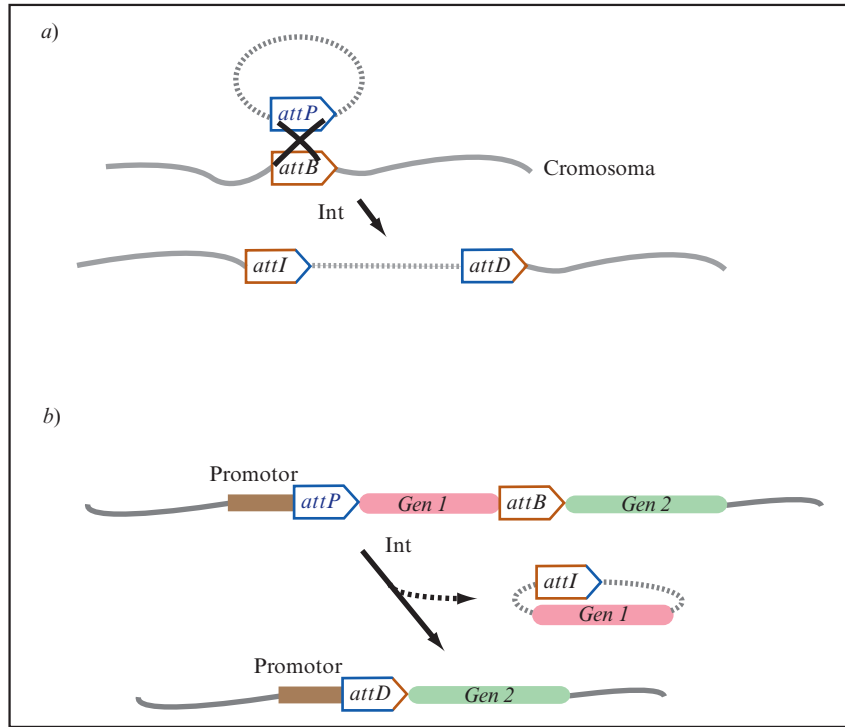


Figura 2.13. Recombinación específica de sitio del fago ϕ C31 de *Streptomyces*. Estrategias generales de inserción (a) y escisión (b) de un gen o secuencia.

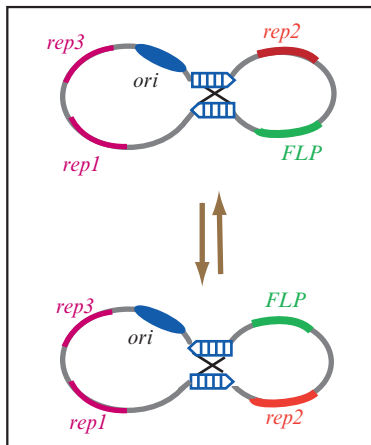


Figura 2.14. Sistema de recombinação específica de sitio del círculo de 2 μ m de *Saccharomyces*. La recombinação provoca la inversión de los genes *rep2* y *FLP*. Este último corresponde a la propia recombinasa FLP.

idénticas a las descritas para *loxP*, la recombinação entre dos secuencias *FRT* flanqueando un gen concreto en presencia de Flp provoca la escisión del gen en forma circular como se vio en la figura 2.12.

4. Los virus adenoasociados y su integración en el cromosoma 19 humano

Los virus adenoasociados (AAV) constituyen una familia de virus defectivos dentro de los parvovirus humanos (dependovirus). Sus distintos miembros (serotipos diferentes) poseen un pequeño genoma de 4,7 kb de ADN monocatenario y no parecen producir ninguna enfermedad en el hombre. Son virus atemperados que pueden llevar a cabo el ciclo lítico o el lisogénico. Cuando el virus infecta una célula humana, adquiere la forma replicativa que es una versión bicatenaria y la

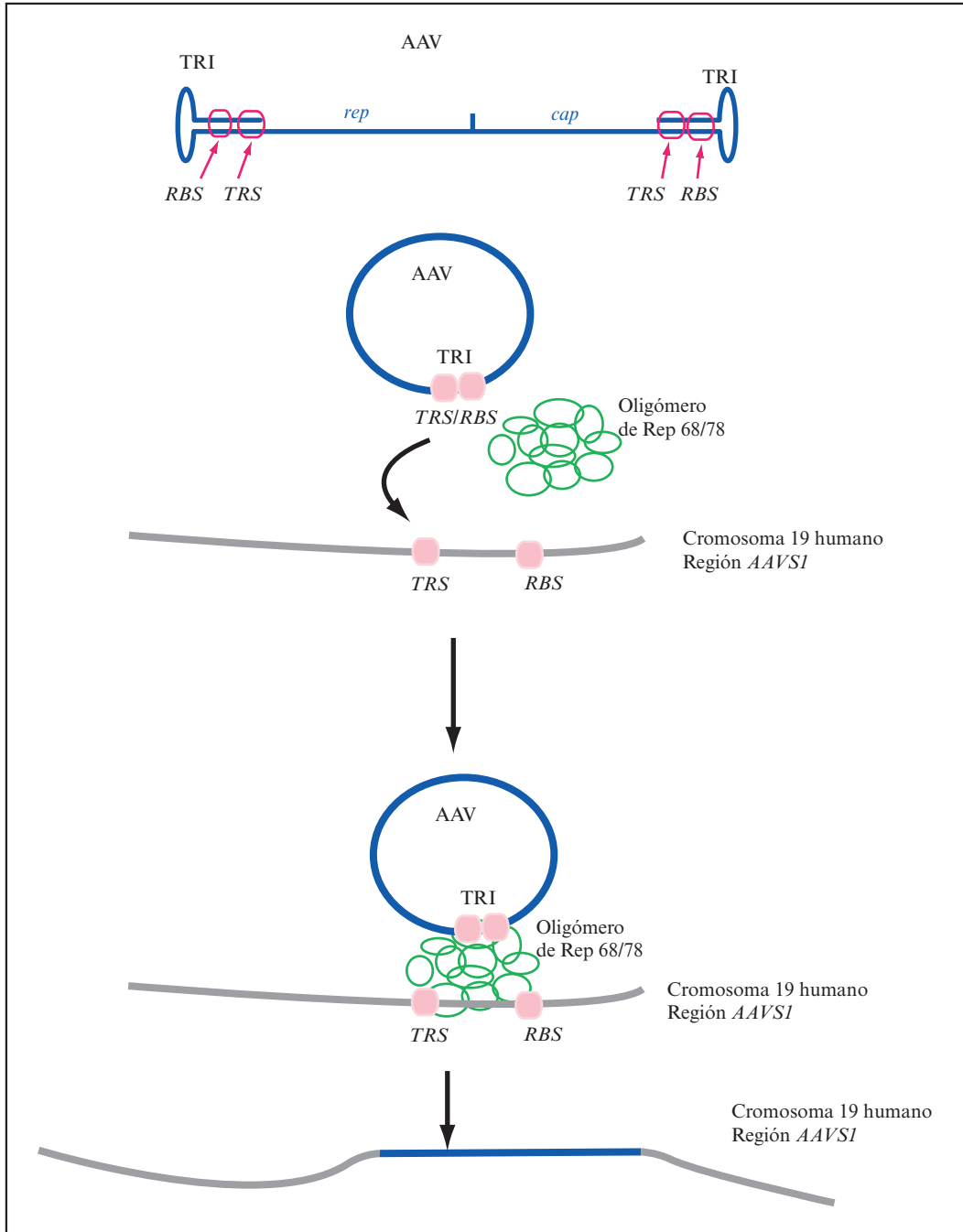


Figura 2.15. Sistema de recombinación específica entre los terminales repetidos invertidos (TRI) del virus adenoasociado (AAV) con la región AAVS1 del cromosoma 19 humano. Participación imprescindible de la recombinasa Rep 68/78. *RBS* es el lugar de unión a la proteína Rep y *TRS* es el lugar implicado en la resolución terminal.

que se integrará en el genoma de la célula. Estos virus necesitan ser auxiliados por un adenovirus o un herpes para llevar a cabo una infección propiamente dicha matando a la célula hospedadora y liberando muchas partículas víricas (ciclo lítico). En ausencia de virus auxiliar, se integran por recombinación específica de sitio en un lugar concreto del cromosoma 19 (región AAVS1, entre q13.3 y qter) quedándose en forma latente hasta que ocurra una infección por un adenovirus o herpes (ciclo lisogénico/infección latente).

El genoma viral contiene dos genes: el gen *rep*, que codifica por cuatro proteínas necesarias para la integración y replicación viral, y el gen *cap*, que lleva la información para sintetizar las dos proteínas de la cápside. Uno de los productos de *rep* es una replicasa/integrasa denominada Rep 68/78 que alinea las secuencias específicas del virus y de la célula que van a recombinar y produce los cortes en las moléculas recombinantes. Los extremos del genoma viral están definidos por dos terminales repetidos e invertidos (TRI) de 145 nucleótidos que son los únicos elementos requeridos en *cis* (en la misma unidad y proximidad física) para la replicación (sirven de origen de replicación), la integración y el empaquetamiento. El proceso de integración está orquestado por la célula hospedadora que contribuye con algunos factores de integración, pero los verdaderos protagonistas son los dos terminales repetidos invertidos (TRI), Rep y la región AAVS1 del cromosoma 19 humano (figura 2.15). El mecanismo en detalle es complejo y está ligado a la replicación de unos pocos genomas del virus y a reparación. Participan los lugares RBS (del inglés: *Rep binding site*) y TRS (del inglés: *terminal resolution site*) que están presentes tanto en el virus como en la región AAVS1 del cromosoma 19.

La propiedad de integrarse en un lugar concreto del cromosoma 19 humano ha permitido a estos virus ser considerados como posibles vectores de genes foráneos, o copias sanas de genes defectuosos, al genoma humano (terapia génica). Desgraciadamente, al convertir el AAV en vector portador, hay que eliminar todo el genoma del AAV para dejar espacio al gen terapéutico y este proceso suele conllevar la pérdida de la capacidad

de integrarse en el cromosoma 19 humano y ya sólo se integrará de forma muy ineficiente y esporádica (véase capítulo 7, apartado 7.3.5).

2.3. TRANSPOSICIÓN

Los transposones son elementos móviles que se desplazan de un sitio a otro del genoma. La entrada y salida del elemento en los cromosomas se realiza por recombinación, pero ni hay homología de secuencias ni sitios concretos; es una recombinación al azar. La transposición constituye una de las mayores causas de mutaciones en los genomas. Los transposones están emparentados con los virus y no está claro si son virus que han perdido la capacidad de salir y entrar en las células o son los predecesores naturales de los virus, antes de que éstos ganaran las competencias de salir y entrar en las células. Estos elementos móviles se clasifican en dos grandes grupos: 1) transposones que se desplazan vía ADN, y 2) los que lo hacen vía ARN, también llamados retrotransposones y muy afines a los retrovirus. Tanto unos como otros pueden ser replicativos o no replicativos dependiendo de que dejen, o no, una copia en el lugar que ocupan antes de desplazarse a un nuevo sitio. Cuando son replicativos, utilizan la maquinaria de síntesis de ADN de la célula, pero además de la transposasa suelen codificar por una resolvasa. Cuando no son replicativos, al salir de un lugar concreto producen una ruptura en el ADN de difícil reparación, que puede acarrear graves consecuencias para la célula hospedadora.

2.3.1. Transposición vía ADN

Los transposones que se movilizan vía ADN se encuentran en procariotas y eucariotas. Los transposones bacterianos transportan el gen o los genes necesarios para la transposición que no encuentren en la célula hospedadora. En eucariotas también encontramos una gran variedad de transposones aunque la gran mayoría de ellos son defectuosos y han perdido la capacidad de moverse de un lugar a otro en el genoma. Sin embargo,

estos transposones no funcionales pueden en ocasiones ser reconocidos por enzimas procedentes de un transposón funcional y ser sustratos de transposición. El genoma de *Drosophila melanogaster*, por ejemplo, posee más de 50 tipos de transposones distintos y un total de varios cientos de elementos móviles individuales. Los primeros transposones se identificaron en bacterias como inserciones espontáneas en regiones reguladoras (operones) impidiendo o modificando la transcripción de los genes asociados. Son componentes habituales de los cromosomas y plásmidos de bacterias. A día de hoy no está clara la posible ventaja evolutiva de estos elementos móviles, ya que, asumiendo que uno o unos pocos de ellos hubiesen llegado a una cepa de bacterias carente de ellos, el número de mutaciones beneficiosas que pudieran derivarse de su presencia es bajo en comparación con las deletéreas y, por tanto, insuficiente para favorecer una selección positiva.

Los transposones más sencillos son las **secuencias de inserción IS** (del inglés: *insertion sequences*). Una cepa de *E. coli* suele tener unas 10 copias de los elementos IS más comunes. La mayoría de estas secuencias de inserción se agrupan en familias relacionadas entre sí, que a su vez poseen subgrupos, y éstos, miembros afines pero no idénticos. Cada secuencia de inserción es una unidad autónoma constituida por dos elementos: 1) dos terminales repetidos en orientación invertida que delimitan el transposón, y 2) el gen de la transposasa, enzima imprescindible en la movilidad del transposón (figura 2.16). Los terminales repetidos no suelen ser idénticos, sino muy parecidos; tienen una longitud variable, entre 10 y 40 pb, pero específica para cada transposón (tabla 2.1), y suelen incluir un promotor débil integrado en el terminal repetido invertido izquierdo; también pueden utilizar promotores vecinos del hospedador para expresar la transposasa. Sin embargo, no podrán expresarla si los terminales repetidos están metilados (silenciados). Como consecuencia de la transposición se produce una duplicación de entre 5 a 9 pares de bases, según el tipo de IS, en el lugar de aterrizaje del transposón. Estas repeticiones tienen la misma orientación denominándose repeticiones directas. En la secuencia de aterrizaje

la transposasa provoca un corte quebrado (dos subunidades, cada subunidad una mella) cuya amplitud de quiebro determinará el tamaño de la secuencia duplicada. El transposón se une a los extremos protuberantes de hebra sencilla resultantes de separar el ADN diana cortado por la transposasa. Los dos huecos resultantes se reparan y sellan. Las transposasas pueden actuar como monómeros, dímeros o multímeros. La transposasa del transposón Tn5 actúa como dímero y cada monómero tiene un centro activo donde se produce la ruptura de cada uno de los extremos del elemento, manteniéndolo protegido (figura 2.16).

Aunque teóricamente las IS se insertan al azar, en la práctica muchas tienen secuencias preferentes que constituyen puntos calientes o lugares donde la llegada de un transposón es más frecuente que la media. En *E. coli* la frecuencia de transposición, en general, es de 10^{-3} a 10^{-4} por elemento por generación, mientras que la frecuencia de que un gen o secuencia reguladora reciba un transposón es de 10^{-5} a 10^{-7} ; por otro lado la frecuencia de reversión se ha estimado entre 10^{-6} y 10^{-10} . Todas las IS, excepto la IS1, poseen una única unidad transcripcional que comienza en un promotor débil parcialmente localizado en el terminal repetido izquierdo y termina en el derecho; el producto es la transposasa. IS1 tiene una organización más compleja con dos pautas de lectura y la necesidad de un salto de una pauta a la otra durante la traducción de la enzima. La transposasa es responsable del reconocimiento de los extremos del transposón, la ruptura de las hebras, la elección de la nueva diana y la inserción en el nuevo sitio. La transposasa provoca tanto las mellas en el transposón como en la secuencia diana. Las mellas en los extremos del transposón pueden: 1) acabar en rupturas totales de las dos hebras, o 2) en círculos de ADN de banda sencilla que se convierten en ADN de hebra doble y se integran (figura 2.17). En el primer caso, cada subunidad de transposasa origina un corte en una de las hebras del ADN en cada extremo del transposón originando extremos 3'OH por un ataque nucleofílico mediado por una molécula de H_2O ; este grupo hidroxilo a su vez ataca a un grupo 5'-P de la hebra intacta complementaria con la que acaba uniéndose formando

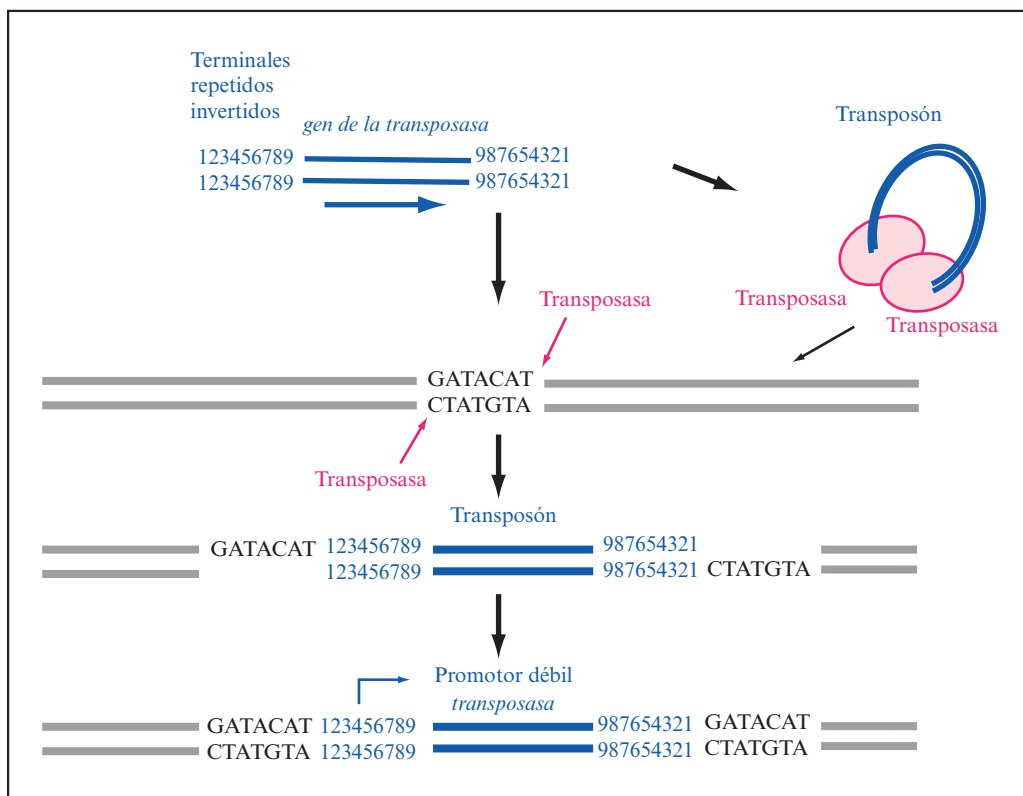


Figura 2.16. Esquema de un transposón. Los terminales repetidos invertidos se representan como números por simplicidad, el de la izquierda suele incluir un promotor débil. La integración provoca una duplicación directa de la secuencia de reconocimiento en el lugar de aterrizaje. Arriba a la derecha se muestran dos monómeros de la transposasa en su interacción con los extremos de un transposón y su protección hasta tomar contacto con la nueva localización.

TABLA 2.1

Distintas familias de transposones sencillos y algunas de sus características

Familia de transposones	Repetición invertida imperfecta	Tamaño del transposón	Repetición directa en la diana
IS1	23 pb	768 pb	9 pb (8-11)*
IS2	41 pb	1.327 pb	5 pb
IS3	20-40 pb*	1.200-1.550 pb*	3-4 pb*
IS4	18 pb	1.300-1.950* pb	9-12 pb
IS5	16 pb	850-1.643* pb	4 pb
IS6	15-20 pb*	789-880* pb	8 pb
IS30	20-30 pb*	1.027-1.221 pb	2-3 pb*
IS903	18 pb	1.057 pb	9 pb

* Rango de tamaños.

una horquilla. Entonces una molécula de agua activada ataca a la horquilla y genera terminales libres en las dos hebras del transposón. La reacción no requiere energía externa. En el segundo caso, el extremo 3'OH generado en el primer corte ataca al extremo opuesto de la misma hebra del transposón,

originándose un círculo de transposición que se duplica (transposición replicativa, véase más adelante) y deja una copia del transposón en el lugar original mientras que la otra copia aterrizará en un lugar nuevo. Esta última se ha observado en algunos miembros de la familia IS3.

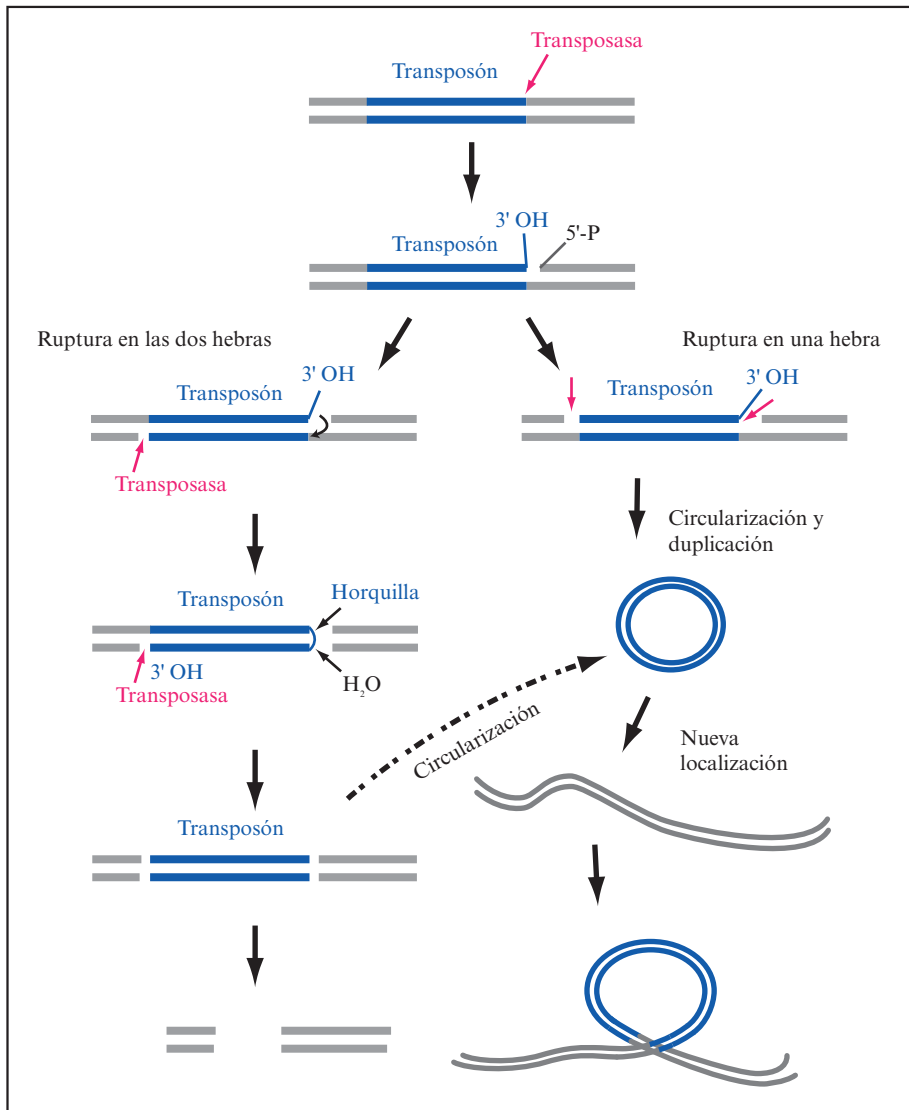


Figura 2.17. Tipos de rupturas de los terminales del transposón mediadas por la transposasa. Por el lado izquierdo, ejemplo de una transposición por ruptura de las dos hebras y circularización. Por el lado derecho, ruptura en una sola hebra, circularización y replicación.

Dos IS cercanos pueden formar un **transposón compuesto** (Tn) y moverse individualmente o conjuntamente (figura 2.18). Cuando lo hacen conjuntamente, la transposasa reconoce el extremo izquierdo del primer transposón y el derecho del segundo, movilizandolo el gen o genes que se encuentren en medio de ambos. Éste parece ser el origen de los transposones compuestos que movilizan genes de resistencia a antibióticos, como

el Tn 903 que, entre dos IS903, lleva el gen de resistencia al antibiótico kanamicina (kan^R), lo que confiere a las células que lo poseen una ventaja selectiva en medios con el antibiótico. Los dos IS que forman un Tn pueden ser iguales o distintos, funcionales o no (tabla 2.2), estar en la misma orientación (Tn 9) o en orientaciones distintas respecto a la región central (Tn 903) (figura 2.18).

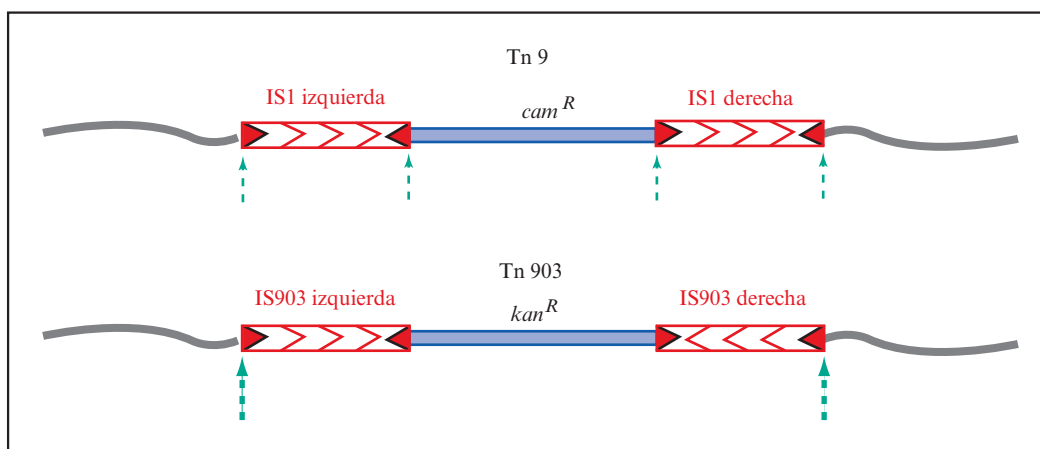


Figura 2.18. Transposones compuestos (Tn 9 y Tn 903) de dos elementos IS y una región central que lleva un gen de resistencia a un antibiótico. Arriba, los dos transposones iguales y en la misma orientación (IS1) del Tn 9. Abajo, los dos transposones iguales y en orientaciones opuestas, respecto a la región central, de Tn 903.

TABLA 2.2

Algunos transposones compuestos y gen de resistencia en el centro de ambos

Transposón	IS izquierda	Gen de resistencia	IS derecha
Tn9	IS1 funcional	<i>cloranfenicol^R</i>	IS1 funcional
Tn903	IS903 funcional	<i>kanamicina^R</i>	IS903 funcional
Tn5	IS50L no funcional	<i>tetraciclina^R</i>	IS50R funcional
Tn10	IS10L no funcional	<i>kanamicina^R</i>	IS10R funcional

Cuando un transposón, simple o compuesto, se mueve de un sitio a otro en el cromosoma, provoca una ruptura en el lugar donde se encontraba de muy difícil reparación y una mutación insercio-

nal en el lugar donde aterrice, por lo que el proceso puede ser letal para la célula (figura 2.17).

Aunque nos hemos centrado en los transposones no replicativos de bacterias, estos elementos

están también muy presentes y activos en plantas y animales de orden inferior; en mamíferos, en cambio, no parecen tener mucha actividad.

Los elementos P de *Drosophila melanogaster*

Ciertas cepas de *Drosophila* tienen dificultades para reproducirse entre ellas y la causa son unos transposones que sólo se encuentran en las cepas denominadas P. Hay de 30 a 50 copias por genoma. En sus extremos, los elementos P tienen unos terminales invertidos repetidos de 31 pares de bases y generan repeticiones directas de 8 pb. El transposón tiene 2,9 kb y 4 pautas abiertas de lectura (ORF, del inglés: *Open Reading Frame*). La fusión de los 4 ORF da como producto la transposasa. Entre el ORF0 y el ORF1 hay un intrón, y entre el 2 y el 3 otro (figura 2.19). El segundo intrón sólo se elimina en células germinales y no en células somáticas. Esto probablemente explica por qué la transposición es tejido-específica y sólo ocurre en células germinales. Las cepas que carecen del transposón se denominan M. Cuando un genoma que posee elementos P (paternos) se fusiona con células germinales M (maternas), se estimula la transposición de estos elementos. Esto sólo ocurre en las células germinales del embrión resultante y origina una serie de efectos genéticos: 1) produce mutaciones, y 2) facilita la recombinación y los reordenamientos cromosómicos. Los embriones se desarrollan a moscas que tienen muchas mutaciones en los gametos y con frecuencia son estériles. El fenómeno se denomina disgenesis híbrida P-M.

y entre el 2 y el 3 otro (figura 2.19). El segundo intrón sólo se elimina en células germinales y no en células somáticas. Esto probablemente explica por qué la transposición es tejido-específica y sólo ocurre en células germinales. Las cepas que carecen del transposón se denominan M. Cuando un genoma que posee elementos P (paternos) se fusiona con células germinales M (maternas), se estimula la transposición de estos elementos. Esto sólo ocurre en las células germinales del embrión resultante y origina una serie de efectos genéticos: 1) produce mutaciones, y 2) facilita la recombinación y los reordenamientos cromosómicos. Los embriones se desarrollan a moscas que tienen muchas mutaciones en los gametos y con frecuencia son estériles. El fenómeno se denomina disgenesis híbrida P-M.

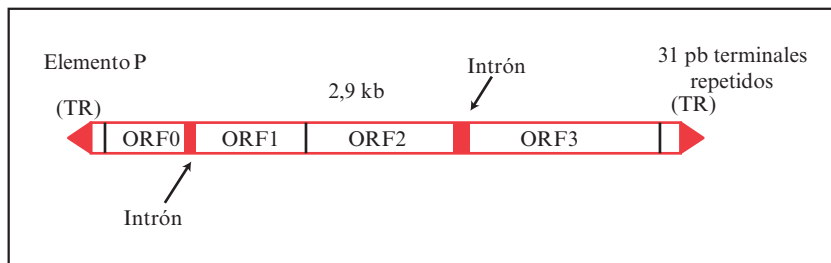


Figura 2.19. Transposones P (paternos) de *Drosophila* con sus regiones con pautas de lectura abiertas (ORF) que dan lugar a la transposasa y sus dos intrones.

La transposición replicativa

La transposición replicativa ocurre cuando un elemento móvil se duplica durante el proceso de transferencia (figura 2.17). Para este tipo de transposición, además de la transposasa que actúa sobre los terminales del transposón y la diana, se necesita una resolvasa que actúa sobre el co-integrado. El daño para la célula es menor, pues se reduce a mutagénesis insercional, pero aumenta el número de transposones en la célula hospedadora y ello puede dar lugar a recombinaciones homólogas que reorganicen drásticamente el genoma. Veremos como ejemplo la transposición del peculiar transposón Mu, un elemento móvil de 38 kb

que también es un bacteriófago. Cuando el bacteriófago Mu infecta una bacteria, se integra por transposición no replicativa, pero durante el ciclo lítico el número de copias se amplifica por transposición replicativa (figura 2.20). Este tipo de transposición se inicia como la no replicativa, la transposasa MuA provoca dos mellas en los dos extremos del transposón y otras dos en la secuencia diana. La proteína MuB (una ATPasa) colabora en la reacción y contribuye en la elección de la secuencia diana que suele ser al azar, pero cuando ocurre en el mismo cromosoma, tiene lugar a una distancia de entre 10 y 15 kb de la inserción original. En este punto ocurre una replicación desde los extremos 3'OH libres que duplican el

transposón con la maquinaria replicativa celular y unen donante y recipiente en un co-integrado que se resuelve con la resolvasa codificada por el

transposón que es una recombinasa. Al resolverse se quedan las repeticiones directas en la localización nueva.

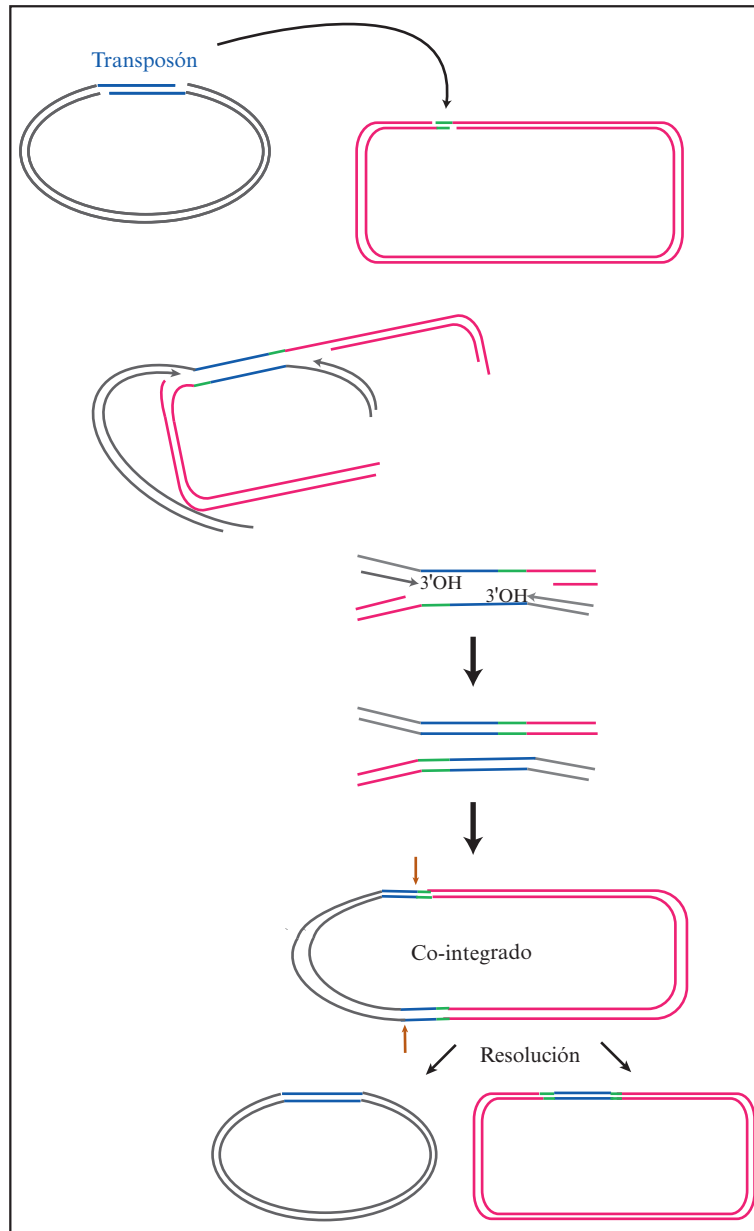


Figura 2.20. Transposición replicativa de Mu entre un transposón de un plásmido a otro con la formación de un co-integrado. En este caso hay duplicación del transposón. El co-integrado posee repeticiones directas del transposón.

La transposición replicativa también es la que utiliza la familia de **transposones TnA**; no son transposones compuestos aunque tienen tres unidades génicas relativamente independientes. Los mejor estudiados de la familia son el Tn3 y Tn1000. Los terminales repetidos tienen alrededor de 38 pb y generan una duplicación de 5 pb allí donde aterrizan. Entre el gen de la transposasa (*tnpA*) y el de la resolvasa (*tnpR*) hay una secuencia específica de 20 pb (*res*) sustrato de la resolvasa y donde tiene lugar la resolución del co-integrado; también es la región donde se inicia la transcripción en los dos sentidos opuestos (figura 2.21). Estos transposones llevan resistencia a antibióticos, generalmente a la ampicilina (*amp^R*). En realidad la resolución viene a ser una recombinación específica de sitio (*res*) al estilo de las ya descritas y sin requerimiento energético. La resolvasa es una especie de topoisomerasa semejante a la integrasa del fago λ ; en la resolución participan 4 unidades de resolvasa, cada unidad responsable de una mella en una de las hebras; seguidamente se produce una reorganización de las unidades entre sí que desplaza físicamente las hebras de ADN entrecruzándolas, las mellas se ligan y la resolvasa se desprende. En ausencia de resolvasa puede tener lugar una recombinación homóloga vía Rec BCD/Rec A, con mucha menor eficiencia.

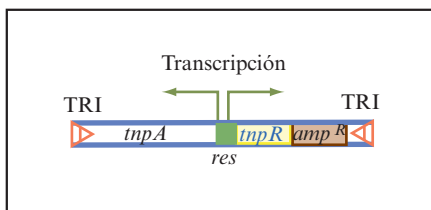


Figura 2.21. Transposón de la familia TnA con tres genes, una región *res* interna y dos terminales repetidos invertidos.

Aunque la mayoría de los transposones se desplazan siempre por un tipo de transposición concreta, bien replicativa (TnA) o no replicativa (Tn5 y Tn10), los hay que pueden utilizar las dos como IS1 o IS903.

Para saber si un transposón utiliza la vía replicativa o la no replicativa se suele introducir una copia del gen *lacZ* de *E. coli* en el transposón. El gen *lacZ* codifica por la β -galactosidasa, enzima que en presencia de un sustrato apropiado (X-gal) produce coloración azul. Sin embargo, la copia del gen *lacZ* que se introduce en el transposón posee en una de sus dos hebras dos mutaciones convenientemente separadas (figura 2.22). Si la transposición es replicativa, se observará color azul en las células en presencia del cromóforo X-gal, al cabo de algunas generaciones, ya que la replicación originará tanto genomas con el gen no mutado como genomas doblemente mutados.

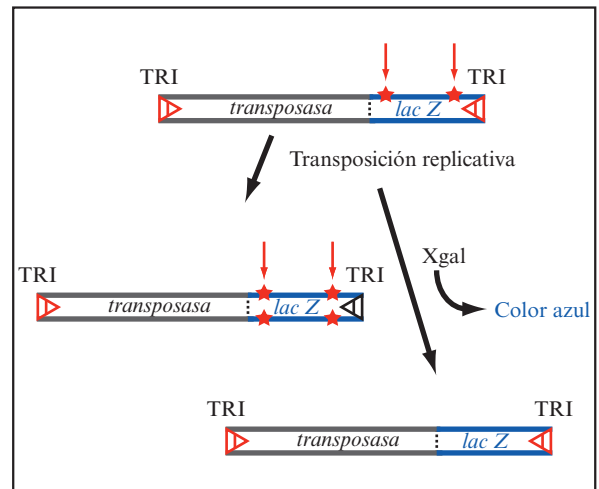


Figura 2.22. Diseño del experimento para saber si una transposición es replicativa o no replicativa mediada por el gen *lacZ* de *E. coli*.

Como se ha mencionado anteriormente, los transposones también se encuentran en plantas y en todos los genomas eucarióticos incluido el humano. Sin embargo, en mamíferos, por ejemplo, la inmensa mayoría de los transposones son defectivos, están silenciados o ambas cosas. Por el contrario, hay veces que elementos no funcionales de una familia de transposones, por carecer o tener alterada la transposasa, pueden ser activados por otros miembros funcionales de la familia, ya que,

aunque las transposasas prefieren actuar en *cis* sobre las secuencias vecinas, pueden actuar en *trans* sobre secuencias más alejadas.

2.3.2. Transposición vía ARN: retrotransposones

Obligatoriamente pasa por un intermediario de ARN y es exclusiva de eucariotas. Cuando se movilizan, generan en el lugar de aterrizaje repeticiones cortas directas. Los retrotransposones están íntimamente relacionados con los retrovirus hasta el punto de que no está claro si los retrovirus actuales son retrotransposones que han ganado la capacidad de salir de la célula para infectar otra, o los retrotransposones son retrovirus que han perdido la capacidad de salir y entrar en las células. Los retrotransposones semejan a retrovirus defectivos que carecen de algunos genes y que sólo se mueven de un sitio a otro del genoma. Unos y otros poseen terminales repetidos y genes que codifican por la enzima transcriptasa inversa que convierte un ARN en ADNc (figura 2.23). Los retrovirus poseen como genoma 2 unidades idénticas de ARN, por lo que necesitan la enzima para convertir este ARN en ADNc una vez que infectan una célula. Los retrotransposones necesi-

tan la enzima para convertir el ARN resultado de la transcripción del transposón en un ADNc que se movilizará a un lugar nuevo (replicativa). Además de la transcriptasa inversa, unos y otros también llevan el gen de la integrasa, que facilita la integración del ADNc en cualquier lugar accesible del genoma. Un retrotransposón también puede salir del lugar que ocupa por recombinación entre los elementos repetidos de los extremos (no replicativa).

Hay una gran variedad de retrotransposones que forman parte de los genomas eucarióticos, desde levaduras al hombre (tabla 2.3). Muchos de los retrotransposones de mamíferos carecen de los terminales repetidos de los extremos, otros carecen de la transcriptasa inversa o de la integrasa, por lo que en realidad lo más abundante en los genomas de mamíferos son los restos de retrotransposones o retrovirus con nula capacidad de moverse. Los *Ty* de levaduras (del inglés: *transposon yeast*) y los *copia* de *Drosophila* poseen terminales repetidos en los extremos y codifican por la transcriptasa inversa y/o la integrasa. Sin embargo, los LINE (del inglés: *long interspersed elements*) carecen de los terminales repetidos aunque la mayoría aún poseen el gen de la transcriptasa inversa. La frecuencia de transposición de los *Ty* es de 10^{-7} - 10^{-8} , algo más baja que la observada en bacterias y hay heterogeneidad manifiesta entre los distintos miembros. El tamaño suele ser de 6,3 kb y los últimos 330 pb de cada extremo son los terminales repetidos, llamados δ que, si bien son idénticos en un elemento dado, pueden variar ligeramente de un elemento a otro. El terminal δ de la izquierda alberga un promotor que inicia la transcripción de dos ARN de 5 y 5,7 kb que incluyen las regiones *TyA* y *TyB*. Esta transcripción es compleja porque poseen distintas pautas de lectura y coinciden en 13 aminoácidos. El resultado son proteínas con actividad transcriptasa inversa e integrasa.

Una de las familias de retrotransposones mejor estudiada de *Drosophila* es *copia*, cuyo nombre hace alusión a la abundancia con que se encuentran los mensajeros de este elemento en la mosca. Suele haber de 20-60 copias del elemento por genoma, pero distintas especies varían considerablemente. Cuando células del insecto se cultivan *in*

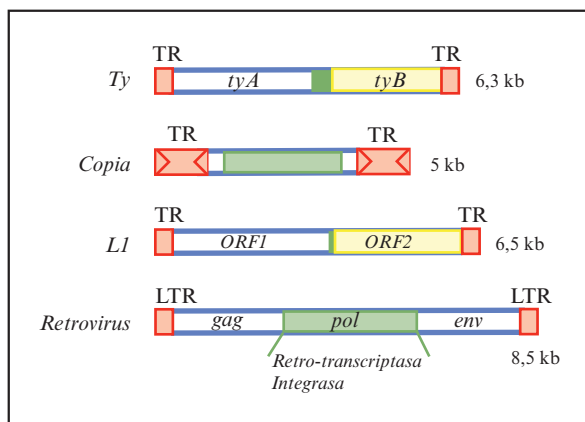


Figura 2.23. Retrotransposones y retrovirus. Organización de sus genomas respectivamente de menor a mayor complejidad.

vitro, el número de elementos *copia* se incrementa 2-3 veces. El tamaño del elemento es de 5 kb con dos terminales repetidos directos en los extremos de 276 pb cada uno. Estos terminales repetidos contienen a su vez dos terminales invertidos no idénticos en sus extremos, algo parecido a las IS de los extremos de los transposones compuestos en bacterias, pero en este caso no se pueden nunca mover independientemente del resto. Al aterrizar en un sitio nuevo, *copia* genera repeticiones directas de 5 pb. El elemento se transcribe a partir del terminal repetido izquierdo a un mensajero de 4.227 pb con terminal de poliA que se traduce en varias proteínas relacionadas con la transcriptasa inversa y la integrasa.

El elemento L1 de humanos tiene 6,5 kb de longitud y dos pautas de lectura abiertas para la transcriptasa inversa y la integrasa cuando está completo, pero en la gran mayoría de los casos son copias truncadas que carecen de actividad. Forman parte del ADN repetido de los genomas y su función es incierta.

Exclusivamente en primates y humanos encontramos la familia de retrotransposones Alu, cuyos miembros no son autónomos y carecen de terminales repetidos largos en los extremos. Tie-

nen un tamaño de 280 pb de longitud y un acusado polimorfismo (capítulo 1, página 27).

Los retrovirus poseen ARN como genoma e infectan muchas especies animales distintas, entre ellas el hombre. Tienen un tamaño aproximado de 8,3 kb, sólo infectan de manera estable células en división y no matan a la célula infectada, sino que se integran en su genoma como provirus, estableciendo una simbiosis indefinida. El virus sintetiza un ADNc que se integra al azar en varios lugares del genoma de la célula infectada (de 1 a 10 copias). La organización del provirus comprende dos regiones con terminales repetidos largos (TRL) de varios cientos de pares de bases. La transcripción se inicia en el terminal 5' a partir de un promotor potente y se termina en la misma región del extremo 3'. En la región interna del provirus hay tres genes: *gag*, que produce un antígeno característico del grupo; *pol*, que codifica por la transcriptasa inversa y la integrasa; finalmente *env*, que codifica por la proteína de la envuelta. Este último gen se traduce a partir de un ayuste que elimina los genes *gag* y *pol*. La integración ocurre en regiones de cromatina en configuración abierta y necesita los extremos TRL y la integrasa asociada al gen *pol*.

TABLA 2.3
Algunos retrotransposones y sus características

Retrotransposón	Terminales repetidos	Repeticiones en la diana	Frecuencia/genoma
<i>Ty (S. cerevisiae)</i>	Sí	4-6 pb	30-40 elementos
<i>Copia (D. melanogaster)</i>	Sí	4-6 pb	20-60
LINE-B1 (<i>M. musculus</i>)	No	7-21 pb	50.000
LINE-L1 (<i>H. sapiens</i>)	No	7-21 pb	850.000

El mecanismo de transposición vía ARN se muestra en la figura 2.24. Tanto los retrovirus como los retrotransposones utilizan la transcriptasa inversa para perpetuarse. Un retrotransposón integrado en el genoma puede transcribirse si no está reprimido y pasar al citoplasma donde podría traducirse y/o pasar a ADNc. El ADNc se

asociaría con integrasas que facilitarían el acceso a una nueva localización en el ADN. Durante la división celular la membrana nuclear desaparece, por lo que es más fácil el acceso al ADN.

Los retrotransposones también son muy abundantes en plantas y aquí el número de elementos es proporcional al tamaño del genoma.

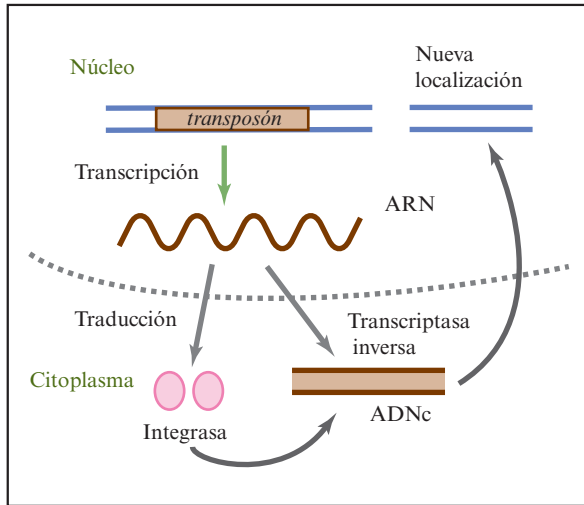


Figura 2.24. Mecanismo de transposición vía ARN. Transcripción en el núcleo del transposón y traslado al citoplasma donde tiene lugar la traducción y la transcripción inversa. Integrasa y ADNc del transposón toman contacto con una nueva localización quizá durante la división celular.

2.3.3. Efectos celulares y control de la transposición

Tanto las células procarióticas como eucarióticas contienen una variedad de transposones que se trasladan o se copian en distintos lugares. Durante la transposición no replicativa se produce una ruptura doble en el lugar original del transposón y una mutación insercional en el lugar de aterrizaje. Además, la presencia de transposones en la célula favorece otro tipo de reorganizaciones derivadas de su múltiple presencia. Así, dos transposones relativamente cercanos y con la misma orientación favorecerán una recombinación homóloga que puede resultar en una delección de la región comprendida entre ambos (figura 2.25). Si, por el contrario, una región genómica está flanqueada por dos transposones con orientaciones opuestas, la recombinación entre ambos puede originar una inversión de la región interna (figura 2.25).

El control de la transposición es importante para la célula y para el transposón. Éste ha de

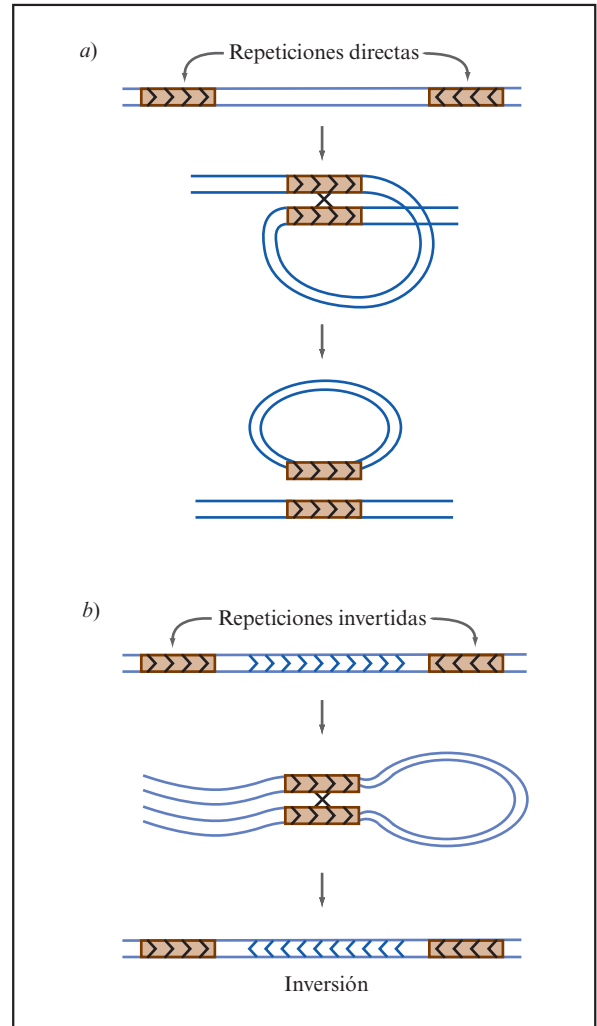


Figura 2.25. Reorganizaciones cromosómicas derivadas de la recombinación entre dos transposones próximos. En a) resultado de la recombinación entre un par de transposones con la misma orientación, y b) con orientaciones opuestas.

mantener un número mínimo para sobrevivir, pero si se hace demasiado abundante puede provocar su forzado suicidio matando a la célula hospedadora. La transposición, tanto en procariotas como en eucariotas, está controlada por ARN de interferencia; los dos mecanismos básicos son 1) metilación de los terminales repetidos,

y 2) ARNi contra la transposasa o el genoma del transposón. Estos mecanismos no son excluyentes. El simple hecho de que el número de genomas del transposón aumente, o se incremente la concentración de los ARN correspondientes a la transposasa, es suficiente para desencadenar una respuesta interferente, también denominada inhibición multicopia.

Si bien la metilación de los terminales repetidos en las dos hebras del ADN de un transposón inhiben la transposición, la metilación de sólo una de las hebras no es tan eficiente. Esta situación de hemi-metilación ocurre inmediatamente después de la replicación y antes de que las metilasas actúen sobre la nueva hebra de ADN. Así la transposición aumenta después de la replicación.

En arqueas y bacterias se ha descrito un sofisticado sistema inmune contra virus, plásmidos y muy probablemente también transposones que se parece mucho a los sistemas de ARN de interferencia que se presume controlan los transposones de eucariotas. En el genoma de *Escherichia coli* se identificaron, allá por 1987, unas secuencias repetidas de 29 nucleótidos separadas por otras únicas de 32 nucleótidos cuya función permaneció desconocida hasta que se demostró que estas secuencias únicas, separadoras de repetidas, eran idénticas a trocitos de plásmidos, bacteriófagos o transposones. La idea que tenemos a día de hoy es que cuando un ADN de dudosa procedencia llega a una bacteria, aunque su ADN se corte y destruya, un trocito de su genoma se mantendrá, se integrará y se almacenará en un lugar concreto del genoma del hospedador, entre unidades repetidas, constituyendo una especie de vacuna inmunológica contra expansiones indeseadas o nuevas infecciones (figura 2.26). En este almacén inmunológico encontraríamos un promotor seguido de una repetición de 29 pb, seguida de un separador de 32 pb idéntico a un bacteriófago, seguido de otra repetición, seguida de otro separador de 32 pb idéntico a un plásmido y así sucesivamente. Este almacén inmunológico se conoce como locus CRISPR (del inglés: *clustered regularly interspaced short palindromic repeat*) y va incrementando el número de separadores de 32 pb con nuevas invasiones de ADN de

procedencias dudosas. Las nuevas adiciones se hacen al lado del promotor, por duplicación de la secuencia repetida e integración a modo de transposición mediada por la endonucleasa Cas1 (del inglés: *CRISPR-associated*). La transcripción del locus CRISPR origina un ARN que puede ser procesado con ligeras variantes (tipos I, II y III), pero en todos los casos originará un ARN de interferencia que lleve la secuencia separadora complementaria al ADN invasor, 8 nucleótidos de la secuencia repetida en el extremo 5' y los 21 nucleótidos restantes de la secuencia repetida en el extremo 3'. La presencia de las secuencias repetidas en el ARNi es muy importante para distinguir entre el trocito de ADN invasor almacenado en el locus CRISPR que se debe respetar y cualquier otro ADN con homología de secuencia y ausencia de secuencia repetida, que ha de ser eliminado. El proceso es igualmente válido para eliminar ARN complementarios a ARNi procedentes, por ejemplo, de virus ARN o retrotransposones. El procesamiento y maduración de los ARNi así como la eliminación de los ADN o ARNs complementarios, corre a cargo de las endonucleasas Cas3, Cas9 y en algunos casos de la ARNasa III que corta ARN de doble hebra, todas ellas codificadas por la célula hospedadora (figura 2.26).

El control de la transposición en eucariotas se ha estudiado principalmente en *Caenorhabditis elegans* (gusano), *Drosophila melanogaster* (mosca del vinagre), el pez cebra y el ratón. De todos estos estudios se deduce que el sistema de ARN de interferencia, conocido como Piwi y semejante al descrito para bacterias, podría ser el responsable del silenciamiento de muchas familias de transposones en estos organismos. Piwi (del inglés: *P-element induced wimpy testis*) se descubre en relación con los elementos móviles P de *Drosophila* y forma parte de una familia de proteínas reguladoras encargadas de mantener las células troncales en estado des-diferenciado. En *Drosophila* hay regiones del genoma particularmente enriquecidas en transposones, restos de transposones y elementos repetidos de dudosa procedencia (al estilo del almacén inmunológico CRISPR). Estas regiones se transcriben, proce-

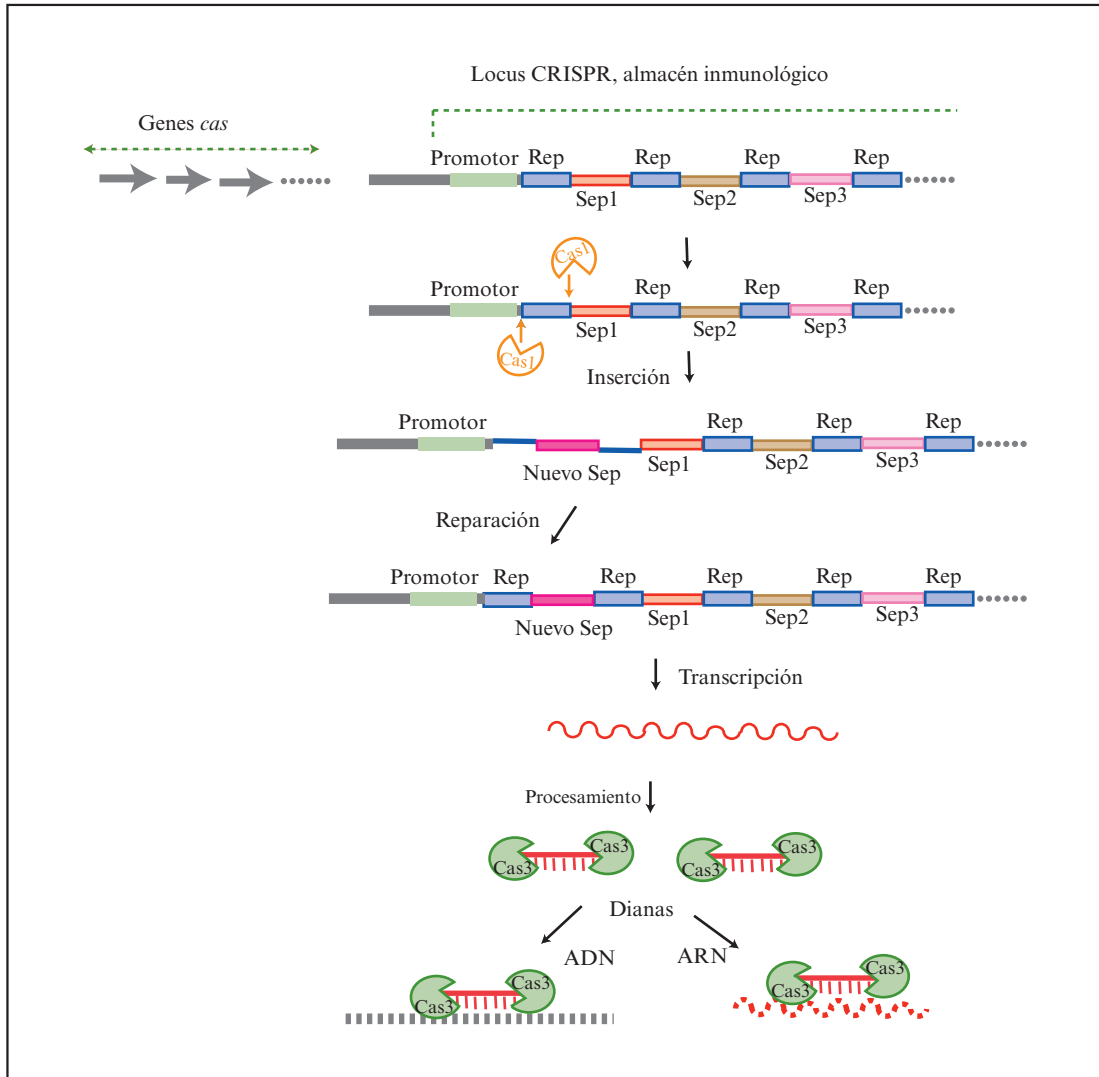


Figura 2.26. Posible control de la transposición mediante el sistema inmunológico bacteriano CRISPR (conjunto de repeticiones palindrómicas cortas regularmente distribuidas). A la izquierda el locus de los genes *cas*, a la derecha el locus CRISPR con un promotor seguido de una serie de secuencias repetidas (**Rep**) interrumpidas por secuencias separadoras (**Sep**) que serían idénticas a un pequeño trozo de un plásmido, bacteriófago o transposón. Esquema de la adquisición de un nuevo **Sep** por ruptura del **Rep** más cercano al promotor, separación de las dos hebras de **Rep**, unión al nuevo **Sep** y reparación. La enzima **Cas1** facilita el proceso. La transcripción y procesamiento del locus CRISPR origina ARN de interferencia que buscan y degradan ADN o ARN complementarios. **Cas3** facilita este proceso.

san y maduran constituyendo los ARN-pi, así llamados porque se asocian con la proteína nuclear Piwi. El sistema Piwi se parece mucho también al descrito para los ARNi mediados por RISC (véase capítulo 1, apartado 1.2.2 y figura 1.6), pero Piwi se asocia preferentemente a ARN de hebra simple. Los ARN-pi tienen entre 24 y 32 nucleótidos de longitud, algo más largos que los asociados a RISC (21 nucleótidos) y más al estilo de CRISPR. Una vez asociados a Piwi, se unen a la proteína HP1a de la heterocromatina. Los mutantes *piwi* acumulan ARN correspondientes a transposones indicando que el silenciamiento de transposones mediado por Piwi actúa a niveles tanto de transcripción como de postranscripción. Este conjunto Piwi/ARN-pi/HP1a se une al ADN complementario al ARN-pi y recluta a metiltransferasas que metilarán la histona H3lisina9 (H3K9me) y silenciarán al transposón (figura 2.27). Piwi/ARN-pi también podría asociarse con el ARN naciente y facilitar su degradación. En el citoplasma, el complejo Piwi/ARN-pi puede interactuar con el mensajero correspondiente y con factores de iniciación de la traducción inhibiendo ésta e induciendo la degradación del mensajero del transposón. Este mecanismo se ha probado tanto para *Drosophila* como para ratón, donde los homólogos de Piwi denominados Mili y Miwi2 son los encargados de unirse a los transcritos de transposones y silenciarlos en el núcleo o degradarlos en el citoplasma. Ambos modos de acción no son excluyentes. En humanos la homóloga de Piwi se llama Hiwi.

Los genomas de los distintos organismos están en lenta pero continua evolución y los transposones han sido un elemento importante para generar variaciones y diversidad. La inserción de un elemento móvil en una región reguladora puede tener consecuencias pleiotrópicas a distintos niveles sobre los genes regulados. La mutagénesis insercional no siempre es deletérea.

La movilidad de transposones somática puede estar relacionada con algunos tipos de cánceres.

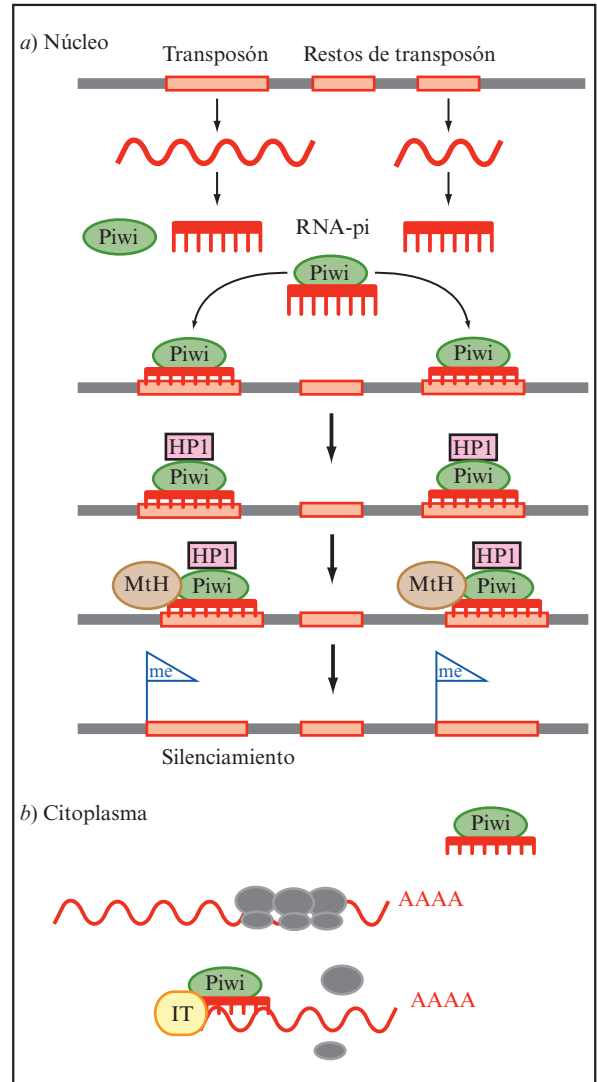


Figura 2.27. Control de la transposición mediada por Piwi (testículos endebles inducidos por elementos P) en *Drosophila*. En el núcleo *a*) los ARN-pi se asocian a la proteína nuclear Piwi y reclutan a la proteína heterocromática HP1 y a metiltransferasas (MtH) que metilarán la histona H3K9me y silenciarán el transposón. En el citoplasma *b*) Piwi/ARNpi puede interactuar con el propio ARN o con factores de iniciación de la traducción (IT) para inhibir el proceso de traducción y degradar el mensajero.

LECTURAS COMPLEMENTARIAS

A major epigenetic programming mechanism guided by
piRNAs **Xiao A Huang, Yin Hang, Sara Sweeney,**

Debasish Raha, Michael Snyder and Haifan Lin.
Developmental Cell, **24**: 502-516. 2013.

3

Enzimas y técnicas de utilización general en manipulación genética

Muchas áreas de la biología molecular se han beneficiado enormemente de las técnicas englobadas dentro del término genérico de ingeniería genética y referidas indistintamente como de clonaje, ADN recombinante o manipulación genética. Estas técnicas permiten: 1) mantener en una misma unidad replicativa genes procarióticos y eucarióticos; 2) crecer grandes cantidades de estas unidades nuevas, y 3) integrar estos genes en el genoma de otro individuo. El descubrimiento de las enzimas de restricción fue en su día un hecho clave para el desarrollo de la ingeniería genética. Las enzimas de restricción representan uno de los sistemas de autoprotección bacteriana a la invasión por fagos (bacteriófagos). Se comprobó que las bacterias podían cortar el ADN de un fago una vez en el interior de la bacteria, siempre que éste hubiera crecido en una cepa bacteriana distinta. Venía a ser un tipo de sistema inmunológico rudimentario. Pero la ruptura del ADN invasor no se realizaba siempre al azar, sino que enzimas concretas cortaban en el mismo lugar, siempre reconocían las mismas dianas. Estas endonucleasas representaron, por tanto, una herramienta de valor incalculable para fragmentar el ADN por sitios fijos. La ingeniería genética ha servido para aislar, clonar y expresar prácticamente cualquier gen eucariótico en un momento en el que no había técnicas alternativas de amplificación génica. Antes del desarrollo de la ingeniería genética no era posible aislar un gen concreto eucariótico en cantidades suficientes para su estudio molecular o el de su producto. En este capítulo estudiaremos las distintas enzimas que se utilizan para realizar

todo tipo de manipulaciones genéticas, así como dos técnicas fundamentales: la reacción en cadena de la polimerasa (PCR) y la secuenciación.

3.1. DEGRADACIÓN Y SÍNTESIS DE ÁCIDOS NUCLEICOS *IN VITRO*

In vivo, la síntesis y degradación de los ácidos nucleicos constituye su metabolismo. *In vitro* se degradan y sintetizan ácidos nucleicos controladamente con el fin de purificar o manipular determinadas clases de moléculas que se utilizarán como base en experimentos diversos de ingeniería genética.

La degradación de los ácidos nucleicos puede ser química o enzimática.

3.1.1. Degradación química

Degradación por pH

A pH muy bajos (1 o menos) se produce una hidrólisis de los grupos fosfodiéster, tanto en la molécula de ADN como en la de ARN, acompañada de la ruptura de enlaces N-glucosídicos entre las bases y los azúcares. Este tipo de degradación se utiliza cuando se quiere determinar la composición de bases de un ácido nucleico.

A pH 4 se produce una depurinización por ruptura de los enlaces N-glucosídicos entre el azúcar y las purinas, afectando más a la molécula de ADN que a la de ARN. A pH 7 se invierte la ten-

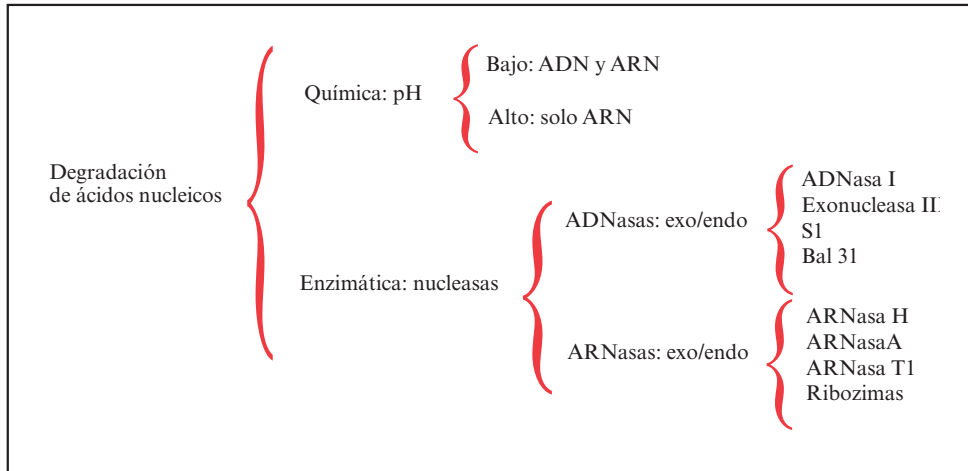


Figura 3.1. Esquema de algunos métodos de degradación de los ácidos nucleicos.

dencia y el ADN se hace unas 200 veces más resistente a la hidrólisis que el ARN; a pH aún más alto el ADN es prácticamente resistente a la hidrólisis, pero no así el ARN, que a pH 11 y 37 °C se hidroliza totalmente en pocos minutos a una

mezcla de nucleótidos 3' y 2' según muestra la figura 3.2. El ADN depurinizado también es sensible a pH elevado y se rompe por el enlace fosfodiéster 5' correspondiente a la desoxirribosa del lugar donde falte la purina.

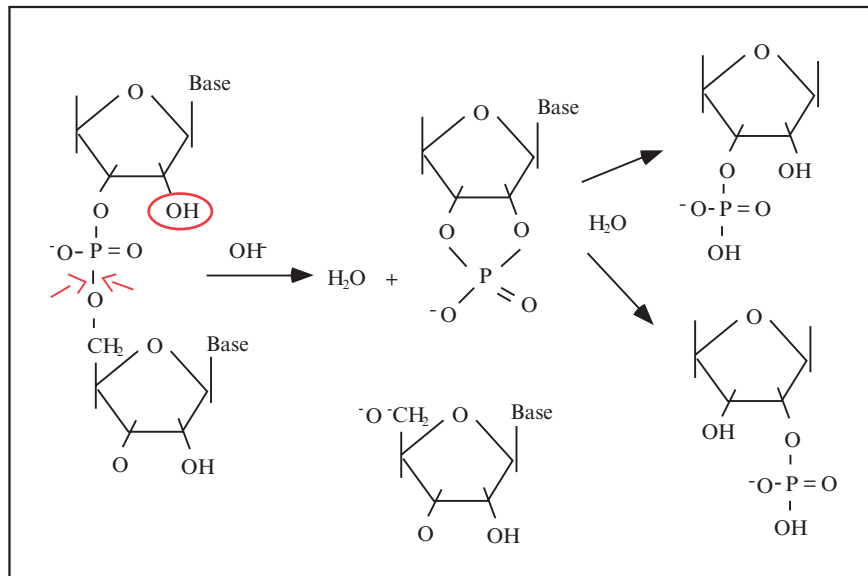


Figura 3.2. Degradación del ARN a pH alto por mediación del grupo 2' hidroxilo libre, creando así un intermediario cíclico de fosfato que se descompone a su vez en una mezcla de nucleótidos 3' y 2'.

Tijeras químicas

Se puede manipular el ADN utilizando una especie de tijeras químicas que cortan el ADN en lugares concretos definidos por oligonucleótidos complementarios a cualquier secuencia. Cada tijera estaría formada por dos componentes: 1) la unidad de reconocimiento formada por un oligonucleótido modificado químicamente, de tamaño variable y secuencia definida y 2) óxido de cerio (CeIV), un potente oxidante, a su vez unido al agente quelante EDTA (etilen/di-amino/tetra-acético) que romperá el enlace fosfo-di-éster del esqueleto de un ADN de hebra sencilla. La modificación del oligonucleótido consiste en sustituir los enlaces fosfodiéster por enlaces amida y las adeninas y timinas por 2,6-diaminopurina y 2-tiouracilo, respectivamente. Este pseudo-oligonucleótido

no tiene carga negativa y puede formar un híbrido estable con una molécula de ADN de hebra sencilla o una triple hélice con un ADN de hebra doble (figura 3.3a). Cuando se quiere cortar un ADN de doble hebra, se utilizan pseudo-oligonucleótidos complementarios a las dos regiones adyacentes al lugar de corte; cuando estas regiones aparean con el ADN correspondiente dejarán una pequeña región de ADN de hebra simple. En este punto se añade el componente de Cerio/EDTA que romperá en la región de ADN de hebra simple de cada hebra del ADN (figura 3.3b). La eficiencia de corte será proporcional a la concentración de Cerio/EDTA añadido. La total complementariedad entre el ADN y los pseudo-oligos es esencial para que se produzca el corte; una variación en un solo nucleótido impedirá la escisión. Estas tijeras pueden usarse tanto en ADN aislados como *in vivo*.

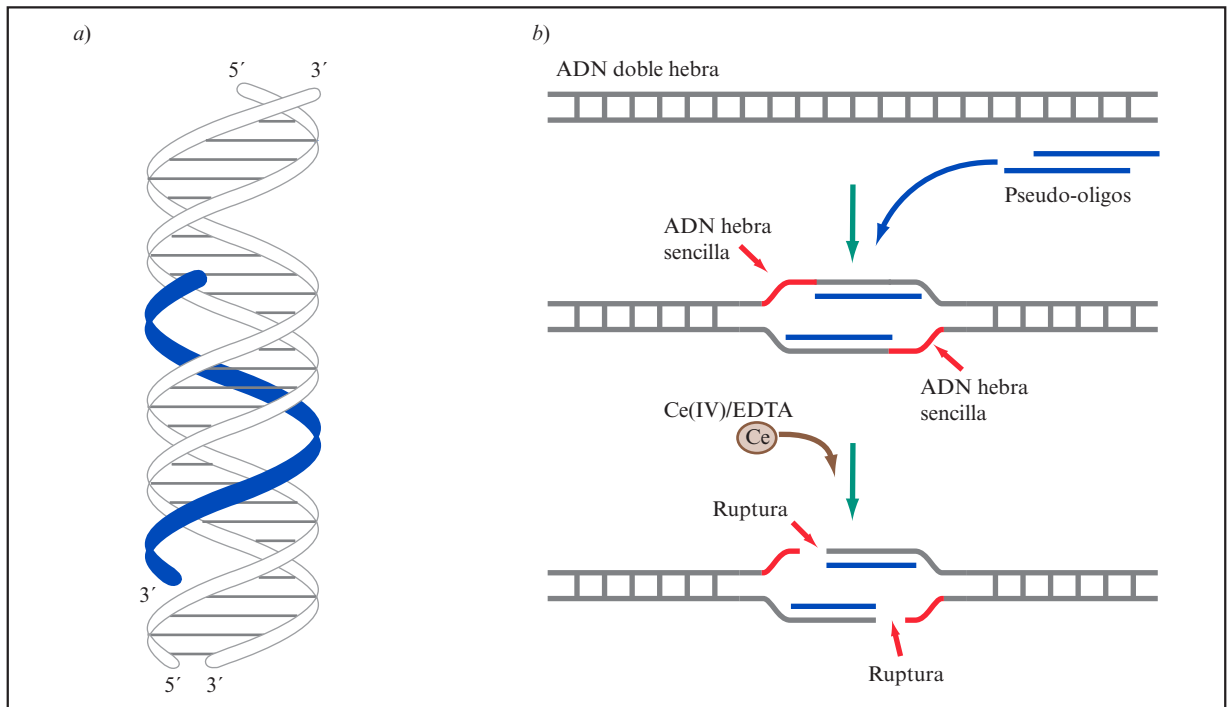


Figura 3.3. Modo de acción de las tijeras químicas. a) Reconocimiento e interacción de un pseudo-oligo con un ADN de doble hebra y b) Mecanismo de corte y reconocimiento: primero se añaden los pseudo-oligos complementarios a las regiones deseadas, después el complejo Cerio/EDTA. Las rupturas se producen en las regiones de ADN de hebra sencilla creadas en los extremos de los pseudo-oligos (flechas).

3.1.2. Degradación enzimática

Las enzimas que degradan los ácidos nucleicos se denominan en general nucleasas y pueden ser de dos tipos dependiendo de si actúan a partir de los extremos (exonucleasas) o dentro de la cadena (endonucleasas).

ADNasa I (endonucleasa I de *E. coli*, endonucleasa I bovina)

Endonucleasa que hidroliza ADN limpio y purificado de hebra simple y doble así como cromatina. Prefiere cortar detrás de pirimidín-nucleótidos y origina extremos 5'-P, 3'OH. En pre-

sencia de Mg^{+2} ataca al azar cualquiera de las dos hebras y produce mellas. En presencia de Mn^{+2} corta ambas hebras aproximadamente en el mismo lugar (figura 3.4).

También se utiliza en estudios de protección de ADN *in vivo* tratando núcleos eucarióticos aislados con concentraciones de ADNasa I suficientes para detectar las regiones sensibles o resistentes al ataque con esta enzima que representarán partes accesibles o inaccesibles, respectivamente, a factores de transcripción. Dentro de las regiones accesibles, las que estén protegidas por los factores de transcripción concretos en un momento dado se harán internamente resistentes (véase apartado 1.8 y figura 1.17).

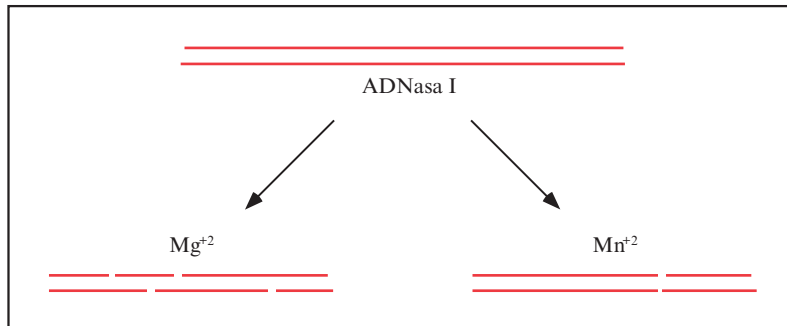


Figura 3.4. Modo de acción de la ADNasa I sobre ADN desnudo, produciendo mellas en presencia de magnesio y cortes limpios con manganeso.

Exonucleasa III de *E. coli*

Actúa sobre extremos 3'OH de ADN de hebras dobles, lineares o circulares. No degrada ADN de hebra sencilla (figura 3.5).

Exonucleasa del fago λ

Esta exonucleasa se purifica de células de *E. coli* infectadas por el bacteriófago λ y actúa sobre extremos 5' fosfato (P), liberando mononucleótidos. No hidroliza extremos 5'OH.

Nucleasa S1

Degrada ADN de hebra sencilla y ARN (es exo y endo). En grandes dosis también degrada

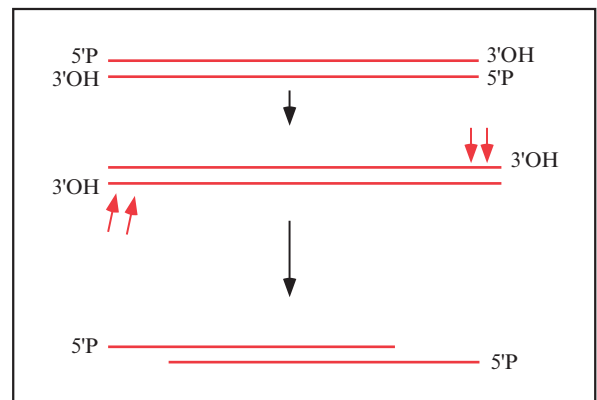


Figura 3.5. Modo de acción de la exonucleasa III de *E. coli* degradando extremos 3'OH de ADN de doble banda.

ADN de hebra doble. Requiere Zn^{+2} y actúa preferentemente a pH bajo (4 es óptimo), lo que es un inconveniente, ya que a este pH hay depuración de las bases del ADN.

Nucleasa BAL31

Es preferentemente una exonucleasa que elimina mononucleótidos a partir de extremos 3'OH, pero también es endonucleasa, por lo que elimina las hebras sencillas creadas por la acción exonucleásica (figura 3.6). La reacción es totalmente dependiente de calcio, por lo que se puede controlar el grado de degradación eliminando éste, a tiempos determinados, con agentes quelantes como el EGTA. Degrada bastante más rápido secuencias ricas en AT que en GC. También degrada con poca eficiencia ARN. Dado que produce terminales de distintos tamaños, algunos de hebra sencilla, será necesario reparar la molécula antes de su posterior utilización con alguna polimerasa que rellene regiones de hebra sencilla (la ADN polimerasa de T4 o el fragmento Klenow de la ADN polimerasa 1 de *E. coli*).

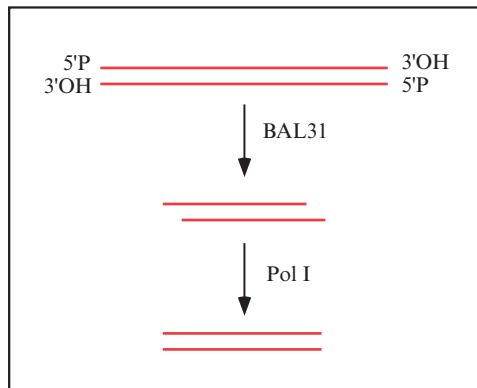


Figura 3.6. Modo de acción de la enzima BAL31 y reparación con la ADN polimerasa I (Pol I) de *E. coli*.

ARNasa H (Ribonucleasa H)

Degrada el ARN de los híbridos ADN-ARN dejando el ADN intacto.

ARNasa A (Ribonucleasa A)

Es una endonucleasa que rompe moléculas de ARN de banda simple por los residuos de pirimidinas. Se aísla del páncreas de muchos mamíferos y se denomina también pancreática.

ARNasa T1

Es también una endorribonucleasa que ataca específicamente guaninas.

ARNasa III

Se une específicamente y corta ARN de doble hebra. En bacterias está implicada en el procesamiento del ARN ribosomal y otros ARN pequeños, como el sistema CRISPR, en eucariotas procesa los ARN de doble hebra largos en ARN de interferencia cortos.

Exorribonucleasas

La fosfodiesterasa de veneno de serpiente es una exorribonucleasa que se enlaza a extremos 3'OH de ARN, pero también degrada ADN.

La fosfodiesterasa de bazo es una exorribonucleasa que requiere extremos 5'OH libres de ARN o ADN.

Ribozimas

Son moléculas de ARN con actividad catalítica que rompen específicamente otros ARN. Hay ribozimas autocatalíticas que actúan sobre su propia molécula eliminando sus intrones (intrones de tipo I); otras actúan sobre otros ARN (en trans) y constan de dos componentes esenciales: secuencias específicas de reconocimiento, que aparean con secuencias complementarias en el ARN diana y región catalítica, que viene a ser un centro activo y produce la ruptura del ARN diana durante el apareamiento. La ruptura trae como consecuencia la actuación sucesiva de otras nucleasas sobre los extremos anormales recién creados en el ARN y que lleva a la destrucción total de la molécula. Son ejemplos de ribozimas la ribonucleasa P humana y varias obtenidas de viroides (partículas

semejantes a virus) de plantas. Muchas de estas últimas son del tipo «cabeza de martillo», así denominadas por la forma del componente catalítico (figura 3.7). El reconocimiento se produce por la complementariedad de las secuencias que flanquean la región catalítica; la secuencia diana debe

incluir el triplete GUC (o una variante muy próxima) en la región central. La ruptura se produce entre la C del triplete y la siguiente base. Las ribozimas pueden ser diseñadas *in vitro* contra ARN específicos y se están utilizando en la actualidad con este fin.

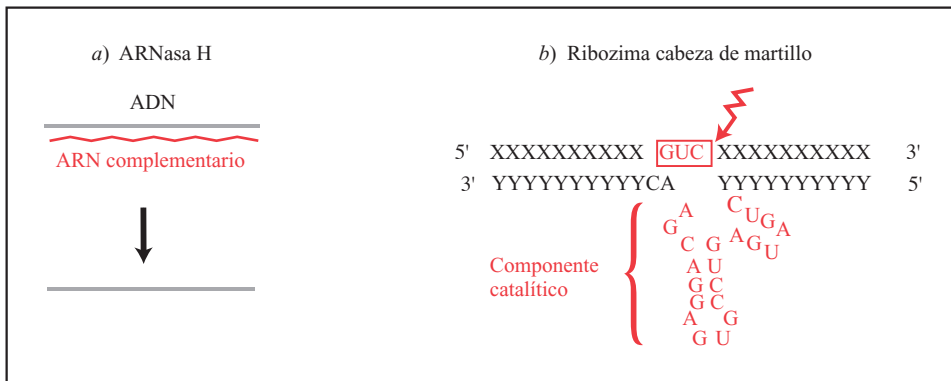


Figura 3.7. Modo de acción de a) la ARNasa H degradando el ARN de un híbrido ADN-ARN, y b) la ribozima cabeza de martillo cortando un ARN diana por el lugar que indica la flecha quebrada.

3.1.3. Síntesis de ácidos nucleicos

Los ácidos nucleicos se sintetizan *in vitro* por una serie de enzimas llamadas genéricamente ADN polimerasas si sintetizan ADN y ARN polimerasas si sintetizan ARN. La mayoría de estas enzimas requieren un molde y sintetizan una molécula complementaria al molde. Las polimerasas de ADN utilizadas con mayor frecuencia son: la ADN polimerasa I de *E. coli* (la holoenzima); la polimerasa Klenow, que es el fragmento de mayor tamaño de la ADN polimerasa I de *E. coli* cuando se trata con la proteasa subtilisina; la polimerasa del bacteriófago T7, la secuenasa, que es una versión modificada de la ADN polimerasa de T7, y la ADN polimerasa termoestable Taq (de *Thermus aquaticus*). La mayoría de las ADN polimerasas tienen una gran preferencia por moldes de ADN, pero también pueden copiar ARN con menor eficiencia, excepto por la transcriptasa inversa que prefiere ARN como molde. Hay también una ADN polimerasa que no necesita molde, ya que

añade nucleótidos exclusivamente a los extremos de las cadenas preexistentes: es la transferasa terminal. Todas ellas requieren iones Mg^{+2} . Las polimerasas de ARN reconocen determinadas secuencias en el ADN denominadas promotores e inician a continuación la síntesis de ARN-complementario a una de las hebras del ADN.

Las características particulares de polimerasas de ADN y ARN se detallan a continuación.

ADN polimerasa I de *E. coli* (holoenzima)

Consta de una cadena polipeptídica que funciona como polimerasa y exonucleasa en la dirección $5' \rightarrow 3'$ y exonucleasa en la $3' \rightarrow 5'$. También tiene actividad de ARNasa H, que es esencial para la viabilidad de la enzima, pero no se utiliza mucho *in vitro*. Esta polimerasa se emplea en el proceso de marcaje radiactivo denominado «traslado de la mella», que es uno de los procesos más comunes de sustituir un ADN por su copia radiactiva (figura 3.8). También se pueden utilizar

marcajes no radiativos empleando biotina o el hapteno esteroideo digoxigenina (DIG) en alguno de los precursores. El producto se detecta por luminiscencia o color.

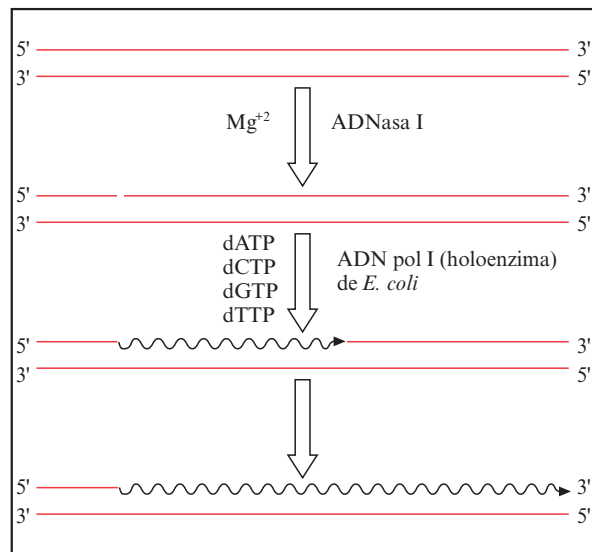


Figura 3.8. Proceso de traslado de la mella. Ruptura por acción de la ADNasa I y síntesis por la ADN pol I de *E. coli*.

El procedimiento se basa en el tratamiento de un ADN de hebra doble con una pequeña cantidad de ADNasa I para que se produzcan algunas mellas al azar que sirvan de enganche a la ADN pol I. En presencia de magnesio y deoxinucleótidos trifosfato (dNTP), la ADN polimerasa va digiriendo nucleótidos en el sentido 5' → 3' (exonucleasa) y sustituyéndolos por los dNTP del medio. Uno o más de estos dNTP pueden ser radiactivos, con lo que marcaríamos el ADN sintetizado, facilitando su utilización como «sonda» en experimentos de hibridación de ácidos nucleicos.

Fragmento Klenow de la ADN polimerasa I de *E. coli*

Representa la holoenzima a la que se le ha eliminado la actividad exonucleasa 5' → 3'. La acti-

vidad exonucleásica de la holoenzima es muy activa y muchas veces elimina los fosfatos 5' de los extremos de moléculas de ADN que se pensaban utilizar como sustratos para uniones con la enzima ligasa, la cual no podrá utilizarlos sin el fosfato 5'. Por otra parte, esta actividad exonucleásica se puede eliminar con relativa facilidad por tratamiento proteolítico de la holoenzima sin afectar la actividad polimerasa. Sintetiza ADN en huecos o terminales de hebra sencilla (figura 3.9a). Esta enzima también se utiliza en la extensión del iniciador o iniciación al azar; el método utiliza ADN circular o lineal de banda doble o sencilla, y consiste en mezclar el ADN con un exceso molar de moléculas iniciadoras o cebadoras. La reacción, en este caso, se lleva a cabo a pH 6,6 porque a este pH la actividad exonucleasa 3' → 5' es muy débil y no se corre peligro de degradar los oligonucleótidos utilizados como iniciadores o cebadores (figura 3.9b).

La enzima Klenow se utiliza también en la síntesis de la segunda hebra del ADNc (ADN copia), en secuenciación del ADN por el método de terminación de cadena por di-deoxinucleótidos (Sanger) y en experimentos de mutagénesis dirigida.

La ADN polimerasa del bacteriófago T7 y la secuenasa

La primera se induce en *E. coli* después de la infección del fago T7. En realidad, es un complejo formado por dos proteínas íntimamente unidas: la proteína del gen 5 del bacteriófago y la tiorredoxina del huésped. Es una polimerasa muy eficiente que posee actividades 5' → 3' y 3' → 5' exonucleasa como las otras polimerasas. La actividad exonucleásica 3' → 5' (correctora de errores), al ser más potente que la de otras polimerasas, pone en peligro la estabilidad de los cebadores que en muchos casos son sintéticos y no aparean con perfección total. La actividad exonucleasa 3' → 5' radica en el extremo amino de la molécula y se puede inactivar incubando la enzima durante varias horas con un agente reductor (oxígeno molecular) a concentraciones bajas del ion de hierro. La enzima sin actividad exonucleásica se denomina secuenasa. También se obtiene

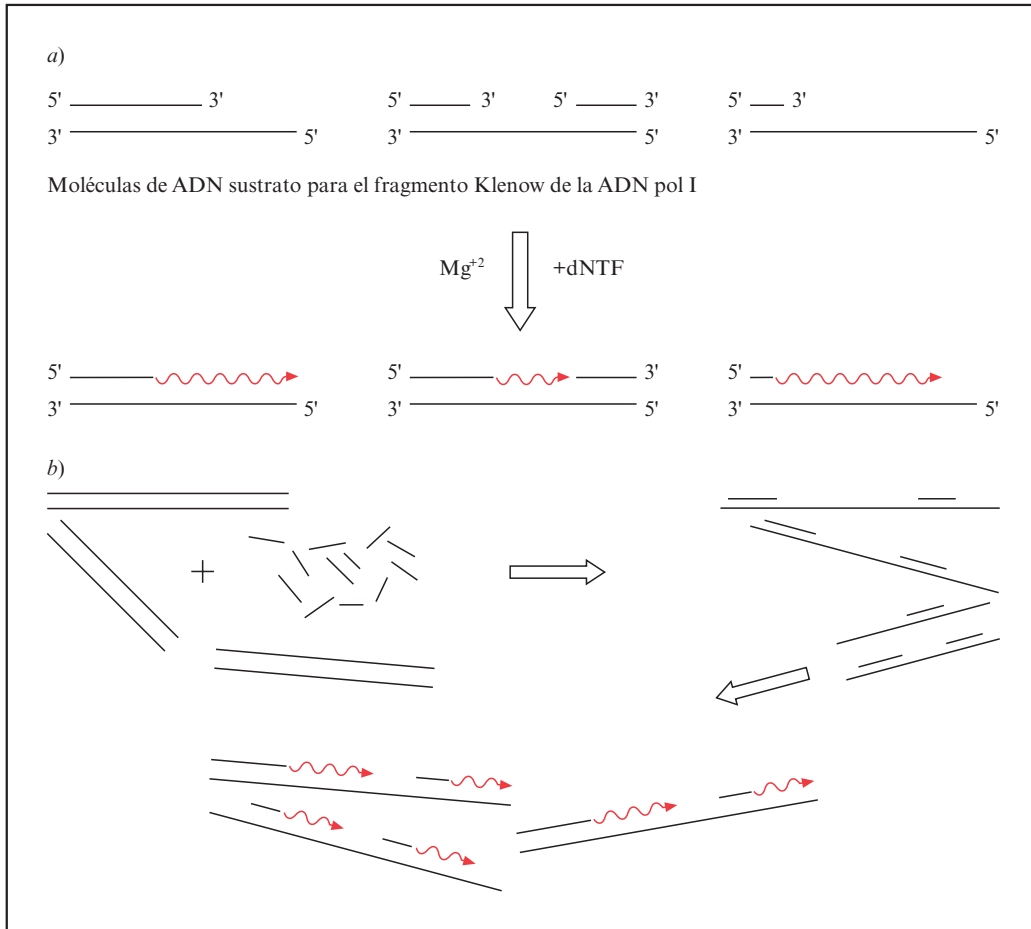


Figura 3.9. a) Modo de acción del fragmento Klenow de la ADN pol I. b) Extensión con iniciadores elaborados con secuencias al azar.

secuencas por ingeniería genética, eliminando del gen los nucleótidos responsables de la actividad exonucleásica.

La ADN polimerasa Taq (de *Thermus aquaticus*) y afines

Es termoestable y, aunque originariamente se aislaba de *Thermus aquaticus*, hoy se sintetiza por ingeniería genética. Su temperatura óptima es entre 75 y 80 °C; a 60 °C retiene sólo la mitad de su actividad, y a 37 °C, una décima parte. Requiere

Mg^{+2} y posee actividad exonucleásica 5' → 3', pero carece de la exonucleásica 3' → 5' correctora de errores. Se utiliza sobre todo para amplificar secuencias en la reacción en cadena de la polimerasa (PCR, del inglés: *polymerase chain reaction*) detallada más adelante en este capítulo.

Existen numerosas enzimas termorresistentes afines a la Taq con propiedades interesantes y utilizables en reacciones de PCR. La enzima Pfu, aislada de *Pyrococcus furiosus*, acumula unas 12 veces menos errores que la Taq. También la Vent (que debe su nombre a haber sido aislada de la

ventana por donde sale el chorro de agua caliente de un submarino) tiene actividad correctora de errores. La enzima Tth, que procede del microorganismo *Thermus thermophilus*, posee actividad de transcriptasa inversa; mientras que la T_f1 de *Thermus flavus* puede leer a través de ADN con estructuras secundarias.

La transcriptasa inversa

Sintetiza ADN utilizando como molde ARN. También necesita un «cebador» que provea un extremo 3'OH donde engancharse para iniciar la síntesis (figura 3.10).

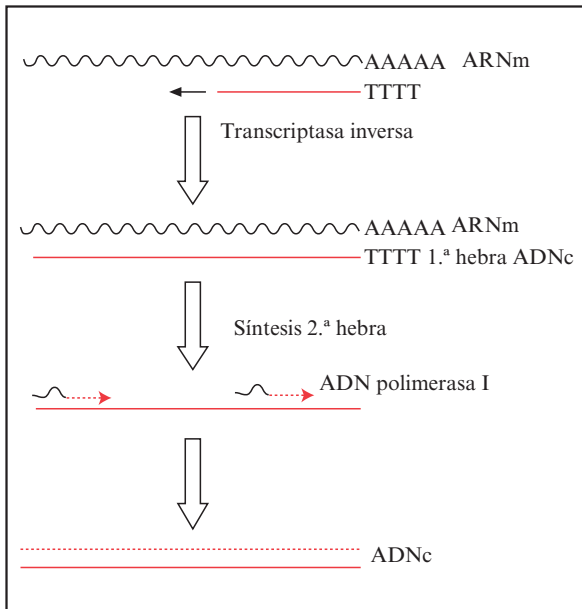


Figura 3.10. Modo de acción de la transcriptasa inversa. Utilización de un oligo-dT como iniciador y síntesis de la primera hebra. Degradación del ARN molde por álcali (OH⁻) y síntesis de la segunda hebra por la ADN polimerasa I.

Carece de actividad exonucleasa 3' → 5', que es la correctora de errores en otras polimerasas, por lo que acumula los fallos que produce; incorpora una base equivocada cada 500 bases aproximada-

mente. Posee actividad ARNasa H que puede degradar el ARN en forma de híbrido ARN-ADN. Durante mucho tiempo se empleó la transcriptasa inversa aislada del virus que produce la mieloblastosis en las aves, pero su poderosa actividad ARNasa interfería muchas veces con la estabilidad del ARN molde antes de que se llegara a completar la copia del mismo. Actualmente se utiliza preferentemente la transcriptasa inversa obtenida por ingeniería genética a partir del gen clonado del virus Moloney que produce la leucemia de ratón; esta última enzima posee una ARNasa H mucho más débil.

La transferasa terminal

Es ésta una extraña ADN polimerasa que a partir de deoxinucleótidos trifosfato (dNTP) añade nucleótidos a extremos 3'OH de cadenas de ADN; prefiere extremos protuberantes, pero también trabaja sobre terminales romos. La enzima solo se encuentra en bacterias, por lo que se piensa que pudo haber tenido importancia desde el punto de vista evolutivo en los orígenes de los primeros ácidos nucleicos. Severo Ochoa fue el descubridor de esta enzima a la que llamó polinucleótido fosforilasa y durante algún tiempo se pensó que era la polimerasa que copiaba el ADN bacteriano. Severo Ochoa recibió el Premio Nobel en 1959, por este y otros estudios; el premio fue compartido con Arthur Kornberg, descubridor de la ADN polimerasa I de *E. coli*. La polinucleótido fosforilasa se utilizó en los primeros experimentos de ingeniería genética para añadir colas de homopolímeros a las moléculas que se deseaba unir: guaninas a unas y citosinas a otras, por ejemplo. Necesita Mg⁺², Mn⁺² o Co⁺², pero curiosamente cuando añade purinas prefiere Mg⁺² y cuando añade pirimidinas funciona mejor con Co⁺², pudiendo añadir hasta varios miles de nucleótidos en cada extremo 3'OH (figura 3.11).

Las polimerasas de ARN

Necesitan determinados factores de transcripción además de Mg⁺² y ribonucleótidos (figu-

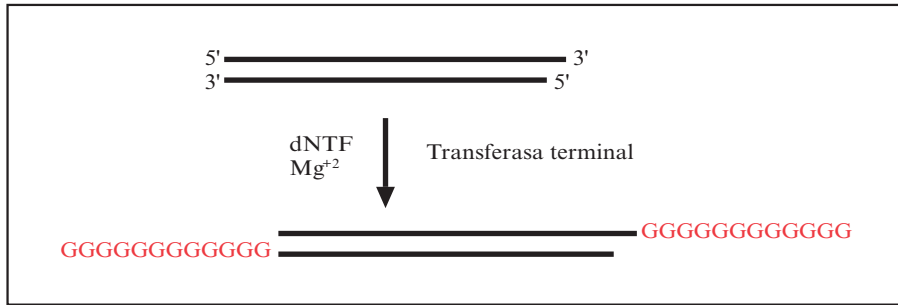


Figura 3.11. Modo de acción de la transferasa terminal, añadiendo en este caso homopolímeros a extremos 3'OH.

ra 3.12). Los promotores tienen un tamaño entre 50 y 100 nucleótidos y son específicos para cada ARN polimerasa. Estas enzimas también fueron estudiadas por Severo Ochoa y sus aportaciones

científicas contribuyeron a la concesión de su Premio Nobel en 1959. *In vitro* funcionan mal por su dependencia de factores de transcripción concretos y específicos para cada una de ellas.

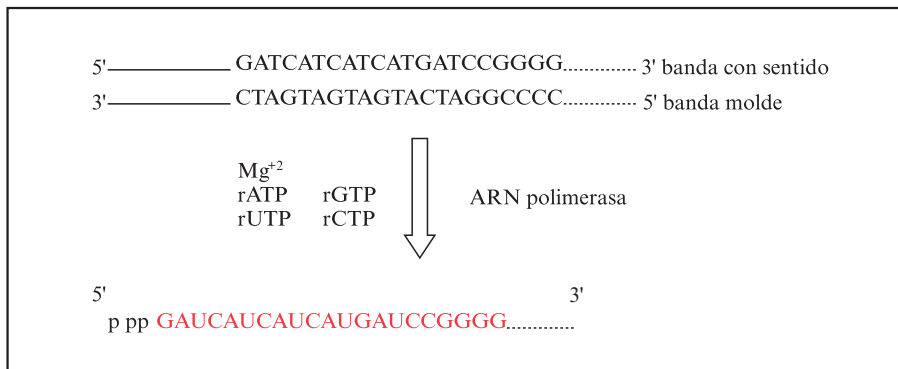


Figura 3.12. Modo de acción de las ARN polimerasas. Reconocimiento de un promotor y síntesis de ARN utilizando una de las hebras del ADN como molde.

En vectores de clonaje se incluyen con mucha frecuencia algún promotor específico de una o dos ARN polimerasas de los bacteriófagos SP6, T7 y T3. Los genes correspondientes a estas polimerasas han de ser también facilitados si se quiere una expresión muy eficiente de un gen de interés. El gen de interés debe colocarse a continuación del promotor en una región de policlonaje con múltiples dianas de diversas enzimas de restricción. SP6 infecta *Salmonella typhimurium*, T7 y T3 *E. coli*. Hay varios vectores de clonaje que se

obtienen comercialmente que poseen alguno de estos promotores en orientaciones opuestas para facilitar la transcripción en la dirección más adecuada en cada caso (figura 3.13). Las ARN polimerasas de los bacteriófagos T7 y T3 también son enormemente eficientes; por ejemplo, la polimerasa T7 sintetiza 200 nucleótidos de ARN por segundo. Todas ellas utilizan exclusivamente como molde ADN de hebra doble.

La ligasa se utiliza normalmente para unir fragmentos de ADN, enlazando extremos 3'OH

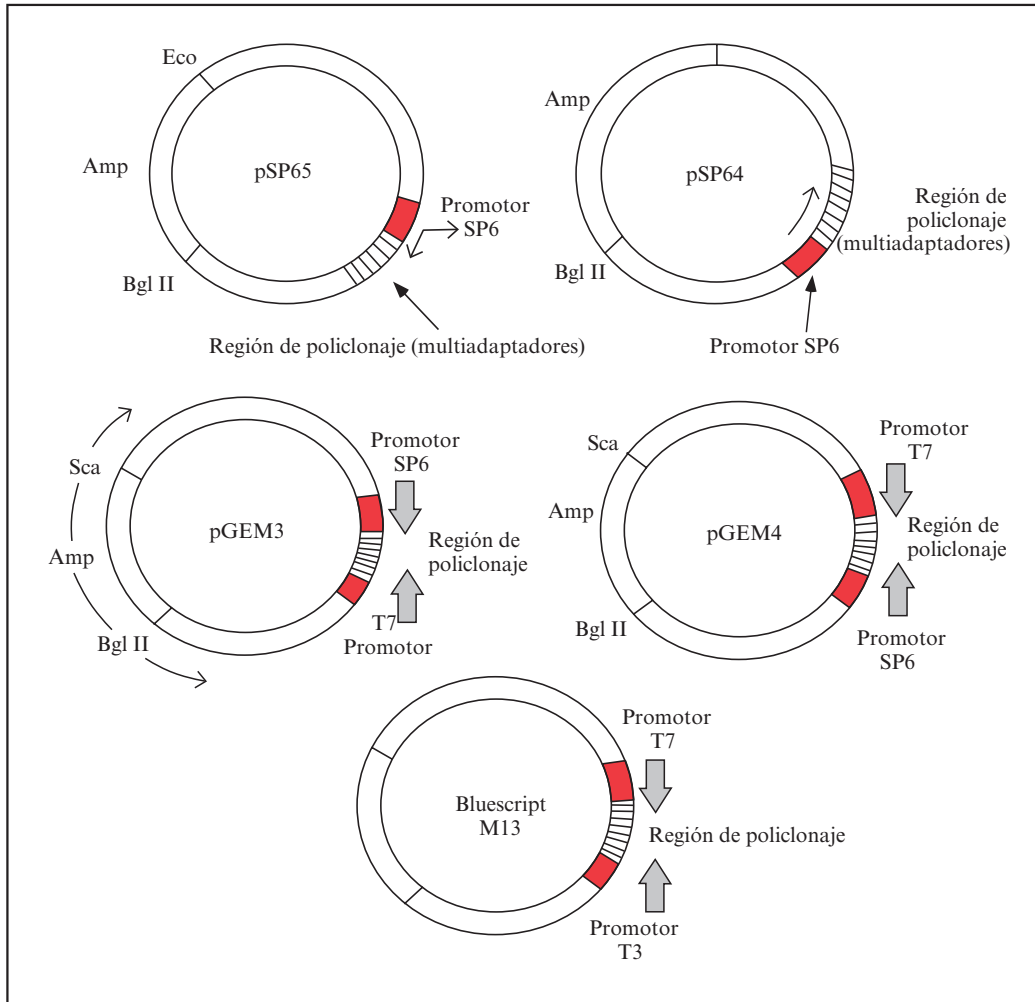


Figura 3.13. Vectores con promotores específicos para determinadas ARN polimerasas. El promotor SP6 en una orientación (pSP 65) o la opuesta (pSP 64) con respecto a la región de policlonaje. Promotores SP6 y T7 en orientaciones opuestas (pGEM3) e invertidas (pGEM4). Vector Bluescript de M13 con los promotores de T7 y T3 en orientaciones opuestas.

con 5'P. Une extremos protuberantes o romos y también puede unir moléculas de ADN con ARN, siendo la del fago T4 la más común (figura 3.14). Restaura mellas y en todos los casos requiere Mg^{+2} .

Las fosfatasas eliminan el fosfato de la región 5' de ADN, ARN y NTF (nucleótidos trifosfato), con lo que ya no serían sustrato para la ligasa (fi-

gura 3.14). Las más utilizadas son la de origen bacteriano y la de intestino de buey.

Las quinasas incorporan de nuevo el fosfato en el extremo 5'OH que han dejado las fosfatasas. Son muy útiles, transfiriendo el fosfato en posición γ del ATP marcado con ^{32}P a extremos 5'OH de moléculas de ADN o ARN que se desea marcar radiactivamente (figura 3.14).

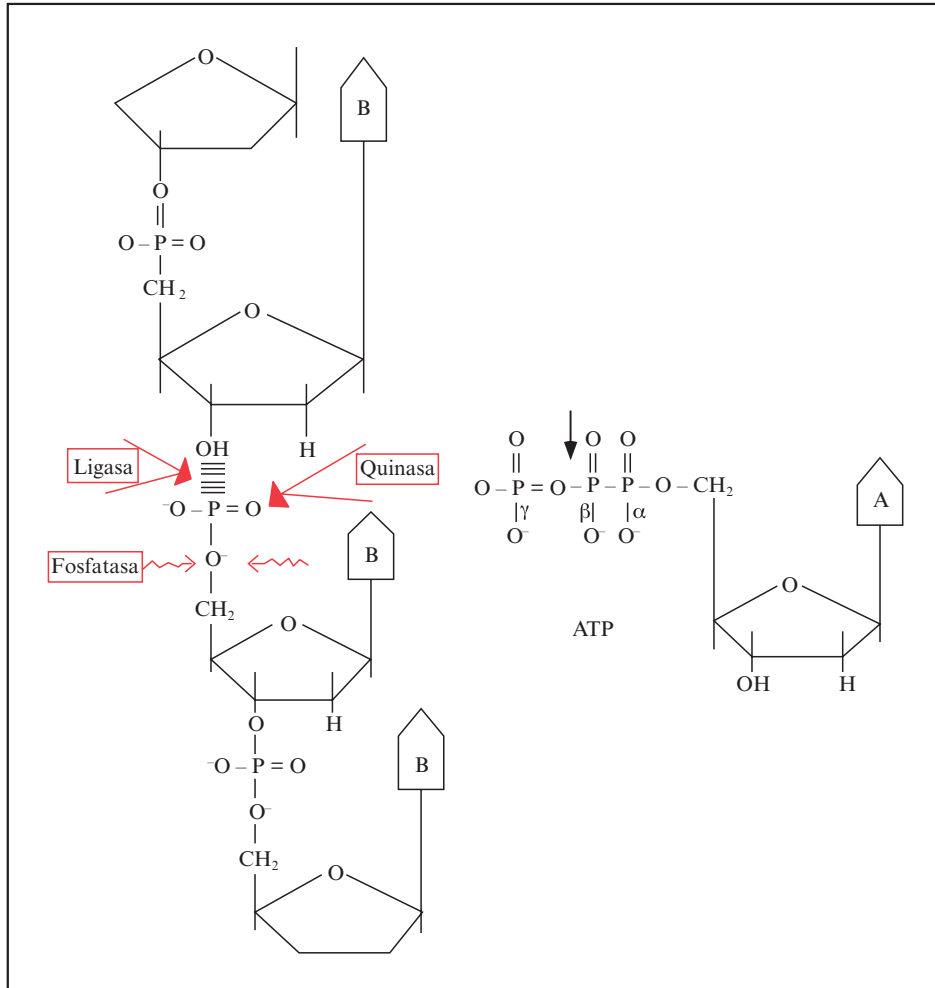


Figura 3.14. Modo de acción de: ligasas, enlazando extremos 3'OH con 5'P; fosfatasa, eliminando el fosfato en posición 5', y quinasas, añadiendo fosfatos a extremos 5'OH.

3.2. ENZIMAS DE RESTRICCIÓN

Son endonucleasas que reconocen dianas específicas en el ADN. Estas enzimas pueden cortar en las propias dianas o a una cierta distancia de ellas, siempre que determinadas bases de la secuencia de reconocimiento no estén metiladas.

Los sistemas de restricción y modificación se deben probablemente a un mecanismo de defensa de las bacterias frente a la invasión de determina-

dos bacteriófagos. Pudiera considerarse análogo al sistema inmune de organismos superiores, ya que protege a cada cepa de bacterias de la entrada de otro ADN que no pertenezca al mismo sistema. Su principal misión *in vivo* es defensiva. Es característico y probablemente exclusivo de bacterias y microorganismos.

La observación inicial que llevó a su descubrimiento se realizó en los años cincuenta (1950) al detectarse que algunas cepas bacterianas eran in-

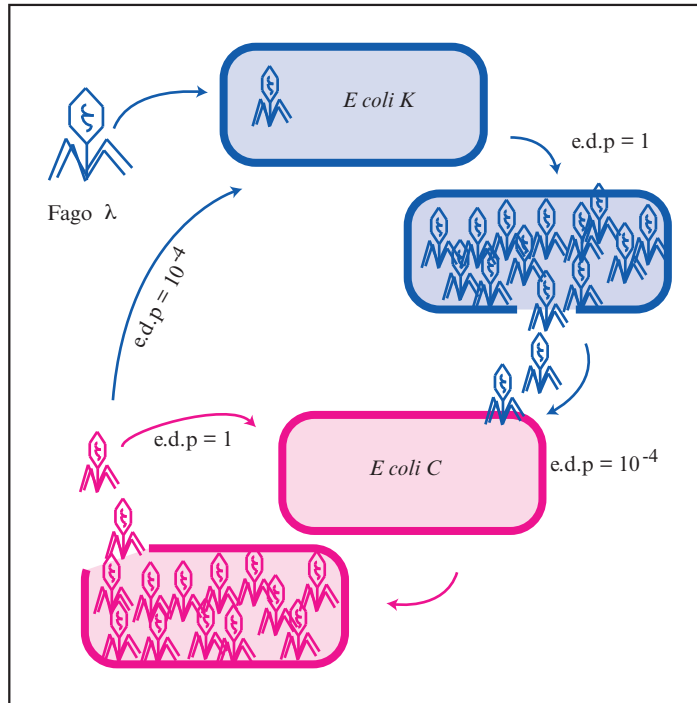


Figura 3.15. Sistema de restricción y modificación del fago λ en *E. coli* cepas K y C. La eficiencia del plaqueo (e.d.p.) del fago λ puede pasar de 1 en *E. coli* K a 10^{-4} en *E. coli* C, si se utilizan como hospedadores sucesivos. La restricción (degradación) en este segundo hospedador no es absoluta y los fagos que sobrevivan estarán modificados (metilados), pudiendo recuperar la e.d.p. que anteriormente tenían.

munes a la infección vírica. Este fenómeno se denominó restricción controlada por el hospedador. No era una propiedad genética y heredable, sino que dependía del último hospedador en el que el huésped (virus) hubiera crecido. Tuvieron que pasar, no obstante, veinte años más hasta que se entendió el mecanismo. Por el descubrimiento y caracterización de las enzimas de restricción, Arber, Smith y Nathans recibieron el Premio Nobel en 1978. La restricción ocurre porque la bacteria produce una enzima que degrada el ADN intruso, mientras que el propio está protegido porque las dianas de restricción están metiladas.

El sistema de restricción y modificación controlado por el hospedador se aprecia claramente cuando un bacteriófago se transfiere de un sistema bacteriano a otro. Si una preparación de fago crece, por ejemplo, en *E. coli* K, y luego se intenta que

crezca en *E. coli* C, es posible que la eficiencia de plaqueo pase de ser 1 a 10^{-4} . El fago ha sido restringido en el segundo hospedador (figura 3.15). La restricción no es absoluta y los fagos que hayan sobrevivido en esta segunda cepa de bacterias plaquearán de nuevo con una eficiencia de 1 en *E. coli* C, siempre que no pasen entremedias por *E. coli* K. Podemos apreciar, pues, que la eficiencia con que el fago plaquea en una cepa determinada depende de la cepa en la que se haya propagado con anterioridad. Esto es debido a que los fagos que no han sido restringidos han sido modificados (metilados) por la cepa hospedadora y son, desde entonces, inmunes al sistema de restricción de esa cepa. Marcando el ADN del fago con ^{32}P , se observó que éste se degradaba al entrar en las cepas restrictivas y la enzima responsable de la degradación se denominó endonucleasa de restricción.

Hay varios tipos de enzimas de restricción que se denominaron, por la sucesión en el tiempo de su descubrimiento y por sus características, como sistemas de tipo I, II y III. Las enzimas de restricción de *E. coli* K fueron las primeras en aislarse (tipo I), pero no son las más sencillas. Con el tiempo han aparecido nuevos sistemas y algunas de las enzimas originalmente clasificadas en un sistema concreto no cumplen los requisitos de la mayoría, por lo que deberían considerarse grupo aparte. De todos los tipos conocidos, las de tipo II son las verdaderas tijeras de la ingeniería genética.

La característica básica del sistema de modificación es que cada cepa bacteriana posee una metilasa con la misma especificidad de secuencia que una endonucleasa de restricción. La metilasa añade grupos metilo a adeninas o citosinas dentro de la secuencia de reconocimiento y unión de la endonucleasa de restricción. La metilación convierte la secuencia diana en resistente a la acción de la endonucleasa y, por tanto, protege el ADN de la degradación. Todas las enzimas de restricción reconocen regiones cortas en el ADN y cortan, bien en el mismo lugar donde reconocen, o a una cierta distancia, dependiendo del tipo de enzima.

3.2.1. Enzimas de restricción de tipo II

Las enzimas de tipo II son las mejor estudiadas. Reconocen secuencias de ADN específicas, generalmente de 4-8 pb y cortan ambas hebras de ADN en un lugar concreto coincidente, en la mayoría de los casos, con el lugar de reconocimiento. La restrictasa y la metilasa son enzimas independientes que actúan sobre la misma secuencia de ADN. La primera enzima de restricción de este tipo que se caracterizó fue la Hind II. H. O. Smith, en 1970, aisló la enzima y estudió los extremos protuberantes que producía en el ADN del fago T7. Identificó la diana de esta enzima como la secuencia G T P_i P_u A C (donde P_i es cualquier pirimidina y P_u cualquier purina). La elección de T7 fue afortunada porque la bacteria tiene otra enzima de restricción bastante más abundante que Hind II y es la que hoy denominamos Hind III, pero que no corta T7 por carecer el ADN del bacteriófago de

la secuencia diana (AAGCTT), de ahí que cualquier contaminación de Hind III en la preparación de Hind II no hubiera afectado al resultado. La restrictasa consta de una sola cadena polipeptídica, actúa en forma de homodímero y requiere magnesio; la metilasa es un monómero que actúa independientemente de la restrictasa y que requiere como cofactor S-adenosil-metionina. En el sistema Eco RI, también la restrictasa es un homodímero, y la metilasa, un monómero independiente.

Las secuencias que reconocen las enzimas de restricción de tipo II son capicúa en el sentido de que la lectura 5' → 3' de una de las hebras es idéntica a la de la banda complementaria, también leída en la dirección 5' → 3'. Forman horquillas monocatenarias. Tienen un tamaño de 4 o 6 pares de bases, y pueden o no estar metiladas en una base de cada una de las hebras (figura 3.16). Cuando la secuencia de reconocimiento o diana está metilada solamente en una hebra del ADN, no es sustrato de la enzima de restricción, pero puede ser reconocida por la metilasa que actúa sobre la hebra aún no metilada. Esto es lo que ocurre al final de la replicación del ADN, donde la hebra recién sintetizada no está aún metilada. Si ninguna de las dos hebras del ADN está metilada, el ADN es restringido, es decir, fragmentado. Si el ADN estuviera metilado en sus dos hebras, no sería sustrato ni de la endonucleasa ni de la metilasa.

Hay enzimas de tipo II que producen cortes quebrados en la secuencia diana originando terminales cohesivos, mientras que en otras el corte es liso, creando terminales planos o romos (figura 3.16).

En cuanto a la nomenclatura, la primera letra procede del género del microorganismo del que se aisló originalmente, y las dos siguientes, de la especie. Por ejemplo, Hind II lleva sus tres primeras letras de *Haemophilus influenza*; la cuarta letra define la cepa (d en este caso). Cuando hay más de un sistema en el mismo organismo se definen por números romanos: I, II, III, etc. Así BamHI procede originalmente de *Bacillus amyloquefaciens* H; EcoRI de *Escherichia coli* RY13; Sau3AI de *Staphylococcus aureus* 3A, y PstI de *Providencia stuartii*.

La frecuencia de reconocimiento puede calcularse matemáticamente. Un tetranucleótido cual-

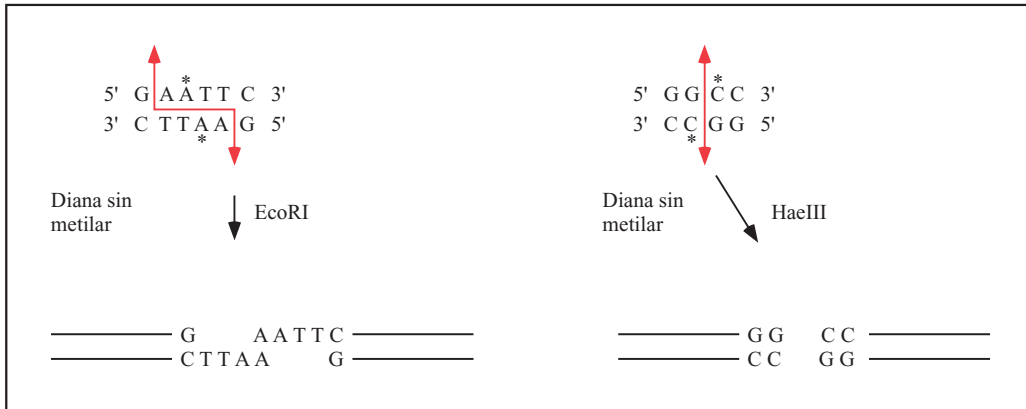


Figura 3.16. Modo de acción de las enzimas de restricción de tipo II. EcoRI reconoce un exanucleótido y origina extremos protuberantes al cortar; HaeIII reconoce un tetranucleótido y origina extremos romos. Los asteriscos marcan las bases que pueden estar metiladas, en cuyo caso la enzima no corta.

quiera ocurre con una frecuencia de $4^4 = 256$ nucleótidos (nuc.), y un exanucleótido (GGATCC), una vez cada $4^6 = 4.096$ nuc. (permutaciones de 4 elementos tomados de 4 en 4 o de 6 en 6 al azar y con repetición). En la práctica hay muchas desviaciones, en parte debidas a que el contenido real de GC en la mayoría de los organismos no es del 50%. Por ejemplo, en el fago λ , con un tamaño de genoma de 49 kb, debería haber unos doce lugares para cada enzima con una secuencia-diana de seis nucleótidos, y, sin embargo, hay sólo cinco para BamHI, o dos para Sall. Las dianas en el genoma tampoco están uniformemente distribuidas, ya que, si lo estuvieran, al tratar un genoma con enzimas de restricción, nos daría fragmentos más o menos del mismo tamaño, y esto no ocurre así. En vertebrados, por ejemplo, la metilación ocurre en el par CpG y, por tanto, la mayoría de las dianas que contengan este par de bases estarán metiladas y serán, al menos parcialmente, resistentes a la restricción. Por ejemplo, en mamíferos, las dianas XmaI (C/GGCG, la barra inclinada representa el lugar de corte) suelen originar fragmentos de 65 kb, y las SstII (CCGC/GG) o PvuI (CGAT/CG), de 150 kb. Estas enzimas son ideales para aislar fragmentos de ADN de gran tamaño.

Una alternativa para conseguir fragmentos grandes es utilizar enzimas que reconozcan oc-

tanucleótidos en vez de exanucleótidos, ya que éstos serán más infrecuentes. Ejemplos de estas enzimas son la NotI que reconoce GC/GGCCGC, y la Sfi que corta en GGCCNNNN/NGGCC.

Las condiciones de sal y magnesio en las que se producen las digestiones pueden modificar la especificidad de las enzimas. Por ejemplo, EcoRI reconoce normalmente GAATTC, pero bajando la concentración de sal y sustituyendo Mg^{+2} por manganeso, reconoce AATT, y se denomina actividad estrella RI*.

Cuando dos enzimas de restricción distintas reconocen, total o parcialmente, la misma secuencia y cortan más o menos en la misma diana, se denominan isoesquizómeras.

En la figura 3.17 vemos los casos de BamHI isoesquizómera de MboI o Sau3AI (isoesquizómeras también entre ellas) o el de Sall y XhoI, donde, una vez reconstruida la diana con ambos cortes, no puede ser cortada ni con una ni con la otra enzima porque se destruyen los exanucleótidos de reconocimiento de ambas.

Quimeras de endonucleasas de restricción y elementos de reconocimiento del ADN

Hay un grupo de enzimas de restricción de tipo II que cortan en un lugar distinto al de reconoci-

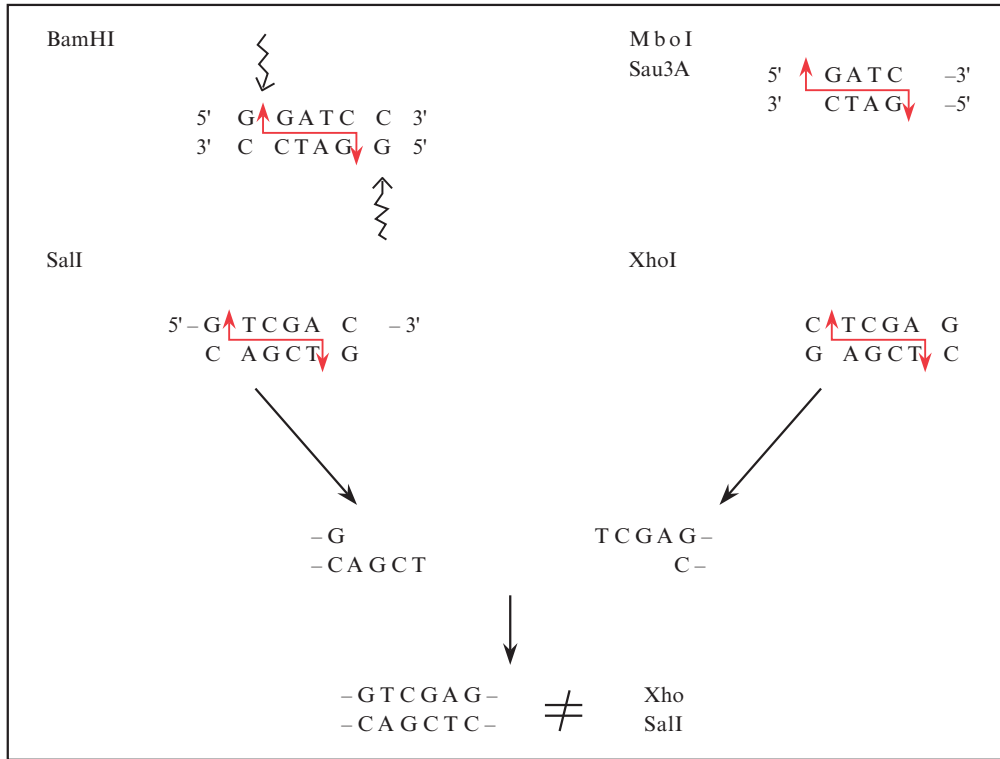


Figura 3.17. Modo de acción de las enzimas isoesquizómeras. BamHI es isoesquizómera de MboI y Sau3A; SalI lo es de XhoI; se representa cómo una diana híbrida de las dos pierde la capacidad de ser reconocida por ambas.

miento, son las de tipo IIS y la enzima FokI es un ejemplo de ellas. FokI se aisló originalmente de *Flavobacterium okeanoikoites*, una bacteria que infecta peces de agua dulce, reconoce «GGATG» y corta la hebra superior e inferior del ADN a 9 y 13 nucleótidos respectivamente del lugar de reconocimiento (figura 3.18). La endonucleasa consiste en una cadena polipeptídica con dos regiones bien diferenciadas: la región responsable del reconocimiento de la diana y la región de corte. FokI se une al ADN como monómero y, sin embargo, su actividad catalítica depende de su dimerización de modo que cada monómero corta una hebra del ADN.

Se pueden construir quimeras entre la enzima FokI y elementos de reconocimiento de secuencias específicas en el ADN para producir cortes específicos en los genomas de organismos superiores. La región de reconocimiento de la enzima

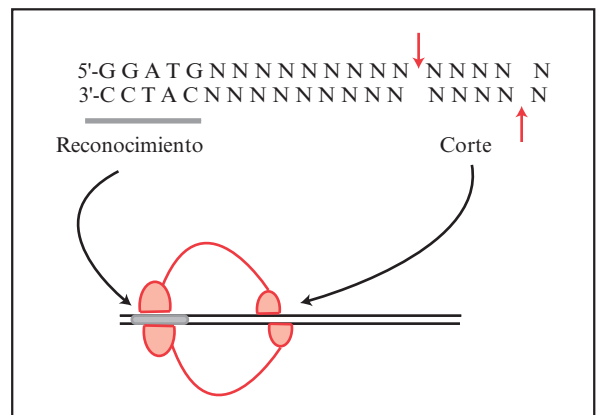


Figura 3.18. Modo de corte de la enzima FokI donde N representa cualquier nucleótido, y las flechas, los lugares de corte. Abajo representación esquemática que muestra los lugares de reconocimiento y corte.

se sustituye por la región de reconocimiento y unión al ADN de algunos factores de transcripción como los dedos de zinc. La quimera constituye las **nucleasas de dedo de zinc** (ZFN, del inglés: *zinc-finger nucleases*). Estas quimeras son capaces de producir un corte en las dos hebras de un ADN de cualquier procedencia en el lugar preciso estipulado por los dedos de zinc. Cada dedo consta de 23 aminoácidos y reconoce 3 pb concretos, hay 7-8 aminoácidos entre dos dedos y los factores de transcripción naturales suelen tener unos 6 dedos consecutivos, lo que les confiere una especificidad de secuencia de 18 pb. La combinación de distintos dedos con un monómero de FokI permite reconocer una secuencia concreta en cualquier ADN y, si tenemos dos elementos

quiméricos convenientemente próximos para que los monómeros de FokI dimericen, provocaremos un corte en las dos hebras del ADN allí donde determine la secuencia reconocida por los dedos (figura 3.19).

De modo similar se pueden usar las **quimeras TALEN** (del inglés: *transcription activator-like effector nucleases*) que combinan la endonucleasa FokI con elementos pertenecientes a unos factores de transcripción de unas bacterias que infectan plantas, donde cada elemento (TAL) reconoce un par de bases apareadas (figura 3.19). Cada módulo TAL consta de 34 aminoácidos comunes y 2 variables en las posiciones 12 y 13. Son estos dos residuos variables los que dan la especificidad de unión preferente a un par de bases apareadas.

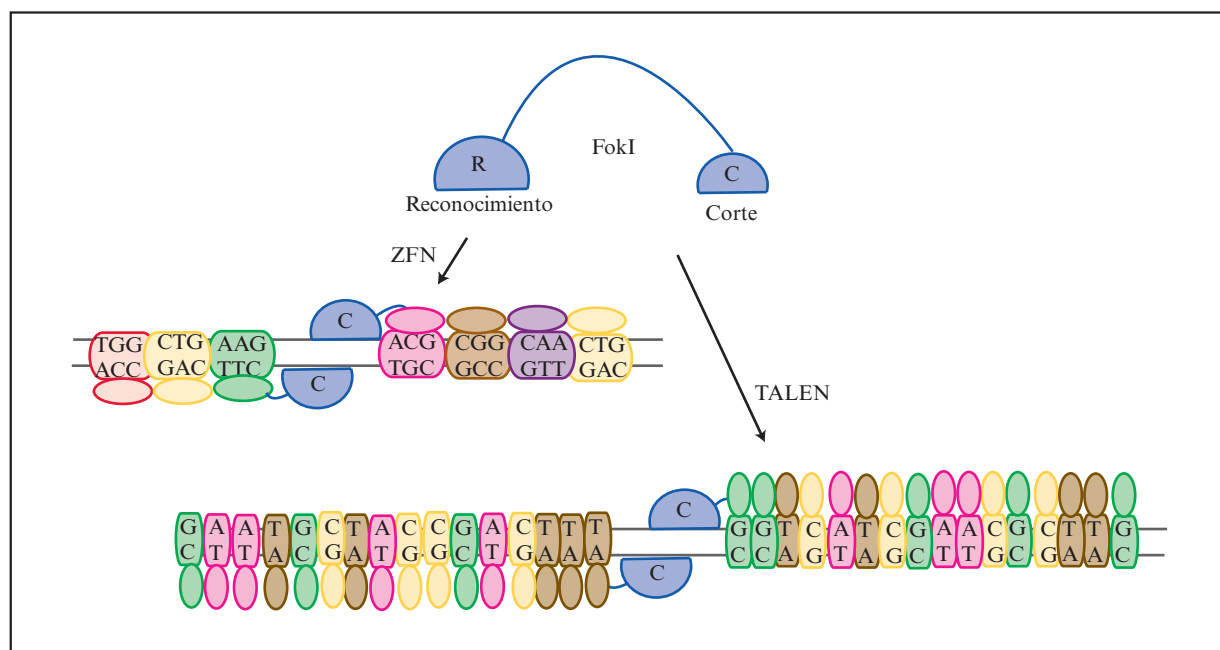


Figura 3.19. Enzimas quiméricas entre FokI y dedos de zinc para constituir las ZFN y FokI y elementos TAL para constituir las TALEN.

Los problemas y limitaciones de las ZFN vienen de que no hay suficientes dedos de zinc naturales para reconocer los 64 tripletes del ADN. Ambas enzimas quiméricas, ZFN y TALEN, os-

tentan una precisión y una eficiencia variables dependiendo de la facilidad para dimerizar, actuaciones fuera de la diana y toxicidad, funcionando ligeramente mejor en todos los casos las TALEN.

Estas quimeras se pueden introducir en las células a nivel de ADN o proteína, al menos las TALEN entran directamente sin mucho problema.

3.2.2. Enzimas de restricción de otros tipos

En un principio las enzimas de restricción se clasificaron en tres únicos tipos: I, II y III, pero realmente esta división no refleja la variabilidad existente. Hemos visto las de tipo II con la variante IIS y ahora veremos algunas otras.

Las enzimas de restricción de tipo I

Las enzimas de restricción de tipo I tienen tres subunidades que llevan a cabo el reconocimiento, la restricción y la metilación. Son las primeras que se descubrieron, pero son bastante menos populares que las de tipo II. La subunidad R es responsable de la restricción; la M de la modificación, y la S del reconocimiento de la secuencia en el ADN. Cada subunidad está codificada por un gen distinto (*r*, *m* y *s*). Los mutantes en *m* han de ser necesariamente mutantes también en *r* para salvar el propio ADN, ya que éste sería destruido por las propias enzimas si fuese *m*⁻, *r*⁺. Entre los sistemas de tipo I mejor estudiados están los EcoB y EcoK. Sus enzimas se unen a la secuencia de reconocimiento en el ADN e inmediatamente interaccionan con ATP. Si la enzima se ha unido a un lugar metilado, el ATP provee la energía necesaria para soltarse del ADN; si no lo está, también utiliza ATP para romper la molécula a una cierta distancia no muy específica del lugar de reconocimiento. Las enzimas EcoB y EcoK son en realidad dos alelos del mismo gen, sus secuencias de reconocimiento son distintas, pero, en diploides, la subunidad S de una cepa puede dirigir las actividades de las subunidades R y M de la otra. Las dianas de reconocimiento para estos dos sistemas constan de dos partes separadas por un octanucleótido no específico (figura 3.20).

La secuencia de ruptura está a unos 1.000 nucleótidos de distancia y aunque no es totalmente al azar, ya que parece haber preferencias, tampoco

TGA* NNNNNNNN TGCT ACT NNNNNNNN *ACGA
--

Figura 3.20. Secuencia de reconocimiento, pero no de corte, para la enzima de tipo I, EcoB. El asterisco señala las bases que pueden estar metiladas. N representa un nucleótido no específico.

es muy específica. El proceso de ruptura en sí se realiza en dos pasos: en el primero se rompe una de las hebras y luego, en un lugar cercano, se rompe la complementaria. Muchas veces la ruptura va acompañada de una cierta hidrólisis en ambas hebras y además se requiere mucho ATP. La enzima permanece fija en el lugar de reconocimiento y desde esta posición tira del ADN, originando una especie de lazo, hasta llegar al lugar de corte.

Las enzimas de restricción de tipo III

Son relativamente raras y no se utilizan en manipulaciones genéticas. Cada enzima tiene dos tipos de subunidades, la R responsable de la acción endonucleásica, y la subunidad MS responsable de metilación y reconocimiento.

La restricción y la modificación se expresan simultáneamente. La enzima se une primeramente al ADN utilizando ATP, y a partir de aquí el requerimiento de ATP es opcional y depende de cada enzima en concreto. A continuación las funciones de metilación y restricción compiten por llevar a cabo sus respectivas reacciones.

Otros sistemas de restricción

Un sistema que se sale de las clasificaciones tradicionales es el Mcr (del inglés: *modified cytosine restriction*) descubierto por la dificultad encontrada en el clonaje de ADN procariótico y eucariótico en ciertas cepas bacterianas comunes. El ADN se degradaba en el huésped siempre que las citosinas estuvieran metiladas. Dentro del tipo Mcr hay, a su vez, al menos dos sistemas: el McrA codificado por un elemento semejante a un profago, y el McrBC codificado por los genes *mcrB* y

mcrC de bacterias. La secuencia de reconocimiento de este último sistema es: A o G, *C (metilada), de 40 a 80 nucleótidos variables, A o G, *C (metilada), que podemos resumir como A/G*C(N₄₀₋₈₀) A/G*C. El corte tiene lugar en múltiples sitios entre las citosinas metiladas y en las dos hebras. El sistema McrBC había sido descrito con el nombre RglB ya en 1952 por Luria y Human, para explicar un fenómeno observado en la serie de los fagos Tpar, es decir, T₂, T₄ y T₆. El ADN de estos fagos tiene la mayoría de las citosinas 5-hidroximetiladas y glucosiladas. Se vio que cuando una mutación, bien en el fago o en el huésped, impedía la glucosilación, quedaba el ADN solo hidroximetilado y era degradado en determinados huéspedes.

Es muy posible que el origen de estas metilaciones en ADN procarióticos esté relacionado con la inmunidad que ello les confiere frente a los sistemas tradicionales de tipos I, II y III.

Los fagos de la serie Tpar no son los únicos que han desarrollado mecanismos antirrestricción y éstos pueden ser tan variados como eliminar de su genoma las dianas para sistemas concretos o codificar proteínas que destruyen cofactores de las enzimas de restricción.

3.3. GENOMAS TRATADOS CON ENZIMAS DE RESTRICCIÓN

El tratamiento *in vitro* de un genoma pequeño (un bacteriófago, por ejemplo) con enzimas de restricción originará una serie de fragmentos discretos de tamaños variables analizables por técnicas de electroforesis. Estos fragmentos lineales de ADN poseerán carga negativa y migrarán al ánodo (+) a una velocidad inversamente proporcional a su tamaño; dentro del rango de resolución del gel empleado. Al subir en la escala evolutiva y analizar, con enzimas de restricción, genomas más complejos (bacterias, levaduras, insectos, mamíferos), los tamaños de los distintos fragmentos originados se solapan y en un gel de agarosa aparecerá como una mancha que abarca todo el surco del pocillo de arriba abajo (figura 3.21).

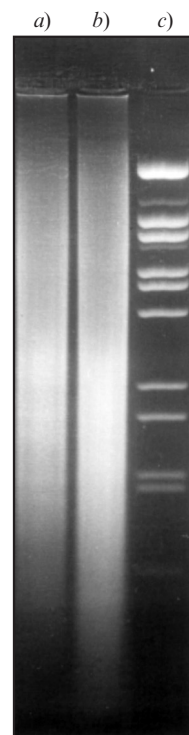


Figura 3.21. Gel de agarosa y tinción con bromuro de etidio para visualizar las bandas al incidir sobre el gel luz ultravioleta. Se muestra la migración de los múltiples fragmentos originados por la digestión total de ADN humano con las enzimas de restricción a) PstI y b) MspI. c) Patrón observado al digerir el fago λ con la enzima BstE2; los fragmentos de esta última digestión sirven como marcadores de peso molecular conocido.

Todo el material genético presente en el gel puede ser transferido a filtros de nailon según la técnica Southern (figura 3.22); así bandas concretas pueden ser identificadas siempre que se disponga de una sonda complementaria marcada. La técnica Southern consiste en desnaturalizar el ADN dentro del propio gel por tratamiento con hidróxido sódico y transferirlo bien por capilaridad o por un impulso eléctrico (electrotransferencia) a una membrana de nailon. El ADN se fija e inmoviliza en el filtro por tratamiento con radiaciones ultravioleta y se hibrida con la sonda marcada cuya región complementaria se desea identificar. Estos filtros, una vez lavados, para eliminar uniones inespecífi-

fragmentos marcados y de tamaños crecientes con respecto a uno de los extremos (figura 3.23a). Puede haber problemas en caso de existir varios

fragmentos de tamaños similares, con lo que habría que elegir otras enzimas de restricción más adecuadas.

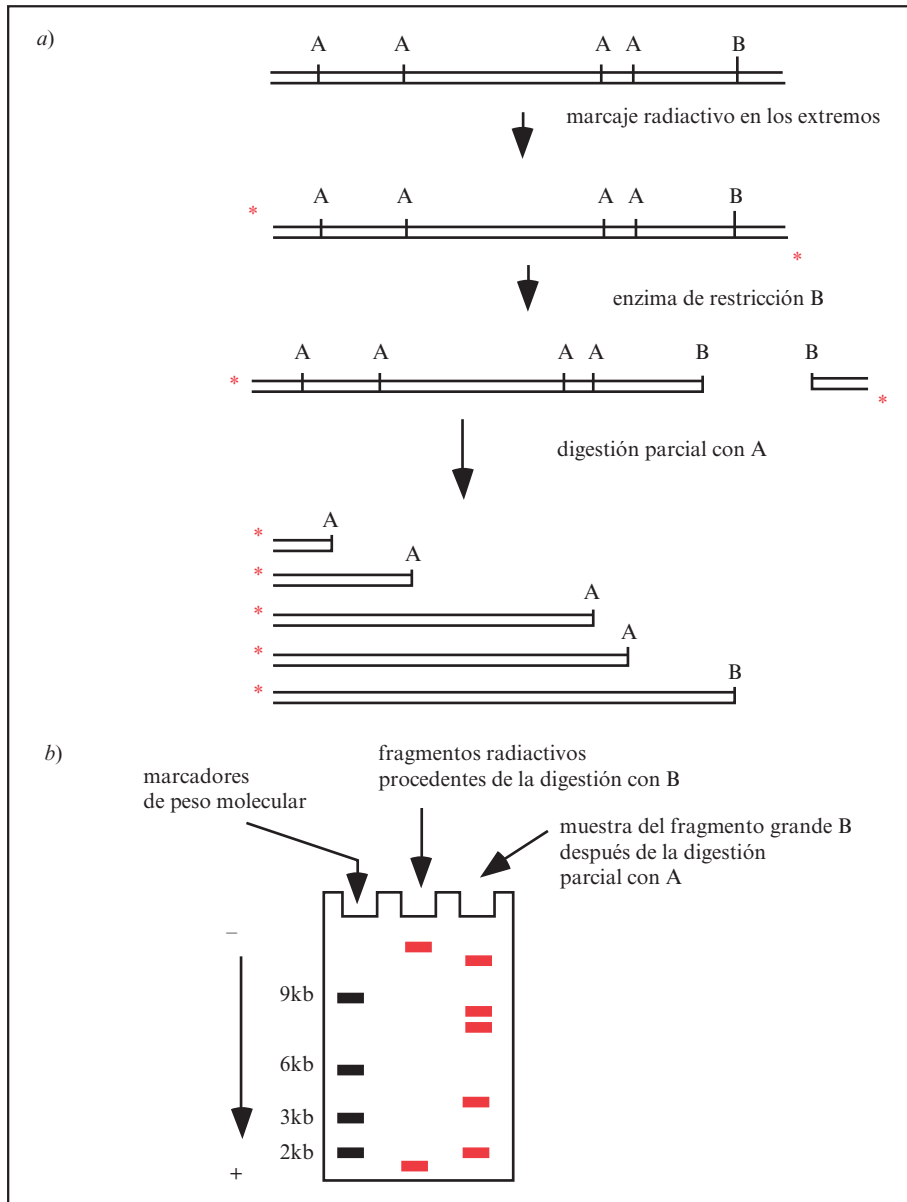


Figura 3.23. a) Estrategia para elaborar mapas genómicos de fragmentos cortos, y b) separación de fragmentos de restricción por electroforesis en geles de agarosa. Determinación del tamaño de muestras desconocidas por autorradiografía y comparación con la movilidad de fragmentos patrón (marcadores no radiactivos y teñidos con bromuro de etidio).

Los fragmentos se pueden separar, según su tamaño, por electroforesis. En esta técnica, como se ha mencionado anteriormente, el ADN segmentado se coloca en un extremo de un gel de agarosa o poliacrilamida. Al pasar una corriente eléctrica a través del gel, cada molécula se desplaza hacia el polo positivo a una velocidad dependiente del logaritmo de su peso molecular. Este movimiento, interrumpido al cabo de un cierto tiempo, produce una serie de bandas, donde cada una es un fragmento de un tamaño concreto.

El tamaño real por aproximación se puede obtener calibrando el gel con una muestra en la que varias moléculas de tamaños conocidos (marca-

dores) se corren en paralelo en un pocillo contiguo a la muestra de tamaño/s desconocido/s (figura 3.23b). La separación de los fragmentos en geles de agarosa o acrilamida y la visualización por autorradiografía nos permitirá establecer el tamaño de cada fragmento y su posición relativa al extremo marcado.

Las digestiones dobles con más de una enzima de restricción también son muy valiosas para establecer mapas de restricción y ordenación de fragmentos. En esta técnica se realizan en paralelo digestiones sencillas y dobles (figura 3.24). Los productos de las digestiones se analizan por electroforesis.

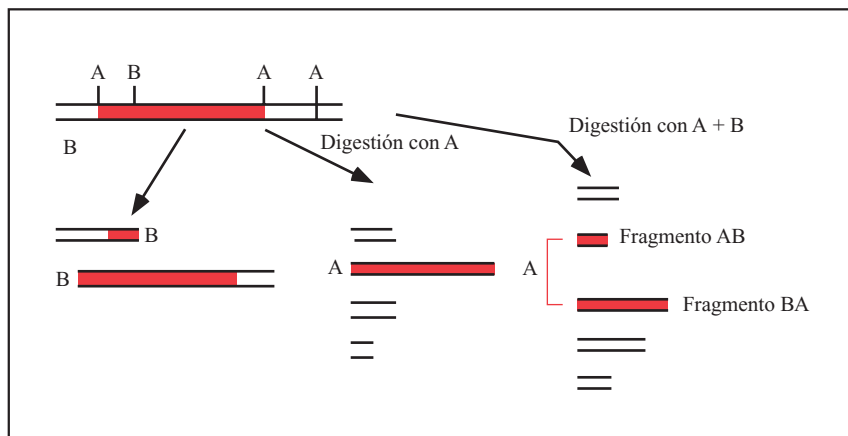


Figura 3.24. Digestión con las enzimas A y B por separado y conjuntamente. La digestión doble pone de manifiesto que la diana B se encuentra dentro del fragmento de mayor tamaño de A y a una distancia de los extremos determinada por la longitud de los fragmentos AB y BA.

Los mapas de restricción así elaborados no muestran genes, sino secuencias, es decir, no corresponden a segmentos de función y no detectan una pequeña mutación o alteración que dentro del genoma afecte a un solo nucleótido o a un número reducido de ellos. Sólo en los casos en los que el nucleótido afectado forme parte de una diana de restricción, o que la alteración sea de grandes proporciones, se verá afectado el patrón de bandas observado en los geles y autorradiografías correspondientes.

En la actualidad quizá la manera más rápida y eficaz de elaborar un mapa de restricción de un fragmento concreto sea su secuenciación directa y análisis por ordenador de la localización de dienas para un gran número de enzimas de restricción.

El análisis de patrones de restricción puede detectar anomalías como amplificaciones génicas (aumento de la intensidad de una banda concreta), translocaciones, deleciones grandes, inserciones y todas aquellas que de algún modo alteren la

movilidad de algún fragmento ocasionado por una enzima de restricción.

3.3.1. Polimorfismos de fragmentos de restricción

Existen muchas pequeñas variaciones entre los genomas de dos individuos que son detectables a nivel de enzimas de restricción, pero que no causan ninguna alteración génica visible o funcional. Alteraciones de este tipo son comunes en intrones con ganancia o pérdida de material genético y, por tanto, de dianas de restricción, sin que ello produzca cambio alguno en ARNm o proteína. Estas

variaciones son equivalentes a los alelos genéticos y se denominan polimorfismos del tamaño de fragmentos de restricción (PTF). El polimorfismo de restricción es mucho más amplio que el alélico, ya que afecta a una gran parte del ADN cuya función no está totalmente caracterizada. Cuando los dos alelos de un individuo difieren en el tamaño de un intrón para un gen dado, el tratamiento con enzimas de restricción de su ADN genómico nos revelará esta diferencia, apareciendo dos fragmentos de distinto tamaño para una sonda que sólo originará un fragmento en individuos con alelos idénticos. Este PTF nos permite diferenciar el alelo de origen paterno del materno, y el patrón de digestión se denomina informativo (figura 3.25).

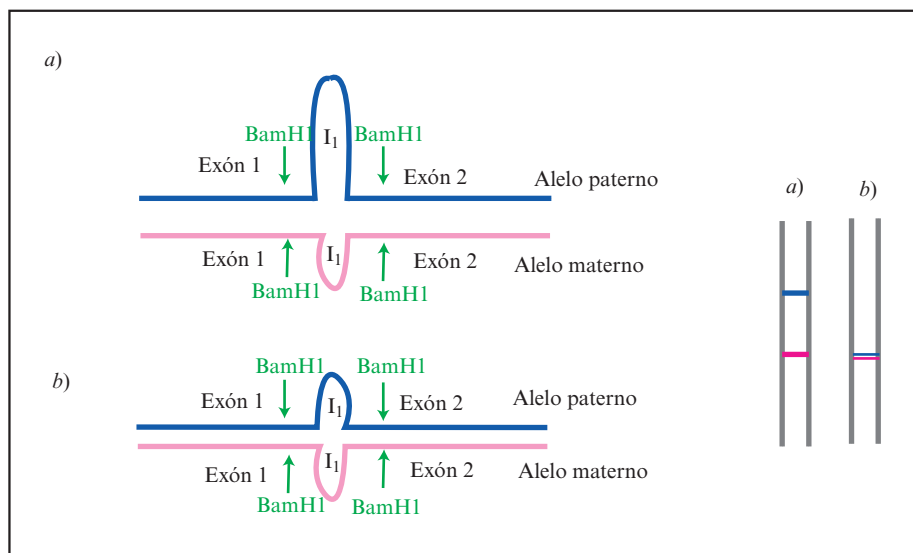


Figura 3.25. PTF debido a la variación en el tamaño de un intrón (I_1) al tratar el genoma de un individuo con la enzima de restricción BamHI que corta en los exones E_1 y E_2 del gen. *a)* Los alelos paterno y materno de un individuo difieren en el tamaño del intrón dando lugar a un patrón de restricción informativo para la enzima BamHI. *b)* Análisis de restricción en otro individuo en el que no hay variación en el tamaño del intrón originando un patrón de restricción no informativo para la misma enzima de restricción.

En ocasiones muy contadas una mutación, que origina una patología concreta, elimina o crea una nueva diana de restricción para alguna enzima conocida. Un caso de este tipo es el de la anemia falciforme, donde la enfermedad está asociada a una

sustitución de una adenina por una timina ($A \rightarrow T$) en el codón número 6 del gen de la β -hemoglobina. El cambio elimina una diana para la enzima MstII permitiendo el diagnóstico de la enfermedad al aparecer un fragmento de 1,4 kilobases (kb) en vez

del normal de 1,2 kb (figura 3.26). En los aminoácidos correspondientes, el cambio sustituye un ácido glutámico polar por una valina en la superficie de la molécula, provocando una mayor adhesión

intermolecular entre deoxi-hemoglobinas que, como consecuencia, se agregan distorsionando el glóbulo rojo. La vida media de estos glóbulos rojos es menor y pueden bloquear capilares.

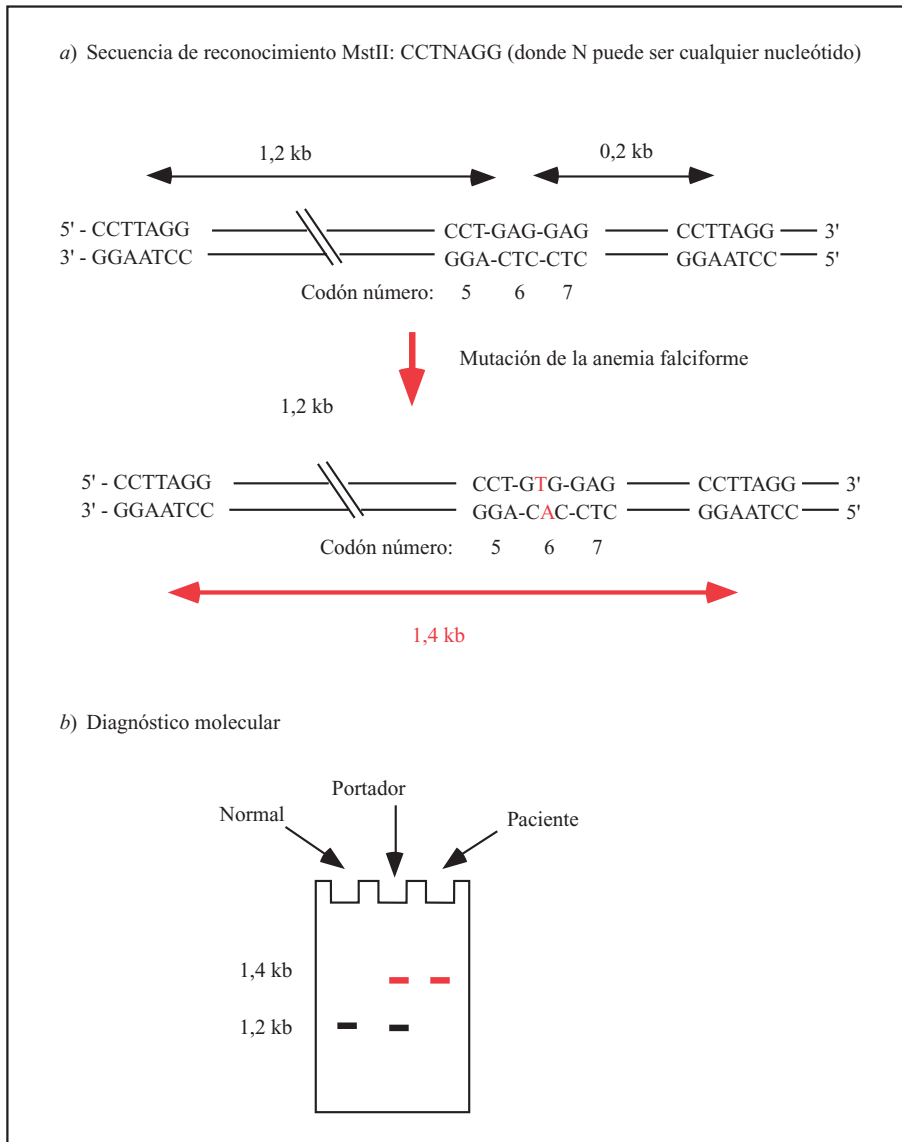


Figura 3.26. a) Mutación en la secuencia de reconocimiento de MstII en el codón número 6 del gen de la β -hemoglobina que convierte un fragmento de 1,2 kb en otro de 1,4 kb. b) Diagnóstico molecular basado en el aislamiento de ADN de posibles portadores y pacientes; tratamiento con la enzima MstII e hibridación con una sonda comprendida en el fragmento 1,2 de un individuo normal.

3.3.2. La huella genética individual

Existen en el genoma humano regiones hipervariables constituidas por secuencias de 10-40 nucleótidos internamente repetidas que pueden encontrarse una vez por genoma haploide (unilocus) o muchas veces (multiloci). La hipervariabilidad de estas secuencias puede alcanzar niveles de frecuencias mutacionales de 8×10^{-3} y es debida en parte al alto porcentaje de recombinaciones acompañadas de sobrecruzamientos desiguales que origina deleciones o ampliaciones en el número de repeticiones internas (figura 3.27). Sin embargo, en algunas regiones formadas por tetranu-

cleótidos repetidos CA/TG no parece que los sobrecruzamientos desiguales hayan sido la causa de nuevos tamaños en la unidad repetida, ya que no se observan cambios en las secuencias colindantes con la propia unidad y la diferencia entre tamaños puede ser de sólo cuatro nucleótidos. Parece más probable que procesos de apareamiento incorrecto en los que se queda fuera una unidad repetida puedan acarrear en la replicación situaciones en las que la unidad desapareada sea ignorada por la ADN polimerasa y quede excluida de la hebra copiada (figura 3.27). El mecanismo se denomina «resbalón de la polimerasa» (*slippage* en inglés). En algunos casos el resbalón provoca

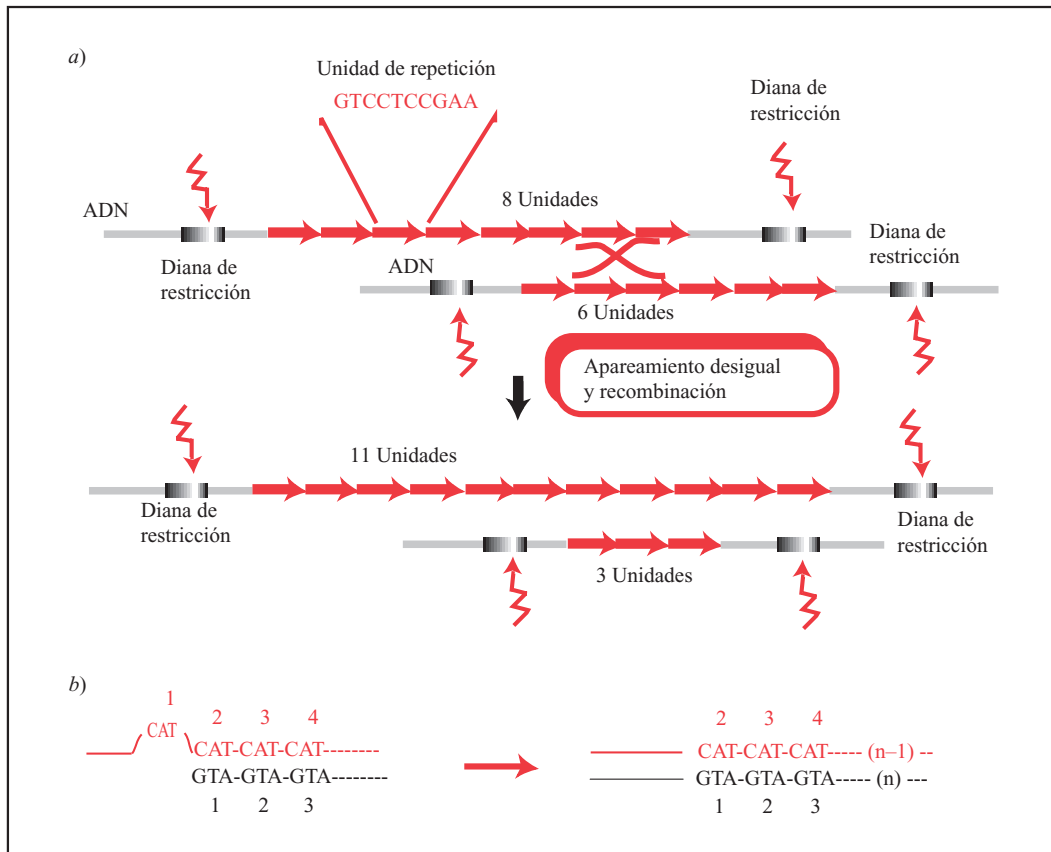


Figura 3.27. a) Origen de la variabilidad en el tamaño de las regiones hipervariables al tratar el genoma con una enzima de restricción que corte fuera de la región repetida, por apareamiento desigual y recombinación. b) Deleción de una unidad corta repetida en tándem por apareamiento incorrecto y resbalón de la ADN polimerasa.

inserciones de unidades; en otros, deleciones, y a veces, la incorporación de la unidad repetida de la hebra complementaria. Dos individuos sin parentesco lo más probable es que tengan dos fragmentos de tamaño dispar al cortar una región hipervariable con una enzima de restricción que no reconozca dianas dentro de la unidad repetida. Esta propiedad puede ser aprovechada para determinar identidad o parentesco entre muestras de ADN y se ha utilizado en medicina forense para identificar sospechosos de asesinatos o violaciones, a partir de muestras de ADN de manchas de sangre, pelos o restos de semen, por amplificación de una o varias regiones hipervariables.

Cuando estas secuencias internamente repetidas están dispersas y se presentan muchas veces en el genoma, aparecen como una banda satélite en determinados gradientes de densidad; de ahí que se las denomine frecuentemente minisatélites. Marcadores genéticos consistentes en la unidad central repetida pueden ser utilizados como sondas radiactivas para detectar un gran número de variantes polimórficas, exclusivas y únicas de cada individuo, pero que a la vez presentan una gran estabilidad somática entre distintos tejidos, incluyendo los germinales. El patrón de restricción observado en un individuo al ser tratado su ADN con una enzima de restricción que no corte dentro de la unidad repetida, y posteriormente transferido el ADN a nailon e hibridado con una sonda radiactiva de un minisatélite, se denomina *la huella genética* o, coloquialmente, el código de barras individual. La probabilidad estimada experimentalmente de que una sola banda de las múltiples que se observan en un individuo esté presente en otro individuo tomado al azar es aproximadamente de 0,2; de esto se deduce que es bajísima la probabilidad (5×10^{-19}) de encontrar otro individuo sin parentesco directo con el primero y con una huella genética idéntica.

La huella genética, que es completamente específica del individuo y sólo compartida por los gemelos univitelinos, tiene aplicación directa en problemas de identificación individual, familiar o de paternidad.

En la figura 3.28a se representa el patrón de restricción de varios individuos no relacionados,

de dos hermanas (H) y de dos muestras del mismo individuo (Daniel). Cuando el ADN de estos individuos se hibrida con el ADN marcado de un minisatélite (33,6), se observa cómo los individuos sin parentesco no poseen casi ninguna banda de restricción en común; las dos hermanas comparten algunas pero no todas, mientras que las dos muestras del mismo individuo son prácticamente idénticas. En criminología, la huella genética de ADN extraída de manchas de sangre, semen, cabellos, etcétera, puede ser comparada con la huella genética de uno o más sospechosos.

Uno de los primeros casos en los que se aplicó esta técnica para un caso de inmigración fue la identificación del niño de Ghana (figura 3.28b).

Un niño nacido en Inglaterra, de padres ghaneses, emigró a Ghana para reunirse con su padre y más tarde volvió solo a Inglaterra para reunirse con su madre y hermanos. Sin embargo, había indicios para sospechar que había ocurrido una sustitución por otro niño no relacionado o por el niño de una hermana de la madre; la madre tiene varias hermanas y todas ellas viven en Ghana. Como consecuencia de todo esto, al niño no se le concedió, de entrada, el permiso de residencia en el Reino Unido hasta que una investigación aclarase el parentesco. Para complicar más el caso, ni el marido ni ninguna de las hermanas de la madre (todos ellos residentes en Ghana) se prestaron para el análisis y, para mayor incertidumbre, la madre no estaba muy segura de si su marido había sido el padre del niño. Se determinó la huella genética de muestras de sangre del niño (X), la madre (M), el hermano (B) y las hermanas (H_1 y H_2). El ADN de todos ellos se digirió con la enzima Hinf I y se hibridó con la sonda de un minisatélite (33,15). Del resultado obtenido podemos deducir primeramente la paternidad de X. Aunque se carece de la muestra paterna, muchas de sus bandas pueden ser deducidas del patrón específico de los hermanos, ausente en la madre. De los 39 fragmentos paternos deducidos, aproximadamente la mitad están presentes en X. Como el patrón de bandas de un individuo sin parentesco es muy distinto (U en la figura 3.28), estos resultados indican claramente que X tiene el mismo padre que sus hermanos. Todas las bandas restantes en X

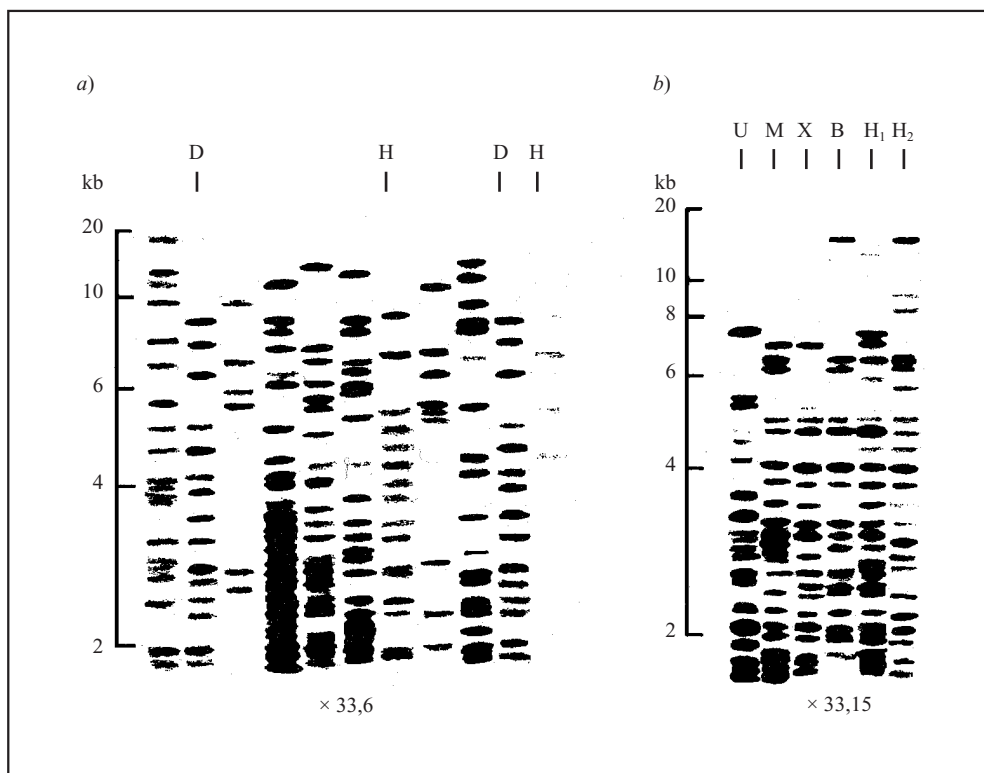


Figura. 3.28. a) Patrón de restricción de ADN obtenido a partir de dos gotas de sangre, tratado con la enzima Hinf I e hibridado con la sonda marcada del minisatélite 33,6. Se analizaron siete individuos no relacionados, dos muestras del mismo individuo (D) y dos hermanas. b) Huella genética de la familia ghanesa: U, sin parentesco; M supuesta madre; X niño a determinar parentesco; B supuesto hermano; H₁ y H₂ supuestas hermanas; $\times 33,6$ y $\times 33,15$ representan la densidad en gradiente de las sondas de minisatélites empleadas en cada caso.

coinciden con las maternas, indicando, asimismo, que M es la verdadera madre.

Como consecuencia de estos resultados, se concedió a X el permiso de residencia indefinida en Gran Bretaña.

Además de las sondas que reconocen muchas regiones hipervariables dispersas por el genoma (multiloci), existen otras que detectan un solo lugar en el genoma, igualmente hipervariable (unilocus). Con las sondas unilocus habrá solamente dos bandas de distinto tamaño por cada individuo (alelo paterno y materno) y se han empleado en casos de paternidad dudosa como el que se muestra en la figura 3.29a). Los alelos paterno y

materno pueden ser identificados en todos los hijos excepto en el primero de ellos, en el que aparece un nuevo alelo como consecuencia de una relación anterior al matrimonio actual de la mujer.

Se puede, asimismo, mezclar varias sondas unilocus (de 5 a 10, por ejemplo) originando un número no muy elevado de bandas específicas que da mayor solidez al diagnóstico. Una mezcla de 6 sondas unilocus se empleó para determinar la paternidad de un niño entre dos posibles padres P1 y P2 (figura 3.29). Uno de los padres muestra total identidad con la mitad de las bandas observadas en el niño.

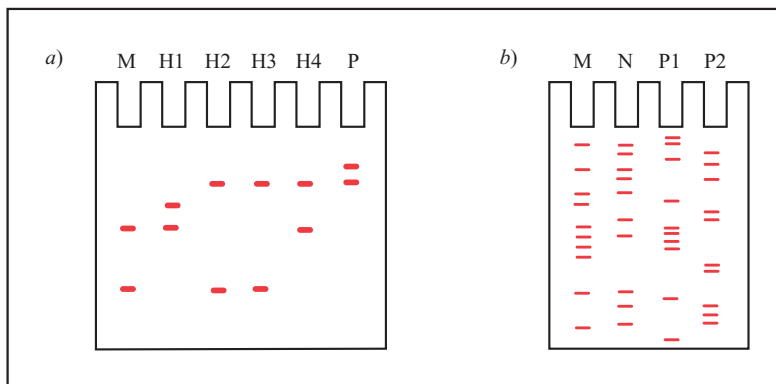


Figura 3.29. a) Análisis de la movilidad de los fragmentos de restricción originados en una región unilocus para identificar los alelos paternos (P) y maternos (M) de los cuatro hijos del matrimonio. Una muestra del ADN de los implicados se trata con una enzima de restricción que no corte dentro de la región hipervariable, se corren las muestras en un gel de agarosa, se desnaturaliza el ADN y se transfiere a nailon. La identificación se hace por hibridación con la sonda unilocus marcada radiactivamente. El primer hijo (H1) es el resultado de una relación anterior al matrimonio actual y no muestra homología con los alelos paternos. b) Determinación de la paternidad entre dos posibles padres, P1 y P2, de un niño (N) con una mezcla de sondas unilocus; el padre P2 queda identificado como el verdadero padre.

3.4. HIBRIDACIÓN DE ÁCIDOS NUCLEICOS EN SOPORTE SÓLIDO ORDENADO EN MATRICES

En esta técnica de hibridación, moléculas cortas (entre 25 y 60 nucleótidos) de ADN de secuencia conocida se unen o sintetizan sobre un soporte de silicio, vidrio u otro material según una disposición ordenada o matriz. La muestra desnaturalizada y marcada con un fluorocromo se vierte sobre la matriz de ADN para que hibride con las secuencias complementarias, a continuación se lava y se detectan las posiciones donde aparece el marcaje. Las matrices de ADN también reciben otros nombres como chips de ADN, microchips, microarrays, biochips o genochips.

Unas 100.000 moléculas de ADN pueden unirse de manera ordenada y robotizada a un portaobjetos de los usados en preparaciones citológicas. Una de las técnicas empleadas consiste en una digitalización de un método fotolitográfico análogo al empleado para imprimir un negativo fotográfico en un papel sensible dejando pasar la luz a través de las partes no opacas de un negativo. Con un gran número de espejos digitalizados

se puede o bien dejar pasar la luz u ocultar las cadenas crecientes de oligonucleótidos que deben incorporar un nucleótido más. La adición de cada nuevo miembro en la cadena requiere cuatro fases de apertura/ocultación (una para cada uno de los cuatro nucleótidos). En una primera fase, sólo se dejarán abiertos (permeables a la luz) los lugares que deben incorporar la guanosina (por ejemplo); en la siguiente fase se ocultarán éstos y se abrirán aquellos lugares donde deba ir la adenosina... y así sucesivamente para la timidina y la citidina. El ADN sintetizado químicamente o unido covalentemente a la matriz es de hebra sencilla y secuencia conocida, representando cada punto de la matriz una secuencia concreta y miles de moléculas idénticas.

La muestra a hibridar con el microchip ha de estar marcada con un fluorocromo. Los fluorocromos son moléculas químicas que absorben luz procedente, por ejemplo, de un láser a una longitud de onda (excitación) y la emiten a otra mayor. Entre los más utilizados para marcar ácidos nucleicos están: el isocianato de fluoresceína (absorbe a 490 nm y emite a 514 nm), Cy3 (absorbe a 550 nm y emite a 570 nm), Cy5 (absorbe a 649 nm

y emite a 670 nm), yoduro de propidio (absorbe a 493 nm y emite a 620 nm), Hoescht 33342 (absorbe a 343 nm y emite a 482 nm), DAPI (absorbe a 350 nm y emite a 470 nm), entre otros. Solo aparecerá la señal fluorescente en aquellos puntos donde se produzca una hibridación positiva entre la muestra marcada y el oligonucleótido complementario, cuya secuencia viene determinada por la posición que ocupa en el microchip (figura 3.30).

La hibridación con microchips nos permitirá detectar el conjunto de genes que se expresan en un momento concreto del desarrollo de un organismo, o en un lugar concreto (como en un borde epitelial), o en una circunstancia especial (como en un cáncer). Con este tipo de hibridaciones podemos identificar, por ejemplo, los mensajeros de una población celular que se expresan en presencia o ausencia de suero, así como los de un tumor frente al tejido normal de un individuo. En el primer caso lo que se debería pegar al vidrio serían moléculas de ARNm correspondientes a todos los genes celulares, o mejor moléculas de ADNcopia obtenidas por transcripción inversa de los ARN mensajeros, definido cada ADNc un punto en un espacio concreto. Por otra parte, se marcan con un fluorocromo verde los ADNc correspondientes a los mensajeros de células crecidas con suero, y con uno rojo los correspondientes a células crecidas sin suero. Cuando se ponen a hibridar cantidades idénticas de los dos últimos ADNc con el microchip, una mancha verde denota una mayor expresión de un gen en células crecidas con suero; una roja indica supremacía del mensajero en las células sin suero, y un color amarillo indica que los niveles de expresión son semejantes (figura 3.30). De manera análoga se pueden extraer los mensajeros de células normales y tumorales de un mismo individuo y pasar las poblaciones a ADNc con la transcriptasa inversa. Los ADNc procedentes de mensajeros de las células cancerosas se marcan con un fluorocromo rojo, y los de las células normales, con uno de color verde. Al combinar cantidades idénticas de las dos poblaciones e hibridar la mezcla con un microchip que contenga ADN de miles de genes que sirvan de gancho y comple-

mento, veremos que rojo indica una elevada expresión del gen en células cancerosas, verde expresión mayoritaria en células normales y amarillo niveles de expresión similares en ambos tipos celulares.

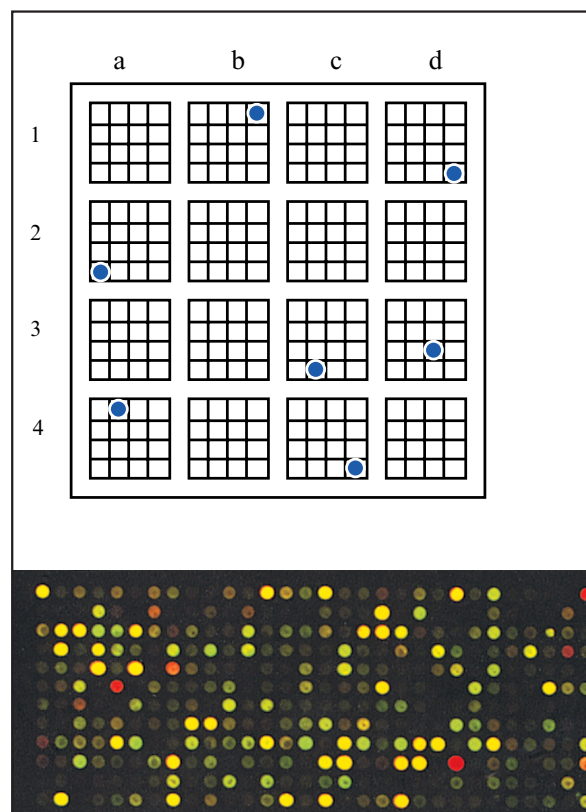


Figura 3.30. En la parte superior, representación esquemática de una hibridación entre una muestra marcada con un fluorocromo y una serie de secuencias conocidas donde cada cuadrante contiene miles de copias de una misma secuencia. La posición del cuadrante en la matriz determina la especificidad de la secuencia. En la parte inferior, un ejemplo concreto: círculos coloreados correspondientes a los lugares de hibridación entre el ADN del microchip y los ADNc con los dos fluorocromos. El color verde indica una mayor hibridación del mensajero ahí localizado en células crecidas con suero. El rojo refleja supremacía del mensajero en las células sin suero y un color amarillo indica que los niveles de expresión son semejantes en ambas situaciones.

3.5. SECUENCIACIÓN DEL ADN: DISTINTOS MÉTODOS Y APROXIMACIONES

La variabilidad de la molécula de ADN no radica ni en la estructura ni en la diversidad de sus componentes, sino en la ordenación y cadencia de los cuatro nucleótidos que forman la molécula, esto es, en su secuencia. Diferencias diminutas en la secuencia de dos moléculas de ADN pueden representar grandes cambios en sus productos. Hoy podemos conocer la secuencia de cualquier ADN aislado de modo rápido y eficaz.

Los dos métodos originales de secuenciación se denominaron químico y enzimático. Ambos comparten algunos principios básicos como, por ejemplo, que han de marcar de algún modo las moléculas a secuenciar, bien radiactivamente o con fluorocromos. Otra de las características compartidas es la de generar fragmentos de ADN de una sola hebra cuyos tamaños sólo se diferencien en una base. Moléculas de ADN con la misma secuencia, y que sólo se diferencien en una base, pueden ser separados en geles desnaturalizantes de acrilamida por electroforesis. Las bandas correspondientes a las moléculas de longitud decreciente forman una escalera que nos permitirá leer en la actualidad, del orden de unos 800 nucleótidos consecutivos a partir de una muestra.

En este apartado repasaremos desde los métodos originales de secuenciación a los más actuales, también llamados de nueva generación y secuenciación masiva.

3.5.1. Método químico

En este método, originalmente descrito por A. Maxam y W. Gilbert en 1977, un fragmento de ADN se marca radiactivamente en sus extremos con γ ^{32}P o γ ^{35}S dATP por acción de la polinucleótido quinasa. Ambos extremos han de ser separados en dos fragmentos distintos (con una enzima de restricción, por ejemplo) y secuenciados independientemente (figura 3.31).

Al comienzo del proceso tendremos un gran número de moléculas idénticas, todas ellas marcadas en el mismo extremo. El paso siguiente consis-

tirá en romper estos fragmentos, no más de una o dos veces por molécula, con reacciones químicas específicas para cada una de las cuatro bases.

Se separan cuatro alícuotas y cada una de ellas será tratada con un compuesto químico que nos modifique sólo una de las cuatro bases, aproximadamente una vez por molécula de ADN. El tratamiento a continuación con piperidina producirá la ruptura de la cadena allí donde la base fue alterada. Existen distintas reacciones de modificación de las bases, y en la técnica que se describe las guaninas se modifican específicamente por tratamiento con dimetil sulfóxido (DMSO) al 1 por 100 en presencia de formato de amonio, pH 3,5, durante 10 minutos. Guaninas y adeninas son atacadas por ácido fórmico al 80 por 100 durante 20 minutos. Las timidinas se degradan al tratarlas con MnO_4K , 0,1 mM, durante 20 minutos (también son atacadas las purinas, pero en menor grado). Las citosinas son eliminadas con hidroxilamina 4 M a pH 6.

Las reacciones se pueden llevar a cabo con el ADN unido a papel (Hybond M&G), lo que facilita el manejo de las muestras (figura 3.31). El tratamiento con piperidina suelta el ADN del papel; el ADN se analiza en geles desnaturalizantes de acrilamida.

Las bandas marcadas se visualizarán por autorradiografía y la secuencia se leerá identificando el tamaño de cada banda en los pocillos correspondientes a cada base (figura 3.32).

3.5.2. Método enzimático

Este método fue desarrollado por F. Sanger en 1981 y ha sido el que más se ha utilizado. Se parte de un ADN, idealmente circular y de banda simple, pero ADN lineal o circular bicatenario desnaturalizado también es buen sustrato (figura 3.33). Necesitamos un cebador que sirva de enganche para el inicio de la síntesis del ADN que vamos a marcar radiactivamente y a secuenciar. El cebador puede ser específico y complementario a una secuencia de localización concreta (o puede representar una pequeña secuencia al azar sin localización definida a priori). La enzima

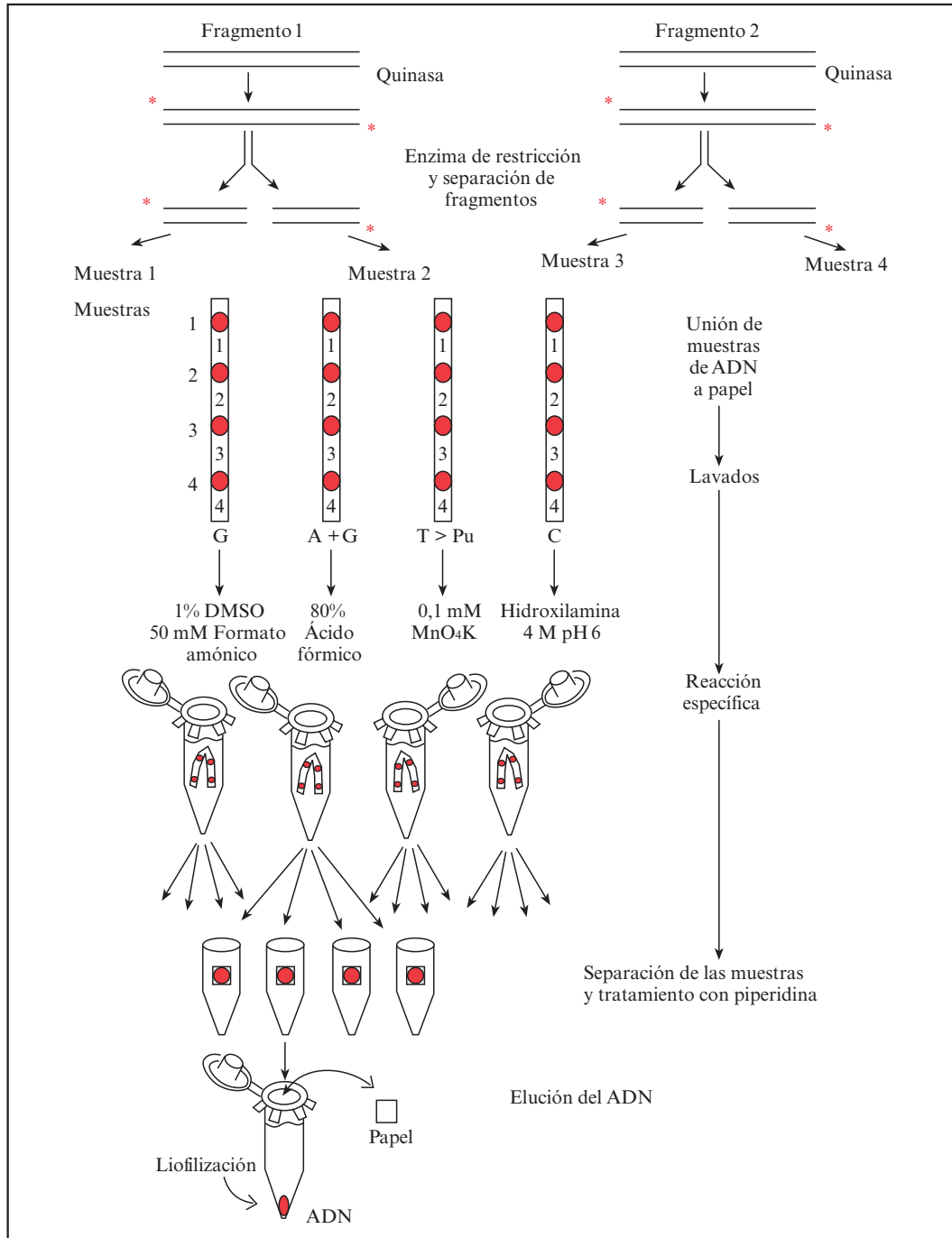


Figura 3.31. Método químico de secuenciación del ADN. Se representa esquemáticamente la secuenciación simultánea de dos fragmentos de ADN, separados en cuatro muestras. DMSO = dimetil sulfóxido.

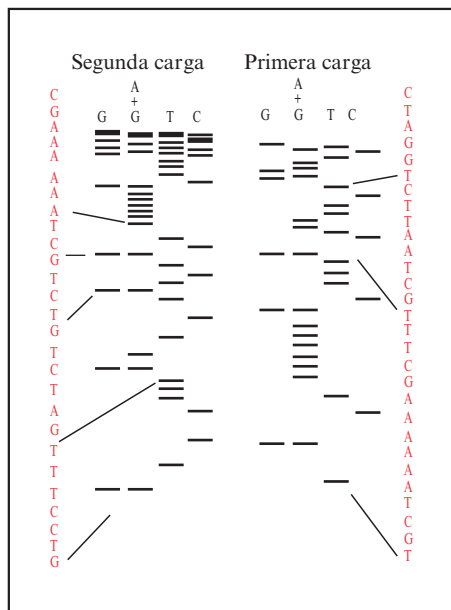


Figura 3.32. Esquema de las bandas de ADN que se observan en una autorradiografía y que permite elaborar la secuencia. Se representan las bandas correspondientes a dos cargas de la misma muestra de ADN introducidas en un gel a tiempos sucesivos para resolver fragmentos de tamaños crecientes.

Klenow o, más recientemente, la secuenasa son las enzimas más empleadas. De modo análogo al método químico, se establecen cuatro alícuotas, una para cada base. La síntesis se lleva a cabo en dos pasos: en la primera parte del proceso el cebador se alarga unos cientos de nucleótidos más, utilizando los cuatro deoxinucleótidos e incluyendo dATP marcado con ^{32}P , con ^{35}S , en el fosfato α como precursor radiactivo o con fluorocromos de un color. En el segundo paso se produce la terminación de cadenas por adición de una pequeña cantidad de un dideoxinucleótido distinto en cada alícuota (figura 3.33). El dideoxido es un nucleótido trifosfato que por carecer de grupos OH, tanto en el carbono 2' como en el 3' de la ribosa, imposibilita el crecimiento de la cadena y paraliza la síntesis. El resultado en cada tubo de ensayo es una colección de fragmentos, de tantos tamaños como unidades de la

base en concreto existan en la secuencia. Los distintos fragmentos se analizarán por electroforesis en geles de acrilamida y autorradiografía como en el caso anterior.

En la actualidad el marcaje radiactivo está sustituido por el marcaje con fluorocromos; se suelen utilizar cuatro fluorocromos distintos: fluoresceína, NBD (4cloro-7nitrobenceno-2-oxal-diazol), tetrametilrodamina y rojo de Texas, uno para cada base. Una vez conjugados a cuatro alícuotas del cebador, se procede a la extensión del iniciador e interrupción con dideoxidos. Los dideoxidos siempre están en cantidad limitante, mucho más baja que los deoxinucleótidos convencionales, para que la interrupción de la extensión sea azarosa y en un momento diferente en cada molécula. Las cuatro muestras se pueden correr en un mismo pocillo por electroforesis capilar, ya que tienen colores distintos y podrán ser distinguidos unos de otros por su diferente emisión (longitud de onda) a medida que van migrando (figura 3.34). Un láser emite un rayo hacia las bandas durante la migración y un detector mide el color; los resultados se transmiten a un ordenador que registra el resultado. Unas 800 bases pueden ser leídas del tirón por este procedimiento.

3.5.3. Secuenciación masiva o de nueva generación

Se denomina así a una serie de técnicas que permiten una secuenciación global por el procesamiento de una o varias muestras en tiempos muy cortos que dan lugar a la obtención de un número ingente de secuencias de fragmentos de ácido nucleico correspondientes a esas muestras. Se precisa de máquinas secuenciadoras y computadoras que permiten automatizar al máximo esta tecnología. Por tanto, hablaremos de secuenciación masiva, por ejemplo, si deseamos secuenciar un genoma entero de un organismo o varios de una especie para su comparación; si pretendemos secuenciar todos los ARNm de un tipo celular (ARN-seq); si deseamos secuenciar todas las regiones genómicas que interaccionan con la cromatina (ChIP-seq) o

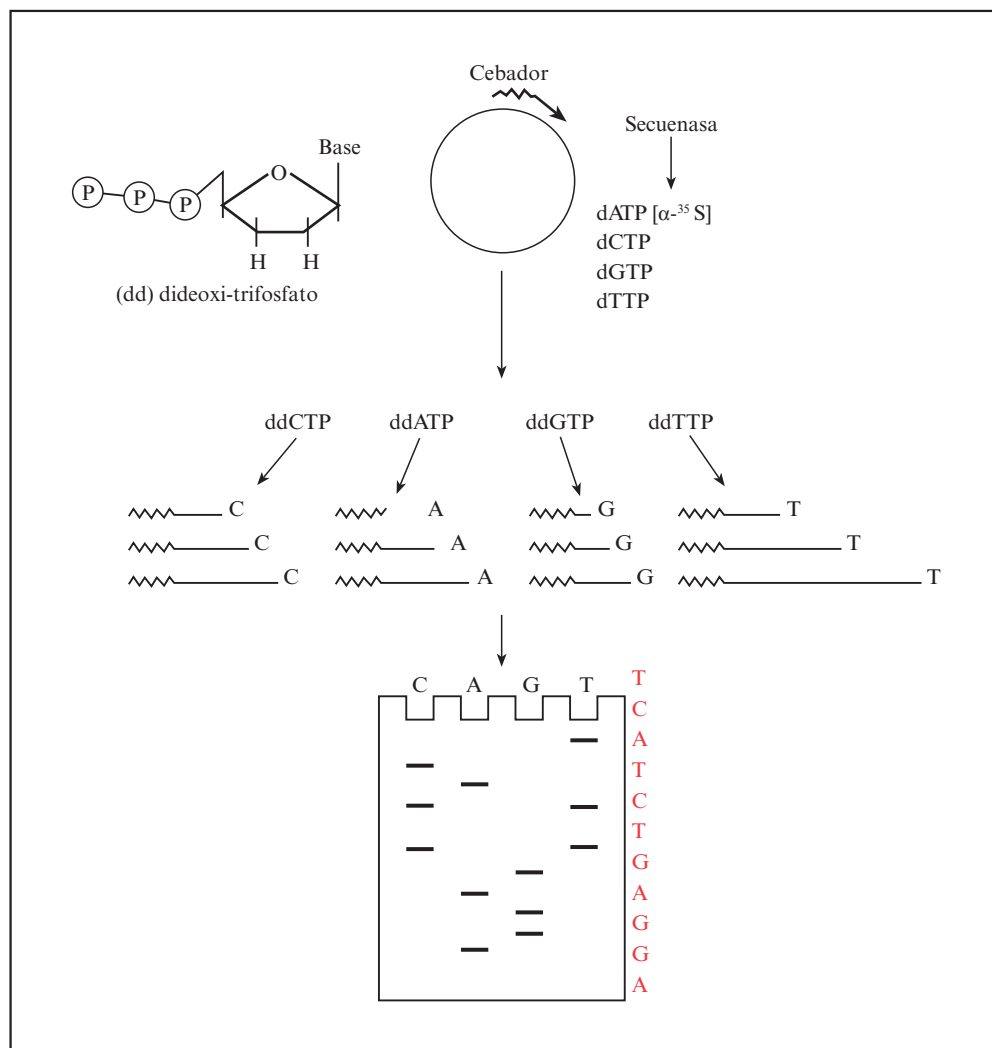


Figura 3.33. Secuenciación por el método enzimático. En la parte superior izquierda se representa la estructura de un dideoxi-trifosfato (dd) que bloquea la elongación de la cadena. Un ADN circular y de banda simple sirve de molde. En cuatro tubos de ensayo se coloca una alícuota del ADN circular, el cebador, la secuenasa y los cuatro dNTP (deoxinucleótidos trifosfato), uno de ellos marcado con el isótopo 35 del azufre en el fosfato alfa de la molécula (α -³⁵S-dATP). Al cabo de un cierto tiempo se bloquea la síntesis por adición de los dideoxidos correspondientes. Las cuatro alícuotas, por cada una de las cuatro bases bloqueadas, se analizan en paralelo en geles de poliacrilamida y el tamaño de los fragmentos se visualiza por autorradiografía.

todas las que están protegidas en un momento dado por factores de transcripción. También entraría en este enunciado la secuenciación de lugares metilados con valor epigenético a nivel de ge-

noma (metil-sec) y el diagnóstico de enfermedades por comparación de las secuencias correspondientes a los loci alterados en comparación con los normales (resecuenciación).

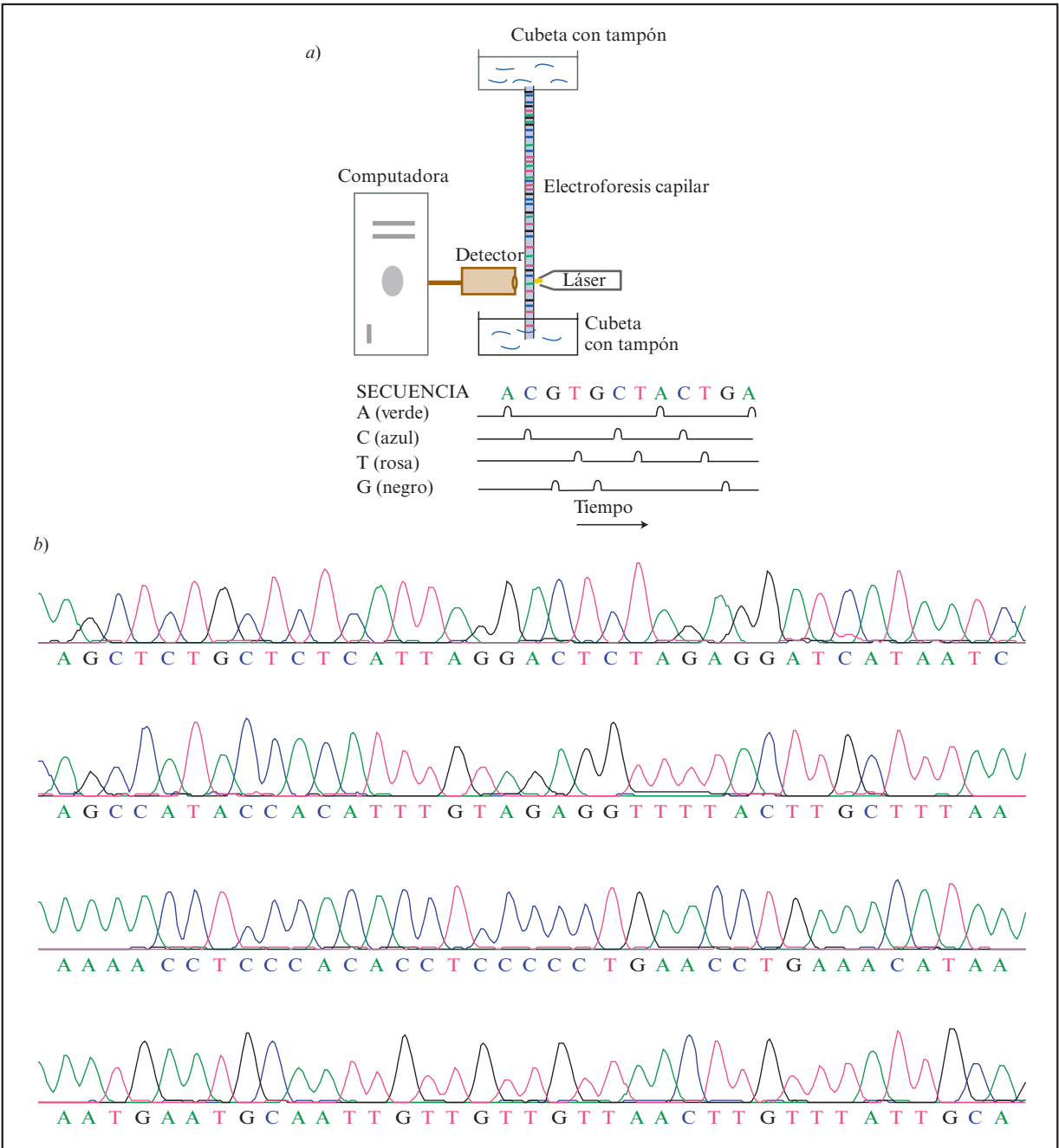


Figura 3.34. Secuenciación automatizada. a) Diagrama de una electroforesis capilar habiendo utilizado cuatro fluorocromos distintos en la que el color de cada banda corresponde al didesoxinucleótido terminal del fragmento; el detector del rayo láser mide la presencia de las bandas al pasar y su color, y b) Registro informatizado del resultado obtenido utilizando cuatro fluorocromos.

Secuenciación de un genoma completo

Partimos de múltiples copias del genoma a secuenciar que se fragmentará al azar tratándolo, por ejemplo, con una enzima de restricción que corte muy frecuentemente, pero limitando sustancialmente el tiempo de digestión (digestión parcial), o empleando métodos químicos de digestión, o bien mediante el uso de equipos especializados (sonicadores). A los distintos fragmentos se les une, mediante la enzima ligasa, unos oligonucleótidos adaptadores de secuencia conocida y universal; se añaden dos adaptadores distintos, uno en cada extremo de cada fragmento (5' o 3') de modo que todos los fragmentos 5' (o 3') tendrán el mismo adaptador universal (figura 3.35). Los fragmentos se fraccionan por tamaño seleccionando el rango de tamaño adecuado (300-800 pb, por ejemplo), y en bastantes casos, se someten a amplificación por PCR (véase capítulo 4, apartado 4.4). Finalmente se secuencian todos los fragmentos en paralelo en máquinas secuenciadoras donde el método enzimático u otros alternativos están automatizados. Los métodos alternativos suelen utilizar bases modificadas al estilo de dideoxis, pero con una modificación química reversible, además cada base terminadora está marcada con un fluorocromo de distinto color (T, verde; G, amarillo; C, azul, y A, rojo, por ejemplo). Así, cada vez que se introduce una base marcada, se termina la cadena momentáneamente, se registra el nucleótido añadido y se continúa añadiendo el siguiente nucleótido; una vez añadido → pausa de nuevo → registro → continuación, etc. Con esta técnica no hay que clonar los fragmentos generados (véase capítulo 4, apartado 4.2); por eso resulta metodológicamente sencillo y tiene una productividad tan desmesurada con un coste relativamente bajo. El paso siguiente es la ordenación de estos fragmentos secuenciados mediante bioinformática con ordenadores que utilizan algoritmos de cálculo basados en el solapamiento de secuencias en la secuenciación *de novo* o alineamiento contra el genoma de referencia, en la re-secuenciación. En la actualidad este paso constituye el cuello de botella de estas tecnologías, debido al enorme volu-

men de datos que deben ser manejados y al coste de las infraestructuras informáticas requeridas para ello (figura 3.35).

Secuenciación de ARNm

La secuenciación de ARNm mensajero es más difícil que la de ADN ya que la molécula es menos estable, por lo que generalmente se pasa a ADN-copia mediante la transcriptasa inversa y luego se procesa el ADNc mediante una metodología de preparación y secuenciación similar a la descrita en el apartado anterior.

Secuenciación de fracciones concretas del genoma

Puede interesar secuenciar, por ejemplo, la fracción del genoma que en un momento dado está asociada a factores de transcripción concretos (véase apartado 1.8, capítulo 1). El procedimiento se denomina coloquialmente como ChIP-seq (del inglés, *chromatin immunoprecipitation*) (figura 3.36). Comienza con el tratamiento de las células con formaldehído para provocar enlaces covalentes entre proteínas y ADN; seguidamente se lisan las células y se fragmenta la cromatina por sonicación a fragmentos de 200 a 1.000 pb; la fragmentación solubiliza la cromatina facilitando la unión con el anticuerpo específico contra un factor de transcripción concreto. La unión con el anticuerpo provoca la inmunoprecipitación arrastrando el ADN asociado. Seguidamente se lava el precipitado con varios tampones para eliminar las moléculas que se han unido de forma inespecífica y se revierten las uniones covalentes para liberar el ADN de los anticuerpos y la cromatina; el uso de bolitas magnéticas, añadidas al mismo tiempo que el anticuerpo, permite una rápida separación posterior con un imán o separador magnético de complejos ADN/Proteína del resto de la mezcla cromatínica. Además hay que revertir los enlaces covalentes entre los residuos de lisina y el ADN provocados por el formaldehído mediante

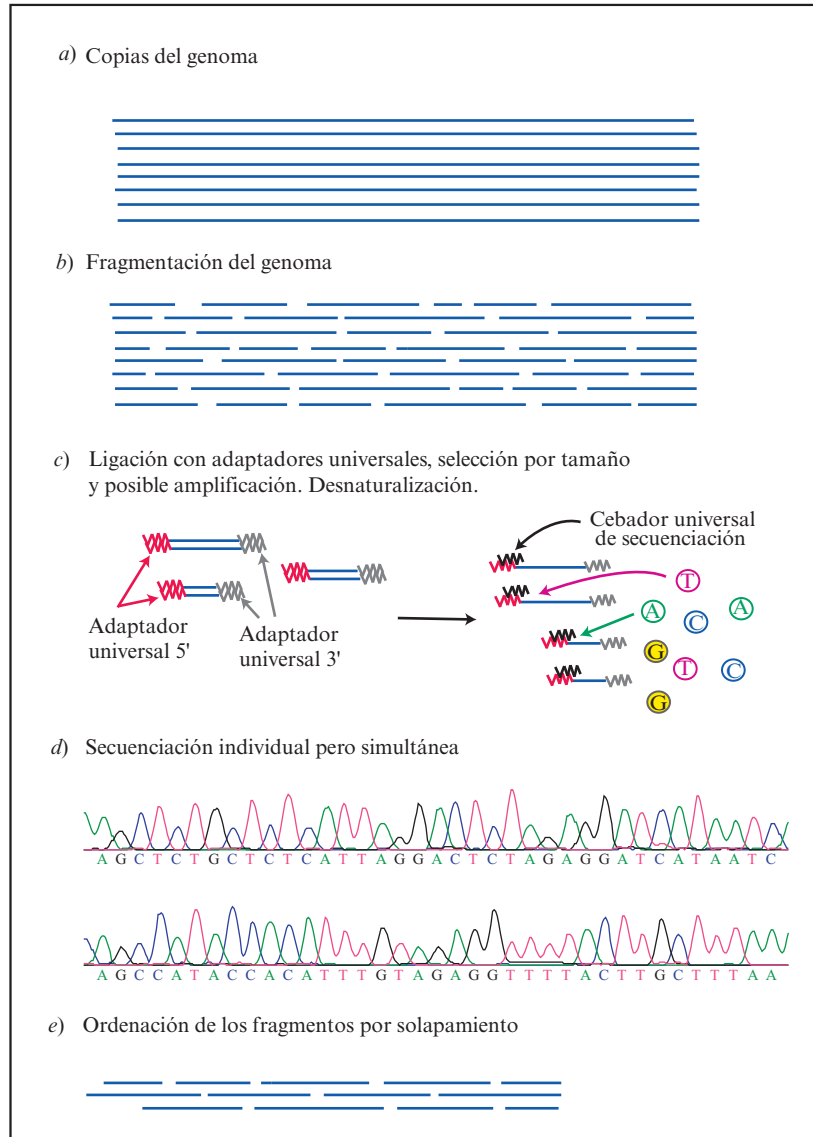


Figura 3.35. Distintas fases en la secuenciación de un genoma completo: a) múltiples copias del genoma a secuenciar, b) fragmentación al azar, c) ligación de adaptadores universales, d) secuenciación individualizada, e) ordenación de los fragmentos secuenciados mediante solapamiento en la secuenciación de novo.

tratamiento con proteinasa K y calor. Finalmente se purifica el ADN por extracción con fenol/cloroformo y se secuencia de modo paralelizado y automático ayudado por secuenciadores y ordenadores.

Las secuencias obtenidas de los distintos organismos incluido el hombre, así como múltiple información relacionada, se encuentra disponible de forma pública y gratuita en numerosas bases de datos en Internet.

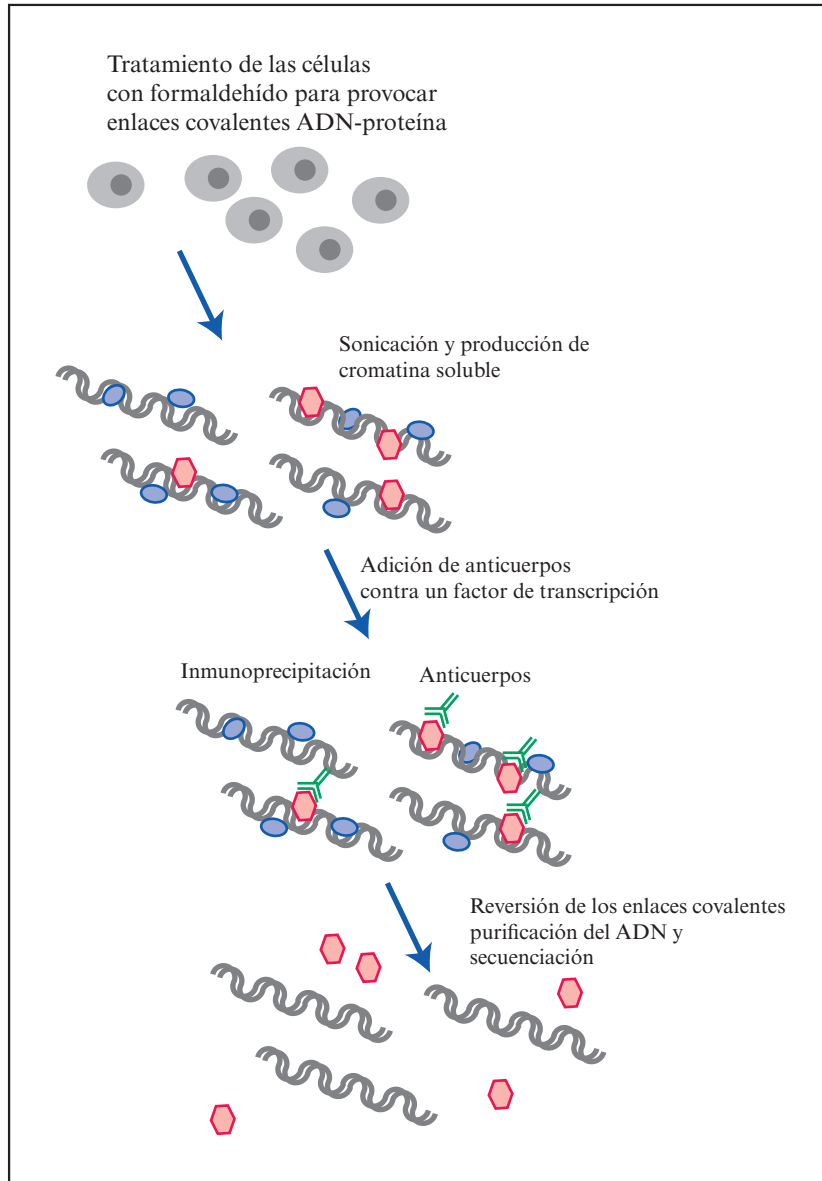


Figura 3.36. Estudios de interacción proteína-ADN y secuenciación de la fracción del genoma implicada (ChIP-seq). Fases consecutivas del procedimiento que se inicia con el tratamiento de las células con formaldehído, aislamiento del complejo cromatínico y sonicación de la muestra. Tratamiento con un anticuerpo contra un factor de transcripción, por ejemplo, e inmunoprecipitación. Reversión del tratamiento con formaldehído y purificación de la fracción de ADN que se secuenciará con la ayuda de secuenciadores y ordenadores.

LECTURAS COMPLEMENTARIAS

Cut-and-Paste of DNA Using an Artificial Restriction DNA Cutter, **Makoto Komiyama**. *Int. J. Mol. Sci.* **14**, 3343-3357. 2013.

Field guide to next-generation DNA sequencers, **Travis C. Glenn**. *Molecular Ecology Resources*, **11**, 759-769. 2011.

Sequencing technologies-the next generation, **Michael L. Metzker**. *Nature reviews Genetics*, **11**, 31-46. 2010.

4

Métodos de amplificación de secuencias: ADN recombinante y reacción en cadena de la polimerasa

Existen dos métodos fundamentales para obtener grandes cantidades de una secuencia concreta del genoma que nos permita su estudio y manipulación: el ADN recombinante y la reacción en cadena de la polimerasa (PCR, del inglés *polymerase chain reaction*). Estas dos técnicas son los pilares de la ingeniería genética.

Anteriormente a las técnicas de ADN recombinante era prácticamente imposible aislar un gen concreto de un genoma complejo, como los de la mayoría de eucarióticas. La pequeñísima proporción de un gen humano por ejemplo (una milonésima parte), con relación a todo el genoma, imposibilitaba su aislamiento en concentraciones suficientes para su posterior caracterización. El ADN recombinante inició su desarrollo a principios de los años setenta del siglo pasado, coincidiendo con el descubrimiento de las enzimas de restricción. Lo verdaderamente importante de estas enzimas es que permiten unir fragmentos de procedencia muy diversa y, por tanto, diseñar construcciones genéticas capaces de replicarse en células hospedadoras para obtener un elevado número de copias idénticas a la original. Las técnicas de ADN recombinante nos permiten aislar fragmentos discretos de ADN procedentes de cualquier genoma y mantenerlos, reproducirlos y expresarlos en sistemas sencillos que faciliten su estudio a nivel molecular. Es posible secuenciar y mutar el gen clonado, ponerlo bajo óptimas condiciones de expresión: en distintos huéspedes, observar su comportamiento, determinar su estructura tridimensional, comprobar sus interacciones

con otras macromoléculas, etc. La ingeniería genética ha roto la barrera de las especies y ha cambiado el concepto de herencia. Genes eucarióticos se expresan en sistemas procarióticos con una gran eficiencia; genes procarióticos se insertan en genomas de mamíferos y también se expresan; un gen humano puede ser aislado, mutado y sustituido.

La finalidad de la obtención de ADN recombinante en bacterias puede deberse a:

1. Obtener grandes cantidades de una secuencia eucariótica para su estudio.
2. La obtención de grandes cantidades de un producto eucariótico con fines quizá comerciales.

Desde su introducción en 1985, la reacción en cadena de la polimerasa ha transformado totalmente los métodos de aislamiento y amplificación del ADN, tanto en laboratorios de investigación básica como aplicada. La técnica la descubrió Kary B. Mullis, que diseñó una manera simple y elegante de sintetizar grandes cantidades de ADN *in vitro* basándose en el procedimiento que la célula emplea *in vivo*. El descubrimiento le valió el Premio Nobel en 1993 (compartido con el canadiense Michael Smith). La PCR ha sustituido en gran medida la amplificación de fragmentos por clonación y tiene numerosísimas aplicaciones en diversos aspectos de la biología molecular. Es una técnica fundamental e imprescindible en ingeniería genética.

4.1. ELEMENTOS DEL ADN RECOMBINANTE

Uno de los objetivos más importantes de la clonación es la utilización de una unidad replicativa procariótica para obtener grandes cantidades de un fragmento de ADN eucariótico. La unidad replicativa constituye el vector, y el fragmento amplificado, el inserto. Toda clonación molecular

implica una serie de procesos y elementos representados en la figura 4.1:

- Un vector autorreplicativo.
- La generación de fragmentos de ADN.
- La unión del vector y los fragmentos.
- La introducción en un hospedador en el que se pueda reproducir.

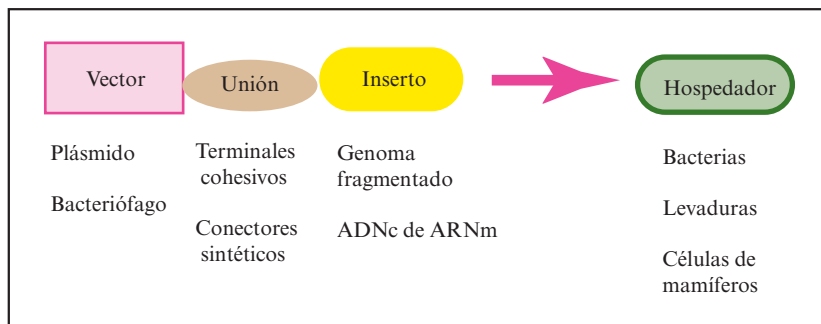


Figura 4.1. Elementos necesarios para la clonación molecular.

4.2. ELEMENTOS DE LA CLONACIÓN MOLECULAR EN BACTERIAS

Las bacterias han sido los primeros hospedadores de la clonación molecular de genes eucarióticos y las primeras en las que estos genes recombinantes se han expresado. En la actualidad siguen ocupando un papel importante en la clonación de genes procedentes de otros organismos. La introducción en una bacteria de una secuencia eucariótica de tamaño medio (10-20 kb) no produce grandes alteraciones a nivel de replicación o viabilidad, por lo que, si deseamos exclusivamente obtener muchas copias del gen o la secuencia clonada, no hay hospedador comparable a una bacteria. Su versatilidad, tiempo de generación y bajo coste coloca a las bacterias en una posición difícil de superar en cuanto a ventajas de clonación.

La introducción de un gen funcional eucariótico con la intención de una expresión fidedigna representa una situación algo más exigente y compleja. La transcripción del inserto requiere la presencia de un promotor reconocible por la ARN

polimerasa de la bacteria. Con este objeto, casi todos los vectores procarióticos llevan el promotor del gen *lac* (β -galactosidasa) u otro promotor procariótico precediendo el lugar de policlonaaje.

Los vectores más importantes en la clonación molecular en bacterias son los plásmidos y el fago λ , pero hay otros también interesantes para casos concretos como los cósmicos para insertos de gran tamaño junto con los cromosomas artificiales de bacteria, o el bacteriófago M13 y sus derivados cuando se quiera productos de ADN de una sola hebra, para secuenciación por ejemplo. Veremos cada uno de estos vectores a continuación.

4.2.1. Plásmidos como vectores

Los plásmidos son replicones que se transmiten y mantienen de manera estable como círculos extracromosómicos. Se replican autónomamente, pero utilizan enzimas de la bacteria: ADNpol I y III, ARNpol, etc. El número de copias del geno-

ma del plásmido por bacteria es variable, y va desde 1 a 3.000 por célula. En muchos, el control del número de copias radica en una región del plásmido, cercana al origen de replicación (*ori*), denominada *rop*; el conjunto *rop-ori* se denomina replicón. La replicación de plásmidos dependientes de la ADN polimerasa I puede continuar en ausencia de síntesis de proteínas; así, en presencia de cloranfenicol la replicación de la bacteria se inhibe, pero la del plásmido puede continuar hasta adquirir de 2.000 a 3.000 copias por célula. El control de *rop* es indirecto. En la región -500 hay un promotor que dirige la síntesis de un ARN II que sirve de cebador en la síntesis de ADN. En la misma región, pero en sentido contrario, también se sintetiza otro pequeño ARN (ARNiI) que es com-

plementario a una pequeña parte del ARN II (figura 4.2). La proteína de 63 aminoácidos Rop favorece la hibridación entre ARNiI y ARN II estabilizando el híbrido I-II. Si el ARN II hibrida con el I, no estará libre para formar el híbrido ADN-ARN que inicia la replicación del ADN. En conclusión, Rop impide la iniciación de la replicación. Mutaciones en Rop impiden que el ARNiI intervenga en el enlace ADN-ARN II, favoreciéndose la replicación. Los vectores plasmídicos más populares, los pUC, son deficientes en Rop y, por tanto, pueden formar muchas copias. Los replicones son característicos de cada tipo de plásmidos, y plásmidos con distintos replicones pueden ser incompatibles. Hay unos treinta tipos de incompatibilidades entre los plásmidos de *E. coli*.

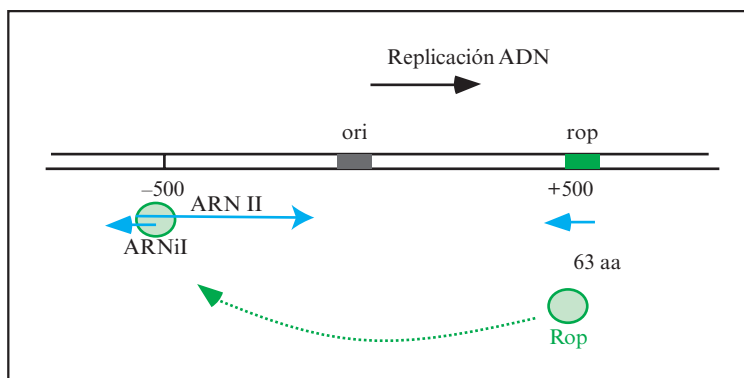


Figura 4.2. Regulación del número de copias por el producto del gen Rop que estabiliza la unión entre ARN II y ARNiI impidiendo que ARN II actúe de cebador en el inicio de la replicación.

Los plásmidos que se utilizan como vectores tienen las siguientes características:

- Tamaño pequeño (alrededor de 3 kb).
- Carecen de los genes *mob* y *tra* de movilidad y transferencia, que serían los que les permitirían moverse por conjugación de una bacteria a otra. Son, pues, inmóviles. También son *rop*⁻.
- Llevan marcadores que permiten la selección de los transformantes. Los más comunes son genes que confieren resistencia a antibióticos: ampicilina, tetraciclina, clo-

ranfenicol y neomicina (kanamicina). **La ampicilina** es una penicilina semisintética, antibiótico β -lactámico de amplio espectro que inhibe la formación de los puentes de péptido-glicanos por las enzimas transpeptidasas y, por tanto, inhibe la síntesis de la pared celular provocando la lisis de la bacteria; el gen que confiere la resistencia a este antibiótico es el de la β -lactamasa que hidroliza la ampicilina. **La tetraciclina** desacopla la fosforilación oxidativa e inhibe la síntesis proteica al unirse a la subunidad 30S del ribosoma; la resistencia al an-

tibiótico procede de una proteína represora que interacciona con el antibiótico, o bien de una disminución de su permeabilidad a que entre en la célula. **El cloranfenicol** inhibe la síntesis proteica de la bacteria al unirse a la subunidad 50S del ribosoma e impedir la formación del enlace peptídico; sin embargo, el plásmido podrá seguir replicándose en su presencia hasta alcanzar 2.000-3.000 copias por bacteria. **La neomicina** es un aminoglicósido que se une irreversiblemente a la subunidad 30S del ribosoma; el gen de resistencia *neo* se traduce en la enzima aminoglicósido-fosfo-ribosil transferasa, originalmente aislada del transposón Tn5, que inhabilita el antibiótico. La misma enzima inhabilita la kanamicina en plantas y el antibiótico G418 en células de mamíferos, estos dos últimos antibióticos también se unen a la subunidad pequeña de los ribosomas en los distintos casos, bloqueando la síntesis de proteínas.

- Llevan adaptadores de policlonaje que son secuencias únicas para 10 o más dianas de restricción, una a continuación de otra.
- La clase de los pUC lleva un segmento de ADN del operón de la lactosa (*lac*) que codifica el extremo amino-terminal de la β -galactosidasa, que en presencia del cromóforo X-gal complementa mutaciones de la bacteria y origina colonias azules. Los mutantes deficientes (y, por tanto, los recombinantes) producen colonias blancas.

Los plásmidos más populares son los de la serie pUC que fueron desarrollados en la Universidad de California (de ahí su nombre) y son totalmente artificiales. Las variaciones entre unos y otros son pequeñas como distintas dianas en la región de policlonaje o ausencia del represor del operón de *lac*. La resistencia al antibiótico ampicilina y el origen de replicación proceden de otro plásmido artificial, el pBR322, construido por los mexicanos Bolívar y Rodríguez (de ahí su nombre), de 4,3 kb y, a su vez, derivado de los plásmidos pSC101 y pMB9. Del pMB9 tienen el origen de replicación; del pSC101 la resistencia a la am-

picilina y a la tetraciclina. El PSC101 adquirió su resistencia a la ampicilina del plásmido RSF2124 y éste, a su vez, de un transposón que porta.

Estos vectores plasmídicos tienen su tamaño reducido al mínimo, para así aumentar su capacidad de aceptar insertos, ya que la transformación de recombinantes con un tamaño superior a las 15 kilobases (kb) en total (vector más inserto) es muy ineficiente. Todos ellos contienen un adaptador de policlonaje con un amplio abanico de dianas para enzimas de restricción que se pueden usar individualmente o combinadas. En este último caso se puede definir la orientación del inserto. En muchos casos la región de policlonaje está situada en medio de un segmento derivado del operón de *lac* de *E. coli*. El gen *lac I* se traduce en el represor del operón, le sigue el promotor/operador y el principio del gen *lac Z* responsable del extremo amino-terminal de la β -galactosidasa. Este extremo amino-terminal es capaz de una complementación intraalélica con una forma defectuosa de β -galactosidasa codificada por la bacteria. La síntesis de β -galactosidasa se induce por isopropil-tio- β -D-galactosido (IPTG). Bacterias expuestas a IPTG sintetizan ambas partes de la enzima y forman colonias de color azul en presencia del sustrato cromógeno X-gal (5-bromo, 4-cloro, 3-indolil, β -D-galactósido). La interrupción del gen por el segmento de policlonaje no afecta a la expresión, pero la inserción de un fragmento grande de ADN en el lugar de policlonaje inactiva la complementación intraalélica. Las bacterias que llevan plásmidos recombinantes son, por tanto, blancas y pueden expresar el gen insertado.

Algunos plásmidos derivados del pUC18 llevan promotores derivados de los bacteriófagos T3, T7 y/o SP6 inmediatamente detrás del lugar de policlonaje (figura 4.3). El ADN insertado en el sitio de policlonaje puede ser muy eficientemente transcrito *in vitro* en presencia de la polimerasa correspondiente y los ribonucleótidos-trifosfato. En algunos casos se pone a cada lado del sitio de policlonaje un promotor distinto, lo que asegura la transcripción de ambas hebras del ADN clonado dependiendo de la polimerasa que se use en cada caso. Los ARN originados de este modo pueden, bien traducirse a proteínas *in vitro*

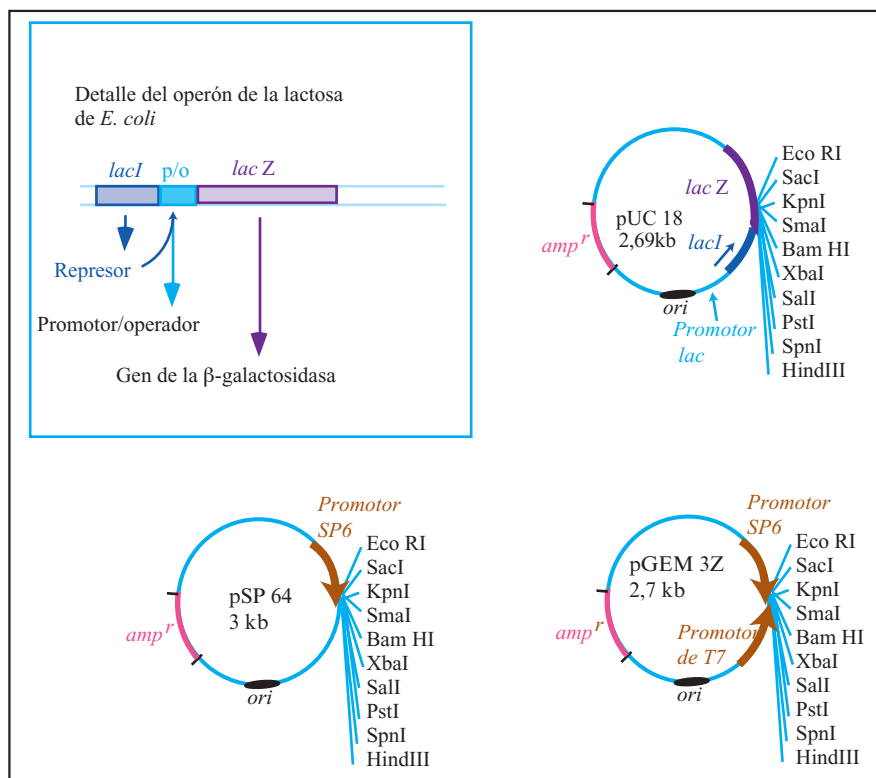


Figura 4.3. Distintos componentes del plásmido pUC y en la parte inferior vectores derivados de pUC con promotores especiales.

(si se transcribe la hebra con sentido), o utilizarse como sondas de hibridación (cualquiera de las hebras).

4.2.2. El bacteriófago λ como vector

Ya hemos descrito las características del bacteriófago λ y cómo se integra en el genoma de la bacteria (capítulo 2, apartado 2.2.1 y figura 2.10). De su mapa genético se deduce que genes relacionados entre sí se localizan próximamente excepto por los dos genes de regulación positiva N y Q. Se han identificado 46 genes, de los cuales siete parecen no ser esenciales. Genes localizados en la parte izquierda del mapa lineal codifican las proteínas de la cabeza y la cola de la partícula fágica

(figura 4.4a). Genes en la región central están implicados en la recombinación y los procesos de lisogenización; estos genes no son esenciales para la infección y crecimiento del fago, por lo que pueden ser sustituidos en los procesos de clonaje. Genes hacia la derecha de la región central son de regulación e inmunidad (*N*, *cro*, *cI*), y a continuación están los relativos a la síntesis de ADN (*O*, *P*). En el extremo derecho se localizan los relativos a las funciones tardías de regulación (*Q*) y a las funciones de lisis de la bacteria.

Durante la fase inicial de la infección, el fago circularizado se replica bidireccionalmente desde un único origen (*ori*) que se activa por los productos de los genes *O* y *P*. Más tarde, empieza la replicación por el círculo rodante y se producen moléculas lineales del profago cabeza-cola (figu-

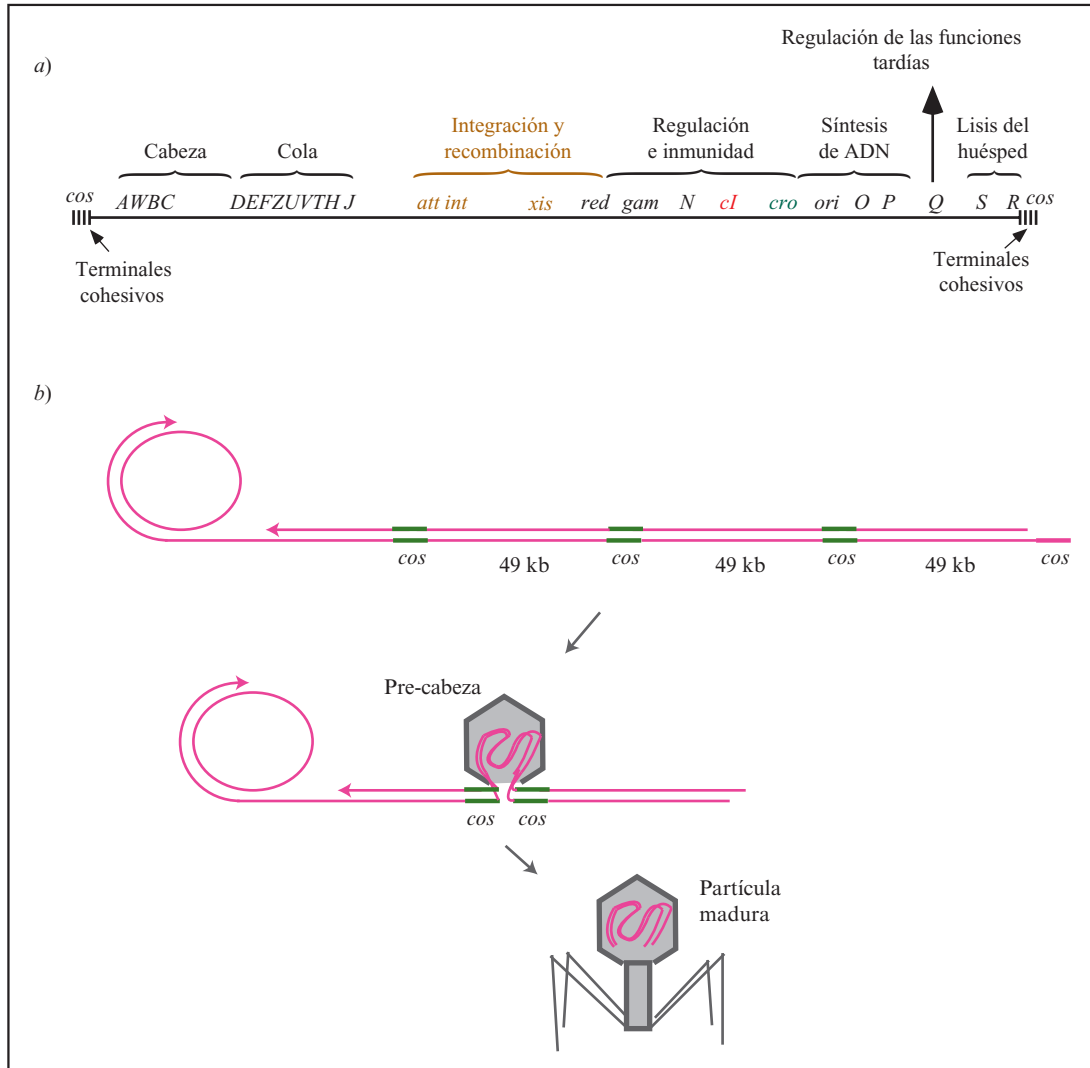


Figura 4.4. Representación esquemática a) del mapa genético del fago λ , y b) duplicación por el círculo rodante y empaquetamiento del bacteriófago.

ra 4.4b). Estas moléculas se cortan en los lugares *cos* por acción de la proteína A (genes de la cabeza).

El fago λ tipo silvestre no es adecuado para su utilización como vector, ya que tiene numerosas dianas de restricción para cualquier enzima y baja capacidad (sin eliminar parte del genoma propio) para insertos de gran tamaño. Las varian-

tes aptas para su utilización como vectores de clonación son aquellas en las que se han eliminado dianas para numerosas enzimas de restricción, dejando sólo las que se encuentran en lugares idóneos. Los vectores de sustitución son aquellos en los que un segmento no esencial del fago es reemplazado por un inserto de ADN heterólogo; los vectores de inserción, en cambio, se originan por

integración directa del inserto en una diana del bacteriófago.

Vectores de inserción

Dos de los vectores de inserción más populares son los λ gt10 y λ gt22.

El λ gt10 permite clonar pequeños fragmentos de ADNc en la región de inmunidad (gen *cI*). El

gen *cI* de lambda se sustituyó por el de un lambda doide afín (λ 434) por poseer una diana única para EcoRI y otra para HindIII (figura 4.5). Los recombinantes, que serán *cI*⁻, se seleccionan en una cepa de *E. coli* con una mutación en el gen *hfl* (alta frecuencia de lisogenizar). Los bacteriófagos que no han adquirido ADN heterólogo lisogenizarán muy eficientemente y no formarán calvas.

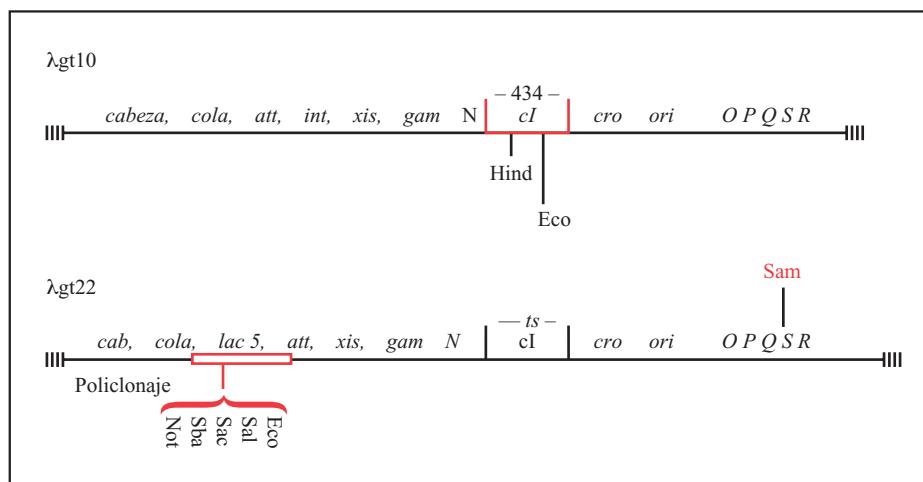


Figura 4.5. Mapas genéticos de los vectores λ gt10 y λ gt22.

El vector λ gt22 posee el gen *lac Z* de *E. coli* y las inserciones se realizan en la región de policlonaje introducida en el gen *lac*; los recombinantes pasan a ser Lac⁻. Además, el λ gt22 es un vector de expresión que produce proteínas con el extremo amino de β -galactosidasa fusionado al producto de la secuencia insertada, lo que permite utilizar métodos inmunológicos (anticuerpos anti-Lac) para detectar su expresión en *E. coli*. Permite insertos de hasta 8 kb y se utiliza frecuentemente para elaborar genotecas de ADNc (apartado 4.3.2). El represor es mutante sensible a la temperatura, por lo que se puede controlar la inducción del ciclo lítico y la síntesis de las proteínas fusión. El vector tiene además una mutación en el gen *S* que afecta a la lisis de la bacteria (Sam100). En un hospedador no supresor, la bacteria acumula el

producto clonado y no es capaz de lisar, con lo cual el producto no se diluye en el medio.

Vectores de sustitución

Ejemplos de vectores de sustitución son las series caronte y EMBL.

Se utilizan preferentemente en la elaboración de librerías genómicas eucarióticas (apartado 4.3.1). Su ventaja con respecto a los plásmidos radica en que se pueden almacenar 10⁹ recombinantes en menos de un mililitro y en la eficiencia de transducción que supera a la de transformación.

Los vectores caronte han tomado el nombre de la mitología griega. Caronte era el barquero que transportaba las almas de los fallecidos de la

orilla del río Aqueronte, donde estaban los vivos, a la orilla opuesta, donde se encontraban los muertos (el infierno, pero sin connotaciones de fuego eterno). Son preferentemente vectores de reemplazamiento aunque también hay algunos de inserción como los caronte 16A y 21A. Ambos tienen dianas únicas para EcoRI que en el caronte 16A se encuentra dentro del gen *lac Z* de *E. coli* previamente introducido en el fago λ por sustitución de las regiones *red* y *gam*. Así los fagos con inserto originarán calvas blancas al infectar un

césped de bacterias, mientras que los que carezcan de inserto darán calvas azules. El vector de inserción caronte 21A carece también de *red* y *gam* admitiendo insertos de hasta 9 kb.

El caronte 35 es un vector de sustitución, posee dos regiones de policlonaje a ambos lados de una región de relleno procedente de *E. coli* y que será sustituida por el inserto eucariótico. Admite fragmentos de hasta 21 kb, utilizando cualquiera de las siete dianas de restricción orientadas en sentido opuesto a ambos lados de las secuencias

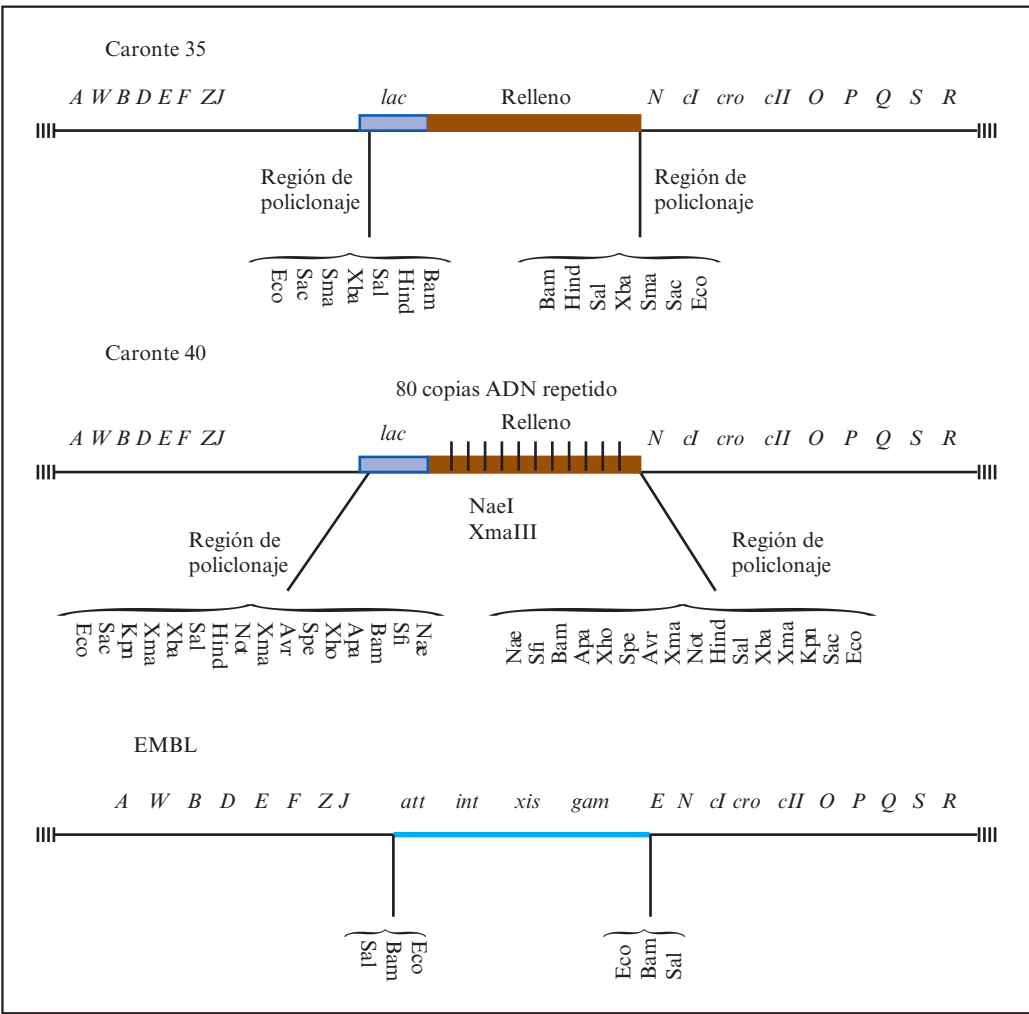


Figura 4.6. Mapas de algunos vectores de sustitución de la serie caronte y EMBL derivados del fago λ .

de relleno. La región de policlonaje izquierda está situada en el gen de la β -galactosidasa de *E. coli*, por lo que se pueden obtener proteínas fusión detectables inmunológicamente con anticuerpos anti-Lac. El único inconveniente de este vector es la necesidad de aislar los brazos y separarlos del relleno antes de recibir al inserto, para así evitar la reconstrucción del fago original; los brazos sin inserto no son capaces de ser empaquetados eficientemente debido a su falta de tamaño crítico (figura 4.6).

El caronte 40 elimina el problema de deshacerse del relleno previo al clonaje. El relleno consiste, en este caso, en unas 80 copias de una secuencia de 235 pares de bases y digerible con las enzimas NaeI, XmaIII; los fragmentos resultantes se pueden separar con una simple precipitación con polietilenglicol. La región de policlonaje tiene 16 dianas de restricción y admite insertos de hasta 24 kb (figura 4.6).

La serie de vectores de sustitución EMBL son particularmente buenos para clonar genotecas de ADN digeridas parcialmente con la enzima Sau3A. Tiene una diana Bam flanqueada por dianas Eco y Sal. Los insertos clonados vía Sau3A pueden recuperarse por digestión con la enzima Sal (figura 4.6). El fago λ tipo silvestre (Spi^+) no puede crecer en cepas lisogénicas para el fago P2. La incompatibilidad radica en los genes *red* y *gam*, que se eliminan con el relleno (la bacteria sin estos genes pasa a ser Spi^-). Un tipo de selección en estos vectores es utilizar precisamente cepas de bacterias lisogénicas para P2, en las que sólo los fagos que hubiesen incorporado un inserto (convertidos en Spi^-) podrían crecer.

4.2.3. Otros vectores para la obtención de ADN recombinante en bacterias

Los cósmidos

Los cósmidos son plásmidos a los que se les ha insertado la región *cos* del fago lambda. Son particularmente útiles para clonar insertos de gran tamaño. La región *cos* confiere al plásmido

la facilidad para ser empaquetado *in vitro* por el sistema del fago siempre que existan de 37 a 52 kb entre *cos* y *cos*, lo que implica una selección efectiva para insertos de entre 40-45 kb (figura 4.7). Una vez que el fago haya inyectado el ADN en la bacteria, éste se comportará como un plásmido y podrá ser seleccionado por resistencia a ampicilina.

Los cósmidos son particularmente atractivos cuando se necesitan insertos de gran tamaño, sin embargo, cabe la posibilidad de que dos o más fragmentos alejados en el genoma se unan durante la ligación y generen un fragmento del tamaño adecuado para su empaquetamiento *in vitro*. El tamaño de los insertos también favorece la posible recombinación homóloga entre secuencias repetidas dentro del inserto, algo que se puede paliar o inhibir utilizando bacterias Rec^- como hospedadoras. El problema con las bacterias Rec^- es que en muchos casos la reparación de daños será deficiente.

El bacteriófago M13 y sus derivados fagómidos

El M13 es un fago filamentosos, su ADN es circular, de hebra sencilla y 6,4 kb. Posee 10 genes, todos ellos esenciales, más una región intergénica donde está el origen de replicación *ori*. Infecta bacterias portadoras del episoma F a través de los filamentos sexuales F o pili que emiten las bacterias macho (capítulo 2, apartado 2.1.1, figura 2.1). La replicación del fago no produce la lisis de la bacteria, sino que ésta continúa creciendo lentamente excretando partículas al medio. Pueden llegar a ser liberadas hasta 1.000 partículas por célula y generación. El proceso total ocurre en veinte minutos y se repetirá indefinidamente.

El ADN viral tiene una fase replicativa en la que se originan unas 100 copias de ADN bicatenario. Posteriormente, se acumulan unas proteínas de unión a ADN de banda simple codificadas por el fago, que impiden la síntesis de la cadena complementaria. La maduración tiene lugar en la membrana, donde se sustituyen las proteínas de unión a ADN de hebra sencilla por las proteínas

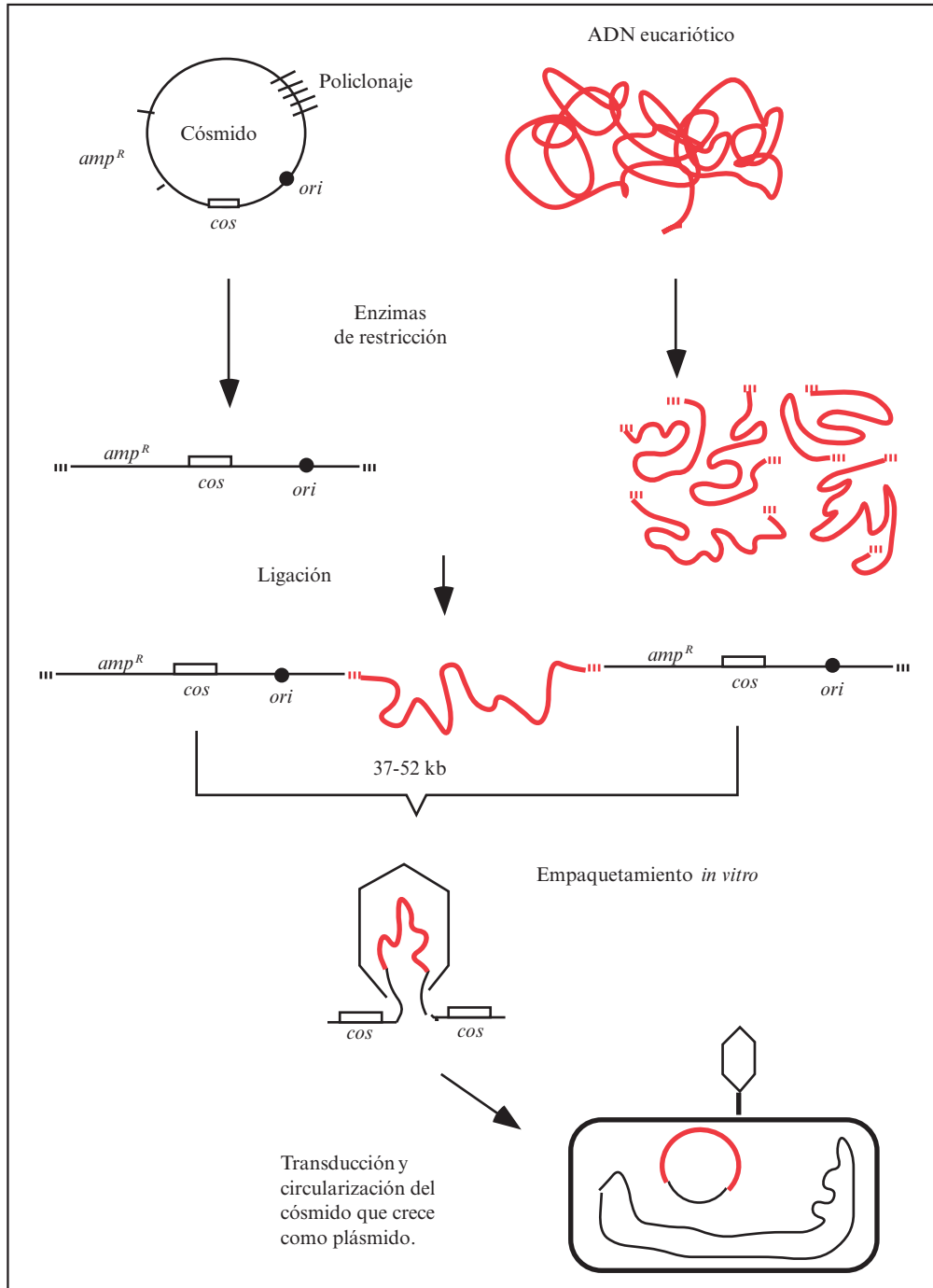


Figura 4.7. Esquema de clonación molecular utilizando un cósmido como vector. El cósmido posee la región *cos* del fago λ insertada en un pUC.

de la cápside (figura 4.8). Tanto el ADN de la forma replicativa como el de hebra sencilla pueden

transformar células competentes de *E. coli* y proseguir con el ciclo vital.

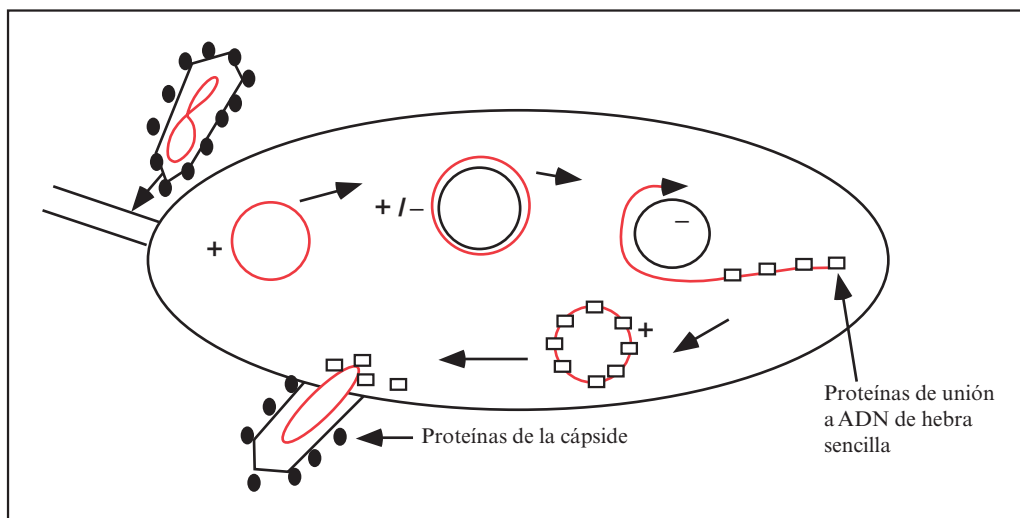


Figura 4.8. Ciclo del bacteriófago M13.

En los procesos de clonación es mucho más sencillo trabajar con ADN bicatenario. Ésta es la razón por la que se utiliza preferentemente la forma replicativa de M13. Esta forma se puede aislar de células infectadas, manipular como cualquier ADN de plásmido y reintroducir en células por transformación. El ADN de doble banda se incorpora al ciclo de replicación y generará partículas fágicas que contendrán sólo una de las dos hebras del ADN (+); la otra (-) nunca se empaqueta. Tanto el ADN de la forma replicativa como el de hebra sencilla del fago pueden transfectar células competentes de *E. coli* y proseguir con el ciclo vital.

En las clonaciones que utilizan como vector el M13 no debería haber grandes restricciones en cuanto al tamaño del inserto, puesto que el tamaño final de la partícula viene determinado por la longitud del ADN; sin embargo, insertos grandes son muy inestables y acaban por ser eliminados o acortados.

En la actualidad los vectores más utilizados no llegan a 3 kb de tamaño, poseen 21 dianas para enzimas de restricción y sólo mantienen un frag-

mento de 1,3 kb del M13 original o de otro fago filamentoso afín (el fago f1, que guarda una homología del 97% con el M13), que contiene las secuencias requeridas para su replicación y morfogénesis. El origen y la terminación de la replicación vienen determinados por dos secuencias distintas pero solapantes. La proteína-producto del gen 2 se une a la secuencia de iniciación originando un corte en la banda (+) de la forma replicativa comenzando la replicación por el círculo rodante. El segundo corte se produce en la secuencia terminal que solapa con la inicial y el ADN es ligado para producir el círculo genómico de banda (+).

Estos vectores, mitad fago y mitad plásmido, se denominan fagómidos y se pueden obtener comercialmente. El clonaje en fagómidos es particularmente atractivo para secuenciar por el método enzimático con dideoxidos (capítulo 3) y para realizar mutagénesis dirigida (capítulo 5). El denominado *bluescript* lleva el origen de replicación de pUC y los primeros 146 aminoácidos del gen *lac-Z* que, por complementación intraalélica, rescata una mutación en el mismo gen de la bacteria y permite se-

leccionar por colonias coloreadas cuando se utiliza como plásmido (figura 4.9). En estos vectores todos los genes de M13 son esenciales, por lo que se necesita un fago auxiliar que facilite en trans las funciones eliminadas. A cada extremo de la región de policlonaje se encuentra el promotor T7 o el T3 en sentidos opuestos para poder transcribir el inserto en cualquier orientación en que entre. En ausencia de un fago auxiliar (M13 o f1) los fagómidos

son indistinguibles de los plásmidos, pero en presencia de uno de ellos se activa el origen de replicación fágico, se produce ADN monocatenario y se exportan viriones al medio; estos viriones serán de dos tipos: recombinantes y auxiliares. En la mayoría de los casos (para secuenciar o realizar mutagénesis dirigida) no será necesario separar los unos de los otros, ya que se emplearán cebadores específicos que sólo están presentes en los recombinantes.

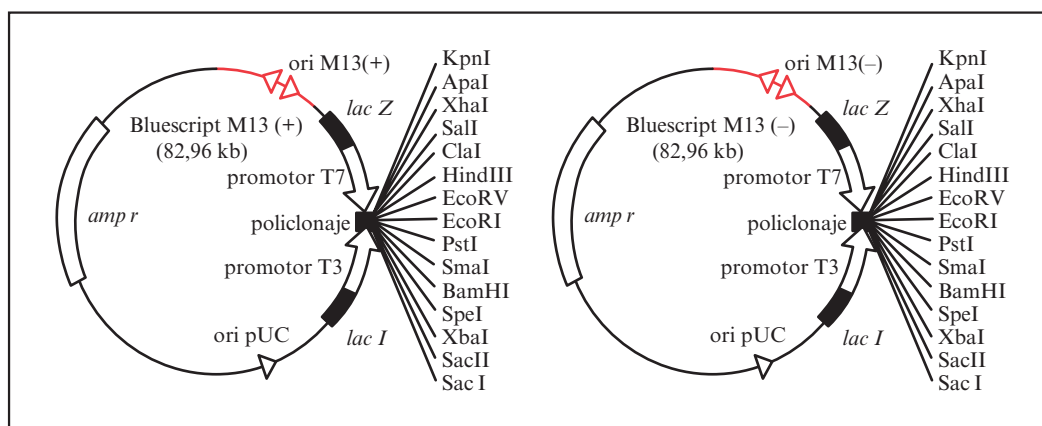


Figura 4.9. Mapa genético del vector bluescript M13, en sus formas replicativas (+) y (-). El origen de replicación es de pUC, pero también lleva las secuencias requeridas para la replicación y morfogénesis de M13 (ori-M13). El vector de la izquierda originará genomas monocatenarios (+) y el de la derecha (-).

Un fagómido complejo y sofisticado es el λ ZAP, particularmente indicado para la clonación de ADNc, ya que no permite insertos de más de unas 10 kb. Es un fago λ con la región de 1,3 kb específica de M13 o f1. Tiene la característica de que el inserto clonado puede liberarse automáticamente del vector fágico *in vivo* con la ayuda de un M13 o f1 auxiliar, quedando el fragmento suelto como un plásmido. Cuando una bacteria que lleva un plásmido F de fertilidad se infecta con el λ ZAP recombinante (figura 4.10) y se superinfecta con un M13 auxiliar, el fago auxiliar facilita una serie de proteínas en trans que reconocen las dos secuencias incorporadas en los brazos de λ ZAP que determinan el inicio y la terminación de la replicación en M13 o f1. El fragmento así replicado se circulariza, se empaqueta como f1 o M13 y se ex-

porta al medio de cultivo de donde se puede aislar y utilizar para infectar nuevas bacterias. En estas nuevas bacterias, y en presencia de M13 o f1 auxiliar, crecerá como plásmido y se podrá seleccionar por resistencia a ampicilina. Los clones con inserto perderán la capacidad de formar colonias azules en una bacteria que complemente la mutación en *lac Z* y en presencia del inductor IPTG y el sustrato cromógeno X-gal.

Cromosomas artificiales de bacterias como vectores

Los cromosomas artificiales de bacterias (CAB, o en versión anglosajona BAC) constituyen unos interesantes vectores con dos características fundamentales: 1) permiten insertos de entre

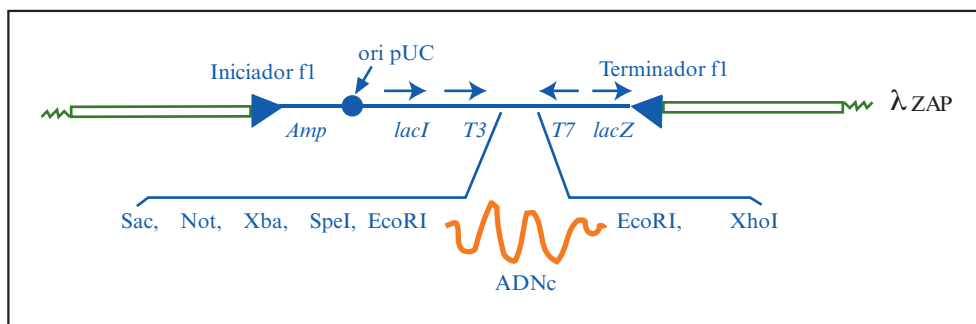


Figura 4.10. Fagómido λZAP. La región comprendida entre el iniciador f1 y el terminador f1 se copia como ADN de banda simple circular y se empaqueta y exporta como un fago M13 o f1. La colocación de los promotores de las ARN polimerasas T3 y T7 en orientaciones opuestas asegura la transcripción del ADN inserto *in vitro*.

100 y 300 kb y 2) los insertos permanecen estables sin reordenaciones o eliminación de material interno. Estos vectores pueden ser derivados bien de los factores de fertilidad F o del fago P1.

CAB contruidos a partir de factores de fertilidad F

Los vectores CAB derivados de los episomas F tienen los genes *OriS*, *repE*, *par A*, *par B*, un gen que confiere resistencia al cloranfenicol (*cmf^r*), la región *cos* del fago λ, que es reconocida por la función *ter* de la proteína A provocando su ruptura, y la región *lox P* del fago P1 que, a su vez, es reconocida de modo similar por la proteína Cre. Además, lleva dianas para las enzimas de restricción HindIII y BamHI para clonar insertos que quedarán flanqueados a derecha e izquierda por los promotores T7 y Sp6 en sentidos opuestos para asegurar la transcripción independientemente de la orientación del inserto (figura 4.11). Por último, incorpora una serie de dianas para enzimas de restricción fuera de los promotores para escindir el inserto cuando sea necesario. Los genes *par A* y *Par B* se denominan genes de partición porque facilitan la homogénea distribución de las copias del CAB después de la duplicación del cromosoma bacteriano y durante la división celular. También mantienen el número de copias en 1 o 2.

Los vectores con inserto transforman la bacteria hospedadora con dificultad debido a su gran

tamaño y se suelen introducir por electroporación. Este método consiste en aplicar descargas eléctricas de alto voltaje a las células competentes para producir poros nanométricos en sus membranas. El ADN entra en la célula bien a través de los poros o al redistribuirse los componentes de la membrana con el cierre de los poros. El método se lleva a cabo colocando la suspensión de las bacterias competentes y el ADN en una pequeña cubeta y aplicando descargas de 2,0 a 4,0 kV (kilovoltios), a 0 °C. La selección de los transformantes se realiza por resistencia a cloranfenicol, y la presencia de insertos se confirma por hibridación. Aproximadamente del 10 al 50% de los transformantes lleva inserto.

Los vectores CAB mantienen insertos estables de hasta 300 kb, debido a la presencia de sólo una copia o, como máximo, dos por bacteria.

CAB contruidos a partir del fago P1

El fago P1 es un fago atemperado, como λ, que puede seguir el ciclo lítico o el lisogénico. Posee 93 kb de tamaño y 112 proteínas, además de 5 genes que no se traducen.

Los CAB derivados de P1 mantienen, de forma estable, insertos de hasta 100 kb y contienen los elementos siguientes (figura 4.12): 1) una región de empaquetamiento del fago P1 denominada *pac* imprescindible para que se pueda empaquetar el fago dentro de la cabeza tanto *in vivo* como

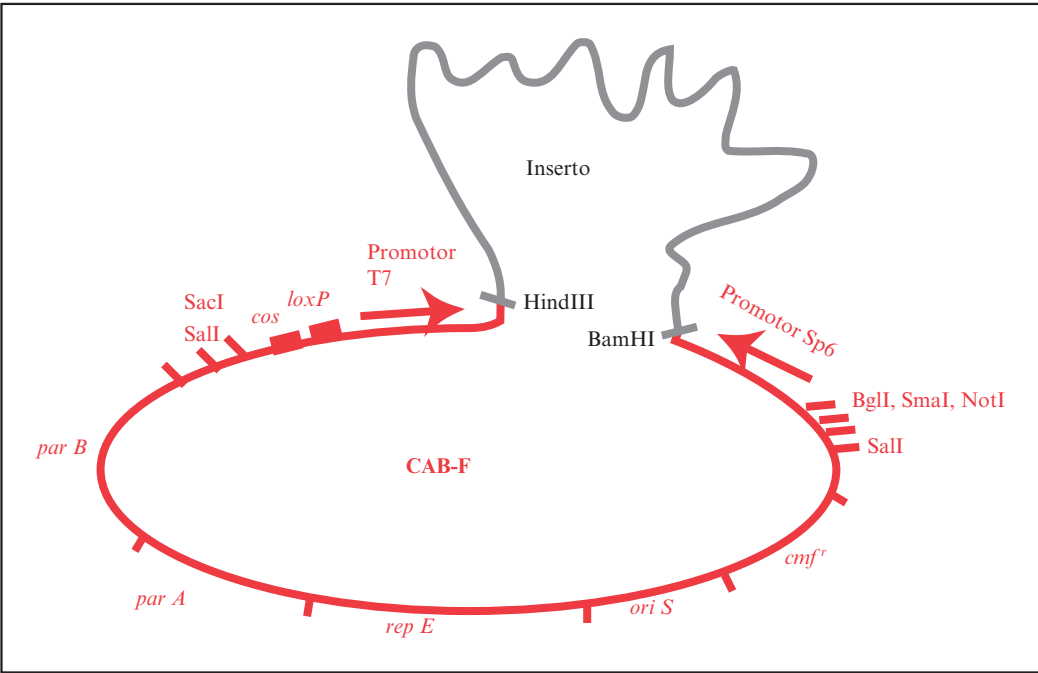


Figura 4.11. Componentes de un vector CAB.

in vitro; 2) dos sitios *loxP* que son lugares de reconocimiento en el fago para la recombinasa, producto del gen *cre*. Al actuar la recombinasa sobre el genoma lineal del fago, se produce la circulariza-

ción de éste y su posible crecimiento como plásmido de bajo número de copias; 3) un gen de selección de la transducción que confiere resistencia a kanamicina; 4) un gen de selección del inserto basado en

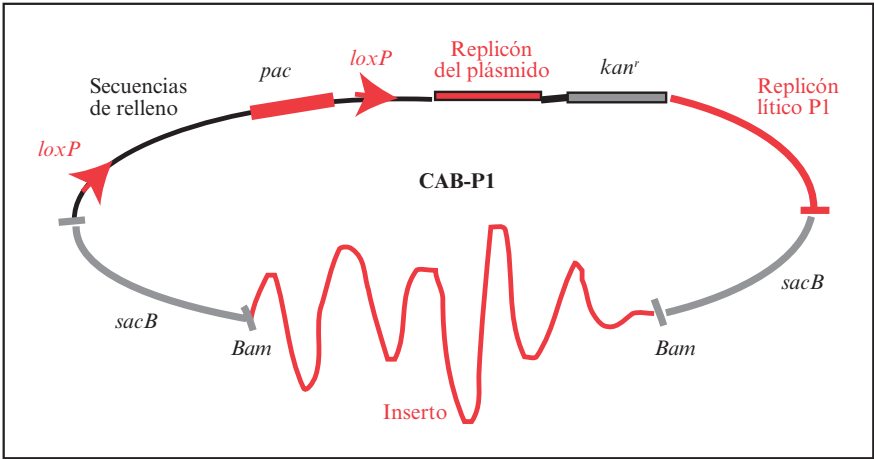


Figura 4.12. Mapa del vector CAB-P1.

las propiedades del gen *sacB* de *Bacillus amyloliquefaciens*. El producto de este gen es la enzima levansacarosa que cataliza la hidrólisis de sacarosa a un producto denominado levano, letal para *E. coli*. El gen tiene una única diana para BamHI y se le han adicionado tres más para las enzimas NotI, Sall y SfiI. También hay versiones con los dos promotores T7 y Sp6 para facilitar la transcripción del inserto. La adquisición del inserto inactiva *sacB* permitiendo el crecimiento de la bacteria en presencia de sacarosa. 5) El replicón P1 lítico permite el salto al ciclo lítico con la producción de un gran número de copias del recombinante.

Un proceso de clonación empezaría, por ejemplo, con la digestión del vector con BamHI que corta en medio del gen *sacB* y su unión con insertos de unas 100 kb obtenidos por digestión parcial con Sau3A de un ADN de mamífero. El empaquetamiento puede realizarse *in vitro*: la enzima pacasa corta en la región *pac*, facilita el empaquetamiento del ADN fágico en el interior de la cabeza, comenzando por la secuencia en que se encuentra *pac*, y vuelve a cortar a una distancia de 115 kb de la incisión inicial, es decir, en las secuencias de relleno. Las colas son entonces añadidas a las cabezas y el fago resultante puede infectar una célula de *E. coli*. En presencia de la recombinasa Cre, se producirá una recombinación específica de sitio que originará un plásmido circular.

4.3. TIPOS DE CLONACIÓN MOLECULAR Y APLICACIONES: GENOTECAS

Uno de los tipos de clonación molecular más frecuentes es la elaboración de *genotecas* (bibliotecas genómicas o mal llamadas librerías genómicas). En este tipo de clonación todo el genoma de un organismo se almacena en múltiples fragmentos individuales constituyendo una *genoteca de ADN* que es característica de la especie. Las genotecas se pueden almacenar durante largos períodos de tiempo y se recurre a ellas cada vez que interese rastrear y aislar una secuencia concreta perteneciente a ese organismo. Para el rastreo es necesario tener una sonda que nos permita «pescar» la se-

cuencia deseada, normalmente por hibridación. Un tipo de clonación relacionado son las *genotecas de ADNc*, referido al conjunto de recombinantes individuales obtenidos a partir de los distintos mensajeros de un tejido determinado o de un organismo completo. Los mensajeros se pasan a ADNcopia de doble hebra mediante la transcriptasa inversa y ADN polimerasa y luego se recombinan con un vector. De manera análoga a las genotecas de ADN, el rastreo y aislamiento de un ADNcopia concreto relativo a un mensajero, requiere una sonda que nos permita aislarlo de entre todos ellos. Las genotecas de ADNc representan el conjunto de genes activos de un tipo celular concreto, con la ventaja de que los ADNcopia no poseen intrones, y por tanto, su tamaño es significativamente inferior a los genes dentro del genoma.

4.3.1. Elaboración de una genoteca de ADN

Una genoteca es el conjunto de recombinantes individuales que abarca todo el genoma de un organismo. El tamaño de la genoteca dependerá del tamaño del genoma y del tamaño medio de los fragmentos individuales (insertos). Así pues, si para tener una probabilidad del 99% de haber incluido una secuencia concreta de levaduras en una genoteca de este organismo se necesitan unos 3.462 recombinantes con un tamaño de inserto medio de 20 kb, en humanos se requieren 690.819 clones individuales. El número de recombinantes necesarios para tener una probabilidad determinada de haber incluido una secuencia concreta se puede estimar según la fórmula que Clarke y Carbon dedujeron en 1976:

$$N = \frac{\ln(1 - P)}{\ln(1 - f)}$$

Siendo:

N = Número de recombinantes necesario.

P = Probabilidad deseada.

f = Fracción del genoma contenida en un solo recombinante.

Y, aplicando el cálculo al genoma humano, el número de recombinantes que debemos tener para aislar un gen con una probabilidad del 99%: Siendo:

Tamaño medio de inserto = 20 kb (2×10^4 pb).
Tamaño del genoma humano = 3×10^9 pb.

$$N = \frac{\ln(1 - 0,99)}{\ln\left[1 - \left(\frac{2 \times 10^4}{3 \times 10^9}\right)\right]} = 690.819$$

Uno de los requisitos es que el tamaño medio de inserto sea homogéneo. Digestiones totales con enzimas de restricción generarían fragmentos cuyo tamaño medio sería de unas 4 kb, excesivamente pequeño. Además, la digestión debe ser al azar para tener algunos clones que posean el gen completo y otros que se extiendan por alguno de los flancos.

Digestiones parciales, a tiempos muy cortos, con una enzima que reconozca cuatro nucleótidos, en vez de seis (Sau3A, por ejemplo), nos producirán fragmentos de tamaño grande y cortados prácticamente al azar.

Los pasos a seguir para la elaboración de una librería serían los que se muestran en la figura 4.13.

ADN genómico y de alto peso molecular se trata con la enzima Sau3A durante tiempos muy cortos para asegurar un tamaño medio de fragmento grande (se comprobará en gel de agarosa). Se prepara, a su vez, ADN del fago lambda (λ) para ser utilizado como vector de sustitución. El ADN del vector se trata con la enzima BamH 1, ya que ésta produce los mismos terminales cohesivos que Sau3A y se podrán unir ambos extremos por complementariedad de las secuencias. Se sustituye un fragmento del vector por uno de inserto, lo que nos permite recuperar el tamaño del genoma del fago original. El fago λ sólo empaqueta ADN de un cierto tamaño, con lo que quedarán excluidos aquellos que unan dos brazos fágicos sin inserto.

La enzima ligasa unirá covalentemente los extremos apareados por las dianas de restricción, y los fagos se empaquetarán *in vitro*.

El fago λ normalmente se duplica por el método del círculo rodante y empaqueta su ADN cortando en las regiones *cos* (figura 4.14). La región *cos*, al ser cortada por la función *ter* de la proteína A, origina unos terminales cohesivos de 12 pares de bases que constituirán los extremos del profago. Para el empaquetamiento *in vitro* de genotecas se utiliza frecuentemente un mutante del fago λ deletado en las secuencias *cos*. Este mutante se mantiene en forma lisogénica (como profago) en una cepa de *E. coli* C. Al inducir el ciclo lítico, se origina una acumulación intracelular de todos los componentes del empaquetamiento, incluyendo las pre-cabezas vacías. El paso siguiente está bloqueado. Lisados de este mutante se pueden añadir a ADN recombinante que posea secuencias *cos*. De este modo, sólo el ADN recombinado será empaquetado en el fago. Existe una gran variedad de genotecas que se pueden obtener comercialmente.

4.3.2. Elaboración de una genoteca de ADNc

Las genotecas de ADNc representan el conjunto de mensajeros de un tejido determinado o de un organismo completo en recombinantes individualizados.

El ADNcopia se sintetiza a partir del mensajero correspondiente y un cebador, por acción de la enzima transcriptasa inversa (véase el capítulo 3, apartado 3.1.3, figura 3.10).

Los recombinantes de ADNc tienen una utilidad especial, ya que, al proceder de los mensajeros correspondientes, carecen de intrones. Los seres procarióticos, por otra parte, carecen de intrones en sus mensajeros y por eso no son capaces de procesar los ARNm eucarióticos correctamente. Si el objetivo final de un experimento es expresar un gen eucariótico en un sistema procariótico, deberemos clonarlo a partir de su ADNc y no de su secuencia genómica, ya que ésta se encontrará probablemente interrumpida por varios intrones.

La mayoría de los genes eucarióticos aislados con fines de caracterización o expresión se han obtenido a partir de genotecas de ADNc. La elaboración de una genoteca de ADNc implica el aislamiento

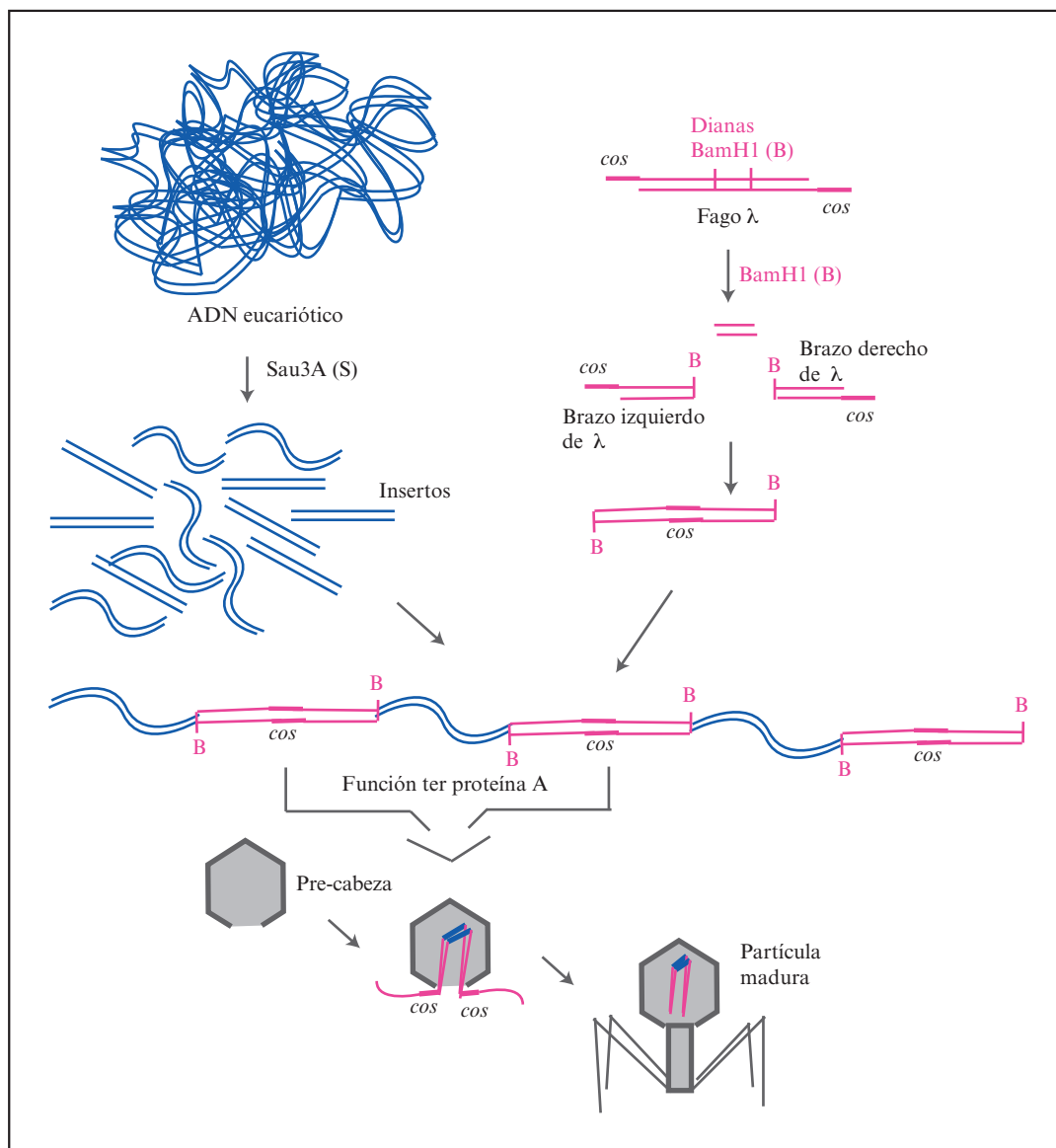


Figura 4.13. Esquema de la elaboración de una genoteca. El ADN de un bacteriófago se corta con la enzima de restricción Bam(B) para eliminar el fragmento interno y dejar dos brazos que se unirán por los terminales cohesivos *cos*. Estos vectores se acoplan con los insertos de ADN eucariótico, que a su vez han sido tratados con la enzima (isoesquizómera de Bam) Sau3A. La enzima ligasa se encargará de unir *in vitro* ambos fragmentos. La función ter (terminales) de la proteína A del fago λ corta en los sitios *cos* y origina fagos recombinantes que podrán ser empaquetados *in vitro* para originar partículas fágicas maduras e infectivas.

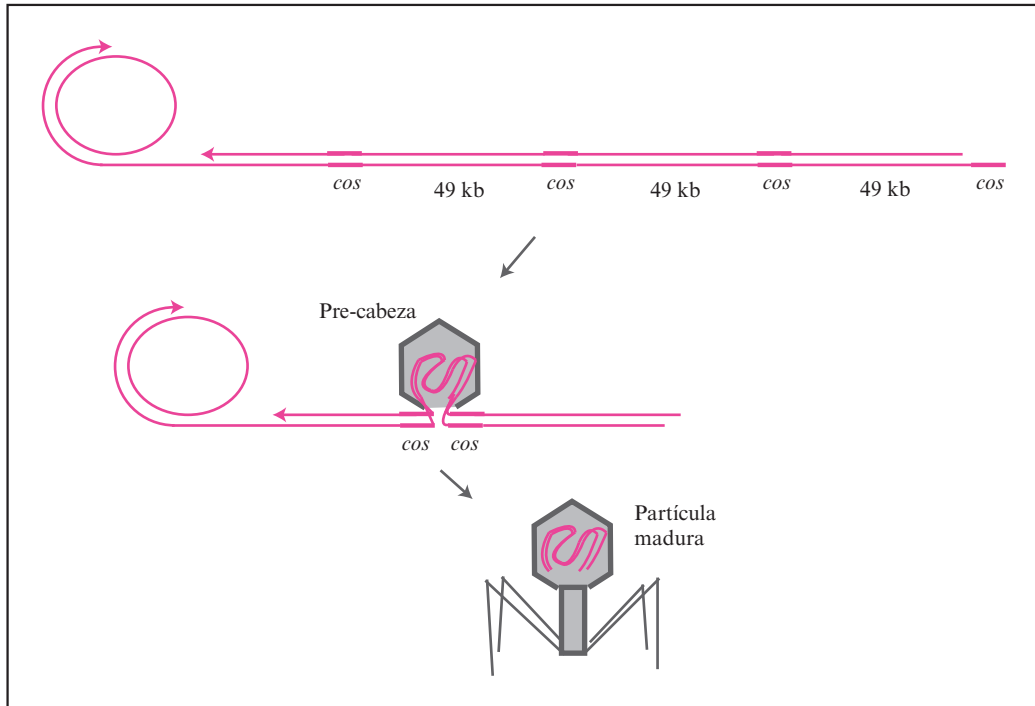


Figura 4.14. Replicación por el círculo rodante y empaquetamiento del fago λ.

to de la población total de mensajeros celulares y su transcripción revertida a ADN. Se suelen utilizar colas de oligo-dT como cebador; al finalizar la copia del ARNm la actividad RNAsa H débil de la transcriptasa inversa del virus Moloney de ratón, por ejemplo, deja restos del molde que sirven de cebadores para la ADN polimerasa I que sintetiza la segunda hebra. El paso siguiente es la adición de conectores sintéticos (dianas de restricción sintéticas) que se enlazan al ADNc con ligasa y se cortan con la enzima para dejar terminales protuberantes. El ADNc debe ser previamente tratado con una metilasa del mismo sistema de restricción para evitar cortes inoportunos en los ADNc. A partir de aquí se introducirán los fragmentos en un vector (derivado del fago λ, preferentemente) previamente tratado con la misma enzima de restricción y se empaquetarán las partículas recombinantes *in vitro* (figura 4.15).

Si el mensajero que deseamos clonar es abundante, no precisaremos de su enriquecimiento. No

obstante, en la mayoría de los casos es conveniente partir de un sistema enriquecido natural o artificialmente. Un enriquecimiento natural lo presentan las células de tejidos especializados o tumores concretos; por ejemplo, las células del páncreas o los insulinosomas representarían un sistema de enriquecimiento para aislar el gen que codifica la insulina. Alternativamente, una reacción en cadena de la polimerasa (más adelante, en este capítulo) podría amplificar artificialmente alguno de los mensajeros, previo paso a ADNc (RT-PCR), facilitando así su posterior clonaje.

Esta técnica de clonación molecular a través del ADNc tiene algunas limitaciones. Por una parte, la transcriptasa inversa carece de función correctora de errores y puede incluir algunos fallos. Por otra, cuando los mensajeros son largos, la síntesis, a menudo, no incluye el extremo 5', ya que la polimerasa en ocasiones se suelta del mensajero y se dobla en forma de gancho antes de lle-

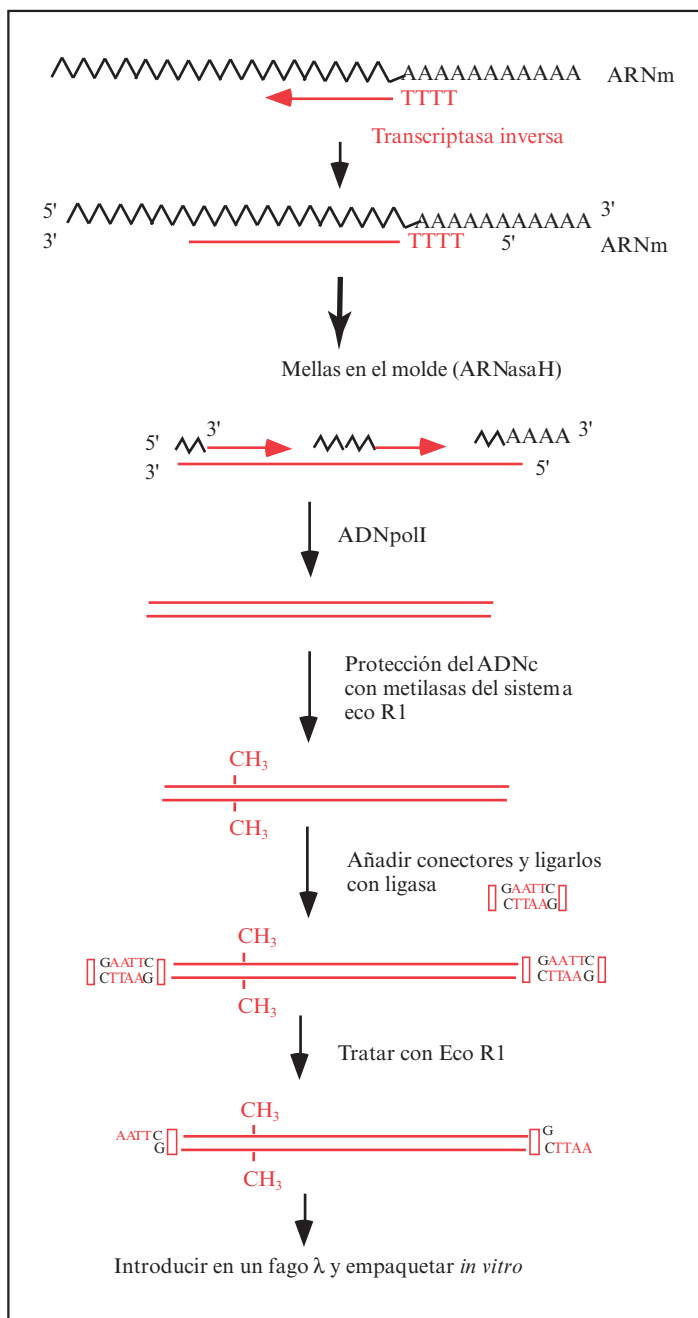


Figura 4.15. Síntesis de ADNc por acción de la transcriptasa inversa utilizando como cebador de la primera hebra oligo-dT. A continuación se utilizan los restos de la acción de la ARNasa H para iniciar la segunda hebra. El ADNc se protege con una metilasa para el sistema Eco R1 antes de añadir conectores Eco R1 y tratar con esta enzima de restricción con el fin de generar extremos protuberantes.

gar a su extremo; por ello las librerías de ADNc están normalmente enriquecidas en secuencias 3'; además, los mensajeros que carezcan de poliA (como los de las histonas) no estarán representados.

Para evitar estas desviaciones se puede utilizar como cebadores una mezcla de hexa-deoxinucleótidos sintéticos con secuencias al azar. Éstos hibridarán donde encuentren complementariedad con los mensajeros e iniciarán la síntesis. Este método constituye la variante de síntesis por iniciación al azar.

La obtención de genotecas de recombinantes específicos de un tejido determinado puede realizarse por un método de sustracción de la siguiente manera: obtendremos ARNm total de un tejido específico, músculo, por ejemplo, y de otro en el que predominen las funciones de mantenimiento, como puede ser una línea celular cultivada *in vitro*. Los ARNm de músculo se pasan a ADNc de hebra sencilla y se hibridan con un exceso de ARNm de la línea celular. El tratamiento con la ARNasa A nos degradará los mensajeros que no hayan hibridado. Quedan libres exclusivamente los ADNc de banda simple músculo-específicos que ahora se podrán pasar a ADNc de doble hebra. Para este último paso se pueden sintetizar colas de oligo-dC por acción de la transferasa terminal y utilizar un oligo-dG como iniciador de la segunda hebra. Los ADNc se unen a vectores del modo ya descrito.

4.3.3. Identificación del recombinante deseado de una genoteca

Cuando un fago infecta una bacteria libera un gran número de partículas fágicas que, a su vez, infectan las bacterias de alrededor. El resultado es una zona clara en un césped (turbio) de bacterias que se denomina calva.

La identificación de una calva de lisis fágica cuyos fagos lleven un gen de interés de entre todas las calvas de lisis crecidas sobre un césped de bacterias, o la de una colonia bacteriana de entre todas las crecidas en una placa Petri, es una práctica común en ingeniería genética.

Los métodos de hibridación son los más comunes para la selección de un recombinante en

particular de entre todos los obtenidos en una genoteca. La utilización de estos métodos precisa del empleo de una sonda, esto es, un ADN o ARN complementario al gen que se desea aislar y que, por alguno de los métodos ya descritos (véase el capítulo 3), podemos marcar radiactivamente, con fluorescencia o biotina. Muchas veces no es posible tener una sonda pura, sino que se parte de una mezcla enriquecida para la secuencia que buscamos, con lo cual el proceso de aislamiento e identificación del gen deseado no es tan directo como en el ejemplo que describiremos a continuación.

Partiendo, por ejemplo, de una librería genómica, el primer paso será crecer las colonias o calvas (dependiendo de que el recombinante sea plásmido o fago) a una gran densidad (casi confluentes) para poder tener el mayor número posible de unidades por placa Petri. A continuación se transfieren las colonias/calvas a papel de nailon, colocando simplemente dicho papel sobre la placa Petri. Una parte de las colonias o calvas permanecerá en la placa Petri y constituirá la placa maestra, a partir de la cual se aislará la colonia/fago de interés (figura 4.16). El filtro con las colonias/calvas es tratado con álcali para lisar las bacterias/fagos y desnaturalizar el ADN. Seguidamente se incubaba con proteinasa K para degradar proteínas y dejar sólo el ADN, que se fija covalentemente al nailon por radiación ultravioleta. Por otro lado, se desnaturaliza la sonda, si es de doble banda, y se hibrida a los filtros en condiciones de sal y temperatura que favorezcan la formación y estabilidad de secuencias complementarias. El resultado de esta hibridación se detecta por autorradiografía. La colonia cuyo ADN da lugar a un positivo en la autorradiografía puede ser seleccionada en la placa Petri maestra de partida (figura 4.16). Este método es idóneo para identificar cualquier secuencia o gen del que se disponga de una sonda. La obtención de la sonda es, pues, el factor limitante.

4.3.4. Aplicaciones de las genotecas de ADN: paseo cromosómico

La técnica del paseo cromosómico consiste en caminar a derecha e izquierda de una secuencia

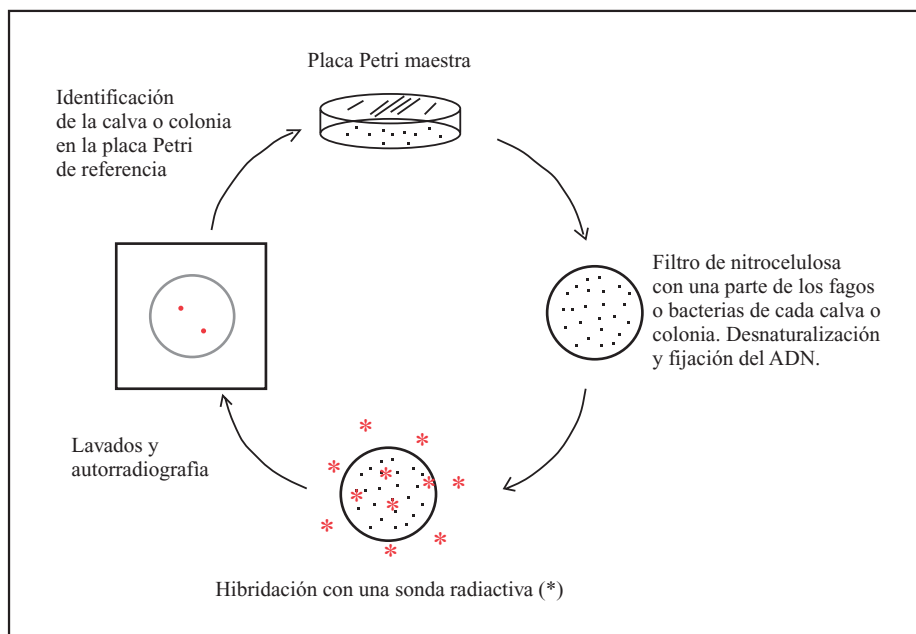


Figura 4.16. Identificación de secuencias complementarias a una sonda específica marcada radiactivamente, por hibridación en colonias o calvas. Muchos de los fagos que forman una calva de lisis o de las bacterias que forman las colonias en una placa Petri se pasan a un filtro de nitrocelulosa por simple contacto; el filtro se separa de la placa y el ADN se desnatura con álcali y se fija por calor o rayos ultravioleta. El filtro se hibrida con la sonda marcada y las colonias o calvas positivas se identifican por autorradiografía. Por último, se aísla la colonia o calva que ha hibridado con la sonda, por identificación en la placa Petri de partida que aún mantiene parte de las bacterias o fagos que formaban las colonias o calvas originales.

conocida por solapamiento de secuencias contiguas. Para realizarlo se suele recurrir al marcaje de los extremos de un fragmento original y subclonaje en varios elementos de menor tamaño; los pequeños fragmentos de los extremos que se encuentran marcados se utilizarán como sonda para rastrear, por hibridación, una genoteca del organismo por el que estemos paseando. Los clones positivos contendrán una parte de la sonda y las secuencias colindantes. Con cada uno de estos clones se repite la estrategia descrita, es decir, se marcan en los extremos, se subclonan y se utilizan los pequeños fragmentos para rastrear secuencias solapantes en la genoteca (figura 4.17).

Este procedimiento nos permite caminar simultáneamente en ambas direcciones a partir del inserto original. Si pretendemos llegar a un lugar

concreto a una cierta distancia del clon original deberemos utilizar, siempre que sea posible, métodos genéticos (mutaciones) o citológicos (hibridación *in situ* a los cromosomas) para confirmar el final de trayecto.

En humanos esta técnica se ha empleado para secuenciar el genoma completo y para aislar varios genes, entre ellos el gen de la fibrosis quística. Esta enfermedad depende de un gen recesivo localizado en el cromosoma 7, y aproximadamente uno de cada veinte individuos de raza caucásica es portador de una alteración. El producto del gen es una proteína transmembrana transportadora de iones cloro cuyo fallo provoca un peligroso desequilibrio entre la concentración de iones dentro y fuera de la célula, acompañado de un endurecimiento de las mucosas, especialmente la pul-

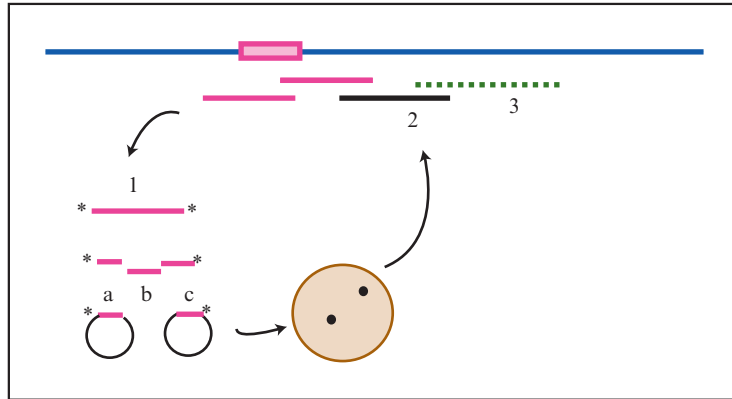


Figura 4.17. Estrategia de paseo cromosómico por solapamiento de secuencias. Marcaje de los extremos de un inserto, subclonaje y rastreo en una genoteca con los fragmentos extremos. Obtención de clones parcialmente solapantes con las sondas y repetición de la estrategia indefinidamente.

monar e intestinal. La localización final del gen se determinó mediante un paseo cromosómico de unas 500 kilobases a partir de marcadores genéticos ligados a la patología y, por tanto, próximos al gen. Se utilizaron genotecas de fagos y cósmidos y se avanzó por solapamiento de clones contiguos. El orden de los fragmentos se verifica por secuenciación y bioinformática. A partir de aquí se aplican una serie de criterios para determinar el lugar preciso donde se ubica el gen. En este caso fueron cinco: 1) identificación de islas CpG que marcan regiones metilables presentes en regiones 5' de muchos genes de expresión ubicua de vertebrados; 2) identificación de pautas de lectura abiertas; 3) hibridación cruzada con el gen de otras especies; 4) comparación entre el ARNm de individuos normales y afectados, y 5) aislamiento de los ADNc correspondientes.

El único inconveniente serio de este procedimiento es encontrar una secuencia repetida en algún lugar del trayecto. Si esto llegara a suceder, habría que recurrir a métodos alternativos como sería el de obtener clones de gran tamaño para saltar la secuencia repetida. En el caso del gen de la fibrosis quística, el salto se ha practicado con mucha frecuencia para evitar este tipo de secuencias repetidas o determinadas secuencias que por razones no claras resultaron ser inclonables.

La estrategia del salto

Para el procedimiento de saltar secuencias se suelen utilizar enzimas de restricción de tipo II que poseen como secuencia de reconocimiento un octanucleótido en vez de exanucleótido. Estas secuencias serán mucho más infrecuentes en el genoma, con lo que los fragmentos que originen serán de mayor tamaño. Ejemplos de estas enzimas son la NotI (GCGGCCGC) y SfiI (GGCCNNN-NNGGCC). Inicialmente se obtienen fragmentos de gran tamaño del ADN del organismo por el que paseamos, por ejemplo, por digestión parcial con la enzima Not I. Cada uno de estos fragmentos enormes (200 kb, por ejemplo) se recombina con un plásmido etiqueta que posea un gen ARNt supresor y una diana Not I (figura 4.18). El ARNt supresor lleva una mutación que cambia la especificidad de reconocimiento del anticodón de manera que suprime mutaciones de terminación de cadena al incorporar un aminoácido (véase el capítulo 5). Los plásmidos híbridos resultantes se digieren con cualquier enzima de corte frecuente, pero que no posea dianas en el plásmido etiqueta (Sau 3 A, por ejemplo). Así se obtendrán muchos plásmidos pequeños que sólo poseen los extremos de los recombinantes originales, ya que habrán perdido las secuencias intermedias en pequeños

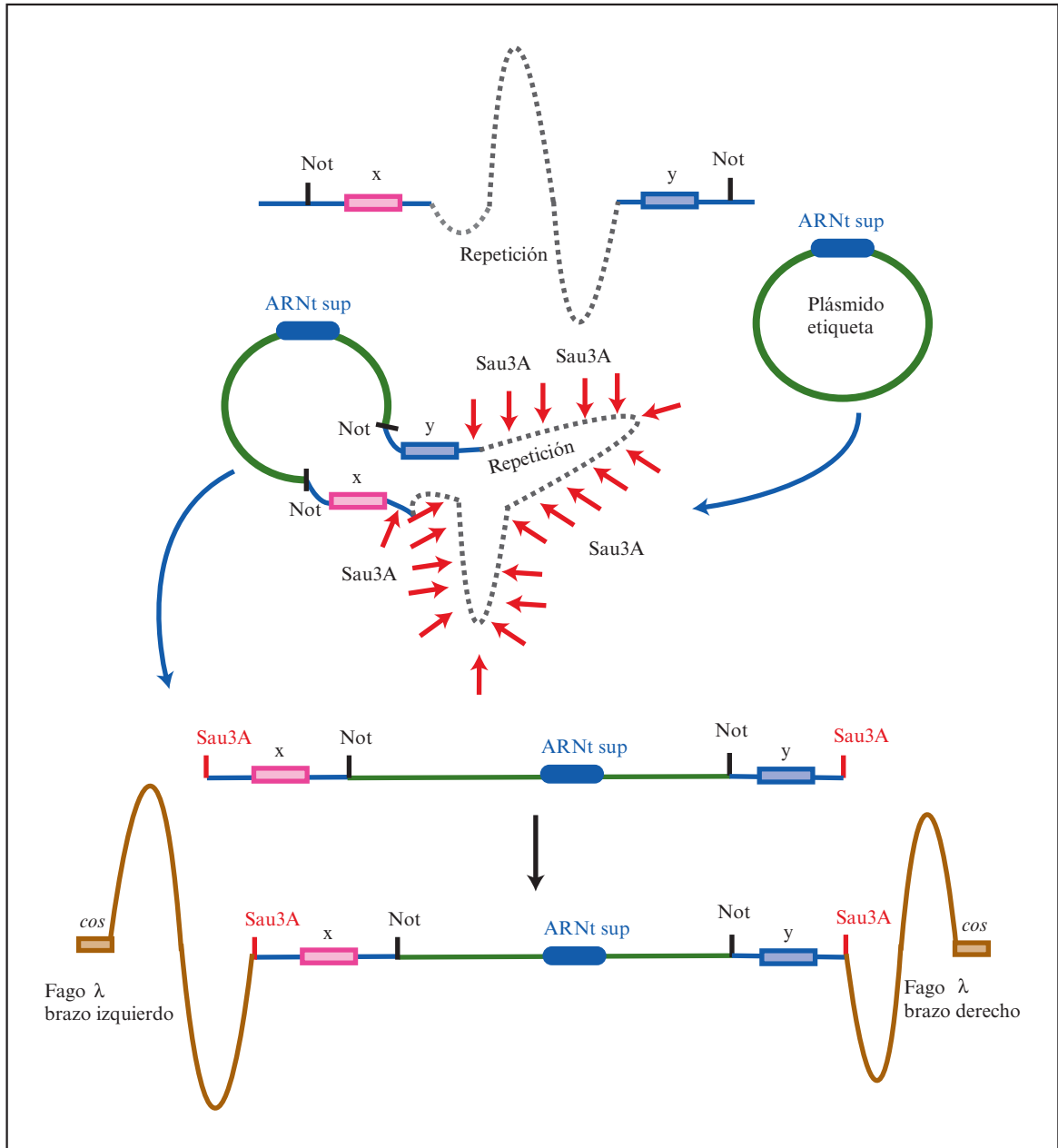


Figura 4.18. Estrategia de salto para eliminar secuencias repetidas en el pase cromosómico. El ADN del organismo por el que se pasea se somete a digestión parcial con la enzima de restricción Not. El plásmido etiqueta también. Una vez constituidos los híbridos, se digieren con otra enzima de corte frecuente para eliminar gran parte de los insertos (Sau3A). Los extremos de los insertos unidos al plásmido constituyen los nuevos insertos en una genoteca, utilizando el fago λ como vector y seleccionando los recombinantes en una bacteria con una mutación importante suprimible por el ARNT aportado por el plásmido etiquetado.

fragmentos. De entre todos estos plásmidos, suponemos que habrá al menos uno que, conteniendo la secuencia única anterior a la repetida (x, en la figura 4.18) que obra en nuestro poder, también posea otra posterior a la secuencia repetida (y, en la figura 4.18). El paso siguiente es elaborar una pequeña genoteca con el fago λ como vector y los distintos plásmidos como insertos. Estos recombinantes han de propagarse en una bacteria con una mutación de terminación de cadena en un gen importante, suprimible por el ARNt que aporta el plásmido etiqueta. Ahora deberemos usar la sonda «x» para obtener el recombinante que posea «y». Así habremos conseguido saltar todas las secuencias intermedias del clon grande inicial.

4.3.5. Elección del vector más idóneo para el tipo de clonación

Hemos descrito numerosos vectores con diversos orígenes y aplicaciones. En todo experimento de ingeniería genética es muy importante la elección del vector más idóneo para el tipo de clonación que se desee, dependiendo del fin último experimental.

En general, los plásmidos están indicados para el subclonaje y la caracterización básica de secuencias, ya que crecen bien y se manejan con facilidad. También podrían utilizarse para poner a punto sistemas de expresión de un gen concreto.

Los bacteriófagos se utilizan preferentemente para la elaboración de genotecas y sistemas de expresión eficientes. Las genotecas pueden amplificarse y almacenarse indefinidamente con perfecta viabilidad y en volúmenes muy reducidos, mientras que las genotecas en colonias de bacterias sufren pérdidas de viabilidad inaceptables y son muy difíciles de amplificar, siendo preferible en la mayoría de los casos volverlas a elaborar. Si el objetivo es una expresión controlable con gran precisión, también utilizaremos bacteriófagos con preferencia a plásmidos.

Para secuenciación y obtención de ADN de banda simple, los derivados del M13 son muy valiosos.

Los cósmidos se utilizarán en la elaboración de genotecas de genomas grandes, pero, una vez aislado el clon deseado, convendrá caracterizarlo y subclonar sus fragmentos en un plásmido convencional y de fácil manejo. La inestabilidad que acompaña a la clonación en cósmidos y posibilidades de reordenamientos entre diversas partes del inserto se supera con los cromosomas artificiales de bacterias (CAB) que, además, admiten insertos bastante más grandes que los clonables en cósmidos. Así pues, podríamos elegir un CAB para paseos cromosómicos en el genoma humano y, en general, para elaborar genotecas de genomas de gran tamaño.

La planificación de un experimento concreto conlleva un estudio exhaustivo de los vectores disponibles y de la elección del más conveniente en cada caso en particular.

4.3.6. Expresión de los productos clonados en bacterias

Una de las grandes aplicaciones de la ingeniería genética y la biotecnología es la producción a gran escala de proteínas con interés comercial y farmacológico. Para tal fin se necesitan plásmidos de alto número de copias, promotores muy potentes y condiciones que favorezcan al máximo la expresión de los genes clonados. Los factores que han de ser tenidos en cuenta para una máxima expresión son:

- Un promotor potente.
- Presencia de secuencias de iniciación de la traducción y unión al ribosoma.
- Implantación de codones de uso más frecuente en el sistema bacteriano.
- Plegamiento y estabilidad de la proteína: proteínas fusión.
- Otras modificaciones como quizá eliminación de la metionina inicial, exportación al espacio periplasmático, etc.

Entre los promotores potentes y regulables que se han utilizado para expresar genes eucarióticos en *E. coli* están:

- El promotor P_I de lambda, que puede ser regulado por un represor sensible a la temperatura (CI ts). Este represor reprime P_I a temperaturas bajas (37 °C), pero es inactivo a temperaturas más elevadas (45 °C).
- Otro promotor muy eficiente es el denominado «*tac*». Este promotor híbrido entre los de «*lac*» y «*trp*» es unas 11 veces más potente que cualquiera de los dos y regulable por el represor de *lac*. Si la proteína que ha de expresarse es tóxica para *E. coli*, se utilizará, de entrada, una cepa que produzca mucho represor, y cuando se quiera inducir el producto de interés se añade el inductor IPTG (isopropil-tio- β -D galactósido).
- Un sistema de expresión aún mejor es el uso del conjunto promotor/ARN polimerasa del fago T7. Esta ARN polimerasa reconoce exclusivamente a su promotor, con lo que se consiguen niveles de expresión específicos muy altos para los genes situados inmediatamente después del promotor. La regulación de la expresión se controla por la presencia de la polimerasa en la bacteria; ésta puede introducirse en otro vector de expresión compatible (M13 o λ). El gen de la ARN polimerasa de T7 también puede mantenerse reprimido a

continuación del represor Lac; la presencia del inductor IPTG disparará su expresión, y con ello la del gen de interés.

La presencia de una región de iniciación de la traducción y unión al ribosoma es otro de los factores requeridos (figura 4.19). En *E. coli*, el promotor tiene dos secuencias consenso, una a -35 y otra a -10 del codón de iniciación (ATG); la secuencia de unión al ribosoma (SUR) incluye una pequeña región de 3-9 nucleótidos localizada entre -3 y -11 nucleótidos antes del codón de iniciación y denominada Shine & Dalgarno (SD). La secuencia SD es complementaria al extremo 3' del ARN ribosomal 16S de *E. coli*; el apareamiento de estas secuencias favorece la unión del ribosoma. La mayoría de los genes eucarióticos requieren su inserción en un vector que incluya un codón de iniciación (ATG) a la distancia apropiada de la secuencia SD. El segundo codón del gen eucariótico debe ser colocado a continuación del ATG, lo que puede requerir manipulaciones adicionales en la preparación del segmento a insertar. Al final de la secuencia codificante, pero sin gran precisión, se coloca un terminador de la transcripción, secuencia formadora de una estructura secundaria de tallo y lazo que aborta toda transcripción independiente del factor de terminación rho.

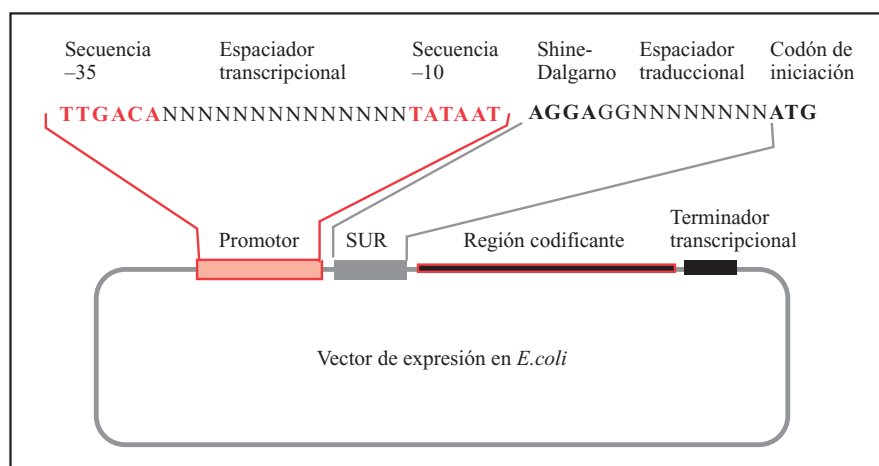


Figura 4.19. Esquema que representa los elementos de control que deben estar presentes en un vector de expresión en *E. coli*.

Hay una gran variedad de vectores de expresión que se pueden obtener comercialmente, un ejemplo de uno de ellos se representa en la figura 4.20 y posee los elementos siguientes: el promotor de la polimerasa del fago T7, lugar donde se incorpora la ARN polimerasa de T7 si está presente; el operador del operón de la lactosa, lugar donde se une el represor de *lac* impidiendo el paso de la ARN polimerasa del fago T7. El gen que codifica por el propio represor de Lac (β -galactosidasa) también está presente (*lac cI*) con un promotor constitutivo en una orientación opuesta. En presencia del inductor IPTG (isopropilotio- β -D-galactósido) el represor queda secuestrado dejando libre el operador a la ARN polimerasa de T7. El gen de interés que queremos que se exprese, con las regiones SUR y SD, se introduce en este vector por recombinación específica de sitio entre *attR1* y *attR2* de modo similar a la recombinación específica de sitio que se ha descrito para la integración del fago λ (véase capítulo 2, apartado 2.2.1 y figura 2.11). Por este procedimiento se sustituyen los genes de resistencia al cloranfenicol y de selección negativa *ccdB* por nuestro gen de interés. El gen *ccdB* ataca la enzima ADN-girasa siendo letal para la bacteria ex-

cepto en cepas con mutaciones en este gen que confieren resistencia y que son las que se deben utilizar aquí. El gen de interés deberá estar flanqueado por regiones *attL* que al recombinar con las *attR* crearán las *attB*; aquí la recombinación está facilitada por una recombinasa (clonasa). El agente comercial facilita los vectores y la clonasa, así como cepas de bacterias resistentes al gen tóxico *ccdB* por poseer un gen mutado de la enzima ADN-girasa. El gen de interés puede llevar una región de parada de la transcripción o puede fusionarse en fase al epítipo V5 que contiene 14 aminoácidos de la proteína V del paramixovirus SV5 que es inmunogénico; así anticuerpos contra el epítipo V5 nos ayudarán a purificar la proteína de interés. También se pueden usar anticuerpos contra la cola de 6 histidinas del terminal carboxilo de la proteína. Estos anticuerpos suelen estar ligados a la peroxidasa de rábano o a la fosfatasa alcalina. En el primer caso, en presencia del sustrato adecuado (benzidina), producirá una coloración marrón allí donde haya habido identificación positiva; en el segundo caso la identificación final se realiza con un segundo anticuerpo marcado radiactivamente o con fluorescencia y dirigido contra el primero.

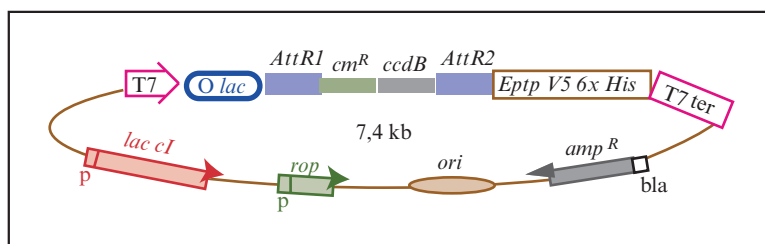


Figura 4.20. Esquema que representa un vector de expresión en *E. coli*. Se observa el promotor de la ARNpol del bacteriófago T7, seguido del operador de *lac*, a continuación el primer sitio de recombinación específica de λ (*attR1*), le sigue el gen de resistencia a cloranfenicol (*cm^R*) y el gen *ccdB*; a continuación el segundo sitio de recombinación específica de λ (*attR2*), el epítipo V5 y la región de polihistidina; a continuación la región de terminación para la polimerasa de T7, el gen de resistencia a ampicilina (*amp^R*) con su propio promotor (*bla* de β -lactamasa) y el origen de replicación de pUC. En orientación opuesta a todo lo anterior, encontramos el represor de *lac* (*lac cI*) con su propio promotor (*p*) y *rop* con el suyo (*p*).

En general, todos los métodos inmunológicos se basan en la elaboración de anticuerpos específicos contra el producto del gen clonado o, en

fusiones génicas, contra la proteína inmunogénicamente activa y fusionada al producto que deseamos aislar.

Hay tres tipos de anticuerpos que pueden reaccionar específicamente con los productos clonados:

- Anticuerpos policlonales que reaccionan con la proteína extraña, cualquiera que sea su conformación. Se preparan inmunizando animales con proteínas parcialmente desnaturalizadas, de modo que conserven partes intactas.
- Anticuerpos monoclonales que reaccionan sólo con epítomos específicos a partes intactas de la proteína.
- Anticuerpos que reaccionan sólo con la proteína desnaturalizada. Se preparan inmunizando animales con proteína totalmente desnaturalizada, y pueden ser policlonales o monoclonales.

La inmunoprecipitación de proteínas expresadas en bacterias se realiza en geles de poliacrilamida-SDS en extractos de proteína total. La identificación de estas nuevas bandas de geles de poliacrilamida se realiza por la técnica denominada Western (del oeste), por homología con la técnica de transferencia de ADN desarrollada por E. Southern (el significado del apellido es «del sur»), del que derivaron los nombres «Northern» (del norte), para la transferencia de ARN, y «Western», para las proteínas. Se separan las proteínas por su movilidad relativa (tamaño) en un gel de poliacrilamida y se transfieren a un papel de nitrocelulosa o nailon, al que se unen fuertemente. Las proteínas en este estado son accesibles al tratamiento y unión con anticuerpos específicos.

Pero no todo son ventajas en *E. coli*: hay proteínas eucarióticas que son tóxicas para bacterias; por el contrario, una expresión eficiente a nivel de traducción no es siempre fácil de conseguir, ya que algunos componentes del complejo aparato biosintético procariótico no funcionan eficazmente con un mensajero eucariótico; los codones utilizados por el gen transferido pueden no ser óptimos en el sistema bacteriano; las bacterias no son capaces de llevar a cabo alteraciones específicas de maduración de ARN como ajustes, imprescindibles para una correcta expresión; tampoco pue-

den realizar algunas modificaciones que tienen lugar durante o después de la traducción; el medio celular para el plegamiento de la proteína recién sintetizada puede no ser adecuado, etc. Todas estas consideraciones tienen consecuencias que afectan a la función del gen de interés. Algunos de estos inconvenientes pueden ser solventados con mayor o menor eficiencia: codones poco utilizados en *E. coli* pueden ser sustituidos por otros más comunes; los ajustes se evitan clonando los ADNc de los genes correspondientes. Pero modificaciones específicas, como fosforilaciones, glicosilaciones, etc., son complicadas de realizar para *E. coli*.

El código genético está degenerado en el sentido de que hay varios codones para el mismo aminoácido, pero distintos organismos utilizan unos con preferencia sobre otros (véase capítulo 1, apartado 1.5.2, y tabla 1.1). La célula carecerá de ARNt de reconocimiento para codones de utilización infrecuente. La implantación de codones de uso más frecuente en la bacteria se suele realizar por manipulación en el gen eucariótico con procedimientos, por ejemplo, de mutagénesis dirigida.

En bacterias el grupo formilo de la metionina inicial lo quita una deformilasa, quedando una metionina normal en el extremo de todas las proteínas. Si la metionina es el aminoácido natural del extremo NH₂ de la proteína que se esté sintetizando, no se necesita ninguna modificación adicional; pero, si no es así, la metionina inicial será eliminada por una aminopeptidasa, lo que ocurre en aproximadamente el 50% de las proteínas de *E. coli*. La permanencia o escisión de la metionina inicial depende del segundo aminoácido: cuando éste es lisina la metionina permanece, pero si el segundo aminoácido es alanina, serina o treonina, la metionina es eliminada de modo rápido y secuencial cuando se han sintetizado unos 15 aminoácidos de la proteína.

Si el plegamiento es incorrecto o la proteína es atacada por las proteasas internas, puede ser conveniente la producción como proteína fusión. Fusiones con el gen de la β -galactosidasa aumentan considerablemente la estabilidad de las proteínas fusionadas y permiten el aislamiento de éstas por

métodos inmunológicos dirigidos contra la β -galactosidasa. La separación de las dos proteínas unidas se puede llevar a cabo por acción de cianuro de bromo, que corta detrás de residuos de metionina siempre que no existan metioninas internas en la proteína de interés, o por la acción de proteasas específicas si existen o se añaden los aminoácidos diana entre la parte procariótica y la eucariótica.

La secreción del producto traducido es una vía alternativa para mejorar la estabilidad y eliminar la metionina inicial. El péptido señal constituye el extremo amino de las proteínas que se exportan y es eliminado por una peptidasa en el espacio periplasmático entre las dos membranas, interior y exterior, de *E. coli*. La secuencia del péptido señal puede ser introducida en el extremo 5' del gen eucariótico. Las proteínas así confeccionadas pasarán al espacio periplasmático donde no hay proteasas y su metionina inicial será eliminada al cortar el péptido señal; además, algunas proteínas que son inactivas por un plegamiento defectuoso intracelular, cuando se excretan, pueden sufrir un replegamiento que les devuelve la función. No obstante, hay muchas proteínas eucarióticas que, a pesar de poseer el péptido señal, no son procesadas correctamente por la peptidasa, o se realiza muy poco eficazmente. Uno de los péptidos señal que más se utilizan es el de la fosfatasa alcalina de *E. coli*.

Cuando los niveles de expresión de una proteína foránea en *E. coli* son elevados, con frecuencia se forman aglomeraciones insolubles denominadas cuerpos de inclusión. Estos cuerpos de inclusión pueden ser aislados como tales y solubilizados para permitir un replegamiento correcto de la proteína. Cada inclusión proteica puede necesitar un método de solubilización distinto y en ningún caso se recupera la actividad del cien por cien de las moléculas.

Existe un método interesante para asignar proteínas a clones concretos en ausencia de sondas específicas, es el método de la *traducción inhibida por el híbrido*, que nos permite correlacionar clones concretos con sus productos. Se basa en el hecho de que un mensajero no podrá traducirse si está secuestrado y formando una molécula híbi-

da con su ADN complementario. La comparación entre los productos de una traducción *in vitro* con mensajeros totales y los mismos mensajeros después de la hibridación con un ADN de un clon concreto nos identificará el producto de dicho clon (figura 4.21). Los experimentos de traducción *in vitro* se llevan a cabo, por ejemplo, en presencia de metionina marcada con un isótopo del azufre (^{35}S), analizándose los productos en geles de poliacrilamida y visualizándolos por autorradiografía.

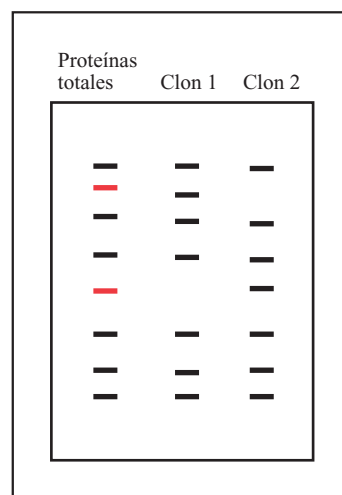


Figura. 4.21. Método de la traducción inhibida por el híbrido para identificar productos de clones concretos.

La caracterización del producto clonado es muy importante en todos aquellos casos en los que la fidelidad entre la proteína original y la producida por ingeniería genética haya de ser máxima. Todas las proteínas para el uso farmacéutico están sometidas a rigurosos controles periódicos en los que se analiza exhaustivamente la similitud entre el producto obtenido *in vitro* y el natural.

La caracterización primaria incluye: tamaño, carga, punto isoelectrico, composición de aminoácidos e hidrofobicidad. Adicionalmente se analiza la conformación estructural y la actividad enzimática (si procede). Especial atención se dedica a los extremos amino (por la presencia de la

metionina inicial) y carboxilo (si está intacto, no ha habido degradación por acción de carboxipeptidasas), así como a glicosilaciones, fosforilaciones, puentes disulfuro y cualquier modificación postraduccional que se requiera en la proteína concreta.

El grado de pureza es otro factor importante. No nos servirá de mucho una proteína muy próxima a la natural, pero contaminada con algún compuesto o péptido bacteriano inmunogénicamente activo para el hombre. Convendrá asegurarse utilizando técnicas fisicoquímicas e inmunológicas. La contaminación por virus o ácidos nucleicos también ha de ser rutinariamente vigilada.

Pruebas en animales de experimentación son el requisito final antes de su utilización generalizada, por ejemplo, en terapia humana.

4.4. AMPLIFICACIÓN POR LA REACCIÓN EN CADENA DE LA POLIMERASA (PCR)

La clonación, que hemos visto en los apartados anteriores, y la reacción en cadena de la polimerasa (PCR), que veremos a continuación, son dos métodos distintos de amplificación de secuencias.

La PCR es posterior en el tiempo a la clonación molecular y tiene algunas ventajas respecto a la tecnología de ADN recombinante: 1) no se necesita purificar el ADN a amplificar pudiendo partir de extractos crudos de tejido, lisados celulares o sangre de un mamífero; 2) es más rápido y versátil.

La PCR es un método enzimático que permite sintetizar numerosas copias de un fragmento concreto de ADN. Utiliza como cebadores dos oligonucleótidos (oligos, de 20 a 30 pb) complementarios a lo que serán los extremos del producto y colocados con sus terminales 3'OH enfrentados. Lo que se amplifica es el ADN comprendido entre ambos oligos. La sucesión de una serie de ciclos en los que tienen lugar 3 fases sucesivas: la desnaturalización del molde, la hibridación con los cebadores y la extensión de la síntesis por acción de la ADN polimerasa, origina una acumulación exponencial de fragmentos específicos cuyos extremos estarán definidos por los extremos 5' de los

cebadores (figura 4.22). Al poder ser utilizados los productos de un ciclo como moldes del ciclo siguiente, el número de copias de ADN producidas crece exponencialmente en vez de linealmente. Así pues, 30 ciclos de PCR producen una amplificación de un millón de veces aproximadamente. En la práctica el rendimiento real es algo inferior al teórico. Se puede calcular el número de copias finales a partir de las copias iniciales, el rendimiento y el número de ciclos según la fórmula siguiente:

$$N = N_0(1 + R)^X$$

donde:

N = Número de copias finales.

N_0 = Número de moléculas a tiempo 0.

R = Rendimiento de cada ciclo expresado en tanto por uno.

X = Número de ciclos.

La reacción se lleva a cabo en termocicladores, pequeños aparatos que controlan con precisión las condiciones de cada etapa y las repite cíclicamente (tiempo, temperatura, número de ciclos, etc.).

Se pueden amplificar con mucha eficiencia fragmentos de tamaños entre 50 pb y 3 kb; asimismo, en condiciones especiales también se pueden amplificar secuencias de 20 a 40 kb.

La temperatura, los tiempos de cada fase del ciclo, el número de ciclos y la concentración de Mg^{+2} se suelen optimizar para cada experimento. Para detectar el producto amplificado se pueden emplear varios métodos: 1) electroforesis en geles de agarosa tiñendo el ADN con bromuro de etidio; 2) marcar el ácido nucleico con fluorescencia y usar geles de poliacrilamida; 3) marcar con radiactividad, separar por electroforesis y pasarlo al papel por la técnica de Southern como se describió en el apartado 3.3 (figura 3.22); 4) usar fluorescencia asociada a la cinética de la reacción, que se describirá en detalle en el próximo apartado 4.4.1 (PCR cuantitativa).

Las polimerasas que se emplean son termorresistentes, es decir, estables a altas temperaturas, para no tener que añadir la enzima manualmente en cada ciclo. La purificada de la bacteria *Thermus aquaticus* (Taq), que vive en manantiales de

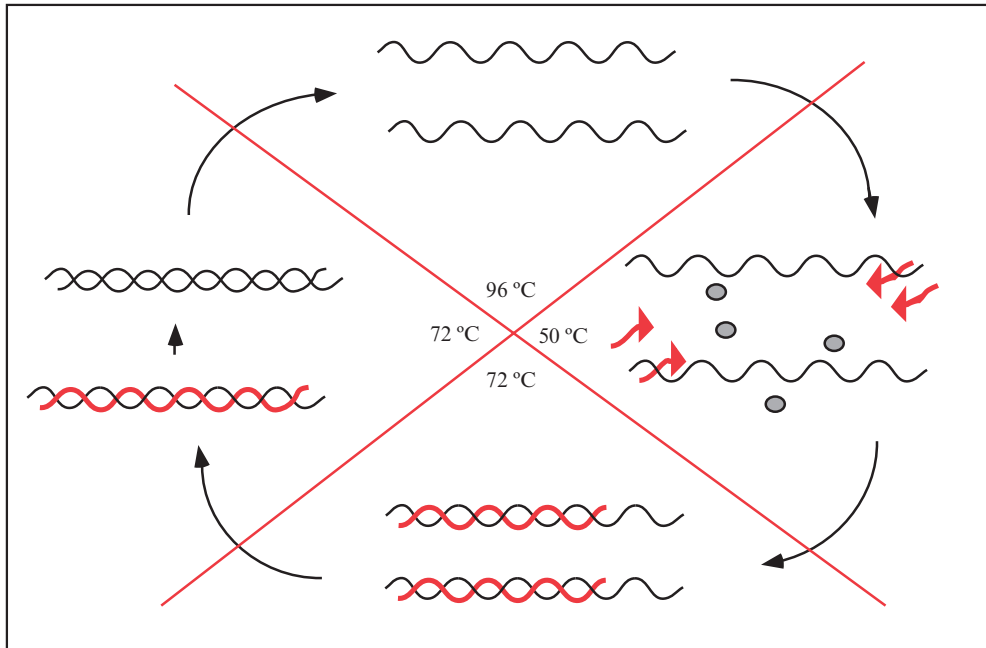


Figura 4.22. Esquema cíclico de la reacción en cadena de la polimerasa (PCR). En cada ciclo la cadena se disocia por calor (96 °C) permitiendo el acceso de oligos específicos, nucleótidos (no se muestran) y polimerasa al ADN; la temperatura se baja a 50 °C para estabilizar los híbridos y, seguidamente, se sube a 72 °C para que la síntesis tenga lugar en las regiones delimitadas por los cebadores (oligos). El tamaño de los oligos suele ser de unos 20 nucleótidos.

agua caliente, es una de las más utilizadas. Sin embargo, esta enzima carece de actividad correctora de errores (exonucleasa 3' → 5'), por lo que puede introducir aproximadamente un nucleótido equivocado por cada 1.000 colocados, a concentraciones altas de magnesio, y uno por cada 1.000.000 colocados, en otras condiciones. Este índice de error no suele presentar problemas en la mayoría de las aplicaciones, pero ha de tenerse en cuenta y la magnitud de la equivocación puede ser importante si ocurre en las primeras rondas. Otras polimerasas, como la obtenida de *Pyrococcus furiosus* (Pfu) o la aislada de la bacteria *Thermococcus litoralis* que se encontró en la ventana de ventilación de un submarino (Vent), incorporan errores con una frecuencia mucho más baja y pueden copiar fragmentos de mayor longitud que la Taq.

Para que la especificidad sea máxima, se debe evitar que la polimerasa y el resto de los compo-

nentes estén a temperatura ambiente antes de iniciar el primer ciclo. La polimerasa puede también añadirse una vez que se ha producido la primera desnaturalización del ADN (*hot start*) o incluso mejor, para no tener que manipular los tubos una vez iniciado el proceso, se pueden añadir anticuerpos anti-Taq, que inhibirán la actividad de la polimerasa hasta que se desnaturalicen por calor. La amplificación de las secuencias requiere cambios rápidos de temperatura, por lo que el proceso se realiza en un termociclador que lleva un microprocesador para programar estos cambios y el número de ciclos deseado.

Hay que tener en cuenta que, si el fragmento a amplificar es de tamaño inferior al molde, sólo empezarán a producirse copias exactas a la diana a partir del segundo ciclo, ya que en el primero no hay oligonucleótidos que sirvan de tope (figura 4.23).

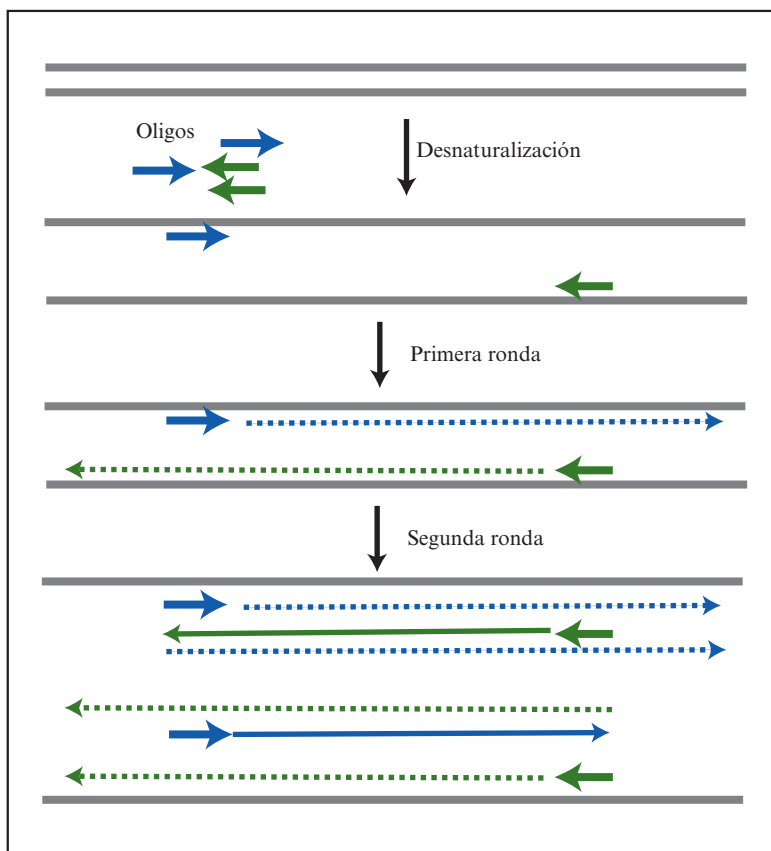


Figura 4.23. Primer y segundo ciclo de la PCR cuando el fragmento a amplificar es de menor tamaño que el molde. En línea discontinua se muestran los intermediarios de mayor tamaño en la primera y segunda ronda.

4.4.1. Variaciones concretas de la técnica básica

Dentro de las variaciones más interesantes están las relativas a la amplificación de moléculas de ARN previo paso a ADNc (RT-PCR o retro-PCR); la PCR cuantitativa a tiempo real; la PCR inversa; la PCR asimétrica; la PCR *in situ* o la adición de diversas secuencias durante la amplificación.

Amplificación de ARNm vía ADNc (RT-PCR o retro-PCR)

El ARN no se puede amplificar por PCR de una manera directa, pero sí se puede pasar previa-

mente a ADN (capítulo 3, apartado 3.1.3 y figura 3.10). Esta amplificación indirecta se puede realizar en un solo tubo de ensayo de dos maneras diferentes:

1. Poniendo las dos polimerasas (transcriptasa inversa y ADN polimerasa termorresistente) desde el comienzo en el mismo tubo, junto con los cebadores de ambos extremos, desoxinucleótidos y sulfato de magnesio. En un principio, durante unos 45 minutos y a 48 °C se sintetiza la primera hebra del ADNc copia del ARNm; a continuación se sube la temperatura a 94 °C durante 2 minutos, con lo que se desnaturaliza el molde, sintetizándose seguidamente la segunda hebra del ADNc y las

múltiples copias de ambas hebras durante 40 ciclos. Se pueden amplificar cantidades de ARNm en el rango de 1 picogramo (10^{-12} gramos) a 100 nanogramos (100×10^{-9} gramos).

2. Utilizando una ADN polimerasa termo-resistente que posee además actividad transcriptasa inversa. Es la polimerasa Tth que se aisló originalmente de la eubacteria termofílica *Thermus thermophilus*. Amplifica bien ARNm de hasta 1 kilobase.

Una variante es la utilización de un oligonucleótido específico dentro del gen y otro inespecífico, con lo que se amplifica no sólo la región codificante de un ARNm, sino también, en muchos casos, la región no traducible 5' o 3' del mensajero.

PCR cuantitativa a tiempo real (qPCR)

Permite el seguimiento de la cinética de amplificación cuando la reacción es exponencial. Para la cuantificación hay muchas variantes y aquí describiremos una de ellas. En esta variante, necesitaremos una sonda que hibride con una parte interna de la secuencia a amplificar. Tras la etapa de desnaturalización, tanto los oligo-cebadores como la sonda interna hibridarán con sus secuencias complementarias. La sonda interna consiste en un pequeño oligonucleótido sintético que posee dos marcas, una en cada extremo (figura 4.24). En el extremo 5', por ejemplo, lleva un carbono marcado con fluorescencia (6-carboxi-fluoresceína), y en el extremo 3' un mitigador de la fluorescencia (6-carboximetil-rodamina), es decir, un compuesto que absorbe la fluorescencia emitida por otro en proximidad física. El efecto del mitigador se debe a la transferencia de la energía fluorescente por resonancia (efecto FRET). Así, mientras la sonda esté intacta, toda la fluorescencia será absorbida por el mitigador, pero durante el proceso de síntesis, la polimerasa en su avance romperá la sonda con su actividad exonucleasa 5' $\rightarrow$ 3' eliminando el efecto mitigador y liberando fluorescencia medible. El aumento de fluorescencia será captado por un detector (a 518 nm para este colorante fluorescente) al tiempo que trans-

curre la reacción de amplificación. Una computadora nos facilita la gráfica de la amplificación, en la que los valores de incremento neto de la fluorescencia se representan respecto al número de ciclos.

Un ejemplo de una gráfica de este tipo se representa en la figura 4.24b. La gráfica representa la amplificación del ADNc correspondiente al ARNm de la interleuquina 4 humana (IL4) clonado en un plásmido. Un ordenador nos facilita la gráfica en la que el incremento de la fluorescencia (ΔR) se calcula según la fórmula: $\Delta R = R_n^+ - R_n^-$, donde R_n^+ es la emisión de fluorescencia en cada punto elegido en el tiempo y R_n^- es la fluorescencia basal. Durante los primeros ciclos de la PCR, los valores de ΔR no superan la línea basal (que se establece arbitrariamente), pero llega un momento en el que los valores de fluorescencia empiezan a crecer despegando de la línea basal. El valor C_t (ciclo umbral, del inglés *cycle threshold*) representa el ciclo en el que ocurre el despegue. Este valor será inversamente proporcional a la cantidad de ADN diana (región a amplificar) inicial.

PCR inversa

La PCR inversa nos permite amplificar secuencias desconocidas vecinas a secuencias genómicas conocidas. Imaginemos que queremos aislar las secuencias adyacentes al lugar donde un retrovirus se ha insertado. Tomaremos ADN genómico y lo trataremos con una enzima de restricción cualquiera, Hind III, por ejemplo (figura 4.25), y permitiremos la formación de círculos por complementariedad de los extremos cohesivos. Se realiza entonces una PCR utilizando como cebadores dos secuencias de la región conocida, pero cuyos extremos 3'OH se dirijan hacia el exterior de la molécula conocida. Así se amplificará un fragmento alrededor del círculo creado, que al ser tratado con la diana Hind nos liberará dos fragmentos que representarán las secuencias anteriores y posteriores al gen concreto. La misma técnica nos permitiría amplificar cualquier región reguladora flanqueante de un gen cuya secuencia sea conocida.

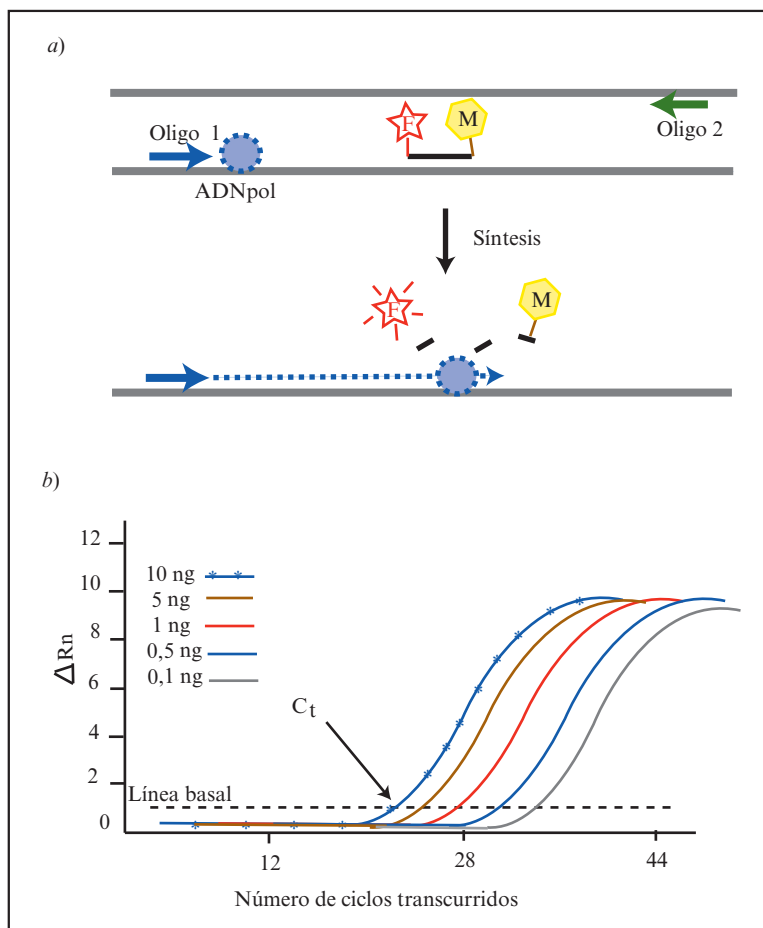


Figura 4.24. Representación esquemática de la PCR cuantitativa a tiempo real (qPCR), *a*) variante en la que se utilizan 3 sondas: dos oligoceadores y una sonda interna que lleve en el extremo 5' un marcaje fluorescente y en el 3' un mitigador que dejará de suprimir la fluorescencia con el paso de la ADN polimerasa. La cantidad de fluorescencia emitida será proporcional al número de moléculas copiadas, *b*) gráfica generada por un ordenador en la que se observa la amplificación de varias diluciones seriadas (5x) del plásmido pADNc-IL4. Los valores de ΔR se representan en ordenadas, y el número de ciclos en abscisas. El ciclo en el que se empieza a emitir fluorescencia representa el valor C_t y será proporcional a la cantidad de ADNc inicial en cada amplificación.

PCR asimétrica

Es ésta una variante que nos permite generar un gran número de moléculas de ADN de banda simple que se pueden utilizar directamente para secuenciar o como sondas para hibridaciones. El método consiste en poner diferente proporción de los dos oligoceadores de manera que uno esté a

concentración limitante. Así, los 20-25 primeros ciclos amplifican las dos hebras del ADN, pero al agotarse uno de los cebadores, los 5-10 ciclos siguientes serán de ADN de hebra sencilla (figura 4.26). Utilizando, por ejemplo, una relación inicial de 50 picomoles de un oligoceador y 0,5 del otro, obtendremos de 1 a 3 picomoles de ADN de banda simple después de 30 ciclos.

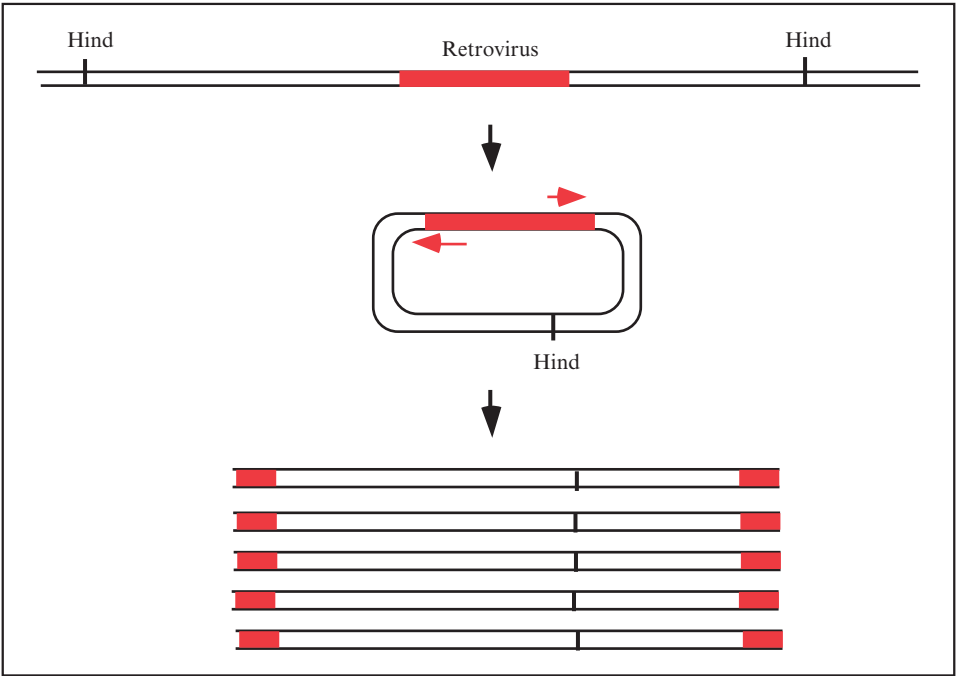


Figura 4.25. Amplificación de las regiones vecinas al lugar de inserción de un retrovirus por PCR inversa.

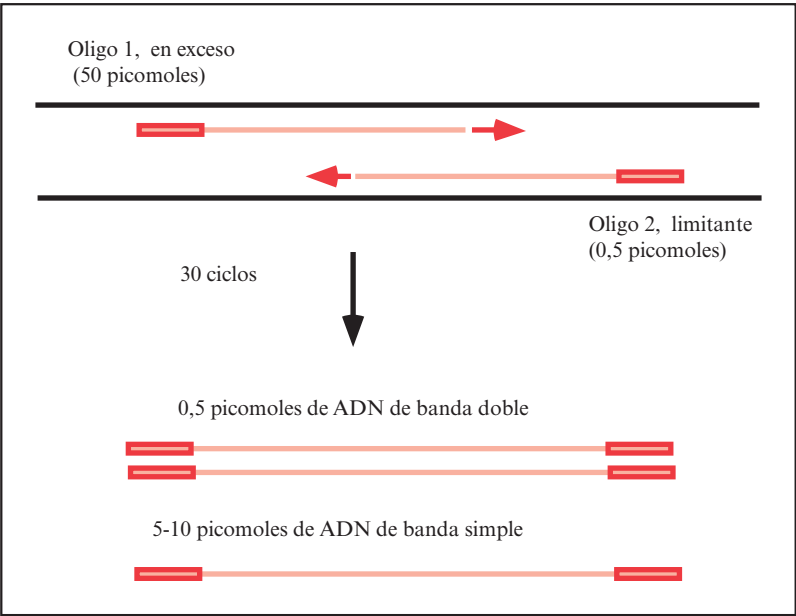


Figura 4.26. Esquema del procedimiento empleado en la PCR asimétrica para generar ADN de hebra sencilla.

PCR anidada

La PCR anidada (en inglés *nested*) mejora la especificidad cuando hay que amplificar muchos ciclos (60, 80). Después de realizada la amplificación

se toma una alícuota y se realiza una segunda amplificación con cebadores internos a los primeros y exclusivos de la región a amplificar. Así se evita la amplificación de fragmentos inespecíficos que pudiera haber en la primera reacción de PCR (figura 4.27).

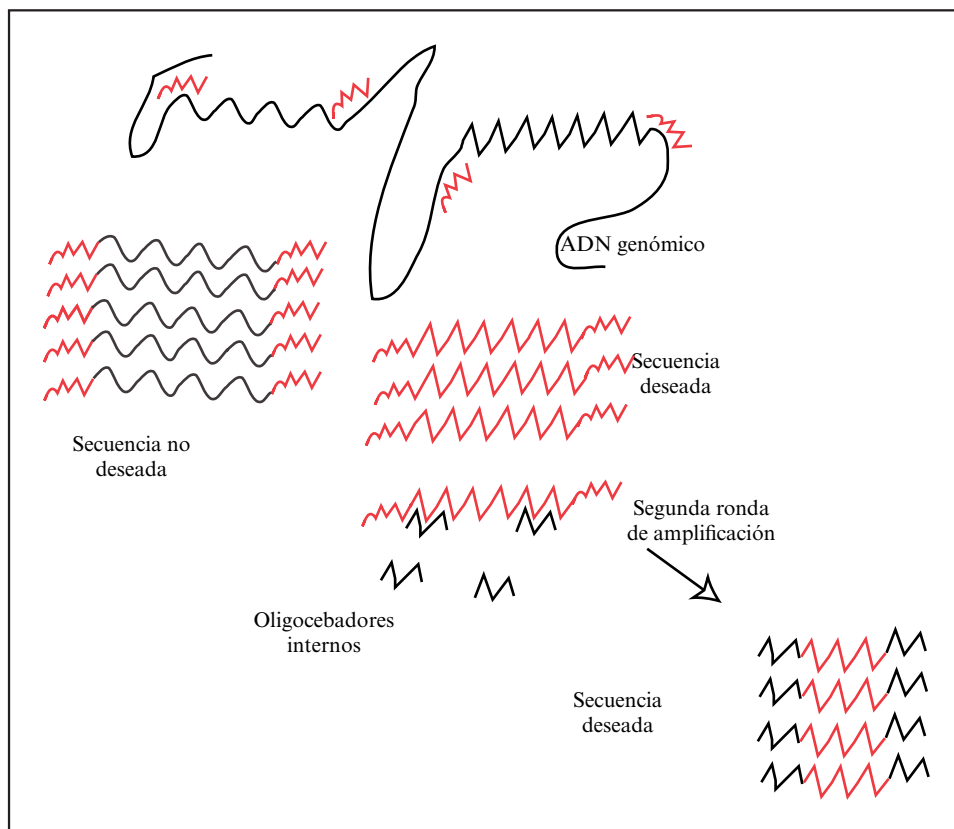


Figura 4.27. PCR anidada. Los productos de una primera amplificación por PCR se someten a una segunda ronda de amplificación utilizando cebadores internos a los primeros y específicos de la secuencia a copiar. Los fragmentos no específicos de la primera ronda no serán copiados en esta segunda.

PCR *in situ*

La utilización de la técnica de la reacción en cadena de la polimerasa para incrementar la sensibilidad de las hibridaciones *in situ* se desarrolló independientemente en varios laboratorios a finales de los ochenta. En la técnica de hibridación *in situ*, una sonda marcada radiactivamente se pone

en contacto durante varias horas con cromosomas cuya estructura permanezca intacta, pero cuyo ADN esté desnaturalizado. La hibridación de la sonda con el ADN cromosómico nos permite localizar la secuencia deseada en los cromosomas. También podemos hibridar la sonda con moléculas de ARN celular complementarias. Desafortunadamente las hibridaciones *in situ* no sue-

len tener sensibilidad suficiente para detectar menos de 10-20 copias de ADN o ARN por célula. La PCR nos amplificaría *in situ* las moléculas cuya detección localizada buscamos.

La reacción de PCR *in situ* se lleva a cabo en células o tejidos fijados en preparaciones citológicas. Las preparaciones se colocan en termocicladores especiales que llevan planchas de calor y

permiten programar los cambios de temperatura asociados a cada ciclo. La mezcla de reacción de la amplificación (nucleótidos, oligos, polimerasa, etcétera) se añade a unos anillos de clonación que se colocan sobre la región citológica (figura 4.28). Uno de los nucleótidos debe marcarse con tritio, fluoresceína o digoxigenina para identificar el lugar de síntesis.

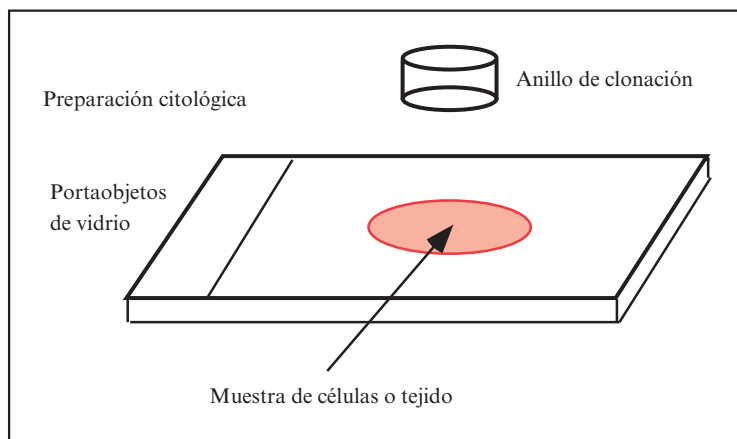


Figura 4.28. Preparación de una muestra para PCR *in situ* en un portaobjetos citológico. Sobre el material biológico se coloca un anillo de clonación y dentro de éste la mezcla de reacción.

PCR con adición de secuencias en los extremos durante la amplificación

La adición de una diana para una enzima de restricción, de adaptadores universales o de cajetines de expresión durante la ampliación de fragmentos es práctica habitual si se desea clonar inmediatamente después el fragmento en el primer caso, amplificar una secuencia desconocida en el segundo o expresar un gen eucariótico clonado, en el sistema bacteriano, en el tercero.

- La adición de una diana para una enzima de restricción se incluye sin dificultad al extremo de los oligocebadores específicos. Un cebador específico suele tener unos 20 nucleótidos de longitud y admite sin dificultad una secuencia diana de una enzima de restricción de 6 nucleótidos, que quedará

desapareada en las primeras rondas de replicación, pero será incorporada a medida que transcurre la reacción. No obstante, el corte sólo será efectivo si añadimos algunos nucleótidos de más en el extremo 5' que permitan a la enzima de restricción asentarse correctamente sobre la diana. En la mayoría de los casos no se observa una disminución de la especificidad o la eficiencia de amplificación (figura 4.29a).

- Los adaptadores universales son oligonucleótidos sintéticos con extremos complementarios a los 4 nucleótidos protuberantes de una determinada diana de restricción (figura 4.29b). Los oligocebadores serán complementarios a los adaptadores y servirán para amplificar una secuencia desconocida que se encuentre entre ambos.

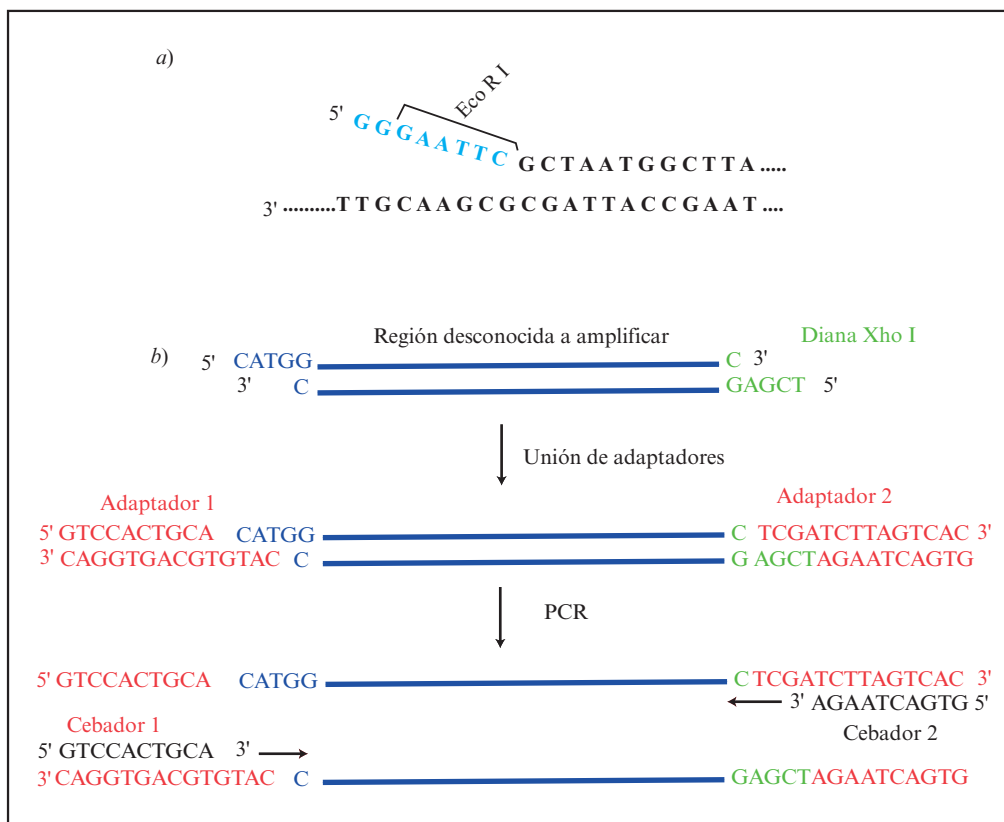


Figura 4.29. Adición de secuencias en los extremos de una PCR, a) adición de la diana de restricción para la enzima Eco RI como parte del oligocebador. Se añaden un par de bases más en el extremo 5' (G G) para asegurar la eficacia de asiento y de corte por la enzima de restricción. b) Adición de adaptadores universales para amplificar secuencias desconocidas.

— Los *cajetines de expresión* del sistema bacteriano se pueden añadir a plásmidos con insertos de ADNc de mensajeros eucarióticos que se deseen expresar en bacterias. En la primera ronda el cajetín quedará colgando para incorporarse a partir de la segunda ronda y ser mayoritario al final (figura 4.30).

Los oligocebadores también puede modificarse de otro modo: con biotina, fluoresceína, radiactividad, etc. Además podemos cambiar algún nucleótido dentro del oligocebador constituyendo un modo de introducir mutaciones específicas.

PCR múltiple

La PCR múltiple hace referencia a la posibilidad de amplificar más de una secuencia o fragmento en una única reacción (varios exones de un mismo gen, por ejemplo). Con el incremento del número de pares de oligocebadores se pueden originar algunos productos espúreos, pero, en general, la especificidad y la eficiencia se mantienen muy próximas a los casos de amplificación única. PCR múltiple ocurre de forma natural con oligocebadores de secuencias repetidas en el genoma, como las secuencias *Alu* o los microsatélites, pero en estos casos se genera un gran número de ampli-

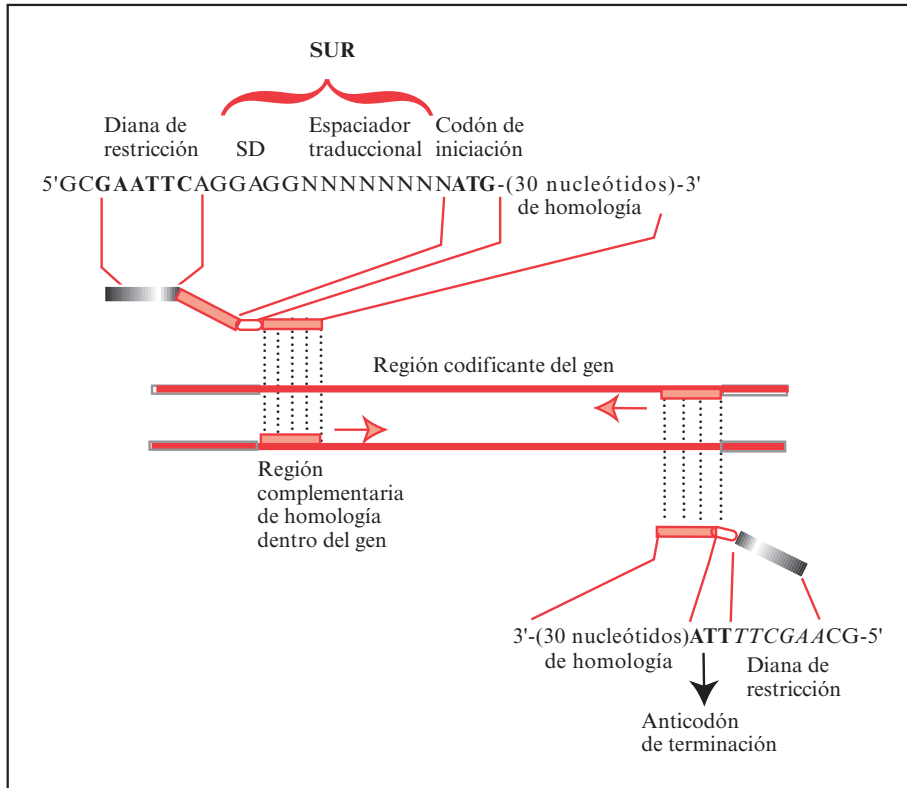


Figura 4.30. Diagrama esquemático de la construcción de un cajetín de expresión por PCR para insertar en un vector con control transcripcional. El oligocebador 5' consta de dos nucleótidos terminales no específicos y requeridos para que se enganche la enzima de restricción; una diana para una enzima de restricción; la secuencia de unión al ribosoma (SUR) formada por la región Shine-Dalgarno (SD) y un espaciador traduccional; el codón de iniciación y 25-30 nucleótidos específicos del ADN codificante. El oligocebador 3' consta de 25-30 nucleótidos específicos de la región 3' codificante seguido de un codón terminación de cadena, diana de restricción y nucleótidos terminales.

ficaciones no específicas de difícil interpretación, por lo que es importante en la PCR múltiple que las secuencias a amplificar no se encuentren repetidas en el genoma haploide.

La amplificación de dos o tres secuencias simultáneas no requiere cambios en las condiciones de reacción generales. Las dificultades surgen cuando el número de fragmentos se incrementa y se observa que algunos de ellos se amplifican con mucha más eficiencia que otros o que aparecen fragmentos totalmente inespecíficos. Una manera de paliar estos efectos es ir añadiendo oligocebadores consecutivamente a la misma reacción, optimizando las

condiciones después de cada nueva adición. La variable más importante suelen ser los propios oligos, y los parámetros a tener más en cuenta: la temperatura, la cantidad de enzima y los tiempos de polimerización. Con las combinaciones apropiadas de oligocebadores y condiciones de la reacción se ha conseguido amplificar simultáneamente una gran variedad de secuencias no repetidas.

PCR arbitraria

Es un tipo de PCR múltiple en la que se utiliza como cebador un único oligonucleótido de banda

simple cuya secuencia, arbitraria, origina múltiples fragmentos anónimos de ADN debido a la relajación en las condiciones de hibridación de los primeros ciclos. Así, la iniciación es prácticamente al azar y dependerá de la secuencia elegida arbitrariamente como cebadora. El resultado es la amplificación de una pequeña muestra al azar del genoma analizado que puede ser directamente comparada con las muestras de otros organismos amplificados con el mismo cebador y en las mismas condiciones. La mayoría de las secuencias amplificadas son únicas y no aparecen desviaciones relacionadas con el origen cromosómico de las bandas observadas.

La PCR arbitraria también se ha utilizado para detectar alteraciones en células tumorales, permitiendo descubrir la existencia de un subgrupo de tumores colorrectales que contienen un gran número de mutaciones somáticas y que probablemente alberguen mutaciones en genes implicados en la reparación de secuencias (genes mutadores).

4.5. ALGUNAS APLICACIONES DE LA PCR

La reacción en cadena de la polimerasa ha facilitado enormemente experimentos largos y tediosos, además de permitir abordar nuevos problemas biológicos. Como ejemplos, pueden citarse: modificación de los extremos de fragmentos clonados, detección de secuencias sin purificación previa, preparación de muestras monocatenarias para secuenciación, detección de polimorfismos de secuencias y de mutaciones, tipificación de ADN para trasplantes, diagnóstico de enfermedades, resolución de problemas forenses, estudios arqueológicos y evolutivos, detección de microorganismos infecciosos, detección de células tumorales, amplificación de un ADN para su posterior clonación.

No se pretende en este apartado describir y analizar cada una de las múltiples aplicaciones de la PCR que se extienden a muchos campos de la biología y la medicina. Así pues, hemos elegido algunas de las aplicaciones básicas y biomédicas que combinan varias de las técnicas ya descritas.

En los capítulos siguientes se hará también referencia a esta versátil técnica que se ha hecho imprescindible en los laboratorios de biología molecular.

Diagnóstico precoz de enfermedades

En ocasiones se puede realizar un diagnóstico genético antes de que aparezcan los síntomas de la enfermedad. Un ejemplo es la enfermedad de Huntington, alteración neurodegenerativa que aparece en uno de cada 10.000 individuos. Los primeros síntomas van asociados a trastornos motores y mentales que aparecen entre los 35 y 40 años cuando el paciente, sin su conocimiento, es probable que lo haya transmitido a sus descendientes. La enfermedad es progresiva causando la muerte a los 10-20 años de la aparición de los primeros síntomas, provoca discapacidad física (baille de San Vito) y mental, con pérdidas del 20% al 30% del peso del cerebro. La alteración genética asociada se transmite con carácter típicamente dominante, donde el homocigótico y el heterocigótico son indistinguibles en sus manifestaciones fenotípicas. El gen se localiza en la región subtelomérica del cromosoma 4 (4p 16.3) y contiene una región de 9-35 repeticiones del triplete CAG en individuos normales. Un número de repeticiones del triplete entre 40 y 100 va asociado a la patología de la enfermedad. CAG se corresponde con el aminoácido glutamina y el resultado es una proteína (huntintina) con una estructura tridimensional distorsionada que se une a sus proteínas asociadas con mucha más fuerza que la proteína normal, impidiendo una función normal del dímero y causando muerte neuronal generalizada.

El gen responsable ha sido clonado y el diagnóstico se lleva a cabo de modo directo amplificando y determinando el tamaño de la región repetida (CAG)_n partiendo de las regiones flanqueantes. Aun pudiendo realizar un diagnóstico, no se conoce una terapia preventiva para evitar el desarrollo de la enfermedad.

En fecundaciones *in vitro* el diagnóstico prenatal de enfermedades hereditarias ligadas al cromosoma X, como la hemofilia o la distrofia muscular, puede realizarse en los primeros momentos

del desarrollo del huevo fecundado. Estas enfermedades, al ser recesivas, sólo afectan a los varones mientras que las hembras son portadoras de la enfermedad, pero no la sufren. En familias con alto riesgo de transmisión de una de estas enfermedades se puede realizar una fecundación *in vitro* y analizar por PCR el sexo de los embriones incipientes. En el caso de **la hemofilia**, si hay sospecha de que una mujer es portadora de un alelo alterado y quiere tener descendencia sana con un

varón no hemofílico, sólo se implantarán en la madre embriones femeninos, ya que el 50% de los varones sufrirán la enfermedad mientras que el 100% de las hembras estarán libres de ella, aunque el 50% de ellas serán portadoras. Si el hemofílico fuera el padre, al tener descendencia con una hembra normal no portadora, se elegirían los embriones varones para la implantación, ya que su cromosoma X procederá forzosamente de la hembra, y así todos los varones estarán libres de la

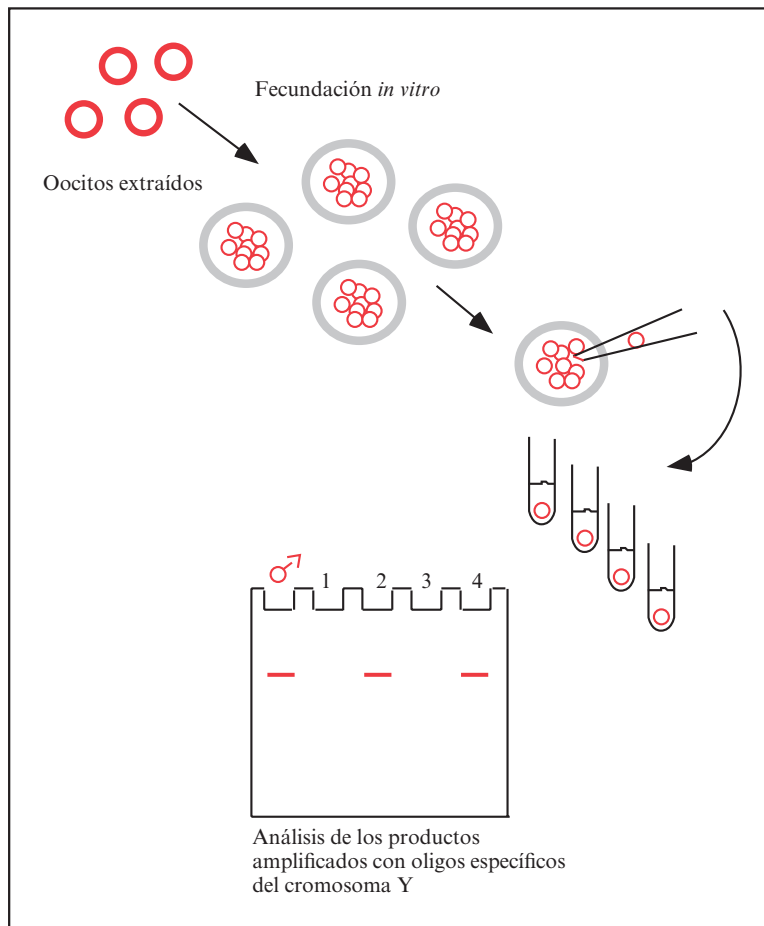


Figura 4.31. Determinación del sexo en embriones con riesgo de heredar enfermedades ligadas al cromosoma X. Extracción de oocitos, fecundación *in vitro* y cultivo hasta el estadio de 6-10 células. Extracción de una célula de cada embrión viable y amplificación por PCR utilizando cebadores específicos del cromosoma Y. El análisis de los productos indica que los embriones 2 y 4 son varones y, dependiendo del caso, deberán o no ser reimplantados para su posterior desarrollo.

enfermedad; las hembras, aunque todas sanas, serán portadoras en todos los casos.

La determinación del sexo se lleva a cabo tomando una célula única de embriones en estadios de 6-10 células, extrayendo su ADN y amplificando regiones específicas del cromosoma Y. La secuencia DYZ1, por ejemplo, tiene 3,5 kb y está repetida 5.000 veces; se puede amplificar esta secuencia del ADN extraído de una sola célula, durante 60 ciclos, dando un producto visible sólo en células masculinas (figura 4.31).

La distrofia muscular (variedad Duchenne, en honor al neurofisiólogo francés que la describió) afecta a uno de cada 3.500 varones y está causada por distintas deleciones en el gen de la distrofina, localizado en el cromosoma X. Este gen tiene un tamaño enorme, con 79 exones y un ARNm de 14 kb. Existen preferencias respecto al lugar, tamaño y cantidad de exones eliminados que se asocian con la patología. El diagnóstico se puede realizar por PCR múltiple en la que se amplifican a la vez los exones que con mayor frecuencia sufren dele-

ciones (figura 4.32). Fallos en la amplificación pueden inducir a falsos positivos, por lo que deleciones de un solo exón o en exones no contiguos deben considerarse como sospechosas y repetir el ensayo. Además, la transmisión de la enfermedad se puede intentar evitar con una fecundación *in vitro* y diagnóstico precoz del sexo de los embriones, del mismo modo que se ha descrito para la hemofilia.

En enfermedades autosómicas se puede amplificar la región implicada y secuenciarla para detectar la presencia o no de la alteración causante de la enfermedad, pero el diagnóstico prenatal es algo más difícil.

PCR en investigación forense

Los investigadores en criminología recogen la evidencia biológica que ha podido quedar asociada al suceso con el fin de identificar al ejecutor. Cada vez con mayor frecuencia se caracterizan muestras forenses a nivel de ADN con el fin de

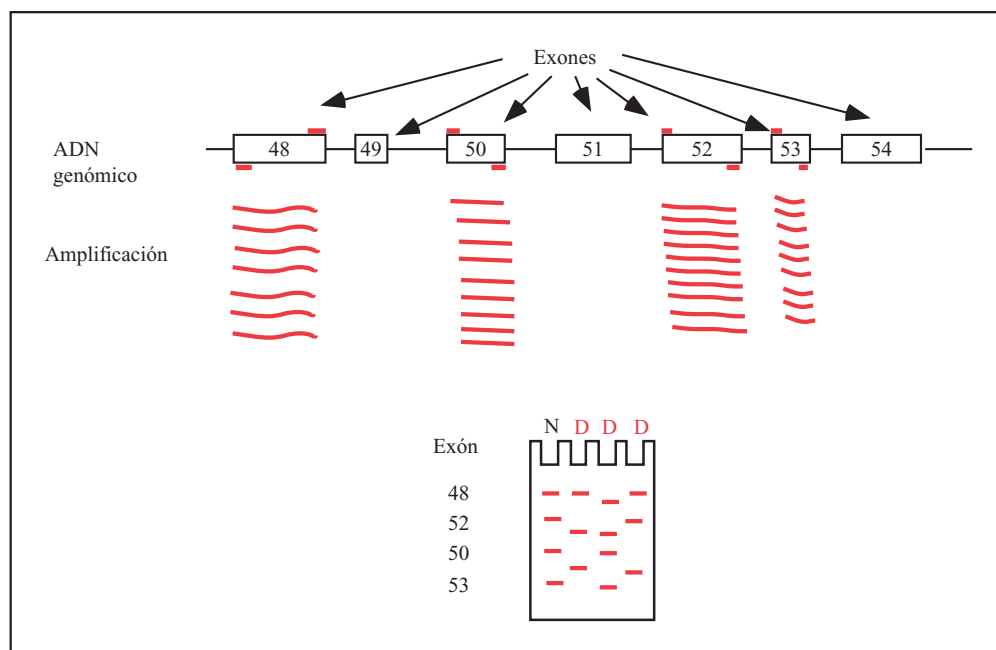


Figura 4.32. PCR múltiple para diagnosticar deleciones en exones del gen de la distrofina.

inculpar o exculpar a un sospechoso. Cuando la muestra es escasa, es necesaria una amplificación por PCR antes de estudiar la variabilidad de fragmentos concretos (véase capítulo 3, apartado 3.3.2). De dos microlitros de sangre, menos de uno de semen, 20 de saliva o un pelo arrancado de raíz, se puede extraer suficiente ADN (50 ng) para amplificarlo por PCR y realizar una prueba con sondas unilocus. Inclusive muestras de ADN degradado pueden ser amplificadas por PCR.

Una vez amplificada la muestra, se procede a la separación en geles de agarosa o poliacrilamida para resolver variantes en el tamaño de fragmentos, o se secuencian.

Se utilizan oligocebadores de las regiones colindantes a una hipervariable y se determina el tamaño de los alelos paterno y materno en cada caso. M representa marcadores de peso molecular; S1, S2 y S3 son sospechosos de una violación, V es la amplificación de la región hipervariable de

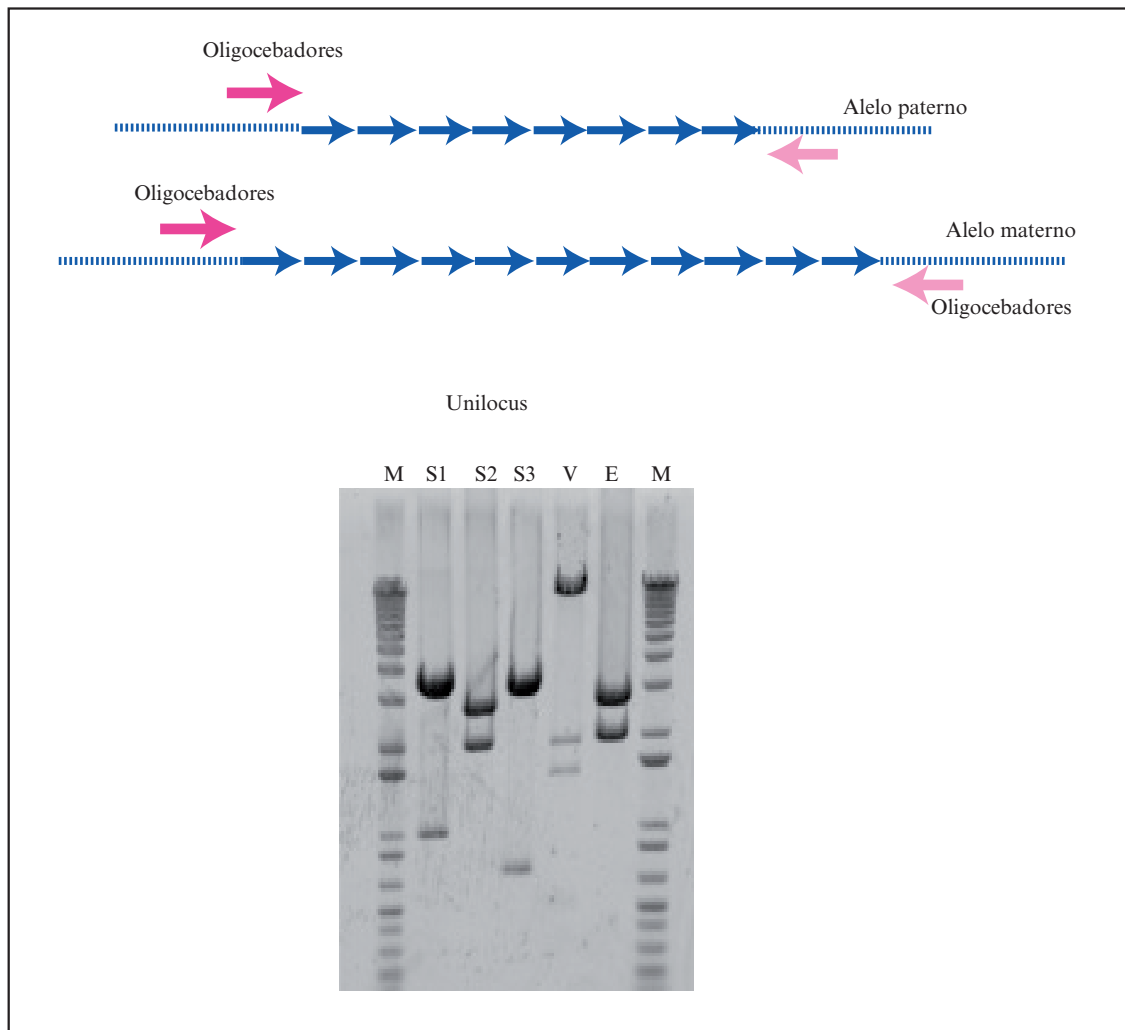


Figura 4.33. PCR de una región hipervariable para determinar el tamaño de la región repetida.

la víctima y E el patrón del semen encontrado en la víctima que coincide con el del sospechoso S2.

El ADN mitocondrial, por otra parte, es un excelente material para el análisis de muestras forenses por varias razones: 1) hay más de 5.000 copias por célula, lo que constituye en sí mismo una amplificación; 2) no se suele encontrar degradado al estar protegido por la mitocondria, y 3) algunas secuencias, dentro de la región del lazo-D no codificante (implicadas en la iniciación de la replicación), son muy hipervariables.

La raíz de un solo pelo suele ser suficiente para amplificar por PCR asimétrica una región hipervariable de ADN mitocondrial y secuenciar a continuación.

También se puede intuir el color de los ojos de un sospechoso a partir de una muestra de sangre o semen, amplificando y secuenciando los polimorfismos en genes relacionados con la pigmentación del iris.

PCR en estudios evolutivos

El grado de relación o parentesco que existe entre dos especies vivas se puede analizar de varias maneras independientemente, por ejemplo: 1) comparando directamente su ADN total por hibridación; 2) comparando el número de mutaciones acumuladas en genes funcionales, o 3) comparando el ADN mitocondrial que no intercambia material genético durante la meiosis y tiene un índice mutacional más elevado que el ADN genómico. Un gen amplificado por PCR servirá de sonda para el aislamiento por hibridación del gen homólogo en casi cualquier otra especie. La relación en las condiciones de hibridación se ajustará al grado de identidad que se presume para este gen entre las distintas especies. Cuando se trata,

sin embargo, de relacionar especies existentes en la actualidad con otras ya extinguidas, tenemos el impedimento del severo deterioro del material genético que se pueda obtener de fósiles, momias o incluso algunas partes blandas preservadas, con un tamaño medio de 100 pares de bases. La PCR, sin embargo, trabaja bien con fragmentos pequeños y deteriorados pudiendo amplificar un porcentaje razonable del genoma, suficiente en la mayoría de los casos para su comparación con especies existentes. ADN mitocondrial de fragmentos óseos o procedentes de piel necrótica de momias o cadáveres de más de diez años, se ha amplificado con éxito. El mayor problema en estos casos es la contaminación con ADN humano del operador o de bacterias que se pueden encontrar en las muestras.

Gracias a la PCR, el ADN mitocondrial de un cerebro humano de hace 7.000 años descubierto en Florida (EE. UU.) pudo ser amplificado; muestras de maíz peruano de hace 4.500 años han servido para extraer ADN genómico amplificado posteriormente. También se han conseguido secuencias de ADN mitocondrial del hombre de Neandertal.

En el área de los fósiles, un caso de plantas de hace 17 millones de años es particularmente llamativo. Se ha podido amplificar un fragmento de 800 pb del genoma del cloroplasto del fósil, lo que ha permitido determinar su probable origen relacionado con la magnolia. Existen, no obstante, serias dudas acerca de una contaminación con ADN bacteriano en este estudio en particular.

Numerosos controles extremadamente rigurosos son imprescindibles en este tipo de estudios, ya que una molécula de ADN intacto bacteriano o humano se amplificará con preferencia a una serie de moléculas viejas degradadas y rehidratadas.

Una mutación en sentido clásico es la alteración de un gen que produce un cambio fenotípico heredable, pero en sentido amplio es cualquier alteración transmisible. Los distintos tipos de mutaciones nos definen la naturaleza de la alteración que ha provocado el cambio.

Las mutaciones aportan la evidencia definitiva de que el ADN es el material genético. Cuando un cambio en el ADN causa una alteración en la secuencia de una proteína, podemos concluir que el ADN codifica esa proteína.

Además de las mutaciones en las que la modificación ocurrida inactiva total o parcialmente la función de un gen, hay toda una gama de variantes dentro de un mismo carácter que implican pequeñas alteraciones funcionales; son las formas alélicas o polimórficas de dicho gen. La palabra alelo significa en griego «el otro» y, por definición, todos los alelos son mutaciones. Las formas alélicas o polimórficas para ser denominadas como tales han de presentarse en las poblaciones con una frecuencia mínima de 0,01%. No es demasiado frecuente que se hable de formas alélicas o polimórficas de las regiones reguladoras, pero en sentido estricto debería aplicárseles la misma terminología.

Las mutaciones son de enorme importancia por ser la base de la evolución de los seres vivos. Los sistemas celulares crean con sus alteraciones nuevas formas de los genes o de regular sus funciones. Hoy se parte de la idea de que el genoma es plástico y tolerante; por una parte, hay redundancia génica que permite alteraciones en una copia asegurándose la función correcta en la otra; por

otra parte, hay redundancia funcional por la que un gen puede suplir la función de otro distinto en determinadas circunstancias. Esta plasticidad favorece la aparición de nuevas funciones por combinatoria de otras ya existentes: exones de distintos genes pueden combinarse en la formación de variantes de sus respectivas funciones de origen. Alteraciones en las secuencias reguladoras pueden aumentar o disminuir el producto de un gen (variaciones en la cantidad), expresarlo o no en determinados tejidos (variaciones en la calidad) y ello puede incidir sobre decisiones de gran trascendencia como la creación de nuevas estructuras o modificaciones sustanciales de las existentes.

La célula o individuo que hereda una mutación es un mutante, mientras que la forma más frecuente de un gen es el tipo silvestre. El proceso de producción de mutantes se denomina mutagénesis y el agente mutagénico sería el compuesto o mecanismo responsable de la alteración.

5.1. TIPOS DE MUTACIONES

Atendiendo a su origen, las mutaciones pueden ser espontáneas, si ocurren de modo natural, o inducidas por agentes específicos. Unas y otras pueden ser puntuales, si afectan solamente a una base, o reordenaciones estructurales, si afectan a una región más amplia.

Las mutaciones puntuales son las sustituciones de una base por otra. Hay dos tipos de mutaciones puntuales por sustituciones de bases: las transiciones y las transversiones.

Transiciones. Son las más corrientes. Consisten en la sustitución de una purina (A o G) por la otra o una pirimidina (C o T) por la otra (figura 5.1).

Transversiones. Son mucho menos comunes y en ellas una purina (A o G) es sustituida por una pirimidina (C o T), y viceversa (figura 5.1).

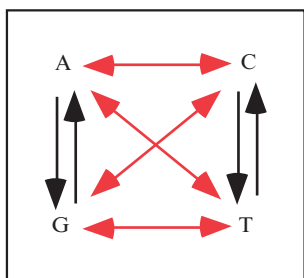


Figura 5.1. Ejemplo de transiciones (flechas negras) y transversiones (flechas de dos cabezas rojas).

Si las mutaciones ocurrieran totalmente al azar y cualquier base tuviera la misma probabilidad de ser cambiada por otra, no tendría sentido encontrar que las transiciones son en muchos casos el doble de frecuentes que las transversiones, cuando por cada transición ($A \rightarrow G$, por ejemplo) pueden ocurrir dos transversiones ($A \rightarrow C$ y $A \rightarrow T$). Las transversiones deberían, en teoría, ser el doble de frecuentes que las transiciones; pero, en la práctica, las deaminaciones espontáneas de citosina metilada del par CpG a timidina ocurren con una frecuencia 8,5 veces superior a la alteración de una base en cualquier otro par de dinucleótidos. También es posible que las actividades correctoras de errores de las polimerasas tengan más facilidad para corregir transversiones que transiciones. En el ADN codificante, las transiciones se toleran mejor que las transversiones porque los cambios originados son más conservadores para la cadena polipeptídica.

Un segundo tipo de mutaciones puntuales son las que se originan por la adición o delección de una base y producen un efecto mucho más devastador que las anteriores, ya que cambia la pauta de lectura de la proteína afectada (figura 5.2).

Sólo en los casos en los que la transición o transversión crearan codones sin sentido se originaría la prematura terminación de la cadena y el efecto sería comparable. Mutaciones puntuales de adición de una base pueden ser recuperadas por una nueva mutación de delección de la misma u otra base en un lugar cercano (figura 5.2), ya que restablecen la pauta de lectura.

Las reordenaciones también afectan a las secuencias, y comprenden: delecciones, inserciones de elementos móviles, inversiones y translocaciones de fragmentos más o menos grandes del ADN.

Las mutaciones pueden dividirse también en: somáticas o germinales, según que afecten a las células de los tejidos del cuerpo o a las células reproductoras. Evidentemente, el alcance de la mutación es muy distinto. Una mutación germinal afectará a todas las células del organismo, mientras que una mutación somática tendrá una penetración muy limitada, especialmente si ocurre en la fase adulta o en la fase final del desarrollo.

Mutaciones por sus efectos:

- *Letales.* Son las que producen la muerte del individuo indicando que han afectado una función vital. Son letales, por ejemplo, las mutaciones auxotróficas en las que se pierde la capacidad para sintetizar un producto metabólico esencial para la supervivencia.
- *Viables.* Afectan: 1) propiedades fenotípicas visibles como color, tamaño, aspecto, etcétera; 2) propiedades fisiológicas que limiten el crecimiento y proliferación celular, y 3) propiedades estructurales.
- *Condicionales.* Son aquellas en las que sólo en determinadas circunstancias se manifiesta la mutación. Los mutantes sensibles a la temperatura son un ejemplo de ellas; estos mutantes, al subir la temperatura unos grados por encima de la normal, muestran el fenotipo mutante; o la letalidad en el caso de ser letales condicionales. La gran ventaja de este tipo de mutantes es que, al no expresar el defecto en condiciones permisivas, se elimina la lentitud y debilidad que suelen acompañar al creci-

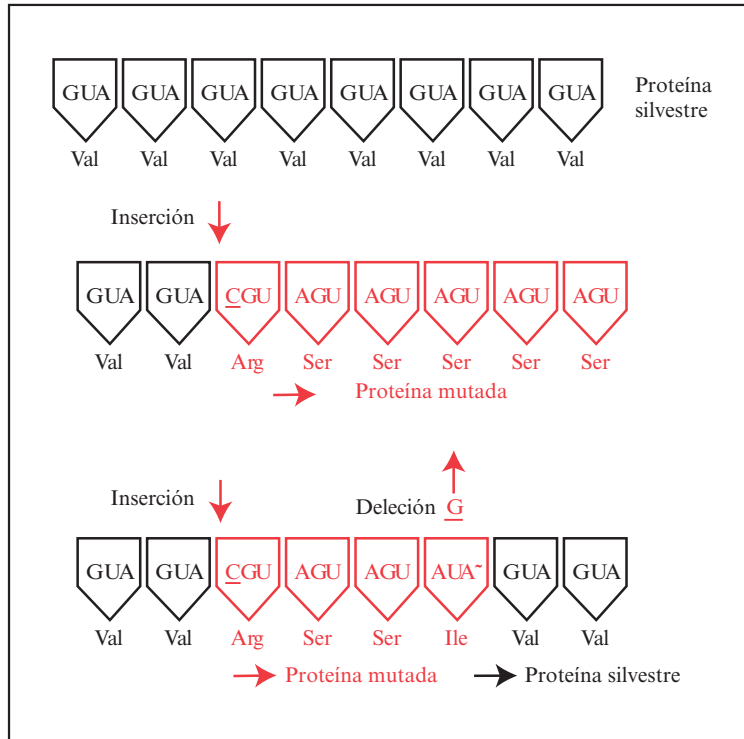


Figura 5.2. Mutaciones puntuales por adición o deleción de una base. Recuperación de la pauta de lectura con una deleción después de una primera inserción.

miento de estos individuos cuando están seriamente afectados en alguna función importante.

- *Resistentes.* Son mutaciones que permiten al individuo crecer en presencia de sustancias tóxicas o letales para la célula en condiciones normales. Ejemplos de este tipo de mutaciones son las bacterias resistentes a antibióticos, a la infección de bacteriófagos, etc.
- *En zonas reguladoras.* Comprende regiones promotoras y potenciadoras de la expresión pudiendo encontrarse próximas o distantes del gen que se regula. Pueden inactivar total o parcialmente la expresión de uno o varios genes.
- *Silenciosas.* No producen ningún cambio apreciable. Cuando aparecen en regiones

codificantes no modifican la actividad del producto y con frecuencia afectan a la tercera base del codón.

- *Supresoras.* Son segundas mutaciones que nos devuelven la función perdida en la primera. Pueden ocurrir en el mismo lugar, cerca o en otro gen. Un ejemplo en el que la segunda mutación está en el mismo lugar que la primera y recupera la función, lo constituiría una proteína con una estructura característica en la que un aminoácido de carácter positivo (una lisina) está apareado con un aminoácido de carácter negativo (ácido aspártico). Si el ácido aspártico muta y se convierte en otro aminoácido de carácter positivo (por ejemplo, arginina o histidina), la proteína no se puede empaquetar de forma adecuada y su

función se ve afectada. Si se produce una segunda mutación en el mismo lugar o en un lugar próximo a la primera, que nos coloque otro aminoácido que tenga carácter negativo (ácido glutámico, por ejemplo), esta segunda mutación es supresora de la primera, y el organismo, un revertiente (figura 5.3). Ejemplos de este tipo de rever-

siones son también las inserciones seguidas de deleciones de una sola base (figura 5.2), o tres inserciones (o deleciones) muy próximas, pues ellas nos devuelven la pauta de lectura y nos dejarán unos pocos aminoácidos erróneos que, si no están en un lugar fundamental para la actividad de la proteína, no dejarán huella.

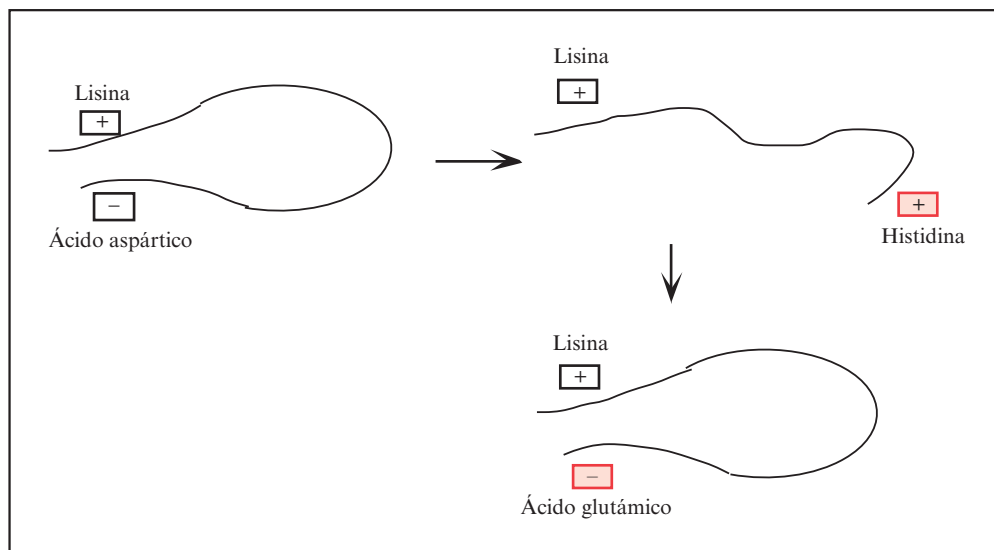


Figura 5.3. Proteína revertiente por mutación supresora. Una mutación nos convierte el ácido aspártico, de marcado carácter negativo, en histidina que tiene carácter positivo. La estructura de la proteína se ve alterada, así como su función. Una segunda mutación que cambia la histidina por ácido glutámico devuelve el carácter negativo al aminoácido y restaura la estructura y la función.

Como ejemplo de una reversión en la que la segunda mutación está en un gen distinto de donde ha ocurrido la primera, tenemos los ARNt supresores. Empecemos con una mutación que origina un codón de terminación de cadena ($\text{UAC} \rightarrow \text{UAG}$) en una proteína importante. La consecuencia es un producto truncado prematuramente y sin actividad (figura 5.4). Esta mutación puede ser rescatada por una alteración en el anticodón de un gen ARNt ($\text{AGC} \rightarrow \text{AUC}$). El nuevo ARNt reconoce un codón de terminación de cadena (como resultado de la mutación) e incorpora el aminoácido que transporta en respuesta a este co-

dón (siempre el mismo, Ser en este caso), eliminando los efectos de una proteína truncada y originando otra con un aminoácido erróneo (Ser por Tyr), pero que presumiblemente no afectará a la actividad del producto final. Este tipo de supresiones podría originar daños generalizados en la célula, ya que impediría la terminación de muchas cadenas polipeptídicas en sus lugares de terminación habitual. En la práctica, esto no parece ser un problema, ya que la mayoría de las proteínas tienen más de un codón de terminación para señalar el fin de la traducción (aun cuando con uno sería suficiente).

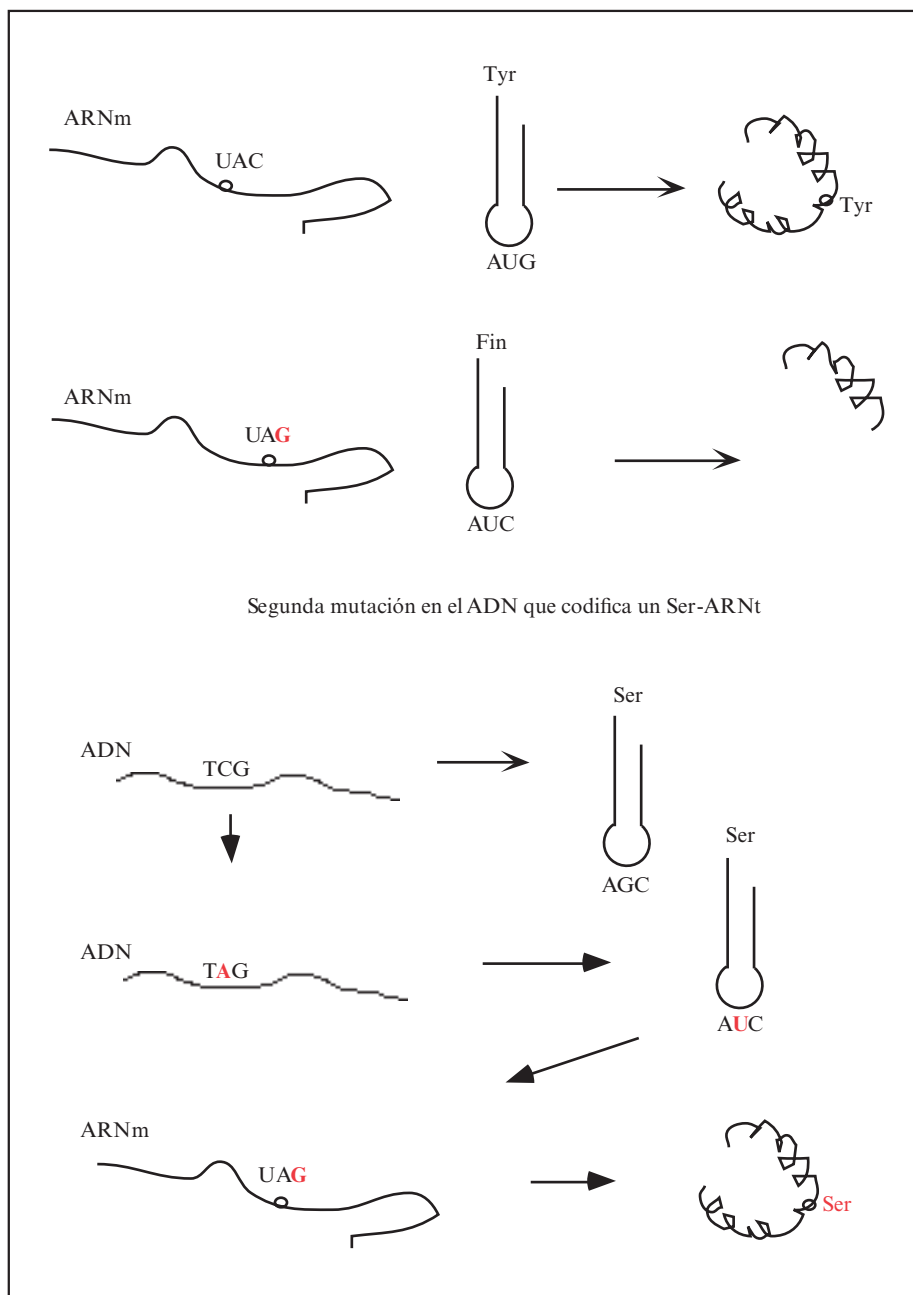


Figura 5.4. Ejemplo de supresión por ARNt. Una mutación puntual sustituye el codón UAC correspondiente a la tirosina (Tyr) por el UAG, que es reconocido por el ARNt de terminación de cadena (Fin); la proteína se termina prematuramente y es inactiva. Una segunda mutación en el anticodón del ARNt correspondiente a la serina (Ser AGC-AUC) reconoce e incorpora serinas en codones de terminación (UAG). El tamaño de la proteína se restaura, así como su actividad.

5.2. MUTACIONES ESPONTÁNEAS: TIPOS Y FRECUENCIA

Las mutaciones espontáneas son alteraciones en la molécula de ADN que ocurren con una cierta frecuencia de manera natural (figura 5.5). Hay enzimas específicas de reparación para todas ellas. Entre los distintos tipos encontramos:

1. *Ruptura de los enlaces glicosílicos* debido a la inestabilidad química del enlace. La

molécula de ARN es más resistente a estas rupturas, es un efecto secundario negativo de la reducción de ribosa a desoxirribosa en los ácidos nucleicos. Ocurre preferentemente en los residuos de purina. La frecuencia en condiciones fisiológicas (pH 7,4; 37 °C) es de 10.000 rupturas por día por genoma humano. No hay enzimas de reparación específicas que reinserten las purinas que faltan en la hebra doble, pero hay enzimas que rompen la hebra de

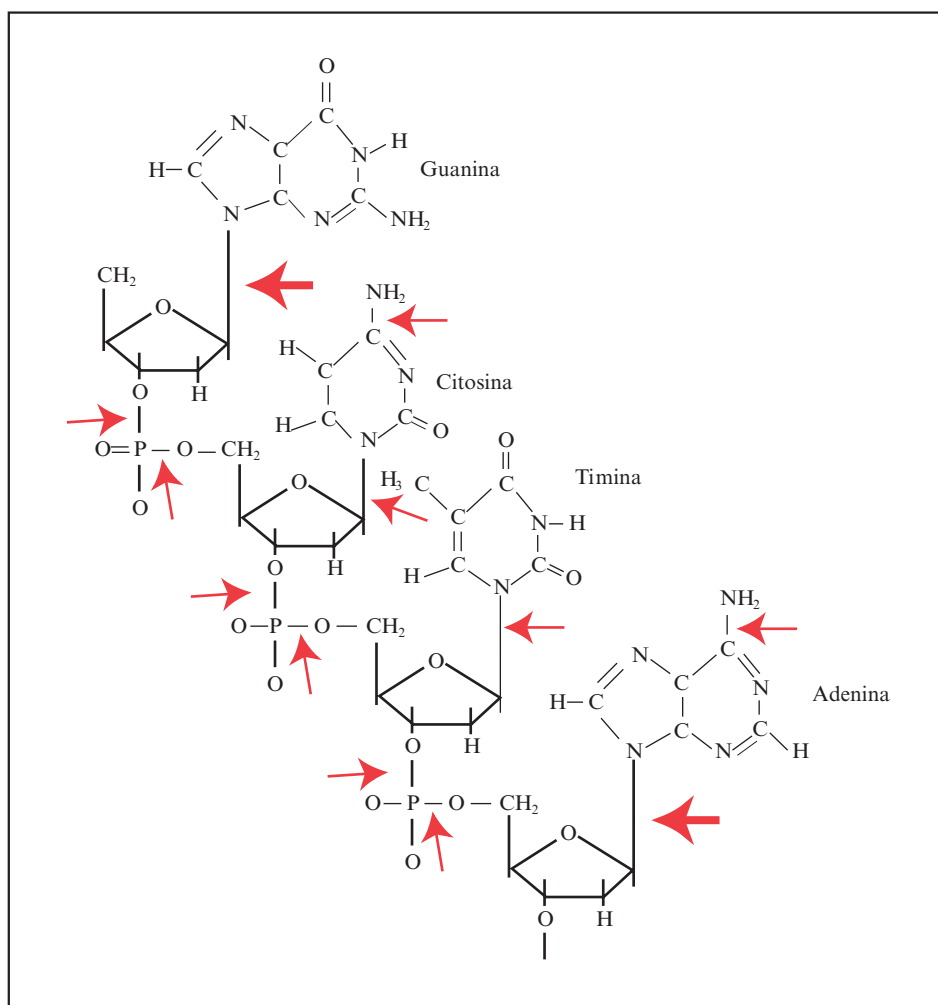


Figura 5.5. Estructura del ADN, mostrando (flechas rojas) los lugares de ruptura espontánea de enlaces más frecuentes.

ADN en lugares despurinizados que serán reconocidos por sistemas reparadores de rupturas (reparación por escisión).

2. *Desaminaciones oxidativas espontáneas.* De citosina a uracilo y menos frecuentemente de adenina a hipoxantina. La enzima uracil-N glicosilasa elimina los uracilos del ADN. En este punto es ventajoso que la molécula de ADN carezca de uracilos, pues si los tuviera no se sabría si son o no desaminaciones de la citosina. La desaminación oxidativa de la citosina metilada origina timina (figura 5.6). Timinas apareadas con guaninas también son eliminadas por un sistema algo lento e ineficiente que consiste en una glicosilasa que reconoce el par T/G y sustituye siempre la timina por citosina.
3. *Hidrólisis de ésteres fosfóricos* (figura 5.5). La molécula de ADN es unas 200 veces más estable que la de ARN a esta hidrólisis. Se cree que esta propiedad pudo haber contribuido de manera decisiva en el establecimiento del ADN frente al ARN como molécula portadora de la información genética.
4. *Replicación del ADN.* La ADN polimerasa de bacterias incorpora aproximadamente 1 error por cada 10.000 nucleótidos copiados, ella misma corrige sus propios fallos hasta una tasa de error de $1/10^7$. En cada ronda de replicación del genoma humano (3×10^9 pares de bases) se escapan aproximadamente de 1 a 3 fallos sin reparar. Un individuo tiene 10^{14} células y durante su vida ocurren una media de entre 10^1 a 10^{17} divisiones.
5. *Radiaciones ultravioleta.* Causan uniones covalentes entre dos timinas consecutivas, lo que paraliza la replicación (fotodímeros de UV). La célula, no obstante, posee un mecanismo de fotorreactivación que le permite utilizar la energía solar (fotones) para romper los enlaces covalentes de los dímeros de timina. La enzima fotorreactivante encargada de reparar los daños se induce por luz solar. También pueden ac-

tuar las enzimas del sistema de reparación por escisión. Este mecanismo se descubrió con la observación de que bacterias sometidas a luz ultravioleta sobrevivían mejor si después de la radiación se las sometía a la luz solar intensa. Es probablemente el primer mecanismo de reparación que existió, ya que está presente en las bacterias más ancestrales donde no hay otros mecanismos de reparación.

Hay un segundo mecanismo de reparación denominado escisión-reparación, que corta la zona dañada en el ADN. Una endonucleasa reconoce la zona afectada porque no se aparea correctamente, la corta y una actividad exonucleasa $5' \rightarrow 3'$ (probablemente de la polimerasa I) elimina la zona dañada, y copiará ahora el fragmento eliminado. En fibroblastos humanos se ha observado que más de la mitad de los dímeros de timina inducidos por tratamientos de luz ultravioleta se reparan en menos de 24 horas. En humanos, hay una enfermedad denominada Xeroderma pigmentosa, cuyos pacientes son extremadamente sensibles a la luz solar y en particular a los rayos ultravioleta, siendo enormemente propensos a adquirir melanomas y otros cánceres de piel. Estos individuos son incapaces de reparar los dímeros de timina producidos por las radiaciones solares.

6. *Radicales libres de la molécula de oxígeno.* El oxígeno activo puede atacar el enlace $3'-4'$ de la desoxirribosa, lo que a su vez provoca una ruptura en la cadena; también puede producir saturación en los anillos de las purinas o pirimidinas, fragmentación y condensación de la molécula. La lesión más frecuente de todas las que produce el oxígeno activo es la de guaninas con un anillo imidazol saturado (8-oxo-7, 8-dihidro-dG). Este residuo en el ADN toma automáticamente la configuración «sin» y aparea con adenina provocando transversiones en la replicación (figura 5.7a). Hay varios mecanismos que mi-

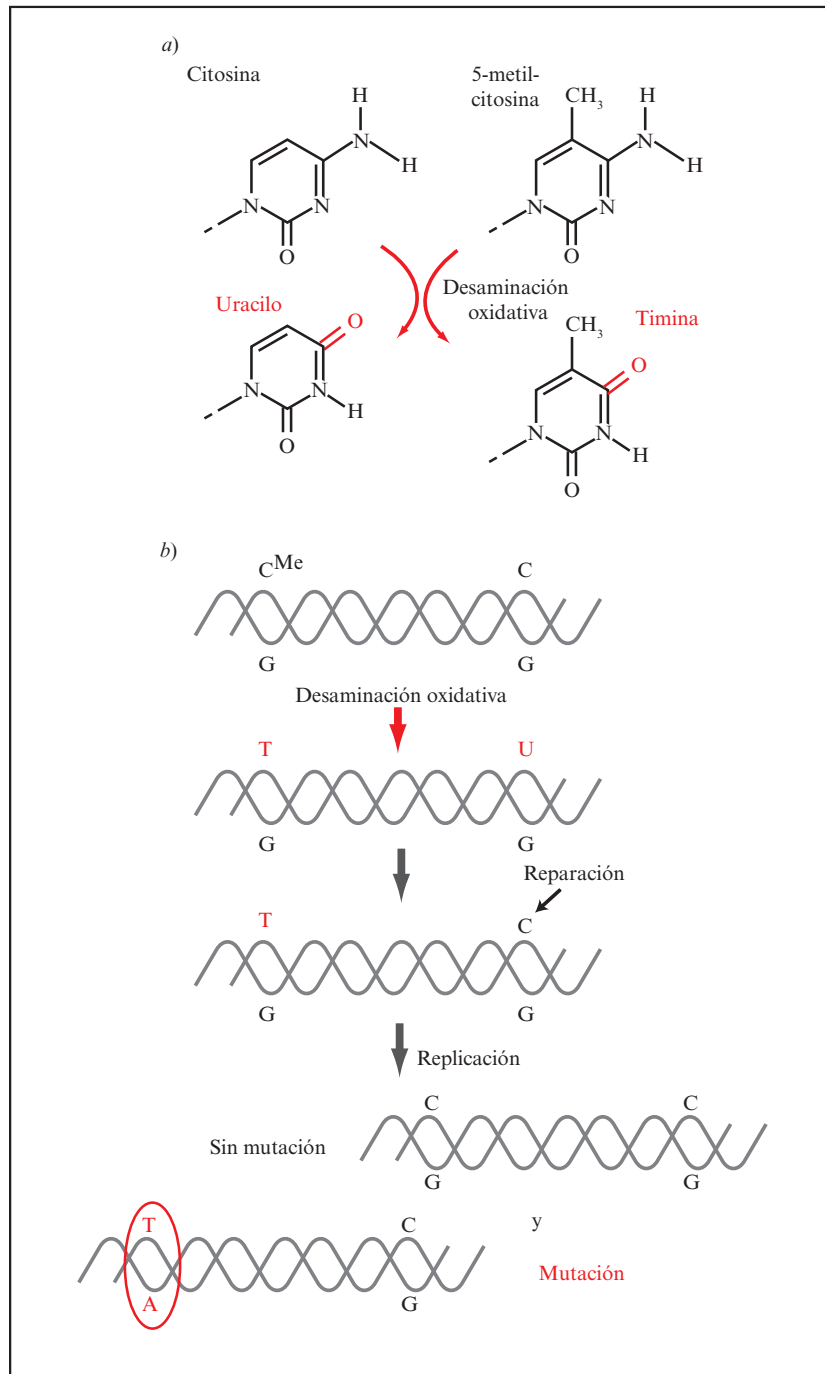


Figura 5.6. Diferencias entre uracilo y timina detectables por la N-glicosilasa y conversión de la citosina metilada en timina por desaminación oxidativa.

nimizan en la célula la acción de los radicales libres de oxígeno; uno de estos mecanismos ni siquiera espera a que el precursor trinucleótido (GTP) se incorpore al ADN, sino que actúa directamente en los depósitos celulares de dGTP. Los depósitos de GTP se mantienen más bajos que los de otros trinucleótidos y la enzima MutT los repasa en busca de 8-oxo-dGTP

que inmediatamente descompone (figura 5.7b). En las células eucarióticas, el mantener el metabolismo del oxígeno en las mitocondrias es una gran medida de protección del ADN nuclear de los radicales libres de oxígeno. El ADN mitocondrial tiene una tasa mutacional unas 10 veces más alta que la nuclear, lo que origina una divergencia del 2-4% por millón de años.

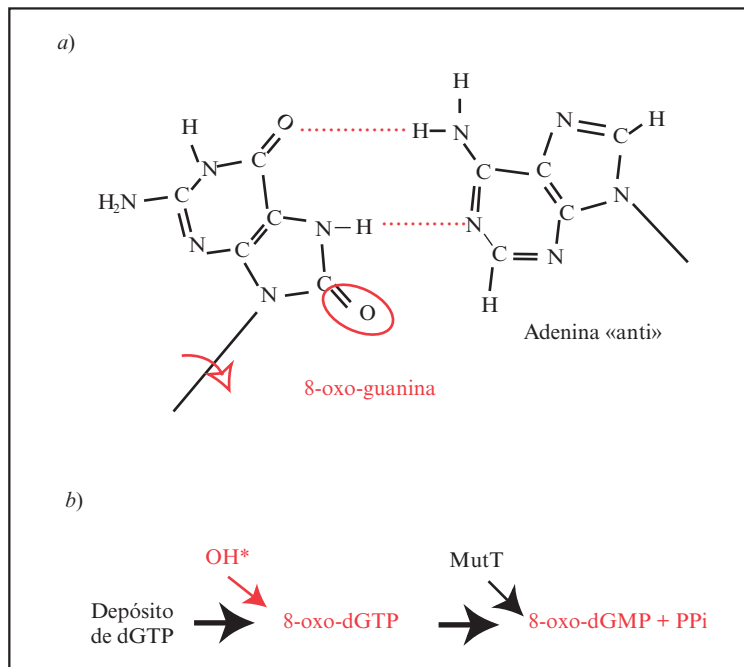


Figura 5.7. Efecto de los radicales libres de oxígeno. a) Apareamiento entre adenina y una guanina convertida en 8-oxi-7,8-dihidroguanina por un radical libre de oxígeno. La flecha indica la rotación de la base, con respecto a la ribosa, de la forma «anti» a la «sin». b) Eliminación de los residuos dGTP oxidados de los depósitos celulares por la enzima MutT.

La frecuencia mutacional espontánea debería ser constante y al azar en todo el genoma pero hay puntos especialmente calientes donde la acumulación de mutaciones es muy superior a la media (10-100 veces más). Las islas CpG, por ejemplo, suelen ser puntos calientes por la frecuencia de desaminaciones oxidativas. Mutaciones espontáneas que inactivan un gen ocurren en bacterias con una frecuencia de entre 10^{-5} y 10^{-6} sucesos por

locus y por generación. No hay datos comparables en humanos.

5.3. MUTACIONES PROVOCADAS: AGENTES MUTAGÉNICOS

La frecuencia de mutaciones se incrementa considerablemente en presencia de ciertos com-

puestos denominados agentes mutagénicos o mutágenos. En este caso las mutaciones son inducidas. La mayoría de los agentes mutagénicos ejercen su acción directamente sobre el ADN. En algunos casos se intercalan ellos mismos entre las bases del ADN, en otros sustituyen alguna de las bases o la modifican irreversiblemente. La eficiencia de estos agentes se mide por el incremento que producen sobre los valores de mutaciones espontáneas.

Los agentes mutagénicos suelen inducir un tipo concreto de mutaciones:

El 5 bromouracilo. Es un análogo de la timina, pero con un grupo bromuro en vez del metilo de la timina, lo que le permite aparear con la guanina además de con la adenina (figura 5.8); provoca transiciones $TA \rightarrow CG$.

El ácido nitroso. Convierte grupos «amino» en grupos «ceto» por desaminación oxidativa. Así pues, convertiremos:

Citosina en uracilo que aparea con adenina: ($C \rightarrow U:A$).

Adenina en hipoxantina que aparea con citosina: ($A \rightarrow HX:C$).

Guanina en xantina que aparea con citosina: ($G \rightarrow X:C$).

Entre los derivados nitrosos utilizados como agentes mutagénicos de modo rutinario en el laboratorio está la nitrosoguanidina, agente mutagénico muy cancerígeno.

Aunque la nitrosoguanidina es un agente alquilante, su efecto mutagénico radica en la producción de transiciones $GC \rightarrow AT$, pero el mecanismo no se conoce bien, ya que en algunos organismos produce otras sustituciones distintas.

Hay muchos derivados nitrosos que son muy cancerígenos. Debido a ello puede ser peligroso ingerir en la alimentación sustancias tratadas con nitritos o nitratos, pues en nuestro hígado una parte de ellos se transforman en ácido nitroso u otros derivados también mutagénicos.

Agentes intercalantes. Son otro tipo de mutágenos, que incluyen la naranja de acridina, proflavina e ICR-170 (2-metoxi-6, cloro-9[3-etil-2cloro-etil] aminopropil amino). Todos ellos se intercalan

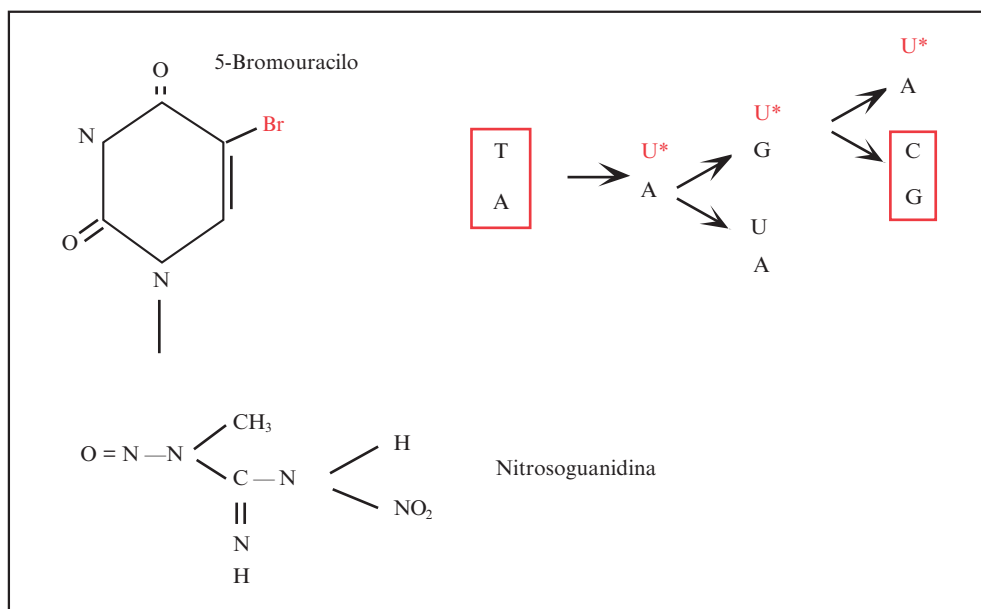


Figura 5.8. Apareamiento del 5-bromouracilo con adenina y guanina, y fórmula de la nitrosoguanidina.

en el ADN, distorsionan la hélice y al llegar la replicación producen la incorporación u omisión de una base por molécula intercalada. Producen muchas mutaciones y supresiones (deleciones seguidas de inserciones que contrarrestarán sus efectos).

Los rayos X producen rupturas cromosómicas induciendo inversiones, translocaciones o deleciones más o menos grandes.

Los *transposones* al moverse causan mutaciones insercionales (véase capítulo 2, apartado 2.3). El gen mutado puede ser clonado utilizando como sonda el elemento móvil. Se les puede inducir a moverse facilitando la transposasa.

Retrovirus. En células de mamíferos se han utilizado retrovirus para inducir mutaciones en genes que se expresan. El retrovirus infecta células en división y se integra preferentemente en genes activos provocando mutagénesis insercional. Se suele sustituir el gen de la envuelta (*env*) por un gen de selección, *puro*, por ejemplo, de modo que, en presencia de puomicina, solo las células que expresan el gen *puro* (N-acetil-transferasa que inactiva al antibiótico) sobrevivirán. Para aislar los genes mutados se utiliza como sonda el propio genoma retroviral en una librería construida a partir del genoma de las células infectadas. Variantes de retrovirus que carecen de su promotor constituyen

los vectores denominados trampas de promotores que producen mutaciones en genes activos en el momento de la inserción.

5.4. MUTAGÉNESIS AL AZAR Y SELECCIÓN ESPECÍFICA

Todos los agentes mutagénicos descritos hasta el momento provocan mutaciones al azar en el genoma del organismo tratado, por lo que será necesario un proceso de selección para aislar mutantes en lugares concretos de entre todos los producidos. Por ejemplo, podemos tratar una cepa de bacterias con nitrosoguanidina y seleccionar colonias mutantes en el gen de la histidina. La selección se realizaría creciendo cada colonia en presencia y ausencia de histidina en el medio.

Si lo que buscamos son mutaciones en el operón *lac* de *E.coli*, concretamente en el gen de la β -galactosidasa, la selección es más sencilla, ya que en presencia del sustrato cromogénico Xgal (5-bromo-4-cloro-3-indolil- β -D-galactóxido) la galactosidasa produce un compuesto coloreado que da color azul intenso a las colonias; al mutar y desaparecer esta función, las colonias permanecen blancas. Así, todas las colonias blancas serán mutantes, y las azules tipo silvestre (figura 5.9).

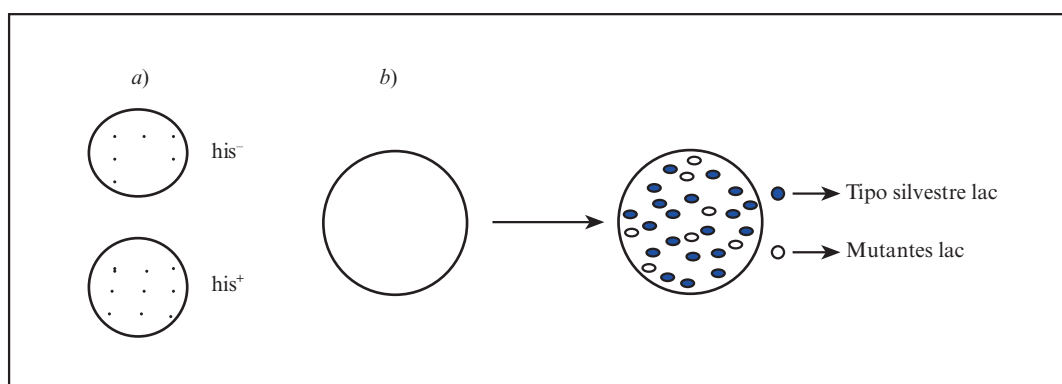


Figura 5.9. Selección de mutaciones al azar. *a)* Mutantes en el operón de la histidina inducidos por nitrosoguanidina y dependientes de la presencia del aminoácido en el medio para sobrevivir. *b)* Selección de mutaciones en el gen de la β -galactosidasa cuyo producto reacciona con un sustrato cromogénico dando color azul. Las colonias mutantes permanecerán de color blanco.

En organismos más grandes, como *Drosophila*, la mutagénesis al azar se hace por radiación o con elementos móviles, como ya se ha indicado. En los experimentos de radiación, se suelen tratar machos con dosis apropiadas de rayos X para inducir mutaciones en el esperma y luego se aparean con hembras normales o especiales. La F_1 poseerá gran número de mutaciones que deberán ser seleccionadas específicamente en cada caso en concreto.

Este tipo de mutagénesis al azar y selección no siempre garantiza que la mutación causante del efecto que vemos en el gen seleccionado haya tenido lugar en el propio gen; hay efectos epistáticos en los que una mutación en un determinado lugar puede afectar a la función de otro u otros genes dependientes del primero.

Otra limitación de esta técnica es la ausencia de métodos de selección específicos para la inmensa mayoría de los genes no muy bien caracterizados e, incluso, de muchos bien estudiados. Otro punto a tener en cuenta es que el mutante obtenido por este método llevará, además de la mutación en el gen seleccionado, un número variable de mutaciones en otros genes que, sin afectar directamente a la viabilidad del organismo, pueden introducir modificaciones indeterminadas respecto al individuo normal. También conviene mencionar que este tipo de mutagénesis requiere una caracterización molecular posterior a nivel de secuenciación para determinar el tipo de mutación originada: cambio de una sola base, pequeña delección, inserción, etc. Por último, hay que tener en cuenta la diploidia de estos organismos, por lo que mutaciones en uno solo de los cromosomas homólogos puede que no se manifiesten.

5.5. MUTAGÉNESIS DIRIGIDA

La mutagénesis dirigida tiene como finalidad modificar un gen en concreto o una parte específica de uno de ellos. Puede llegar a modificar o destruir exclusivamente un nucleótido dentro de un gen concreto. Hay varios métodos para producir mutagénesis dirigida sobre cualquier gen clonado y aislado. Una vez mutado este gen, se puede

de estudiar *in vitro* o reintroducirlo en células en cultivo para su estudio *in vivo*. Podemos modificar desde un solo nucleótido o varios, hasta zonas más o menos extensas dentro de la región clonada. La importancia de la mutagénesis dirigida es grande, ya que permite modificar a voluntad cualquier gen o su expresión. La mutagénesis permitirá idealmente mejorar las prestaciones de muchos productos génicos; sería como mejorar la naturaleza.

5.5.1. Estrategia de provocar delecciones en secuencias clonadas

La secuencia a mutagenizar debe estar clonada. Como ejemplo veremos cómo se crearon *in vitro* delecciones laterales e internas para estudiar los elementos de control de la transcripción en el promotor del virus de los simios: SV 40 (figura 5.10). Se utilizó la exonucleasa Bal 31. Esta nucleasa elimina nucleótidos a partir de extremos 3' OH. Por secuenciación del ADN se había ya localizado el promotor y visto la existencia de dos bloques de 72 pares de bases repetidos, tres de 21 pares de bases también repetidos y ricos en GC y, por último, la región TATA. Mutaciones en esta región van a permitir determinar el efecto de cada uno de estos elementos en la transcripción de cualquier gen puesto a continuación. Debemos tener clonada la región reguladora que uniremos a un gen señal (luciferasa, por ejemplo) de fácil detección en células eucarióticas. La enzima luciferasa se aisló originalmente de la luciérnaga *Photinus pyralis* y emite luz amarilla al unirse a la luciferina, su sustrato. La reacción requiere ATP, Mg^{2+} y O_2 molecular, produciéndose un fogonazo de luz que decae rápidamente (en menos de un segundo), pero que puede ser captado por un luminómetro. La emisión total de luz es proporcional a la actividad luciferasa en la muestra.

Antes de realizar la mutagénesis propiamente dicha trataremos el recombinante representado en la figura 5.10 con las dianas EcoRI (R) y BamHI (B), para obtener una cierta cantidad de la región invariante (B → R) que comprende el vector plasmídico y el gen señal *luc* (luciferasa). Las delecciones en el flanco 5' de la región reguladora

se obtienen tratando una porción del recombinante original con EcoRI y Bal 31; a continuación tratamos con BamHI para eliminar las deleciones en la región plasmídica y uniremos esta molécula a la región invariante para obtener el recombinante de partida, pero delecionado. De

manera análoga, por tratamiento con BamHI (B) y Bal 31, obtendremos recombinantes con deleciones en el flanco 3'. La transfección en células de mamífero y determinación de los niveles de luciferasa nos indicará el efecto que la deleción ha tenido en su expresión.

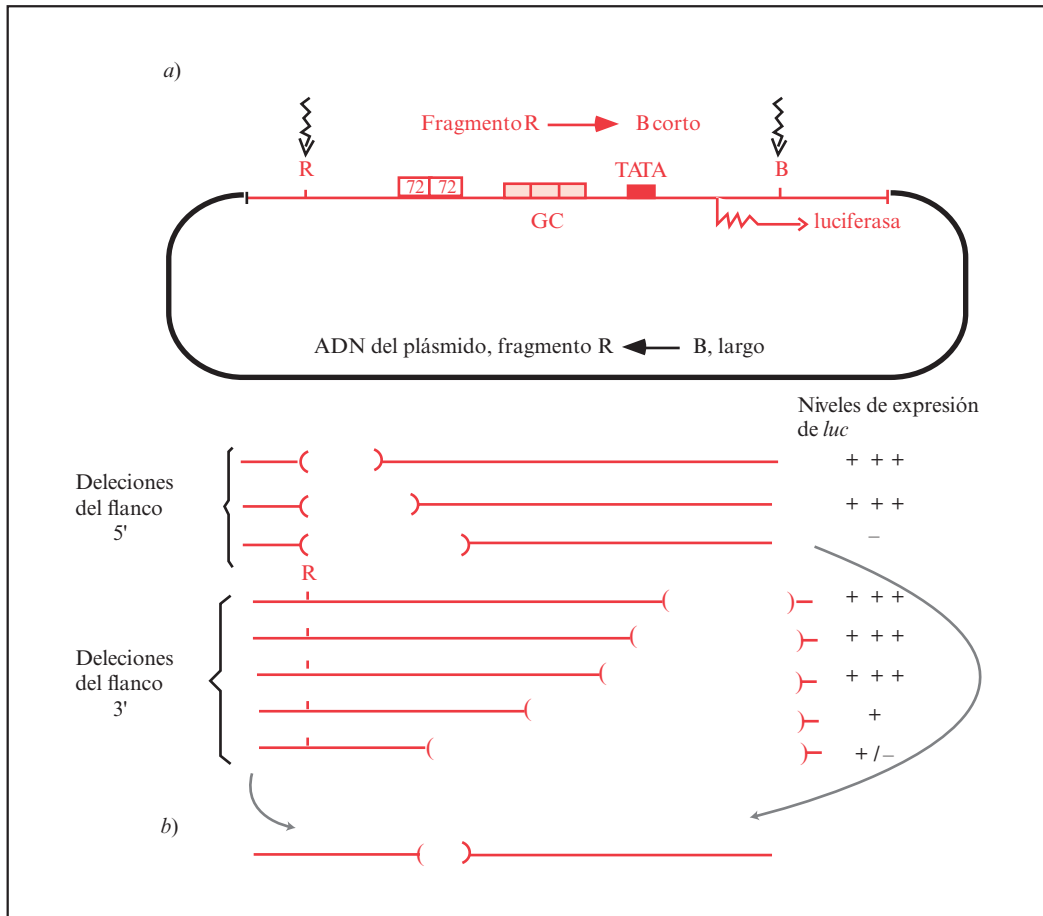


Figura 5.10. Mutaciones internas para el estudio de elementos de control. En la parte superior se representa el mapa del promotor de la región temprana de SV40 y los elementos de control cuya influencia en los niveles de transcripción del gen señal *luc* se desea estimar. a) El tratamiento con la enzima de restricción EcoRI (R) y la exonucleasa Bal31 origina una serie de deleciones en el flanco 5' (fragmento R → B corto), que se unen a un fragmento B → R largo intacto. La transformación de células de mamíferos con estas construcciones nos permiten medir los niveles de transcripción del gen *luc* controlado por esta región promotora. De modo análogo se crean las deleciones en el flanco 3' al tratar con la enzima de restricción Bam (B) seguida de Bal 31. b) Combinación de dos deleciones flanqueantes para obtener una deleción interna al unir una deleción en el extremo 5' que aporta el extremo 3', con otra en el extremo 3' que aporta el extremo 5' (flechas grises). Los huecos representan las regiones eliminadas.

Las deleciones en la región 5' afectan a uno o los dos elementos de 72 nucleótidos, mientras que las de la región 3' afectan a la secuencia TATA, y a las regiones de 21 nucleótidos próximas a ella. Con estas series de deleciones parciales se transforman células eucarióticas y se mide la expresión de luciferasa. Se observó que cuando faltan los dos elementos de 72 nucleótidos se afecta seriamente la transcripción por lo que son potenciadores; en cambio, los efectos de las deleciones de los 21 pares de bases no son tan drásticos. También se pueden elaborar deleciones internas combinando deleciones de ambos flancos. Deleciones en la región 5' (que aportará la región 3') con deleciones en la región 3' (que aporta la región 5') de modo que se elimina exclusivamente la región central (figura 5.10). Hay versiones de este método en las que los segmentos eliminados de la región promotora se sustituyen por otros para que las distancias concretas se mantengan intactas; es análogo al método anterior, pero con mayor número de deleciones para combinar. En este caso se producen deleciones de unos 10 nucleótidos que se rellenan con un adaptador sintético que contiene una diana de restricción y algún nucleótido más, hasta conseguir el mismo tamaño que la deleción.

5.5.2. Estrategia de mutaciones puntuales en secuencias clonadas vía M13

La región que se desea mutar debe clonarse en el fago M13 por ser un fago de banda simple que pasa por una forma replicativa de banda doble. Se obtiene comercialmente, asimismo, un oligocebador que lleve la mutación puntual deseada. Este oligocebador iniciará la síntesis *in vitro* de la banda (–) en presencia de polimerasa, ligasa y los cuatro nucleótidos trifosfato. La forma replicativa sintetizada se utiliza en ensayos de transfección en *E. coli*, que en presencia de un M13 auxiliar originará fagos (+) portadores de la mutación por haber utilizado como molde la (–) que lleva el nucleótido alterado (figura 5.11). También se producirán fagos tipo silvestre (auxiliar) que no suelen

interferir con los que llevan inserto. La utilización de estos M13 (+) en una segunda infección *in vivo* nos permite obtener formas replicativas de ADN de doble hebra con la mutación en las dos hebras.

5.5.3. Obtención de mutaciones puntuales por PCR

En cualquier secuencia de ADN, los oligonucleótidos que sirven de cebadores pueden llevar algún cambio en su secuencia que será incorporado en el extremo del producto amplificado. Una variación de este método permite introducir una mutación en cualquier lugar de una secuencia clonada. Para ello transformaremos una PCR en dos y utilizaremos la región a mutar como oligocebadores internos solapantes (figura 5.12). Si, por ejemplo, queremos mutar un nucleótido en la región central de un gen conocido, planificaremos realizar dos PCR separadas simultáneas: una que nos amplifique desde el extremo izquierdo hasta la mutación y la otra que lo haga desde el extremo derecho también hasta la mutación. Se utilizarán cuatro oligocebadores, dos externos y los otros dos solapantes y portadores de la mutación. Los productos de estas reacciones simultáneas se desnaturalizan, se mezclan y se renaturalizan por las partes solapantes. Las moléculas con extremos 3'OH se terminarán de sintetizar con la polimerasa y podrán ser amplificadas posteriormente con los oligocebadores externos de izquierda y derecha. Estos mismos oligocebadores aportarán extremos 3'OH para la extensión de los híbridos complementarios (5'-P) y su amplificación. Hay una variante de este método en la que primero amplificamos sólo uno de los lados (izquierdo o derecho) y el producto amplificado sirve de megacebador para la amplificación del producto final (figura 5.13).

Un inconveniente de los métodos basados en la PCR es que la mayoría de las polimerasas termorresistentes carecen de actividad correctora de errores pudiendo introducir alteraciones adicionales a las planificadas, por lo que conviene secuenciar los productos obtenidos.

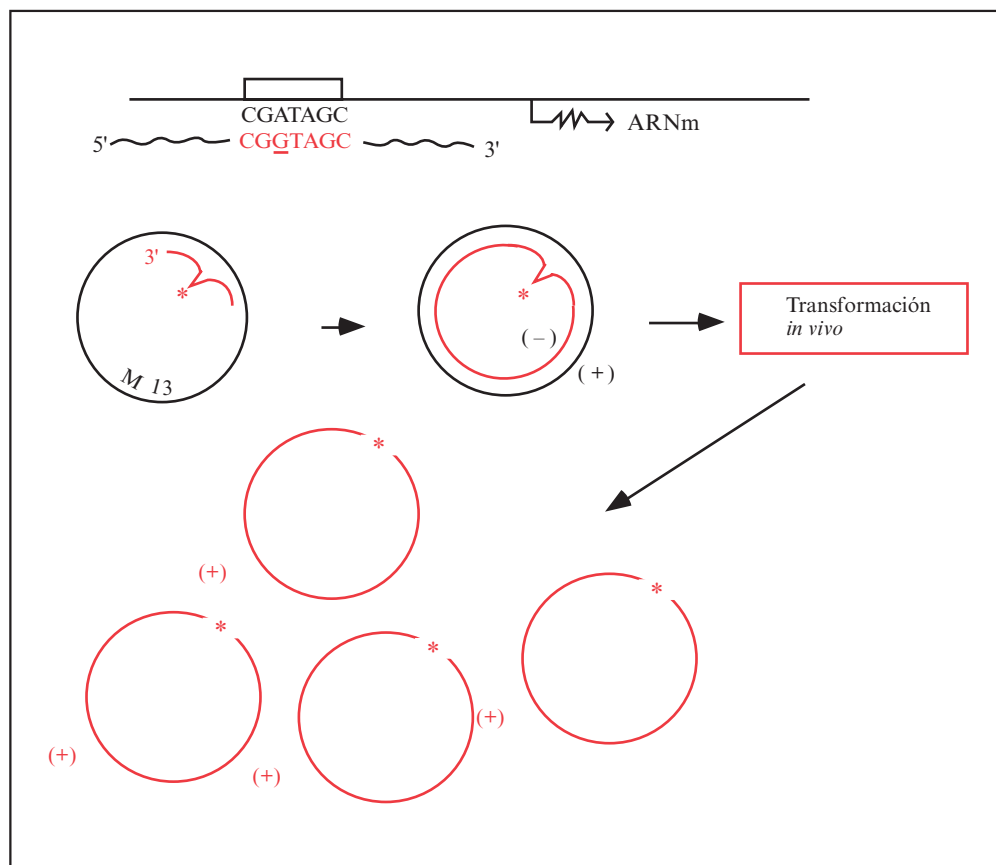


Figura 5.11. Mutaciones con oligocebadores y M13. Utilización de un pequeño oligocebador obtenido comercialmente correspondiente a la región que se desea mutar en la que se ha cambiado un único nucleótido A/G; apareamiento con M13 y síntesis de la banda complementaria (-). La forma replicativa creada *in vitro* es transferida a *E. coli* en presencia de un fago M13 auxiliar. Se obtendrán fagos M13 que habrán incorporado la mutación deseada.

Construcción de genes quimera

La reacción en cadena de la polimerasa puede también utilizarse como método de obtención de alteraciones en determinados genes al unir partes de dos genes distintos en una misma unidad. Hay muchas situaciones en las que la construcción de genes quimera puede ser deseable, por ejemplo, para facilitar su purificación o para crear funciones adicionales a las ya realizadas por una proteína. En estos casos se utilizan oligocebadores mixtos obtenidos comercialmente con secuencias de

ambos genes y dos extremos 3'OH (figura 5.14). Se amplifica primero la parte correspondiente al gen A, se desnaturaliza el producto y se utiliza como megacebador para la amplificación de la parte que interese del gen B.

5.6. MUTAGÉNESIS *IN VIVO*

Existen varios métodos de introducir una mutación en un gen concreto *in vivo*. En este apartado analizaremos algunos de estos métodos.

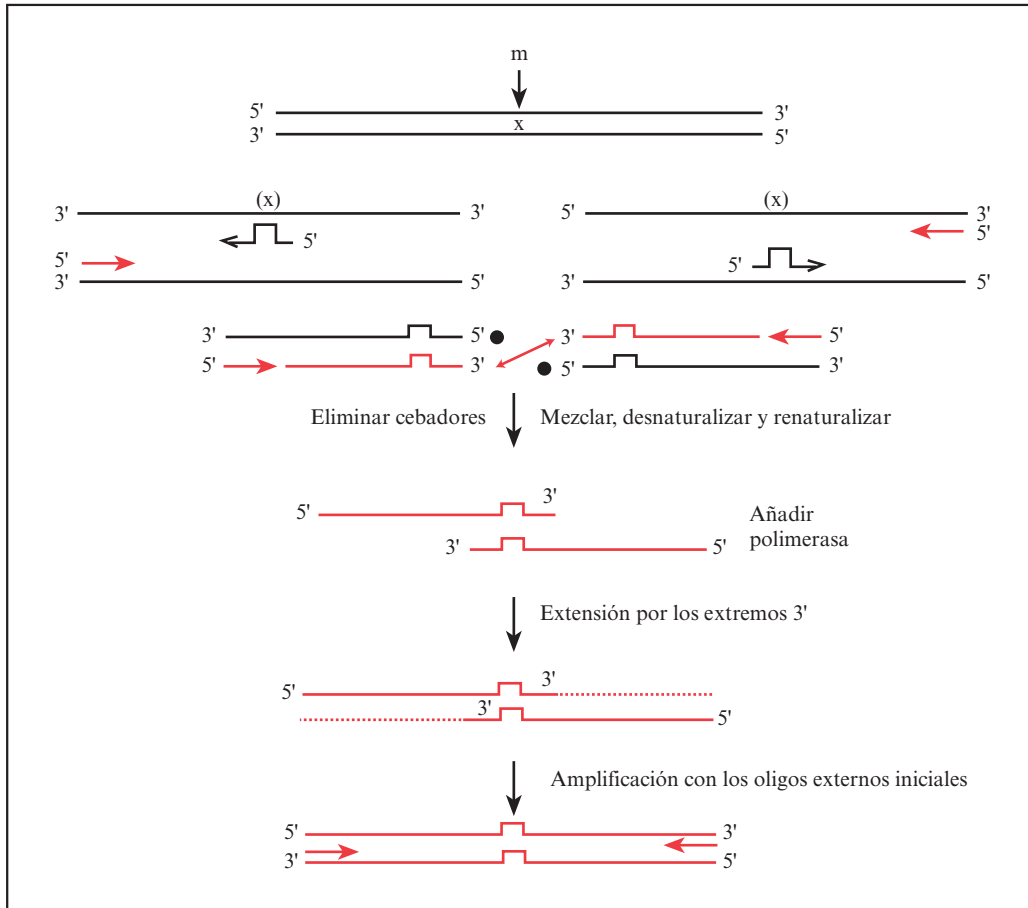


Figura 5.12. Mutagénesis mediante reacción en cadena de la polimerasa. Si deseamos producir una mutación en el centro del ADN que se muestra en la parte superior (x), procederemos del modo siguiente: se parte de dos reacciones de PCR (derecha e izquierda) y dos oligocebadores externos (flechas); también se sintetiza un oligo en el que hayamos introducido un nucleótido erróneo que constituirá la mutación (flecha quebrada en la figura). Las dos reacciones de PCR se llevan a cabo por separado, pero sus productos se juntan para formar híbridos con dos extremos 3'OH que podrán ser extendidos y amplificados utilizando los oligocebadores exteriores de la primera reacción.

Mutagénesis por recombinación homóloga: dianas génicas

Uno de los métodos mejores y más tradicionales para determinar la función de un gen es identificar o producir una mutación y ver el efecto que ello genera en la célula o el organismo. La mutagénesis por recombinación homóloga permite sustituir cualquier gen por una versión truncada del

mismo gen. Cuando esto ocurre hablamos de una diana génica (figura 5.15). La forma más sencilla de modificar la copia del gen transferido es mediante un gen de selección positiva, como *neo* medido en su interior (vector de reemplazamiento). El gen de selección confiere resistencia a los antibióticos neomicina, kanamicina y G 418, todos ellos aminoglicósidos que se unen a determinadas proteínas ribosomales bloqueando la síntesis de proteí-

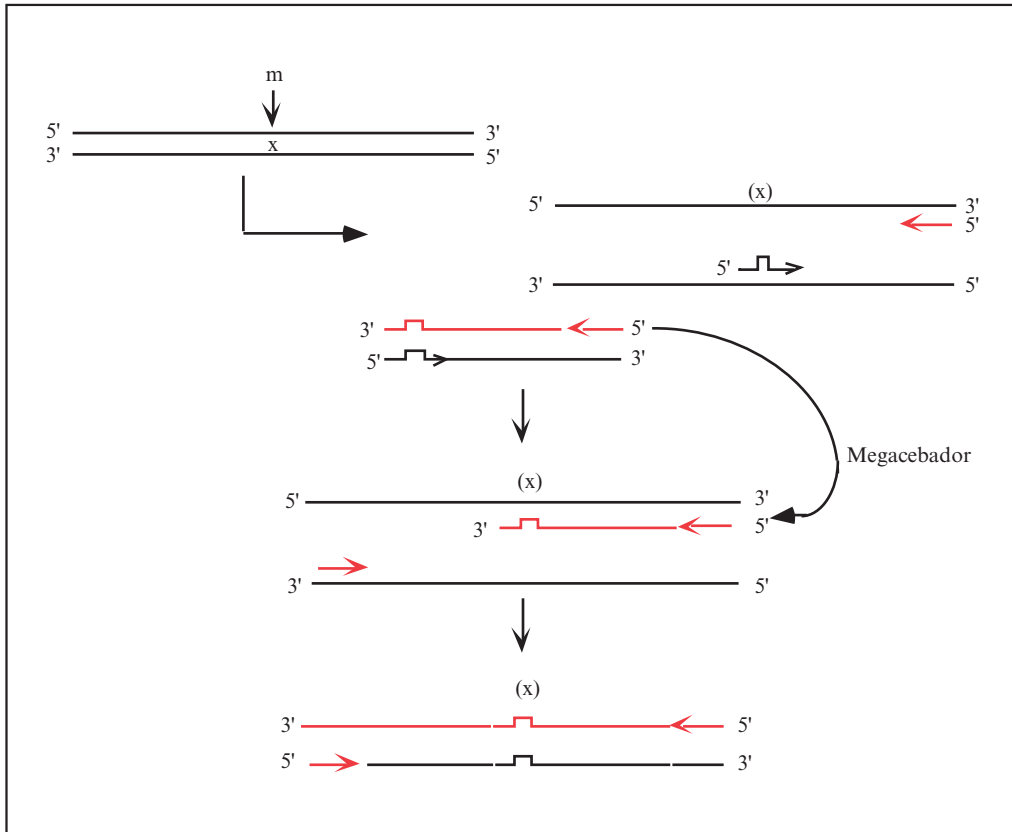


Figura 5.13. Variante de mutagénesis por PCR utilizando el producto de una primera amplificación de la parte derecha como megacebador del fragmento total.

nas. El gen *neo* codifica la enzima aminoglucósido-fosforribosil-transferasa, que inactiva el antibiótico. En sistemas bacterianos la neomicina es el antibiótico más eficaz; en mamíferos en cambio es el G418, y en plantas, la kanamicina. La construcción que se introduce en las células posee además el gen de la *quinasa de timidinas (tk)* virus *Herpes simplex* (VHS) en uno de sus extremos.

El tipo de selección aplicado en estos experimentos es doble: por una parte, se selecciona positivamente por la presencia del gen *neo*, y, por otra, seleccionaremos negativamente por la ausencia del gen *tk* del virus *Herpes simplex*, lo que nos indicará que la recombinación ha tenido lugar en la diana y no al azar. La enzima Tk codifi-

cada por el VHS no es tan estricta como la de la célula y puede incorporar análogos de nucleósidos como el aciclovir o ganciclovir (análogo de la guanina) que son tóxicos letales para las células en división. El vector recombinante, una vez en el interior del núcleo, tiene dos opciones, la integración en el primer lugar accesible o la búsqueda de su homólogo. Si la recombinación es al azar, ésta normalmente tiene lugar por los extremos del fragmento y, por tanto, incluirá el gen de la *tk* del VHS (figura 5.15); si, por el contrario, la recombinación ocurre en el locus diana, el gen de la *tk* de VHS será excluido.

El método aquí descrito de recombinación homóloga también puede utilizarse para crear muta-

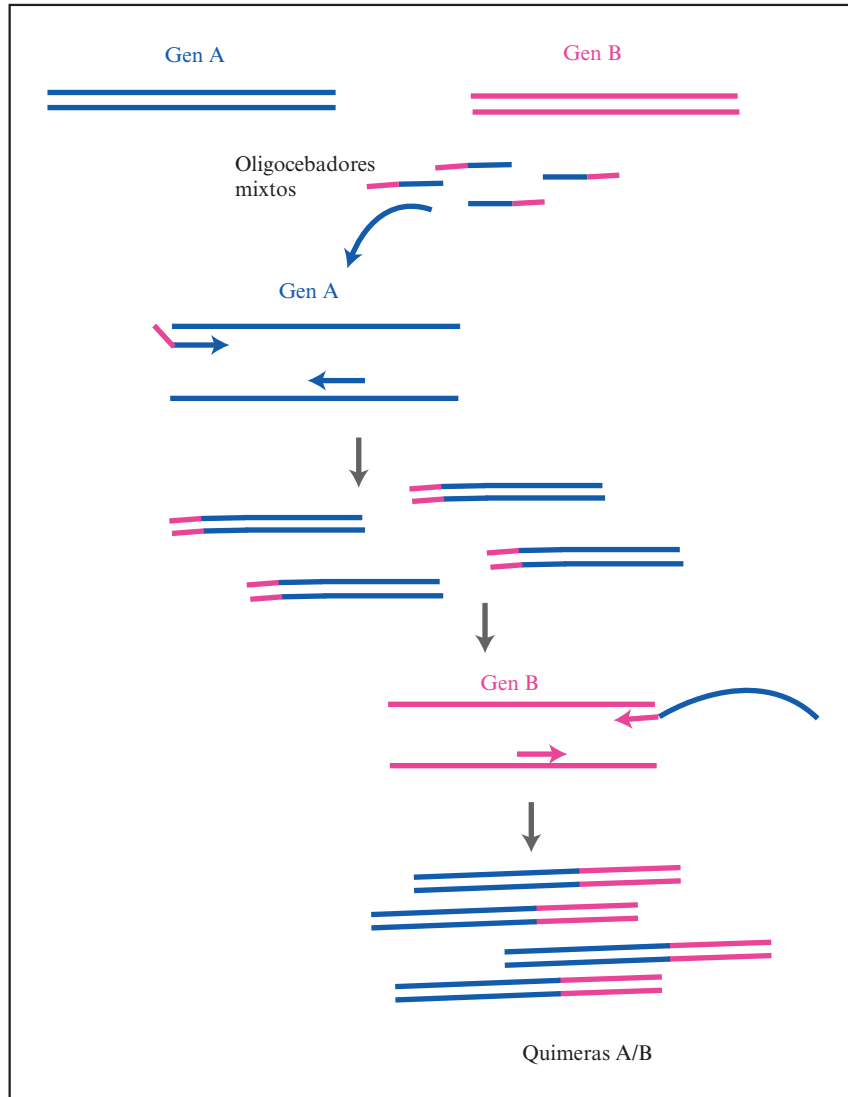


Figura 5.14. Construcción de proteínas mutantes quiméricas por PCR. Los oligonucleótidos que se emplean para la amplificación del gen A son mixtos y llevan una pequeña región complementaria al gen B. El producto de la primera amplificación se desnatura e hibrida con el gen B; se amplifica seguidamente la parte correspondiente al gen B utilizando como megacebador el producto de la primera amplificación.

ciones puntuales o deleciones muy pequeñas en genes homólogos. En estos casos no tenemos selección positiva dentro del gen y dependemos de la negativa que se encuentra fuera de la región de homología.

La recombinación homóloga (o dianas génicas) se produce con una frecuencia aproximada de una por cada 3×10^6 células de mamíferos transfectadas. Esta baja frecuencia puede incrementarse dependiendo, no tanto de la concentra-

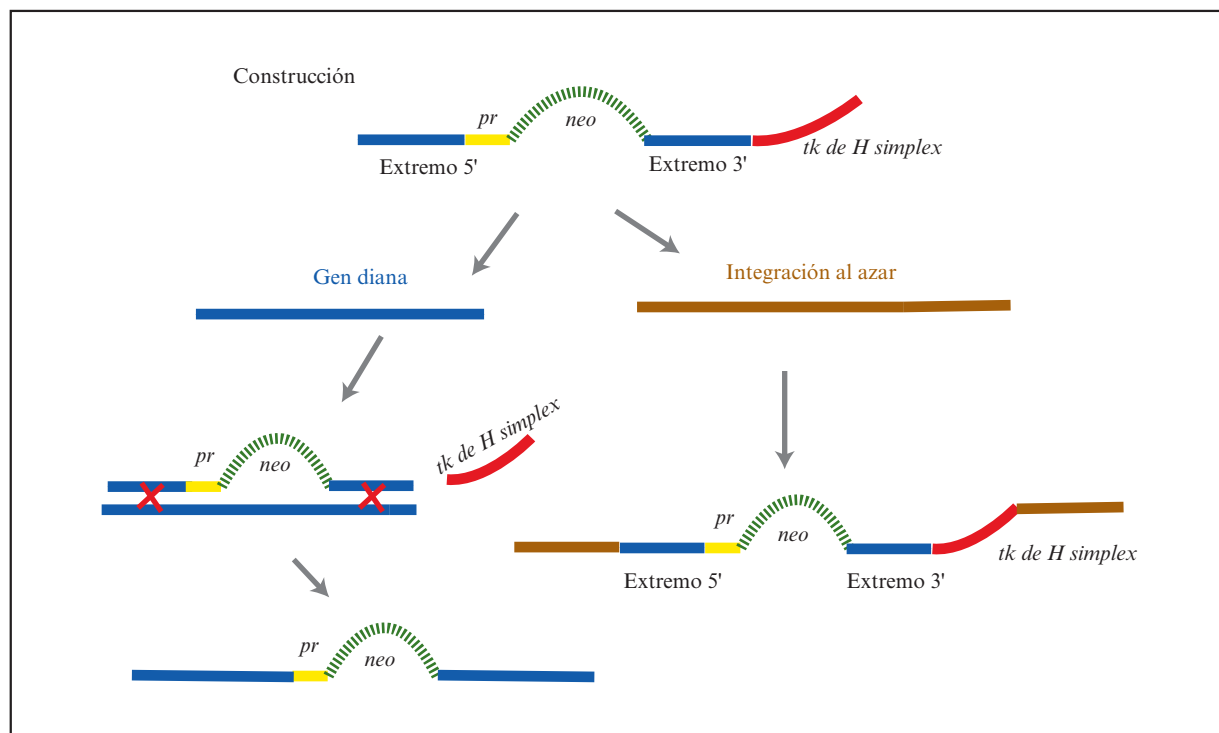


Figura 5.15. Estrategia para modificar genes por recombinación homóloga (diana génica). El gen *neo* y un promotor (*pr*) sustituyen una parte del gen diana y constituyen la selección positiva del evento. En versión esquemática se representan las diferencias de integración en la recombinación al azar y la homóloga. La integración al azar (a la derecha) incluirá el gen *tk*, las células serán Tk^+ y morirán en presencia de ganciclovir. La recombinación homóloga (a la izquierda) convertirá las células en deficientes para la *Tk* del virus *Herpes simplex* con lo que serán resistentes al ganciclovir.

ción del vector de reemplazamiento, como de la longitud de la región homóloga. Varios grupos de investigación independientes han verificado que el incremento de la región de homología de 2 a 6 kb aumenta la frecuencia de dianas génicas en, aproximadamente, 4 veces. En células de mamíferos, excepto en los cromosomas sexuales, tendremos dos copias del gen a mutar, con lo que requeriremos dos sucesos de recombinación homóloga, uno en cada homólogo para obtener una célula con inactivación total de expresión (*knock-out*). La obtención de dianas génicas por recombinación homóloga también se ha aplicado para mutagenizar extensas regiones clonadas en cromosomas artificiales de bacterias (300 kb) procedentes de organismos con pocos mutantes natu-

rales accesibles para su estudio. El sistema se ha empleado con éxito en el pez cebra, donde se han obtenido mutaciones en diversos genes y regiones reguladoras previamente clonadas en un cromosoma artificial. La copia modificada lleva, además de la interrupción por el gen de selección *neo*, dos segmentos *chi* (5'GCTGGTGG 3') convenientemente orientados al inicio y al final del segmento de homología. Como se ha mencionado en el capítulo 2 (apartado 2.1.1), las secuencias *chi* estimulan la recombinación homóloga en bacterias mediada por el sistema RecBCD (figura 2.5). Con este sistema se han obtenido recombinaciones muy precisas en el lugar deseado sin incrementar las reordenaciones en otros lugares del cromosoma artificial.

Mutagénesis por quimeras ADN-ARN

La técnica se basa en la introducción en la célula de una quimera: ADN-ARN de hebra sencilla circular, que lleve la secuencia diana con la mutación deseada en la parte de ARN (figura 5.16). El método se basa en la alta eficiencia con que se forman complejos ADN/ARN desplazando a los ADN-ADN endógenos. El apareamiento favorece la recombinación homóloga en la que la molécula de ARN sustituye a la de ADN idéntica. El cambio de un codón funcional por uno de terminación (stop) nos dejará fuera de combate la función del gen por KO (del inglés, *knock-out*) siempre que sea único o que ocurran tantas dianas génicas como múltiplos haploides del genoma. Después de una serie de reparaciones y divisiones en las que el ARN se convertirá en ADN y se reparará cambiando el nucleótido de la hebra celular original, la mutación estará totalmente integrada.

Mutagénesis por quimeras enzimáticas entre Fok I y nucleasas de dedos de zinc o TALEN

Como se ha descrito en el capítulo 3, apartado 3.2.1 y figura 3.19, las quimeras entre la región de corte de la enzima de restricción Fok I y elementos de reconocimiento de secuencias específicas en el ADN permiten realizar cortes específicos en los genomas de organismos superiores. Los dos tipos de quimeras más populares son las **nucleasas de dedo de zinc** y las **quimeras TALEN**. En ambos casos se produce un corte mediado por la enzima Fok I en las dos hebras de un ADN de cualquier procedencia, en el lugar preciso estipulado por los dedos de zinc o los módulos TAL.

Se ha observado que, como consecuencia de este corte, al menos en células de mamíferos se desencadenan dos posibles mecanismos de reparación: 1) el inicio de una recombinación homóloga con la cromátida hermana sustituyendo así su propio ADN por el de la otra cromátida, y 2) reparación de los terminales cortados. En este segundo caso la reparación es complicada, ya que se podría iniciar una degradación nucleolítica letal para la célula; para evitarlo, los terminales rotos se unen a factores proteicos especiales que se en-

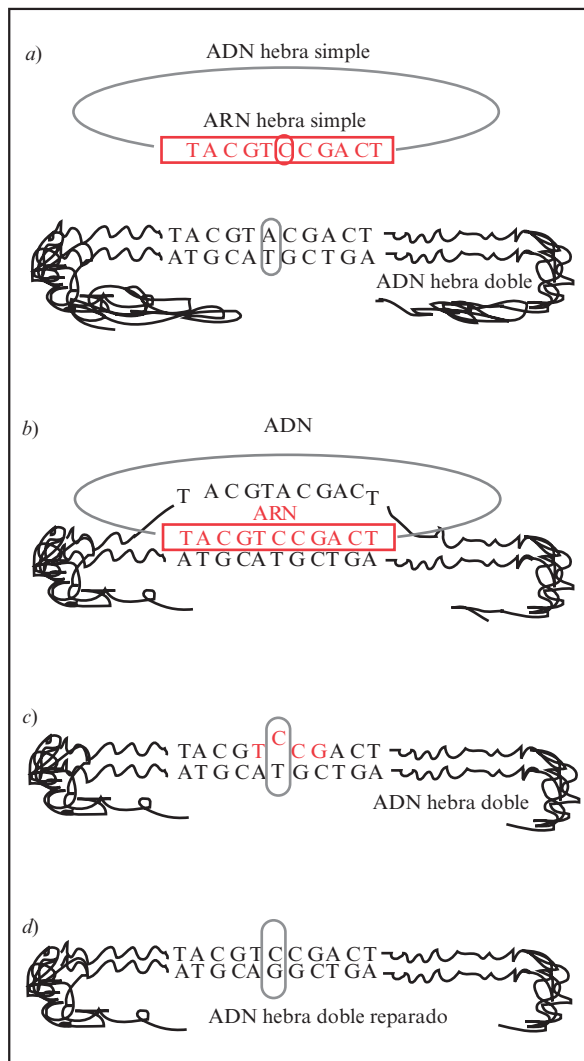


Figura 5.16. Mutagénesis por quimeras ADN-ARN. *a)* Quimera sintética formada por ADN-ARN de banda simple que se introduce en células, y contiene la mutación en la parte de ARN. *b)* Desplazamiento de la banda idéntica de ADN y recombinación. *c)* Resultado de la recombinación y sustrato de la reparación. *d)* Molécula de ADN restaurada.

cargan de acercar y ligar los terminales. El proceso es mutagénico, pues se suelen introducir pequeñas deleciones o inserciones en el lugar de unión de las partes (figura 5.17).

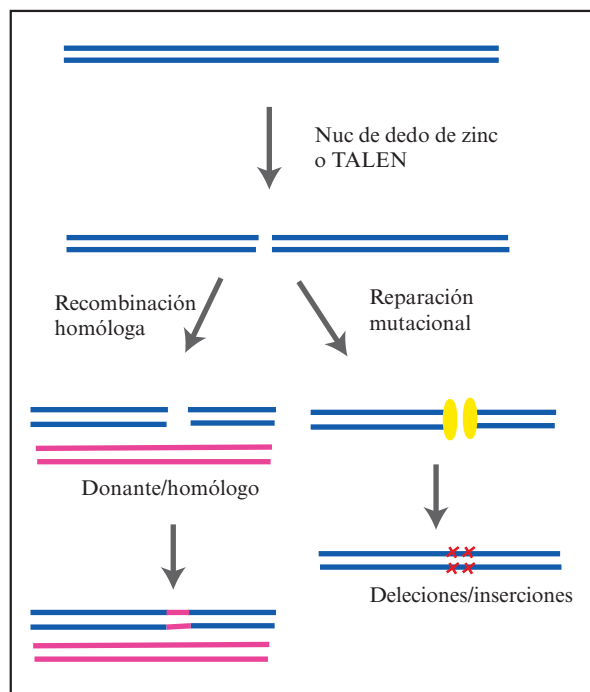


Figura 5.17. Mutagénesis por quimeras de nucleasas de dedo de zinc o TALEN. El dedo de zinc o los módulos TAL reconocerán una secuencia que la enzima de restricción Fok I modificada cortará produciendo una ruptura en las dos hebras de ADN de difícil reparación. Como consecuencia se desencadenará uno de los dos mecanismos siguientes: a la izquierda, recombinación homóloga si en las proximidades hay un ADN donante, y a la derecha, restauración con introducción de pequeñas deleciones o inserciones.

El mecanismo produce dianas génicas por KO y también se puede utilizar para terapias génicas. En enfermedades dominantes, por ejemplo, se podría sustituir el alelo mutado por el normal. También se podría hacer llegar al entorno de la ruptura un ADN donante que se incorporará durante la recombinación homóloga.

Mutagénesis por tijeras químicas

Por último, volveremos a mencionar las tijeras químicas (capítulo 3, apartado 3.1.1, figura 3.3) para cortar un ADN de doble hebra utilizando

pseudo-oligonucleótidos complementarios a las dos regiones adyacentes al lugar de corte; cuando estas regiones aparecen con el ADN correspondiente, dejarán una pequeña región de ADN de hebra simple. En este punto se añade el componente de Cerio/EDTA que romperá en la región de ADN de hebra simple de cada hebra del ADN. El resultado se resolverá como se ha explicado en el caso anterior con dos posibles mecanismos de reparación: la recombinación homóloga con un ADN donante cerca, o la mutagénesis reparadora.

5.7. MUTAGÉNESIS CONTROLADA A NIVEL DE GENOMAS COMPLETOS: BACTERIAS SINTÉTICAS

Cuando la alteración de una pequeña parte del genoma pasa a gran escala, tendríamos la modificación de todo el genoma de un organismo. Se han obtenido bacterias sintéticas por modificaciones parciales ampliadas hasta lograr la sustitución total del genoma de partida. El genoma sintético puede ser idéntico al de partida o, durante el proceso de síntesis, podemos modificarlo a voluntad.

Hasta el momento se han utilizado dos métodos para la obtención de bacterias con el genoma remodelado: 1) el método químico o síntesis de novo, que consiste en sintetizar nucleótido a nucleótido fragmentos de ADN que, a su vez, se van ensamblando en unidades cada vez más grandes hasta conseguir el genoma completo, y 2) el método biológico que consiste en sustituir paulatinamente, por recombinaciones homólogas, el genoma natural por oligonucleótidos sintéticos.

En ambos casos podremos introducir los cambios que deseemos durante el proceso de síntesis. Veremos en detalle un ejemplo del método biológico en el que el genoma natural de *E. coli* se ha sustituido por otro totalmente sintético y equivalente en el que todos los codones «ámbar» de terminación (TAG, 314 en total) se han cambiado por los codones «ocre» (TAA) de función idéntica. El procedimiento consistió en dividir el genoma natural de *E. coli* en unas 32 regiones de 100 kb cada una (100.000 nucleótidos) de modo que

cada región llevara aproximadamente unas 10 secuencias TAG. Para cada división se sintetizaron numerosos oligonucleótidos pequeños (20-50 nucleótidos) que cubrían toda la zona y que poseían los codones stop TAA en sustitución de los TAG. La primera fase consistió en obtener 32 variantes de *E. coli*, cada una con una región concreta sustituida por haber forzado la recombinación homóloga entre los oligonucleótidos sintéticos y los naturales (figura 5.18a). La verificación de la sustitución se realiza por secuenciación. El paso siguiente consistió en favorecer la conjugación y recombinación entre bacterias con regiones sustituidas adyacentes hasta conseguir cuatro cepas en las que un cuarto del genoma estaba ya sustituido y carecían de codones TAG. El paso final fue la conjugación y recombinación de las cuatro cepas, dos a dos, hasta completar la mitad del genoma sustituido en las dos finalistas que dieron lugar, por el mismo procedimiento, a una última cepa de *E. coli* con todo el genoma artificial y ningún codón stop TAG (figura 5.18). Como ya no era necesario, también se eliminó el gen correspondiente a la proteína que lee el codón ámbar. La bacteria resultante es viable y aparentemente normal.

Entre las aplicaciones más interesantes de esta metodología está la de conseguir, por ejemplo, bacterias resistentes a virus, lo que beneficiaría mucho a empresas biotecnológicas que emplean grandes cantidades de bacterias recombinantes para producir un gran número de productos. Una infección viral durante la producción obliga a cancelar la elaboración del compuesto con los consiguientes retrasos y pérdidas asociados. Se podría manipular el genoma de una bacteria de modo análogo al ya explicado, teniendo en cuenta que el código genético está degenerado y que existen varios codones para un mismo aminoácido. Se podría, por ejemplo, eliminar por completo del genoma de una bacteria uno de los codones redundantes del aminoácido prolina (CCC o CCG), sustituyéndolo por el otro (CCC → CCG). Así la célula no necesitaría el aminoacil-ARN de transferencia que lleva CCC como anticodón y que estaría cargado con prolina (figura 5.18b), por lo que se puede eliminar el gen correspondiente. Sin embargo, cualquier virus que infecte la bacteria

necesitará el aminoacil-ARNt que tenga como anticodón CCC y esté cargado con prolina, para llevar a cabo la síntesis de sus proteínas. La infección será abortiva si la síntesis de las proteínas virales está bloqueada.

5.8. MUTACIONES Y CÁNCER

El cáncer es una enfermedad genética que puede tener un fuerte componente epigenético y que se caracteriza por un crecimiento incontrolado de células pertenecientes a un tejido; estas células pueden abandonar el núcleo tumoral inicial y alcanzar nuevos destinos (metástasis).

La enfermedad se produce por la acumulación de mutaciones en el genoma y alteraciones epigenéticas relativas a los niveles de expresión de determinados genes. Dependiendo del tejido donde se produzca originalmente, puede tener características distintas y malignidad dispar. Es una enfermedad multigénica, es decir, que la patología está producida por la alteración de varios genes, pero éstos pueden variar de un caso a otro.

Los genes cuyas mutaciones contribuyen al desarrollo de un cáncer se denominan **oncogenes** y los que de manera natural impiden la proliferación celular se llaman **oncosupresores**. Un cáncer se desarrolla por una combinación de fallos en oncogenes y oncosupresores, tanto a nivel genético como epigenético. Dentro de las modificaciones epigenéticas de oncogenes y oncosupresores que contribuyen al desarrollo de un cáncer, encontramos la metilación del ADN, modificaciones de las histonas y remodelación de la cromatina.

La masa de células que prolifera en exceso respecto a sus homólogas tisulares constituye el tumor primario. Cuando las células de este tumor se mantienen compactas, adheridas entre sí y quizá rodeadas por una membrana basal que las separa del resto de las células, se denomina *tumor benigno* y suele eliminarse por cirugía, sin mayores consecuencias. Los *tumores malignos*, en cambio, van acompañados de nuevas alteraciones que implican fallos adicionales en las proteínas que mantienen la adhesión entre las células, como las integritinas, o exceso de función de metaloproteasas que

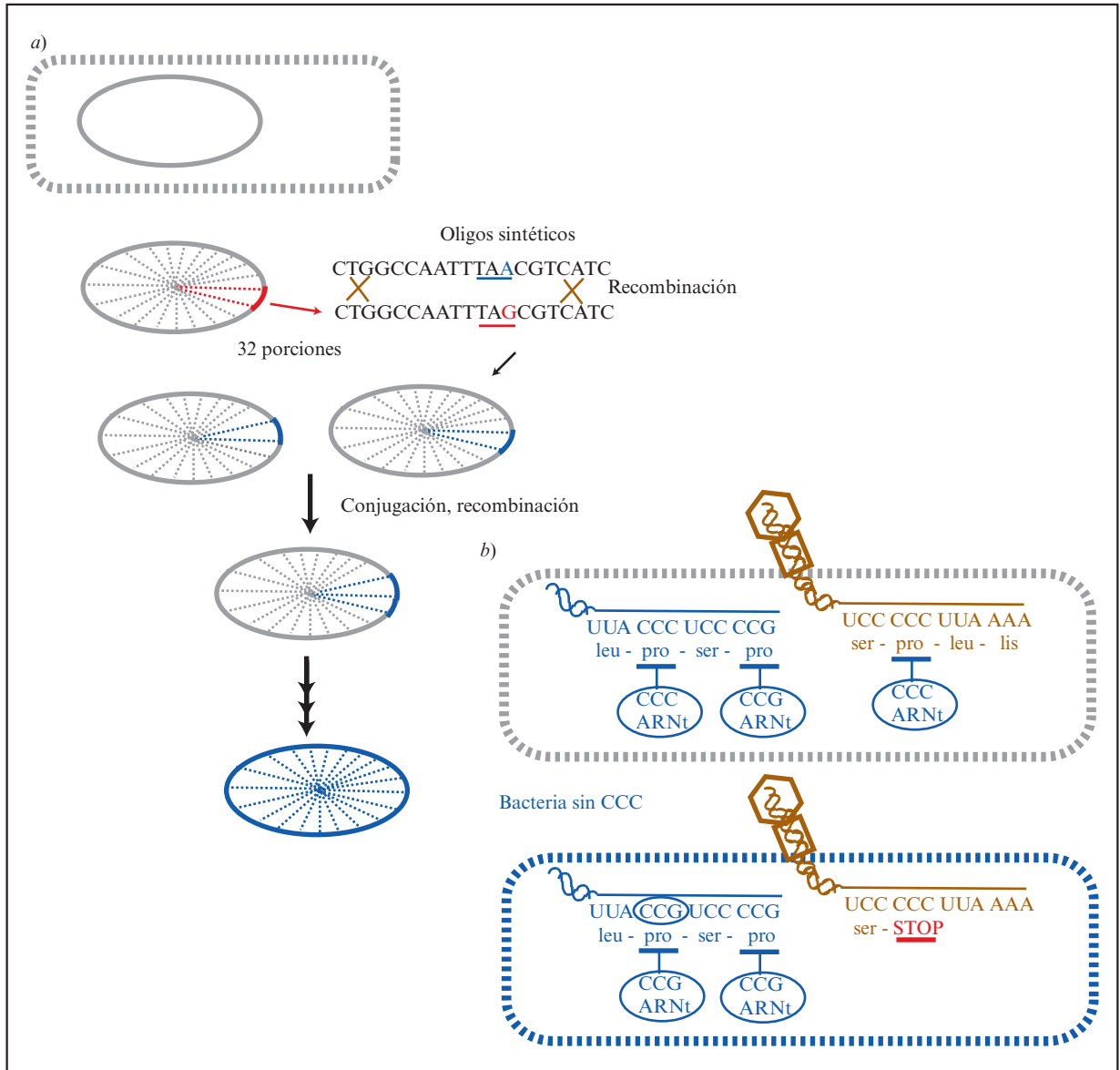


Figura 5.18. Construcción de una bacteria sintética por el método biológico de reemplazamiento de fragmentos por recombinación homóloga y posible aplicación. *a)* El genoma se divide en 32 regiones consecutivas y a cada cepa de *E. coli* se le asigna una región. Cada región se bombardea con oligonucleótidos sintéticos que llevan sustituidos los codones stop TAG por sus equivalentes TAA y que por recombinación homóloga suplantarán al genoma original en la región concreta. Por conjugación y recombinación entre cepas con regiones adyacentes sustituidas se obtendrán 16 cepas con fragmentos sintéticos adyacentes sustituidos. El proceso se repite hasta obtener 8 cepas con un octavo del genoma sustituido primero, 4 cepas con un cuarto del genoma sustituido después y así hasta obtener en pasos sucesivos una bacteria con todo el genoma sustituido. *b)* construcción de una bacteria resistente a virus sustituyendo en el genoma los codones CCC por CCG y eliminando el gen del aminoacil-ARNt cargado con prolina; cualquier virus que infecte la bacteria no podrá sintetizar proteínas y morirá.

degradan los componentes de la matriz extracelular. En muchos tipos de cáncer, como en colon o en cerebro, se puede producir el paso secuencial de un tumor benigno a maligno, lo que requiere tiempo en el que las nuevas modificaciones ocurren. En estos casos se observa que las primeras mutaciones confieren una ventaja proliferativa y pueden ser de dos tipos: conductoras y pasajeras; *las mutaciones conductoras* son esenciales para el incremento de la proliferación, pero *las mutaciones pasajeras* van de acompañantes, probablemente ocurren en regiones del ADN sin una función esencial en el proceso de desarrollo del tumor.

Para que un tumor sobreviva es necesario un aporte de oxígeno y nutrientes, lo que se consigue por la formación de nuevos vasos sanguíneos como respuesta a la secreción de factores angiogénicos por las propias células tumorales. Asimismo, es muy probable que nuestro sistema inmune esté eliminando sistemáticamente posibles células de características distintas a las del tejido donde surgen y que pudieran ser tumorales; fallos en esta vigilancia antitumoral también contribuirían al desarrollo de un tumor. El papel del sistema inmune en cáncer es comprometido, ya que debe eliminar células propias «ligeramente distintas», algo para lo que quizá no esté programado.

Si una de las causas del desarrollo de un cáncer es una acumulación de mutaciones, parece una buena práctica preventiva evitar en la medida de lo posible los agentes mutagénicos.

Test de Ames para detectar agentes mutagénicos

Bruce Ames demostró en el test que lleva su nombre (test de Ames) que una inmensa mayoría (90%) de agentes mutagénicos son carcinogénicos. El test se basa en la utilización de una cepa mutante de la bacteria *Salmonella typhimurium*, auxótrofa para histidina, y determina el porcentaje de revertientes en distintas condiciones o en presencia de determinadas sustancias. Se preparó un homogenado de hígado de rata que se añadió al medio y se probó el efecto de sustancias mutagénicas, carcinogénicas o sustancias sospechosas de serlo. Se observó que en un gran número de casos la proporción de revertientes his^- a his^+ se incrementaba enormemente en las proximidades al lugar donde se había colocado la sustancia mutagénica/carcinogénica (figura 5.19). El suplemento de hígado de rata tiene como misión facilitar todas aquellas enzimas que pueden transformar compuestos precursores en otros verdaderamente carcinogénicos.

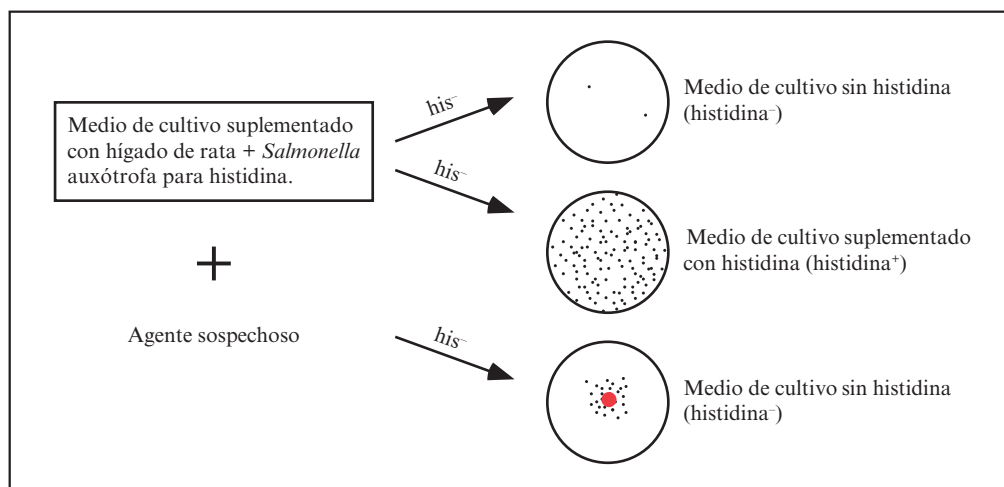


Figura 5.19. Test de Ames. Se determina la frecuencia de revertientes his^- a his^+ en presencia de un agente sospechoso de ser carcinogénico colocado en el centro del medio de cultivo.

El paisaje genómico humano del cáncer

Un estudio global sobre el genoma humano en el que se han secuenciado y analizado más de 100 tumores de varios tipos concretos de cáncer ha puesto de manifiesto una serie de características novedosas de gran interés.

En tumores sólidos de colon, mama, cerebro o páncreas se encuentra una media de 33 a 66 genes alterados portadores de mutaciones en su mayoría puntuales. Por el contrario, los melanomas y tumores de pulmón acumulan muchas más mutaciones (unas 200) quizá debido al efecto que sobre estos tejidos tienen potentes agentes mutagénicos como son la luz ultravioleta y el tabaco, respectivamente. En consonancia con estos datos, se ha observado que los tumores de fumadores tienen diez veces más mutaciones que los de no fumadores. Asimismo, los tumores de cualquier tipo, con mutaciones en alguna de las enzimas de reparación del ADN o mutaciones en la actividad correctora de errores de las ADN polimerasas, suelen acumular miles de mutaciones. Por el contrario, los tumores pediátricos y las leucemias suelen tener una media de 9,6 mutaciones puntuales por neoplasia.

En la mayoría de los tumores sólidos también se han detectado docenas de translocaciones, pero estas alteraciones parecen ir más bien de pasajeras que de conductoras, lo que parece indicar que las rupturas ocurren con mayor frecuencia en regiones no relacionadas con proteínas implicadas en la proliferación celular.

Hemos distinguido entre mutaciones conductoras y pasajeras dependiendo de que originen

una ventaja proliferativa a la célula o no, respectivamente; debemos aclarar que cualquier oncogén puede albergar los dos tipos de mutaciones, por lo que en algunos casos funcionará como conductor y en otros como oncogén pasajero. En el paisaje genómico humano del cáncer, se definen como «montañas» los oncogenes que acumulan un gran número de mutaciones y supuestamente funcionan como conductores; se definen como «colinas» los oncogenes que acumulan un número menor de mutaciones, pero que también pueden albergar mutaciones conductoras. En la mayoría de los cánceres nos topamos con más colinas que montañas. Cuando un oncogén o un oncosupresor no está mutado, pero contribuye al desarrollo de un cáncer por modificaciones epigenéticas en las que se altera sustancialmente la expresión debido a metilaciones o modificaciones de la cromatina, también se le considera conductor, aunque para distinguirlo se le suele llamar epi-conductor. Un recuento de todos los genes conductores encontrados en el estudio realizado arroja un valor de 125 genes conductores distintos, de los cuales 71 son oncosupresores y 54 oncogenes. Así, volviendo a los datos sobre número de mutaciones en tumores sólidos de colon, mama y cerebro mencionados al principio del apartado, ahora puntualizaremos que la media de genes conductores en este tipo de tumores es de 3 a 6; siendo 0 para los tumores pediátricos. Mucho más complicado es determinar el número de epi-conductores, ya que el estado de metilación de un gen puede variar dependiendo del tipo de célula o de la edad del paciente.

LECTURAS COMPLEMENTARIAS

Cancer genome landscapes. Artículo de revisión. Bert Vogelstein, Nickolas Papadopoulos, Victor E. Velcu-

lescu, Shibin Zhou, Luis A. Díaz, Kenneth W. Kinzler. *Science*, vol. 339, 1546-1558. 2013.

Las levaduras mantienen las ventajas de los microorganismos en cuanto a que son unicelulares, de fácil manipulación y crecimiento rápido. Su organización celular es, sin embargo, eucariótica, permitiendo la realización de procesos de expresión y maduración característicos de células animales y vegetales.

6.1. CLONACIÓN EN LEVADURAS: *SACCHAROMYCES CEREVISIAE*

Las levaduras, en su calidad de eucariotas, pero a la vez unicelulares, representan un puente biológico entre bacterias y organismos superiores. Son la alternativa natural en la producción a escala industrial de proteínas con interés farmacológico para terapia humana. Las levaduras son capaces de llevar a cabo procesos que no realizan las bacterias y que pueden ser importantes para la expresión de un gen eucariótico en un sistema distinto. Realizan acetilaciones N-terminales, N-glicosilaciones, fosforilaciones, eliminación de la metionina inicial en la proteína madura, secreción proteica, procesamiento proteolítico, y permiten la unión y ensamblaje de subunidades en macromoléculas. Además, son fáciles de cultivar y baratas de mantener. Por todo esto las levaduras podrían utilizarse para la expresión de proteínas humanas con interés biotecnológico. Pero las levaduras presentan limitaciones en cuanto a la eficiencia con la que realizan estas funciones sobre proteínas procedentes de organismos evolutivamente lejanos. El patrón de glicosilación puede no

ser totalmente fiel al original, la eliminación de la metionina puede no alcanzar a todas las moléculas, etc.

Las levaduras también nos permiten clonar fragmentos de ADN de gran tamaño. Insertos de 100 a 2.000 kb (2 megabases; mgb) de cualquier organismo pueden mantenerse en los vectores YAC (del inglés: *Yeast Artificial Chromosome*), que describiremos con detalle en este capítulo. Esta posibilidad tiene singular importancia cuando interese aislar un gen completo que posea muchos intrones y esté muy extendido en el genoma original.

Las levaduras constituyen los vehículos de clonación eucarióticos más versátiles y en los que se han llevado a cabo manipulaciones más sofisticadas.

Saccharomyces cerevisiae es la levadura mejor estudiada. Su nombre se refiere a su papel en la fermentación de la cerveza, pero también interviene en la fermentación de frutas en vinos y en la elaboración del pan. Hay, pues, una larga tradición en la explotación de este organismo por industrias biotecnológicas. En realidad, uno de los primeros productos comerciales elaborado con técnicas de ingeniería genética fue una vacuna contra el virus de la hepatitis B y el proceso se llevó a cabo en *S. cerevisiae*.

Saccharomyces es uno de los pocos organismos eucarióticos que posee un plásmido natural no vírico, denominado el «círculo de 2 μ m», que ha facilitado enormemente la construcción de vectores y sistemas de expresión. Su genoma está secuenciado y todos sus genes han sido identifica-

dos. Éstos pueden ser directamente sustituidos y mutagenizados y, además, se pueden aislar estructuras cromosómicas, como centrómeros o telómeros, para manipularlos y formar cromosomas artificiales de gran utilidad en la clonación de grandes fragmentos de ADN.

S. cerevisiae es un organismo unicelular, de forma elíptica, con un espacio periplasmático entre la membrana citoplasmática y la pared celular externa. Su genoma tiene 12.000 kb (kilobases) de ADN organizadas en 16 cromosomas portadores de 5.500 genes. Los cromosomas son separables por electroforesis en campo pulsante. Este tipo de electroforesis fuerza a las moléculas de ADN a cambiar la dirección en la que se están moviendo al someterlas a campos eléctricos alternantes con intervalos de tiempo muy cortos. Las moléculas pasan la mayor parte del tiempo reorientándose; la movilidad es muy pequeña y prácticamente independiente del tamaño de las moléculas.

Aunque la mayoría de los procesos celulares en levaduras son muy parecidos a los de células de mamíferos, hay algunas diferencias notables. La glicosilación, por ejemplo, es más simple que en mamíferos. En levaduras sólo se añaden residuos de manosa formando estructuras ramificadas en asparraginas, serinas o treoninas; en mamíferos, las estructuras ramificadas pueden tener hasta 150 residuos e incluir una gran variedad de azúcares complejos. Otra diferencia radica en la permanencia de la membrana nuclear durante la mitosis, mientras que en mamíferos se destruye. Por último, las secuencias consenso que determinan el borde 3' de los ajustes es distinto en levaduras, con lo que la maduración de ARN de organismos superiores en éstas no es muy precisa.

El ciclo vital de *S. cerevisiae* tiene una fase haploide y otra diploide, ambas estables, siendo el tiempo de generación de una hora y media en buenas condiciones de cultivo. Las células se dividen por gemación. Células haploides de distinto tipo de apareamiento (a o α) pueden unirse para formar diploides (figura 6.1). El apareamiento está controlado por el locus MAT (del inglés *mating type*), situado en el cromosoma 3. En células haploides, si se expresa el alelo α del gen MAT, la célula será alfa (α), y si se expresa el alelo a será

«a». En realidad, MAT controla la expresión de los genes específicos de α o a. Estos genes son, por ejemplo, el MF α que produce el factor alfa, o el MFa que produce el factor a. Alfa y a son pequeños péptidos semejantes a hormonas que se excretan promoviendo el apareamiento con las células de tipo contrario. Alfa estimulará a las células de tipo a, y viceversa. La unión de células α/a produce células diploides que, bien pueden propagarse indefinidamente por mitosis, o pueden ser inducidas a entrar en meiosis y a esporular en un medio deficiente en compuestos nitrogenados. La esporulación origina cuatro esporas haploides reunidas en un ascus, que pueden aislarse y producir colonias individuales, pudiéndose analizar separadamente los productos directos de la meiosis.

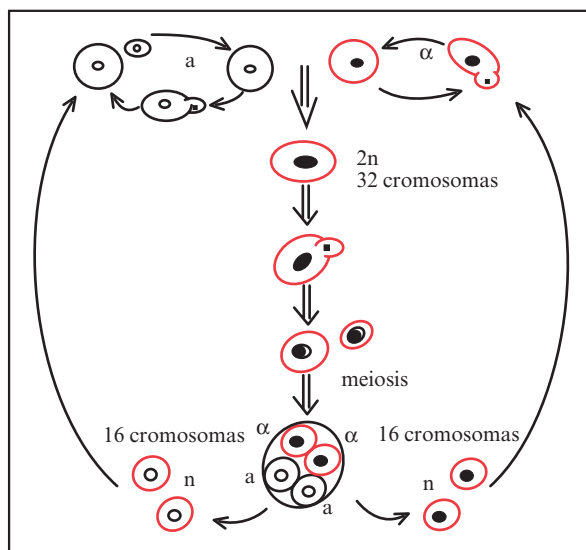


Figura 6.1. Ciclo vital de *S. cerevisiae*. Las células haploides de tipo «a» o « α » se fusionan para formar diploides; éstas, a su vez, especialmente en condiciones adversas, entran en meiosis, esporulan y originan esporas nuevamente haploides.

6.2. MÉTODOS DE TRANSFORMACIÓN Y SELECCIÓN EN LEVADURAS

La transformación en levaduras se lleva a cabo por tratamiento con sales de litio seguido de

una solución de PEG 4000 (polietilenglicol) que favorece la entrada del ADN y un choque térmico para facilitar el proceso. El procedimiento en sí es mutagénico; y, por tanto, si se introduce un gen seleccionable, es importante verificar que la reversión a tipo silvestre ha sido debida a la transformación y no a una segunda mutación.

La selección de levaduras que han adquirido un ADN recombinante frente a las que no lo han hecho es relativamente sencilla, puesto que se conocen cepas mutantes o auxótrofas y marcadores seleccionables que se incorporan al vector y suplen las carencias o mutaciones. En la tabla 6.1 se muestran los marcadores más usados para la selección. También se utilizan marcadores dominantes que se manifiestan sobre un genotipo sil-

vestre, como los genes que inducen resistencia a antibióticos. Por ejemplo, el gen de la enzima cloranfenicol acetiltransferasa confiere resistencia al cloranfenicol (*cf^R*) y el de la neomicin-fosforribosil transferasa (*neo*) al antibiótico G418 (véase capítulo 4, apartado 4.2.1). Este tipo de marcadores se utiliza preferentemente para transformar cepas de *Saccharomyces* de interés industrial, ya que suelen ser poliploides, y las variantes auxotróficas son muy difíciles de aislar. El gen de selección determinante de la puromicina-acetiltransferasa (*puro*) no es válido para conferir resistencia a puromicina en el sistema de levaduras, ya que el antibiótico es automáticamente expulsado de la levadura en cuanto entra, confiriendo al organismo una resistencia natural al compuesto.

TABLA 6.1
Mutaciones de genes de levaduras utilizadas en sistemas de selección y gen neo

Gen*	Enzima	Mutación en <i>E. coli</i>
<i>HIS3</i>	Imidazol glicerolfosfato IGP dehidratasa	<i>hisB</i>
<i>TRP5</i>	Triptófano-sintetasa	<i>trpAB</i>
<i>LEU2</i>	β-isopriomalatodehidrogenasa	<i>leuB</i>
<i>URA3</i>	Orotidín 5'fosfato decarboxilasa	<i>pyrF</i>
<i>neo</i>	Neomicina-fosforribosil transferasa	<i>neo</i>

* En levaduras el gen silvestre se escribe, con frecuencia pero no siempre, con mayúsculas y cursiva, mientras que el gen mutante se suele escribir con minúsculas y cursiva.

6.3. TIPOS DE VECTORES EN LEVADURAS

En levaduras, careciendo de vectores de origen vírico, podemos distinguir dos tipos de portadores génicos:

- Vectores de integración cromosómica.
- Vectores de replicación autónoma. Éstos, a su vez, se dividen en:
 - Vectores de clonación: autónomos basados en componentes cromosómicos y

cromosomas artificiales (CAL, en versión anglosajona YAC).

- Vectores de expresión: derivados del plásmido de 2 μm y vectores mixtos.

Todos estos vectores ofrecen la posibilidad de amplificar el vector en procariotas antes de ser introducidos en la levadura. El sistema de levaduras se utiliza fundamentalmente con la finalidad de expresar el gen clonado en un sistema eucariótico, con la excepción de los CAL o YAC, donde la

posibilidad de grandes insertos supera el interés tradicional.

6.3.1. Vectores de integración cromosómica

Los vectores de integración son moléculas híbridas que contienen un plásmido procariótico y secuencias de levadura. La presencia del plásmido procariótico facilita el crecimiento y preparación de grandes cantidades del vector en *E. coli*. Las secuencias de levadura permiten una integración en su propio genoma por recombinación homóloga. Cualquiera de los genes descritos en la tabla 6.1 es apto para crear la homología necesaria para la recombinación (*HIS*, por ejemplo), pero en la actualidad se prefiere provocar la integra-

ción en un lugar cromosómico más neutro como es el gen *HO*. Este gen determina una recombinasa de «homotalismo» que realiza el cambio de tipo sexual de α a «a», o viceversa, en células haploides; así, en las levaduras con interés industrial no se expresa por ser éstas, al menos, diploides.

Los vectores de integración se utilizan principalmente con dos finalidades:

- Introducir un gen de interés en la levadura.
- Provocar deleciones en genes concretos por recombinación homóloga (dianas génicas).

Un vector de integración de levaduras típico, también llamado vector lanzadera, es el **pFA6-kanMX3** (figura 6.2a) que tiene un esqueleto procariótico del tipo bluescript de 2,5 kb (capítulo 4,

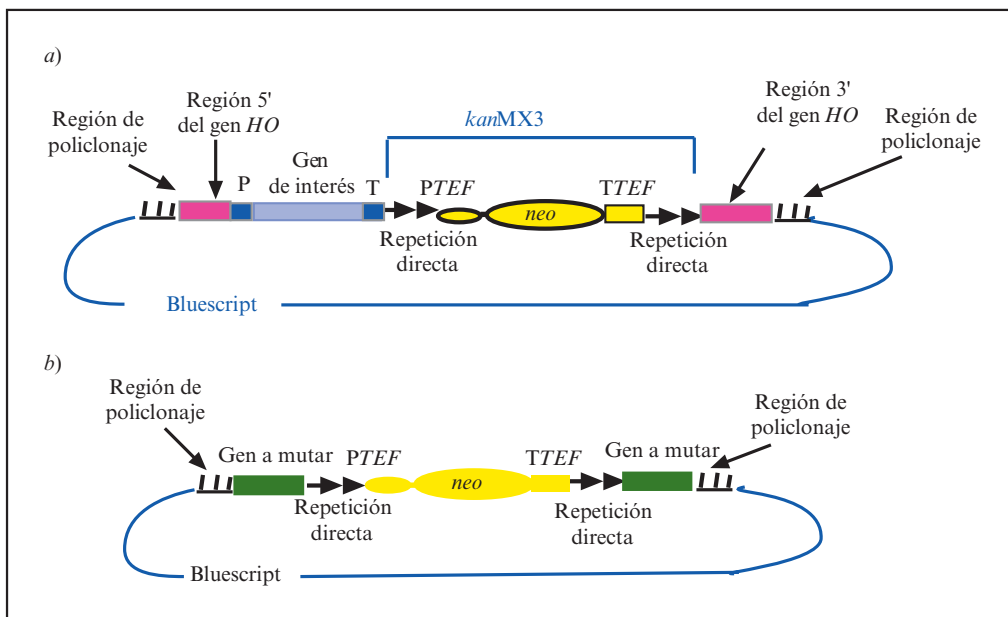


Figura 6.2. Representación esquemática de un vector lanzadera de integración de levaduras. *a)* A partir de un esqueleto bluescript, podemos apreciar dentro de la región de policlonaje, de izquierda a derecha: la región 5' del gen *HO* (recombinasa no esencial), el gen de interés (con promotor y terminador de levadura, P y T) y el elemento *kanMX3* que a su vez incluye las dos repeticiones directas terminales de *A. gossipii* y el gen bacteriano *neo* flanqueado por el potente promotor y el terminador del gen *TEF* también de *A. gossipii*. Finalmente, la región 3' del gen *HO* que crea la homología terminal para la integración. *b)* Vector para provocar dianas génicas y eliminar la expresión de un gen concreto en el que los extremos 3' y 5' del gen que deseamos mutar flanquean el elemento *kanMX3*.

apartado 4.2.3 y figura 4.9) y dentro de la región de policlonaje tiene insertada la región 5' del gen *HO* seguida del gen de interés, con su promotor y terminador; a continuación encontramos el elemento **kanMX3** que consta de dos terminales con repeticiones directas en los extremos, procedentes del hongo filamentoso *Ashbya gossipii* y el gen *neo* (originalmente del transposón Tn5 de bacterias) flanqueado por el potente promotor y el terminador del gen *TEF* (responsable del factor de translocación peptídica en el ribosoma) también de *A. gossipii*. Finalmente está colocada la región 3' del gen *HO*. Las regiones 5' y 3' del gen *HO* serán las responsables de la recombinación homóloga e integración del inserto del vector en este locus. Los terminales repetidos directos de **kanMX3** tienen como finalidad eliminar el gen de selección *neo* una vez que se hayan obtenido las levaduras con el gen de interés integrado y antes de su posible utilización comercial (si la proteína de interés se va a usar en humanos, por ejemplo). Con una frecuencia de recombinación de 10^{-3} - 10^{-4} , se producirá de nuevo una recombinación homóloga entre los dos terminales repetidos eliminando el gen de selección y devolviendo a la levadura su sensibilidad a G418. Esto se verificará haciendo crecer una muestra de cada clon de levaduras en presencia y ausencia del antibiótico (se realiza en medio sólido y rico como si fueran bacterias; véase la figura 5.9a).

Si lo que queremos es eliminar la función de un gen de la levadura por una diana génica, utilizaremos una versión del mismo vector en el que los extremos 3' y 5' del gen que deseamos mutar flanquean el elemento **kanMX3** (figura 6.2b). Al producirse la recombinación homóloga se sustituirá el gen endógeno por la versión delecionada e interrumpida.

Una aplicación biotecnológica de estos vectores en la producción de cerveza es, por ejemplo, la introducción en *S. cerevisiae* del gen bacteriano determinante de la α -acetolactato descarboxilasa. En la vía biosintética de la valina, el piruvato pasa a α -acetolactato y éste puede convertirse en diacetilo produciendo un desagradable sabor a mantequilla. Este sabor lo tendrían las cervezas de no ser maduras convenientemente. El sabor será

mucho menor (unas 200 veces) si la enzima bacteriana convierte el α -acetolactato procedente del piruvato en acetoína (figura 6.3). El método alternativo más natural y utilizado es mantener las cervezas madurando durante varias semanas a 4 °C, lo cual alarga el proceso de producción y encarece considerablemente el producto. En nuestro país este método puede resultar muy caro y complicado teniendo en cuenta el tamaño de los tanques de fermentación y la temperatura ambiental de los meses estivales en muchos lugares. En presencia de la enzima bacteriana, la conversión de α -acetolactato en acetoína es inmediata.

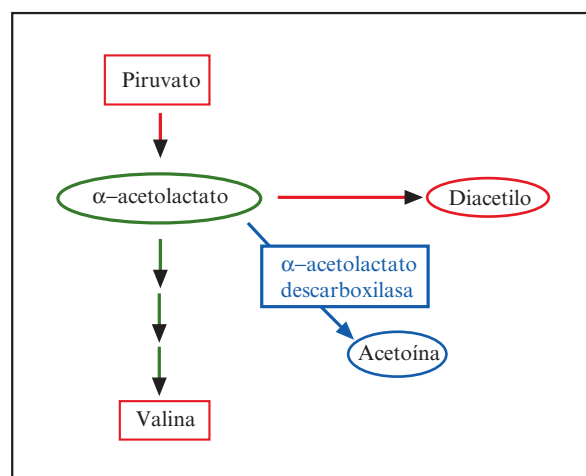


Figura 6.3. Esquema de la vía biosintética de la valina en *S. cerevisiae* y lugar donde actúa la enzima α -acetolactato descarboxilasa.

Otra aplicación interesante es la introducción de enzimas amilolíticas bifuncionales, como la glucoamilasa de algunos hongos que tiene la doble actividad α , 1 $\rightarrow$ 6 desramificante y α , 1 $\rightarrow$ 4 exoamilasa y que degrada amilodextrinas a maltosa y glucosa, ambas fermentables por la levadura. Así, durante la fermentación se produce un mayor contenido alcohólico reduciendo el contenido calórico, ya que la glucosa es más energética, en términos de moléculas de ATP generadas por molécula de glucosa, que el alcohol.

Otro gen que puede interesar añadir al genoma de la levadura cervecera es el de una β -glucanasa de origen bacteriano o fúngico. Esta enzima degrada los β -glucanos de la pared de levaduras muertas durante la fermentación y de las células del grano de la cebada. La mejora es doble: por una parte, propicia una fermentación adicional de la glucosa liberada, aumentando la producción de etanol y disminuyendo el aporte en calorías; por otro lado, evita que estos glucanos procedentes de levaduras muertas y otras fuentes obstruyan los filtros que separan los restos sólidos y de alto peso molecular del producto final.

También se han modificado levaduras panaderas industriales con la llegada de copias del gen *HAP4* precedidas por el promotor constitutivo pTEF2 para facilitar su sobre-expresión. *HAP4* es un factor de transcripción que activa la transcripción de genes implicados en la respiración y redirige el flujo de fermentación al de respiración. Esta modificación se ha realizado en diversas levaduras usadas en repostería para preparar masas de dulces y bollería, consiguiéndose un incremento de la velocidad de formación y del volumen de la biomasa final.

6.3.2. Vectores de replicación autónoma

Son vectores que se pueden replicar en forma libre sin necesidad de integración en el genoma de la levadura para perpetuarse. Dentro de estos vectores hay varios tipos. Puede interesar la clonación del ADN de interés, para crecer el recombinante, como en el caso de insertos de gran tamaño en cuyo caso elegiremos un vector basado en componentes cromosómicos como el cromosoma artificial, o bien lo que se busca es la expresión del inserto en cuyo caso elegiríamos un derivado del círculo de 2 μ m.

Vectores autónomos basados en componentes cromosómicos: cromosomas artificiales

Las secuencias de replicación autónoma (*SRA*) fueron descubiertas por casualidad en los

primeros experimentos de transformación y recombinación *in vitro*. Al transformar con los distintos vectores de integración mencionados (tabla 6.1), se observó que, cuando se utilizaba el *TRP1* en vez de *TRP5* con algunas secuencias flanqueantes del propio gen, la frecuencia de transformación aumentaba de 100 a 1.000 veces y además no ocurría la integración en el cromosoma. Al lado de *TRP1*, gen que codifica la enzima N-(5' fosforribosil)-antranilato isomerasa, había una secuencia de replicación autónoma (*SRA1*); pero la región *SRA1* no es la única secuencia de replicación autónoma en levaduras, ya que se pueden aislar otras simplemente seleccionando clones con una gran eficiencia de transformación. El problema de todas es, sin embargo, la inestabilidad y pérdida al cabo de unas pocas generaciones por falta de segregación equitativa durante la gemación, la mayoría de las yemas hijas carecen del plásmido. Es de suponer que, durante la división, el plásmido tiende a asociarse con algún compartimento celular que se queda predominantemente en la célula madre. Los vectores *SRA* mejoran sustancialmente con la adición de secuencias centroméricas (*CEN*) para aumentar su estabilidad. La región centromérica es responsable de la segregación a las células hijas al crear una estructura cromatínica especial; no parece que la secuencia de ADN sea el factor más importante. Los centrómeros son específicos de la especie y pueden tener tamaños muy diversos; el de *Saccharomyces cerevisiae* es de tipo puntual y sólo ocupa 250 nucleótidos, mientras que el de *Schizosaccharomyces pombe* es regional y abarca 40-100 kb. La estabilidad del vector, sin embargo, no se consigue plenamente. Se siguen perdiendo paulatinamente transformantes aún bajo selección, pero disminuye al 1% por generación. El número de copias por célula es de 1 o 2 copias por levadura transformada, por lo que queda descartado como vector de expresión (figura 6.4).

Cromosomas artificiales

Los vectores *SRA/CEN/TEL*: denominados CAL o YAC, son vectores *SRA/CEN* a los que se les han añadido telómeros cromosómicos.

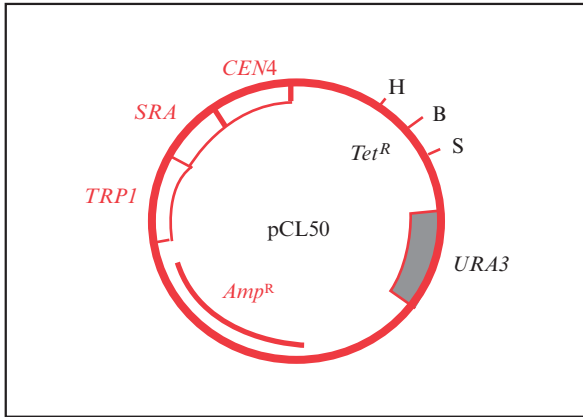


Figura 6.4. Ejemplo de plásmido *SRA/CEN*: pCL50 (plásmido centromérico de levadura). Además de la región *SRA* y la centromérica *CEN4*, tiene el gen *URA3* para la selección de transformantes. Posee dianas Sal (S) y Bam (B) en el gen que confiere resistencia a la tetraciclina.

Los telómeros son estructuras cromosómicas terminales muy conservadas; su secuencia de ADN está constituida por repeticiones cortas (6-25 pares de bases), pero la secuencia en sí no hace el telómero, pues se puede encontrar en otros lugares no teloméricos. Los telómeros son indispensables para mantener los cromosomas individuales; sin ellos, en general, los cromosomas se fusionan y degradan paulatinamente, pero las funciones pueden variar de unos organismos a otros. Los telómeros que se utilizan para construir los CAL proceden del protozoo ciliado *Tetrahymena piriformis* por estar mejor caracterizados que los de *Saccharomyces cerevisiae*. El resultado es una especie de minicromosoma artificial que permite clonar fragmentos de ADN de hasta 2 mgb (2.000 kb). Son los vectores que permiten clonar los fragmentos más grandes de ADN, siendo especialmente útiles cuando interesa aislar genes con muchos intrones que ocupan una gran extensión en el genoma. Básicamente, los vectores CAL (YAC) incluyen un plásmido procariótico para su propagación y preparación en *E. coli*. Se puede cortar con *SmaI* en dos fragmentos que constituyen los brazos cromosómicos. El brazo izquierdo tiene un telómero, un centrómero, una secuencia

SRA y un gen (*TRP*) para la selección (figura 6.5). El brazo derecho tiene otro telómero y otro gen de selección (*URA3*). La diana de restricción que separa los dos brazos se encuentra situada en medio del gen supresor 4 ocre, que codifica un alelo del ARNt de tirosina en levaduras y suprime las mutaciones *ade2* ocre, que afectan a la síntesis de adenina. Durante la clonación, los brazos se ponen en contacto con el ADN que se desea insertar; éste deberá tener un tamaño de entre 100 y 2.000 kb. Se transforman levaduras y se selecciona por *TRP* y *URA*, para asegurarse de la presencia de los dos brazos. También se puede verificar la presencia o ausencia del supresor 4. Una cepa de levaduras con la mutación *ade2* producirá colonias rojas, frente a colonias blancas si la mutación es suprimida. En los mutantes *ade2* la adenina se sintetiza por una vía metabólica alternativa y el color rojizo es debido a la formación de un compuesto coloreado como subproducto.

La eficiencia de clonaje con este sistema es de unas 100 colonias por μg de ADN, el tamaño medio de inserto, de 500 kb, y los recombinantes se analizan por electroforesis en campo pulsante que nos permite separar fragmentos de tamaños entre 100 y 2.000 kb.

El sistema CAL se emplea preferentemente para elaborar librerías de genomas muy grandes, como el humano, o para mantener un gen completo con muchos intrones en un solo clon. También se ha utilizado en la creación de bacterias sintéticas para mantener y crecer fragmentos cromosómicos de gran tamaño, así como para construir cromosomas artificiales humanos (capítulo 7).

Dentro de las limitaciones, tenemos el posible enlace de secuencias adyacentes en el vector, que correspondan a segmentos distantes en el cromosoma original; estas quimeras artificiales aparecen en el 35% de los casos. También es una limitación la posible reordenación o delección espontánea de parte del ADN inserto por recombinación homóloga entre secuencias repetidas o palindrómicas.

Vectores de expresión

Utilizaremos un vector de expresión de levaduras si la finalidad de la clonación es obtener un

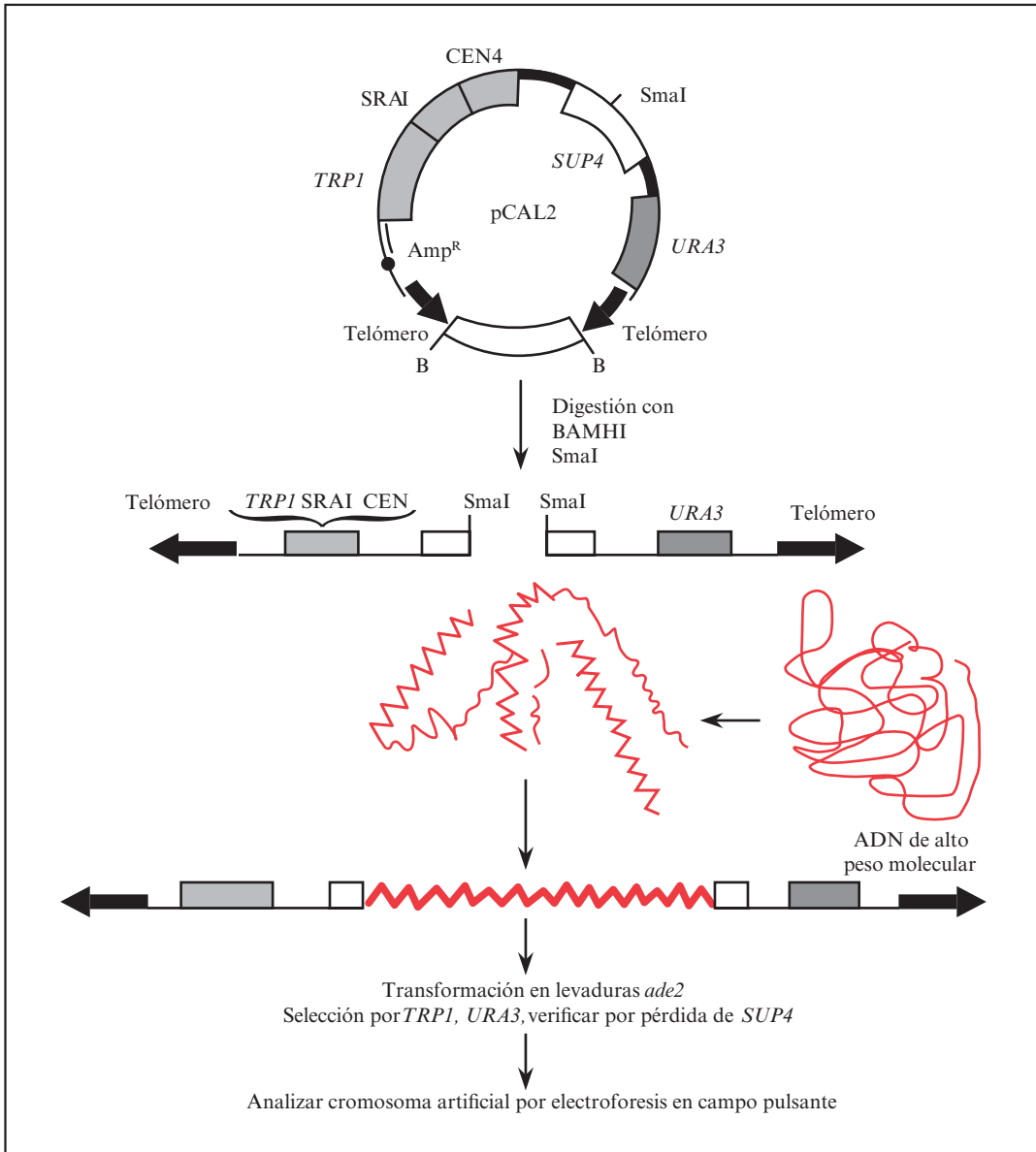


Figura 6.5. Vector CAL (YAC) para insertos muy grandes de ADN. La digestión con Bam y Sma elimina la región procariótica (Bam) y origina dos brazos cromosómicos entre los que se inserta el macrofragmento a clonar. El ADN de alto peso molecular se origina por digestión parcial con SmaI.

producto de origen distinto, humano por ejemplo. La región promotora en levaduras es algo distinta a la típica reconocida por la polimerasa II en la

mayoría de los genes eucarióticos (capítulo 1, apartado 1.5, figura 1.11), por lo que en los vectores conviene mantener la propia de levaduras.

Las dos regiones importantes en la expresión de cualquier gen en levaduras son: 1) la región SIT que comprende el complejo de pre-iniciación que es donde se inicia realmente la transcripción, y 2) la secuencia activadora (SA) ubicada a una distancia de entre 200 y 1.400 pares de bases, corriendo arriba, que es el equivalente a los potenciadores de eucariotas superiores.

Círculo de 2 μ m

S. cerevisiae posee un plásmido endógeno de 2 μ m de longitud, que no confiere a la célula ninguna característica especial. Se le denomina simplemente el círculo de 2 μ m. El número de copias es de 50 a 100 unidades por célula; es estable y se comporta como un minicromosoma circular. Posee la estructura típica de la cromatina, es nuclear y su replicación parece estar bajo el mismo control genético que el resto de los cromosomas. Su

presencia no es esencial para la levadura; las cepas que lo poseen se denominan *cir*⁺, y las que carecen de él *cir*^o. Su tamaño original es de 6,3 kb, pero se ha manipulado considerablemente variando sus dimensiones y contenido.

El plásmido original posee los genes *REP* importantes para la replicación del plásmido, de entre ellos *REP3* actúa como centrómero. El genoma está dividido por dos repeticiones invertidas perfectas de 599 pb que separan dos regiones distintas, cada una con dos genes. El gen *FLP* se corresponde con la recombinasa flipasa (capítulo 2, apartado 2.2.2, y figura 2.14), que es la responsable de la recombinación entre las dos secuencias repetidas y que cambia de orientación a *REP2* y *FLP* (figura 6.6). La función de este cambio de orientación es desconocida. Las versiones modificadas tienen el origen de replicación y la resistencia a ampicilina de pBR para su selección y crecimiento en *E. coli*; el gen *LEU2* como marcador

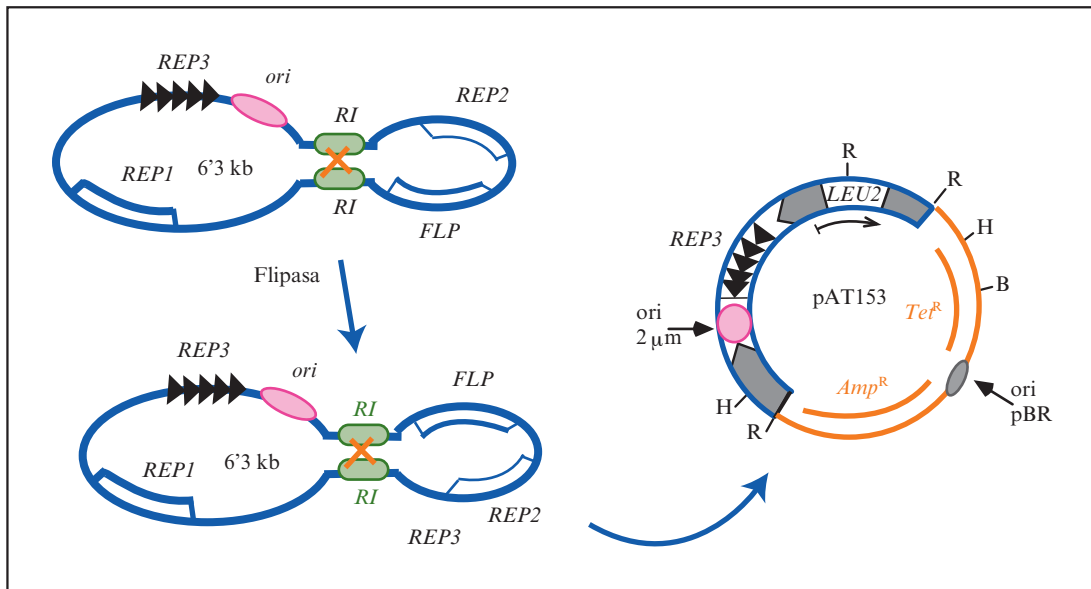


Figura 6.6. Círculo original de 2 μ m y vector modificado basado en el original. A la izquierda el círculo original con las dos repeticiones invertidas (RI), los genes *REP* y el gen de la flipasa (*FLP*). Las cabezas de flecha simbolizan las secuencias repetidas *REP3*. La región procariótica comprende el plásmido pAT153 que confiere resistencia a tetraciclina (*Tet*^R) y ampicilina (*Amp*^R) y porta el origen de replicación de pBR322. Las dianas HindIII (H) y BamHI (B) son idóneas para la inserción de genes de interés. El inserto eliminará la resistencia a tetraciclina manteniendo la de ampicilina.

selectivo de transformantes en levaduras, y el gen de resistencia a la tetraciclina con dianas Hind III y Bam HI para la clonación de genes de interés. Mantiene el origen de replicación del círculo primitivo y la región repetida *REP3* (5,5 copias de una secuencia de 62 nucleótidos) que estabiliza la replicación y distribución de copias a la progenie, pero le faltan varios de los genes que originalmente poseía el plásmido y que también contribuyen a la estabilidad de la replicación (*REP1* y *REP2*). Los productos de estos genes deben ser facilitados en trans, por lo que estos vectores sólo se pueden mantener en cepas *cir*⁺. La eficiencia de transformación es alta, produciendo de 10.000 a 100.000 transformantes por µg de ADN.

Los vectores basados en el círculo de 2 µm tienen como finalidad la expresión del inserto, no tanto a niveles de explotación farmacéutica o industrial de la proteína recombinante, como de sistema modelo en investigación básica o aplicada. Uno de los primeros genes foráneos expresado en levaduras fue el de la β-hemoglobina de conejo insertado en un vector derivado del círculo de 2 µm en 1980; el gran número de copias por célula garantiza niveles muy elevados del mensajero y de la proteína correspondiente. Pero en aquella ocasión lo más que se llegó a detectar fue un mensajero defectuoso. El promotor del gen eucariótico no había sido reconocido eficientemente por la polimerasa de *S. cerevisiae*, los intrones no habían sido procesados y la transcripción había terminado prematuramente en una región rica en AT del segundo intrón. Sin embargo, en la actualidad se sintetizan con mucha eficiencia multitud de genes foráneos en este sistema. Uno de los promotores con frecuencia utilizados en vectores derivados del círculo de 2 µm es el del gen fosfoglicerato kinasa (*PGK*), así como su región de terminación de la transcripción. Otros promotores, como el de la respuesta al choque térmico o los basados en el gen *GAL*, implicados en el metabolismo de la galactosa, son los más populares. Los genes *GAL1* y *GAL10* se inducen unas 1.000 veces por galactosa.

Vectores derivados del círculo de 2 µm se han utilizado también mucho para identificar y clonar proteínas de cualquier organismo que interaccio-

nen entre sí, por el sistema del doble híbrido o del cebo.

Sistema del doble híbrido o del cebo

Cuando tenemos clonado un gen de un organismo cuyo producto presumimos o tenemos la certeza de que interacciona con otro, se hace necesario, en muchos casos, identificar este segundo producto y aislar su ADN correspondiente. Teniendo en cuenta que las interacciones proteína-proteína están implicadas en todas las cascadas enzimáticas y funciones celulares conocidas, la probabilidad de que un gen de interés hable con otro es muy alta.

La proteína conocida y clonada es el cebo, su ADN se fusiona al de una proteína de unión a ADN de levaduras como «LexA» o «Gal4». El cebo puede ser, por ejemplo, una quinasa de la que no conocemos el ligando o una fosfatasa, pero en todos los casos debe cumplir dos premisas: 1) que por sí mismo el ADN del cebo no corresponda a un factor de transcripción activo y 2) que el producto del cebo no deba forzosamente ser excluido del núcleo de la levadura. La fusión del cebo y *lexA* (por ejemplo) estará clonada en un vector derivado del círculo de 2 µm como el pEG 202 de 9,8 kb (figura 6.7). Se inserta además en el genoma de la levadura el gen señal *lacZ* con su correspondiente secuencia de activación SA (en inglés UAS por *upstream activating sequences*), reconocibles por LexA. Esta misma secuencia SA se introduce en el extremo 5' del gen *leu* para controlar su transcripción. La levadura portadora del plásmido ha de ser, pues, auxótrofa para la leucina (*LEU2*⁻), además de para la histidina (*His3*⁻) y *cir*⁺. Además, ha de ser auxótrofa para el triptófano (*TRP1*⁻) porque el paso siguiente es la transformación de esta cepa de levaduras que ya lleva y expresa el cebo, con los componentes de una librería de ADNc elaborada, en otro derivado del círculo de 2 µm, como el pJG 4-5 de 6,2 kb, cuya expresión se controla e induce con galactosa y lleva el gen de la biosíntesis del triptófano *Trip 1* que suple la mutación *TRP1*⁻ de la levadura (figura 6.7). Todos los ADNc de esta librería están fusionados en fase con un activador de la transcripción que por sí

solo no puede unirse a los promotores, si no es mediante la interacción de la proteína codificada por el ADNc con la proteína fusionada cebo-lexA que, a su vez, estará fijada a los promotores con SA. Así, levaduras con fragmentos de la librería de

ADNc que no expresan proteínas que interaccionan con el cebo no crecerán en ausencia de leucina. Sin embargo, las que expresen la proteína de interacción crecerán en un medio sin leucina y además se colorearán de azul en medio con X-gal.

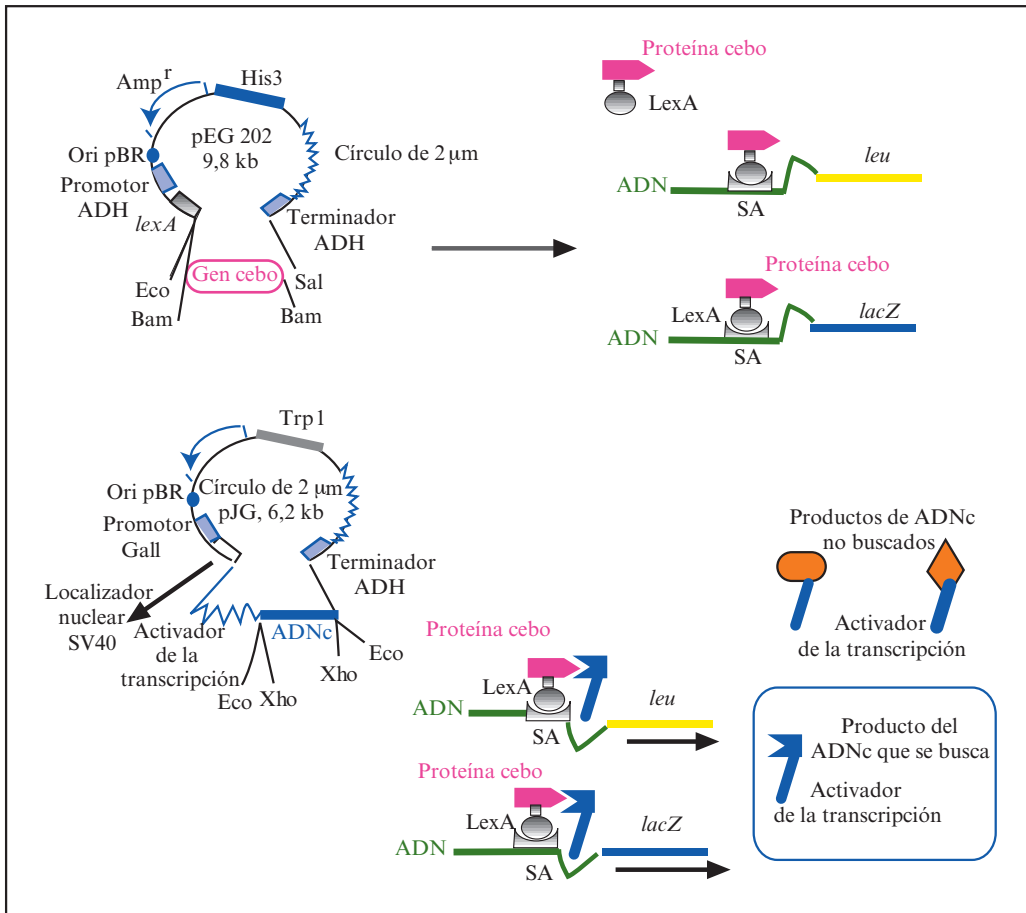


Figura 6.7. Sistema del doble híbrido o del cebo para identificar y aislar el ADN correspondiente a proteínas que interaccionen con una ya conocida y clonada. Arriba, a la izquierda, el plásmido de 2 µm portador del cebo fusionado a *lexA*. Abajo, a la izquierda, el segundo plásmido de 2 µm portador de la secuencia que va a hablar con el cebo, fusionada a un activador de la transcripción.

Vectores de origen mixto

El vector de expresión λ ZAP-pEL-URA3 (plásmido de expresión de levadura URA3) es

bastante sofisticado. Este vector es derivado del λ ZAP (véase capítulo 4, apartado 4.2.3 y figura 4.10). Recordemos que en el λ ZAP en presencia de un virus auxiliar, F1 o M13, la parte de blues-

cript puede escindirse *in vivo* de λ y crecer como plásmido. Dentro del bluescript, se han insertado las secuencias de levaduras que le convierten en un excelente vector de expresión. Inmediatamente detrás del promotor *GAL1* inducible, se ha incorporado otro promotor inducible *GAL10*, orientado en sentido opuesto al anterior y limitado por una diana *EcoRI*. El gen *URA3*, para seleccionar por transformantes en levaduras; las secuencias *SRA1* y *CEN4*, para estabilizar dentro de lo posi-

ble el plásmido en levaduras. Este vector de expresión de 6,4 kb queda incorporado en el λ ZAP una vez que tenga el inserto del gen eucariótico (clones de ADNc en la mayoría de los casos). La ventaja de esta inserción es la gran eficiencia de transducción fágica en bacterias. Una vez en la bacteria, se puede escindir con ayuda de un fago auxiliar y reproducirse, circularizado, como plásmido (figura 6.8). Este vector se puede obtener comercialmente.

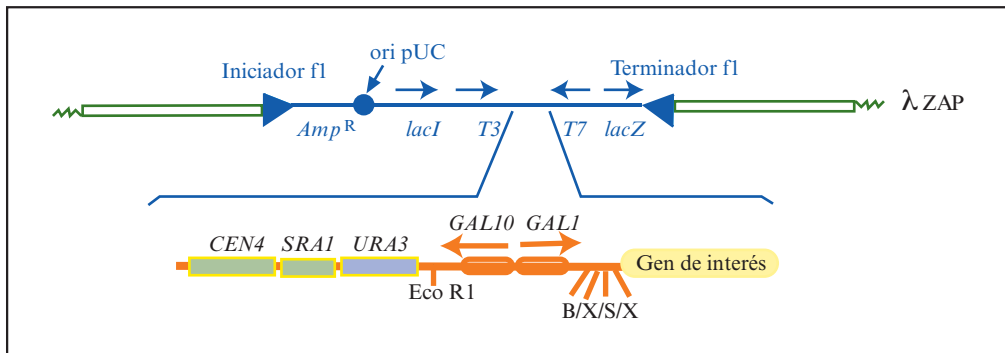


Figura 6.8. Vector de expresión de levaduras λ ZAP-pEL-URA3. Es un vector λ de inserción que lleva en su interior otro vector procariótico, el bluescript M13 (apartado 4.2.3, figura 4.9), constituyendo el λ ZAP. En un extremo de la región de policlonaje del bluescript se han introducido secuencias de levadura centroméricas (CEN), autorreplicativas (SRA), necesarias para la selección en levaduras (URA3) y los promotores inducibles *GAL1* y *GAL10*, transformándose en el vector λ ZAP-pEL-URA3. El gen de interés se inserta en el resto de la región de policlonaje, que a su vez ha sido modificada, manteniendo sólo dianas *BamHI*, *XbaI*, *SalI* y *XhoI*.

6.4. TRADUCCIÓN Y MODIFICACIONES POSTERIORES COMO GLICOSILACIÓN Y SECRECIÓN DE PRODUCTOS

Un factor que puede disminuir considerablemente la traducción de mensajeros de mamíferos es el uso sesgado de determinados codones. Por ejemplo, el codón UUG de la leucina se usa en el 91 % de los casos, frente a CUG que nunca es utilizado; los genes que utilicen CUG para sus leucinas se traducirán muy lentamente, debido a la escasez de los correspondientes ARNt isoaceptores en levaduras. Una solución es alterar *in vitro* codones poco utilizados por los preferidos en este sistema. La composición de bases alrededor de la

región de iniciación de la traducción puede también afectar a la eficiencia de iniciación: una región rica en A+T cercana al ATG inicial es unas tres veces más eficiente que una región rica en G+C. La adquisición de determinadas estructuras secundarias también influye en la velocidad de traducción y pueden enlentecer o incluso desmontar en subunidades a los ribosomas cuando pasan.

Las levaduras pueden realizar un gran número de modificaciones en productos de genes humanos como fosforilar la proteína del oncogén *c-myc*, acetilar la proteína del oncogén *ras* y glicosilar muchas otras. Las proteínas normalmente tienen una localización precisa dentro de la célula y esta localización está determinada por signos de carácter es-

tructural que muchas veces son correctamente interpretados en levaduras. Algunas señales en el extremo N-terminal dirigirán la proteína a la secreción; la proteína C-myc se localizó correctamente en el núcleo y Ras en la membrana. Incluso el receptor de la acetilcolina en el *Torpedo California* o la hemaglutinina del virus de la gripe se localizaron correctamente en la membrana citoplasmática.

Las levaduras también excretan correctamente proteínas de mamíferos. La secreción facilita la producción de extremos amino sin la metionina inicial, facilita la purificación, evita la degradación por proteasas intracelulares y favorece el correcto plegamiento de la proteína. El proceso de secreción en levaduras es esencialmente idéntico al de otros organismos superiores. Las proteínas de levaduras que se excretan pueden tener distintos destinos: pueden acabar en el medio de cultivo, como el factor sexual α ; pueden acumularse en el espacio periplasmático, como la invertasa o la fosfatasa ácida; en vacuolas, como la carboxipeptidasa, o, incluso, en la membrana citoplasmática, como las permeasas. Para que una proteína entre en la vía de secreción necesita tener en su extremo 5' un **péptido señal** de unos 20 aminoácidos aproximadamente (varía ligeramente de unas proteínas a otras), que durante el proceso de secreción será eliminado. El péptido señal dirige la proteína hacia el interior del retículo endoplasmático, donde el propio péptido es eliminado, y la proteína, enviada a su último destino celular. Los distintos péptidos señal no presentan homología entre ellos, pero mantienen tres regiones características: una región N-terminal hidrofílica, una región central hidrofóbica y una región terminal de reconocimiento para la peptidasa.

La glicosilación tiene lugar en el retículo endoplasmático y puede ser específicamente inhibida por tunicamicina. La glicosilación estimula la secreción, y en algunos casos puede ser imprescindible para la función. Desgraciadamente la glicosilación de *S. cerevisiae* es mucho más simple que la de otros organismos superiores. En general, los patrones de glicosilación son difíciles de imitar, pues incluso dentro de un mismo organismo pueden variar de un tejido a otro.

A pesar de las limitaciones mencionadas, se han podido producir y excretar en levaduras un

número aceptable de proteínas de diverso origen con eficiencias variables. La secreción de la lisozi-ma humana es muy eficiente, pero la del interferón baja considerablemente. En general, el modo más eficaz consiste en acoplar un péptido señal de levaduras con el inicio del gen de interés en el vector. El péptido señal de la invertasa es uno de los más utilizados. En la elaboración del vector de secreción hay que tener en cuenta acoplar el péptido señal sin alterar la fase de lectura del gen; además, hay que mantener intacta la región de ruptura del péptido señal (que precisa un residuo de alanina) para dejar posteriormente la proteína de interés libre.

6.5. OTRAS LEVADURAS DE INTERÉS

Hay otras levaduras, más o menos relacionadas con el género *Saccharomyces*, que son mejores para la producción de proteínas de interés. Las especies más utilizadas son: *Schizosaccharomyces pombe*, *Pichia pastoris*, *Hansenula polymorpha*, *Schwanniomyces occidentalis* y *Kluyveromyces lactis*.

Schizosaccharomyces pombe se divide produciendo un *septum* en vez de por gemación y su ciclo celular es más parecido a eucariotas que al de *S. cerevisiae*, por lo que se utiliza preferentemente para estudios del ciclo celular. *S. pombe* es capaz de procesar correctamente algunos intrones eucarióticos (no todos).

Pichia pastoris es interesante porque produce grandes cantidades de la enzima alcohol oxidasa (AO), llegando a representar el 35% de la proteína total celular. La expresión de la enzima está regulada de modo que no se sintetiza en presencia de etanol o glucosa, pero se estimula cuando la fuente de carbono es metanol. La facilidad con que *P. pastoris* produce grandes cantidades de proteína, con una fuente de carbono tan barata como el metanol, coloca a esta levadura en una posición muy atractiva para la producción de otras proteínas eucarióticas de interés biotecnológico. Tanto *S. cerevisiae* como *P. pastoris* se han utilizado para la expresión con carácter comercial del antígeno S utilizado como vacuna de la hepatitis B. Los niveles del antígeno S en una cepa de *S. cerevisiae*, con más de 50

copias del gen heterólogo integrado, son comparables a los producidos en *P. pastoris* con una sola copia integrada. Ambas levaduras se han utilizado en la producción de otras muchas proteínas (el factor de crecimiento epidérmico de ratón, la albúmina del suero humana, la inmunoglobulina F-1 humana, etc.) y en todos los casos la productividad de *P. pastoris* ha sido muy superior a la de *S. cerevisiae*.

Hansenula polymorfa y *Schwanniomyces occidentalis* son capaces de excretar numerosas proteínas muy eficientemente al medio de cultivo aunque su utilización práctica no está todavía bien desarrollada.

Cluyveromyces lactis se ha utilizado para la síntesis de interleuquina-1 β y de la albúmina del suero humano. En situaciones experimentales muy semejantes, *K. lactis* produce 80 mg de interleuquina por litro, mientras que *S. cerevisiae* sólo logra 1-2 mg/litro. La albúmina del suero también se excretó directamente al medio de cultivo en concentraciones muy superiores en *K. lactis* que en *S. cerevisiae*.

6.6. TRANSFERENCIA GÉNICA EN INSECTOS COMO SISTEMA DE EXPRESIÓN ALTERNATIVO AL DE LEVADURAS: LOS BACULOVIRUS COMO VECTORES

La expresión de proteínas recombinantes eucarióticas en insectos puede ser comparable o superior al sistema de levaduras, con producciones de 1-100 mg por litro de cultivo. Las células de insecto cultivadas *in vitro* pliegan correctamente las proteínas, las glicosilan y forman los puentes disulfuro correspondientes. La propia larva del insecto puede convertirse en una fábrica del producto clonado. Los vectores son **los baculovirus**, familia de virus que infectan exclusivamente a invertebrados y predominantemente a insectos; sus miembros son numerosos (unos 500) y poseen como genoma un ADN circular de unas 128 kb que mantienen dentro de una cápside tubular. Deben su nombre a que los viriones tienen forma de báculo o bastón. El prototipo de la familia es el virus de la poliedrosis nuclear de *Autographa californica* (VPNAc) que se propaga con facilidad en líneas celulares de los

insectos *Spodoptera frugiperda* (Sf, el gusano cogollero del maíz), *Trichoplusia ni* (Tn), *Mamestra brassicae* y *Estigmene acrea*. En menor medida también se ha utilizado el virus de la poliedrosis nuclear de *Bombyx mori* (el gusano de seda). Las larvas de los insectos son muy voraces y pueden tragarse el virus que se encuentre en la superficie de las hojas, donde pueden pasar meses o años como cuerpos de inclusión de forma poliédrica. Estos cuerpos de inclusión están formados por virus y poliedrina, donde la poliedrina es la proteína soporte de los virus. En el intestino del insecto, las condiciones alcalinas disuelven la poliedrina y liberan el virus a la hemolinfa provocando una infección primaria en la que nuevos virus abandonan la célula infectada por gemación. El virus no se integra en el genoma de la célula infectada. Los virus liberados provocan una infección secundaria sistémica, en el núcleo de las células infectadas se acumulan muchos cuerpos de inclusión de forma poliédrica, la lisis celular es entonces inevitable, con liberación de miles de partículas.

Los genes de los baculovirus se expresan en fases temprana, tardía y muy tardía; uno de los genes muy tardíos es el gen de la poliedrina. El gen está bajo el control de un promotor muy potente, pero puede ser delecionado sin que afecte a la replicación o infectividad de las partículas. Puede introducirse un gen de interés en sustitución del de la poliedrina, expresándose con gran eficiencia durante la infección y causando la ausencia de cuerpos de inclusión en el núcleo como criterio de identificación del recombinante.

El tamaño excesivo del ADN viral dificulta su manipulación como vector. La metodología que se emplea es la de eliminar un gen esencial que le convierta en no-infeccioso a no ser que sea facilitado en trans por un plásmido de transferencia también introducido en las células de insecto. El plásmido lanzadera lleva además un gen señal de selección positiva, como lo es el de la proteína verde (*gfp*, del inglés: *green fluorescent protein*) que fluoresce con luz ultravioleta de longitud de onda larga y que se aisló originalmente de la medusa *Aequorea victoria*. Finalmente, lleva el gen de interés colocado dentro del gen de la poliedrina (figura 6.9). Las células del insecto cultivadas *in vi-*

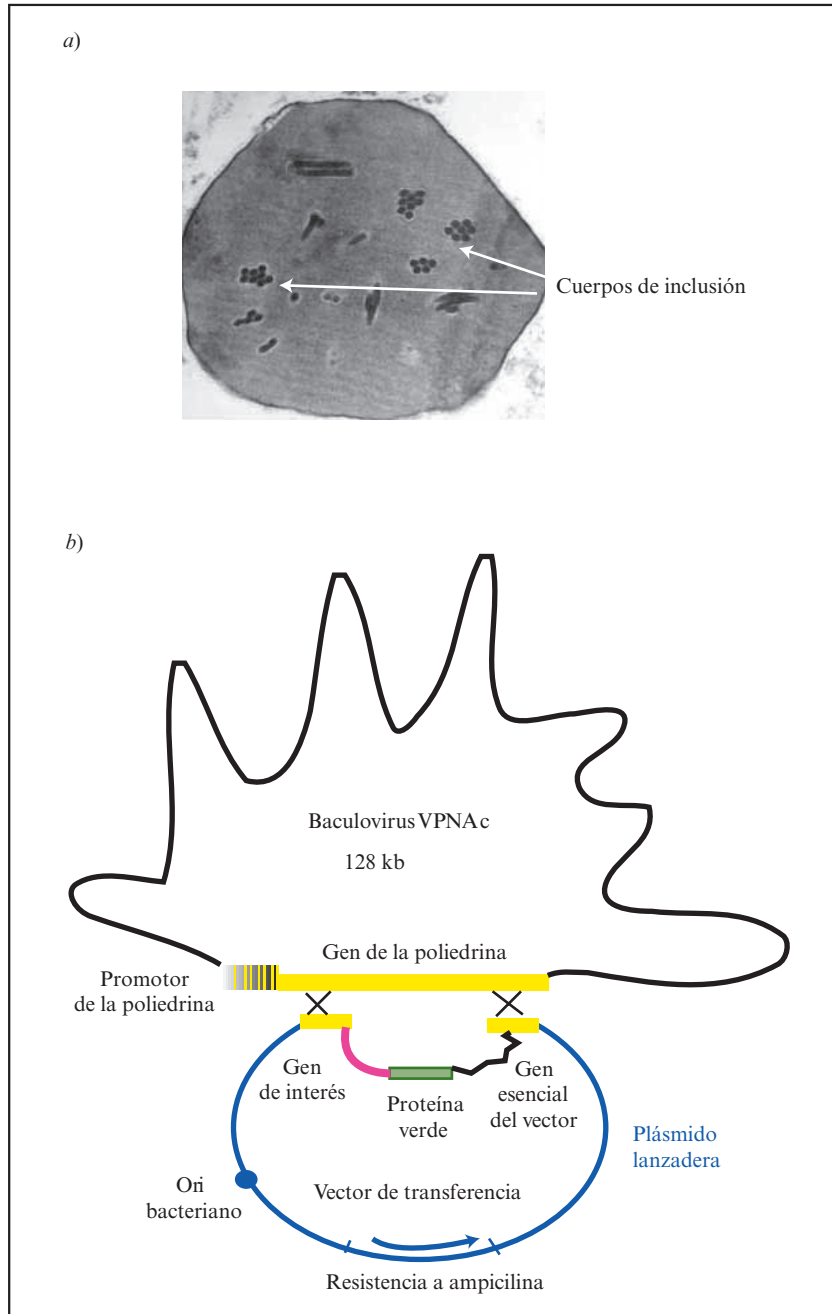


Figura 6.9. Baculovirus. a) Corte de un núcleo infectado por baculovirus. b) Cointegración por recombinación homóloga en el gen de la poliedrina, entre el baculovirus defectivo y el plásmido lanzadera. El baculovirus resultante es viable, carece de inclusiones poliedrónicas nucleares y toma color verde con luz ultravioleta; además, produce grandes cantidades del gen de interés insertado.

tro se cotransfectan con el vector viral defectivo y con el plásmido de transferencia. La recombinación homóloga por el gen de la poliedrina elimina la función de este gen que puede ser monitorizada, recupera la viabilidad del virus y las células adquieren color verde al ser iluminadas con luz ultravioleta. Cuando las células lisan, al cabo de unos tres días, los niveles de producción son del orden de 1-100 mg de la proteína de interés por litro de cultivo. Este sistema se ha utilizado con éxito para producciones a gran escala del interferón α humano, el factor de crecimiento epidérmico o polimerasas para uso comercial.

Los insectos realizan mejor los plegamientos de proteínas de mamíferos y la formación de puentes disulfuro que las levaduras, pero el sistema de expresión en baculovirus tiene la limitación de no ser capaz de realizar una N-glicosilación similar a la de mamíferos. Esencialmente las células de insectos no producen oligosacáridos complejos y ramificados con residuos de galactosa en penúltima posición y ácido siálico en última, excepto en la línea celular derivada de *Estigmene acrea* que, por poseer la enzima N-acetil- β -glucosidasa, produce un patrón de glicosilación muy similar al de los mamíferos.

El gusano de seda (*Bombix mori*) es una alternativa más barata que los cultivos celulares en gran escala. La cotransfección y recombinación homóloga se lleva a cabo en células de *Bombix* cultivadas *in vitro*; pero los virus recombinantes producidos en éstas se inyectan en orugas del gusano de seda, llegando a producir 50 mg de la proteína de interés por larva. La industria relativa a cultivos en gran escala de gusanos de seda está ya muy establecida para la producción de seda, por lo que las adaptaciones que se requieren son mínimas.

Los baculovirus también se han utilizado como vectores de genes terapéuticos en mamíferos. Estos virus pueden entrar en numerosos tipos celulares de mamíferos; una vez en el citoplasma, inducen la formación de microfilamentos de actina que facilitan el transporte al núcleo. Una vez en el núcleo, pueden permanecer expresando el gen de interés durante algún tiempo. Entre las ventajas de este vector para terapias génicas, está que no se divide ni produce infección viral en mamíferos o en el hombre. Entre los inconvenientes está el que provoca una respuesta inmunológica al cabo de una o dos semanas y que crece mal a 37 °C, ya que su temperatura óptima es 27 °C.

Transferencia génica a células de mamíferos. Terapia génica

Aunque las células bacterianas son enormemente útiles para clonar, propagar y expresar genes de organismos superiores, las células de mamíferos siguen siendo las más idóneas para la producción de proteínas, por ejemplo, humanas, en forma biológicamente activa. Este tipo de células son el medio natural para la producción de proteínas con aplicación terapéutica. Quizá el mayor riesgo que conlleva este tipo de producción a gran escala es que los vectores, o algunos de los componentes de los vectores usados en estos experimentos, son oncogénicos o proceden de virus oncogénicos, con lo que, en teoría, podrían inducir transformaciones de tipo tumoral. En el presente capítulo se estudiarán los distintos procedimientos existentes para introducir y expresar genes eucarióticos en células de mamíferos, las limitaciones de estos métodos, así como las perspectivas de futuro.

En células de mamíferos el término «transformación», que significaría la introducción de ADN desnudo en la célula, puede inducir a error, ya que el paso de una célula normal a tumoral también se denomina «transformación». Utilizaremos el término «transfección» para la introducción de ADN desnudo o unido a un plásmido procariótico; cuando la entrada está mediada por un virus hablaremos de «transducción» o, más coloquialmente, de infección.

La terapia génica es un tratamiento de enfermedades con alteraciones genéticas incurables que actúa a nivel celular y utiliza como agente terapéutico un ácido nucleico. La terapia génica, como fenómeno biológico, no es algo extraordi-

nario ni siquiera muy raro. Hoy por hoy es sólo somática, puesto que no intenta corregir la mutación en la línea germinal para suprimir la enfermedad en la descendencia del individuo afectado. Existen en la actualidad unas 3.000 enfermedades genéticas descritas que serían susceptibles de terapia génica. De estas enfermedades, las denominadas **monogénicas** son aquellas en las que la alteración en un solo gen causa la patología. Su terapia requeriría transferir a las células deficientes una copia intacta del gen, con lo que se restauraría el fallo celular y clínico. Son ejemplo de este tipo de enfermedades la inmunodeficiencia combinada severa (ADA o SCID-X1), la fibrosis quística, la distrofia muscular o la amaurosis congénita de Leber. La esquizofrenia, algunas formas de diabetes o el cáncer son, por el contrario, enfermedades **multigénicas** donde alteraciones en diversos genes son la causa de la patología. El cáncer es, sin duda, la enfermedad con más ensayos clínicos realizados a pesar de su complejidad y ello es debido a su alta incidencia entre la población, con unos 150.000 casos diagnosticados por año en España y de los que sólo aproximadamente la mitad sobrevivirán a la enfermedad.

7.1. TIPOS Y MÉTODOS DE TRANSFECCIÓN

Hay dos tipos de transfección en células eucarióticas: **transitoria** y **estable**. En la transfección transitoria el ADN, una vez introducido, permanece en forma de plásmido en el citoplasma celular, expresándose y llegando a un máximo de ex-

presión a las 48 horas después de la entrada, pero desapareciendo al cabo de algunas generaciones por no transmitirse a las células hijas. En la transfección estable o permanente, en cambio, la integración del ADN en el genoma de las células transformadas asegura la perpetuidad de las secuencias introducidas.

Método del fosfato cálcico. Fue durante mucho tiempo el método más extendido, aunque ya no se utiliza por ser demasiado laborioso. No se conocen las causas por las que el tratamiento con este compuesto químico predispone a las células a ingerir ADN por endocitosis:



Dependiendo del tipo de células, se puede conseguir que hasta un 90% incorporen ADN. Se utiliza tanto para transfecciones transitorias como para permanentes, pero sólo del 1 al 5% de las moléculas que entran en la célula llegan a integrarse en el ADN.

El polibreno-DMSO. El polibreno es un polímero que arropa y cubre el ADN, favoreciendo su adsorción en la membrana citoplasmática. El DMSO (dimetil sulfóxido) al 25-30% se añade como tratamiento breve de choque para permeabilizar la membrana y favorecer la entrada del ADN. Este método se ha utilizado con éxito en células embrionarias de pollo y fibroblastos de ratón.

El DEAE dextrano es un método similar al anterior, donde el ADN interacciona y se cubre con el polímero de alto peso molecular: dietil aminoetil dextrano. Es específico para transfecciones transitorias. Si el tratamiento se completa con un choque de DMSO, se consigue que hasta un 80% de las células transformadas expresen los genes introducidos. Una vez que ha penetrado el ADN, a las 24 horas de la transfección, suele adoptar la estructura de nucleosomas que a su vez forman minicromosomas, pero no da lugar a transformaciones estables.

La electroporación consiste en aplicar descargas eléctricas de alto voltaje a las células para pro-

ducir poros nanométricos en sus membranas. El ADN entra en la célula, bien a través de los poros, o al redistribuirse los componentes de la membrana con el cierre de éstos. El método se realiza introduciendo las células en suspensión y el ADN en una pequeña cubeta y aplicando descargas de 2,0 a 4,0 kV (kilovoltios), a 0 °C. Del 60 al 90% de las células suelen sobrevivir al tratamiento, pero conviene hacer una curva de calibración donde se determinen los porcentajes de muerte celular respecto a la diferencia de potencial aplicada, la duración de ésta y la concentración de ADN óptima en cada caso. El ADN en forma lineal es unas veinte veces más eficiente que superenrollado, probablemente debido a que de esta forma se integra más fácilmente en el cromosoma (los extremos son recombinogénicos). El método tiene una eficiencia de unos 25 a 30 transformantes por cada 10⁵ células vivas. Se ha utilizado con éxito en una gran variedad de células, incluyendo linfocitos B y T de ratón, células linfáticas humanas, fibroblastos de diversos orígenes y células de pituitaria de rata.

Con *rayos láser* también es posible permeabilizar membranas celulares. Se han aplicado sobre células humanas en cultivo obteniéndose transfectantes con una frecuencia cien veces superior a la obtenida con el método del fosfato cálcico. El rayo láser produce una radiación intensa, monocromática y electromagnética, que puede dirigirse a una estructura subcelular de modo controlable y consistente. Se pueden confinar lesiones a un área de 0,1 µm de diámetro. En la membrana puede producir numerosos poros microscópicos por donde se introduce el ADN, que parece estar ya en el interior de la célula cuando tienen lugar los procesos de reorganización de la membrana y cierre. Al poder concentrar el haz de rayos láser sobre un espacio menor que los límites de difracción del microscopio óptico, es muy importante obtener imágenes nítidas y de mucha calidad. La utilización de una cámara de televisión y una microcomputadora que amplifique y digitalice la imagen observada se hace imprescindible en este tipo de experimentos. Con ello se consigue no sólo aumentar considerablemente el contraste de los contornos, sino también disminuir el fondo

inespecífico, avivar los colores, calcular áreas, longitud de superficies, distancias intracelulares, etcétera. Se pueden irradiar unas 1.000 células por hora, enfocando cada una en particular antes de emitir la descarga correspondiente con un disparador electrónico asociado al láser. El único inconveniente de este método es la necesidad de poseer un equipo bastante sofisticado para su realización.

Los *liposomas* son vesículas artificiales preparadas con lípidos catiónicos que espontáneamente arropan al ADN en disolución y actúan como vehículos para introducirlo en la célula (figura 7.1). Generalmente no se preparan en el laboratorio, sino que se obtienen comercialmente. Se cree que los grupos fosfato cargados negativamente en el ADN se unen a las cargas positivas de la superficie de los liposomas y las cargas residuales positivas del liposoma lo hacen con los residuos de ácido siálico cargados negativamente de la membrana celular. Existen también liposomas de carácter aniónico que, en vez de arropar al ADN, lo engullen en su interior. La consecuencia práctica en todos los casos parece ser una fusión entre los lípidos artificiales y los componentes lipídicos de la membrana que provocan la internalización del ADN. Muy pequeñas cantidades de ADN (2 μ g) son suficientes para obtener eficiencias de transfección cercanas al 90%, sin embargo, altas concentraciones de ADN o liposomas pueden ser tóxicas.

Otros métodos, como la microinyección o el bombardeo genético, se utilizan preferentemente en determinados contextos. La microinyección nos permite introducir ADN, de manera directa e individual, en el interior de un núcleo celular ayudados por un micromanipulador y visualizando la operación en el microscopio. Es un método muy eficiente que se emplea para insertar material genético en óvulos fecundados de mamíferos con el fin de obtener animales transgénicos (capítulo 8). El bombardeo genético, por el contrario, se utiliza preferentemente en plantas o células con una pared celular consistente (capítulo 9, apartado 9.3 y figura 9.9); se basa en disparar con una pistola una nube de partículas de oro o tungsteno recubiertas de ADN.

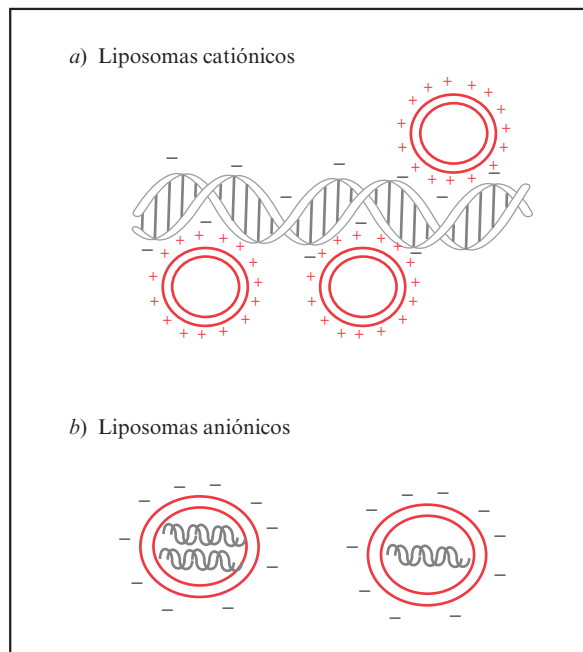


Figura 7.1. Estructura de los liposomas a) catiónicos que arropan el ADN, y b) aniónicos que engullen el ADN.

7.2. MÉTODOS DE SELECCIÓN Y GENES SEÑAL PARA TRANSFORMACIONES ESTABLES Y TRANSITORIAS

La llegada a una célula en cultivo de un gen foráneo ha de poder ser monitorizada de algún modo sencillo y eficaz. La mayoría de los genes que se introducen en células de mamíferos, no obstante, no confieren ninguna característica especial que distinga a la célula que lo ha incorporado. Se han desarrollado, por tanto, una serie de marcadores que ofrecen a la célula receptora una característica singular y diferencial.

Hay dos tipos de genes que nos permiten monitorizar su propia presencia en la célula recipiente: los genes de selección y los genes señal.

Los **genes de selección** confieren una ventaja de supervivencia a la célula portadora que no tendrá la célula que carezca de dicho gen. Ejemplos conocidos son *neo* (neomicina fosforribosiltransferasa)

o *puro* (puromicina acetiltransferasa). Estos últimos se denominan genes de selección positiva porque la presencia del gen permite la supervivencia en presencia del agente tóxico (G418 o puromicina, respectivamente); un gen de selección negativa sería *tkHs* (gen de la quinasa de timidinas del virus *Herpes simplex*) cuya presencia confiere sensibilidad (en vez de resistencia) al antibiótico ganciclovir.

Los **genes señal**, en cambio, sólo confieren una característica diferencial, que puede ser un color distinto, a la célula portadora como es el caso de *lac* (β -galactosidasa), *gus* (β -glucuronidasa) o *gfp* (del inglés: *green fluorescent protein*, proteína verde fluorescente). El gen de interés puede estar físicamente ligado o no al gen de selección o señal, aunque la primera opción es preferible.

7.2.1. Genes de selección estable

Es una selección a largo plazo donde las células transformantes han incorporado en su genoma el gen de interés.

a) Selección positiva por resistencia a antibióticos

Muchos de los antibióticos son aminoglicósidos relacionados entre sí que se unen a determinadas proteínas ribosomales bloqueando la síntesis proteica. Todos los aminoglicósidos pueden ser inactivados por las enzimas correspondientes aminoglicósido N-acetil fosfotransferasa.

El gen *puro* (de 0,6 kb, también denominado *pac*) codifica la enzima puromicina N-acetiltransferasa de *Streptomyces alboniger* y se usa como gen de selección en células de mamíferos, confiriendo resistencia al antibiótico puromicina (2 μ g/ml). La puromicina inhibe la síntesis proteica de células animales y bacterianas al instalarse en el sitio «A» del ribosoma (subunidad grande) como si fuera un aminoacil-ARNt; el enlace peptídico se realiza, pero se inhibe la elongación de la cadena y causa una terminación prematura.

El gen *neo* codifica la enzima neomicina fosforiltransferasa del transposón Tn5 y confiere resistencia a la neomicina en bacterias, kanamici-

na en plantas y G418 en mamíferos. Estos antibióticos se unen irreversiblemente a la subunidad pequeña del ribosoma. Las células de mamíferos transfectadas con *neo* sobreviven en presencia de hasta 400 μ g/ml de G418.

El gen *hygro*, obtenido originalmente de un plásmido de *E. coli* (pJR 225), confiere resistencia a la higromicina B (400 μ g/ml) al codificar una quinasa que inactiva el antibiótico por fosforilación. La higromicina B la produce la bacteria *Streptomyces hygroscopicus* e inhibe el proceso de translocación en la síntesis de proteínas de procariontes y eucariotas.

El gen *bs^R* codifica la enzima blasticidina deaminasa de *Bacillus cereus* y confiere resistencia a la blasticidina S. Este antibiótico es un peptidil-nucleósido aislado de *Streptomyces griseocromogenes* que bloquea la síntesis de proteínas al inhibir la formación del enlace peptídico en el ribosoma.

b) Selección por interferencias con el metabolismo de purinas y pirimidinas

El gen de la quinasa de timidinas, *tk*, dicta la síntesis de una enzima que fosforila timidinas y guaninas y las convierte en sus correspondientes derivados monofosfóricos (TMP y GMP). Éste ha sido el primer gen de uso generalizado para la selección en células de mamíferos.

Se pueden obtener células deficientes para esta quinasa de timidinas (*tk⁻*) por mutagénesis y posterior selección por resistencia al análogo tóxico 5' bromodesoxiuridina (BrdU). Estas células, en contraste con las de tipo silvestre, crecerán en un medio que contenga el análogo de timina BrdU, ya que no lo podrán incorporar en su ADN (figura 7.2). De este modo se pueden establecer líneas celulares *tk⁻* que no suelen revertir. Estas células son utilizadas en los experimentos de transfección en los que los vectores llevan el gen *tk*; y se bloquea la vía alternativa de síntesis de TMP con aminopterina, lo que se consigue creciendo las células en medio HAT: hipoxantina, aminiopterina y timina. Los mayores inconvenientes para su uso son: la necesidad de obtener una población receptora *tk⁻*, el que las necesidades celulares se suplan con cantidades mínimas del enzima y la frecuen-

cia con la que aparecen mutaciones espontáneas en este gen. De modo análogo se puede utilizar como gen de selección el correspondiente a la enzima que convierte la guanina en GMP. La vía alternativa de producción de GMP a partir de XMP se puede bloquear con 6-tioguanina cuando se use la enzima HPRT como selección (figura 7.2).

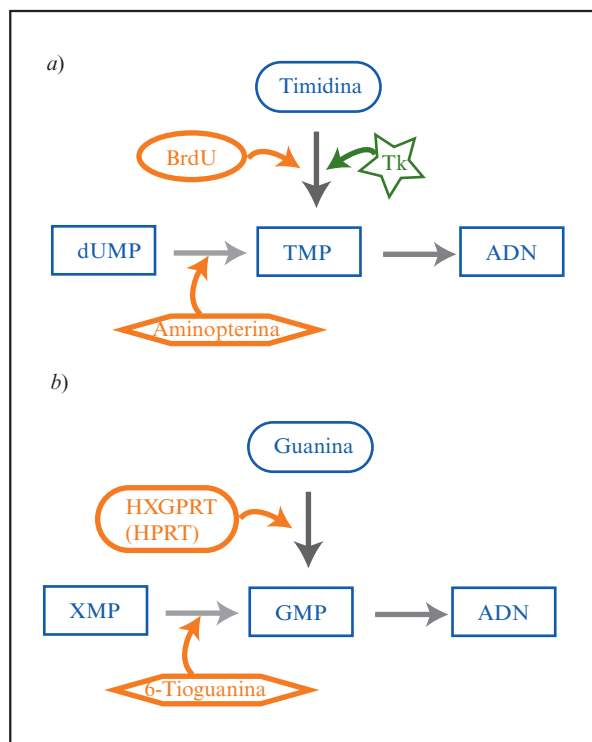


Figura 7.2. Esquema del lugar de acción de los distintos componentes de selección basado en: a) la quinasa de timidinas, Tk y 5 bromodesoxiuridina, BrdU; dUMP: deoxiuridina monofosfato; TMP: timidina monofosfato; b) hipoxantina-guanina fosforribosiltransferasa (HPRT); GMP: guanina monofosfato; XMP: xantina monofosfato.

El gen de la quinasa de timidinas procedente del virus *Herpes simplex* (*tkHs*) protagoniza un tipo de selección negativa algo distinta a la descrita. La enzima de origen viral no es tan estricta como su homóloga celular, y puede fosforilar aná-

logos de nucleósidos que no serían sustrato de la quinasa celular, como el ganciclovir o aciclovir (análogos de la guanina). Los precursores fosforilados se incorporan en el ADN durante la replicación paralizándola y causando la muerte celular. Así, células portadoras de vectores con el gen *tkHs* morirán en presencia del análogo tóxico con un IC_{50} de entre 0,1 y 1 $\mu\text{g/ml}$ del compuesto. El IC_{50} es la concentración a la cual la mitad de las células en cultivo no sobreviven.

7.2.2. Genes señal para transferencias estables

El producto del gen *gfp* (del inglés: *green fluorescent protein*) es la proteína verde fluorescente, un polipéptido de 238 aminoácidos muy fluorescente y muy estable en condiciones diversas. La proteína muestra una toxicidad muy baja y no interfiere con las actividades celulares rutinarias; tampoco precisa de cofactores o sustratos exógenos. Su detección es bastante simple en un microscopio de fluorescencia, la proteína absorbe luz azul fluorescente de 395 nm y la emite verde de 509 nm. Posee un fluoróforo cíclico responsable de la estimulación, formado por la serina en posición 65, la tirosina de la posición 60 y la glicina 67. La fluorescencia que emite el gen original no es excesivamente luminosa, por lo que se ha empleado mutagénesis dirigida para cambiar la serina 65 por una treonina, además de sustituir varios codones por otros equivalentes de uso más frecuente en células de mamíferos. La nueva proteína verde fluorescente se denomina potenciada (*enhanced* en inglés). En mezclas celulares, se pueden separar las células verdes de las no verdes con un separador de células (FACS; del inglés *fluorescence-activated cell sorting*). El gen *gfp* se aisló originalmente de la medusa *Aequorea victoria* y se expresa muy eficientemente en bacterias, plantas y mamíferos.

El gen *lacZ* de *E. coli* se traduce en la enzima β -galactosidasa, cuya actividad puede determinarse en procariotas y eucariotas, en presencia del sustrato X-gal. La detección de un color azul intenso con este sustrato precisa de la fijación de la

muestra celular con anterioridad al ensayo, lo que imposibilita la utilización de muestras vivas. Un sustrato alternativo es el di- β -D-galactopiranosido fluoresceína, que al ser procesado por la enzima permanece atrapado en la célula pudiendo detectarse por fluorescencia *in vivo*.

El gen que confiere resistencia al antibiótico ampicilina (*amp^R*) en *E. coli* codifica una β -lactamasa que rompe muy eficientemente penicilinas y cefalosporinas. Existe un sustrato fluorescente y permeable (CCF2/AM) muy sensible a la hidrólisis por esta β -lactamasa. Cada unidad de enzima puede actuar sobre muchas moléculas de sustrato cambiando la longitud de onda de su emisión fluorescente de verde a azul. El gen se expresa con eficiencia en células de mamíferos con promotores y terminadores apropiados.

7.2.3. Selección transitoria

Este tipo de selección se suele utilizar para determinar la eficacia de promotores que posteriormente se utilizarán con genes de interés. La selección transitoria nos permite detectar expresión del gen señal entre 24 y 72 horas después de la transfección. Puede combinarse con otra selección estable, ya que la transitoria se aplica antes de que sean efectivos los medios selectivos convencionales.

El gen *luc* es de los más utilizados como gen señal transitorio en células de mamíferos. Su producto, la enzima luciferasa, se aisló de la luciérnaga *Photinus pyralis* y emite luz amarilla al unirse a la luciferina, su sustrato. La reacción requiere ATP, Mg^{2+} y O_2 molecular, produciéndose un fogonazo de luz que decae rápidamente (en menos de un segundo), pero que puede ser captado por un luminómetro. La emisión total de luz es proporcional a la actividad luciferasa en la muestra.

7.3. ADN DESNUDO Y VECTORES

En células en cultivo podremos introducir ADN desnudo o cubierto con liposomas en ausencia total de vector. Es una práctica habitual

utilizar plásmidos procarióticos para la clonación y manipulación de los genes de interés procedentes de mamíferos. Estos plásmidos se introducen por transfección mediada por liposomas en células en las que la expresión del gen de interés sea requerida y se denominan coloquialmente vectores mixtos de expresión; una pequeña parte de estos recombinantes se integran en el genoma de la célula dando lugar a una transferencia génica estable y permanente.

No se conocen plásmidos que se mantengan de forma autónoma en células de mamíferos. Los «amplicones» constituyen lo más parecido al sistema de plásmidos en bacterias y son virus defectivos que no se integran en el genoma de la célula infectada y que pueden mantenerse como episomas citoplasmáticos durante algún tiempo (*SV40* o *Herpes simplex*). Los cromosomas artificiales también son vectores autónomos y constituyen la gran esperanza para introducir genes que podríamos llamar terapéuticos con sus propias secuencias reguladoras, a células de mamíferos.

Los vectores más característicos son derivados de virus. Se han elegido y desarrollado como vectores preferentemente virus con genomas de ADN o con una fase de su ciclo vital como ADN de hebra doble. Es conveniente, además, que tengan orígenes de replicación que multipliquen eficientemente sus genomas y promotores potentes, para conseguir un elevado porcentaje del gen clonado o de su producto y que, además, permitan una selección específica y eficiente en la célula infectada. Entre los virus más utilizados como vectores eucarióticos veremos los retrovirus, los lentivirus, los adenovirus y los adenoasociados; todos ellos se utilizan con frecuencia en diversas terapias génicas.

La transferencia de genes a células de mamíferos persigue varios objetivos: primero, el aporte de nuevo material genético que supla o complemente el endógeno; segundo, que el gen de interés llegue a un gran número de células específicas, y, por último, que se consiga una eficiente y exclusiva expresión en las células diana. Los vectores diseñados para conseguir esta triple finalidad que veremos en este apartado son: SV40 y vectores mixtos derivados, retrovirus, lentivirus, adenovirus, adenoasociados, derivados del herpes, el virus

de la vacuna y los cromosomas artificiales de mamíferos. Los requerimientos pueden variar ligeramente dependiendo de si los experimentos se realizan *in vitro* (células en cultivo) o *in vivo* (organismo completo).

7.3.1. SV40 y vectores mixtos derivados

SV40 (virus de los simios) tiene un pequeño genoma bicatenario de 5,2 kb superenrollado negativamente formando un minicromosoma de cromatina que carece de la histona H1. Pertenecce al género de los polyomavirus y a la familia de los papovavirus. Se le ha llamado el virus del poliovirus (del griego, *poli* = muchos, *oma* = tumores) porque produce tumores en hámsteres, pero no produce cáncer en el hombre. El genoma se divide en dos regiones: temprana y tardía, dependiendo del tiempo transcurrido para su expresión después de la infección. La región temprana comprende los dos genes que producen los antígenos T y t; T es imprescindible para la replicación del genoma. Puede seguir el ciclo lítico infectando células de mono como las de riñón COS, CV-1 o AGMK. En estos casos el virus se puede encontrar en el núcleo, pero no se integra. Mata a la célula infectada al producir numerosas partículas hijas que, a su vez, son altamente infecciosas. Las células se denominan entonces permisivas. El virus, no obstante, puede infectar células de algunos roedores, hámsteres por ejemplo, y no producir partículas infecciosas. El genoma del virus se integra, en estos casos, por recombinación ilegítima; el antígeno T sigue produciéndose (ya que regula su propia síntesis), y las células se transforman en tumorales. Las células del roedor se denominan, en esta situación, no permisivas para la replicación. Las células humanas son semipermisivas, ya que sólo del 1 al 2% de las infectadas producen partículas infecciosas. Hay dos ventajas en la utilización de este sistema de introducción de ADN recombinante en células de mamíferos: una de ellas es la eficiencia de transducción (infección), unas 100 veces mejor que los sistemas de transfección, y la segunda es que se pueden obtener entre 5×10^4 y 4×10^5 copias del ADN recombinante por célula,

acompañado de altos niveles de expresión de estas secuencias. El sistema de expresión es transitorio, puesto que acaba matando la célula huésped, pero durante 4 o 6 días produce grandes cantidades de la proteína clonada.

La mayoría de los vectores de expresión basados en SV40 carecen de casi todas las secuencias del genoma vírico original. Son minirreplicones a los que se les ha añadido secuencias procarióticas de replicación (origen de replicación) y selección (resistencia a la ampicilina) para poder crecerlos en un sistema procariótico. Poseen el origen de replicación de SV40 y los promotores temprano y tardío en orientaciones opuestas para la iniciación de transcripción; tienen incorporadas varias dianas de restricción para insertar ADN de interés y se pueden adquirir comercialmente. Los genes insertos en estos vectores se expresan transitoriamente y no hay producción de partículas virales. También se pueden obtener pseudopartículas por empaquetamiento *in vitro* del plásmido de expresión con proteínas de la cápside de SV40. Las proteínas de la cápside se obtienen, a su vez, de un extracto nuclear crudo de células de insecto infectadas con baculovirus portadores del gen de la proteína de la cápside sustituyendo al de la poliedrina (véase capítulo 6, apartado 6.6). El ensamblaje tiene lugar en un entorno controlado con componentes purificados eliminando el riesgo de contaminantes y la presencia de virus auxiliares infectivos.

Para estudiar la función de muchas proteínas es necesario diseñar su sistema de expresión en el medio habitual: la célula de mamíferos. Los **vectores mixtos** derivados de SV40 utilizan secuencias procarióticas para su crecimiento y secuencias de diversos virus animales para la expresión del gen de interés (figura 7.3). Poseen un promotor de diseño a escoger entre: el promotor de SV40, el de citomegalovirus, el de un retrovirus o el estimulable por tetraciclina; un epítipo o apéndice denominado FLAG que se fusiona al extremo 5' o 3' (en este caso al 5') y que permite localizar e inmunoprecipitar la proteína fusionada. FLAG es un heptaaminoácido: Asp-Tyr-Lys-Asp-Asp-Asp-Lys, reconocido por varios anticuerpos monoclonales; además los 5 últimos aminoácidos

son diana para la proteasa enteroquinasa que elimina el apéndice del producto final. El gen de interés se coloca a continuación en la región de policlonaaje y estará seguido por una región de poliadenilación (de SV40). Tendremos además tres orígenes de replicación compatibles: dos procarióticos (pBR322 y f1) y uno eucariótico (SV40). El gen de resistencia a ampicilina permitirá crecer el vector en bacterias. Si este gen fuera precedido por un promotor eucariótico, también podría usarse como gen señal en presencia del sustrato CCF2/AM.

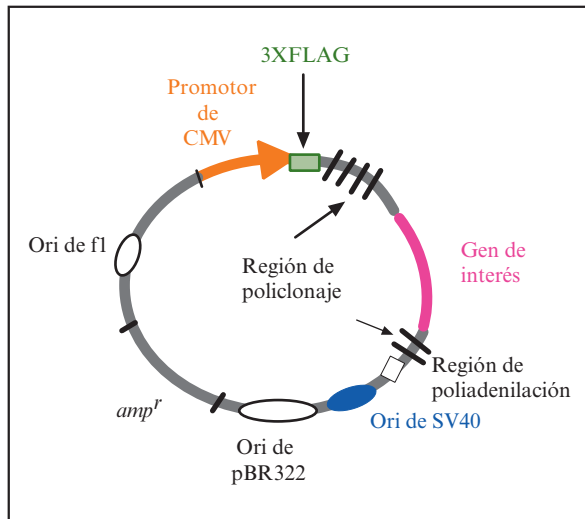


Figura 7.3. Vector mixto de expresión transitoria en mamíferos. El gen a expresar se coloca en la región de policlonaaje que está precedida por tres unidades FLAG y el potente promotor de citomegalovirus (CMV). A continuación del gen está la región de poliadenilación de SV40 y los orígenes de replicación del mismo virus y el de pBR322 para su crecimiento en eucariotes y procariotes, respectivamente. Además, posee el gen que confiere resistencia a ampicilina y el origen de replicación de f1.

Es interesante tener la posibilidad de regular la expresión del gen de interés con un promotor a medida, ya que el producto pudiera ser tóxico, por ejemplo. Idealmente se debería controlar desde una expresión nula del gen en condiciones

de represión a un elevado grado de expresión en condiciones de activación. El sistema que más se aproxima a esta situación ideal es el del promotor de la tetraciclina, aislado originalmente del transposón Tn 10 de *E. coli* y adaptado para su uso en células de mamíferos. La tetraciclina es un antibiótico producido por varias especies de *Streptomyces* y con efecto antibacteriano. El operón de la tetraciclina, tal como lo encontramos en *E. coli*, consta del gen *i-tet r* que codifica el represor (R) cuya unión al operador *tet-O* impide la expresión del gen tetraciclín-acetiltransferasa (*tat*) que inactivaría la tetraciclina. El antibiótico, sin embargo, tiene una gran afinidad hacia el represor, inactivándolo y permitiendo la expresión del gen de resistencia (figura 7.4a). Así pues, en presencia de tetraciclina se activará la expresión del gen que confiere a la bacteria resistencia al antibiótico. En las adaptaciones para su utilización en células de mamíferos se ha fusionado el gen *i-tet r* con el cajetín de activación de la proteína VP16 del virus *Herpes simplex*, convirtiendo así el represor en activador (*tTA*). Esta quimera transactivadora permite la expresión de cualquier gen situado a continuación de un promotor sintético que consta de varias copias del operador *tet-O* fusionado al promotor eucariótico de citomegalovirus. En presencia de tetraciclina, se produce un cambio conformacional en el transactivador, perdiendo éste su capacidad estimuladora e inhibiendo eficientemente la expresión del gen transferido (figura 7.4b). También existen variantes genéticas de *tTA*, obtenidas por mutagénesis dirigida y denominadas *rtTA*, que sólo se unen a la región operadora en presencia de tetraciclina.

Estos sistemas responden tanto a tetraciclina como a otros derivados, por ejemplo, la doxiciiclina. Se pueden obtener comercialmente plásmidos o retrovirus que llevan el promotor sintético de respuesta a la tetraciclina seguido de una región de policlonaaje. La fusión del represor con VP16 puede estar en otro plásmido o incluso ya incorporado al genoma de las células en cultivo. Las inhibiciones nunca llegan a ser totales, pero se logran estimulaciones de hasta 1.000 veces con relación a la situación reprimida.

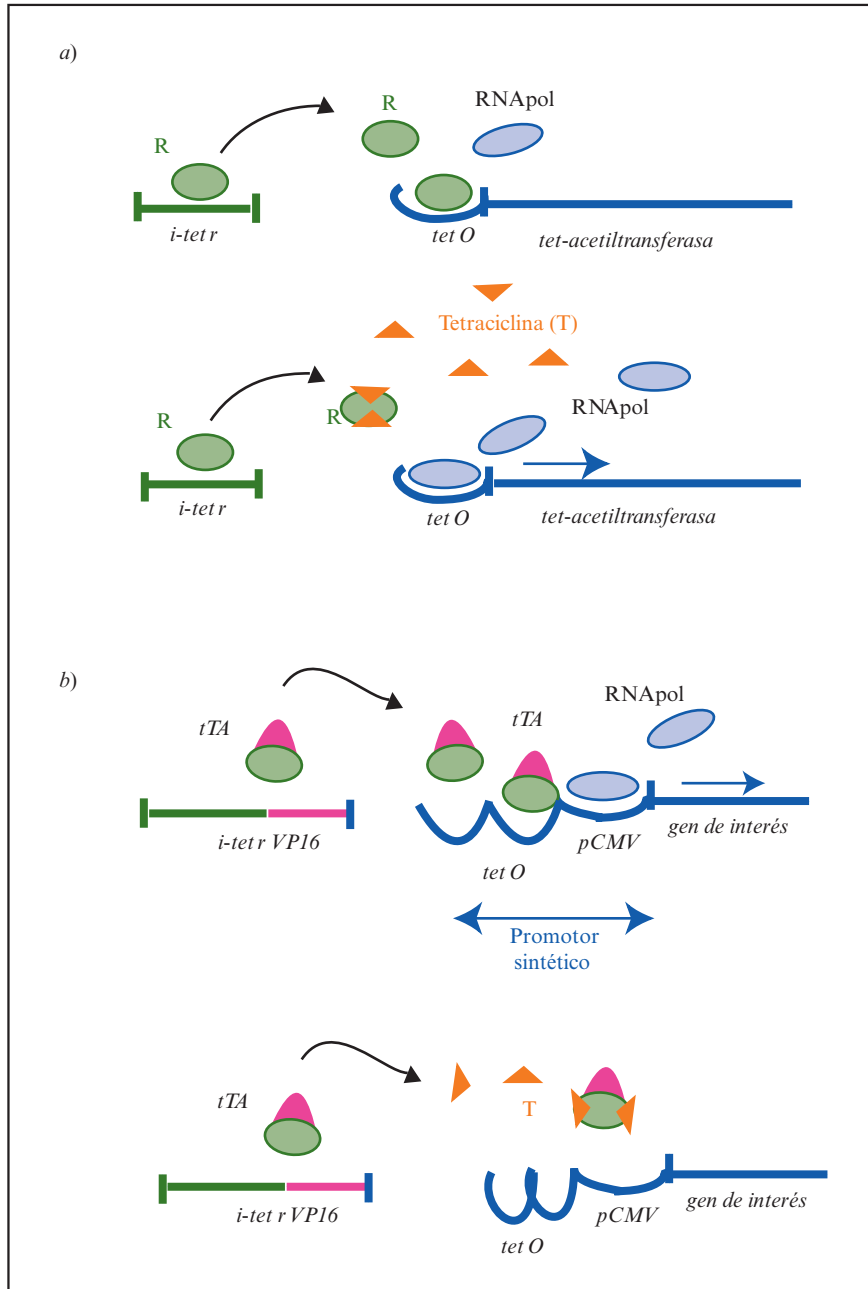


Figura 7.4. Promotor de expresión controlable positiva y negativamente por tetraciclina. *a)* Control positivo por tetraciclina tal como se encuentra en bacterias, y *b)* negativo adaptado a mamíferos, donde la presencia de tetraciclina inhibe la expresión del gen asociado. *tTA* es la proteína fusión entre el represor de tetraciclina (*R*) y el activador del gen 16 de *Herpes simplex* (VP 16). El promotor sintético consta de varias unidades del operador de *tet* (aquí se representan dos) fusionadas al promotor de citomegalovirus.

7.3.3. Retrovirus y lentivirus

La familia *Retroviridae* se divide en distintos géneros dependiendo de su estructura genómica y secuencia de ADN. Aquí veremos los dos géneros que se utilizan como vectores: 1) los **gammaretrovirus** con el virus de la leucemia murina como especie tipo (MLV) y 2) los **lentivirus** con el virus de la inmunodeficiencia humana tipo 1 como modelo (HIV-1).

γ -Retrovirus

El retrovirus prototipo de los γ -retrovirus posee una envuelta con una estructura de bicapa lipoproteica que incluye glicoproteínas codificadas por el gen *env* del virus. La nucleocápsida es esférica y está formada por tres subunidades proteicas codificadas por el gen *gag*. Su genoma son dos moléculas idénticas de ARN de cadena sencilla de un tamaño aproximado de 8 kb con estructura *cap* en el extremo 5' y una cola de poli (A) en el extremo 3'. El virus puede empaquetar genomas de entre 7 y 10 kb.

Se introduce en las células a través de los transportadores de fosfato de las membranas, pero no tiene capacidad para llegar al núcleo y, por tanto, ha de esperar a que la célula infectada entre en división en las próximas horas a su llegada al citoplasma, para acceder al genoma de la célula hospedadora donde se integra. Antes de su integración ha de convertir su genoma, que es ARN, en ADNc, lo que hace por mediación de la transcriptasa inversa codificada por uno de los genes que forman el complejo génico *pol*. También están codificadas por *pol* una integrasa y una proteasa. La síntesis del ADNc se inicia en la secuencia redundante (R) utilizando como cebador una molécula de ARN de transferencia. Tanto este cebador como la transcriptasa inversa no necesitan ser traducidos por el genoma viral, sino que algunas moléculas están incluidas en la nucleocápsida procedentes de la infección anterior. El ADNc se circulariza e integra preferentemente en genes activos, ya que éstos tienen una configuración de la cromatina mucho más accesible (de 1 a 10 copias). El ARN genómico posee

una secuencia redundante (R) de 10-80 pares de bases en ambos extremos, seguida por secuencias únicas características de cada extremo: U5, única del extremo 5' y U3, única del extremo 3'. U5 es necesaria para la retrotranscripción a ADNc y U3 para la integración. Sin embargo, el provirus tiene como extremos dos terminales repetidos largos (TRL) de varios cientos de pares de bases, en los que U3 está copiado en el extremo 5' y U5 en el 3' (figura 7.5). Así el tamaño del provirus es siempre mayor que el del genoma viral.

La transcripción se inicia en la región R del terminal 5' a partir de un promotor de U3, y se termina en la misma región R del extremo 3'. El promotor es muy potente y se ha incorporado en numerosos vectores mixtos donde se persigue preferentemente la expresión del gen clonado.

Los tres paquetes de genes codificados por el genoma retroviral son: *gag*, que produce 3 proteínas: matriz, cápside y nucleocápside, estas proteínas son características del grupo; *pol*, que codifica la polimerasa de transcripción invertida, una proteasa y la integrasa, y *env* responsable de dos proteínas de la envuelta. Este último gen se traduce a partir de un ajuste que elimina los genes *gag* y *pol*.

El retrovirus no mata a la célula infectada, sino que establece una simbiosis indefinida con la liberación de una a diez partículas virales por día en cultivos celulares.

Los vectores más utilizados son derivados del virus de la leucemia murina. Estos vectores se desarrollaron con tres finalidades: 1) aumentar la eficiencia de transfección; 2) obtener un mayor número de transformantes estables, ya que el virus se integra en el genoma de la célula, y 3) aumentar la diversidad de células transfectables, pues estos virus pueden infectar una gran variedad de células. El mayor inconveniente es que potencialmente pueden producir diversas enfermedades, entre ellas cáncer.

La manipulación de los vectores derivados de retrovirus se realiza sobre la forma replicativa, que es ADN de doble hebra. Los vectores resultantes carecen prácticamente de todos sus genes originales (*gag*, *pol* y *env*), pero mantienen la se-

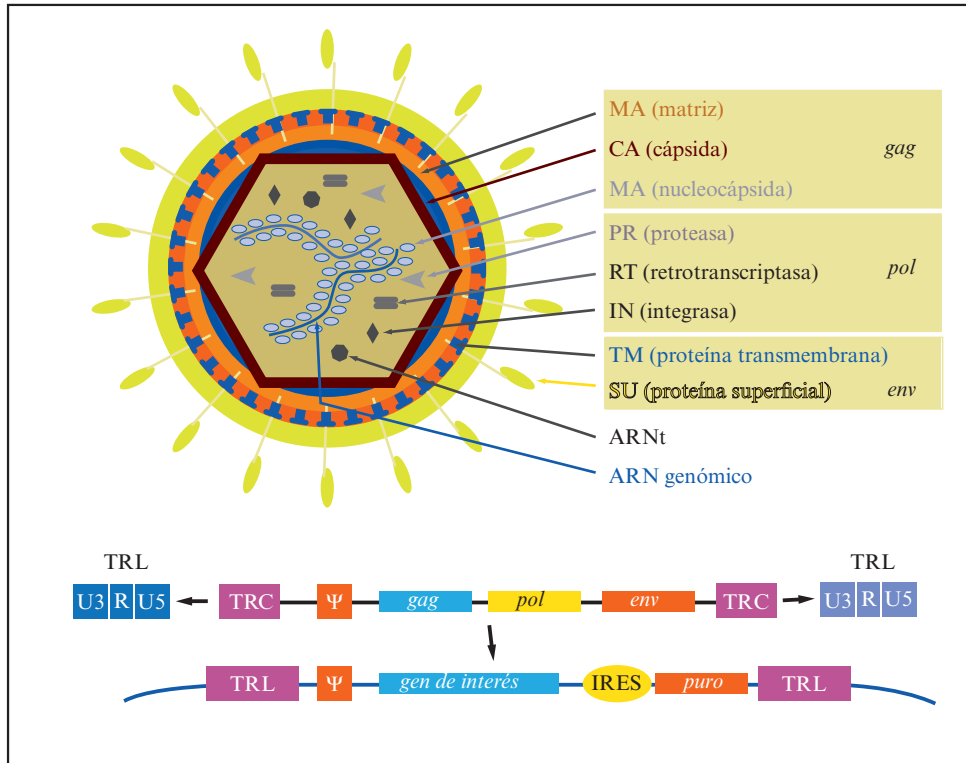


Figura 7.5. Esquema de una partícula retroviral derivada del virus de la leucemia murina. Cada partícula posee dos moléculas de ARN viral que constituye el genoma, la transcriptasa inversa que le permite copiarse a ADNc y unas moléculas de ARNt que utiliza como cebadores o iniciadores en la síntesis de ADNc. La partícula está recubierta por una envuelta proteica con proteínas específicas que permiten reconocer y asociarse a moléculas transportadoras de fosfatos que actúan como receptores de entrada de retrovirus en numerosas células de mamíferos, entre ellas las humanas. Se muestran los tres paquetes de genes *gag*, *pol* y *env* que se sustituyen por el gen terapéutico, un IRES y un gen de selección (*puro*) en terapias génicas. También se muestra la región *psi* (Ψ, del inglés *packaging site*), que es una secuencia de reconocimiento indispensable en el empaquetamiento.

cuencia *psi* Ψ (del inglés: *packaging site*). En sustitución de los genes propios se coloca el gen de interés cuya transcripción queda controlada por el potente promotor del extremo 5' TRL y un gen que nos permita una selección eficiente como *puro* (puromicina acetiltransferasa) precedido por una secuencia de iniciación interna para ribosomas (IRES, del inglés *internal ribosomal entry site*) que garantiza una expresión igual a la del gen de interés.

El provirus manipulado es insertado en un plásmido para su propagación y crecimiento en

E. coli y se utiliza para transfecciones permanentes en células de mamíferos (plásmido retroviral). Se trata de un vector defectivo e incapaz de llevar a cabo una infección con producción de partículas víricas, por lo que necesita un virus auxiliar que aporte los productos *gag*, *pol* y *env* en trans. La presencia de un virus auxiliar no está indicada si se utilizan estos vectores para terapias *in vivo*, ya que tendríamos una proporción variable de virus no defectivos capaces de extenderse a diversas partes del organismo. Las denominadas células empaquetadoras suplen la presencia de los

virus auxiliares porque llevan en su propio genoma *gag*, *pol* y *env* procedentes del γ retrovirus; los genes pueden estar en localizaciones dispersas para evitar recombinaciones. Así, células empaquetadoras infectadas o transfectadas por un plásmido retroviral se convierten en «productoras» de retrovirus defectivos portadores del gen de interés

(figura 7.6). Las células productoras son como fábricas del retrovirus recombinante. En cultivos, se pueden alcanzar unos títulos en el medio de cultivo de 10^6 a 10^7 partículas infectivas por mililitro. Las ventajas e inconvenientes de los retrovirus como vectores se presentan más adelante en la tabla 7.1.

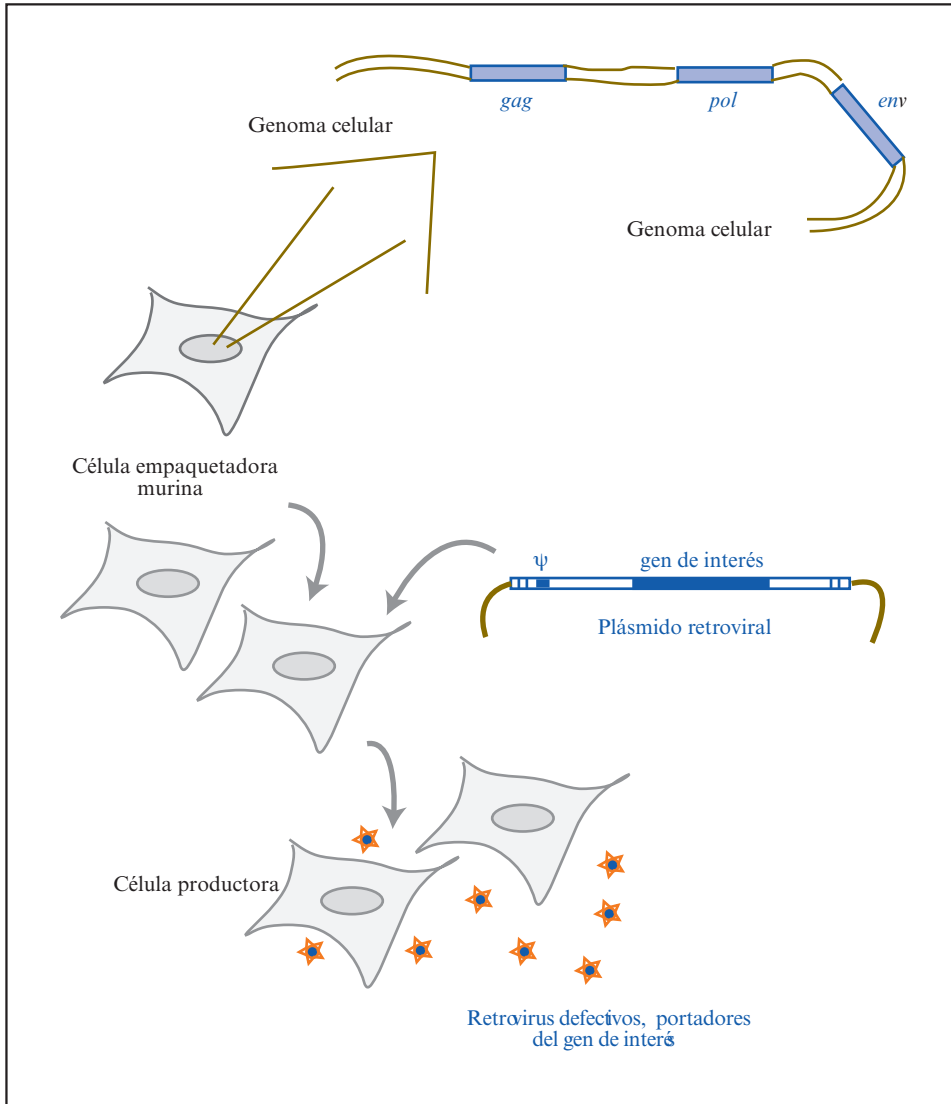


Figura 7.6. Conversión de células empaquetadoras murinas en células «productoras» de retrovirus defectivos portadores del gen de interés por transfección con un plásmido retroviral crecido en bacterias.

Lentivirus

El género lentiviral de la familia *Retroviridae*, es algo más evolucionado que el γ -retroviral, ya que sus miembros tienen la capacidad de llegar al núcleo y, por tanto, no dependen de que la célula infectada entre en división para acceder al genoma celular como les ocurre a los retrovirus. Se denominan lentivirus porque tienen periodos de incubación largos. Sus miembros se asocian con enfermedades crónicas del sistema inmune y sistema nervioso central, siendo la más conocida en humanos el síndrome de inmunodeficiencia adquirida (SIDA). El virión del VIH tiene una envuelta formada por una bicapa lipídica que deriva de la membrana de la célula previamente infectada. El núcleo de la partícula contiene, además de las dos moléculas de ARN de 9,2 kb cada una que constituyen el genoma, algunas moléculas de proteasa, transcriptasa inversa e integrasa como los retrovirus, pero el genoma de los lentivirus es más complejo que el de los retrovirus. Además de los 3 complejos génicos típicos de los retrovirus: *gag*, *pol* y *env*, los lentivirus poseen 6 genes reguladores: *tat*, *rev*, *nef*, *vif*, *vpr* y *vpu*. Las funciones más importantes de cada una de estas proteínas son: Tat, regula la transcripción y elongación del ARN mensajero; Rev, regula el ajuste, transporte, procesamiento y traducción de los ARNm virales; Nef, aumenta la infectividad viral, regula la entrada en linfocitos T e induce apoptosis; Vif, regula

la infectividad e integración; Vpr, regula el transporte del ADNc al núcleo, y Vpu, regula y facilita la liberación de los viriones. Dentro de la partícula viral, van incluidas algunas moléculas de los productos de los genes Nef, Vif y Vpr procedentes de la infección anterior. En el hombre VIH infecta los órganos linfoides donde los virus permanecen durante largos periodos de tiempo. Tiene tropismo hacia los linfocitos T-CD4⁺, siendo éstos los responsables de la producción de los viriones circulantes de los pacientes infectados. El virus, sin embargo, destruye los linfocitos T, ya que alguno de sus productos, como la proteína de la envuelta, es tóxico para la célula, con lo que se liberan nuevas partículas que a su vez infectan los nuevos linfocitos T que se producen para suplir los destruidos. Además los T-CD4⁺ infectados se transforman en dianas del propio sistema inmune, ya que expresan péptidos virales en sus moléculas del sistema de incompatibilidad de tipo I (HLA) y son reconocidos y destruidos por linfocitos citotóxicos y células asesinas natas (NK, del inglés *natural killers*).

Se pueden obtener comercialmente plásmidos lentivirales derivados del VIH-1 en los que, como inserto en un plásmido de tipo pUC, encontramos un provirus con los siguientes componentes (figura 7.7): dos terminales repetidos largos en los extremos; la región de empaquetamiento ψ ; un elemento que suple al ajuste eliminado dando estabilidad a los mensajeros y que se denomina

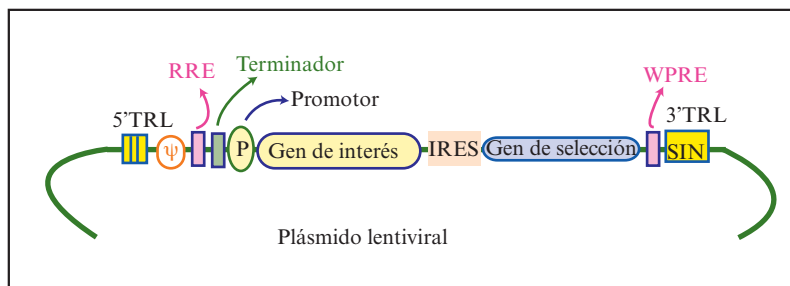


Figura 7.7. Vector lentiviral. Desde el extremo 5' al 3' encontramos los elementos: terminal repetido largo 5', la región de empaquetamiento ψ ; un elemento RRE que suple al ajuste; una secuencia de terminación de la transcripción; promotor de diseño; el gen de interés; IRES; un gen de selección a elegir; WPRE, elemento que actúa en cis, y terminal repetido largo 3' deleciónado (SIN) para eliminar el promotor.

RRE (del inglés, *Rev response element*); una secuencia de terminación de la transcripción para eliminar cualquier ARN que se pudiera originar en el promotor del TRL no controlado; un promotor de diseño a elegir entre varios posibles como el de citomegalovirus, tetraciclina, etc.; el gen de interés; IRES o un segundo promotor; un gen de selección a elegir entre *neo*, *puro* o *bs^R* y WPRE (del inglés, *Woodchuck hepatitis posttranscriptional regulatory element*) que es un elemento que actúa en cis y junto con RRE estabilizan el mensajero y aumentan su expresión. Quizá en este vector un IRES es mejor elección que un segundo promotor, ya que dos promotores cercanos pueden inducir una cierta interferencia.

El gen *rev* se facilita in trans en las empaquetadoras junto con los otros genes esenciales (*gag*, *pol* y *env*). La glicoproteína Env de los vectores VIH-1 tiene serias restricciones de hospedadores, por lo que en las empaquetadoras el gen *env* propio se ha sustituido por el del virus de la estomatitis vesicular (lo que se denomina pseudotipado), aumentando así su rango de organismos a infectar y facilitando su concentración por ultracentrifugación. Los títulos de estos vectores en el medio de cultivo son de 10^6 - 10^7 partículas infectivas por mililitro.

7.3.4. Adenovirus

Los adenovirus deben su nombre a haberse encontrado por primera vez en tejido adenoide humano (1953). Por aquellos años se utilizaba como fuente de tejido humano, fácil de cultivar *in vitro*, el tejido adenoide procedente de niños (o adultos) a los que se les quitaban las vegetaciones o amígdalas. En muchas ocasiones, sin embargo, las células morían misteriosamente al cabo de un cierto tiempo en cultivo. Se hizo evidente que la causa de las misteriosas muertes era un grupo de virus relacionados, residentes habituales de las vías respiratorias. Se los denominó adenovirus. Estos virus producen con frecuencia infecciones asintomáticas o muy leves, como catarros, conjuntivitis o gastroenteritis, que nuestro sistema inmune elimina en unos pocos días.

Existen unos 42 tipos distintos de adenovirus, denominados serotipos porque estimulan la producción de anticuerpos distintos y se relacionan con afecciones también diferentes. Se dividen en subgrupos (A-F) dependiendo de sus propiedades físicas, químicas, biológicas e inmunológicas. El subgrupo A, por ejemplo, produce tumores en hámsteres con una alta frecuencia (80%); el B también lo hace, aunque con menor frecuencia (20-60%), y el C no induce tumores. Todos ellos tienen ADN como genoma, pero pueden variar en tamaño desde 4 kb a 200 kb; el ADN puede ser de banda simple, doble, lineal o circular.

De todos ellos, sólo se utilizan como vectores los serotipos Ad 2, Ad 4, Ad 5 y Ad 7, pertenecientes todos al grupo C. El Ad 4 y el Ad 7 se han utilizado como vacunas en la prevención de enfermedades respiratorias agudas, ya que se ingieren en una pastilla y provocan una infección asintomática a la vez que confieren inmunidad para las afecciones respiratorias provocadas por otros adenovirus. Son icosaédricos, carecen de envuelta y poseen unas fibras terminales por las que toman contacto con las células a infectar. El receptor primario es el elemento CAR (del inglés, *coxsacki adenovirus receptor*), que es una proteína que atraviesa la membrana, de la familia de las inmunoglobulinas, que funciona como molécula de adhesión en células epiteliales. CAR está presente en muchas células de distintos tejidos como corazón, pulmones, hígado y cerebro.

El más utilizado y manipulado como vector de transferencia génica en mamíferos y humanos es el **Ad 5**. Este virus tiene 36 kb de ADN de doble banda como genoma y admite insertos de 7,5 kb. El virus entra por endocitosis, favoreciendo el proceso la presencia de integrinas y heparán sulfatos, la diferencia de pH con el medio celular provoca la liberación del contenido genómico que se asocia con microtúbulos y viaja al núcleo donde se transcribe (figura 7.8). Lo primero que se transcribe es la región E1A (E, del inglés *early*) cuyo producto se une a las proteínas celulares retinoblastoma (Rb) y p53 que desencadenan la fase de síntesis de ADN, lo que favorece la replicación del ADN viral. Además, el producto de E1A estimula la transcripción de las regiones E1B, E2A,

E2B, E3 y E4 (figura 7.9a). Las proteínas tardías (unas 17) salen todas ellas de un único mensajero de 28 kb transcrito de la banda complementaria a la que origina E2 y E3. Se denominan L1 a L5 (del inglés *late*), siendo las proteínas de la cápsida algunas de ellas. Los virus se ensamblan en el núcleo produciéndose la lisis de la célula que libera

las partículas. En ningún momento se integra en el genoma de la célula infectada. Posee, además, una región de empaquetamiento ψ que coincide parcialmente con E1A y en los extremos tiene unas terminaciones repetidas invertidas (TRI) de 100-140 pb. Produce títulos de 10^{11} partículas formadoras de placas de lisis, por mililitro (pfl/ml).

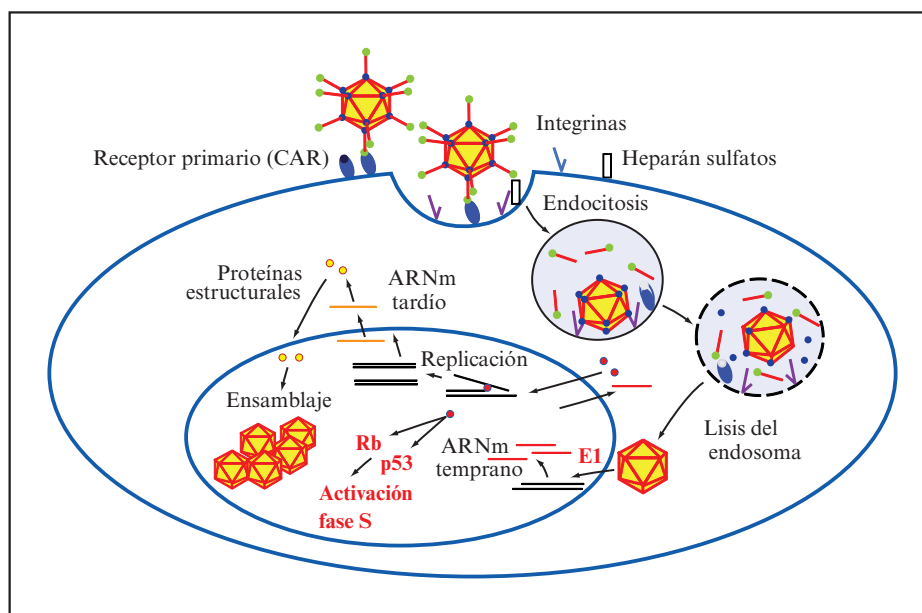


Figura 7.8. Ciclo vital del adenovirus Ad5 (cortesía de Ramón Alemany) Entrada mediada por el receptor CAR, integrinas y heparán sulfato. Lisis del endosoma y liberación de las partículas en el citoplasma. Desplazamiento al núcleo y transcripción de los genes tempranos (E_1). Inhibición de RB y p53 de la célula por E_1 . Activación de la fase S y de la replicación del virus. Traducción de las proteínas estructurales del virus en el citoplasma y ensamblaje en el núcleo. Finalmente, lisis de la célula y liberación de partículas virales.

Para utilizarlo como vector conviene, por razones de seguridad, eliminar la región E1A, con lo que se convierte en defectivo y ya sólo podrá replicarse en presencia de un virus auxiliar o en células empaquetadoras que lleven el gen o los genes delecionados. E3 y/o E4 también se eliminan en muchos vectores para crear espacio ocupable por el gen de interés (figura 7.9b). La función de E3 es reducir la inmunidad que provoca en el organismo infectado ya que interfiere con la síntesis o exposición de las proteínas inmunogénicas que

saca a la superficie el sistema de histocompatibilidad de tipo I (HLA). Sería de gran interés para las terapias *in vivo* no eliminar este gen, pero desafortunadamente al eliminar E1, E3 no se puede expresar porque está regulado por E1. Se han construido algunos vectores para uso de terapias génicas en humanos que delecionan E1 y E4 manteniendo E3 bajo el control de un promotor independiente. También se han desarrollado los vectores sin entrañas (*gutless*, en inglés), en los que se han eliminado todos los genes del adenovirus para

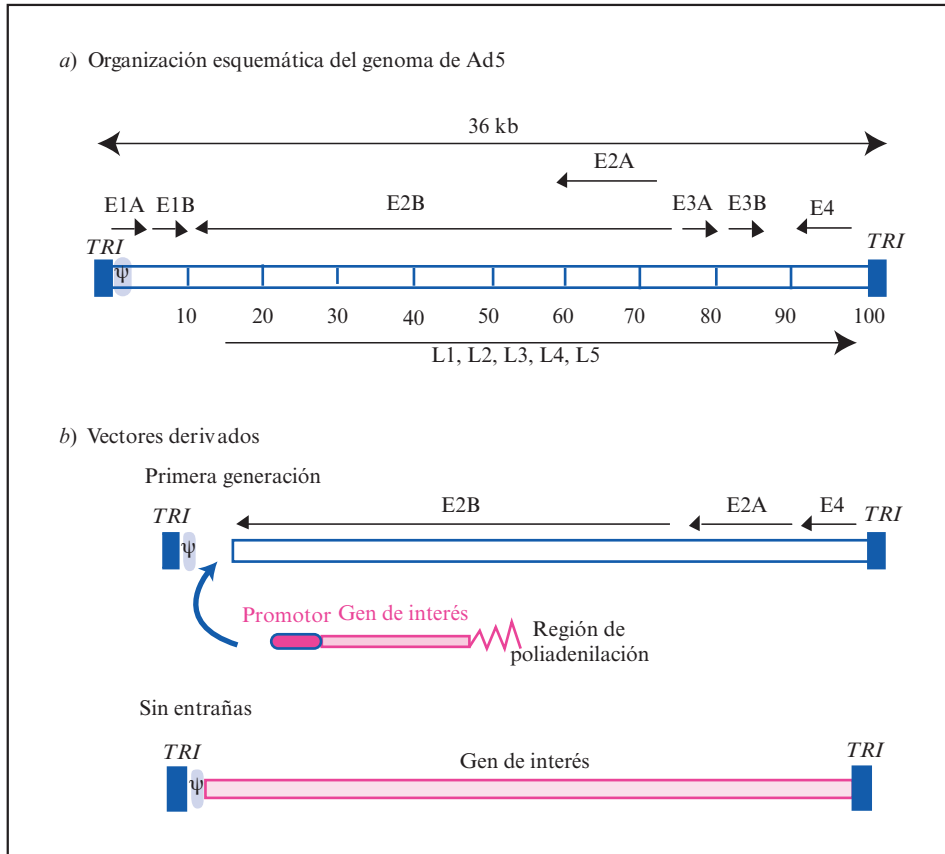


Figura 7.9. Los adenovirus y su genoma. a) Organización esquemática del genoma de Ad5 representado en porcentaje del tamaño total (36 kb) y mostrando la localización de los genes tempranos (E1-E4) y los tardíos (L1-L5); b) vectores derivados: de primera generación y sin entrañas.

dejar sitio al gen de interés. Estos vectores precisan de un virus auxiliar que les facilite las funciones eliminadas, pero admiten insertos de 36 kb.

El gran inconveniente de los adenovirus como vectores *in vivo* para terapias génicas, es que generan una fuerte respuesta inmunológica.

7.3.5. Adenoasociados

Los virus adenoasociados pertenecen a la familia *Parvoviridae* y al género *Depedovirus* porque requieren la coinfección con un adenovirus o herpesvirus, por ejemplo, para desarrollar un ciclo

productivo. Son virus sin envuelta que poseen una cápsida icosaédrica y un pequeño genoma de 4,7-5,1 kb de ADN con una sola hebra. Se han aislado del tracto respiratorio o gastrointestinal humano no conociéndose patologías asociadas a su presencia. Pueden infectar células quiescentes utilizando heparán sulfato como receptor primario y el factor de crecimiento de fibroblastos y la integrina $\alpha\beta_3$ como correceptores. En ausencia de virus auxiliar, se integran, por recombinación específica de sitio, en un lugar del brazo largo del cromosoma 19 (19q13.3-qter) quedándose en forma latente hasta que ocurra una infección por un adenovirus, herpes, vaccinia o papilomavirus. El

genoma viral posee en sus extremos secuencias palindrómicas de 100 a 250 nucleótidos que forman horquillas terminales relativamente estables en una determinada configuración. Las horquillas contienen el origen de replicación del ADN y las señales para el empaquetamiento del genoma viral en las cápsides. Forman unos terminales repetidos invertidos (*TRI*) de 145 nucleótidos que son los únicos elementos requeridos en *cis* (en la misma unidad y proximidad física) para la replicación, la integración y el empaquetamiento. Ade-

más poseen 3 promotores: dos (*P5* y *P19*) desde los que se transcribe el gen *rep* que dará lugar a los mensajeros de las proteínas Rep, necesarias para la integración y replicación viral, y un promotor más tardío (*P40*) desde donde se transcribe el gen *cap* responsable de las proteínas de la cápsida (*VP1*, *VP2* y *VP3*).

La construcción de un vector adenoasociado (*VAA*) se realiza en la forma replicativa de banda doble, eliminándose las regiones *cap* y *rep* que se sustituyen por el gen deseado (figura 7.10). Todo

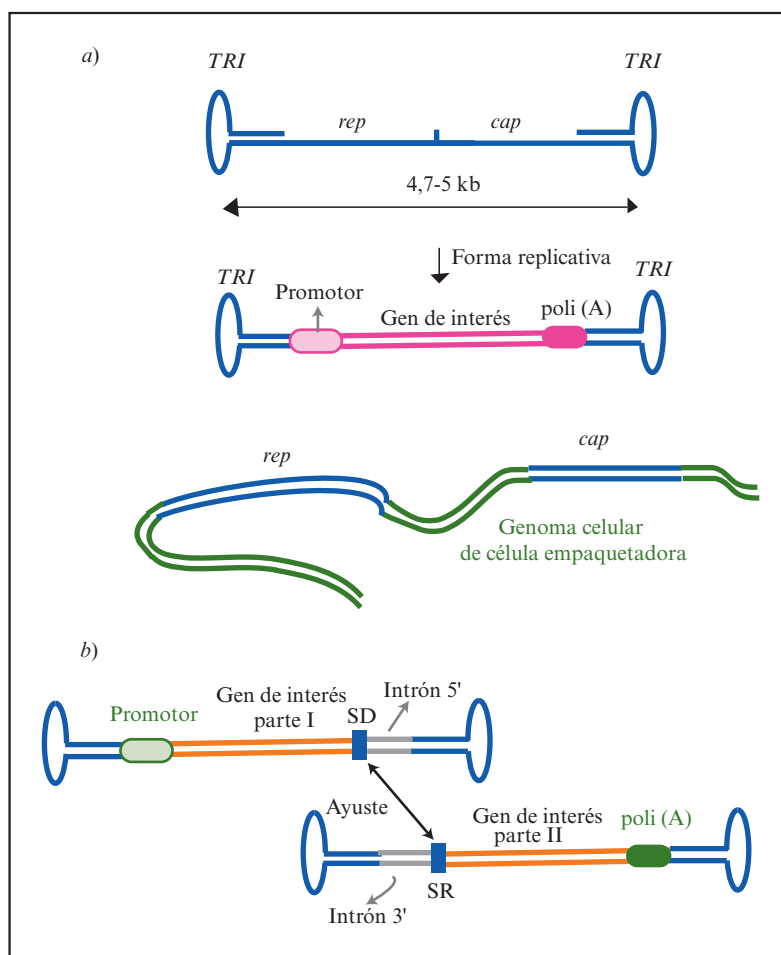


Figura 7.10. Virus adenoasociado y vectores derivados. *a)* Esquema del virus con *TRI* en forma de cabeza de martillo en los extremos, genes *cap* y *rep* en el interior y vector derivado de la forma replicativa portador de un gen de interés. *b)* Estrategia de un gen de interés en dos mitades: *SD* y *SR* son las secuencias donante y receptora de ajuste, respectivamente.

ello se mantiene en un plásmido procariótico para su crecimiento y manipulación. Con el vector recombinante VAA, se transfectan células empaquetadoras (Hela o 293) que faciliten en trans los genes *rep* y *cap* eliminados del vector y presentes en el genoma de la empaquetadora. Además es necesaria una coinfección con un virus auxiliar (adeno o herpes) o un plásmido que aporte los genes esenciales de éstos. El resultado es una mezcla de vector, virus auxiliar y proteínas contaminantes. El vector portador del gen de interés puede ser aislado del resto por centrifugación en gradientes isopícnicos, pero es difícil conseguir títulos altos de este recombinante (bajan de 10^{10} a 10^6). Una de las ventajas de los adenoasociados frente a los adenovirus es su menor capacidad para inducir una respuesta inmune, aunque ésta no es nula. La integración específica en el cromosoma 19 humano se pierde al convertirlo en vector, con lo que el vector rara vez se integra en el genoma.

La poca capacidad de empaquetamiento de estos vectores (5 kb) se puede suplir con la estrategia de formación espontánea de genomas dimericos en tándem a lo largo de la infección. Se utilizan dos plásmidos que contienen cada uno una mitad del gen de interés y dos *TRI*. La formación del dímero provoca, mediante ajuste, un ARNm correspondiente al gen de interés completo y, por tanto, de tamaño superior a 5 kb y menor de 10 kb.

7.3.6. Vectores derivados del Herpes

La familia de los *Herpes* incluye, entre otros, el *Herpes simplex* tipos 1 y 2; el *Varicella-zoster*, el *Epstein-Barr* y el *Citomegalovirus*. Los vectores derivados de *Herpes simplex* (VHS) son interesantes porque tienen tropismo hacia el sistema nervioso que permitirá utilizarlos como vehículos para depositar material genético en estas células. Además permiten insertos de más de 30 kb. El virus invade neuronas de los ganglios sensoriales donde se puede establecer como un elemento extracromosómico (episoma, plásmido, amplicón) durante largos períodos de latencia. *Herpes* deriva del griego y significa «reptar o gatear» haciendo

alusión al aspecto que toma la piel con algunas de las lesiones comunes. En el hombre produce patologías variadas: calenturas en la comisura de los labios (tipo 1), conjuntivitis, meningitis, infecciones genitales (tipo 2) o incluso encefalitis; pero tiene una gran capacidad para esconderse y pasar inadvertido durante años. Aunque es muy inmunogénico cuando produce infecciones, en los estados de latencia no estimula respuesta alguna, tampoco interfiere con las funciones celulares. Más del 80% de la población ha estado en contacto con el virus en algún momento de su vida y probablemente lo albergue de por vida tras la infección primaria; el contagio más frecuente es oral (a través de la saliva). El virus ataca los terminales sensoriales nerviosos (neuronas) que existen bajo la piel en la cavidad oral y que terminan en el ganglio trigémino. También infecta la mucosa genital y se transmite vía sexual, permaneciendo en muchos casos asintomático. El virus entra por interacción de las glicoproteínas de la envuelta viral con receptores de la membrana celular y se dirige al núcleo donde transcurre la mayor parte de su ciclo vital. Aquí se circulariza y se transcribe con la polimerasa II de la célula infectada, pero con factores de transcripción del propio virus que controlan todos los estadios del ciclo de infección. Este virus posee algunos genes de funciones muy parecidas o idénticas a otros genes celulares, como es el caso de la quinasa de timidinas (*tkHs*) o ribonucleótido reductasa. Nos podríamos preguntar por qué este virus no utiliza los productos celulares en vez de utilizar el propio genoma para codificarlos. Una posible explicación es que, al no estar las células en reposo expresando estos genes ligados a la replicación, el virus prefiera llevarse los puestos a tener que inducirlos, lo cual retrasaría su replicación.

Uno de los grandes inconvenientes de este vector es su tamaño, de 160 kb, y su difícil manejo en el laboratorio. Su genoma es un ADN lineal, bicatenario protegido por una cápsida. Alberga 81 genes de los cuales la mitad no son esenciales, pudiendo introducir hasta 50 kb de inserto. Otro inconveniente es la citotoxicidad asociada a este virus que acaba matando todo tipo de células donde entra. Esta toxicidad puede disminuir en

mutantes delecionados en genes esenciales inmediatamente tempranos (del inglés: *immediate early*, IE) y que se deben suplir en trans en la línea celular. Al convertirlo en vector también se eliminan muchos otros genes no esenciales que reprimen la expresión de genes en la célula infectada o tienen funciones redundantes. El recombinante debe, asimismo, tener la capacidad de expresar el gen de interés en condiciones de latencia en las que la gran mayoría del genoma del *Herpes* está silenciado; para ello se ha colocado dicho gen detrás de uno de los dos promotores de latencia naturales (LAP1 o LAP2) que no son muy eficientes (LAP1 más fuerte que LAP2), pero permiten una expresión prolongada (meses) en la célula.

Los amplicones son un tipo de vectores derivados del genoma de *Herpes* que contienen, junto con el gen de interés, los orígenes de replicación de *Escherichia coli* y del virus, así como las secuencias de empaquetamiento también de *Herpes*. El amplicón en su conjunto tiene unas 15 kb, se crece en bacterias y se cotransfecta con un VHS defectivo que actúe de virus auxiliar. La cotransfección se lleva a cabo en células empaquetadoras que complementan la deficiencia. *Herpes* sólo empaqueta genomas de su tamaño, por lo que nos empaquetará, por una parte, concatémeros del genoma del amplicón que a su vez contienen el gen de interés, y, por otra, el virus auxiliar defectivo. Este virus auxiliar defectivo puede ser sustituido por otro plásmido que lleve las funciones requeridas para el empaquetamiento y la replicación, pero en estos casos el título (partículas liberadas por mililitro de cultivo) baja considerablemente (10^4 - 10^5 unidades formadoras de colonias por mililitro).

Los PAM (plásmidos artificiales de mamíferos) son derivados de los cromosomas artificiales de bacterias (CAB), a los que se les añade el origen de replicación del estado latente del virus herpético *Epstein-Barr* (ori P) y el transactivador EBNA-1. EBNA-1 se une de manera específica a *oriP* en dos lugares distintos iniciando la replicación y puede incluirse en el plásmido o ser facilitado en trans por previa integración en el genoma de la célula hospedadora. También se incluye un gen de selección para células de mamíferos. Un

inserto de 185 kb que albergaba el pequeño gen de la β -hemoglobina y su enorme región reguladora se ha podido mantener establemente en un PAM durante más de un año en células humanas transformadas vía liposomas. El número de copias por núcleo se mantuvo entre 11 y 15, expresándose el gen a niveles relativamente altos.

7.3.7. El virus de la vacuna

El virus de la vacuna se ha utilizado durante 200 años para vacunar contra el virus de la viruela. La viruela no existía en América y fue introducida por los conquistadores españoles causando miles de muertes entre los indígenas de México, Perú y el Caribe; de aquí se extendió por Brasil y toda Latinoamérica. El descubrimiento y aplicación del virus de la vacuna contra la viruela se debe a Jenner quien, en 1776, observó cómo las mozas de cuadra que ordeñaban vacas en Gloucester (Inglaterra) adquirían una forma muy atenuada de la viruela que atacaba a las vacas (viruela vacuna). También observó que estas mozas nunca contraían la viruela humana cuando ocurrían grandes epidemias. Algunos años más tarde se aisló el virus que causaba la viruela vacuna y se utilizó para inmunizar al ser humano frente al virus de la viruela. La palabra «vacuna» ha permanecido como referencia de inmunización frente a virus o bacterias de diversa índole y, curiosamente, al estar la viruela totalmente erradicada, ya no se precisa la utilización del virus de la vacuna como agente inmunogénico contra su pariente próximo.

El genoma del virus de la vacuna es muy grande, posee 180 kb de ADN bicatenario y no se integra en el genoma de la célula que infecta, ya que se divide en el citoplasma celular. Hay ciertos aspectos del virus de la vacuna que le hacen particularmente importante en ingeniería genética: es probablemente un virus seguro y su abanico de hospedadores es muy extenso, ya que infecta tanto a mamíferos como a una gran variedad de invertebrados. Sin embargo, provoca una genuina infección viral y, por tanto, acaba destruyendo la célula infectada. Otra limitación es la necesidad de introducir el gen o los genes deseados a través

de un proceso de cointegración similar al descrito para los baculovirus. Primero se introducen los genes en un vector procariótico denominado vec-

tor de transferencia o lanzadera y éste se introduce en el genoma del virus por recombinación homóloga *in vivo* (figura 7.11).

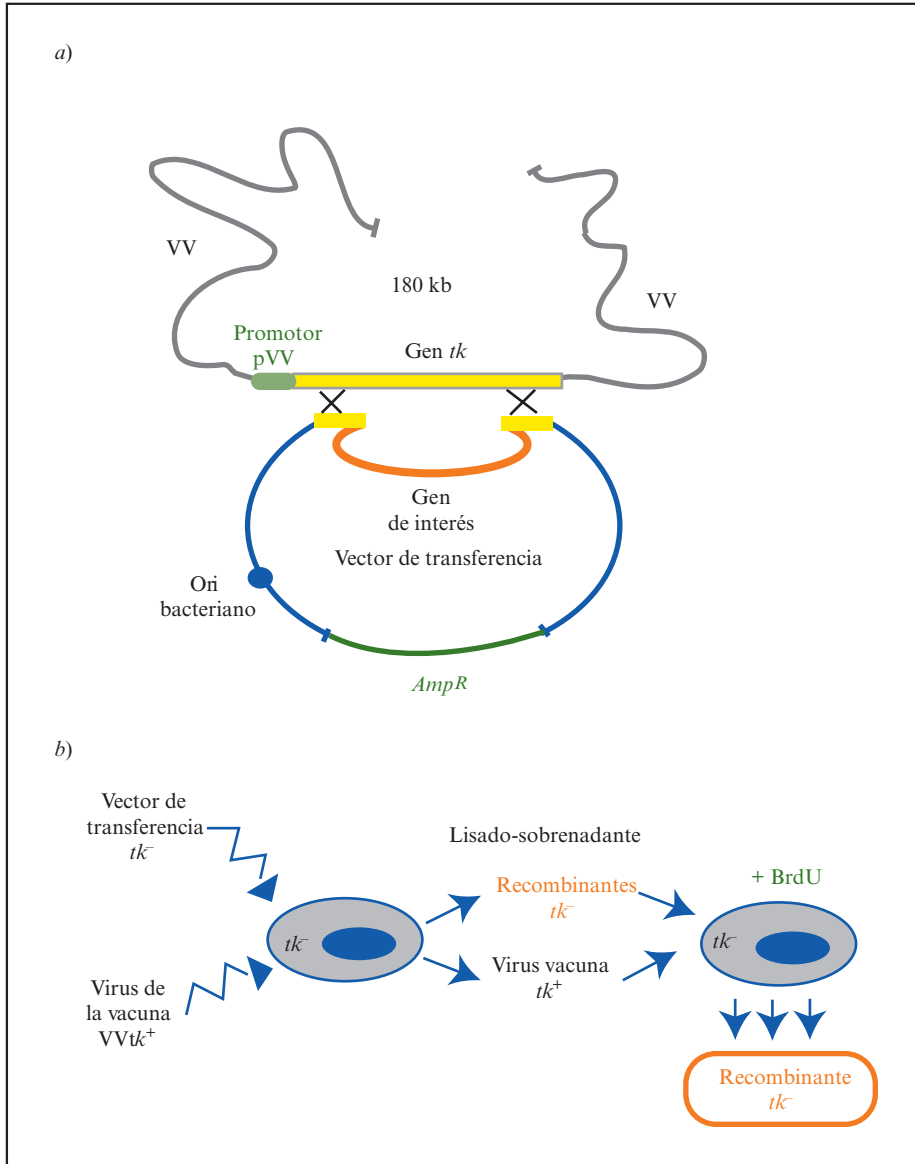


Figura 7.11. Virus de la vacuna. *a)* Clonaje por cointegración en el virus de la vacuna (VV). Un gen eucariótico de interés se introduce en un vector procariótico flanqueado por la región inicial y final del gen de la quinasa de timidina (*tk*). *b)* Este vector se introduce en células *tk*⁻ a la par que el virus de la vacuna tipo silvestre. El lisado-sobrenadante se pone en contacto con células frescas en presencia de BrdU; en estas condiciones sólo crecerán los recombinantes.

Para la recombinación homóloga se puede utilizar, entre otros, el gen de la quinasa de timidina (*tk*) del virus, ya que no es esencial para la infectividad. El procedimiento consiste en flanquear el gen de interés con dos regiones homólogas al inicio y al final del gen de la quinasa de timidinas del virus de la vacuna en el vector de transferencia que llevará también los componentes bacterianos necesarios para su crecimiento en bacterias. Este vector de transferencia se introduce en células *tk*⁻, como, por ejemplo, las 143 humanas, a la par que el virus de la vacuna (VV) tipo silvestre (*tk*⁺). Al tener lugar la recombinación homóloga entre ambas moléculas, se producirán virus deficientes para *tk* que contienen el gen de interés (figura 7.11). Las placas del virus se recogen y el lisado se utiliza para infectar células frescas *tk*⁻, pero

esta vez en presencia de BrdU (5-bromodesoxiuridina). En estas condiciones, sólo los virus *tk*⁻ podrán replicar su ADN seleccionándose frente a los no recombinantes. El mecanismo molecular es el siguiente: el compuesto BrdU, normalmente, se convierte en un análogo de TMP (timidina monofosfato) por acción de la quinasa de timidina y se incorpora en el ADN. Una vez incorporado en el ADN, produce un aumento en la sensibilidad a los rayos ultravioleta (UV) y, por tanto, fallos drásticos y letales en la replicación. Así pues, los virus y células que sean *tk*⁻ no podrán incorporar el análogo suicida y sobrevivirán, mientras que los *tk*⁺ morirán. El único inconveniente de esta selección es la aparición con una frecuencia bastante elevada (10^{-3} - 10^{-9}) de mutantes *tk*⁻ espontáneos. Admite insertos de unas 25 kb de ADN.

TABLA 7.1

Resumen de las ventajas e inconvenientes de algunos vectores virales descritos

Vectores	Ventajas	Inconvenientes
SV40 y derivados	— Pequeños y de fácil manejo. — Infectan células en división y quiescentes.	— Expresión transitoria. — Poca capacidad para inserto. — Títulos bajos (10^5 u.f.c/ml)*.
Retrovirus	— Transferencia muy eficiente. — Pequeños y de fácil manejo. — Se integran en el genoma.	— Sólo infectan células en división. — Poca capacidad para inserto. — Mutagénesis insercional/leucemia. — Títulos bajos (10^6 u.f.c/ml)*
Adenovirus	— Infectan células en división y quiescentes. — Títulos altos (10^{11} u.f.p./ml)**.	— Inmunogénicos. — Manipulación compleja. — Expresión transitoria.
Adenoasociados	— Pequeños y de fácil manejo. — Poca inmunogenicidad.	— Capacidad de inserto limitada.
Herpes	— Tropismo hacia el sistema nervioso. — Admiten insertos grandes.	— Tamaño muy grande. — Tóxicos.
Virus de la vacuna	— No producen patología. — Admiten insertos grandes.	— Son inmunogénicos. — Tamaño muy grande.

* Unidades formadoras de colonias por mililitro.

** Unidades formadoras de placas por mililitro.

7.3.8. Cromosomas artificiales de mamíferos

Los cromosomas artificiales de mamíferos (CAM, MAC en versión anglosajona, o HAC aplicado a humanos) son unidades lineales con todos los componentes cromosómicos indispensables: un centrómero, dos telómeros y, al menos, un origen de replicación. Las ventajas de los cromosomas artificiales como vectores son triples: 1) permiten la transferencia de insertos de tamaño prácticamente ilimitado; 2) se mantienen de forma estable en el núcleo, y 3) no provocan rechazo inmunológico.

Existen dos maneras de construir un CAM:

1. **De abajo arriba** o síntesis de novo a partir de los distintos componentes manteniéndolos por separado en cromosomas artificiales de levaduras.
2. **De arriba abajo**, se va truncando un cromosoma completo hasta sustituirlo por uno artificial.

Cromosomas artificiales contruidos de abajo arriba

La *región centromérica* de los cromosomas de mamíferos es el lugar por donde permanecen unidas las cromátidas hermanas, apareciendo muy condensada en cromosomas metafásicos. Desde el punto de vista genético, el centrómero es responsable de la segregación del cromosoma en meiosis y mitosis, asegurando que el número de copias no exceda de una o dos por célula. El centrómero de *S. cerevisiae* sólo tiene 125 pares de bases; el de *S. pombe* tiene ya 40 kb; el de ratón, 150 kb. El centrómero del cromosoma Y humano ha sido el más estudiado. Consta de un largo trecho de ADN satélite (altamente repetido, con una unidad de 171 pb) que puede ocupar de 240 a 1.600 kb dependiendo del individuo; se denomina ADN alfoide. En los flancos del ADN alfoide hay otras familias de ADN repetido que ocupan hasta 300 kb. Funcionalmente relacionada con el ADN alfoide está la proteína centromérica CENP-B que ha sido reconocida como antigénica para el suero de

pacientes con la enfermedad inmune esclerodermia CREST (calcinosis, fenómeno de Raynaud, dismovilidad esofágica, esclerodactilia, telangiectasia). Además de la región centromérica típica descrita, pueden aparecer neo-centrómeros, es decir, regiones cromosómicas que carecen de ADN alfoide o unión a CENP-B, pero que funcionan como verdaderos centrómeros en cuanto a segregación y estabilidad. Se ha descrito un neo-centrómero en medio del brazo largo del cromosoma 10 humano y ocho más no muy bien caracterizados. Estos neo-centrómeros cuestionan la necesidad de la presencia de ADN alfoide para la función de segregación y estabilidad.

Los telómeros (del griego, *telos* = terminal y *meros* = parte), están formados por una hilera de ADN repetido de secuencias cortas (TTAGGG)_n seguido por otra de secuencias moderadamente repetidas (región subtelomérica). El tamaño del telómero es muy variable siendo sólo de 200 pares de bases en levaduras y 150 kb en ratón. Su papel biológico es proteger los extremos evitando que se fusionen los cromosomas entre sí. La replicación del telómero requiere una transcriptasa inversa llamada telomerasa que es una ribonucleoproteína presente desde los protozoos hasta los humanos. La telomerasa utiliza como molde su propio ARN (8-30 bases) y como enganche el extremo 3'OH de la molécula de ADN telomérico. En tejidos diferenciados de mamíferos la actividad telomerasa deja de detectarse, mientras que está presente en tumores.

Los orígenes de replicación de los cromosomas de metazoos no son fijos ni precisos, es muy posible que haya uno por cada 30-40 kb, por lo que formarían parte de los insertos.

Los cromosomas humanos tienen tamaños de 50-250 mgb. Artificialmente se han obtenido cromosomas lineales de 6-10 mgb transfectando mediante liposomas los distintos componentes cromosómicos: ADN alfoide, ADN telomérico e inserto, a células HT1080 humanas. Estos cromosomas de primera generación parecen originarse dentro de la célula por un proceso de recombinación no bien definido, pero son mitóticamente y citogenéticamente estables durante seis meses en cultivo sin necesidad de selección. Algunos de es-

tos cromosomas contienen ADN alfoide exógeno (de otros cromosomas), además del transfectado o secuencias no centroméricas (en uno de los casos aparecieron secuencias del cromosoma 8). Esta modalidad de construir cromosomas artificiales de mamíferos es de muy baja eficiencia, además de originar un producto de tamaño y contenido impredecibles.

Cromosomas artificiales contruidos de arriba abajo

Utilizando este método, se parte de un cromosoma completo y se van eliminando regiones de los dos brazos para sustituirlas por genes de selección y genes de interés. Se han obtenido cromosomas artificiales a partir de varios cromosomas humanos como el 14 y el 21. Los mejores resultados se obtienen cuando todos los genes en el cromosoma original son eliminados, ya que así no hay posibilidad de recombinación homóloga cuando se reincorpora el cromosoma artificial a la célula de procedencia. La metodología es laboriosa y vamos a seguir los pasos empleados para la obtención del cromosoma artificial humano 21 (HAC21). En primer lugar hubo que marcar y aislar el cromosoma humano 21. El marcaje se realizó por recombinación homóloga utilizando como diana un gen del cromosoma 21 en el que se introdujo el gen de selección *neo*. Para su aislamiento se trataron las células donantes con colcemida, un análogo sintético de la colchicina para bloquear las células en mitosis, y con citocalasina B, que induce la división del núcleo provocando minicélulas con sólo un cromosoma y un poquito de citoplasma. El paso siguiente fue fusionar estas minicélulas con células DT40 de pollo y seleccionar por G418. Las células DT40 de pollo exhiben una frecuencia de recombinación alta que nos facilitará los sucesivos pasos siguientes consistentes en sustituir fracciones del cromosoma 21 por un *loxP* para introducir el inserto y genes de selección (figura 7.12). Lo primero que se introdujo fue *loxP*, para lo cual se forzó una recombinación homóloga entre dos regiones que denominaremos *a* y *b* del cromosoma 21 y las mismas regiones *a* y *b* introducidas en un vector flanqueando *loxP/hprt*, donde *hprt* es el gen de

selección positiva (véase apartado 7.2.1). Para eliminar el brazo pequeño del cromosoma se facilitó una recombinación homóloga entre la región peritelomérica del cromosoma 21 por un lado y la región peri-centromérica del otro, con un CAL que llevaba como inserto: una pequeña región del peritelómero, un promotor potente, el gen de selección positiva *puro*, una región homóloga al peri-centrómero y el gen de selección negativa *tkHs*. La recombinación homóloga nos sustituirá todos los genes del brazo pequeño por el gen *puro*. Debemos crecer las células DT40 en presencia de puromicina y ganciclovir para seleccionar por los recombinantes. De manera análoga se sustituyeron los genes del brazo largo del cromosoma 21 por el gen de resistencia a la higromicina.

Una vez finalizada la sustitución de los genes del cromosoma 21 por genes de selección, aún hay que introducir el gen de interés mediante *loxP*, seleccionar por este evento y devolver el nuevo cromosoma 21 a las células humanas. El nuevo HAC21 llevará numerosos marcadores de selección que facilitarán el mantenimiento del cromosoma en las células receptoras humanas.

7.4. TERAPIA GÉNICA

La terapia génica es la utilización de ácidos nucleicos como agentes terapéuticos. Existen dos modalidades de terapias génicas: *ex vivo* e *in vivo*. En la terapia *ex vivo* células del paciente son manipuladas *in vitro* y devueltas al organismo de partida, con lo que evitamos un rechazo inmunológico. Con la terapia *in vivo* se introduce directamente en el paciente el gen terapéutico desnudo o en el interior de un vector viral. Una variante de esta segunda modalidad es la introducción en el paciente de células productoras de virus terapéuticos.

En enfermedades hematopoyéticas es frecuente la terapia *ex vivo* utilizando células troncales linfoides de la médula ósea, cordón umbilical o sangre periférica. Estas células son totipotentes y pueden ser aisladas del individuo, manipuladas *in vitro* y devueltas al órgano hematopoyético. Un inconveniente de estas células troncales es que representan sólo el 0,01 % de las células de la médu-

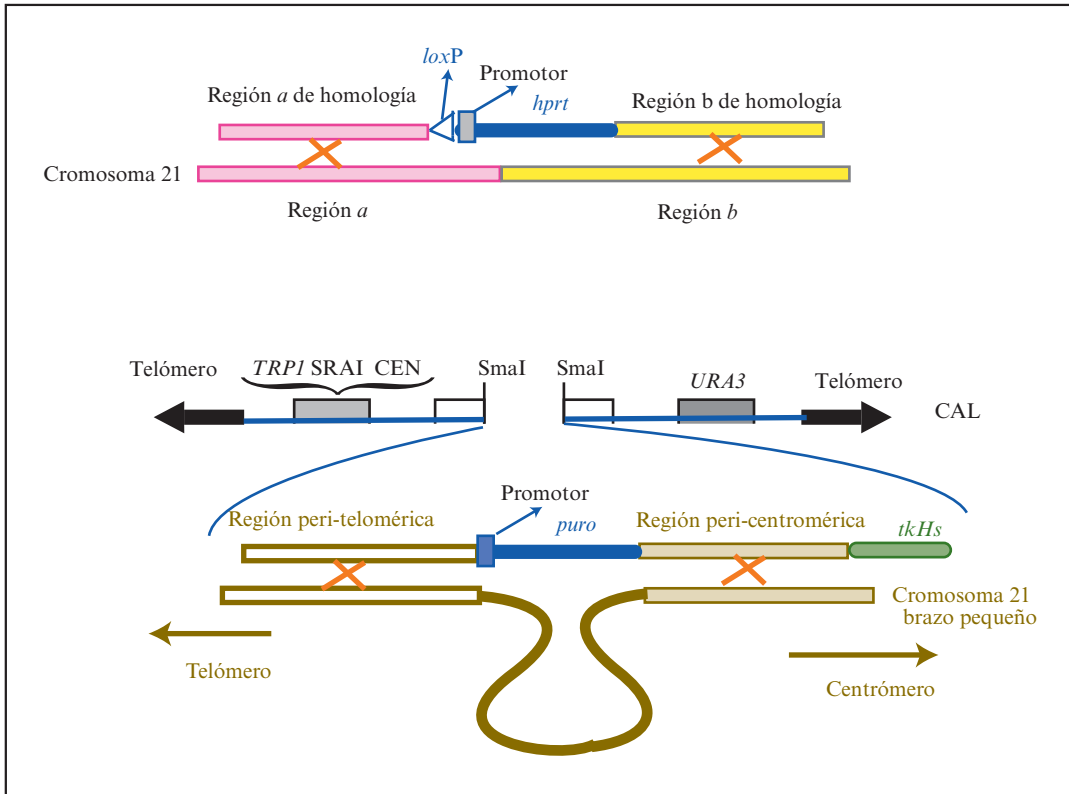


Figura 7.12. Obtención de cromosomas artificiales de mamíferos por el método de arriba abajo. En la parte superior, esquema que muestra la recombinação homóloga entre las regiones *a* y *b* del cromosoma 21 y las mismas regiones flanqueando *loxP*/*hpert* introducidas en un vector. En la parte inferior, recombinação homóloga entre un CAL portador de la región peri-telomérica y peri-centromérica del cromosoma 21 que flanquean el gen de selección *puro* y un promotor, con su región homóloga en el brazo pequeño del cromosoma.

la y no hay realmente manera eficiente de aislarlas del resto. Expresan el antígeno de superficie CD34⁺ y su crecimiento es dependiente de factores de crecimiento o interleuquinas.

Para el tratamiento de las distintas enfermedades se elige el tipo de vector más adecuado dependiendo de las características del mismo y de la enfermedad. En la fibrosis quística, por ejemplo, emplearemos adenovirus por tener éstos un tropismo hacia las vías respiratorias, mientras que para la terapia del cáncer será muy apropiado utilizar vectores retrovirales, que sólo infectan células en división; pero unos y otros tienen sus correspondientes inconvenientes, como son el rechazo inmu-

nológico que provocan los primeros y la mutagénesis insercional que pueden producir los segundos. Los virus en general son buenos vectores, ya que tienen una innata habilidad de introducir material genético en la célula infectada, siendo su mayor preocupación encontrar la manera de dividirse y producir copiosa descendencia. La mayor parte de los virus no ganan nada matando la célula hospedadora, pero desafortunadamente muchas veces es inevitable debido a la toxicidad o patogénesis de algunos productos virales. Para transformar un virus en un vector es indispensable convertirlo en defectivo condicional como medida de seguridad y crear «sitio» para albergar el gen terapéutico.

No hay un vector perfecto y en todos podemos encontrar ventajas e inconvenientes. El vector menos malo es el que mejor sorteaba una serie de obstáculos comunes a todos ellos, siendo el primer obstáculo el de la producción a gran escala. Un buen vector debería poderse obtener en forma concentrada y pura con una metodología asequible y reproducible. Esto sigue siendo problemático en algunos vectores (retrovirus), pero está prácticamente superado en algunos otros (adenovirus). El segundo obstáculo a salvar es el hacer llegar y facilitar la entrada en el tejido más apropiado para la curación de una enfermedad concreta, bien sea una célula quiescente o en división. En muchos casos es necesario conseguir una expresión estable y sostenida durante largos períodos de tiempo (enfermedades crónicas), lo que requerirá la integración del gen terapéutico en el genoma hospedador, mientras que en otros se premiará una alta expresión por un tiempo corto (eliminación de la célula neoplásica), lo que se puede conseguir con una expresión transitoria elevada. La integración en el hospedador puede producir mutagénesis insercional si la inclusión se realiza preferentemente en genes activos (retrovirus), por lo que vectores con sitios concretos y definidos de integración son preferidos aunque no hay muchos.

En muchos casos una regulación exquisita de la expresión del gen terapéutico es necesaria, por lo que el fragmento a insertar deberá incluir regiones de carácter regulador que sean capaces de responder a elementos o moléculas activadoras y represoras. Las regiones reguladoras, a su vez, pueden superar en muchas kilobases el tamaño del transgén, añadiendo un nuevo obstáculo al diseño y elección del vector. Finalmente, el vector no debe ser patógeno en sí mismo, ni desencadenar una respuesta inmune, que suele ser adversa en la mayoría de los casos.

El desarrollo de una terapia génica, desde su diseño hasta su posible opción terapéutica en la práctica médica, conlleva una serie de pasos:

1. Probar la terapia en células en cultivo.
2. Utilizarla en animales de experimentación y verificar, por un lado, su eficiencia y,

por otro, la ausencia de efectos secundarios o toxicidad.

3. Probar en un pequeño número de pacientes voluntarios.

7.4.1. Terapia génica de enfermedades monogénicas

Hablamos de una enfermedad monogénica cuando la patología de la enfermedad es atribuible a un solo gen. Sin embargo, el término merece aclaración, puesto que hay variantes dentro de muchas enfermedades en las que distintos genes causan las distintas variantes de la enfermedad, pero en todos los casos un gen específico es responsable de la variante concreta de la enfermedad. Veremos como ejemplos de terapias monogénicas exitosas, las desarrolladas contra varias enfermedades sanguíneas como las inmunodeficiencias combinadas severas ADA y SCID-X1, la leucodistrofia metacromática o la inmunodeficiencias primaria de WAS, y en terapias oculares nos centraremos en la terapia génica contra la amaurosis congénita de Leber.

La inmunodeficiencia combinada severa

Hay dos variantes de esta enfermedad abordables por terapia génica. El gen implicado en la **variante ADA** es el gen de la deaminasa de adenosinas. La alteración en el gen elimina la actividad de la enzima y como consecuencia se acumula adenosina y 2' desoxiadenosín trifosfato en la sangre de los pacientes. La consecuencia inmediata es la destrucción masiva de los linfocitos T para los que estos compuestos son particularmente tóxicos. Los linfocitos T son, por otra parte, imprescindibles para desarrollar inmunidad hacia cualquier patógeno. La segunda variante de la enfermedad es la inmunodeficiencia combinada severa-X1 (SCID, del inglés, *severe combined immunodeficiency*). En **la SCID-X1** el gen afectado es el receptor γ C de las interleuquinas 2, 4, 7, 9, 15 y 21, y el resultado de su alteración, un bloqueo en la diferenciación de linfocitos T y células asesinas natas naturales (NK, del inglés *natural killers*).

Ambas variantes de la enfermedad son letales desde el nacimiento a no ser que los niños afectados se mantengan en cubículos aislados y estériles (niños burbuja). Pero la inmunodeficiencia combinada severa es una de las pocas enfermedades en las que la terapia génica es ya una opción terapéutica.

Atrás han quedado los experimentos en los que se eliminaban por KO el gen de la deaminasa de adenosina implicado en la variante ADA, o el del receptor γ C de las diversas interleuquinas, para simular en células en cultivo la enfermedad. También en cultivo, se administró un vector retroviral portador del gen dañado que recuperó la funcionalidad del gen en cuestión. También se llevó a cabo con éxito la curación de ratones que mimetizaban la enfermedad humana y, por tanto, hace ya algunos años se procedió a numerosos ensayos clínicos en los que se verificó su potencialidad como opción terapéutica.

La variedad ADA

El tratamiento antes de la terapia génica para la variedad ADA consistía en la inyección repetida de células de la médula ósea procedentes de un donante compatible en HLA (sistema de incompatibilidad) o inyecciones musculares múltiples y

frecuentes (semanales) de la enzima unida a polietilenglicol para estabilizarla. En estos últimos años se han tratado varios niños por terapia génica *ex vivo* introduciendo un gen funcional *ada*, insertado en un vector retroviral, en células troncales linfoides extraídas de la médula ósea y sangre del paciente y devueltas al organismo una vez modificadas genéticamente (figura 7.13). En los primeros experimentos se observó una baja eficiencia de la transferencia génica (aproximadamente sólo un 1 % de la población celular había sido modificada), con una severa limitación del beneficio terapéutico. La modificación posterior de la metodología vino acompañada del éxito terapéutico. Se incluyó un tratamiento quimioterápico de eliminación de la mayoría de las células de la médula ósea antes de la restitución de las células del paciente tratadas *ex vivo* con el gen terapéutico y estimuladas a proliferar en presencia de interleuquinas antes de ser repuestas al paciente. Así se eliminaba una población de precursores no tratados y se dejaba sitio para los nuevos precursores modificados genéticamente. De unos 40 pacientes tratados en estos últimos años, el 70 % mostró beneficio clínico y no necesitó terapias de soporte adicionales como administración oral de la enzima.

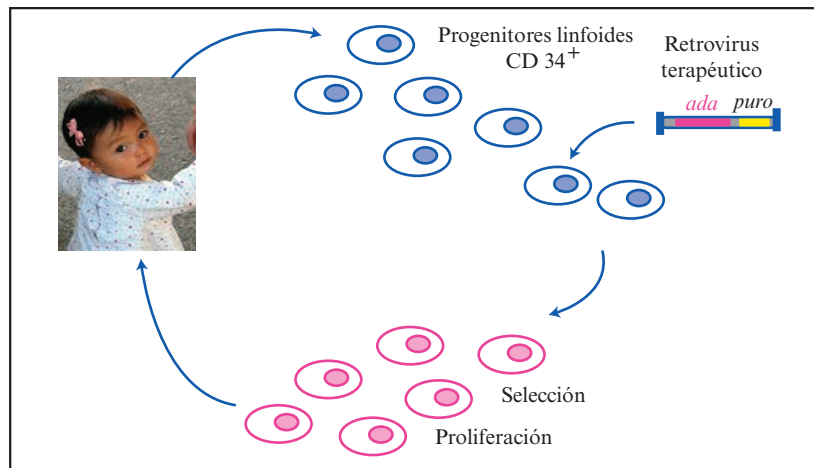


Figura 7.13. Terapia *ex vivo* de la inmunodeficiencia combinada severa SCID-ADA. Obtención de progenitores linfoides del paciente: CD34⁺, infección con retrovirus terapéuticos portadores de un gen funcional *ada* y un gen de selección *puro*. Selección y crecimiento *in vitro* y devolución a la médula ósea del paciente.

La variedad SCID-X1

En la terapia génica del gen del receptor γ C de las interleuquinas 2, 4, 7, 9, 15 y 21 se utilizó asimismo un vector retroviral y un procedimiento similar al descrito para SCID-ADA con una variante que consistió en omitir el tratamiento quimioterápico, ya que los progenitores linfoides CD34⁺ que recibieron el retrovirus desarrollaron una ventaja proliferativa sobre los que no lo recibieron. El receptor γ C al unirse a sus ligandos (interleuquinas 2, 4, 7, 9, 15 y 21) provoca una señal de proliferación a las células progenitoras linfoides T y NK. Así, al cabo de unos meses, la población de linfocitos modificados genéticamente era muy superior a la de los no modificados y los niños tratados, en su mayoría, se curaron. Sin embargo, hubo un 20% de los niños tratados que al cabo de un tiempo (2-5 años) desarrollaron leucemias que respondieron bien, en todos los casos menos uno, a la quimioterapia. La mayoría de estas leucemias se han producido por la integración del retrovirus en los alelos del gen *LMO2* correspondiente a un factor de transcripción regulador de la hematopoyesis. Parece existir una interacción indeterminada entre *LMO2* y γ C. Como medida de seguridad adicional, se están utilizando los vectores retrovirales autoinactivantes al estilo de los lentivirales ya descritos (ver figura 7.7). En estos vectores se imposibilita la transcripción a partir del promotor del TRL por una delección en la región U3, con lo que el transgén ha de llevar un promotor independiente; esto, a su vez, abre posibilidades de controlar la expresión del gen terapéutico con promotores estimulables o específicos de tejido.

En Europa, en la actualidad, aún se prefiere el trasplante de progenitores linfoides de donantes compatibles cuando ello es posible. El trasplante garantiza una supervivencia del 90% a los 5 años.

Si se analiza críticamente la relación riesgo/beneficio del tratamiento génico de SCID-X1, tendríamos que admitir que, dada la pésima calidad de vida de los pacientes sin posibilidades de un trasplante de médula de un donante compatible y su mermada esperanza de vida, la posible adquisición de una leucemia es mejor que la ausencia de tratamiento génico. Las leucemias infantiles tienen

un porcentaje de supervivencia relativamente alto y la calidad de vida de estos pacientes es mejor que la de los inmunodeficientes severos SCID-X1.

La terapia génica es una verdadera opción terapéutica y se está haciendo un hueco en el tratamiento de las inmunodeficiencias combinadas severas, especialmente cuando no se encuentran donantes compatibles.

Dos enfermedades que afectan a los leucocitos con potencial terapéutico

La leucodistrofia metacromática y el síndrome de Wiskott-Aldrich son dos enfermedades genéticas raras que afectan también a niños desde el nacimiento por la falta de dos proteínas esenciales en los primeros años de vida. La terapia convencional para estas dos enfermedades es la de trasplante de médula ósea cuando hay donantes compatibles, pero la terapia génica puede en breve ser una opción terapéutica.

La leucodistrofia metacromática (LDM)

La leucodistrofia metacromática es una enfermedad autosómica recesiva que afecta a los lisosomas y el gen implicado es el correspondiente a la enzima arilsulfatasa A (*arsa*). La ausencia de la enzima provoca una acumulación tóxica de sulfátidos que origina daños en la mielina que rodea las neuronas y en todo el sistema nervioso, además de afectar a los riñones y a la vesícula biliar. Los niños con esta enfermedad parecen normales al nacer, pero van perdiendo gradualmente las facultades cognitivas, neurológicas y motoras.

El síndrome de Wiskott-Aldrich

Los niños afectados por la inmunodeficiencia primaria de WAS, o síndrome de Wiskott-Aldrich, carecen de un sistema inmunológico eficiente. El gen afectado (*was*) se localiza en el cromosoma X y se traduce a una proteína (WasP) que regula el citoesqueleto. Los pacientes sufren infecciones frecuentes, microtrombocitopenia (bajo número de plaquetas en sangre), eczemas, disminución de linfocitos T y predisposición a cáncer.

Un procedimiento *ex vivo* muy similar se ha empleado para el tratamiento génico de estas dos enfermedades. Tras un tratamiento suave de quimioterapia para destruir precursores hematopoyéticos como hemos descrito para ADA, se extrajeron células troncales hematopoyéticas (CD34⁺) de la médula ósea de 16 pacientes (6 de WAS y 10 de LDM) y se infectaron con lentivirus sin promotores en los TRL (SIN), portadores de copias sanas

de los genes *arsa* o *was* según el caso; una vez crecidas y seleccionadas, se devolvieron las células manipuladas a los pacientes (figura 7.14). Al menos 6 niños de los tratados (3 de cada enfermedad) han mostrado un claro beneficio terapéutico a los 2 años de realizada la terapia. El análisis de los lugares en los que se insertó el lentivirus, por secuenciación masiva, no reveló lugares preferenciales ni se desarrollaron leucemias en ningún caso.

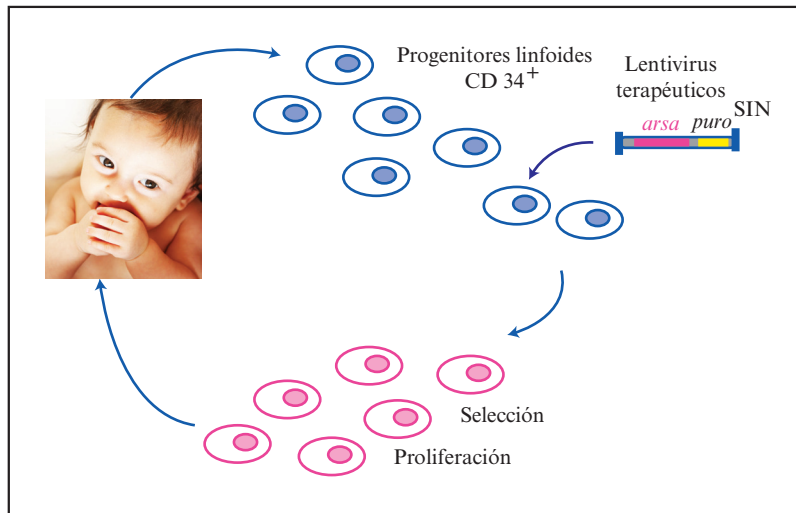


Figura 7.14. Esquema de terapia *ex vivo* utilizada en ensayos clínicos de la leucodistrofia metacromática y el síndrome de Wiskott-Aldrich. Después de una ameloablación suave, se obtuvieron progenitores linfoides CD34⁺ de los niños afectados, se infectaron con lentivirus (SIN) terapéuticos portadores de un gen funcional *arsa* o *was* según el caso, además de un gen de selección (véase figura 7.7) se seleccionaron y crecieron *in vitro* las células troncales manipuladas y se devolvieron a la médula ósea de los niños.

La amaurosis congénita de Leber

El ojo es un órgano idóneo para terapias génicas por varios motivos: 1) la retina es accesible a inyecciones locales, 2) posee un estatus inmunológico bastante privilegiado, 3) las barreras internas del ojo dificultan la invasión sistémica de los vectores virales, 4) se puede medir fácilmente el grado de ganancia visual obtenida con la terapia y 5) se puede tratar sólo uno de los ojos, sirviendo el otro de control.

Como animales de experimentación se suelen usar, bien ratones transgénicos o perros, en los

que esta enfermedad ocular ocurre con una frecuencia bastante elevada. Como vectores, los virus adenoasociados son los que más se han utilizado en esta enfermedad y el tipo de terapia es *in vivo*. La amaurosis congénita de Leber (ACL, o LCA en versión anglosajona) es una enfermedad autosómica recesiva que afecta a los conos y bastones de la retina del ojo. Se han identificado unos 14 genes cuyas deficiencias provocan distintos tipos de amaurosis congénita. Uno de estos genes es el *rpe65* (del inglés, *retinal pigment epithelium 65*) que se localiza en el cromosoma 17 humano y es responsable de la síntesis de la proteína asocia-

da a membrana Rpe65. Esta proteína convierte los ésteres *trans*-retinil en *cis*-retinol en el epitelio pigmentado de la retina. Alteraciones en el gen provocan acumulación de los ésteres *trans*-retinil lo que produce daños en los fotorreceptores y pérdida de visión o ceguera. Alteraciones en el gen *rpe65* representan el 6% de todos los casos de ACL. Experimentos preclínicos en ratones y perros *rpe65*^{-/-} a los que se trató con vectores adenoasociados (VAA) portadores de una copia normal del gen *rpe65* restauraron la visión de modo efectivo y estable. En estos experimentos se ha puesto de manifiesto que el tratamiento es más eficiente si se realiza durante los primeros estadios de la enfermedad. En pacientes humanos se han

llevado a cabo varios ensayos clínicos utilizando vectores adenoasociados portadores de *rpe65* y realizando la terapia mediante una sola inyección subretinal (figura 7.15). La mayoría de los pacientes mejoraron sensiblemente su visión, especialmente en penumbra. No se observaron efectos tóxicos o adversos al menos en los dos años siguientes al tratamiento. Se detectaron vectores virales en lágrimas y sangre después de la cirugía, que remitieron al cabo del tiempo. No se observó respuesta inmune adversa. Las mejorías observadas en los pacientes también correlacionaron con la edad en el sentido de que los pacientes jóvenes respondieron mejor al tratamiento que los mayores.

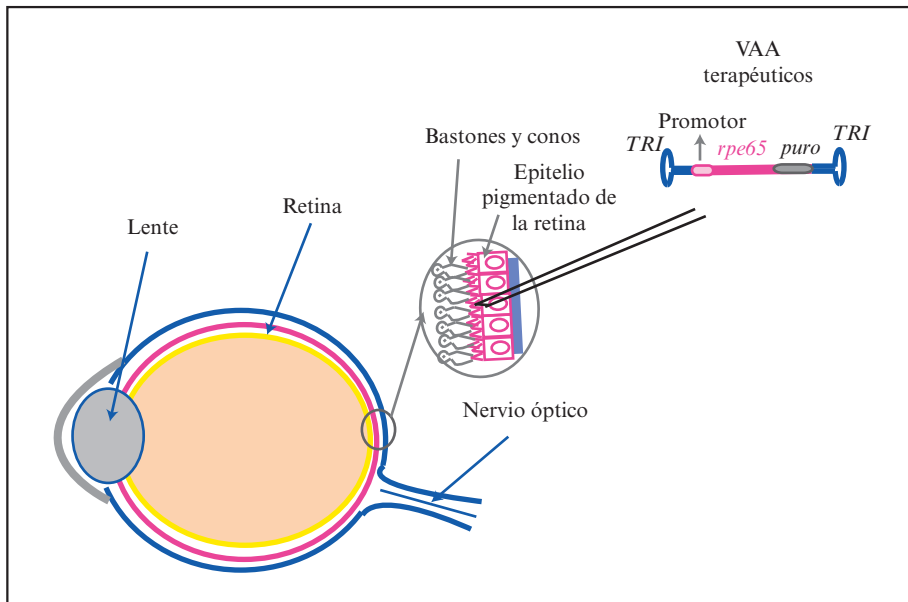


Figura 7.15. Terapia *in vivo* de la amaurosis congénita de Leber. Inyección subretinal de vectores VAA portadores del gen *rpe65*.

7.4.2. Terapia génica de enfermedades multigénicas: cáncer

El cáncer es una enfermedad multigénica y de modo intuitivo pensaríamos que la eliminación de la célula tumoral puede ser la estrategia más

segura, simple y eficiente contra la enfermedad. La naturaleza autopropagativa y expansiva de la célula tumoral nos alerta sobre la conveniencia de que las células neoplásicas sean eliminadas tan eficientemente como sea posible frente a alternativas de carácter corrector.

La conversión de una célula normal en una célula tumoral es un proceso gradual en el que se acumulan sucesivas alteraciones genéticas. Por otra parte, hay mucha evidencia de que los sucesos moleculares y celulares que desencadenan un tumor están orquestados por células cancerosas de tipo troncal. La denominación célula troncal tumoral (CSC, del inglés *cancer stem cell*) define una célula neoplásica capaz no sólo de autorreplificarse indefinidamente, sino también de generar la diversidad del tumor. La mayor parte de los cánceres están constituidos por una población heterogénea de células, con diferencias tanto proliferativas como de capacidad de reproducir el tumor cuando se trasplantan a modelos animales inmunodeficientes (por ejemplo, ratones desnudos). Se ha demostrado la existencia de células troncales cancerosas en prácticamente todos los tipos de tumores en los que se han buscado: leucemias mieloides, cánceres de mama, melanomas, tumores de páncreas, tumores cerebrales, etc. Las CSC forman esferas no adherentes en cultivos y dan lugar a distintos tipos de células tumorales con marcadores de superficie característicos. La inoculación de células troncales tumorales en animales de experimentación origina tumores con una mezcla de células troncales tumorales y células no troncales con reminiscencias del tejido de procedencia. No está claro si las CSC derivan de células troncales autóctonas tisulares, de células troncales de la médula ósea o de células diferenciadas que han sufrido un proceso de dediferenciación o una transdiferenciación.

La mayoría de las estrategias terapéuticas que se aplican contra las células tumorales no hacen distinción alguna entre las CSC y el resto de las células tumorales. Tradicionalmente los tratamientos contra el cáncer se basaban en la eliminación de las células en división, pero curiosamente las CSC se dividen con baja frecuencia, con lo que estas terapias tenderían al fracaso. Además, las CSC tienen una mayor capacidad de reparar su ADN, algo que, unido a su poca actividad proliferativa, las hace menos sensibles a terapias de tipo radioterapéutico o quimioterapéutico. Teniendo en cuenta estos conocimientos, podríamos predecir que, para conseguir la erradicación total

de un tumor e impedir las recidivas, necesitaríamos un sistema que eliminara no sólo las células en división sino también las quiescentes del entorno tumoral, en especial las células troncales.

Hace algunos años, se crearon grandes expectativas sobre la terapia génica contra el cáncer usando vectores retrovirales portadores de genes potencialmente asesino-suicidas o tratamientos génicos que estimulaban el propio sistema inmunológico para rechazar tumores. También se utilizaron virus para aportar al tumor oncosupresores, inductores de apoptosis, factores antiangiogénicos, citoquinas o diversos ARN de interferencia. Haremos un resumen de algunas de estas estrategias, aunque en la actualidad no hay ningún tratamiento génico contra el cáncer que constituya una verdadera opción terapéutica.

Sistemas asesino-suicidas condicionales

Las terapias asesino-suicidas consisten en la utilización de un gen cuyo producto convierte un compuesto inocuo en otro tóxico y mortal para la célula. Los vectores suicidas que se emplean para erradicar tumores son preferentemente retrovirus defectivos, en los que los genes propios del virus (*gag*, *pol* y *env*) han sido sustituidos por un gen de selección (*neo* o *puro*) y por otro gen capaz de inducir su propia muerte. Un ejemplo de gen suicida es el gen de la quinasa de timidina (*tk*) del *Herpes simplex* (*Hs*), que al expresarse en la célula tumoral confiere sensibilidad al antibiótico antiherpético ganciclovir. Como hemos comentado con anterioridad, la quinasa del *Herpes* no es tan estricta como su homóloga celular y fosforila análogos de la guanina (ganciclovir o aciclovir) que pueden así ser incorporados en la molécula del ADN paralizando la replicación y provocando la muerte celular. Solamente aquellas células que hayan adquirido la quinasa del *Herpes* incorporarán el ganciclovir fosforilado en su ADN. Los retrovirus infectan exclusivamente células en división, característica diferencial de las células tumorales que se desarrollan en tejidos especializados, los cuales normalmente ya no se dividen. Los retrovirus portadores del gen *tk* del *Hs* se introducen en células murinas empaquetadoras que poseen en su geno-

ma un retrovirus defectivo. Este retrovirus defectivo carece de la región (ψ) imprescindible para el empaquetamiento de la partícula; estas células murinas se convierten entonces en fábricas de retrovirus terapéuticos y se denominan células productoras (apartado 7.3.3, figura 7.6). La introducción en el seno de un tumor de células productoras origina un cierto número de retrovirus que infectarán las células en división del tumor. Al cabo de unos días, en los que cada vez más células tumorales adquieren el retrovirus terapéutico, se inicia el tratamiento generalizado con el antibiótico ganciclovir, que provocará el suicidio de las células tumorales. El sistema ha sido probado con éxito en tumores cerebrales de ratas y con menos éxito en pacientes de glioblastoma. Ninguno de los pacientes se curó, aunque se observó una reducción significativa del tumor en el 30% de los casos.

Otro ejemplo de estos sistemas asesino-suicidas es el de la linamarasa/linamarina (*lis/Lin*) que se empleó para erradicar tumores cerebrales malignos (glioblastomas) de forma experimental. El gen asesino-suicida era el gen de origen vegetal *linamarasa* (*lis*) que convierte el sustrato inocuo linamarina (*Lin*) en glucosa y cianuro. La linamarina es una β -glucosidasa particularmente abundante en las semillas del lino (de ahí su nombre) y en la «casava» *Manihot esculenta* (también llamada yuca o mandioca). La transducción del gen de la *linamarasa* a células de mamíferos las sensibiliza a la presencia de linamarina y la liberación del cianuro producido puede matar las células del entorno provocando así un efecto colateral importante. Sin embargo, cuando la terapia se probó en tumores cerebrales desarrollados en el cerebro de ratas inmunodeficientes, el cianuro se difundió rápidamente desde el lugar de producción, observándose una sorprendente capacidad de la célula tumoral para sobrevivir en condiciones de intoxicación temporal.

La utilización de vectores retrovirales portadores de genes potencialmente letales para la célula tumoral tiene la ventaja de que los retrovirus sólo infectan células en división, por lo que no hay peligro de que células quiescentes normales, próximas al tumor, sean infectadas y eliminadas. Sin embargo, esta misma propiedad genera algún

problema, ya que no todas las células del tumor estarán en división en el momento del tratamiento, especialmente las células troncales tumorales, por lo que habrá un porcentaje variable de células que no serán infectadas por el retrovirus y sobrevivirán, pudiendo en un futuro regenerar el tumor.

Inmunoterapia génica

Existen numerosos sistemas de terapia génica en los que se modifican genéticamente células tumorales de modo que expresen antígenos nuevos o ciertas interleuquinas para que aumente la inmunogenicidad del tumor. Al inyectar el gen de las interleuquinas en células tumorales *ex vivo* y luego reintroducirlas en un organismo, se observa una respuesta local inflamatoria que puede provocar el rechazo de un tumor, más por la acción de leucocitos (antiinflamatorios) que por las células T citolíticas. Por otra parte, se puede incrementar la co-estimulación y el rechazo inmunológico asumiendo que la presencia de más de una señal instigante aumenta la eficiencia de la respuesta. Las señales co-estimulantes pueden ser ligandos unidos a la membrana (como B7) o solubles como las interleuquinas. Una estrategia alternativa es la utilización de células dendríticas ya cargadas con posibles antígenos tumorales. Las células dendríticas son las presentadoras oficiales de antígenos (péptidos) e iniciadoras naturales de la respuesta inmune mediada por ambas células T: CD4⁺ auxiliares y CD8⁺ citolíticas. Estas células también pueden activar células B de memoria y células asesinas natas (NK). Por todo ello las células dendríticas se han convertido en una poderosa arma para elaborar vacunas contra el cáncer. La metodología consiste en extraer precursores o progenitores de estas células del paciente (por ejemplo, a partir de sangre periférica), cargarlas *ex vivo* al ponerlas en contacto con lisados, moléculas de los sistemas de incompatibilidad de tipos I y II (MHCI y II) que llevan péptidos, o ARNm tumorales, y devolverlas al organismo de partida. Un posible contratiempo es que las células dendríticas pueden inducir y mantener tolerancia, especialmente, pero no exclusivamente, las inmaduras. Uno de los sistemas modelo más utilizado para

estos estudios es el de melanoma murino (células B-16), por poseer, entre otros, un antígeno (mTRP-2) reconocido por células T citolíticas. Sin embargo, la incorporación del ARNm correspondiente a este péptido en células dendríticas no inhibe el crecimiento de tumores inducidos en ratones con células tumorales ni la aparición de metástasis en pulmón. La respuesta inmunológica cuando no hay tolerancia suele ser débil y transitoria con peligro de autoinmunidad. Aunque el sistema de incompatibilidad de tipo II está más implicado en la presentación de antígenos exógenos a la célula, cada vez hay más evidencia de que, de una forma indirecta y mediada probablemente por células dendríticas e interleuquinas, también este sistema es importante para el rechazo tumoral vía determinantes CD4⁺. El análisis de los resultados de más de 1.000 pacientes con cáncer metastático (principalmente melanomas) tratados con una variedad de estrategias de vacunas, incluyendo péptidos, virus recombinantes, toda la célula tumoral, células dendríticas y proteínas de choque térmico, muestra un decepcionante porcentaje de regresión tumoral de alrededor del 3%. Los tratamientos son personalizados y no suelen ser baratos.

ARN de interferencia

El ARN de interferencia (ARNi), como ya hemos apuntado en otras partes de este libro, se define como un conjunto de mecanismos de silenciamiento génico relacionados entre sí, que utiliza como molécula ejecutora final un pequeño ARN antisentido. Para tener una inhibición del 80% o más, de la proteína diana, hay que diseñar con gran precisión el ADN que servirá de molde *in vivo* del ARNi. El ARNi de diseño no debe tener más de 21 pares de bases para evitar la respuesta del interferón si se va a utilizar *in vivo*. En general se deben evitar regiones donde haya, en los mensajeros correspondientes, estructuras secundarias de horquillas o lazos, así como homologías con otras secuencias genómicas. Siendo el ARNi una molécula de doble banda que sólo carga una de las dos hebras en RISC, conviene asegurarse de que la hebra cargada es la antisentido al ARNm

diana y no su complementaria (que sería idéntica al ARNm). Esto se consigue eligiendo una secuencia que tenga el extremo 5' antisentido rico en A/U, ya que RISC elegirá la molécula con el extremo 5' más inestable. Otras reglas, como empezar con una G o C; tener una U en posición 10; A en posiciones 3 y 19; no incluir más de 4 A o T seguidas para impedir una terminación prematura de la polimerasa III, también son importantes.

Vectores retrovirales o adenovirales portadores de ADN transcribibles a ARNsi, así como moléculas de ARNi sintéticas, se han utilizado contra muchos mensajeros de genes que tienen un papel relevante en el control de la progresión del ciclo celular, la viabilidad celular, el desarrollo de algunos tumores concretos o la invasión metastática. Se presume que la efectividad que puede ser alcanzada con los ARNsi es mucho mayor que con otras estrategias alternativas como los ARN antisentido o los fármacos antimitóticos que suelen acarrear numerosos efectos secundarios o de toxicidad.

Una buena diana es un mensajero correspondiente a una proteína que se sobre-expresa y es responsable del fenotipo tumoral, pero también son dianas interesantes cualquier proteína esencial para la célula neoplásica cuya destrucción pueda acarrear la eliminación del tumor. El melanoma es un cáncer de piel agresivo, resistente a la mayoría de los tratamientos conocidos una vez que las células han empezado a extenderse a otras partes. El oncogén *B-raf* que está alterado en el 59% de las líneas celulares derivadas de melanomas y en el 66% de los melanomas, es una buena diana contra este tipo de tumores. Se han diseñado retrovirus y lentivirus portadores de secuencias que se transcriben a ARNsi contra la mutación activante *B-raf*^{V599E} presente en más del 60% de los melanomas malignos. La supresión de la variante mutada causó una parada en el crecimiento e inducción de apoptosis en líneas celulares e *in vivo*. En carcinomas de células pequeñas de pulmón, es frecuente la sobre-expresión del oncogén *skp-2* implicado en la degradación de reguladores del ciclo celular como p27Kip1, p21 y C-myc. El tratamiento con adenovirus y lentivirus portadores de secuencias convertibles en ARNi contra

skp-2 inhibió el crecimiento celular *in vitro* e *in vivo*. También se ha utilizado ARNi para silenciar el ARNm correspondiente al oncogén *her2/neu* (*erbB-2*) que se sobre-expresa en el 30% de los cánceres de mama y ovario. Se han descrito algunos éxitos rotundos en la destrucción de la célula tumoral por estos métodos en células en cultivo y en animales de experimentación, fase imprescindible para pasar a realizar ensayos clínicos con pacientes voluntarios. Algunos ensayos clínicos que utilizaron ARNi han tenido un éxito sólo parcial; quizá sea necesario afinar el vector a utilizar, o el modo en que una notable concentración de ARNs ha de llegar a las proximidades del tumor, si no es por inyección directa.

7.4.3. Terapia génica por cirugía genómica

Existe una variedad de enfermedades causadas por una simple mutación de un solo nucleótido en un gen concreto. Antes de poder reparar esta mutación en el genoma del paciente, hay que provocar un solo corte específico en el lugar exacto de la mutación. Esta cirugía genómica necesita una enzima de corte asociada a una secuencia de reconocimiento eficaz. Ya hemos visto varios de estos sofisticados métodos de cortes específicos en el genoma basados en nucleasas de dedos de zinc o TALEN (capítulo 3, apartado 3. 2.1, figura 3.19) o tijeras químicas (capítulo 3, apartado 3.1.1). Además están las nucleasas guiadas por ARN descubiertas en bacterias y arqueobacterias CRISPR (del inglés: *clustered regularly interspaced short palindromic repeat*) asociadas a la endonucleasa Cas (del inglés: CRISPR-associated) que vimos en el capítulo 2, apartado 2.3.3 y figura 2.26. Las enzimas de restricción clásicas no valen para la cirugía genómica porque las secuencias de reconocimiento son muy cortas (4-8 pb) y son demasiado frecuentes en cualquier genoma de mamíferos. Las nucleasas de dedos de zinc en su forma dimerica reconocen 36 pb, y las TALEN (del inglés: *transcription activator-like effector nucleases*) unos 24 pb. La especificidad de CRISPR-Cas se basa en el ARN unido fuertemente a CRISPR

(crARN) que guía a la endonucleasa Cas al lugar complementario en el ADN y que será el lugar de corte. La endonucleasa Cas-9 producirá un corte en las dos hebras del ADN reconocido. La eficiencia de CRISPR-Cas-9 se ha verificado de momento en células humanas en cultivo. Una vez que se ha producido el corte por cualquiera de los mecanismos expuestos, se producirá una reparación como se ha descrito anteriormente, bien por recombinación homóloga si facilitamos en las proximidades una molécula homóloga, o por reparación inexacta incluyendo pequeñas delecciones o mutaciones (capítulo 5, apartado 5.6, figura 5.17).

Estas técnicas están aún en fase experimental en células en cultivo, pero prometen intentar curar, por ejemplo, enfermedades heterocigóticas dominantes, en las que la ruptura del alelo mutado favorecerá una recombinación homóloga con el alelo sano, que eliminará la mutación causante de la patología.

7.4.4. Terapia génica, consideraciones finales

Unos 1.700 ensayos clínicos se están llevando a cabo en la actualidad contando las enfermedades genéticas hereditarias o adquiridas (cáncer). Estos ensayos están siendo seguidos muy de cerca por la comunidad científica y cualquier efecto secundario o suceso adverso se analiza profundamente.

En 1999 la Universidad de Pensilvania inició un ensayo clínico para el tratamiento de pacientes con alteraciones en el gen de la enzima transcarbamilasa de ornitina. El vector era un adenovirus Ad5 delecionado en E1 y E4. Un voluntario de 18 años (Jessie Gelsinger) recibió la dosis máxima y 4 días más tarde falleció como consecuencia de una fuerte reacción contra el vector. Las leucemias aquí descritas también son consecuencia del vector (en este caso retroviral), pero no por ello la terapia génica ha de ser abandonada, sino perfeccionada para que pueda convertirse en un futuro próximo en una verdadera opción terapéutica para una variedad de enfermedades. En los próximos años se precisan mejoras en el diseño de vectores que permitan una integración controla-

da de retrovirus y lentivirus que garanticen una mayor bioseguridad. El ARNi representa un frente adicional en la guerra contra el cáncer y contra algunas enfermedades producidas por exceso de función de genes concretos. La cirugía genómica promete ser crucial para enfermedades producidas por mutaciones dominantes negativas. Estamos en un momento crítico de identificación de los obstáculos que impiden la generalización de estas terapias en el hombre. Se necesita una estrecha colaboración entre disciplinas afines que aporten sus conocimientos para solventar los problemas observados. Genéticos y biólogos moleculares que caractericen las alteraciones génicas a nivel molecular y celular; virólogos que desarrollen vectores más eficientes, específicos y seguros, y médicos que en último término lleven a

cabo los ensayos clínicos pertinentes optimizados para cada tipo de enfermedad y paciente.

En los próximos años podríamos ver terapias *ex vivo* en las que el gen defectivo se corrija por cirugía genómica, mediada por recombinación homóloga con una copia normal, antes de devolver las células al paciente, evitando así problemas de mutagénesis insercional del vector, o cáncer.

La terapia génica ha sufrido altos y bajos continuos desde sus comienzos. Los investigadores trabajan por una continua mejora de esta poderosa baza. Es prioridad que los nuevos ensayos clínicos eviten los fallos ya observados y garanticen cada vez más la calidad del ensayo y la protección del paciente. Sólo así se podrá avanzar en esta nueva edad de oro de la biomedicina sin precedentes históricos.

LECTURAS COMPLEMENTARIAS

Gene therapy that works. Comentario de **Inder M. Verma** en la revista *Science*, **341**: 853-855. 2013.
New tool for genome surgery Comentario de **John van der Oost** en la revista *Science*, **339**: 768-770. 2013.

Refined human artificial chromosome vectors for gene therapy and animal transgenesis. **Kazuki, Y. et al.** *Gene Therapy*, **18**: 384-393. 2011.
Virus patógenos. Coordinadores, **Luis Carrasco** y **José M.ª Almendral del Río**. Fundación BBVA. Editorial Hélice. España. 2006.

Individuos transgénicos y clónicos, clonación terapéutica

Las células en cultivo no pueden representar todos los aspectos de las células en el organismo vivo. Las interacciones entre distintos tejidos, por ejemplo, no se pueden reproducir en cultivos *in vitro*. El modo más completo de analizar la expresión de un gen y su influencia en otros sistemas es la transferencia de distintas versiones, mutaciones o modificaciones del gen a la célula huevo y analizar su función a lo largo del desarrollo a nivel molecular y celular. Así, la introducción de un gen foráneo en un organismo puede tener como finalidad el estudio del comportamiento, expresión o función del gen introducido. También puede perseguir una mejora genética en el individuo afectado, con fines terapéuticos o de mejora animal. La inserción de genes en células germinales produce un adulto en el que la expresión del gen transferido puede analizarse en todos los tejidos y además se transmitirá a la progenie. Los organismos que han sido manipulados de este modo se denominan **transgénicos**, y el gen introducido, **transgén**. Otra manera de producir individuos transgénicos consiste en introducir un transgén concreto en unas pocas células embrionarias y éstas posteriormente en el embrión. Los tejidos del individuo resultante pueden derivarse de las células manipuladas o de las endógenas, constituyendo una quimera transgénica. Esta quimera puede producir descendencia totalmente modificada, si las células introducidas contribuyen a la formación de la línea germinal.

La clonación, en sentido general, es la obtención de seres genéticamente idénticos. Los **individuos clónicos** contienen la misma información

genética y, por tanto, el mismo genotipo. Cada individuo clónico constituye un clon. La clonación es un proceso fisiológico normal, ya que son clónicos los gemelos monocigóticos o univitelinos, procedentes de un solo óvulo fecundado.

La clonación terapéutica es una metodología desarrollada con la finalidad de obtener células madre a partir de un núcleo de una célula adulta somática que permitan, mediante trasplante, curar enfermedades del donante del núcleo. La metodología consiste en extraer el núcleo de una célula somática de un paciente e introducirlo en una célula huevo desnucleada o irradiada. A este embrión se le permitirá desarrollarse hasta la obtención de células madre (troncales embrionarias), que serán aisladas y usadas en terapias somáticas celulares. Una alternativa muy sólida a la clonación son las **células madre iPS** o de pluripotencia inducida, que son células diferenciadas, de la piel por ejemplo, que se han desdiferenciado a células madre.

8.1. MÉTODOS DE OBTENCIÓN DE ANIMALES TRANSGÉNICOS

Los animales transgénicos son organismos con genes de otros organismos (foráneos). El estudio de animales transgénicos ha tenido múltiples aplicaciones tanto en investigación básica como aplicada. Mucho de lo que hoy sabemos sobre la regulación de genes específicos de tejido durante el desarrollo, en el adulto y en diversas patologías, proviene de experimentos realizados en moscas y ratones transgénicos; la obtención de

productos farmacéuticos en la leche de animales de granja también se debe a esta tecnología; la creación de modelos animales que mimetizan enfermedades hereditarias humanas es otra de sus aplicaciones, así como la posibilidad de controlar la expresión de un transgén en el espacio y en el tiempo.

Existen dos métodos fundamentales para la obtención de animales transgénicos: la microinyección de ADN a huevos fecundados y la introducción de ADN en embrioblastos del blastocisto. El primer método nos da una alta proporción de transgénicos integrales, es decir, animales que han introducido el transgén en todas las células de su organismo, y más baja de quimeras transgénicas o animales que sólo han sido modificados en algunos tejidos. Cuando uno de los tejidos afectados son las gónadas, al menos la mitad de la descendencia será transgénica integral. Las predicciones de segregación mendeliana en animales transgénicos pueden estar alteradas, ya que no sabemos a priori si se ha incorporado más de una copia del transgén, ni los lugares cromosómicos donde la inserción ha tenido lugar.

8.1.1. Microinyección de ADN a huevos fecundados

Entre los animales predilectos para la realización de este tipo de experimentos, están: el sapo africano (*Xenopus laevis*), la mosca del vinagre (*Drosophila melanogaster*) y sobre todo el ratón casero (*Mus musculus*).

Xenopus transgénicos se obtienen fácilmente inyectando el ADN lineal foráneo a núcleos de esperma y éstos a huevos no fecundados. Para la expresión del transgén se puede utilizar, entre otros, un promotor-potenciador del propio *Xenopus* específico de choque térmico (*Xhsp70*), lo que garantiza su correcta expresión al subir la temperatura (de 16 °C a 34 °C) durante 30 minutos.

Drosophila melanogaster es otro de los organismos modelo por lo bien que se conoce a nivel genético, citológico y de desarrollo. *Drosophila* posee cromosomas gigantes en las glándulas salivares, y los distintos estadios del desarrollo desde

el embrión al adulto son externos (sin conexión con la madre), pudiendo manipularse fácilmente. La transferencia génica en *Drosophila* es algo más complicada que en otros organismos, ya que está ligada a los transposones o elementos P (capítulo 2, apartado 2.3.1 y figura 2.19).

Se han elaborado vectores basados en elementos P en los que se han sustituido los genes del transposón por el gen *neo* bacteriano que confiere resistencia al antibiótico G418, entre otros. El gen se ha hecho inducible por temperatura a través del promotor de la proteína 70 de choque térmico. Los transformantes se seleccionan incluyendo en la comida de las larvas 1 mg/ml del antibiótico G418, concentración normalmente tóxica. La ventaja de este sistema selectivo es que no hay que inspeccionar individualmente cada transformante; el inconveniente es que puede haber transformantes que no expresen el gen *neo*, o que los niveles de expresión no sean suficientes para conferir la resistencia deseada. Se puede usar cualquier otro marcador genético. El vector ha de ser inyectado junto con un elemento P auxiliar que aporte los genes eliminados en el vector y que son imprescindibles para la movilización. El embrión ha de ser de tipo M para permitir la transposición del elemento P inyectado. Inmediatamente después de la inyección el elemento P salta del vector, utilizando la transposasa que provee el elemento auxiliar en trans. Pueden ocurrir una o varias inserciones que se fijan y estabilizan en las siguientes generaciones (donde ya no hay transposasa).

La microinyección se lleva a cabo en la región polar del embrión que originará las células polares, que son las precursoras de la línea germinal (figura 8.1). Los embriones se desarrollan hasta moscas adultas (G0 = generación 0). En esta generación el vector P estará integrado únicamente en las células germinales, por lo que no interesa aún seleccionar por resistencia a G418. Los niveles de expresión serán muy bajos, al ser las células germinales las únicas capaces de expresar el gen introducido. Más aún, hay del orden de 30-50 células polares y puede que no todas hayan adquirido el elemento, constituyendo un mosaico o quimera transgénica. Los individuos de la G0 se

cruzan entre sí o con los padres para dar la F_1 . La selección se puede usar en las larvas F_1 . Sólo aquellos individuos que procedan de células ger-

minales con elemento P serán transgénicos y suelen representar el 10% de la progenie (figura 8.1). Una cepa se establece y mantiene por cruzamien-

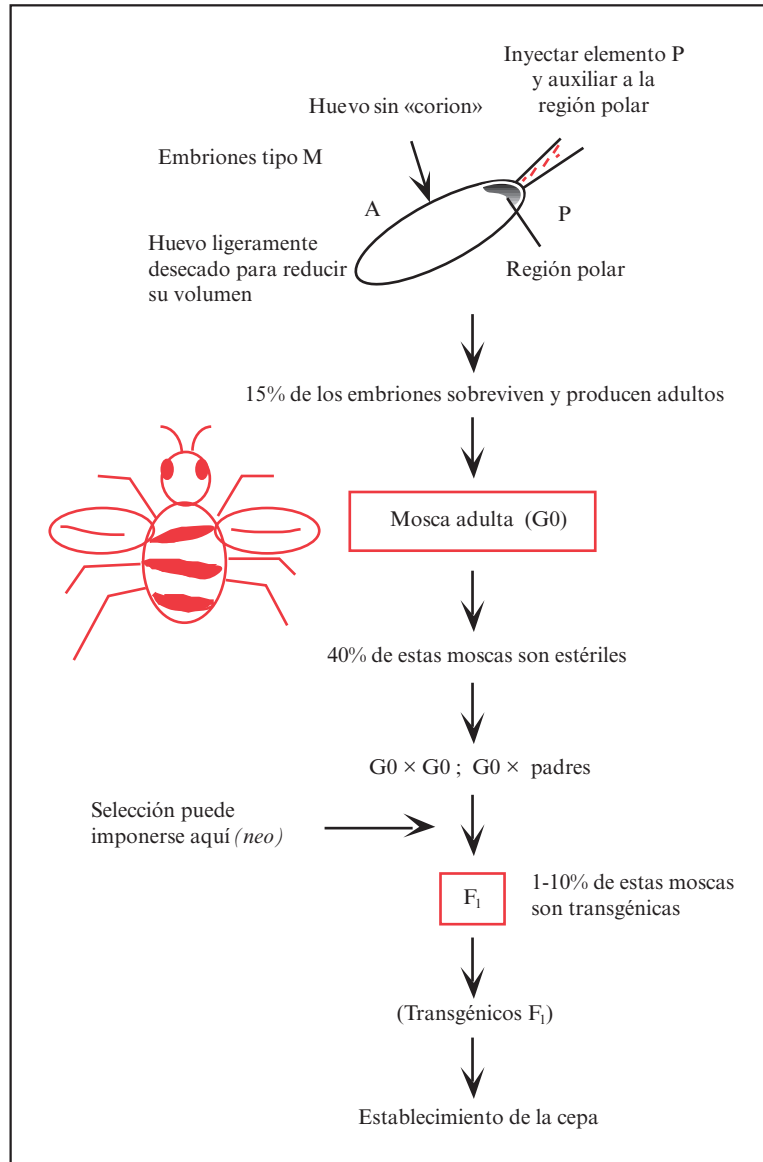


Figura 8.1. Selección de individuos transgénicos en *Drosophila*. El huevo de *Drosophila* posee una región anterior (A) y otra posterior (P). En esta última se encuentra la región polar que es donde se llevan a cabo las inyecciones. Las moscas que llegan a adultos constituyen la generación (G0). A los individuos descendientes de esta generación se les aplica la selección específica de transgénesis; los adultos constituirán la F_1 .

tos con los padres o hermanos. El elemento P se integra al azar en el genoma de la célula embrionaria hospedadora. El lugar donde ha ocurrido la inserción puede ser identificado por hibridaciones *in situ* a los cromosomas politénicos de las larvas. Se puede utilizar como sonda radiactiva el elemento P y aparecerá un precipitado de puntitos negros sobre la banda cromosómica donde haya tenido lugar la hibridación. Este método de hibridación *in situ* se ha utilizado para el aislamiento de genes colindantes o muy próximos a alguno de los lugares de integración del elemento P. Un pequeño paseo cromosómico nos llevará al gen deseado.

Mus musculus, el ratón casero, tiene grandes ventajas para realizar experimentos de laboratorio. El 99% del genoma que se expresa es idéntico al humano; además, su pequeño tamaño, resistencia a infecciones, camadas de 4-8 ratoncitos y un período de gestación medio relativamente corto (19-20 días) lo convierten en el organismo predilecto para la inducción de transgénicos. El uso de ratones también está favorecido por la existencia de numerosas mutaciones que afectan al color de la piel (albino, negro, agutí, pardo, etc.). Hay varios métodos para introducir ADN heterólogo en la línea germinal de ratones que son, en general, aplicables a cualquier mamífero. Todas las técnicas se basan en el aislamiento de huevos fecundados o embriones tempranos, un breve cultivo *in vitro* para su manipulación y su reimplantación en madres nodrizas para que prosigan su desarrollo.

Durante la formación del embrión en los mamíferos, unas seis horas después de la fecundación, se pueden distinguir con claridad los pronúcleos masculino y femenino antes de su fusión. En ratones no es difícil manipular los pronúcleos masculino y femenino en la célula huevo. En condiciones naturales, el óvulo fecundado crece hasta formar una mórula de 20-50 células, dentro de la cual algunas células se separan para formar el embrión propiamente dicho, son las células madre, también llamadas: células troncales embrionarias, blastómeros o embrioblastos; otras formarán las membranas (amnios y placenta) mediante las que el embrión permanece unido a la pared del útero

para su alimentación. Esta colección de células se denomina blastocisto (figura 8.2).

La microinyección de ADN es un método muy eficiente. Consiste esencialmente en la introducción del ADN lineal que se desee insertar en pronúcleos de huevos de ratón fecundados (figura 8.3).

La ovulación en el ratón ocurre de manera natural como respuesta a la cópula, pero puede inducirse por tratamiento con la hormona luteinizante. En la mayoría de los experimentos se induce a superovular a ratones hembras por acción hormonal y se cruzan con machos adecuados. A las 22 horas del apareamiento se extraen los huevos fecundados de sus oviductos, y se mantienen en un medio de cultivo. Al microscopio se reconocen los huevos fecundados por la presencia claramente visible de sus dos pronúcleos. La inyección se realiza en el pronúcleo masculino, que es ligeramente más grande que el femenino. Con la llegada del ADN heterólogo, este pronúcleo aumenta aún más de volumen (figura 8.3).

Con algo de práctica es posible inyectar unos 60 huevos por hora. El 80% de los huevos inyectados sobreviven al pinchazo. La reimplantación en madres nodrizas pseudopreñadas puede llevarse a cabo inmediatamente, o bien pueden cultivarse *in vitro* los huevos durante la noche (estadio de dos células) e implantarlos al día siguiente. Se pseudopreñan ratonas jóvenes copulándolas con machos estériles y en cada una de ellas se depositan de 10 a 20 embriones. Sólo de la mitad a una tercera parte de los mismos llevan a término la gestación y, de nuevo, sólo de la mitad a un tercio de estos supervivientes son transgénicos. Estos transgénicos teóricamente deberían tener el transgén integrado en el genoma de todas las células del organismo (generación G0, transgénicos integrales). Sin embargo, aproximadamente el 20% de los individuos inyectados producen quimeras transgénicas, lo cual implica que la integración del ADN ha ocurrido en estadios algo posteriores al de la inyección y no ha afectado a todos los tejidos del ratón. El número de copias que se suelen integrar varía entre 1 y 100. Cuando son muchas, suelen formar concatémeros cabeza-cola y localizarse en un único sitio, aunque algunas veces, de

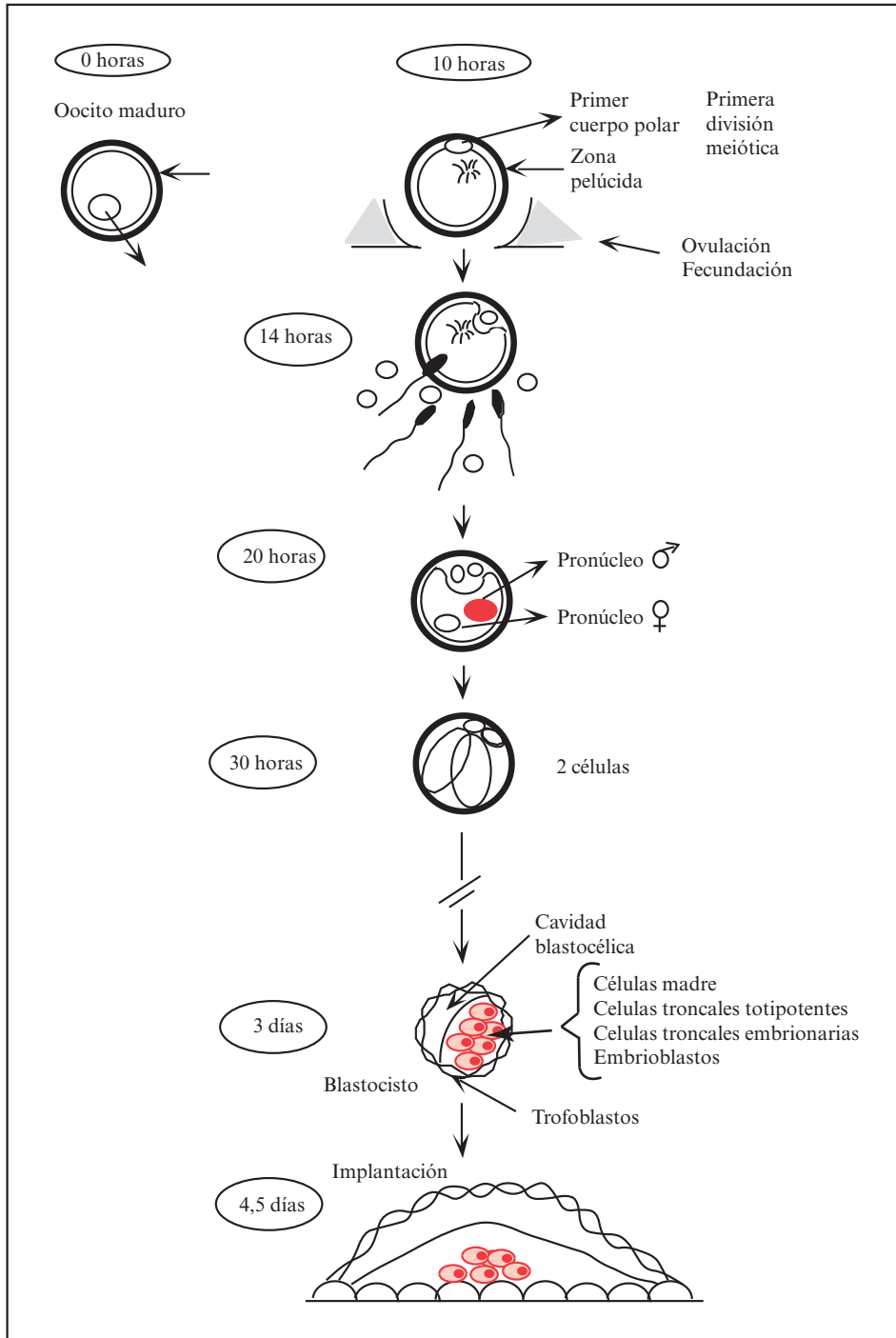


Figura 8.2. Ovulación, fecundación y desarrollo embrionario temprano en el ratón.

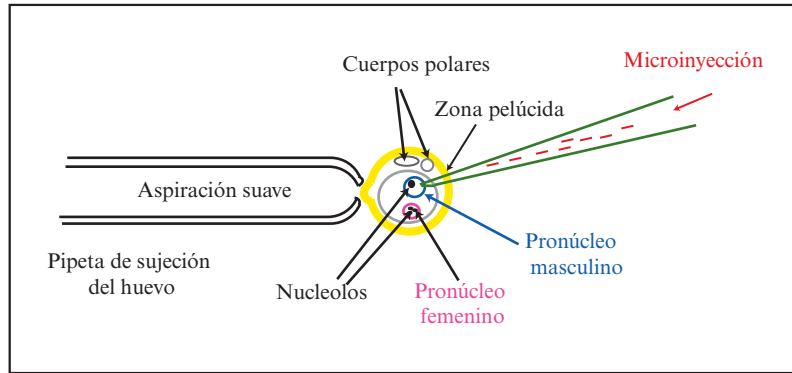


Figura 8.3. Microinyección en el pronúcleo masculino de ratón. Entre la zona pelúcida y la membrana del huevo están los cuerpos polares: uno es el resultado de la primera división meiótica, y otro, de la fecundación y segunda división meiótica.

modo extraordinario, pueden integrarse en varios sitios cromosómicos.

La confirmación de que un individuo ha incorporado el ADN inyectado y es transgénico se obtiene analizando el ADN de todos los individuos de la progenie. Unas gotas de sangre y un trocito de cola son básicamente las muestras que se toman para extraer el ADN de cada transgénico potencial y analizarlo por PCR. La primera progenie que se obtiene a partir de la microinyección constituye los fundadores transgénicos G0. En la mayoría de los casos el ADN adquirido se transmite como un carácter mendeliano y no es difícil establecer líneas estables de estos individuos (figura 8.4).

Para obtener una línea transgénica, los individuos de la G0 se cruzan con la cepa paterna por retrocruzamiento para evitar variabilidad genética adicional, con el padre si son hembras y con una hembra de la misma camada que la madre si son machos. Si el transgén se comporta como un carácter mendeliano, el 50% de su progenie será heterocigótico para el ADN insertado (hemicigóticos). Estos individuos constituyen la F₁ y la línea transgénica se puede mantener indefinidamente apareando hemicigóticos de la F₁ con individuos silvestres no transgénicos. Alternativamente, individuos de la F₁ pueden cruzarse entre sí para producir un 25% de homocigóticos transgénicos, un 50% de hemicigóticos y un 25% sin ADN heteró-

logo, siempre que el transgén no afecte a la viabilidad de la descendencia.

8.1.2. Obtención de transgénicos por introducción de ADN a embrioblastos del blastocisto

Es un método alternativo a la microinyección que produce transgénicos integrales a partir de quimeras transgénicas. Se pueden aislar los blastocistos del útero materno entre 3,5 y 4,5 días después de la cópula. Del blastocisto se aíslan células madre troncales totipotentes (embrioblastos o blastómeros), que se cultivan sobre una capa de células alimentadoras que actúan de soporte. La introducción del ADN puede realizarse por infección con retrovirus u otros tipos de transfección como la mediada por liposomas (figura 8.5).

Las células troncales manipuladas son reinsertadas en un nuevo blastocisto y éste en una madre pseudopreñada por tratamiento hormonal y apareamiento con machos vasectomizados. Al cabo de unos 18 días aparecerá la G0, cuyos individuos transgénicos serán todos quimeras, que, por retrocruzamiento con la cepa paterna, darán lugar a la F₁ que, a su vez, se analizará para detectar la presencia de hemicigotos transgénicos. Un cruceamiento entre hermanos de la F₁ dará lugar a homocigotos transgénicos que permitirán establecer

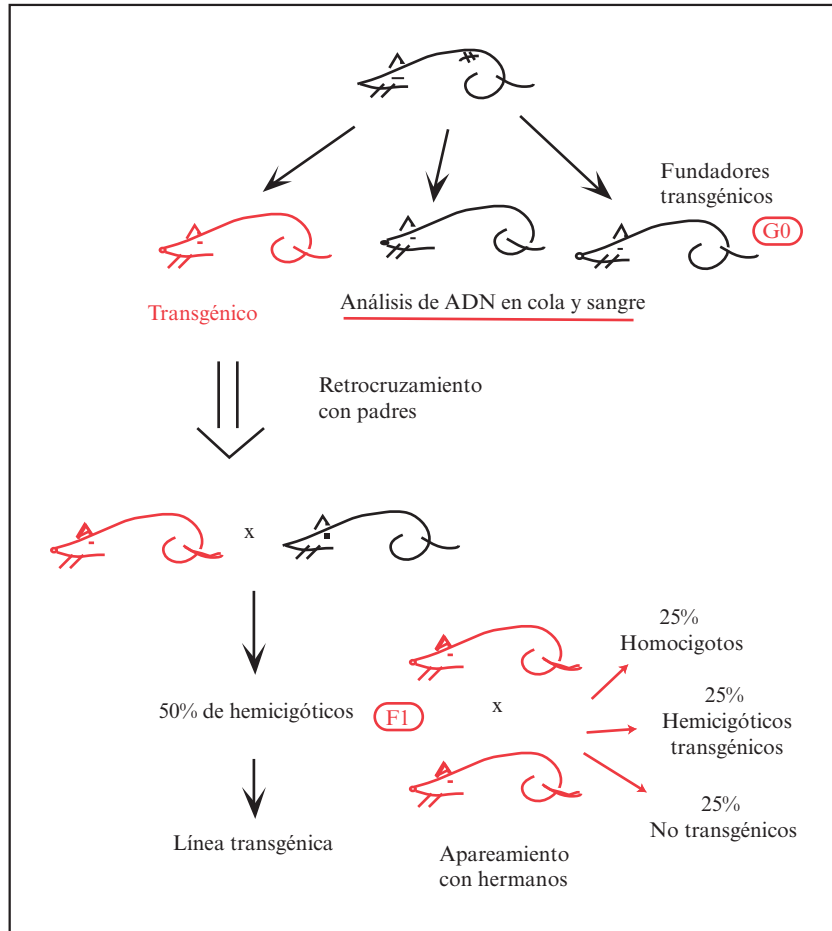


Figura 8.4. Segregación de transgénicos en ratón y establecimiento de líneas estables. La generación G0 la constituyen los adultos de huevos injectados; el análisis de su ADN por PCR determinará si son transgénicos. La generación F₁ son los descendientes de la G0 por retrocruzamiento con los padres. La línea transgénica se mantiene apareando hemicígóticos de la F₁ con individuos normales. Se pueden obtener homocigotos (25%) cruzando hermanos de la F₁.

una línea pura. Sólo la mitad de aquellos individuos cuyas células germinales hayan adquirido el gen injectado en la G0, serán transgénicos (figura 8.5).

Para la infección con retrovirus se utilizan como células soporte células productoras de retrovirus defectivos (derivados del virus de la leucemia murina) que lleven el transgén. La forma replicativa del retrovirus se suele integrar en una sola copia, y si esto ocurre en células de la línea

germinal, el provirus se transmite como un carácter mendeliano. A veces existen problemas con la expresión de genes introducidos en uno de estos vectores retrovíricos, debido a metilación de su ADN y, por tanto, bloqueo de su expresión. Las secuencias de la célula infectada que flanquean el lugar de inserción del virus también son metiladas e inactivadas. Esta metilación puede desaparecer en períodos posteriores del desarrollo del animal, y la expresión variará dependiendo de la localiza-

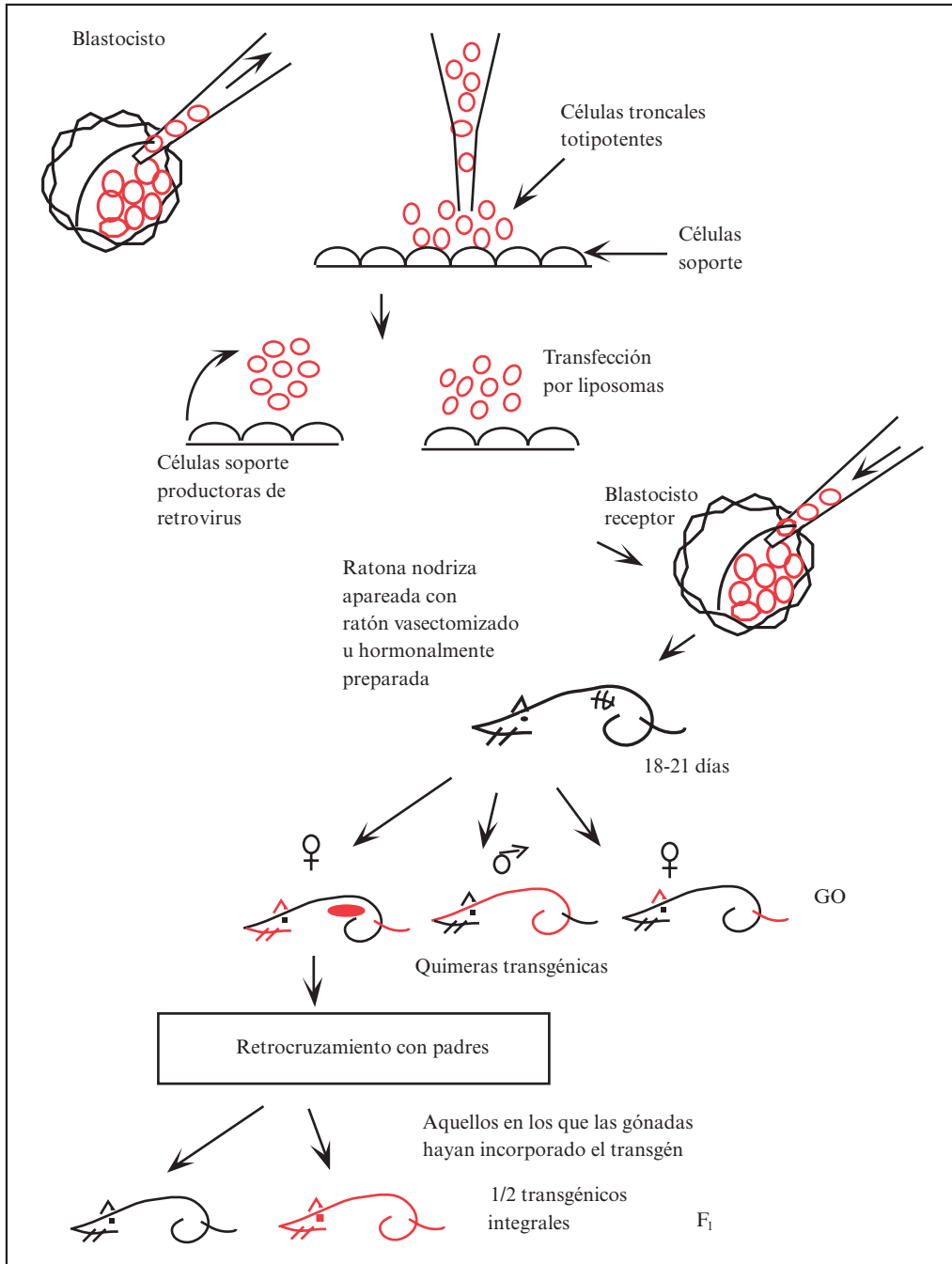


Figura 8.5. Manipulación de células troncales procedentes de blastocistos de ratón. Aislamiento de células troncales totipotentes y transfección pasiva o inyección con el retrovirus MLV (virus de la leucemia murina) *in vitro*. Reimplantación en ratonas nodrizas y análisis de la progenie.

ción cromosómica donde se encuentre. Como la integración ocurre en un lugar distinto en cada embrioblasto, la expresión diferirá de célula a célula en el blastocisto receptor.

No siempre se controla el sexo de los embrioblastos donantes y receptores, por lo que se pueden añadir blastómeros masculinos a blastocistos de hembras, provocando una quimera sexual en la que ocurrirá un cambio de sexo ($XX \rightarrow XY$) si las células XY forman la mayoría de los tejidos en el embrión que contribuyen a la determinación del sexo. También puede aparecer un individuo hermafrodita con órganos sexuales masculinos y femeninos.

La utilización de células troncales totipotentes embrionarias tiene una ventaja fundamental frente a la microinyección y es la posibilidad de seleccionar *in vitro* aquellas células que han incorporado el transgén, antes de su reincorporación a un blastocisto. La selección nos permite obtener animales transgénicos producidos por recombinación homóloga y dianas génicas, algo que no podemos llevar a cabo por microinyección en huevos fecundados.

8.2. EXPRESIÓN DEL TRANSGÉN Y APLICACIONES DE ANIMALES TRANSGÉNICOS

En la mayoría de los casos, el ADN transferido se expresa en los tejidos del animal transgénico. La cantidad y calidad de esta expresión varía considerablemente de unos casos a otros.

Ratones transgénicos

Uno de los primeros experimentos transgénicos con expresión espectacular del gen inyectado fue la introducción en el genoma de ratón de la hormona del crecimiento de la rata y el desarrollo de superratones o ratones gigantes. El vector utilizado (pMGH) tenía 8,9 kb, llevaba secuencias del plásmido pBR para su crecimiento y obtención de grandes cantidades en un sistema bacteriano y el promotor de la metalotioneína que es inducible por cobre, zinc y otros metales. Una de las funciones de la metalotioneína *in vivo* es la de unirse a metales para detoxificar el organismo.

Por último, el vector llevaba el gen de la hormona del crecimiento de rata. Este vector linealizado fue inyectado en pronúcleos de ratón. De los 170 huevos inyectados, se desarrollaron 21 ratoncitos (G0) de los que 7 resultaron transgénicos. Se detectaron de 1 a 35 copias insertadas por ratón transgénico, siendo el tamaño del ratón proporcional al número de copias adquiridas. La expresión se indujo con Zn^{+2} , pero en realidad no habría sido necesario, pues en este caso se expresó constitutivamente (no se sabe la razón). Los niveles de la hormona de crecimiento llegaron a ser 800 veces mayores que los niveles normales, y el tamaño de algunos de estos ratones fue el doble que el de sus hermanos no transgénicos.

También se han obtenido ratones transgénicos portadores de un potente oncogén, como es el antígeno T del virus SV40, bajo el control del promotor del gen de la insulina de la rata, que confiere la especificidad de tejido. Los transgénicos resultantes desarrollaron insulinomas pancreáticos pudiendo estudiarse el desarrollo, angiogénesis y progresión de este tipo de tumores. En el caso del oncogén *fos* la expresión de niveles altos del oncogén producida por copias adicionales transgénicas conlleva el desarrollo de osteosarcomas y condrosarcomas, dos tipos distintos de tumores óseos que surgen de la proliferación incontrolada de células implicadas en el desarrollo del hueso y cartílago, respectivamente.

Ratones KO

Una de las aplicaciones más interesantes de los animales transgénicos, y en particular de los ratones, es la posibilidad de eliminar la función de un gen (KO, del inglés *knock-out*), para observar el efecto de la falta de función en el organismo completo. La producción de ratones KO se realiza por el procedimiento de introducir ADN en embrioblastos del blastocisto. Las células madre (embrioblastos) son infectadas con vectores que lleven, por una parte, un fragmento del gen que se desea eliminar, interrumpido por un gen de selección positiva (*puro*, *neo*), y, por otra, un segundo gen de selección negativa contra las integraciones al azar, fuera de la región de homología (*tkHs*). Esta estra-

tegia permite seleccionar las células en las que el ADN exógeno se ha incorporado por una recombinación homóloga (diana génica) sustituyendo el gen endógeno por el interrumpido (véase capítulo 5, apartado 5.6 y figura 5.15). Estos experimentos son difíciles de realizar, ya que la frecuencia de éxito es francamente baja (una diana por 3×10^6 células transfectadas), necesitando tratar un gran número de embrioblastos. Si el gen que se desea deleccionar se encuentra en el cromosoma X y los embrioblastos son masculinos, sólo necesitaremos una diana génica para tener un individuo KO, pero, si la región de interés es autosómica, se deberán realizar cruzamientos entre hermanos procedentes de la línea transgénica establecida para obtener un 25% de homocigóticos para el transgén. No es posible llevar a cabo esta tecnología en huevos fecundados microinyectados, ya que tendríamos que utilizar millones de huevos y no podríamos seleccionar por el suceso infrecuente en que hubiese ocurrido una diana génica.

Un ejemplo de ratones KO son los ratones SCID (síndrome de la inmunodeficiencia combinada severa), transgénicos a los que se les ha inducido una mutación en el gen que codifica la enzima deaminasa de adenosinas (ADA) o el receptor γ C de las interleuquinas 2, 4, 7, 9, 15 y 21 (capítulo 7, apartado 7.4.1). Son modelos animales excelentes para la enfermedad homóloga humana que afecta a los denominados niños burbuja, porque han de ser mantenidos en esterilidad casi absoluta.

En cáncer, se ha obtenido un gran número de ratones transgénicos KO para las dos copias de algún oncosupresor (p53, por ejemplo). Estos transgénicos desarrollan tumores con una frecuencia muy alta sirviendo de modelo para terapias antitumorales.

La inducción de ratones KO para cientos de genes en un pasado reciente ha puesto de manifiesto la gran redundancia funcional que parece existir en el genoma de mamíferos. Muchos ratones a los que se les elimina una función localizada en un gen concreto único no parecen sufrir ninguna disfunción detectable. Es frecuente que determinados genes sean capaces de suplir la función de otros, constituyendo una verdadera redundancia funcio-

nal. En este sentido conviene recordar que han ocurrido a lo largo de la evolución dos pasos sucesivos de ploidía desde los cordados primitivos a los vertebrados. Es muy posible que familias de genes relacionados que pudieron tener un único antecesor aún mantengan suficiente similitud para suplir o compensar sus respectivas funciones.

Cerdos transgénicos

Los resultados de los superratones animaron a los granjeros estadounidenses y, durante los años 1988 y 1989, se inyectó a cientos de embriones de cerdo con construcciones de promotores constitutivos y genes de hormonas de crecimiento (somatotropinas) que asegurasen niveles altos de la hormona en todo momento. Contrariamente a lo que sucedió con los superratones, no se obtuvieron supercerdos. La calidad de los jamones era, sin embargo, muy buena, con poca grasa y mucho magro. Desafortunadamente estos cerdos transgénicos empezaron a desarrollar diabetes, esterilidad y otras patologías. No se había prestado suficiente atención a la endocrinología del animal, y en cerdos es importante que los niveles de la hormona oscilen periódicamente, ya que los niveles altos y constitutivos inducen letalidad. Estos experimentos ponen de manifiesto lo importante que puede ser el control de la expresión del gen introducido.

Se han obtenido cerdos transgénicos que producen grandes cantidades de eritropoyetina humana en la leche (figura 8.6). La eritropoyetina (EPO) es una hormona glicoproteica de 165 aminoácidos que se sintetiza preferentemente en el riñón y circula por el torrente sanguíneo. Estimula la eritropoyesis, incrementando la producción de glóbulos rojos, por lo que se utiliza como fármaco contra anemias y enfermedades renales. Pero es también la hormona del dopaje, ya que eleva el hematocrito aumentando considerablemente la cantidad de eritrocitos en sangre, permitiendo así una mayor resistencia durante el ejercicio físico. Se utilizó el promotor de la proteína ácida de la leche, seguido del gen de la eritropoyetina humana de 7,8 kb con dos intrones para dar estabilidad al transgén y una región de polia-

denilación al final de la construcción transgénica (figura 8.6). Los transgénicos se verificaron por la presencia de tres bandas de 900, 300 y 500 nu-

cleótidos, amplificadas por PCR y correspondientes al gen de la eritropoyetina humana. El gen endógeno de cerdo tiene tres intrones.

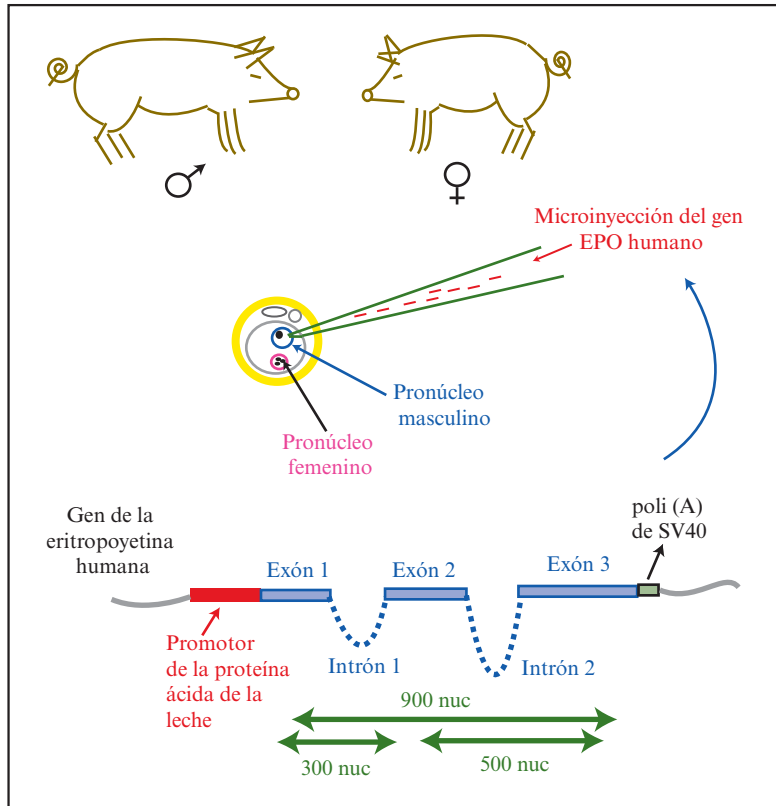


Figura 8.6. Cerdos transgénicos que producen eritropoyetina humana en la leche de las hembras. Microinyección del transgén en el pronúcleo masculino de embrión. Las dobles flechas de la parte inferior indican el tamaño de los fragmentos amplificados por PCR, que verifican la inserción del transgén humano en el genoma del cerdo.

Salmones transgénicos

En salmones, se han obtenido salmones transgénicos de 2 a 13 veces el tamaño de sus hermanos no transgénicos. El gen de la hormona del crecimiento del salmón *chinook* se introdujo en huevos fecundados del salmón *atlántico*. Se microinyectaron 800 huevos en el micropilo con 10^6 copias de una construcción que llevaba el promotor de la proteína anticongelante de la faneca unido al ADNc de la hormona de crecimiento

(figura 8.7). La presencia del transgén se detectó por el tamaño de los salmones y se confirmó en sangre y escamas por PCR usando cebadores específicos de la construcción. El gen de la hormona de crecimiento del salmón (*hc*) posee un intrón que no estará presente en la versión de ADNc. La diferencia de tamaño entre ambos se puede observar por PCR utilizando oligocebadores al principio y final de la secuencia codificante. Sólo los transgénicos poseerán la banda de menor tamaño.

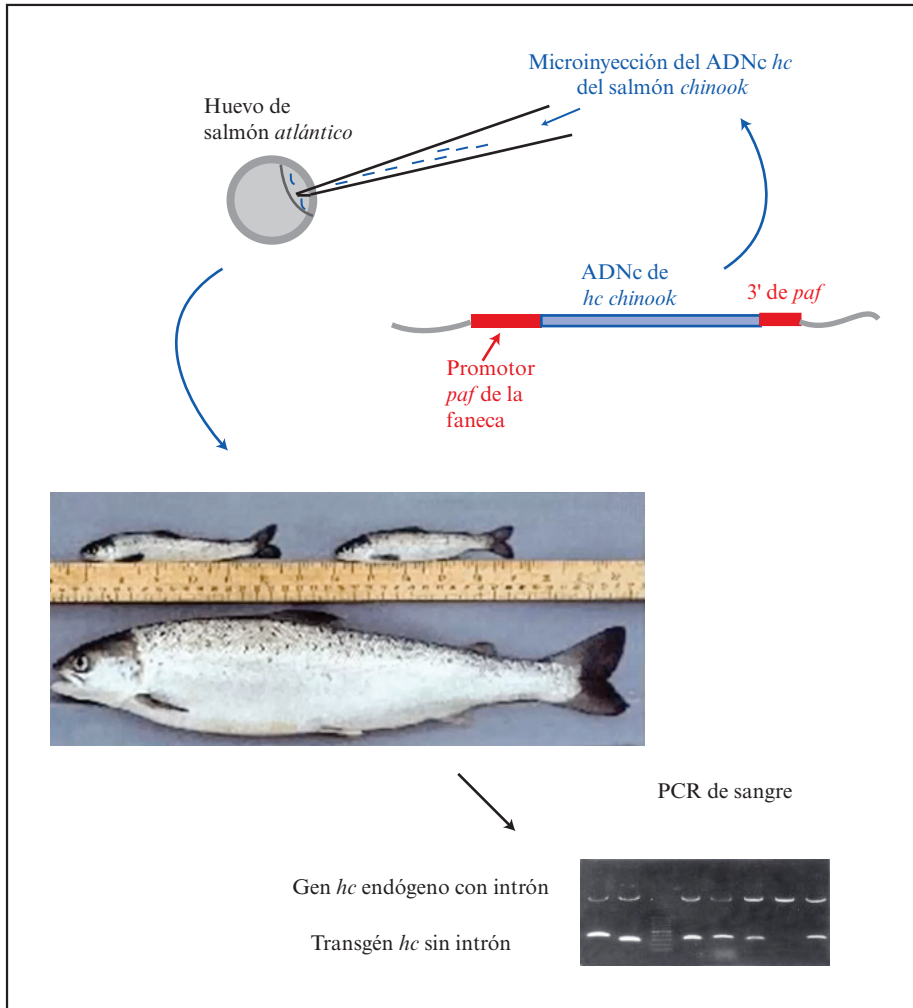


Figura 8.7. Salmones transgénicos. Microinyección de huevos con una construcción que lleva el promotor de la proteína anticongelante de la faneca (*Ppaf*) unido al ADNc del gen de la hormona de crecimiento (*hc*) del salmón chinook. Tamaño de un salmón transgénico y de 2 hermanos silvestres en la parte superior. Verificación por PCR del tamaño del gen endógeno y del transgén que carece del intrón que posee el endógeno.

Vacas transgénicas

En algunos países, como Estados Unidos, se comercializa desde 2008 carne de vacas transgénicas portadoras de la hormona de crecimiento.

También se comercializa en varios países la lactoferrina humana producida en la leche de vacas transgénicas. La proteína se usa como su-

plemento alimenticio, ya que parece tener propiedades inmunoestimulantes, antibacterianas y antivirales. La microinyección de un fragmento de ADNc obtenido a partir del mensajero de la lactoferrina humana (LHh, proteína que se une a hierro), fusionado al promotor del gen de la caseína $\alpha S1$, permitió la expresión de este gen en la glándula mamaria y la obtención de la proteína

de la leche. Como alternativa a otros sistemas de producción en células de mamíferos, la glándula mamaria es capaz de realizar todas las modificaciones pertinentes postranscripcionales o a nivel de proteína. Además, el sistema es enormemente atractivo por la simplicidad de acceso a la proteína, la potencialidad productiva del tejido y el bajo coste una vez que se obtienen los animales transgénicos. Típicamente, el gen de interés, tanto en la versión genómica como en ADNc, se fusiona con las secuencias reguladoras del gen de una de las proteínas de la leche (caseína, proteína ácida del suero o β -lactoglobulina) para generar transgénicos. El producto deseado se recoge de la leche.

Control del transgén en el espacio y en el tiempo

El control del transgén en el espacio consiste en poder expresarlo sólo en determinados tejidos y ello se consigue en construcciones en las que el gen de interés va precedido por un promotor específico de tejido (tejido-específico) que garantiza, en gran medida, una expresión confinada al tejido diana. No existe en la actualidad ningún promotor tejido-específico que sea realmente eficiente en el tejido propio y verdaderamente nulo en otros tejidos, pero hay varios que son capaces de modular una expresión de hasta 50 veces superior en el tejido diana con respecto a otros. El promotor denominado específico de hígado se transcribe con clara preferencia en este órgano; el correspondiente a la proteína básica de la mielina lo hace con preferencia en astrocitos, y los de las proteínas de la leche (lactoalbúmina, proteína ácida del suero o caseína) son muy eficientes en glándula mamaria, pero en ningún caso los resultados son contundentemente negativos en otros órganos.

Una variante que permite la eliminación de la expresión de un gen en algún órgano (sería un KO tejido-específico) se consigue utilizando el sistema de recombinación específica de sitio Cre/*loxP* del bacteriófago P₁ (capítulo 2, apartado 2.2.2). Cre es una recombinasa capaz de eliminar cualquier secuencia localizada entre dos sitios *loxP*. Así podemos introducir por recombinación homóloga sitios *loxP* en los extremos de un gen de interés y

facilitar el gen *cre* bajo el control de un promotor tejido-específico. Se producirá la eliminación exclusiva del gen diana en el tejido controlado por el promotor (figura 8.8). Para este tipo de experimentos se emplean dos cepas de ratones: una que lleva el gen de interés flanqueado por *loxP* obtenido por diana génica en embrioblastos, y la otra que lleva *cre* como transgén controlado por el promotor tejido-específico (específico de hígado, por ejemplo). La progenie de cruces entre estas dos cepas eliminará el gen de interés sólo en hígado.

El control del transgén en el tiempo se puede hacer con el sistema de resistencia a tetraciclina bacteriano ya estudiado (véase capítulo 7, apartado 7.3.1 y figura 7.4), en sus dos variantes, positivo y negativo. Un promotor sintético formado por varias subunidades del operador de la tetraciclina y el potente promotor eucariótico de citomegalovirus constituye el elemento de respuesta a la tetraciclina que se fusiona al gen de interés. Por otra parte, el animal debe ser ya transgénico para el gen fusión del represor de la tetraciclina con el activador *VPI6* (*i-tetrVPI6*). La presencia de tetraciclina o doxiciclina (análogo de la tetraciclina) inhibe la expresión del gen, mientras que en su ausencia se expresa. El sistema no consigue inhibiciones totales, sino más bien inducciones de la expresión sobre un cierto nivel basal.

8.3. INDIVIDUOS CLÓNICOS POR TRASPLANTE DE NÚCLEOS

Los primeros experimentos de trasplante de núcleos se realizaron en anfibios (*Rana pipiens* y *Xenopus laevis*) debido preferentemente a que los huevos fecundados son de gran tamaño (1-2 mm) y la embriogénesis ocurre fuera del cuerpo materno. En la década de los cincuenta se transfirieron con éxito núcleos de células en el estadio de blástula a huevos desnucleados. La desnucleación se conseguía pinchando ligeramente el huevo en la región del polo animal donde se encuentra el huso acromático meiótico y los cromosomas. El pinchazo no sólo facilitaba la salida del núcleo del oocito, sino que inducía la activación del hue-

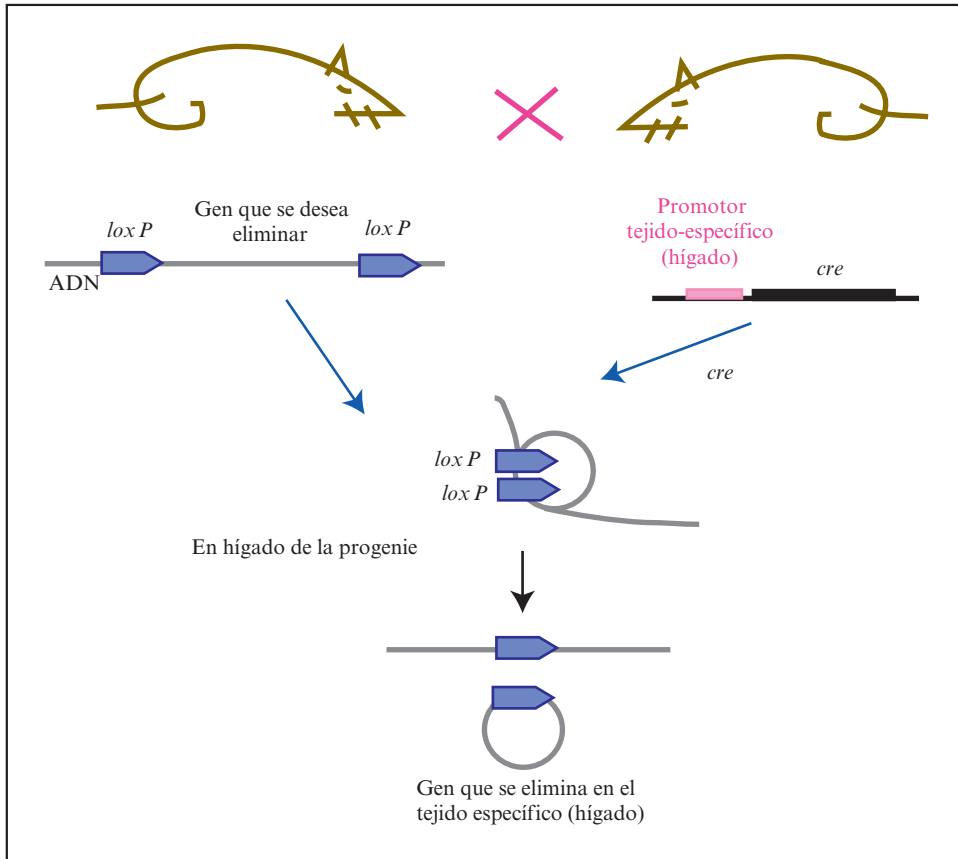


Figura 8.8. Obtención de ratones transgénicos con un gen eliminado en un tejido concreto. Se cruzan dos ratones, de los cuales uno lleva el gen a eliminar flanqueado por dos elementos *loxP* introducidos previamente por recombinación homóloga. El otro ratón tiene en su genoma integrado el gen *cre* bajo el control de un promotor específico de tejido (hígado, por ejemplo). La progenie carecerá de este gen sólo en hígado pudiendo observarse el efecto que ello produce en todo el organismo.

vo iniciándose todos los cambios citológicos y bioquímicos asociados a la fecundación. Seguidamente, mediante un nuevo pinchazo, con una micropipeta se introducía el núcleo procedente de células embrionarias. De esta forma se produjeron los primeros renacuajos y ranas adultas con material genético de un solo progenitor. Los mismos investigadores intentaron sin éxito obtener renacuajos a partir de núcleos procedentes de células de una rana adulta en vez de los núcleos embrionarios de blástula. Este fallo experimental parecía indicar que, cuando las células habían ad-

quirido un grado de diferenciación y especialización considerable, ya no eran capaces de reprogramarse para dirigir el desarrollo total de una célula huevo a adulto. Algunos años más tarde, no obstante, John Gurdon logró obtener algunos sapos adultos a partir de núcleos de células epiteliales de intestino de renacuajo introducidos en huevos de sapo (*Xenopus laevis*). El procedimiento consistió en irradiar el huevo con luz ultravioleta para destruir el material genético materno antes de inyectar los núcleos donantes de intestino de renacuajo. La eficiencia obtenida fue muy

baja y la mayoría de los 726 núcleos transferidos a huevos irradiados dieron lugar a embriones anormales que no prosperaron o murieron como renacuajos; siete de los huevos manipulados, sin embargo, lograron desarrollarse totalmente llegando a ser sapos adultos y fértiles (figura 8.9). Los núcleos de intestino de los renacuajos son, pues, totipotentes y su material genético es capaz de orquestar el desarrollo completo del individuo. Por estos experimentos John Gurdon recibió el Premio Nobel de Fisiología y Medicina en el año 2012; premio que compartió con Shinya Ya-

manaka, que estudió la clonación terapéutica con células iPS, es decir, la posibilidad de obtener células madre a partir de células adultas especializadas desdiferenciándolas que veremos más adelante en este capítulo (apartado 8.5).

En el ratón, intentos de clonación en los que el núcleo donante procediera de embriones de ratón en los estadios de 4 u 8 células embrionarias fallaron estrepitosamente. Sin embargo, no todas las especies de mamíferos se comportan de la misma manera en los primeros estadios del desarrollo y en la implantación.

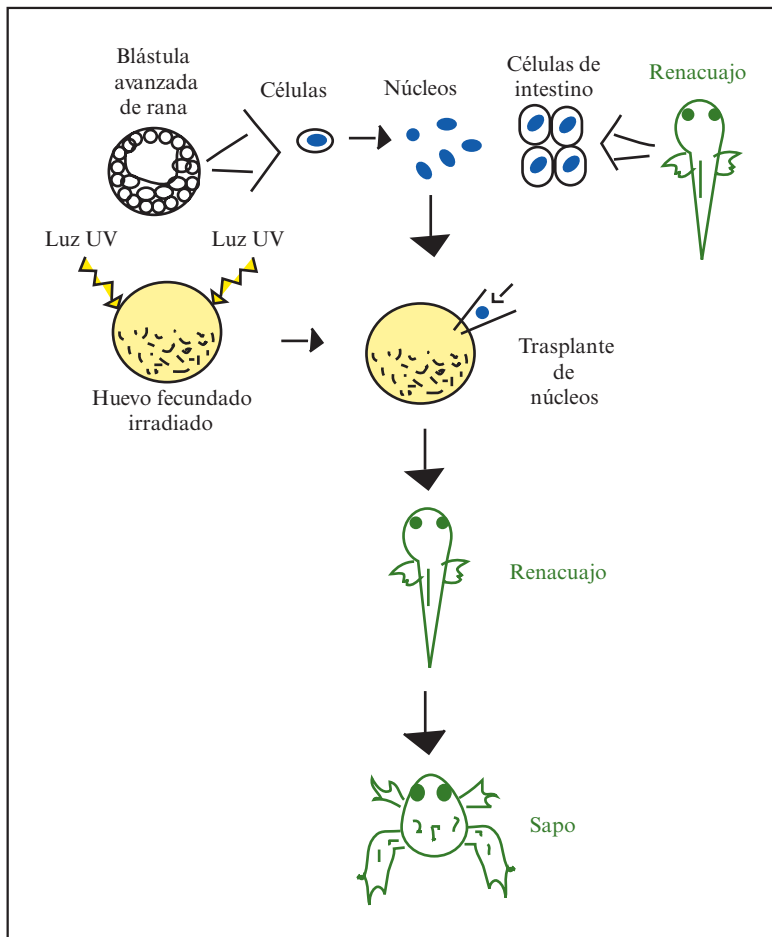


Figura 8.9. Trasplante de núcleos en anfibios. Núcleos procedentes de células indiferenciadas de blástula avanzada o diferenciadas de intestino de renacuajo se trasplantan a huevos fecundados irradiados, originando renacuajos y sapos adultos.

Los estudios de trasplantes de núcleos, tanto en anfibios como en mamíferos, indicaron que la causa de las limitaciones observadas en la aplicación de esta técnica eran las incompatibilidades entre el estado del ciclo celular del núcleo introducido y del citoplasma recipiente. El ciclo celular (figura 8.10) se divide en 4 períodos: G1 (del inglés *gap*), intervalo entre la división nuclear previa y el inicio de la síntesis de ADN; S, el período de síntesis de ADN; G2 intervalo entre la duplicación del ADN y la división nuclear, y M, el período de mitosis durante el cual los dos nuevos genomas se separan en dos células hijas. Las células quiescentes por largos períodos de tiempo se consideran en G0, pudiendo reincorporarse al inicio de G1 en determinadas circunstancias. Así como durante los períodos G0 y G1 las células son diploides en cuanto a su contenido de material genético ($2n$), tanto en S como en G2 y hasta el final de M, las células tienen 4 veces el contenido haploide ($4n$). La mayoría de los núcleos usados para trasplantes estaban en fases S o G2 (tetraploide), incompatible con el citoplasma diploide de la célula huevo, porque el núcleo recién llegado tiende a replicarse de nuevo siguiendo las órdenes dictadas por el citoplasma receptor, y, si ya era tetraploide, una nueva división ocasionaría un

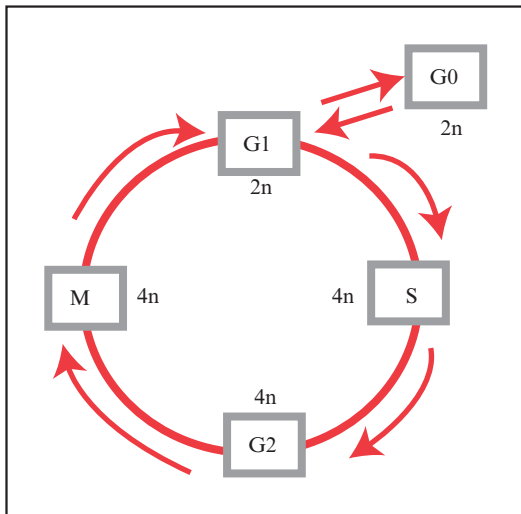


Figura 8.10. Fases sucesivas del ciclo celular.

gran número de aneuploidías y otras aberraciones cromosómicas que, con una gran probabilidad, causarían letalidad en el embrión durante el desarrollo. Las dificultades para la clonación en ratones quedaron así explicadas conociendo que los núcleos en los estados de preimplantación se encontraban preferentemente en S y G2, por lo que habría sido muy difícil obtener células embrionarias (blastómeros) en G1. Analizando los resultados obtenidos en clonaciones de ovejas, se observó que el desarrollo llegaba mucho más lejos si los núcleos donantes procedían de blastómeros al final de la mitosis o en G1 que si éstos procedían de células en fases S o G2.

Así, aunque la elección del ratón y las condiciones experimentales iniciales resultaron ser desafortunadas, con algunas modificaciones sobre el procedimiento descrito para ellos, se pudieron obtener ovejas, cerdos, conejos y vacas adultas a partir de núcleos procedentes de células embrionarias en estadios tan tardíos como mórula temprana (8-16 células). También se ha obtenido descendencia viable en ovejas y vacas utilizando, como núcleos donantes del material genético, blastómeros mantenidos en cultivos celulares por algunas generaciones.

En 1997 Ian Wilmut y colaboradores logran por vez primera una oveja clónica (conocida como Dolly, en honor a los exuberantes pechos de la cantante norteamericana Dolly Parton) a partir del núcleo de una célula totalmente diferenciada procedente de unas células congeladas de la ubre de otra oveja adulta que ya había fallecido. El éxito se debió fundamentalmente a que las células que aportaron el núcleo (células de la glándula mamaria de una oveja de 6 años en su último trimestre de preñez) fueron obligadas (por privación de suero) a salirse del ciclo celular y entrar en la fase G0 antes de la transferencia nuclear. Esta situación indujo a las células donantes a entrar en un estado de quiescencia, provocando presumiblemente cambios en la estructura de la cromatina que facilitan la reprogramación de la expresión génica a su llegada al citoplasma del oocito. Es muy posible, además, que en el estadio G0 o G1 del ciclo celular, el material genético tenga un acceso más directo a los factores de transcripción, proteí-

nas de unión a cromatina, etc. También hay que considerar el hecho de que en ovejas la transcripción no comienza hasta la mórula de 8-16 células, mientras que en el ratón hay transcripción en el estado de sólo dos células. Esto daría la oportunidad a los núcleos recién trasplantados de oveja a dividirse dos veces y reprogramarse, antes de empezar a expresar sus propios genes específicos de desarrollo. En los experimentos descritos por Wilmut y colaboradores, la gestación concluyó con

normalidad aunque con muy baja eficiencia. De 277 núcleos de glándula mamaria trasplantados, 29 llegaron al estado de mórula y 13 ovejas nodrizas resultaron preñadas, pero solamente una oveja (Dolly) nació viva, mostrando las características (fenotipo) del núcleo donante (cara blanca) frente a las del huevo receptor (cara negra). La oveja cuyo núcleo celular fue trasplantado a un huevo desnucleado es genéticamente idéntica a Dolly (figura 8.11). Esta oveja constituyó el primer mamí-

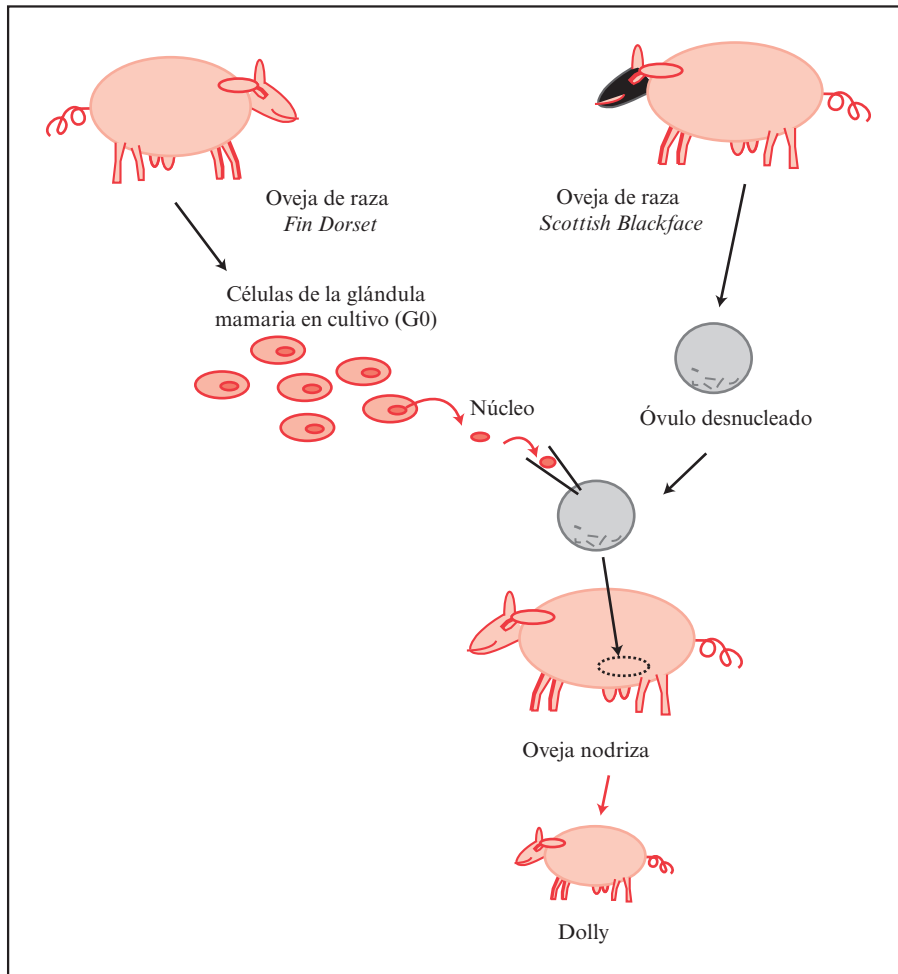


Figura 8.11. Clonación de la oveja Dolly. A la izquierda la oveja donante del núcleo y a la derecha la oveja de cara negra, donante del óvulo desnucleado. El núcleo resultante se implanta en una oveja nodriza. Dolly es un clon de la oveja donante del núcleo.

fero que se desarrolló felizmente a término a partir de una célula diferenciada (glándula mamaria) de otro adulto. Dolly, sin embargo, murió prematuramente a los 6 años de edad de un adenocarcinoma pulmonar ovino, un tipo de cáncer de origen vírico muy común en Europa y que ataca a ovejas jóvenes. La vida media de las ovejas es de unos 11 años en condiciones óptimas, y lo más probable es que la muerte de Dolly no tuviera nada que ver con la clonación, o al menos es lo que opinan sus creadores.

Desde entonces se han clonado unas 20 especies animales, entre ellas decenas de ratones, ratas, gatos, cerdos, conejos, venados, terneros, una yegua de carreras y un toro de lidia. En muchos casos los animales son, además, transgénicos. Aunque la mayoría de estos experimentos tienen objetivos científicos, en algunos casos el objetivo es claramente económico.

La evolución ha estado siempre ligada a la generación de variabilidad genética debido a que durante la meiosis se forman nuevas combinaciones de alelos que se transmiten a la descendencia y esto genera individuos más viables y mejor adaptados. El proceso de clonación tendería a la homogeneidad y probablemente si se aplica en exceso sea, a largo plazo, más bien deletéreo para las especies.

Una de las primeras conclusiones a sacar de estos experimentos de clonación es que las células diferenciadas de mamíferos mantienen la capacidad de cambiar su programación genética y que, por tanto, durante la especialización a tejidos concretos no hay pérdida sustancial de información genética.

Otra conclusión importante de los experimentos de clonación es la posibilidad de obtener mamíferos adultos por vía asexual. El único papel de la célula huevo (femenina en todos los casos) es permitir que su citoplasma reprogramme la expresión genética nuclear y contribuya con orgánulos celulares (mitocondrias preferentemente) al futuro organismo.

No parece lejano el día en que se clone el ser humano, pero surge la pregunta: ¿hay alguna justificación científica, médica o ética para generalizar la clonación en la especie humana? La justifi-

cación científica se basa en el propio avance del conocimiento que supondría el estudio de hombres y mujeres genéticamente idénticos. Ello permitiría determinar, por ejemplo, hasta dónde pueden influir los factores no genético-nucleares (ambientales o herencia mitocondrial) para crear variabilidad y diversidad en la especie humana. También se podrían establecer los límites de la influencia genómica nuclear en el comportamiento de personas genéticamente idénticas. Sin embargo, aproximaciones a estas interesantes preguntas pueden ser contestadas con estudios en gemelos univitelinos o con clonaciones en animales más o menos cercanos al hombre, por lo que la justificación científica es solamente parcial. Con los conocimientos biológicos que se tienen en el presente, no se puede afirmar que los individuos están determinados totalmente por su material genético. Lo que denominamos personalidad de un individuo está influenciada, además de por su genoma, por el ambiente físico, cultural y social en el que se desarrolla. Los gemelos idénticos (univitelinos) crecidos por separado constituyen el ejemplo más cercano a lo que pudiera ser la clonación en seres humanos (las similitudes y coincidencias entre este tipo de individuos pueden llegar a ser asombrosas, como la de acudir a una cita, ya de mayores, ambos con 5 anillos distribuidos en los mismos dedos, o poner el mismo nombre a su perro aun sin haberse conocido nunca). Estos individuos también genéticamente idénticos (aunque uno no es descendiente del otro) pueden ser bastante diferentes. Dos clones humanos, separados por varias décadas (en el tiempo), probablemente diferirán psicológicamente mucho más que dos gemelos idénticos que viven en el seno de la misma familia.

Médicamente, con la clonación, podrían desaparecer algunas esterilidades conocidas. Hombres y mujeres incapaces de producir gametos podrían tener descendencia. La reproducción podría ser asexual y monoica, ya que todo el material genético procede de un único individuo. El caso de identidad mayor se conseguiría con una mujer cuyo óvulo desnucleado fuese inoculado con un núcleo diploide de un tejido diferenciado propio, y ella misma gestara el embrión. La reproducción

ya no tendría que estar asociada a parejas (macho-hembra), sino a individuos, ya que hasta donde sabemos el descendiente sería genéticamente idéntico al donante del núcleo celular. Dos mujeres lesbianas podrían contribuir una con el citoplasma y la otra con su material genético nuclear a una hija de ambas.

Una segunda aplicación médica de la clonación, muy interesante y actual, es la clonación con fines terapéuticos, que veremos en detalle más adelante en este capítulo (apartado 8.5) y que consiste en la utilización del embrión para obtener de él células madre (troncales totipotentes) que pueden, en principio, generar cualquier tejido para trasplantarlo al individuo del que se obtuvo el clon, eliminando así cualquier problema de rechazo inmunológico.

La sociedad en que vivimos ha condenado la clonación en seres humanos si no es una clonación con fines terapéuticos. Existe un miedo genuino, y probablemente justificado, a que estas tecnologías se nos vayan de las manos y degeneren por derroteros incontrollables. Entre los muchos casos posibles, tendríamos el de la clonación con fines de inmortalización de líderes, etc.

La clonación en animales puede considerarse como un avance científico con aplicaciones múltiples en investigación básica y en biotecnología. La clonación humana no parece tener en la actualidad suficiente justificación científica, médica o ética para llevarse a cabo; además en muchos países, entre ellos España, está terminantemente prohibida.

8.4. ANIMALES CLÓNICO-TRANSGÉNICOS DE INTERÉS FARMACOLÓGICO Y BIOMÉDICO

Veremos como ejemplo la obtención de ovejas clónico-transgénicas para el factor IX anticoagulante humano. En los experimentos originales, fibroblastos fetales se co-transfectaron con el gen *neo*, que confiere resistencia a G418, y con el ADNc del factor IX anticoagulante humano fusionado a la región promotora de la β -lactoglobulina, para su expresión exclusiva en la glándula

mamaria del animal. Después de seleccionadas las células por resistencia a G418, los núcleos de algunas de ellas fueron transferidos a oocitos desnucleados (figura 8.12) y éstos a ovejas nodrizas. En este experimento se obtuvieron seis ovejas clónicas transgénicas: tres de ellas para ambos genes *neo* y *factor IX*, y las otras tres, sólo para el marcador *neo*. La producción de ovejas clónico-transgénicas por el método de trasplante de núcleos que acabamos de describir es, aproximadamente, el doble de eficiente que la microinyección y obtención de ovejas sólo transgénicas.

Las aplicaciones prácticas de los animales de granja clónico-transgénicos han generado un gran abanico de posibilidades comerciales respecto a la calidad sistemática y homogeneidad de productos derivados de estos animales, con implicaciones en ganadería, farmacología y biomedicina. Las denominadas «farmogranjas» son granjas dedicadas a la producción de fármacos a partir de la leche de animales domésticos como ovejas, cabras y vacas clónico-transgénicas. Un inconveniente de los rebaños clónicos pudiera ser, no obstante, una susceptibilidad generalizada a una infección determinada. Parece razonable que los ganaderos mantengan varios grupos de ovejas o vacas procedentes de clones independientes en su rebaño, para poder aprovecharse de la uniformidad clónica sin sufrir sus riesgos.

Cerdos clónico-transgénicos para xenotrasplantes

Órganos humanos para trasplantes de riñón y corazón escasean, por lo que muchos pacientes se mueren literalmente esperando un donante apropiado. Los xenotrasplantes son la sustitución de un órgano humano (dañado) por el equivalente procedente de un animal. A pesar de la enorme barrera inmunológica que nos separa de los cerdos, son éstos animales bastante más indicados que los primates para xenotrasplantes por razones prácticas, éticas y de seguridad. Uno de los problemas inmediatos con los xenotrasplantes de órganos es el rechazo inmunológico y vascular agudo provocado por los xenoantígenos, que son en su mayoría carbohidratos asociados a proteínas

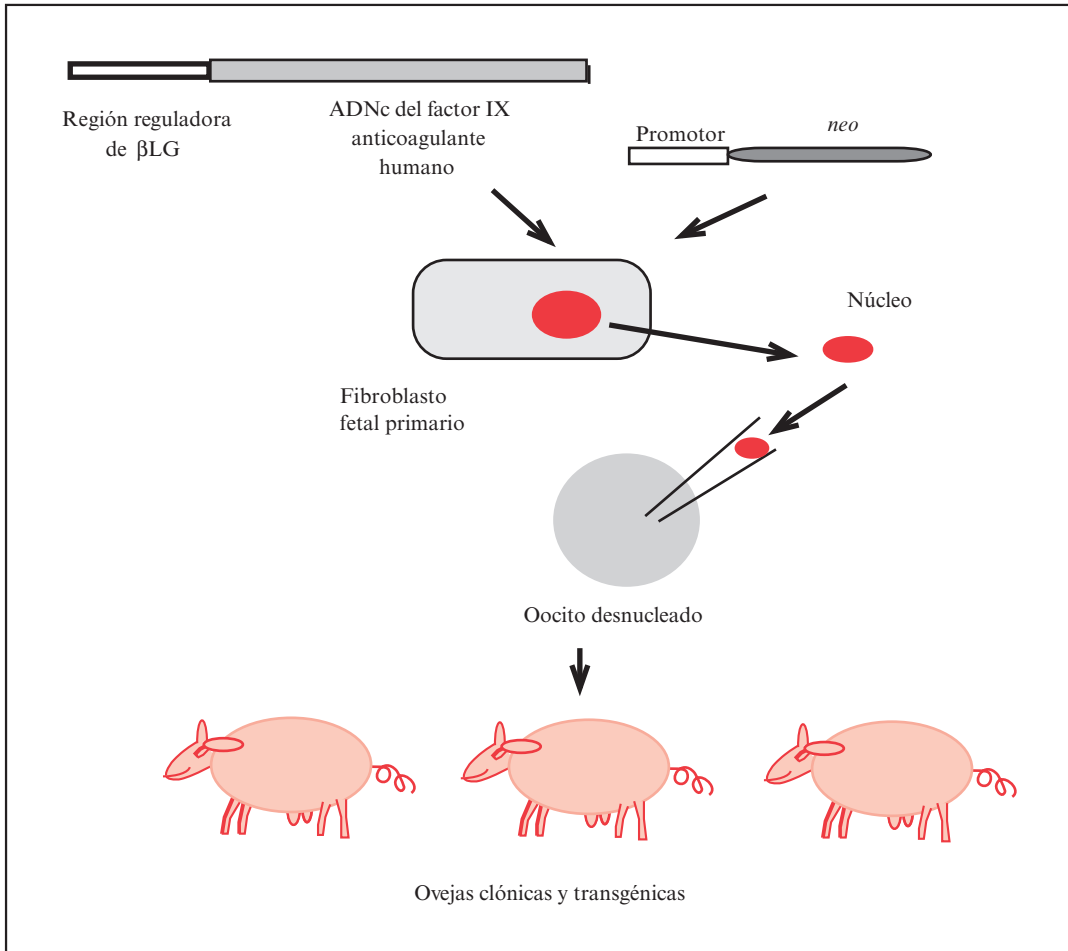


Figura 8.12. Obtención de ovejas clónico-transgénicas para el gen del factor IX anticoagulante humano. Fibroblastos fetales primarios se cotransfectan con el gen *neo*, que confiere resistencia a G418, y con el ADNc del factor IX anticoagulante humano fusionado a la región promotora de la β -lactoglobulina (β LG), para su expresión exclusiva en la glándula mamaria del animal. El núcleo de una de estas células se transfiere a un oocito desnucleado que se inserta en una oveja nodriza. De tres ovejas nodrizas nacieron tres ovejas clónicas portadoras de los dos transgenes *neo* y *factor IX*.

(mayoritariamente α 1, 3 galactosa) que se expresan en las células endoteliales que revisten los vasos sanguíneos porcinos. Era por tanto necesaria la obtención de cerdos transgénicos KO para el gen *GGTA1* que codifica la enzima α 1, 3-galactosil-transferasa (α 1,3GT) responsable de la síntesis del antígeno. Dado que la obtención de animales clónico-transgénicos es más eficiente que la de los simplemente transgénicos, se eligió esta metodo-

logía para la obtención de cerdos KO para las dos copias del gen *GGTA1* (figura 8.13). Se usó una construcción del transgén consistente en una parte del intrón 8 y otra del exón 9 del gen *GGTA1* que aportaron la homología necesaria para provocar una diana génica; además, la construcción llevaba un gen de selección *neo* precedido de un IRES que permitió la entrada de ribosomas y evitó la presencia de un promotor adicional; un gen

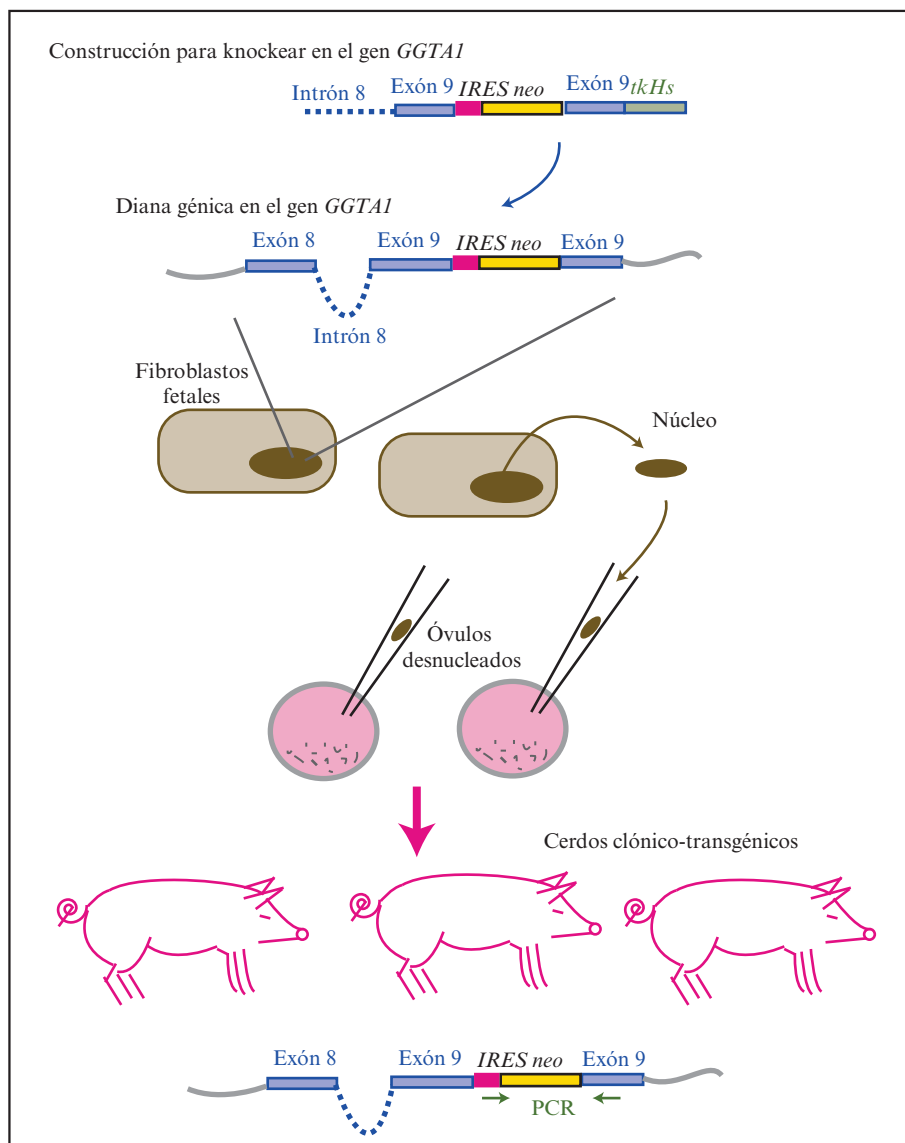


Figura 8.13. Obtención de cerdos clónico-transgénicos para xenotrasplantes de órganos a humanos. Fibroblastos fetales primarios se transfectan con una construcción que sustituye el exón 9 del gen *GGTA1* por otro que lleva el gen de selección *neo* que confiere resistencia a G418. Núcleos de estas células se transfieren a oocitos desnucleados que se insertan en cerdas nodrizas. Los clónico-transgénicos se verificaron por PCR; las flechas horizontales muestran los oligonucleótidos empleados.

de selección negativa como el gen *tkHs* en uno de los extremos del transgén y fuera de la región de homología. Este transgén se introdujo por electro-

poración en fibroblastos fetales (día 33 de la gestación) y se seleccionó por G418 (selección positiva) y ganciclovir (selección negativa). Núcleos de

las células sobrevivientes se inyectaron en óvulos no fecundados maduros desnucleados y se introdujeron en cerdas nodrizas. En este experimento, de 150 núcleos trasplantados se obtuvieron 5 transgénicos clónicos, verificados por PCR. El cruce entre dos hermanos de la G0 produjo homocigóticos KO para la enzima. Sin embargo, hay otros factores que también son responsables del rechazo agudo como la activación del complemento. Estos clónico-transgénicos necesitarían adquirir nuevos transgenes que inhiban la activación del complemento, conjunto de proteínas del suero y de la superficie celular que, cuando se activan, desencadenan una cascada secuencial de proteólisis que amplifica la respuesta inmune y provoca el rechazo del trasplante.

Otro problema derivado de los xenotrasplantes es el riesgo de transmisión de retrovirus endógenos porcinos. Se calcula que el genoma del cerdo posee unos 50 provirus defectivos integrados establemente y que se transmiten de forma mendeliana. Existe el riesgo real de que un paciente-receptor pueda ser infectado, bien por el retrovirus porcino o por un recombinante entre el provirus porcino y secuencias retrovirales humanas. El virus resultante podría extenderse por la población humana en un período de tiempo indefinido. Esta posibilidad parece remota, ya que ha habido 160 personas que en distintos ensayos clínicos han tenido contacto repetido con células vivas de cerdo y ninguna se ha infectado. Sin embargo, la cautela está justificada, ya que el propio virus del SIDA probablemente haya llegado al hombre a través del chimpancé en un único suceso inter-especies y han bastado veinte años para una amplia propagación en la especie humana. El seguimiento médico de los pacientes xenotrasplantados puede ser el único método preventivo aceptable. Así, los xenotrasplantes basados en cerdos transgénicos nos ofrecen en la actualidad ciertas promesas a costa de inciertos riesgos que no se pueden desestimar.

8.5. CLONACIÓN TERAPÉUTICA

La clonación terapéutica consiste en utilizar un núcleo o una célula totalmente diferenciada de

un paciente, como, por ejemplo, de la piel (fibroblastos de la dermis), para insertarlo o fusionarla con el óvulo de una donante en el que el núcleo está irradiado o extraído con una micropipeta. La activación de esta célula huevo quimérica, por electroporación por ejemplo, inicia el desarrollo embrionario y cuando se llega a la fase de blastocisto, se obtienen las células madre (también denominadas troncales totipotentes, troncales embrionarias o embrioblastos). Estas células madre embrionarias son capaces de generar cualquier tejido del cuerpo *in vitro* en las condiciones adecuadas (figura 8.14). Han sido detalles importantes para el éxito del experimento que la célula donante del núcleo estuviese en G0/G1; que el óvulo recipiente hubiese sido extraído recientemente de una donante joven; que se extrajera el núcleo del óvulo en metafase; que se fusionara la célula donante del núcleo con el óvulo, en vez de extraer solo el núcleo, y que se utilizara electroporación (descargas eléctricas) para activar el huevo tras la fusión. Una vez realizadas todas estas fases, el huevo inició su desarrollo que culminó con la formación de un blastocisto a los 5 días. Alternativamente se podría haber implantado el huevo en una hembra nodriza que llevara a término el embarazo y produjera un clon humano del donante del núcleo, algo que no está contemplado en la actualidad por su estricta prohibición. En el experimento real que estamos describiendo, las células madre cultivadas se introdujeron en ratones inmunodeficientes (SCID) donde produjeron tumores en los que se pudieron distinguir células pertenecientes a las 3 capas embrionarias: 1) endodermo, representado por células del epitelio intestinal, 2) mesodermo, representado por células musculares de tipo cardiomiocitos contráctiles y 3) ectodermo, con células neuroectodérmicas.

En fecundaciones *in vitro* también se han obtenido cultivos de células madre a partir de una sola célula extraída de la mórula de 8 células y antes de su implantación en el útero (ver figura 4.31). Este cultivo mantenido en nitrógeno líquido representa una reserva de células madre para su posterior uso, en caso de necesidad.

En la actualidad el paso más problemático es el de la diferenciación a células de los distintos

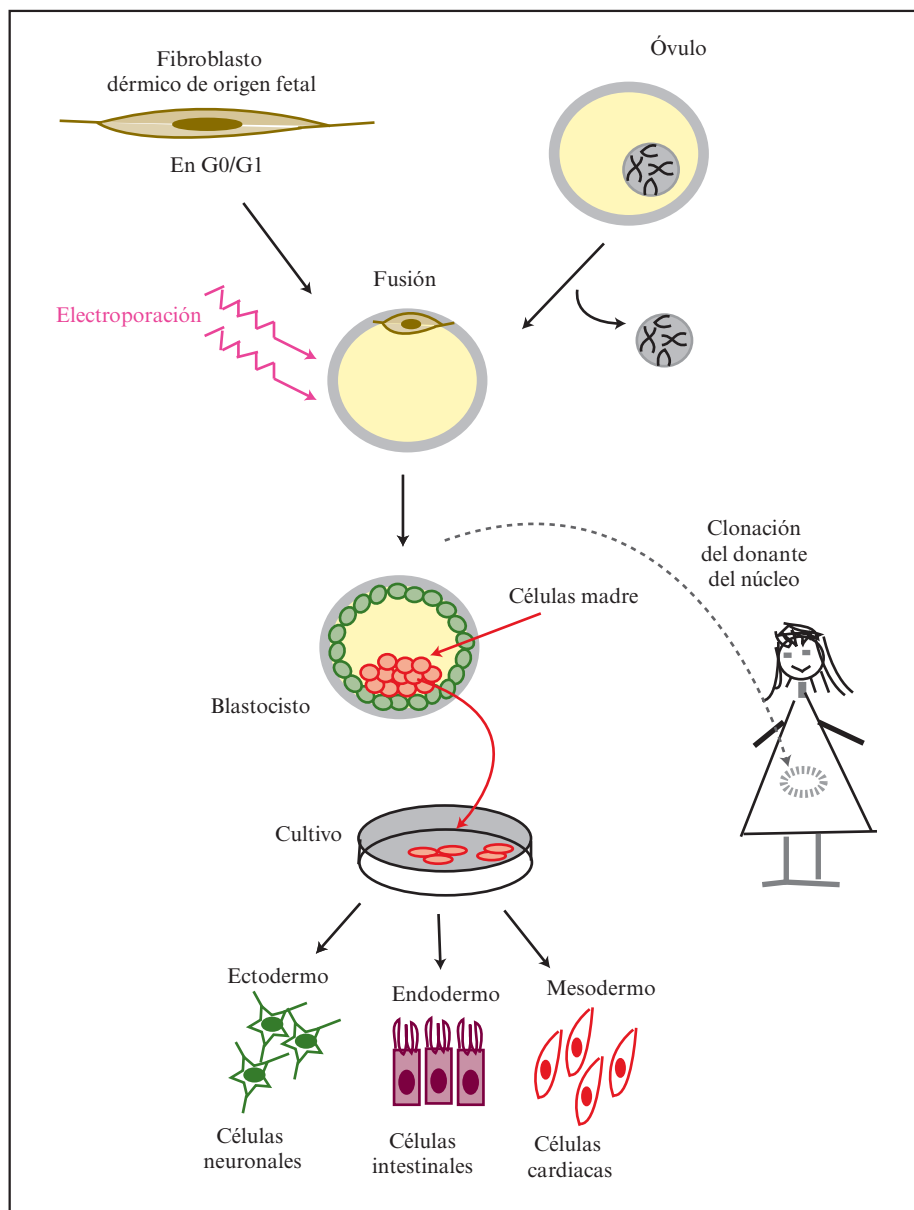


Figura 8.14. Clonación terapéutica en humanos. Una célula de la piel humana (fibroblasto de la dermis de origen fetal) se fusionó con un óvulo de una donante al que se le extrajo el núcleo en metafase con una micropipeta por microinyección. Tras la fusión se estimuló el óvulo por electroporación y se inició el desarrollo embrionario. El desarrollo se detuvo en la fase de blastocisto para la extracción de las células madre. Las células madre cultivadas se introdujeron en ratones inmunodeficientes (SCID) donde produjeron tumores en los que se pudo distinguir células pertenecientes a las 3 capas embrionarias: endodermo (células epiteliales intestinales), mesodermo (cardiomiocitos contráctiles) y ectodermo (células neuroectodérmicas). También se muestra la posibilidad hipotética de implantar la célula huevo en una hembra nodriza donde se desarrollaría un clon humano del donante del núcleo.

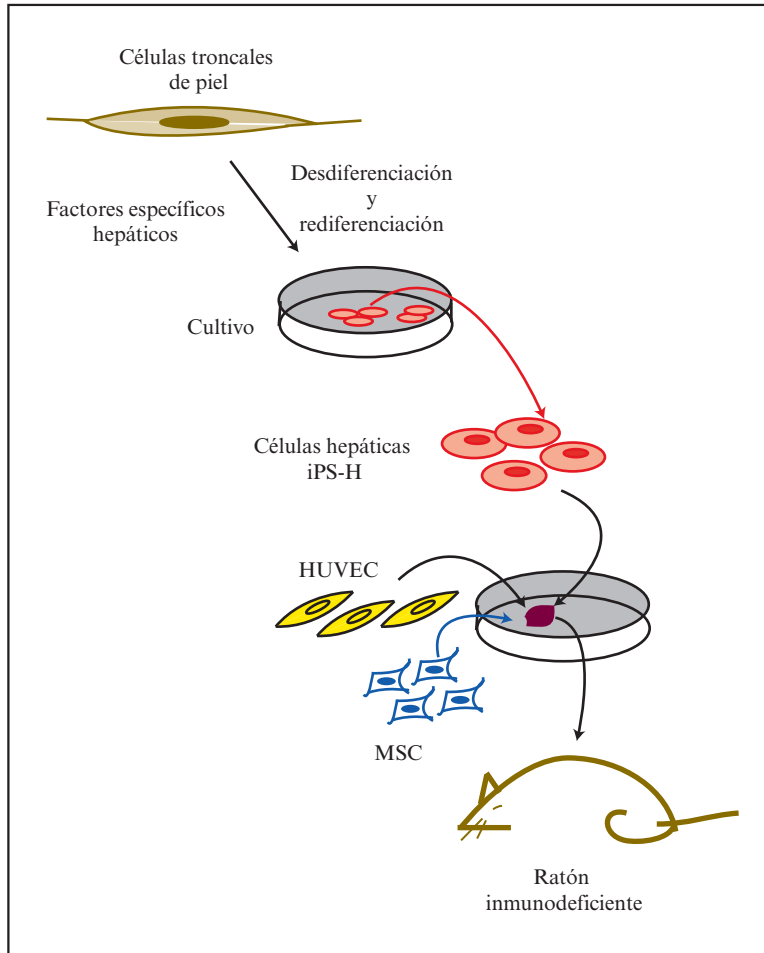


Figura 8.15. Creación de un hígado artificial a partir de células madre iPS de piel. Reprogramación de fibroblastos humanos a células hepáticas iPS (iPS-H) por desdiferenciación con factores de transcripción y diferenciación dirigida. Las iPS-H se cultivaron con células endoteliales de cordón umbilical humanas (HUVEC) y células troncales mesenquimales humanas (MSC). Los primordios de hígado humano resultantes se trasplantaron ectópicamente a ratones inmunodeficientes.

tejidos, a partir de las células madre. Es preciso añadir a las células indiferenciadas distintos factores de crecimiento y diferenciación, que han de ser escrupulosamente distintos y en concentraciones exquisitamente específicas para obtener primordios de órganos concretos y alejar la posibilidad de inducir un tumor.

La medicina regenerativa se perfila como la gran beneficiaria de la clonación terapéutica, esto

es, la fabricación de órganos y tejidos que puedan trasplantarse a pacientes y puedan curar una amplia gama de enfermedades. Utilizando células madre embrionarias humanas, se han llegado a crear *in vitro* minicerebros del tamaño de un guisante que espontáneamente desarrollan las estructuras y capas neuronales del córtex, mimetizando fielmente el desarrollo normal *in vivo*. Ir más lejos en el tubo de ensayo precisaría de células endote-

liales de cordón umbilical que formaran un sistema circulatorio rudimentario y aportaran nutrientes y oxígeno a todas las células del órgano incipiente. En el momento actual, la utilización de células cardíacas o células neuronales a partir de células madre, contra infartos o enfermedades neurodegenerativas, respectivamente, puede que esté a la vuelta de la esquina.

Células troncales pluripotentes inducidas (iPS)

La alternativa a la clonación terapéutica son las células troncales pluripotentes inducidas (iPS), que se desarrollaron a partir de los estudios realizados en células madre embrionarias, pero que no precisan de la implicación de un embrión, ya que se basan en la desdiferenciación de células totalmente diferenciadas. El Premio Nobel de Fisiología y Medicina de 2012, Shinya Yamanaka, ha sido el gran impulsor de estas células cuyo único problema es evitar que al desdiferenciarlas se conviertan en potencialmente tumorales.

Entre los resultados espectaculares que llenan los periódicos actuales, nos encontramos con la fabricación de retinas sintéticas para tratar cegueras o la creación de primordios de hígado a partir de células de piel. Un minihígado en tres dimensiones y vascularizado fue capaz de realizar funciones he-

páticas al trasplantarlo ectópicamente a un ratón inmunodeficiente (figura 8.15). La metodología consistió en reprogramar fibroblastos humanos a células hepáticas iPS por desdiferenciación al tratar estas células con los 4 factores de transcripción: OCT4, NANOG, SOX2, y LIN28 para reprogramarlas a iPS-hepáticas (iPS-H) con nuevos factores de diferenciación. Las iPS-H se cultivaron con células endoteliales de cordón umbilical humanas (HUVEC) y células troncales mesenquimales humanas (MSC) y al cabo de unos días (4-6 días) se observó una reorganización de todas estas células en un primordio de hígado en tres dimensiones. De momento, estos primordios de hígado se están utilizando para examinar la toxicidad de nuevos fármacos, ya que es el hígado el órgano encargado de metabolizar las sustancias extrañas al organismo.

La manipulación de células iPS, junto con las células madre embrionarias, son la gran promesa de la medicina regenerativa; sin embargo, los ensayos clínicos realizados hasta el momento trasplantando estas células a pacientes no han dado aún resultados claramente satisfactorios. El potencial terapéutico de estas nuevas tecnologías está aún por conocer; los próximos años serán testigo de hasta dónde se puede llegar en su aplicación clínica.

LECTURAS COMPLEMENTARIAS

Human embryonic stem cells derived by somatic cell nuclear transfer. **Tachibana, M. et al.** *Cell*, **153**: 1228-1238. 2013.

Refined human artificial chromosome vectors for gene therapy and animal transgenesis. **Kazuki, Y. et al.** *Gene Therapy*, **18**: 384-393. 2011.

Vascularized and functional human liver from an iPSC-derived organ bud transplant. **Takebe, T. et al.** *Nature*, **499**: 481-485. 2013.

Virus patógenos. Coordinadores **Luis Carrasco** y **José M.ª Almendral del Río**. Fundación BBVA. Editorial Hélice. España. 2006.

Los vegetales son uno de los grandes componentes nutricionales del hombre, y en algunos casos, el único. Las plantas representan un enorme potencial para aplicaciones de la ingeniería genética, con vertientes tan dispares como la mejora en la producción de frutos, semillas y vegetales, la creación de nuevas variedades de flores y plantas ornamentales o la producción de tejidos, vacunas y biocombustibles.

En general, las plantas tienen un poder regenerativo muy superior a los animales. En muchos casos, una sola célula somática puede regenerar la planta completa. En principio, la manipulación genética de una sola célula vegetal es suficiente para la rápida producción de una nueva variedad.

En cuanto a la expresión, las plantas pueden expresar genes de mamíferos abriendo grandes posibilidades para su utilización como biorreactores para producir proteínas, carbohidratos o lípidos de origen animal, incluyendo hasta vacunas que pudieran ser ingeridas en nuestra dieta. Existen, no obstante, mecanismos que silencian con gran frecuencia y eficiencia los transgenes y que pudieran representar una defensa por parte de la planta contra la invasión por ADN foráneo. El ARN de interferencia se descubrió en petunias cuando, al tratar de incrementar el color rojo de sus flores con el transgén responsable de la enzima dihidroflavonol-4-reductasa, se observó una supresión total o parcial de la coloración de las flores.

El hombre ha manipulado durante miles de años el contenido genético de plantas con interés

agrícola, por sucesivos cruces y selección de semillas. Lo que ahora puede alterarse es la velocidad de producción de esos cambios al aplicar nuevas tecnologías que permiten la formación de genomas híbridos de orígenes muy dispares. En estos últimos años las plantas transgénicas han pasado a formar parte del consumo humano. Trigo o maíz resistentes a herbicidas, tomates que maduran a voluntad o arroz resistente a patógenos son algunos de los productos comercializados en muchos países, incluido el nuestro.

En investigación básica, la manipulación genética y la secuenciación han permitido aislar y caracterizar un gran número de genes de origen vegetal, estudiar su expresión, su regulación y sus peculiaridades. *Arabidopsis thaliana* sigue siendo la planta modelo de referencia obligada. Su genoma (7×10^7 pb, 25.498 genes) está secuenciado y saturado con mutaciones, se pueden obtener dianas génicas por recombinación homóloga con frecuencias de 10^{-3} y algunos de sus genes presentan gran movilidad. Esta información está permitiendo a muchos científicos profundizar en el conocimiento del crecimiento, desarrollo, expresión génica y regulación de plantas superiores. Todo esto, unido a la estimulante habilidad que están demostrando los vegetales de sintetizar antibióticos, inmunoglobulinas o plásticos biodegradables, convierten la transferencia génica a plantas en una de las áreas más interesantes de la biología molecular y de la ingeniería genética actuales.

9.1. VECTORES

Veremos dos tipos de vectores en este apartado:

- Plásmidos: plásmidos Ti y Ri de *Agrobacterium*.
- Virus: virus del mosaico de la coliflor, virus del mosaico del tabaco, virus X de la patata, entre otros.

9.1.1. Plásmidos

El plásmido Ti de *Agrobacterium tumefaciens*

Fue en este sistema vegetal en el que se observó que la ingeniería genética, de hecho, ocurre en la naturaleza. Entre la bacteria *Agrobacterium tumefaciens* y muchas plantas dicotiledóneas hay una simbiosis que permite algo muy semejante a lo que *in vitro* catalogamos como ADN recombinante: un plásmido de la bacteria lleva algunos genes que originalmente han sido de la planta. Este vector natural ha sido modificado y manipulado y, como en el caso del círculo de 2 μ m de levaduras, sigue siendo uno de los vectores más usados para transferir genes heterólogos al genoma de las plantas. La transferencia génica basada en el plásmido Ti es particularmente idónea para plantas dicotiledóneas, que incluyen las legumbres (garbanzos, lentejas, guisantes, habas, etc.), además de frutas (peras, manzanas, melocotones, etc.), cacahuetes, almendras, avellanas, etc., además de *Arabidopsis thaliana* y *Nicotiana tabacum*. Sin embargo, su uso en monocotiledóneas es más problemático; así, vegetales de gran interés agrícola como trigo, arroz, maíz y centeno no se transforman fácilmente con el plásmido Ti. El problema parece radicar en la dificultad para cultivar células de monocotiledóneas en general, y detectar células transformadas en particular, y no en una incompatibilidad biológica más profunda.

Agrobacterium es una bacteria Gram (–), es decir, que no coge el color azul de la tinción de Gram, porque la pared celular es fina y no retiene

el colorante; son también Gram (–): *E. coli*, *Salmonella*, *Haemophilus influenza*, entre otras muchas. Esta bacteria se encuentra en el suelo y puede infectar más de 330 géneros de dicotiledóneas. La infección se realiza a través de una herida o incisión. Una pequeña parte de la población de *Agrobacterium* que vive en el suelo alberga un plásmido que induce la formación de un tumor calloso denominado tumor de la corona. El plásmido se denomina Ti (inductor de tumores). Este plásmido es de gran tamaño: 150-200 kb, de las cuales 20-23 kb pueden escindirse del plásmido y pasar de la bacteria a la planta, donde se integran en uno o varios lugares del genoma, aparentemente al azar, constituyendo el ADN T (transferido).

El ADN T (figura 9.1) induce la síntesis de unos compuestos raros nitrogenados denomina-

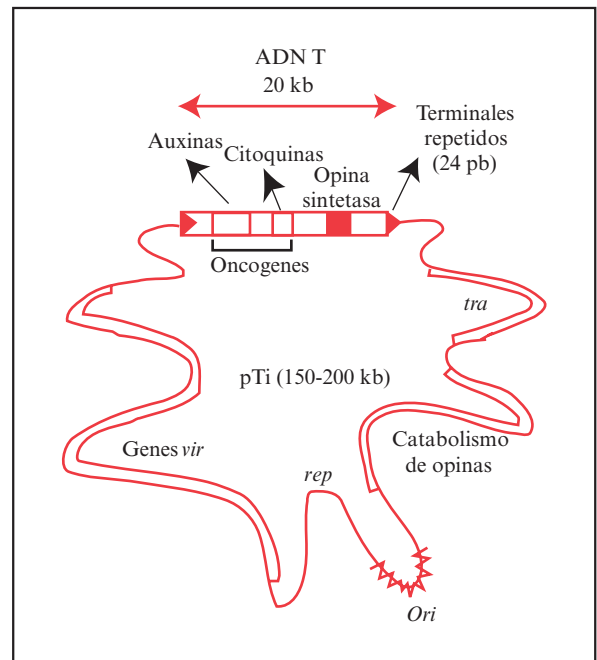


Figura 9.1. Mapa del plásmido natural Ti de *Agrobacterium*. Se muestra esquemáticamente el ADN T y los genes *vir* implicados en la ruptura del ADN T y activación de otros genes adyacentes; el origen de replicación (*ori*), los genes implicados en el catabolismo de la octopina, por ejemplo, y los *tra* responsables de la transferencia por conjugación.

dos «opinas». Esto se lleva a cabo mediante la enzima «opina sintetasa» codificada por el propio ADN T y utilizando aminoácidos básicos, como arginina, ornitina, lisina o histidina como sustrato. El tumor de la corona que se forma está constituido por células transformadas, inmortales, que crecen en ausencia de hormonas vegetales (auxinas, citoquinas, giberelinas y ácido abscísico).

Las bacterias que poseen plásmidos Ti pueden metabolizar opinas como única fuente de carbono y nitrógeno, debido a que los genes de las enzimas necesarias para su catabolismo están en el plásmido. Además, las opinas estimulan la transferencia por conjugación, con lo que esta propiedad puede extenderse a toda la población de bacterias. Son genes esenciales del plásmido los necesarios para la replicación, *ori* y *rep*, así como los implicados en la transferencia por conjugación *tra*, que afectan al contacto célula-célula, y los *vir*, que intervienen en la ruptura y transporte del ADN transferido (figura 9.1). El ADN que se transfiere al genoma de la planta (ADN T) lleva los genes necesarios para la síntesis de una opina y los genes que le confieren las propiedades oncogénicas a la planta. Hay al menos seis clases diferentes de opinas, y algunos tumores sintetizan más de una. Los plásmidos más estudiados son los Ti que llevan el gen correspondiente a la enzima nopalina sintetasa *nos* o el de la octopina sintetasa *ocs*. Los promotores de estos genes son eucarióticos y también lo son las señales de poliadenilación que se encuentran en sus extremos 3'; esto hace pensar que probablemente tengan un origen eucariótico y que hayan quedado atrapados en el plásmido. Los extremos del ADN T son dos repeticiones directas imperfectas de 24 nucleótidos. El corte en el extremo derecho del ADN T es preciso y constante, pero el izquierdo puede variar en unos 100 nucleótidos, por lo que puede o no incluir el terminal repetido de este lado. El ADN T también incluye genes de enzimas implicadas en la biosíntesis de fitohormonas que desestabilizan el equilibrio hormonal de la planta infectada e inducen la proliferación de tejido indiferenciado en la zona de la herida.

Fuera del ADN T está la región de los genes *vir* (virulencia), responsables de la activación de

genes adyacentes. También algunos de los genes *vir* son responsables de la ruptura de la región T y de su migración e integración en el genoma de la planta. El proceso de transferencia se inicia con la captación, por parte de la bacteria, de compuestos fenólicos específicos exudados por las células de la región cercana a la herida de la planta; compuestos que son detectados por el producto del gen *vir A*, que constituye el sistema sensorial bacteriano. Reconocida la señal, Vir A fosforila al producto del gen *vir G* (Vir G) produciéndose la inducción de otros operones *vir* encargados, como ya dijimos, de la escisión y movilización del ADN T. Como respuesta a la señal de activación, los productos de los genes *vir D1/D2* identifican los 24 pares de bases en el borde interno del ADN T y producen dos incisiones en la misma hebra del ADN del plásmido Ti. Se sintetiza una copia de la región T, siendo este ADN de banda sencilla el que se transporta a la célula vegetal. El extremo 5' repetido (derecho) se une covalentemente a la proteína VirD2. Este ADN también está asociado no covalentemente con la proteína de unión a ADN de hebra sencilla (que no es específica de secuencia) producto del gen *virE2*. VirE2, además, mantiene el ADN de forma extendida con un diámetro muy pequeño (2 nm), lo que va a facilitar el paso a través de las numerosas membranas y espacios intercelulares antes de llegar al núcleo de la célula vegetal. El primer obstáculo de los múltiples existentes son las membranas interna y externa, así como la pared celular, bacterianas; aquí parecen intervenir las proteínas VirB formando una especie de canal de paso o complejo de transporte (figura 9.2). La segunda barrera es la pared celular y la membrana de la célula vegetal y, por último, la membrana nuclear que tiene poros. La superación tanto de este último paso como del anterior no se conoce bien. Ya en el núcleo, el ADN T se integra al azar en el genoma de la célula vegetal. En algún momento de la transferencia, el ADN de hebra sencilla se convierte en ADN de doble hebra, y si el proceso es parejo a lo que ocurre en la conjugación bacteriana con la que tiene una gran similitud, esta conversión ocurre al llegar a la célula receptora.

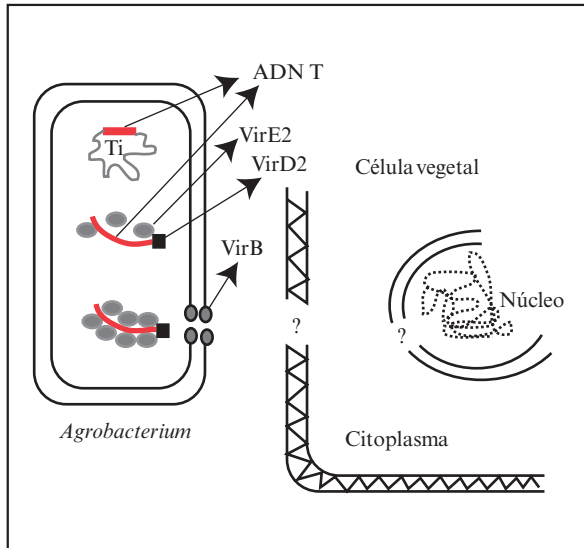


Figura 9.2. Proceso de transferencia en pasos sucesivos del ADN T de la bacteria al núcleo de la célula vegetal.

Como en otras ocasiones, este vector natural ha sido ampliamente modificado *in vitro*. Los genes oncogénicos del ADN T se reemplazaron por los de pBR. Ello permite, por una parte, poder regenerar plantas normales y, por otra, crecer el recombinante en *E. coli*. La versión no oncogénica del plásmido se denomina «desarmado». En un principio se mantuvieron los genes de la síntesis de las opinas porque podían usarse como método de selección (sólo las células transformadas pueden sintetizarlas), pero más tarde se eliminaron también dejando, algunas veces, sus promotores y sustituyendo la región codificante por un gen de selección como el gen bacteriano *neo* (neomicina-fosforribosil-transferasa) que confiere resistencia al antibiótico kanamicina, entre otros.

El número de copias del plásmido Ti que se insertan en la célula vegetal es variable. La integración puede ocurrir en un locus o en varios y en ocasiones la orientación de alguna unidad se invierte.

La propiedad de integración al azar observada en el ADN T puede ser utilizada para el aislamiento de genes de la planta. El ADN T produci-

rá mutaciones allí donde se integre, inactivando el producto o la función del gen invadido. Si esta falta de función puede ser reconocida fenotípicamente, podremos aislar el gen con sondas específicas para el ADN T; el gen deseado será colindante. Alternativamente, la utilización del ADN T como trampa de promotores se ha empleado para detectar genes estructurales activos en las especies *Nicotiana* y *Arabidopsis*. El método es el siguiente: un gen señal, esto es, aquel cuyo producto es fácilmente detectable histológicamente: β -glucuronidasa o proteína verde por ejemplo, se incorpora sin promotor en el ADN T. Los tejidos que expresen β -glucuronidasa (GUS) o proteína verde habrán insertado el ADN T vector en una región próxima a un promotor celular.

El mayor inconveniente de los plásmidos Ti es su gran tamaño. La inserción de genes heterólogos se puede realizar de modo indirecto por cointegración como se ha explicado para los baculovirus (véase el capítulo 6, apartado 6.6 y figura 6.9), pero en la mayoría de los casos se utiliza el denominado sistema binario.

En el **sistema binario** se emplean dos vectores. Uno de ellos es un plásmido Ti desarmado que aportará las funciones *vir* indispensables para la transferencia e integración del ADN T en el genoma de la planta. El otro vector de este sistema binario, denominado mini T, aporta la región de ADN T; así en la región donde originalmente se encontraba el gen de la opina sintetasa, se ha insertado el gen señal: *lac* o *gus*, y una región de policlonaje situada en su interior. De este modo será posible una selección por color de los plásmidos que hayan adquirido el inserto (figura 9.3). Ambos vectores poseen un origen de replicación, por lo que la cointegración no es necesaria. El mini T puede introducirse en *Agrobacterium*, bien por transformación (electroporación) o por conjugación con la ayuda de un plásmido auxiliar que facilite los genes *mob* y *tra*. También se pueden suplir las funciones del operón *vir* del ADN T por las del *mob* de bacterias, y las de la secuencia líder del extremo repetido derecho por las del origen de conjugación *ori* de los plásmidos de fertilidad F de *E. coli*. Hay, sin embargo, otras funciones de *A. tumefaciens* implicadas en el contacto

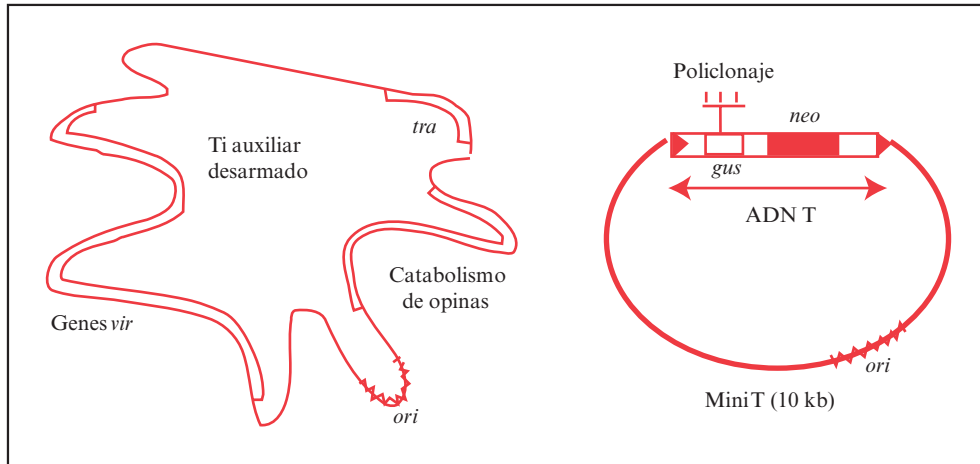


Figura 9.3. Sistema binario de transferencia de genes heterólogos a plantas. El mini T es un plásmido bacteriano con algo de ADN T, el gen heterólogo insertado en la región de policlonaje y los genes *gus* y *neo* para seleccionar por el inserto. El Ti auxiliar carece de ADN T (desarmado), pero provee todas las funciones para la transferencia de ADN T. Cuando ambos están en *Agrobacterium*, el Ti auxiliar facilita la transmisión del ADN T del mini T.

célula a célula que hasta el momento son indispensables para que se lleve a cabo la transferencia.

El plásmido Ri de *Agrobacterium rhizogenes*

Cuando *Agrobacterium rhizogenes* infecta plantas, en vez de inducir tumores induce un gran número de raíces adventicias que forman como una cabellera (enfermedad de las raíces peludas). Este fenómeno está mediado por el plásmido Ri (inductor de raíces) que es muy semejante al Ti. Al inducir raíces en vez de masas indiferenciadas de células, no es preciso eliminar los genes oncogénicos y se pueden obtener plantas adultas con los genes heterólogos insertados. La planta retiene su dependencia de fitohormonas y es fértil.

El plásmido Ri es más grande (250 kb) y está peor caracterizado que el Ti, pero las regiones que contienen el ADN T y los genes *vir* están identificadas. En el ADN T se han localizado opinas y genes que codifican β -glucosidasas y otras enzimas que hidrolizan enzimas, por lo que se piensa que estos genes convierten las formas inactivas de algunas hormonas en activas y así cambian la fisiología de la planta.

9.1.2. Virus como vectores

La utilización de virus de plantas como vectores es de gran interés, pero desafortunadamente aún no se ha encontrado un virus vegetal de fácil conversión en el vector ideal.

Los virus ofrecen numerosas ventajas en la expresión transitoria de genes de interés. Son enormemente eficientes en la multiplicación, originando un gran número de copias del transgén así como en su expresión, que alcanza valores máximos entre una y dos semanas después de la inoculación. Nos ofrecen la posibilidad de inducir su expresión en una fase precisa del desarrollo de la planta, eligiendo el momento que se considere más oportuno para la inoculación del vector viral. Los virus, además, pueden extenderse rápidamente a otras partes de la planta, a otras plantas y a otras especies.

Más del 80% de los virus en el reino vegetal tienen como genoma ARN de una sola hebra con polaridad de mensajero (virus ARN+). No se han podido identificar en plantas ni virus ADN de hebra doble que sólo se reproduzcan vía intermediarios de ADN, ni retrovirus que se integren en el genoma.

Describiremos primeramente el virus del mosaico de la coliflor (CaMV, del inglés *Cauliflower mosaic virus*), que tiene por genoma ADN de hebra doble, y los geminivirus que tienen por genoma ADN de hebra sencilla. Seguidamente veremos varios ejemplos de virus ARN+, como el virus del mosaico del tabaco y el virus X de la patata, entre otros.

Virus que poseen ADN por genoma

El virus del mosaico de la coliflor (caulimovirus)

Se usa como vector de reemplazamiento; se consideró inicialmente ideal por ser de pequeño tamaño (8 kb) y tener ADN de doble hebra que puede ser manipulado más fácilmente *in vitro*. Es muy infeccioso, pero la organización de su genoma es compleja y compacta. Sólo se han identificado dos genes como no-esenciales y éstos no son grandes; además, la cápsida es de pequeño tamaño y no admite ADN extra; recombinaciones moleculares del propio genoma eliminan rápidamente insertos foráneos que pudieran afectar su estabilidad. La replicación también es compleja y requiere un paso a través de ARN, lo cual puede introducir errores puesto que la transcriptasa inversa carece de actividad correctora. Por último, los caulimovirus tampoco tienen un amplio abanico de hospedadores, limitándose a unas pocas especies. Un aspecto positivo es que el promotor del ARN genómico, de tamaño 35S, es muy eficiente en una gran variedad de células vegetales, utilizándose para la expresión de muchos transgenes u otros sistemas virales.

Los geminivirus

Son atractivos porque algunos miembros del grupo infectan monocotiledóneas además de dicotiledóneas. En la naturaleza producen infecciones serias.

Su genoma es ADN de una sola hebra y la replicación ocurre en el núcleo por el método del círculo rodante a través de un intermediario de hebra doble que permanece infeccioso al clonarlo. No hay integración en el genoma de la célula infectada.

Los geminivirus tienen genomas muy pequeños, lo que implica una severa limitación en el tamaño de ADN que puede ser sustituido o insertado. Además, la propagación de muchos geminivirus está restringida al floema y células adyacentes, lo que constituye otra importante limitación de los vectores de expresión basados en esta familia de virus.

Hay varios géneros de geminivirus, entre ellos los más interesantes como vectores son: los begomovirus, y los mastrevirus.

Los begomovirus infectan dicotiledóneas y están transmitidos por las moscas blancas (*Bemisia tabaci* entre ellas). Pertenecen a este grupo: 1) el virus del mosaico de la *Cassava africana*, y 2) el virus del mosaico dorado del tomate. Sus genomas están bipartidos en dos piezas circulares de 2,7 kb cada una, denominadas A y B. El genoma A es suficiente para la replicación autónoma y el ensamblaje de las partículas maduras, mientras que el B contiene las proteínas del movimiento que facilitan el traslado de los virus del núcleo al citoplasma y de una célula a otra (figura 9.4).

La proteína de la cubierta (o cápsida) del virus del mosaico de la *Cassava africana* puede ser sustituida por un gen de interés que, de modo transitorio, se expresará en grandes cantidades, a la vez que el vector se divide y expande por las plantas infectadas. Sin embargo, y a pesar de que en estas circunstancias el ADN no se encapsula, parecen existir estructuras secundarias tales que provocan delecciones del inserto si éste es mínimamente superior en tamaño al gen de la cubierta. Este fenómeno se ha verificado con el gen *gus* (β -glucuronidasa) que es ligeramente mayor que la proteína de la cápsida; la sustitución no ha provocado problemas de replicación pero sí de estabilidad, induciéndose delecciones hasta el tamaño del gen autóctono.

Los mastrevirus son geminivirus de una sola pieza de 2,68 kb (forma A, figura 9.4). Pertenecen a este grupo el virus de la hoja estriada del maíz que infecta monocotiledóneas. Este último virus se transmite a la planta *in vivo* solamente a través de cicadelas (chinchas de campo), produciendo unas estriaciones amarillentas en las hojas del maíz. En estos virus la proteína de la cubierta se

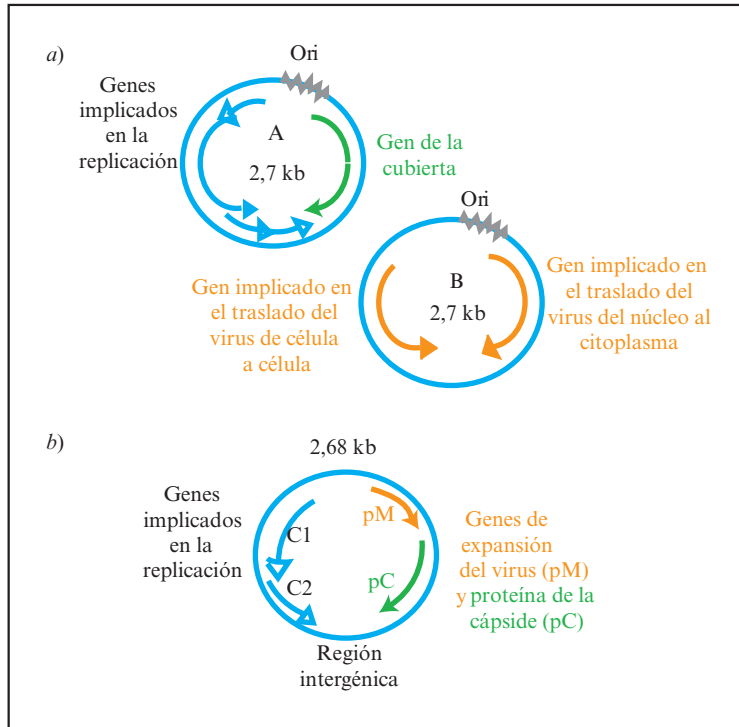


Figura 9.4. Representación esquemática de los genomas de dos geminivirus. a) Genoma bipartido del begomovirus del mosaico de la *Cassava africana*, y b) El mastrevirus de la hoja estriada de maíz.

necesita para el movimiento y la expansión viral, imposibilitando su reemplazamiento. Como alternativa se han utilizado estrategias de inserción génica en vez de sustitución. El pequeño gen *bar* (fosfinotricina acetiltransferasa), que confiere resistencia al herbicida Basta (fosfinotricina), se ha introducido en la región intergénica del virus. Las hojas inoculadas expresan el gen, pero se bloquea la expansión del mismo al resto de la planta hasta que el vector revierte a su tipo silvestre.

El dímero de la forma replicativa de hebra doble del virus se puede introducir en la planta por bombardeo con pistolas o agroinfecciones (más adelante en este capítulo). *Agrobacterium*, portando el virus como parte del ADN T, ha sido capaz de infectar heridas de hojas de maíz, lo cual abre la posibilidad de que pueda transferir ADN a gramináceas. Este tipo de infecciones también se han conseguido en petunias agroinfectando el bego-

movirus del mosaico dorado del tomate. Si bien existían grandes esperanzas de convertir este virus en un buen vector, ya que puede llegar a detectarse una sola célula infectada debido a la enorme amplificación que tiene lugar durante la replicación, los problemas de estabilidad han mermado notablemente el entusiasmo inicial.

Virus que poseen ARN+ por genoma

La manipulación de los virus ARN+ es compleja, hay que clonar copias de ADN de los genomas de ARN, manipularlas adecuadamente *in vitro* y convertirlas de nuevo en ARN infeccioso por transcripción *in vitro* (usando, por ejemplo, promotores de bacteriófagos) o *in vivo* (empleando promotores reconocidos por una ARN polimerasa de la planta) (figura 9.5).

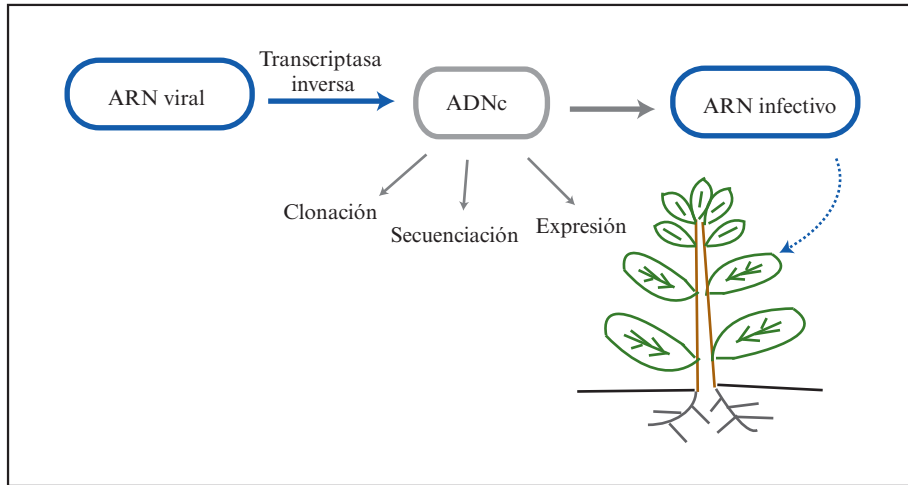


Figura 9.5. Representación esquemática de la manipulación asociada a los virus ARN.

El virus del mosaico del tabaco y otros virus con forma de varilla

Los virus con forma de varilla pueden superar las limitaciones de espacio que se imponen en los virus esféricos, siendo más flexibles y estables a la hora de expresar genes de interés.

El virus del mosaico del tabaco (TMV) es un virus ARN de hebra sencilla (+) en una sola pieza. El ARN actúa como ARNm y codifica cuatro polipéptidos. Produce dos ARN subgenómicos a partir de dos promotores diferentes. La proteína de la cápsida (o cubierta) es muy abundante. Si se coloca el gen bacteriano *cat* (cloranfenicol acetiltransferasa) justamente a continuación del codón de iniciación de la proteína de la cápsida y se inoculan hojas de plantas de tabaco con la construcción, se detecta actividad CAT, pero la infección no se extiende al resto de la planta. En algunos casos se ha observado que el gen introducido ha sido eliminado de modo preciso para readquirir el virus su forma original. En este virus, su propagación al resto de la planta depende de una proteína de la cápsida intacta y de la proteína de movimiento (pM) del virus. La pM tiene 268 aminoácidos y aunque es responsable de la extensión de la infección dentro de la hoja, no está claro si tam-

bién es necesaria para la propagación por la planta. El éxito mayor de los vectores basados en TMV han sido los que producen un tercer ARN subgenómico que se sintetiza a partir del promotor de la proteína de la cápsida del TMV. En estos casos, se sustituye el gen de la proteína de la cápsida del TMV, por la de otro TMV con su propio promotor.

Como método de infección se puede usar la **magnificación**, que consiste en hacer llegar el genoma viral como ADNc dentro del ADN T de un plásmido Ti e infectar la planta a través de *Agrobacterium tumefaciens* por infiltración al vacío (apartado 9.3 y figura 9.8) provocando una infección en todas las hojas maduras de la planta simultáneamente (transfección sistémica). En esta situación el virus se replica muy eficientemente y no necesita la proteína de movimiento o la cápsida para moverse de un sitio a otro de la planta. Con esta metodología y vectores modificados basados en el TMV, se han obtenido grandes cantidades de antígenos de agentes patógenos como el virus causante de la hepatitis B, la bacteria *Yersinia pestis*, causante de la peste, o el parásito protozoo *Plasmodium yoeli*, del mismo género que el causante de la malaria. Todos ellos podrían usarse en un futuro no muy lejano para

la preparación de vacunas contra estas enfermedades.

El virus del mosaico del caupí (CPMV, del inglés *cowpea mosaic virus*) es un buen vector de expresión en plantas; ha sido utilizado como vector de fusión para la expresión de péptidos en la superficie de la partícula viral. Se ha comprobado que la viabilidad y estabilidad de los virus recombinantes dependen fundamentalmente del tamaño y punto isoeléctrico del péptido expresado. También se ha usado como vector de inserción, clonando el gen de interés en el ARN 2 viral, entre las secuencias codificantes de la proteína de movimiento y de la cápsida. Hay también una versión modificada no replicante (y, por tanto, con un elevado nivel de seguridad biológica) del ARN 2 del CPMV (CPMV-HT), sin aparentes restricciones en el tamaño del inserto y con una eficacia de traducción muy buena. Además, se ha comprobado que los niveles de expresión del sistema CPMV-HT podían mejorarse co-expresando un supresor de silenciamiento de ARN modificado para no causar daños al desarrollo de la planta. Estos vectores modernos derivados del CPMV se están usando para la producción de grandes cantidades de una amplia variedad de productos incluido un anticuerpo contra el virus del SIDA.

El virus X de la patata es otro posible buen vector en forma de varilla. La proteína de la cápsida es esencial. Los vectores tienen duplicado el promotor de esta proteína de la cápsida, de manera que con el segundo pueden sintetizar un producto de interés. El factor limitante suele ser la estabilidad del inserto, habiéndose descrito en numerosos casos que el gen insertado se delecionaba de forma parcial o total según avanzaba la infección, terminando por imponerse variantes virales que no expresaban el gen extraño. Otra estrategia que se ensayó para facilitar la expresión de genes de interés fue la traducción del gen de la cápsida viral a partir de un elemento IRES (sitio de entrada interna de ribosomas) en un ARN subgenómico bicistrónico. También se usaron vectores de expresión basados en este virus para la producción de pequeños péptidos fusionados a la proteína de la cápsida viral o en estudios destinados a usar las plantas como fábricas biológicas de pro-

ductos de interés. Mediante la coinfección de dos virus recombinantes basados en dos virus X de la patata, o en un virus X de la patata y un TMV, usando un sistema de magnificación mixto, se ha conseguido co-expresar las cadenas ligera y pesada de anticuerpos y demostrar su funcionalidad.

El virus del grabado (jaspeado) del tabaco es un potyvirus bastante bueno como vector de inserción. Los potyvirus tienen un sistema de traducción complejo, con poliproteínas que son procesadas en varios pasos, por lo que no se debe insertar un gen de interés en un sitio arbitrario. Hay, sin embargo, lugares concretos donde sí se puede introducir un gen heterólogo que quedaría, bien como gen fusión con otro endógeno del virus o llevando su propia secuencia diana para el procesamiento que permita la separación del producto de interés. Los genes señal *gus* y *gfp* (proteína verde fluorescente) han podido expresarse muy eficientemente con este vector aunque se producen delecciones en el gen inserto con cierta frecuencia.

El virus de la Sharka es también un potyvirus que produce la enfermedad de la viruela del ciruelo, afectando también a melocotoneros y albaricoqueros. En condiciones experimentales se puede propagar en diferentes especies herbáceas, como guisantes, *Nicotiana* o *Arabidopsis thaliana*. Se han desarrollado vectores basados en el virus de la sharka para la producción tanto de pequeños péptidos fusionados a la proteína de la cápsida como de proteínas completas independientes. Se ha observado que diferentes vectores para la expresión de epítomos de agentes patógenos en la superficie de las cápsidas del virus mostraban diferencias en su tolerancia a secuencias insertadas y en la antigenicidad e inmunogenicidad de los epítomos expresados. Se han construido vectores con dos sitios de clonación distintos para producir dos proteínas. Estos vectores se han usado tanto para expresar genes señal que facilitan la monitorización de la infección viral, como para producir proteínas antigénicas para la obtención de vacunas recombinantes. El tamaño del inserto sigue siendo uno de los elementos restrictivos para la expresión de genes foráneos en este vector. La planta se transforma con un ADNc correspon-

diente al genoma viral. Estos sistemas, conocidos como amplicones, combinan la estabilidad genética de las plantas transgénicas con las altas tasas de replicación de los virus. Se han desarrollado diferentes amplicones basados en el virus de la sharka. Uno de ellos se ha usado para diseñar un método de control de la expresión mediante la regulación de la temperatura durante la transformación y el subsecuente cultivo de las plantas transgénicas resultantes.

Como vemos, el mayor problema para la utilización generalizada de virus de plantas como vectores, hoy por hoy, es la inestabilidad y eliminación del inserto, sobre todo si el gen de interés tiene un tamaño superior a la kilobase. Desgraciadamente no se conocen los mecanismos moleculares que provocan estas delecciones, pero, probablemente, sean similares a los descritos para recombinaciones homólogas o heterólogas de muchos virus. Cualquiera que sea el mecanismo responsable, es muy posible que la facilidad mostrada por estos virus de plantas (tanto ADN como ARN) para eliminar insertos refleje una estrategia general forjada a lo largo de la evolución, contra la incorporación promiscua de genes foráneos que no aportan ventaja alguna al virus. La habilidad para eliminar genes no deseados es particularmente importante cuando existen restricciones de espacio para el empaquetamiento de la partícula viral. La estructura secundaria que probablemente adquieran estos genomas virales también se verá afectada por la sustitución o inserción de un gen foráneo. Otro problema es la pérdida de infectividad más o menos acusada, asociada a los vectores con insertos o sustituciones, respecto al virus intacto. Esta propiedad, que puede ser una gran desventaja para usos comerciales de estos vectores, puede ser beneficiosa ecológicamente, ya que impediría que una variante manipulada pudiera llegar a suplantar al original en la naturaleza.

9.2. MÉTODOS DE CULTIVO

Dentro de las especies de plantas que más se han utilizado para la manipulación genética expe-

rimental, destacan *Arabidopsis thaliana*, *Nicotiana benthamiana* y *Nicotiana tabacum*.

Callos

Para obtener buenos cultivos de células vegetales se debe partir de un explante rico en células capaces de una rápida proliferación, como las células apicales de la raíz, las yemas o semillas en germinación. A partir de ellas se puede obtener una masa indiferenciada denominada «callo». El callo tisular puede propagarse indefinidamente por subdivisiones en un medio nutricional básico con fitohormonas (una auxina, una citoquinina y una giberelina, por ejemplo). Si el callo se cultiva en medio líquido con agitación, producirá grumos de células y agregados de gran tamaño junto con células individuales. Esta heterogeneidad morfológica está acompañada de gran inestabilidad genética.

El estado de diferenciación puede ser mantenido o alterado variando el tipo y concentración de fitohormonas en el medio: una concentración alta de auxinas frente a citoquininas inducirá la formación de raíces, mientras que concentraciones altas de citoquininas frente a auxinas induce la formación del tallo (figura 9.6). Las concentraciones de hormonas adecuadas varían de una planta a otra, por lo que hay que determinarlas en cada caso experimental.

Protoplastos

Los protoplastos son células a las que se les ha eliminado la pared celular por tratamiento enzimático con celulasa y pectinasa. Se suelen emplear células de las hojas porque son fáciles de dispersar y regeneran muy bien, pero pueden obtenerse de numerosos tipos celulares. La pared celular se re-sintetiza al cabo de 5-10 días y empiezan las divisiones celulares. La facilidad con la que los protoplastos regeneran la planta completa es muy variable. En general, es difícil regenerar plantas monocotiledóneas a partir de protoplastos, aunque se han obtenido algunas plantas de maíz (variedades silvestres) y plantas de arroz.

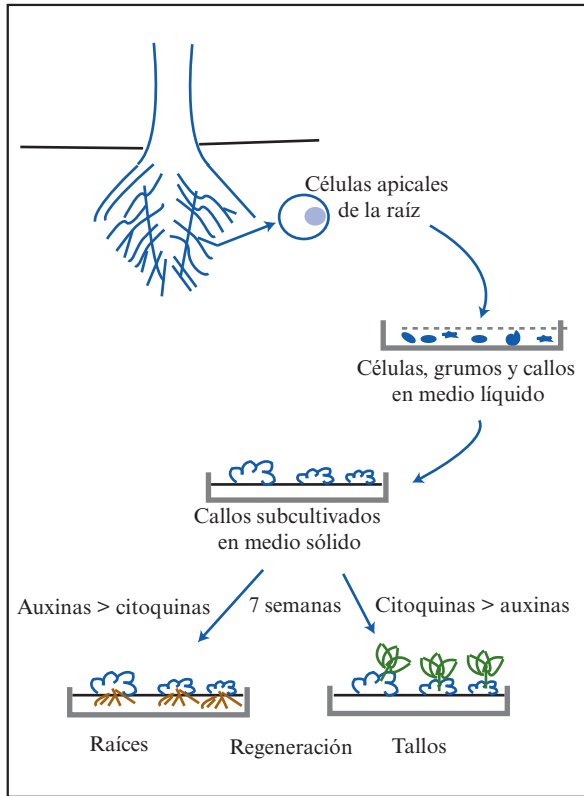


Figura 9.6. Cultivos de callos y variaciones dependiendo de la textura del medio de cultivo y de la concentración de fitohormonas.

También se pueden cultivar protoplastos haploides de polen. La regeneración de éstos origina plantas haploides que pueden ser de gran utilidad para aislar plantas mutantes o analizar genes reguladores.

9.3. MÉTODOS DE TRANSFERENCIA DEL ADN A LA PLANTA

Transferencia mediada por *Agrobacterium*

Aunque en sistemas de transferencia de ADN por *Agrobacterium* el ADN T se integra al azar, si hay homología de secuencias entre alguna región del genoma y el inserto del ADN T, también se pueden obtener dianas génicas por recombinación

homóloga con una frecuencia aproximada, relativa al número de sucesos recombinantes no-homólogos, de 10^{-3} para *Arabidopsis* y entre 10^{-4} y 10^{-6} para otras especies. Se ha conseguido, asimismo, mediada por una recombinación homóloga, la introducción de genes foráneos al genoma de cloroplastos.

Agrobacterium se ha utilizado como mediadora en la transferencia de ADN T manipulado genéticamente a hojas de tabaco, petunia, tomate y otras dicotiledóneas. El procedimiento consiste en seccionar, con una taladradora de papel, por ejemplo, una hoja joven de la planta y obtener pequeños discos de unos 2 mm de diámetro. Estos discos se ponen en contacto con la bacteria agitando suavemente (figura 9.7). La bacteria infectará las superficies cortadas de la hoja y, al cabo de unas horas (una noche), se mata la bacteria con un antibiótico como la carbenicilina. Los discos se transfieren a papel de filtro, que se pone sobre una lámina de células alimentadoras durante 48 horas. Pasado este período de tiempo, el papel de filtro se pasa a un medio de regeneración de brotes con kanamicina u otro agente de selección. Los brotes se transferirán a un medio estimulador de raíces y de ahí a tierra.

La transferencia mediada por *Agrobacterium* se ha simplificado mucho con el método de **infiltración al vacío** puesto a punto para *Arabidopsis* (figura 9.8), pero válido para otras especies, entre ellas, *Nicotiana benthamiana* y *Nicotiana tabacum*. En *Arabidopsis*, la planta erguida se sumerge en un cultivo fresco de *Agrobacterium* y se establece el vacío. La bacteria es succionada por la planta y, por un proceso no muy claro, se transforman las células germinales de los órganos florales en desarrollo. Las plantas transgénicas se seleccionan de las semillas resultantes de la autopolinización de la planta tratada. La verificación de que la planta es verdaderamente transgénica se puede realizar por PCR. En *Nicotiana*, además de las plantas con flores, se han utilizado plantas adultas sin flores; en este caso, las partes aéreas de la planta se sumergen en un cultivo fresco de *Agrobacterium* y se establece un vacío no muy potente (0,5-1 bar) durante 1-2 minutos seguido de un retorno gradual y suave a la presión atmosférica. La

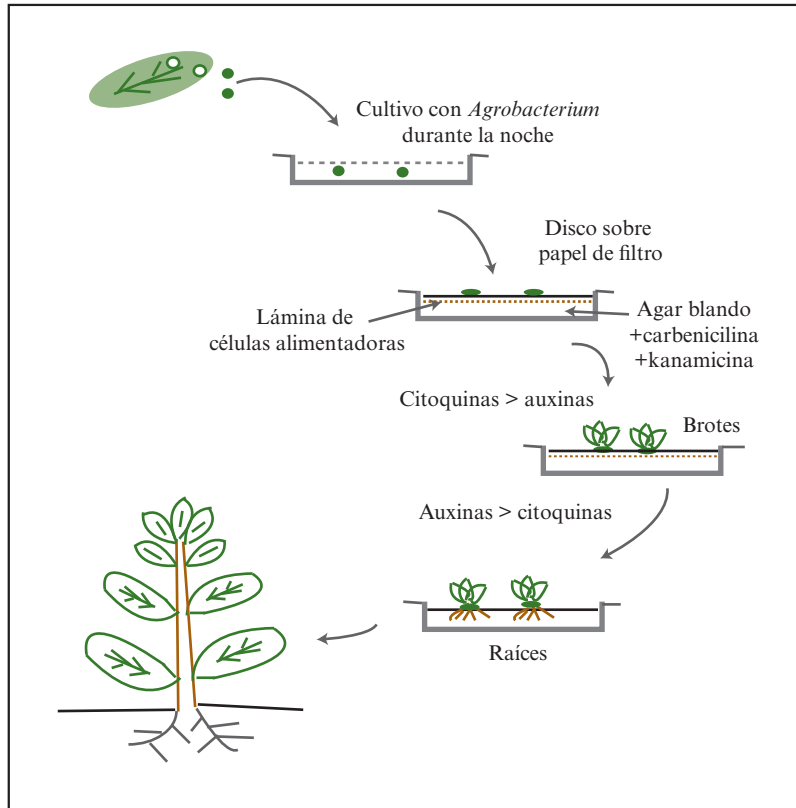


Figura 9.7. Infección de discos de hoja por *Agrobacterium* y transformación por plásmido Ti.

bacteria es succionada por la planta al espacio intercelular de todas las hojas maduras de la planta. El método es también válido para petunia, pepino, girasol, remolacha, espinaca, *Chenopodium capitatum* y *Tegragonia expansa*, entre otros.

Transferencia mediada por virus

Los virus se transmiten en la naturaleza por diferentes vías, incluyendo diversos tipos de vectores específicos. Sin embargo, en un gran número de casos es posible la inoculación experimental simplemente por aplicación manual de extractos de plantas previamente infectadas, o de virus o ácido nucleico viral purificados, junto con un abrasivo. Para comenzar la infección con virus modificados genéticamente, como es el caso de vectores virales

que expresan genes foráneos, se pueden emplear estrategias similares, aplicando en este caso ADN (virus con este tipo de genoma) o ADNc del ARN genómico (virus con genoma ARN) clonados bajo el control de promotores activos en plantas, como el 35S del virus del mosaico de la coliflor. Asimismo, es posible aplicar directamente ARN sintetizado previamente *in vitro* a partir de ADNc clonado bajo el control del promotor de la ARN polimerasa de un bacteriofago, como T7 o SP6. Se han desarrollado también métodos de inoculación más eficaces, como el disparo con pistolas génicas (véase más adelante) de microprojectiles recubiertos con el ADN o ADNc virales, o la infiltración con cultivos de *Agrobacterium tumefaciens* portadores de sistemas Ti que llevan en su ADN T copias completas del ADN o ADNc del virus.

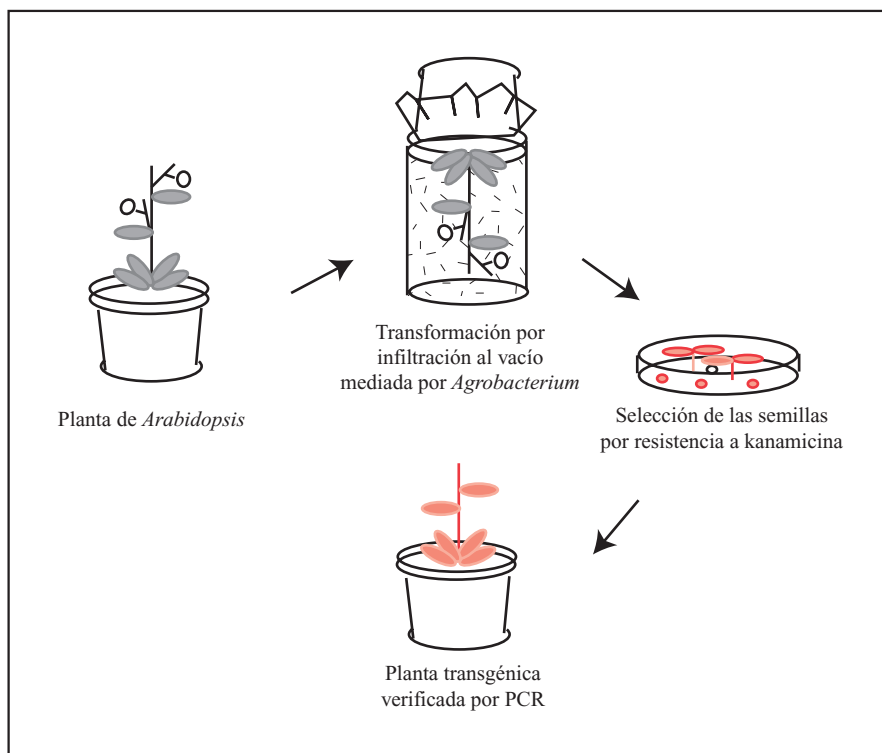


Figura 9.8. Técnica de infiltración al vacío puesta a punto para *Arabidopsis*. La planta se somete al vacío en una solución de *Agrobacterium*. Las semillas transformadas se seleccionan por resistencia a kanamicina y se plantan en tierra. La presencia del transgén se verifica por PCR.

Transferencia a protoplastos

Los protoplastos se comportan como una célula animal y se les puede incorporar ADN por transfección, transducción o electroporación. Es necesaria la presencia de un gen de selección bajo el control de un promotor de plantas potente (el 35S del virus del mosaico de la coliflor, por ejemplo). Los protoplastos, una vez que han ingerido el ADN, se colocan en un medio nutricional adecuado donde regeneran la pared e inician la división celular. En una gran variedad de dicotiledóneas se han inducido tallos y raíces que, posteriormente, han dado lugar a plantas completas transgénicas y fértiles. La técnica ha dado buenos resultados en petunias, maíz, trigo y arroz, pero en muchas otras especies de cereales no se

han podido obtener plantas transgénicas fértiles a partir de ellos.

Pistolas génicas, microproyectiles y ultrasonidos

Las pistolas génicas introducen microproyectiles de oro o tungsteno (4 μm de diámetro) cubiertos de ADN directamente en las células de la planta. Hay dos tipos de pistolas: 1) El cañón de la pistola está taponado por una membrana, que se rompe cuando se aumenta la presión de helio dentro de la pistola y se hace vacío fuera de ella; cuando se rompe la membrana, sale un chorro de helio que arrastra a una placa portadora sobre la que se depositan los microproyectiles recubiertos de oro; más allá existe una pantalla de parada que retiene a la placa portadora y deja pasar a los mi-

croproyectiles (figura 9.9). 2) Pistola en la que el helio sale a presión a través de un tubo de plástico en cuya cara interior se depositan las partículas de oro cubiertas de ADN; el chorro de helio arrastra a los microproyectiles hacia el exterior, que se encuentra a presión ambiental; así pues, aquí no hay placa portadora que salga disparada ni pantalla de parada. Este segundo tipo de pistola es el que se usa para inocular virus a plantas.

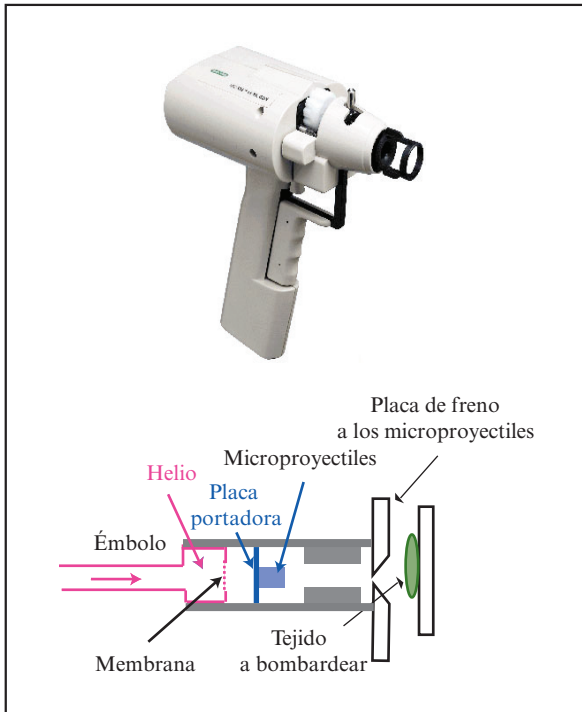


Figura 9.9. Pistola génica que dispara microproyectiles recubiertos de ADN a tejidos de plantas.

Se ha logrado transformar por estos métodos células epidérmicas de cebolla (*Allium cepa*), alcanzando eficiencias de hasta el 44%. En general, se ha observado que si las células receptoras están en las fases M o G2, el bombardeo es entre 4 y 6 veces más eficiente que si están en S o G1. La eficiencia, en este caso, baja considerablemente para transformaciones estables y obtención de plantas fértiles derivadas. También se han obtenido con

estos métodos plantas transgénicas de arroz, incluyendo algunas especies que hasta ahora eran recalcitrantes a la manipulación. El método consistió en bombardear células embrionarias de unos 12 días con partículas de oro recubiertas de ADN expulsadas por un acelerador de partículas eléctrico.

Entre la variedad de tejidos y especies que se han bombardeado con éxito están: embriones de soja, guisante, cacahuete, papaya, arroz y algodón; callos de trigo y caña de azúcar; tallos de arándano y raíces de *Arabidopsis*. En cuanto a árboles, se han obtenido chopos (*Populus*) transgénicos a partir del bombardeo de tres tejidos: callos, nódulos y tallos.

El bombardeo de partículas también puede utilizarse en las infecciones con *Agrobacterium* como método de producir las heridas asociadas a la infección. En tabaco se ha visto que los bombardeos con *Agrobacterium* producen unas 100 veces más plantas transgénicas que el bombardeo normal de ADN desnudo.

9.4. EXPRESIÓN DEL TRANSGÉN

La expresión del transgén una vez ya en la célula vegetal no está garantizada. Las plantas poseen un mecanismo de regulación génica y defensa con tendencia nata al silenciamiento de genes de procedencia dudosa. El análisis de las modificaciones que sufren los transgenes en plantas está desvelando los procesos generales de silenciamiento génico vía ARNi. El ARNi puede inducir metilaciones y represión transcripcional del transgén y su entorno. Es evidente que los genomas de las plantas pueden detectar y responder a la presencia de ADN extraño en el contexto de ciertas secuencias y en numerosos lugares.

Distintos tipos de ARNi regulan en la planta transcripción y traducción, así como la eliminación de ARN de infecciones virales. La variedad y complejidad de los ARNi en estos seres vivos es enorme. Las plantas poseen muchos tipos de ARNi endógenos, es decir, codificados por el propio genoma de la planta, entre ellos se encuentran los miARN, como los que hemos visto en anima-

les. La función de estos miARN en plantas es controlar la función de genes a su vez reguladores, como factores de transcripción, especialmente durante el desarrollo. Estos miARN se transcriben por la polimerasa II y generan estructuras en horquilla al plegarse sobre sí mismos; seguidamente son procesados por complejos enzimáticos del tipo RISC (capítulo 1, apartado 1.2.2). Los miARN reconocen transcritos homólogos y producen un solo corte entre las bases 10 y 11 relativas al extremo 5' del miARN; los dos fragmentos resultantes en el ARN diana son sustrato de exorribonucleasas que terminarán de seccionarlo totalmente. Los miARN también pueden impedir la traducción.

Como ya hemos indicado, el ARN de interferencia se descubre en petunias, donde el silenciamiento no se limita al transgén, sino que afecta al gen homólogo endógeno, sugiriendo que el locus silenciado puede imponer en trans un grado de metilación similar al suyo en cualquier región homóloga.

En *Arabidopsis* se han generado distintas líneas transgénicas con un número variable de copias del transgén en un mismo locus, observándose que las líneas con copias repetidas del transgén estaban silenciadas y metiladas, mientras que las copias únicas se expresaban y no estaban metiladas. Las secuencias CpG y CpNpG representan los sustratos preferentes de metilación de citosinas en plantas, pero no son los únicos. Se sabe que la metilación de un promotor bloquea el inicio de la transcripción y causa silenciamiento génico transcripcional, pero no está muy claro el mecanismo por el que la planta reconoce el gen en exceso. El modelo de metilación de ADN vía ARN postula que la inducción de la metilación en el ADN es una concentración límite de ARN, por encima de la cual se dispararía un mecanismo para metilar de novo el gen correspondiente y/o su promotor.

El silenciamiento puede ser transcripcional o postranscripcional. En el primero se bloquea la síntesis de ARN, mientras que en el segundo se observa una reducción paulatina en los niveles de ARNm sin afectar a los niveles de transcripción. Ambos casos son frecuentes en plantas transgénicas con múltiples copias de un gen endógeno o un transgén.

La mayoría de los virus que infectan plantas son virus ARN que tienen como genoma un ARN monocatenario. Las plantas poseen una ARN polimerasa que utiliza como molde un ARN monocatenario y que lo convierte en intermediario de replicación bicatenario. Una vez como ARN de doble hebra sería sustrato de las enzimas tipo dicer que al romperlo en forma de dado (21 nucleótidos) lo convertirían en ARNi sustrato de RISC y de aquí pasaría a buscar y eliminar los ARN virales genómicos.

Hay otros factores que también pueden determinar o contribuir al silenciamiento de un transgén. El lugar de integración es uno de ellos (efecto de posición), que se refiere a la influencia represiva que ejercen sobre un transgén las secuencias cromosomales inactivas vecinas al inserto; sólo quedarían como transgenes funcionales aquellos que se hubieran introducido en una región transcripcionalmente activa de ADN. La acetilación de la cromatina precede en un gran número de casos a su activación transcripcional, y la desacetilación, a su silenciamiento; se ha postulado que la inserción de un transgén en determinados lugares provocaría una desacetilación de la cromatina causando el silenciamiento.

9.5. PLANTAS TRANSGÉNICAS EN LA AGRICULTURA

La población mundial crece de forma alarmante y los recursos naturales son limitados. Necesitamos cosechas con más cantidad y más calidad de productos para conseguir una agricultura sostenible en las próximas décadas. La ingeniería genética y la biotecnología asociada permiten combatir las deficiencias nutricionales de muchos vegetales aumentando su contenido proteico, lipídico, de hidratos de carbono o vitamínico. Desde 1990 ha habido un énfasis especial en ingeniería genética y biotecnología agrícola para desarrollar y obtener cultivos y cosechas de mayor calidad nutricional, resistentes a herbicidas, resistentes a plagas de insectos o virus, así como plantas productoras de biocombustibles o vacunas. Los tomates fueron los primeros vegetales transgénicos

que se comercializaron, para el consumo humano, en Estados Unidos de Norteamérica en 1994. Desde entonces, el camino no ha sido fácil, ya que no siempre la ingeniería genética asociada ha podido apoyarse sobre un conocimiento básico de la biología molecular y la genética de las especies de interés agrícola. Además, la sociedad se mantiene recelosa sobre la posibilidad de que los productos transgénicos puedan producir alergias, puedan transferir genes de resistencia a antibióticos al genoma humano o de que las plantas transgénicas puedan expandirse sustituyendo y suplantando a sus homólogas silvestres. A pesar de todas estas barreras, muchísimos países incluyendo China, Canadá, Estados Unidos de Norteamérica, Brasil, Argentina, India, Australia, Méjico, Chile, Alemania, Polonia, Portugal y España producen cosechas transgénicas, principalmente de maíz, algodón, remolacha, trigo, arroz, soja, tomate, petunias y claveles, entre otros.

En cuanto a los retos pendientes, está la producción de biocarburantes o de vacunas. Se están desarrollando por ingeniería genética variedades del alga verde-azulada (cianobacterias) que producen biocarburantes o ácidos grasos a escala piloto y hay que tener en cuenta que estas algas no compiten en nutrientes con otros cultivos, ya que sólo necesitan agua y sol para crecer. Otra estrategia emergente para producir biocombustibles es la eliminación por KO del gen que codifica la enzima CSE (cafeoil shikimate esterasa) implicada en la síntesis de la lignina y cuyas mutaciones reducen considerablemente la cantidad de este compuesto indigerible y así multiplican por cuatro la eficacia de su digestión para hacer biocarburante y para fabricar papel. De momento los estudios se han llevado a cabo en *Arabidopsis* con bastante éxito.

Aunque se han realizado muchos esfuerzos para aislar y caracterizar los genes responsables de la tolerancia a la salinidad, la aridez del terreno o a los extremos térmicos, así como la resistencia a insectos, pesticidas, etc., la complejidad de la fisiología relacionada con el estrés que provocan estas situaciones en la planta dificulta el desarrollo de las aplicaciones prácticas de muchos de estos proyectos.

De entre las numerosas aplicaciones de plantas transgénicas con interés agrícola-comercial, veremos en los siguientes subapartados el estado actual de: la obtención de semillas híbridas a gran escala; la maduración controlada de frutos y flores; la concepción de plantas resistentes a herbicidas; la construcción de plantas resistentes a patógenos (virus, bacterias, hongos o insectos); las plantas tolerantes a las situaciones estresantes de salinidad, aridez o frío y, por último, las plantas como fábricas de plásticos, pinturas, anticuerpos o enzimas.

Obtención de semillas híbridas a gran escala

La mejora de plantas con interés agrícola está en muchos casos asociada a la producción de variedades híbridas. Las plantas autofecundadas son más propensas al fenómeno de la homocigosis de factores deletéreos responsable de que estas plantas produzcan descendientes más débiles que las cruzadas (vigor del híbrido). La producción de semillas híbridas a gran escala es un reto porque muchas plantas de cultivo tienen los estambres y pistilos en la misma planta, bien en la misma flor (como la planta de la colza o el tomate), o en flores distintas (como el maíz). Esta disposición favorece la autopolinización impidiendo los cruces entre distintas líneas de plantas. Para garantizar una polinización cruzada, agricultores o productores de semillas debían quitar los estambres manual o mecánicamente, lo cual es bastante difícil cuando las flores son muy pequeñas como en la planta de la colza. El estudio de genes que producen esterilidad masculina por inhibición de la producción de polen y de otros genes que restauran la fertilidad a plantas estériles masculinas, así como su aplicación en plantas transgénicas, está facilitando el desarrollo de cosechas híbridas a gran escala. La estrategia, que coloquialmente se ha denominado «terminator», consiste en combinar el promotor específico del tejido tapetal (pTA) con el gen citotóxico de la ARNasa T. El tejido tapetal sirve de soporte a las microesporas y el polen de los sacos polínicos del estambre y su promotor se aisló originalmente de *Nicotiana tabacum* (pTA-29). El gen de la ARNasa T procede de

Bacillus amyloliquefaciens y se denomina *barnasa* (figura 9.10). La expresión selectiva de este gen en los estambres inhibe la formación del tapetal y del polen causando esterilidad masculina. El gen correspondiente al inhibidor de la ARNasa en la bacteria, *barstar*, está asimismo aislado y caracte-

rizado. Así, podemos obtener también plantas transgénicas portadoras del gen *barstar* unido al mismo promotor específico TA-29. El cruce de estas últimas, con las primeras estériles masculinas produce progenie fértil, ya que *barstar* es un supresor dominante de *barnasa*.

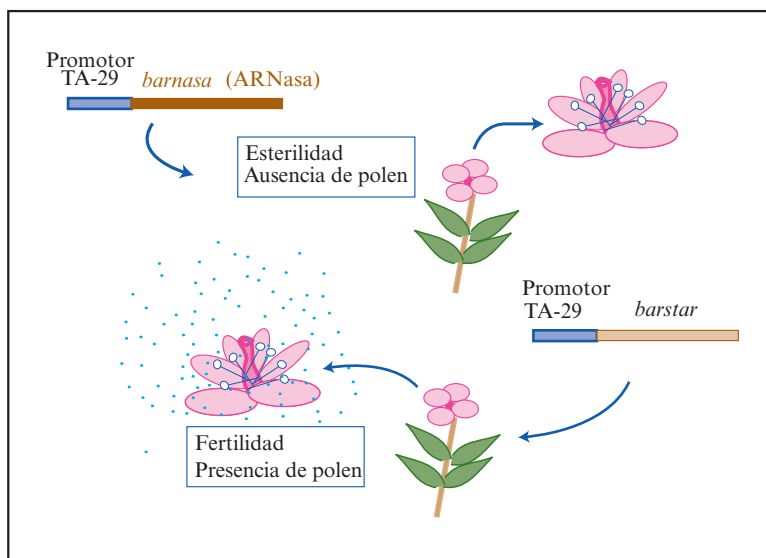


Figura 9.10. Estrategia para producir semillas híbridas a gran escala. El promotor específico de los sacos polínicos TA-29 controla la síntesis del gen citotóxico *barnasa* y se producen flores sin polen. Recuperación de la fertilidad al cruzar las primeras plantas con otras que lleven el inhibidor *barstar*.

Maduración controlada de frutos y flores

El saber popular tiene conocimiento desde hace décadas de que las frutas y legumbres maduran más rápidamente en una bolsa de plástico bien cerrada, y en la antigua Grecia (3 siglos antes de Cristo) Teófrates describe cómo hay que raspar con una uña de hierro los higos del sicomoro para que maduren. Hoy sabemos que las heridas inducen la producción de etileno (C_2H_4 , $CH_2 = CH_2$), verdadera hormona de la maduración de la fruta. En las plantas, la maduración permite una mejor dispersión de las semillas; la producción de etileno es la causa del proceso.

Se han aislado los genes que se activan durante el proceso de maduración y que provocan un

aumento respiratorio, además de la producción de etileno. Uno de estos genes es el correspondiente a la enzima ACCS (aminociclopropano-1-carboxilato sintasa), fundamental en el proceso de síntesis de etileno. Se han diseñado dos procedimientos para bloquear la enzima, el primero consiste en introducir en las plantas una construcción en la que se transcriba un ARN antisentido respecto al del gen *accs*, bajo el control del promotor 35S del virus del mosaico de la coliflor. Otra ruta ha sido la de clonar el gen bacteriano de la *acc deaminasa* que metaboliza y degrada ACC (aminociclopropano-1-carboxilato). Este segundo procedimiento es el que se ha empleado en la obtención de tomates transgénicos que tienen una inhibición del 90 al 97% en la producción de eti-

leno y cuya maduración se retrasa entre uno y tres meses. Los tomates se recogen de la planta en el momento deseado y se maduran por tratamiento externo con etileno o propileno durante 6 días. Los frutos tratados tienen el mismo color, textura, aroma, complexión y sabor que los madurados en la propia mata.

Muchas flores ornamentales, como las rosas, los claveles o los tulipanes, tardan muy poco en marchitarse una vez cortadas. Se han utilizado claveles transgénicos portadores del gen *aco*, correspondiente a la enzima ACC oxidasa, en versión antisentido para bloquear la síntesis de etileno. Los claveles producidos mostraron un retraso apreciable en marchitarse.

Plantas resistentes a herbicidas

Los herbicidas en principio no distinguen unas plantas de otras, por lo que su acción es generalizada destruyendo tanto las malas hierbas como las plantas de interés. **El glifosato** (fosfometil glicina) es un buen herbicida porque los microorganismos del suelo lo pueden degradar con bastante rapidez. Es un herbicida de amplio es-

pectro de nombre comercial «Round up». Si introducimos genes que generen resistencia al glifosato en las plantas de interés, podremos utilizar el herbicida durante la fase de crecimiento de las plantas y cosechar las de interés, sin miedo a la competencia de nutrientes y abonos con las malas hierbas. El glifosato actúa inhibiendo la síntesis del ácido corísmico y aminoácidos aromáticos. Esta misma vía metabólica existe en la bacteria *Salmonella typhimurium*, pero no se encuentra en mamíferos. El glifosato es un inhibidor competitivo de la enolpiruvato shiquimato 3-fosfato (EPSP) sintetasa, por lo que plantas transgénicas con copias extra del gen que produce la enzima superan la inhibición. Existe también un mutante de *S. typhimurium*, con una sustitución de una prolina por una serina en el gen de la EPSP sintetasa (*aro A*), que produce una enzima con baja afinidad para el glifosato, manteniendo la actividad específica intacta (figura 9.11). Se han obtenido plantas de maíz, tabaco, petunia y soja resistentes a glifosato utilizando la enzima resistente o varias copias de la enzima en vectores derivados de los plásmidos Ti o Ri y la técnica de transformación de discos de hoja.

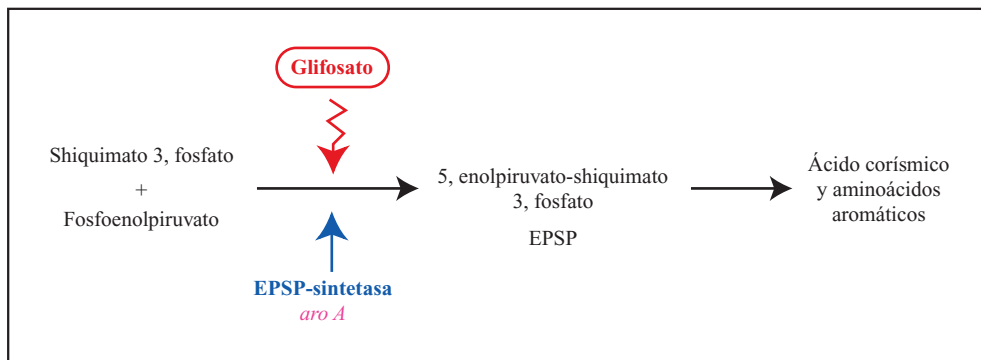


Figura 9.11. Inhibición del glifosato en la vía metabólica de síntesis del ácido corísmico y lugar de acción del gen bacteriano mutado *aro A*.

También se ha obtenido trigo (*Triticum aestivum*) transgénico y fértil resistente al herbicida **Basta**. La fosfinotricina, el componente activo de Basta, es un análogo del glutamato que inhibe la

glutamina sintasa provocando una acumulación letal de amonio (figura 9.12). Al ser esta especie monocotiledónea, se eligió el bombardeo de microproyectiles a callos embrionarios como méto-

do de introducir el gen *bar* (fosfinotricina acetiltransferasa) originalmente aislado de *Streptomyces hygroscopicus*, que acetila e inactiva la fosfinotricina. El ADN recombinante bombardeado utiliza el promotor 35S del virus del mosaico de la coliflor para dirigir la expresión de *bar* y, en sentido contrario, el promotor del gen correspondiente a la enzima alcohol deshidrogenasa (*Adh*) para transcribir el gen señal *gus*. También se incluyen dos regiones de terminación de la transcripción del gen de la nopalina sintetasa (*nos*) y dos intrones del gen *adh* para dar estabilidad a los mensajeros. Se comercializan numerosas plantas transgénicas que llevan el gen de resistencia *bar*, entre ellas: tomate, remolacha, patata, alfalfa y tabaco. En la actualidad, el 71 % de todas las cosechas a nivel mundial son resistentes a herbicidas; además de los cultivos ya mencionados, también se usa en maíz, soja, tabaco, colza y algodón.

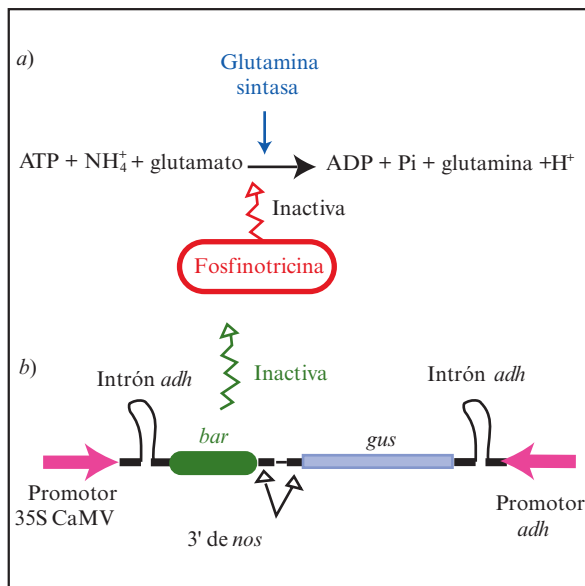


Figura 9.12. Mecanismo de resistencia al herbicida Basta. *a)* Lugar de acción de la fosfinotricina en la biosíntesis de glutamina como inhibidor de la glutamina sintasa. *b)* Construcción que confiere resistencia al herbicida Basta. CaMV = virus del mosaico de la coliflor; *adh* = gen de la alcohol deshidrogenasa; *bar* = gen de la fosfinotricina acetiltransferasa.

Plantas resistentes a patógenos y plagas: virus, bacterias, hongos, nemátodos e insectos

Los patógenos causan pérdidas multimillonarias en cosechas de todo el mundo, por lo que las plantas transgénicas resistentes a los diversos patógenos se han explorado a fondo. La planta posee diversos mecanismos naturales de defensa frente a la infección, y una de las estrategias empleadas es inducir esta respuesta sin perturbar el crecimiento general. Es conocido el fenómeno de que la expresión, por parte de la célula vegetal, de material genético del patógeno desestabiliza los procesos secuenciales y esenciales del agente agresor, desencadenando una resistencia contra dicho patógeno.

Resistencia a virus

Como ya hemos comentado, la mayor parte de los virus de plantas tienen ARN como genoma, no despiertan un gran rechazo por parte de la célula vegetal y tienen frecuencias de recombinación muy altas que producen múltiples variantes.

Se han obtenido plantas resistentes a virus utilizando el gen de la proteína de la cubierta (o cápsida), la replicasa (o parte del gen), la proteína de movimiento y algunas otras, concluyéndose que cualquier parte del genoma viral puede inducir resistencia. Aunque en la primera planta transgénica resistente a virus que se construyó la proteína producto del transgén era el agente inductor de la resistencia, en la mayoría de los casos siguientes se confirmó, usando construcciones no traducibles y por otros medios, que simplemente el ARN es suficiente. La protección en la mayoría de los casos no es de amplio espectro y se limita a virus muy relacionados. La resistencia mediada por la proteína de la cubierta se describió por primera vez en el virus del mosaico del tabaco. El ADNc del ARNm de la proteína de la cubierta del virus, clonado bajo el control del promotor 35S del virus del mosaico de la coliflor, se insertó en el genoma nuclear de plantas de tabaco a través del sistema Ti y *Agrobacterium*. El ADN T utilizado llevaba también el gen *neo*. Las células transformadas se seleccionaron por resistencia a la kanamicina y regeneraron plantas completas. Las

plantas contenían de 1 a 5 copias del ADNc y la proteína de la cápsida llegó a constituir el 0,1% de la proteína total de las hojas. Descendientes de estas plantas se inocularon con TMV y en un alto porcentaje de los casos, dependiendo de la línea transgénica, no hubo infección. En los que se produjo la infección, ésta apareció mucho más tarde que en las plantas control. No se ha determinado con total exactitud las bases moleculares de esta resistencia, pero se sabe que está inducida fundamentalmente por el producto proteico del transgén. Es posible que la proteína de la cápsida reempaque el ARN viral convirtiéndolo en no infeccioso, o que dicha proteína bloquee los receptores internos que intervienen en la liberación del ARN viral (desencapsulación).

Los virus ARN también desencadenan el silenciamiento mediado por ARNi a través de la ARN polimerasa dependiente de ARN, que sintetizará un ARN bicatenario a partir del infectivo viral, como hemos mencionado con anterioridad (apartado 9.4). La mayoría de las resistencias a virus generadas por la expresión de fragmentos del genoma viral se deben a la potenciación del sistema endógeno de ARNi antiviral. Además, las plantas tienen otras variantes y mecanismos más complejos de ARNi, que no detallaremos, para deshacerse de ARN de procedencia dudosa. No debemos olvidar que el fenómeno de silenciamiento de transgenes tiene su origen en este sistema natural de defensa contra las infecciones virales.

El fenómeno de la resistencia mediada por proteínas o ARN virales se ha demostrado en más de 10 grupos de virus distintos y plantas correspondientes, entre las que se encuentran el tabaco, la alfalfa, la petunia, el tomate, la cebada, la patata, el trigo y el maíz.

Se han descrito casos en especies de *Nicotiana* o *Brassica* en los que plantas silvestres infectadas con virus son capaces de superar la infección haciéndose resistentes a un segundo inóculo.

Resistencia a nemátodos

Las infecciones de plantas por nemátodos causan grandes pérdidas económicas en muchos

cultivos preferentemente de patata, remolacha, soja, cebada, trigo, tomate y cacahuete. Algunas plantas exhiben una notable resistencia a nemátodos y otras han perdido esta resistencia. El análisis de reordenaciones cromosómicas asociadas a marcadores (ADN satélite) que segregan con la resistencia ha permitido el clonaje de algunos genes implicados en la protección frente a nemátodos. Los géneros más peligrosos y devastadores son *Heterodera*, *Globodera* y *Meloidogine*, todos ellos endoparásitos que viven dentro de la raíz. Típicamente estos nemátodos tienen un amplio abanico de hospedadores y sus huevos pueden sobrevivir durante años en la tierra.

Uno de los genes clonados es el *Rhg4* (resistencia a *Heterodera glycines* 4), obtenido de la soja (*Glycine max*), que confiere resistencia al nemátodo microscópico *Heterodera glycines*. El gen *Rhg4* codifica por una serina-hidroximetiltransferasa responsable de la conversión de serina en glicina e indispensable en el metabolismo y ciclo de los folatos. Los alelos que confieren resistencia se diferencian de los que no lo hacen en dos polimorfismos próximos al centro activo de la enzima que alteran una propiedad reguladora clave. La deficiencia de folato en la planta puede desencadenar una carencia nutricional letal en el desarrollo del nemátodo. Con carácter experimental se ha obtenido soja transgénica resistente a la infección.

Resistencia a insectos

Las plantas pueden ser devastadas por plagas de insectos, cuyas larvas suelen ser muy voraces. Todos los vegetales poseen en mayor o menor medida resistencia a insectos mediada por metabolitos secundarios nocivos. Hoy se conocen más de 40 genes, muchos de ellos de origen bacteriano, que confieren resistencia a insectos y algunos de estos genes se han introducido en plantas de interés agrícola que se han comercializado. Esta tecnología ofrece algunas ventajas biológicas y ecológicas frente a insecticidas convencionales, como son la protección individual de la planta transgénica, la biodegradabilidad del compuesto, la independencia de las condiciones climáticas en su efectividad, la protección del medio ambiente evi-

tando la dispersi3n de insecticidas y el ahorro econ3mico.

La mayor3a de los genes introducidos en plantas proceden de microorganismos y producen toxinas letales para algunos de estos insectos; adem3s, suelen ser espec3ficas contra el aparato digestivo de la larva.

Bacillus thuringiensis es un ejemplo de bacteria productora de una toxina letal. Cuando esporula en la superficie de las hojas, produce unos cristales que se convierten en p3ptidos t3xicos al ser atacados por las proteasas del tracto intestinal de las larvas. Las orugas que ingieren la toxina quedan paralizadas. Los cristales de la toxina pueden ser aislados y utilizados para vaporizar las plantas, pero esto es muy costoso y suele dejar abundantes residuos en la planta. Es m3s f3cil y efectivo introducir el gen *bt*, que codifica la toxina de *B. thuringiensis*, en el genoma de la planta. De esta manera se han obtenido numerosas plantas transg3nicas que expresan la endotoxina. Las primeras plantas de este tipo se originaron en tabaco introduciendo diversas construcciones de ADN T. El aspecto exterior de las plantas era indistinguible de las no-transg3nicas. Los niveles de la endotoxina Bt producida fueron suficientes para matar larvas de primer estadio de *Manduca sexta* y la propiedad insecticida era heredable. Algunas de las plantas que produc3an niveles elevados de la toxina tambi3n engendraron semillas resistentes a kanamicina, originando plantas con 5 o m3s copias del ADN T por genoma. Desde estos primeros experimentos, se han caracterizado al menos 10 genes que codifican distintas toxinas Bt contra lepid3pteros (*cry1Aa*, *cry1Ab*, etc.). Como la frecuencia de uso de codones es muy distinta entre bacterias y plantas, ha habido que modificar y optimizar la secuencia original de los genes naturales *bt* bacterianos. Algunos de estos genes *bt* se han introducido en una gran variedad de plantas transg3nicas como: tomate, ar3ndanos, algod3n, ma3z, patata, arroz, br3col, repollo, vid, colza, cacahuete, soja, alfalfa, ca3a de az3car, manzano, peral, nogal y *Arabidopsis*, entre otras. Tambi3n se han obtenido plantas resistentes a larvas de cole3pteros con el gen *cry3A* en berenjenas, patata, tabaco y alfalfa.

En la actualidad se buscan promotores espec3ficos que permitan la expresi3n de genes de resistencia s3lo en aquellos tejidos susceptibles de ser infectados, ya que una expresi3n potente y generalizada de un agente nocivo para el insecto puede provocar en 3ste la aparici3n de productos que contrarresten su acci3n y aun en el hombre que ingiera la planta transg3nica. Ejemplos de estos promotores son: el promotor espec3fico de expresi3n en el floema para su uso contra pulgones que lo extraen, los promotores inducidos por heridas o los promotores confinados al polen.

Plantas resistentes al estr3s generado por salinidad, aridez o fr3o

La planta est3 sometida a numerosos cambios radicales de temperatura (heladas, calores), de irrigaci3n (inundaciones, sequ3as), de salinidad en el suelo, etc., y hay genes espec3ficos que facilitan a la planta la adaptaci3n a nuevas condiciones ambientales. La caracterizaci3n de estos genes persigue crear plantas m3s tolerantes a altas concentraciones de sal, temperaturas extremas, inundaciones o sequ3a.

La acumulaci3n de glicina-beta3na protege a la planta contra los cambios osm3ticos, produciendo tolerancia a la salinidad y a la deshidrataci3n del suelo. El problema b3sico es la capacidad para retener agua. Hay algunos compuestos org3nicos llamados osmoprotectores que favorecen esta retenci3n. El amonio cuaternario y compuestos derivados, como la glicina-beta3na, son parte de ellos. Est3 presente en bacterias, plantas, algas y animales. Hay otros, como los iones Na^+ , Cl^- o SO_4^{2-} , pero 3stos son t3xicos si se acumulan en el citoplasma. Otros derivados, como la prolina-beta3na, se encuentran en las fab3ceas y lamil3ceas. La glicina-beta3na se acumula proporcionalmente a la sal existente en el suelo. Experimentos con espinaca, planta que tolera moderadamente la sal, muestran que al aumentar la salinidad aumenta la concentraci3n de beta3na, junto con la de otros osmoprotectores.

La colina deshidrogenasa de *E. coli* cataliza la conversi3n de colina en glicina-beta3na v3a aldeh3do de beta3na como intermediario (figura 9.13).

Se han obtenido plantas transgénicas de tabaco y patata, tolerantes a sal y heladas, respectivamente, portadoras del gen *betA*, responsable de la enzima colina deshidrogenasa de *E. coli*, pero no está muy claro si las concentraciones de glicina-betaína alcanzadas (5 mmol por gramo de peso seco en la patata) son suficientes para facilitar los reajustes osmóticos, o se ha producido algún tipo de estabilización celular de estructuras o macromoléculas común a ambos fenotipos.

Otro transgén que se puede utilizar con los mismos fines es el de la colina oxidasa de *Arthrobacter globiformis* o *A. pascens*. Esta enzima cataliza los dos pasos de conversión de colina a glicina-betaína (figura 9.13). Plantas transgénicas de *Arabidopsis* portadoras del gen resultaron tolerantes a heladas y salinidad. Utilizando un péptido de tránsito se pudo dirigir la proteína a los cloroplastos, lugar habitual de localización, y se observaron concentraciones de 50 mM, lo que ya parece suficiente para facilitar los ajustes osmóticos necesarios.

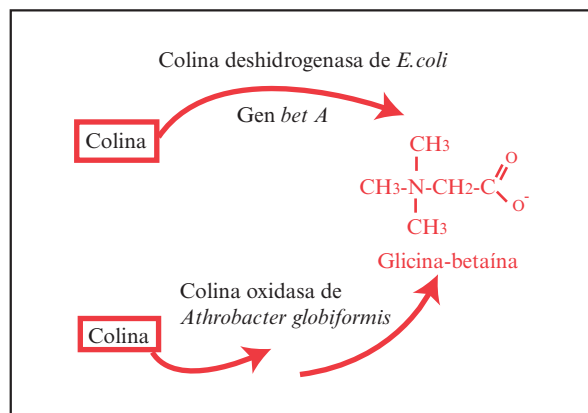


Figura 9.13. Plantas tolerantes a la salinidad. Obtención de glicina-betaína por dos rutas: 1) a través de la enzima colina deshidrogenasa de *E. coli* y 2) en dos pasos mediante la colina oxidasa de *Arthrobacter globiformis*.

Las plantas, además de acumular compuestos osmoprotectores, secuestran directamente iones Na^+ en las grandes vacuolas intracelulares. En

Arabidopsis las proteínas transportadoras de iones implicadas en la tolerancia a la sal están bastante bien caracterizadas (figura 9.14). La más importante es la antiportadora sodio/protón *AtNHX1*, localizada en las membranas de la gran vacuola y prevacuolas.

En algunos géneros de plantas, como en *Plantago*, se ha observado que distintas especies poseen notables diferencias respecto a la tolerancia a la sal. Así, mientras *P. maritima* es muy tolerante, *P. media* lo es mucho menos, debiéndose la diferencia a la presencia en la primera de la proteína antiportadora de sodio/protón (Na^+/H^+), ausente en *P. media*. Se han obtenido plantas de *Arabidopsis* resistentes a la salinidad del suelo simplemente induciendo una expresión exagerada del gen *AtNHX1*. Plantas transgénicas con varias copias de este gen también son significativamente más tolerantes a la sal y crecen bien aunque se rieguen con agua con 200 mM de ClNa de salinidad. Con esta estrategia parece muy probable que se puedan manipular toda una gama de plantas con interés comercial que se podrían regar con agua salada procedente del mar. Este tipo de manipulación genética es muy probable que tenga una mejor acogida social que las plantas resistentes a herbicidas o pesticidas.

Además de los genes osmoprotectores y los implicados en el secuestro de iones descritos, existen otros que también contribuyen a la tolerancia al estrés y que han sido caracterizados y utilizados en plantas transgénicas a nivel experimental: genes de enzimas que catalizan la formación de dobles enlaces-*cis* en ácidos grasos saturados o monosaturados de membrana provocan resistencia a las heladas, por bajar la temperatura de transición de fases a 0°C ; o genes que codifican proteínas limpiadoras de radicales libres de oxígeno como las superóxido dismutasas, peroxidasas o catalasas.

Aunque sería deseable utilizar promotores específicos que sólo estuvieran activos en condiciones de estrés, éstos aún no son asequibles, por lo que puede haber efectos secundarios no deseables derivados de mantener en la planta en todo momento (de forma constitutiva) la expresión de genes de emergencia. Se necesitaría saber cómo detectan el estrés las raíces de la planta y qué tipo de

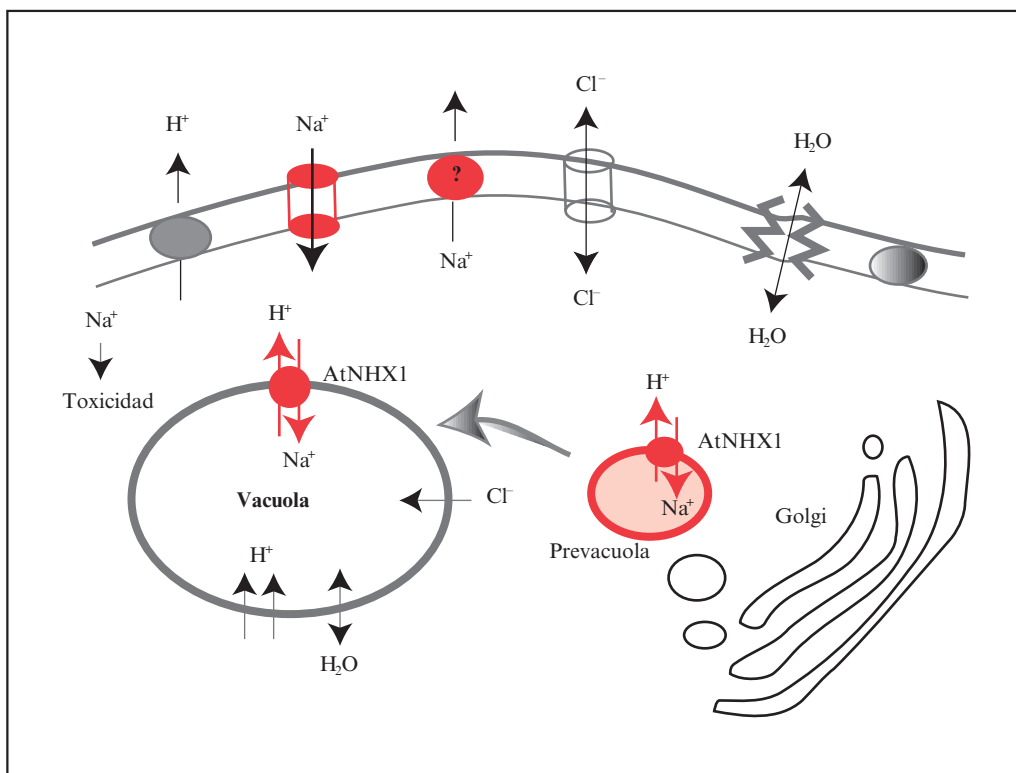


Figura 9.14. Proteínas transportadoras de iones implicadas en la tolerancia a la sal en *Arabidopsis*. La antiportadora sodio/protón AtNHX1 se localiza en las membranas de la gran vacuola y prevacuolas.

señales se transmiten para inducir la expresión de genes de tolerancia sólo en los momentos necesarios.

Plantas ricas en aminoácidos o vitaminas

A escala mundial, hay aproximadamente medio millón de niños de países subdesarrollados o en vías de desarrollo que tienen deficiencias en la vitamina A, lo que les puede acarrear ceguera. La base nutricional de estos niños es el arroz, que carece de vitamina A. Se ha obtenido arroz transgénico que produce provitamina A en el endospermo. Para ello se han transformado embriones inmaduros de arroz cultivados con *Agrobacterium* portador en el ADN T de dos genes suficientes y necesarios para desencadenar la producción de

provitamina A (β -caroteno) en el endospermo. Unos 800 embriones de arroz se transformaron en este experimento, para obtener 50 plantas transgénicas. El ADNc del gen *sintasa de phitoene* (*psy*) del narciso (*Narcissus pseudonarcissus*) se puso bajo el control del promotor específico del endospermo de la glutelina, y el gen de la *desaturasa de phitoene* (*crtI*) de la bacteria *Erwinia uredovora*, bajo el promotor constitutivo 35S del virus del mosaico de la coliflor (figura 9.15). El arroz de estas plantas posee un característico color amarillento que les ha valido el apodo de «arroz dorado». El consumo de 75 gramos de arroz dorado al día es suficiente para eliminar la carencia de provitamina A. Hay una variedad de arroz transgénico que además de provitamina A es rico en hierro y se comercializa en India, Vietnam y Filipinas.

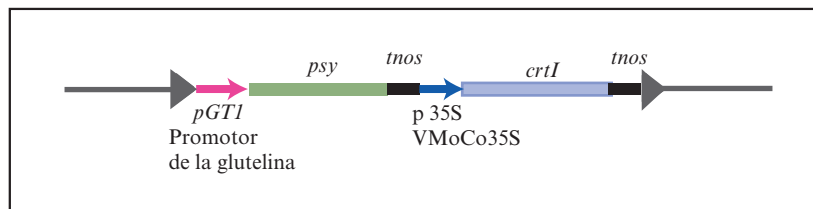


Figura 9.15. ADN T portador de los dos genes suficientes y necesarios para desencadenar en el endospermo del arroz la síntesis de provitamina A (β -caroteno). Entre los extremos izquierdo y derecho encontramos el promotor específico del endospermo de la glutelina (pGt1), el ADNc del gen *sintasa de fitoene* (*psy*), la secuencia de terminación de la transcripción de *nos* (nopalina sintasa), el gen de la *desaturasa de fitoene* (*crtI*) de la bacteria *Erwinia uredovora* bajo el promotor constitutivo 35S del virus del mosaico de la coliflor y de nuevo el terminador *nos*.

Otros ejemplos de mejoras nutricionales en cosechas transgénicas son: cereales ricos en lisina y treonina, leguminosas ricas en metionina o crucíferas ricas en vitaminas A y E.

Plantas como fábricas de plásticos, anticuerpos o vacunas

Las plantas transgénicas constituyen una alternativa barata y atractiva a los sistemas microbianos de producción de biomoléculas de interés. Las ventajas de la célula vegetal frente a la bacteriana son múltiples: los sistemas microbianos no realizan eficazmente las modificaciones postraduccionales de muchas proteínas eucarióticas; tampoco tienen una gran capacidad de almacenamiento del producto, siendo frecuente la formación de cuerpos de inclusión insolubles que han de ser solubilizados y replegados con una cierta pérdida de actividad; los fermentadores de carácter industrial son caros y requieren medios de cultivo especiales que se pueden contaminar con facilidad; por último, el uso de plantas no provoca tanto rechazo social como lo haría el uso de animales transgénicos.

Los plásticos no suelen ser biodegradables y puede ser problemática su destrucción; la producción de plásticos biodegradables en plantas transgénicas se ve como una alternativa ecológica a los plásticos que polucionan. Por otra parte, la expresión de proteínas heterólogas en vegetales que se usan en el consumo humano o de animales permite su ingestión oral o uso tópico sin necesidad de purificarlas.

Se llama «biofarma» a la producción de fármacos en plantas. Se pueden producir tanto anticuerpos (planticuerpos) o antígenos en plantas; estos últimos servirán para producir vacunas contra varias enfermedades.

Producción de plásticos biodegradables

La sociedad es hoy día muy sensible a la contaminación creada por los plásticos no degradables que contaminan campos y ciudades. Por el contrario, los plásticos biodegradables están constituidos total o parcialmente por material destruible por hidrólisis no-enzimática o por la acción de enzimas liberadas por microorganismos.

El poli-(R)-3-hidroxibutirato (PHB) y otros polihidroxialcanoatos son poliésteres alifáticos con propiedades termoplásticas totalmente biodegradables. La bacteria *Alcaligenes eutrophus* y especies afines producen estos polímeros como material de reserva de carbono cuando se les crece en exceso de este elemento y defecto de otro componente nutricional esencial como el nitrógeno o el fósforo. PHB se sintetiza como gránulos y puede llegar a constituir el 80% del peso seco de estos microorganismos. La síntesis tiene lugar por la acción consecutiva de tres enzimas a partir de acetil-CoA: 3-cetotiolasa, acetoacetil-CoA reductasa y PHB sintasa, codificadas por los tres genes: *PhbA*, *PhbB* y *PhbC*, respectivamente. Por otra parte, una gran variedad de microorganismos pueden hidrolizar PHB a monómeros y metabolizar éstos como fuente de carbono. El coste, no obstante, es

superior en la producción biológica al de la síntesis química. Las plantas nos permiten combinar las ventajas de la síntesis biológica a precios de síntesis química. Se han conseguido plantas transgénicas de *Arabidopsis* que expresan las tres enzimas en plastos y acumulan grandes niveles de PHB sin deterioro aparente en el crecimiento de la planta o en la producción de semillas. Se utilizó el sistema binario de plásmidos Ti con la construcción de tres mini T que llevan cada uno de los genes bajo el control del promotor 35S del CaMV. En principio, se establecieron tres líneas de *Arabidopsis* transformadas cada una con uno de los genes de la ruta; posteriormente, por una serie de cruces, se han obtenido líneas de plantas que producen las tres enzimas a niveles altos.

Arabidopsis no es una planta con interés comercial, pero es un buen sistema modelo para demostrar la viabilidad del método y detectar los posibles efectos secundarios que pudiera ocasionar en cuanto al crecimiento y desarrollo de la planta. La soja, la colza o el girasol son las especies agrícolas que podrían transformarse más fácilmente en productoras de plásticos, ya que sus semillas son ricas en aceites y tanto el aceite como el PHB son derivados del acetil-CoA. En estas plantas se podría bloquear la acetil-CoA carboxilasa con construcciones antisentido, para derivar todo el acetil-CoA hacia la síntesis del plástico, impidiendo la elaboración de ácidos grasos que se acumulan en las semillas.

PHB no es el homopolímero con las mejores propiedades físicas para su uso en la elaboración de utensilios de plástico, por lo que hay mucho interés en extender esta tecnología ya desarrollada a la producción de otros polímeros.

Un copolímero interesante es el PHBV, poly(3-hidroxibutirato-co-3-hidroxivalerato), que, al incorporar una pequeña cantidad de 3-hidroxivalerato al polímero PHB, mejora su flexibilidad haciéndolo menos quebradizo. Este copolímero se ha podido sintetizar en las hojas de *Arabidopsis thaliana* y en las semillas de *Brassica napus*. El proceso requiere que la planta sintetice grandes cantidades de ambos, acetil-CoA y propionil-CoA. Pero mientras que el acetil-CoA es muy abundante en los plastos de muchas plantas olea-

ginasas como precursor de ácidos grasos, el propionil-CoA se encuentra en muy pequeñas cantidades en peroxisomas como intermediario en la degradación de aminoácidos. Se hizo necesario modificar la vía biosintética de la isoleucina para generar un depósito de propionil-CoA en el plastido. La estrategia consistió en incrementar la concentración de 2-cetobutirato, sustrato convertible en propionil-CoA, por la piruvato deshidrogenasa del plastido (figura 9.16). La producción de grandes cantidades de 2-cetobutirato se consiguió a su vez expresando en exceso la enzima treonín-desaminasa de *E. coli* en la planta. Este experimento representa una de las manipulaciones de ingeniería genética y transferencia génica más complejas que se han llevado a cabo en plantas, necesitando la expresión de cuatro genes bacterianos y la modificación de dos vías metabólicas (biosíntesis de ácidos grasos y aminoácidos). La cantidad de copolímero sintetizado en *Arabidopsis* sólo representó el 2,5% del peso seco de la planta, que compara mal con el 14% de PHB que se puede obtener, pero no deja de ser una primera aproximación a un reto metabólico y biotecnológico para la producción de bioplásticos.

Producción de anticuerpos (inmunización pasiva)

Los anticuerpos se componen de dos cadenas pesadas y dos ligeras unidas covalentemente por puentes disulfuro; dentro de cada cadena hay una región constante y otra variable dando diversidad a su estructura final.

Se han obtenido plantas transgénicas que sintetizan anticuerpos (planticuerpos) funcionales complejos. Estas plantas usan las mismas vías de ensamblaje para las cadenas pesadas y ligeras que las usadas en mamíferos, utilizando los mismos péptidos señal y chaperonas para el plegamiento. También, como en mamíferos, los planticuerpos se excretan después del procesamiento y ensamblaje final. La única diferencia está en el patrón de glicosilación, ya que los anticuerpos producidos en plantas incorporan residuos de xilosa y de fucosa, ausentes en los anticuerpos naturales, y carecen de ácido N-acetil neuramínico, que es el

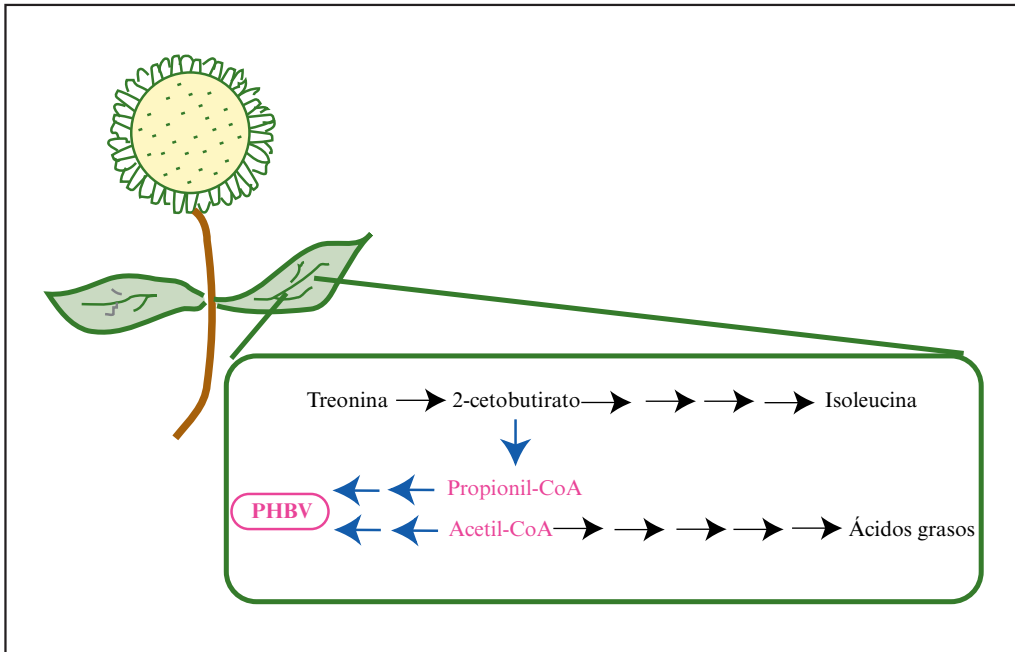


Figura 9.16. Síntesis de PHBV en una planta transgénica (girasol) por modificación de las vías metabólicas de biosíntesis de isoleucina y de ácidos grasos.

azúcar terminal predominante en los glicanos complejos de mamíferos. Sin embargo, se están desarrollando diferentes plantas transgénicas capaces de mimetizar los patrones de glicosilación de mamíferos.

En plantas de tabaco se ha podido sintetizar un anticuerpo monoclonal recombinante que reconoce la proteína de adhesión de *Streptococcus mutans* impidiendo la colonización de la bacteria y el desarrollo de caries en dentaduras de monos y hombres. El anticuerpo (SIgA/G) lleva la cadena pesada de ratón con las regiones $\gamma 1$ y α , la cadena ligera κ y la cadena de unión J también de ratón, mientras que el componente secretor (SC) es de conejo. Los genes correspondientes a las cadenas pesadas y ligeras, a la cadena J y al componente secretor se introdujeron en plantas de tabaco por separado. Por polinización cruzada se obtuvieron plantas que expresaban simultáneamente todos los componentes y producían anticuerpos funcionales. El anticuerpo se purificó a partir de plantas de

tabaco adultas por precipitación con sulfato amónico y cromatografía de afinidad con la proteína G (aglutinina de *Streptococcus* de gran afinidad por IgG). La producción fue de entre 10 y 80 mg/kg de peso de material vegetal fresco. La comparación entre este anticuerpo de origen vegetal (SIgA/G) y el parental de origen murino (IgG) reveló una constante de afinidad similar para ambos anticuerpos, pero el vegetal sobrevivió durante más tiempo (3 días) que el animal (1 día) en la cavidad oral humana dando lugar a una protección bucal contra caries de, al menos, 4 meses.

Vacunas (inmunización activa)

¿Podrá la gente vacunarse tomando un plátano o un tomate producido por una planta transgénica?

La mayoría de las vacunas tienen como componente activo una proteína capaz de provocar una respuesta inmune en animales. La vacuna

puede ser administrada por inyección o por vía oral, siendo esta última alternativa la más cómoda y barata, aunque habitualmente la menos eficaz. Las vacunas orales actúan estimulando el sistema inmune a través del tejido linfoide de la garganta y las mucosas del sistema gástrico.

La producción de vacunas en plantas tiene algunas ventajas con respecto a los métodos habituales. Son más baratas, lo cual puede ser muy importante en la vacunación de gentes en países subdesarrollados, que a menudo no pueden comprarlas. Además, elimina el miedo a contaminaciones con virus animales que se puede producir cuando se obtienen de células de mamíferos. Los virus de plantas no infectan al hombre.

Hay muchas enfermedades gastrointestinales, como el cólera y otras infecciones que cursan con diarreas, producidas por las toxinas que liberan algunas cepas de *Escherichia coli*, *Shigella* o *Salmonella*, siendo las diarreas bacterianas la primera causa de mortandad infantil en países pobres.

La enterotoxina de *Escherichia coli* consta de 6 subunidades: una proteína enzimáticamente activa (LT-A) que se introduce en las células epiteliales de la garganta y provoca una serie de cambios metabólicos que conducen a la deshidratación celular, y cinco proteínas idénticas, enzimáticamente inactivas (LT-B), que se unen a los gangliosidos de las membranas de las células epiteliales. Cuando estas últimas proteínas se facilitan oralmente a ratones o humanos, se provoca una fuerte reacción inmunológica de rechazo sin síntomas de la enfermedad. El gen bacteriano (o sintético) de LT-B optimizado para su expresión en el sistema vegetal (codones de uso más frecuente), y bajo el promotor 35S del CaMV, se introdujo en células de patata de las que se obtuvieron plantas transgénicas. Las patatas obtenidas de las plantas

se dieron a comer (crudas) a ratones y a 14 personas voluntarias. En todos los casos se observó una clara respuesta inmunológica con anticuerpos en las mucosas del aparato digestivo.

El ADN correspondiente al antígeno de superficie del virus de la hepatitis B se ha incorporado a plantas de tabaco, patata, plátano, lechuga y tomate por transformación mediada por *Agrobacterium*. Cuando estos productos se han dado a comer a ratones, produjeron en todos los casos una respuesta inmune. Las lechugas transgénicas fueron ingeridas por tres voluntarios adultos, pero sólo dos mostraron una respuesta inmunológica.

El virus del papiloma humano causa cáncer cervical en mujeres y ya existen comercialmente vacunas para evitar este mal. Sin embargo, los países en desarrollo no tienen acceso a estas vacunas por su elevado precio. Las plantas transgénicas son la alternativa idónea para estos casos, la proteína L1 del virus ha sido generada en plantas de patata y tabaco. Las patatas transgénicas han sido capaces de inmunizar animales de experimentación.

Respecto a nuestra pregunta inicial, no parece probable que se permita tomar plátanos, tomates o lechugas transgénicas como vacunas, más que nada por la posible confusión que se podría crear entre la población, al ser aparentemente indistinguibles de los productos naturales.

Las plantas transgénicas están devolviendo el protagonismo que un día tuvieron estos organismos, como productoras de compuestos medicinales, y que más tarde les fue arrebatado por la síntesis química.

Sin duda, las plantas transgénicas no van a resolver todos los grandes problemas que en la actualidad tiene la humanidad, pero pueden ayudar a paliar algunos.

LECTURAS COMPLEMENTARIAS

Role of transgenic plants in agriculture and biopharming, **Ahmad, P., Ashraf, M., Younis, M.** et al. *Biotechnology Advances*, **30**: 524-540. 2012.

Índice analítico

3-cetotiolasa, 280, 281, 285, 287, 291, 292
35S del virus del mosaico de la coliflor, 280, 281, 285, 287, 291, 292
5' bromodesoxiuridina, 212
5 bromouracilo, 176
6-tioguanina, 213
 λ gt10, 129
 λ gt22, 129

α -acetolactato descarboxilasa, 197
AAV, 67, 68, 69
AAVS1, 68, 69
acc deaminasa, 285
ACC (aminociclopropano-1-carboxilato), 285
ACC oxidasa, 286
accs, 285
Acetil-CoA, 292, 293, 294
Acetilación de la cromatina, 283
Acetoacetil-CoA reductasa, 292
Aciclovir, 183, 213, 238
Ácido desoxirribonucleico, 17, 19
Ácido micofenólico,
Ácido nitroso, 176
Ácidos nucleicos, 91, 93, 112, 151, 172, 231
ACL, 236, 237
Activación transcripcional, 283
ADA, 209, 233, 234, 235, 236, 252
ada, 234
Adaptador de policlonaje, 126
Adaptador sintético, 180
Adenina, 19, 20, 21, 31, 35, 48, 107, 172, 173, 175, 176, 199
Adenoasociados, 67, 214, 224, 226, 229, 236, 237
Adenosina, 112, 233, 234
Adenovirales, 240
Adenovirus, 51, 69, 214, 222, 223, 224, 226, 229, 232, 233, 240, 241
Adenovirus como vectores, 224
Adición de la diana de restricción para la enzima, 159

Adición de secuencias, 158, 159, 198
ADN, 11, 13, 17-21, 23-35, 38, 39, 41-51, 54-63, 67, 69, 70, 76, 78, 80, 81, 82, 85-104, 106-117, 119-121, 123-135, 137-161, 163-165, 167, 168, 171-173, 175-180, 182, 186-188, 190, 191, 193, 194, 195, 198, 199, 200, 202, 203, 206, 209-215, 218, 222, 224, 225, 227, 229, 230, 231, 238, 240, 241, 244, 246, 248, 249, 251, 252, 256, 258, 269-276, 278-283, 287-289, 291, 292, 295
ADN alfoide, 230, 231
ADN de banda simple-ARN, 131, 135, 146, 155, 156
ADN desnaturalizado, 24
ADN desnudo, 47, 88, 209, 214, 282
ADN genómico, 27, 107, 138, 154, 157, 163, 165
ADN mitocondrial, 165, 175
ADN nativo,
ADN pol I, 91, 92
ADN polimerasa, 17, 27, 89-93, 104, 109, 125, 137, 140, 151, 153-155, 173
ADN polimerasa del bacteriófago λ , 91
ADN polimerasa I de *E. coli* (holoenzima), 90
ADN polimerasas, 18, 21, 90, 191
ADN recombinante, 85, 123, 124, 131, 138, 151, 195, 215, 270, 287
ADN renaturalizado, 25
ADN satélite, 26, 230, 288
ADN T, 186, 270-273, 275, 276, 279, 280, 287, 289, 291, 292
ADN telomérico, 230
ADN transferido, 50, 151, 171
ADN viral, 131, 206, 222
ADNasa I, 30-34, 42, 47, 86, 88, 91
ADNc, 77-79, 91, 93, 113, 119, 124, 129, 134, 135, 137, 138, 140-142, 144, 149, 153-155, 159, 202-204, 218, 219, 221, 253-255, 261, 262, 276, 277, 280, 287, 288, 291, 292
Agente mutagénico, 167, 176
Agente terapéutico, 209
Agentes intercalantes, 176
Agentes mutagénicos, 175-177, 190, 191

- Agrobacterium*, 270, 272, 273, 275, 276, 279-282, 287, 291, 295
Agrobacterium rhizogenes, 273
Agrobacterium tumefaciens, 270, 276, 280
Alcaligenes eutrophus, 292
 Alelo, 33, 107, 111, 162, 164, 167, 187, 194, 199, 241
 Alfalfa, 187-189
 Algodón, 282, 284, 288, 289
 Alteraciones, 43, 49, 107, 124, 149, 161, 167, 172, 180, 181, 188, 191, 209, 237, 238, 241, 242
 ALU, 27, 78, 159
Allium cepa, 282
 Amaurosis congénita de Leber, 209, 233, 236, 237
 Amida, 87
 Aminociclopropano-1-carboxilato sintasa, 285
 Aminoglicósido fosfo-ribosil-transferasa, 126
 Aminopterina, 212, 213
 Amnios, 246
 Ampicilina, 76, 125, 126, 131, 134, 148, 201, 202, 207, 214-216, 228
 Amplicones, 214, 227, 278
 Amplificación, 48, 85, 110, 119, 120, 123, 151-165, 180-184, 275
 Amplificación de ARNm vía ADNc, 153
 Amplificación del ADN, 123
 Amplificaciones génicas, 106
 Anemia falciforme, 107, 108
 Angiogénesis, 251
antennapedia (Ant), 43, 44
 Anillos de clonación, 158
 Animales transgénicos, 211, 243, 244, 251, 255, 292
 Animales transgénicos y clónicos, 11
 Anticuerpos, 119, 121, 129, 131, 148, 149, 152, 222, 277, 284, 292-295
 Anticuerpos (planticuerpos), 292, 293
 Anticuerpos monoclonales, 149
 Anticuerpos policlonales, 149
 Aplicación médica de la clonación, 261
 Aplicaciones de la PCR, 161
 Aplicaciones de la PCR en estudios evolutivos, 165
 Aplicaciones prácticas, 261, 284
Arabidopsis, 272, 279, 281-284, 289-291, 293
Arabidopsis thaliana, 39, 41, 269, 270, 277, 278, 293
 Aridez, 284, 289
 ARN, 18, 19, 21-24, 27, 31, 33-35, 37-39, 42-47, 62, 69, 77-82, 85, 86, 88-90, 93-95, 113, 116, 124-126, 135, 142, 147-149, 153, 157, 158, 172, 173, 186, 188, 194, 201, 218, 219, 221, 222, 230, 238, 240, 241, 269, 273-278, 280, 282, 283, 285, 287, 288
 ARN de interferencia, 23, 38, 62, 79-81, 89, 238, 240, 269, 283
 ARNi, 23, 24, 38, 39, 45, 80, 82, 240-242, 282, 283, 288, 289
 ARN mensajero, 45, 221
 ARNncs, 39
 ARN-PET, 46
 ARN-pi, 82
 ARN polimerasa, 31, 33, 34, 44, 94, 124, 147, 148, 201, 275, 280, 283, 288
 ARN polimerasas, 31, 33, 47, 90, 94, 95, 135
 ARN polimerasas de los bacteriófagos SP6, T7 y T3, 94
 ARN viral, 219, 276, 288
 ARNasa A (Ribonucleasa A), 89
 ARNasa H (Ribonucleasa H), 89
 ARNasa T barnasa, 285
 ARNasa T1, 86, 89
 ARNm, 19, 21, 23, 24, 27, 28, 34, 36, 37, 45, 93, 107, 113, 116, 119, 124, 138, 140-142, 144, 153, 154, 163, 171, 181, 221, 226, 239-241, 276, 283, 287
 ARNm antisentido, 240
 ARNt, 23, 33, 37, 45, 144-146, 149, 170, 171, 188, 189, 199, 204, 212, 219
 Arroz, 39, 41, 269, 270, 278, 281, 282, 284, 289, 291, 292
 ATG, 147, 204
 AUG, 37
 Autorradiografía, 104-106, 114, 116, 117, 142, 143, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300
 Auxiliar, 69, 134, 180, 181, 203, 204, 219, 223, 224, 226, 227, 244, 245, 272, 273
 Auxotróficas, 168, 195
 Ajuste, 22, 35, 36, 42, 45, 46, 218, 221, 225, 226
 Azúcares, 85, 194
 β-galactosidasa, 76, 124, 126, 129, 131, 148-150, 212, 213
 β-glucanasa, 198
 β-glucanos, 198
 β-glucosidasas, 339
 β-glucuronidasa (GUS), 212
 β-hemoglobina, 108, 227
 β-isopriomalatodehidrogenasa, 195
 β-lactamasa, 125, 148, 214
 β-lactoglobulina, 262
 BAC, 134
Bacillus amyloliquefaciens, 285
Bacillus thuringiensis, 289
 Bacterias, 11, 17, 18, 26, 28, 37, 40, 50, 51, 54, 59, 61, 70, 73, 77, 78, 80, 85, 89, 93, 96, 97, 101, 103, 123, 124, 126, 130, 131, 134, 135, 142, 143, 146, 148, 149, 159, 165, 169, 173, 175, 177, 185, 187, 188, 193, 197, 199, 204, 212-214, 216, 217, 220, 227, 229, 241, 271, 272, 284, 287, 289
 Bacterias sintéticas, 187, 199
 Bacteriófago, 57, 61, 63, 66, 74, 80, 81, 88, 90, 91, 97, 98, 103, 124, 127, 128, 129, 131, 133, 139, 148, 255, 280
 Bacteriófago *lambda* (fago λ), 61, 127, 132
 Bacteriófago M13, 124, 131, 133
 Bacteriófago P₁, 25
 Bacteriófago SP6, 94, 95, 126, 127, 135-137, 280
 Bacteriófago T3, 94, 95, 126, 134, 135, 204

- Bacteriófago T7, 57, 90, 91, 94, 95, 98, 126, 127, 134-137, 147, 148, 204, 280
 Bacteriófagos, 50, 61, 80, 85, 94, 96, 126, 129, 146, 169, 275
 Baculovirus, 206-208, 215, 228, 272
 Baculovirus como vectores, 206
 BamH I, 138
 BamHI, 98-100, 134-137, 178, 179, 200, 201, 204
 Banda satélite, 110
bar, 275, 279, 287
 Barnasa, 285
barstar, 285
 Bases, 17-25, 27, 28, 39, 40, 56, 61, 70, 74, 78, 85, 89, 93, 96, 98, 99, 101, 102, 114, 116, 117, 119, 120, 131, 138, 159, 165, 167, 173, 176, 178, 180, 199, 201, 204, 218, 230, 240, 271, 283, 288
 Basta, 275, 286, 287
 Begomovirus, 274, 275
 Berenjenas, 289
 Betaina, 289, 290
 Biología molecular, 11, 12, 42, 48, 85, 123, 161, 269, 284
 Biomedicina, 242, 261
 Biorreactores, 269
 Biotecnología, 48, 146, 161, 283
 Biotina, 91, 142, 159
bithorax, 43
 Blastocisto, 244, 246-248, 250, 251, 264, 265
 Blastómeros, 246, 248, 251, 258
 Bluescript de M13, 95
 Bombardeo de partículas, 282
 Bombardeo genético, 211
Bombix mori, 208
Bos taurus, 39, 41
Brassica napus, 293
 BrdU, 212, 213, 228, 229

 CAB, 129, 134, 135, 136, 146, 227
 CAB derivados de P1, 135
 Cabeza de martillo, 90, 225
 Cacahuete, 282, 288, 289
Caenorhabditis elegans, 41, 61, 80
 CAGE, 46
 Cajetín de activación de la proteína VP16, 216
 Cajetín de expresión, 160
 CAL, 195, 198, 199, 200, 231, 232
 Calva, 142, 143
 Callos, 278, 189, 282, 286
 Callos de trigo, 282
 CAM, 73, 230
 CaMV, 274, 287, 293, 295
 Cáncer, 113, 167, 188, 190, 191, 209, 215, 218, 232, 235, 237-242, 252, 260, 295
 Caña de azúcar, 282, 289

 CAR, 222, 223
 Caracterización del producto clonado, 150
 Carbenicilina, 279, 280
 Carboxipeptidasa, 205
 Caronte, 129-131
 Cas1, 81
 Cas3, 80, 81
Cassava africana, 274, 275
 CAT, 109, 276
 cat, 276
 Cationicos, 211
 Caulimovirus, 274
 Cebada, 198, 288
 Cebadores, 91, 134, 140, 142, 151-155, 162, 180, 182, 219, 253
 Cebo, 202, 203
 Célula huevo, 43, 243, 246, 256, 258, 260, 264, 265
 Células, 27, 80, 124, 126, 177, 178, 179, 184-186, 194, 209, 211-216, 219, 220, 223, 226, 227, 230, 231, 233, 234, 236, 238, 239, 238, 241, 243, 244, 246-248, 250, 251, 255, 260-262, 264, 266, 267, 295
 Células de la médula ósea, 234
 Células de mamíferos, 27, 124, 126, 177, 179, 184-186, 194, 209, 211-216, 219, 227, 239, 255, 295
 Células diferenciadas de mamíferos, 260
 Células empaquetadoras, 219, 220, 223, 226, 227
 Células en cultivo, 178, 213-216, 233, 234, 241, 243
 Células endoteliales, 262, 266, 267
 Células hepáticas, 266, 267
 Células HT1080 humanas, 230
 Células madre, 238, 243, 246-248, 251, 257, 261, 264-267
 Células madre iPS, 243, 266
 Células musculares, 264
 Células polares, 244
 Células troncales, 80, 231, 234, 236, 238, 239, 246-248, 250, 251, 266, 267
 Células troncales linfoides, 231, 234
 Células troncales totipotentes, 247, 250
 Células troncales totipotentes embrionarias, 251
 Centrómero, 199, 201, 230-232
 Centrómero del cromosoma Y humano, 230
 Cerdos, 252, 253, 258, 260-264
 Cerdos transgénicos, 252, 253, 262, 264
 Cerio/EDTA, 87, 187
 Cerveza, 193, 197
 Chips, 112
 Chips de ADN, 112
 ChIP-seq, 47
 cI, 62, 127-130, 147, 148
 CII, CIII, 62
 Cianuro de bromo, 150
 Ciclo celular, 205, 240, 258
 Ciclo lítico, 61-63, 67, 69, 74, 129, 135, 137, 138, 215
 Ciclo vital de *S. cerevisiae*, 194

- Ciclo vital del retrovirus, 218
 Científicas, 94
 CII, 62, 63, 130
 CIII, 62
 Círculo de 2 μ m, 201
 Círculo rodante, 127, 128, 133, 138, 140, 274
 Cirugía genómica, 11, 241, 242
 Citidina, 112
 Citomegalovirus, 215-217, 222, 226, 255
 Citoplasma receptor «universal», 258
 Citosina, 19-21, 168, 172, 173, 174, 176
 Citotoxicidad, 226
 Claveles transgénicos, 286
 Clonación, 11, 123, 124, 128, 132-134, 137, 140, 146, 151, 158, 161, 193-195, 198, 199, 202, 214, 257, 259-261, 264-267, 276, 277
 Clonación en animales, 261
 Clonación en fagómidos, 134
 Clonación en M13, 134
 Clonación en ratones, 258
 Clonación en seres humanos, 260, 261
 Clonación terapéutica, 11, 143, 157, 164-167
 Clonación en seres humanos, 260, 261
 Clonaciones de ovejas, 258
 Clonaciones en animales, 260
 Clonar, 85, 119, 129, 131, 135, 140, 158, 193, 199, 200, 202, 209, 275, 285
 Clones, 134, 137, 138, 143, 144, 150, 198, 204, 260, 261
 Clones contiguos, 144
 Clones de ADNc, 204
 Clones humanos, 260
 Clónicos, 11, 243, 255, 261, 264
 Clónico-transgénicos:
 Ovejas: 261
 Cerdos, 261-264
 Cloranfenicol, 125, 126, 135, 148, 195, 276
 CMP (o dCMP), 20
 Código, 19, 37, 110, 149, 188
 Código de barras individual, 110
 Código genético, 37, 149, 188
 Codón de iniciación, 147, 160, 276
 Codones, 37, 39, 48, 146, 149, 168, 181, 187, 188, 189, 204, 213, 289, 295
 Codones de uso más frecuente, 146, 149, 295
 Cointegración, 207, 228, 272
 Cointegración por recombinación homóloga en el gen, 207
 Cólera, 295
 Colina, 289, 290
 Colina deshidrogenasa, 289, 290
 Colina oxidasa, 290
 Colonias, 126, 134, 142, 143, 146, 177, 194, 199, 227, 229
 Colza, 284, 287, 289, 293
 Complejo de iniciación, 31, 33, 34, 37
 Complejo sinaptonémico, 60, 61
 Compuestos osmoprotectores, 290
 Concepción de plantas resistentes a herbicidas, 284
 Conejos, 258, 260
 Conjugación, 50, 51, 125, 188, 189, 270-272
 Construcción de genes quimera, 181
 Construcción de plantas resistentes a patógenos, 284
 Control del número de copias, 125
 Control del transgén en el espacio y en el tiempo, 255
 Control del transgén en el tiempo, 255
 Conversión génica, 59
 Corte y empalme, 35, 46
 Cos, 61, 62, 128, 131, 132, 135, 136, 138-140, 145, 215
 Cósmido, 132
 Cósmidos, 131, 144, 146
 Cot, 26, 28
 CPMV, 277
 Cre, *cre*, 63, 65, 66, 135-137, 255, 256
 Criminología, 110, 163
 CRISPR, 80, 81, 82, 89, 241
 Cro, *cro*, 62, 127-130
 Cromátidas, 50, 59, 60, 230
 Cromatina, 2, 28, 29-35, 42, 44, 45, 47, 59, 66, 78, 88, 116, 119, 121, 188, 191, 201, 215, 218, 258, 259, 283
 Cromatina activa, 24, 30-32, 42, 47
 Cromatina inactiva, 30
 Cromosoma Artificial de Levaduras, 195, 200
 Cromosomas, 17, 28, 30, 32, 42, 45, 49, 50, 52, 57, 60, 66, 69, 70, 124, 134, 143, 146, 157, 178, 185, 194, 195, 198, 199, 201, 214, 215, 227, 230-232, 244, 246, 255
 Cromosomas artificiales, 124, 134, 146, 185, 194, 195, 198, 199, 214, 215, 227, 230-232
 Cromosomas artificiales de bacterias, 134, 146, 185, 227
 Cromosomas artificiales de mamíferos, 215, 230-232
 Cromosomas politénicos, 32, 246
 Cruzamiento, 248
 CSC, 238
 Cuerpos de inclusión, 150, 206, 207, 292
 Cultivos de células vegetales, 278
 Curva de reasociación, 26
 DEAE dextrano, 210
 Dedos, 33, 45, 101, 186, 241, 260
 Dedos de zinc, 33, 45, 101, 186, 241
 Degeneración del código genético, 37
 Degradación de los ácidos nucleicos, 85, 86
 Degradación enzimática, 88
 Degradación química, 85
 Dehidrofolato reductasa, 13
 Elecciones, 18, 65, 66, 106, 109, 110, 163, 168, 170, 177-180, 184, 186, 196, 241, 274, 277, 278
 Elecciones en el flanco 3', 179
 Elecciones en el flanco 5', 179

- Deleciones en genes, 196
 Deleciones grandes, 106
 Depurinización, 85
 Derivados del herpes, 214, 226
 Derivados del plásmido de 2 µm, 195
 Desacetilación, 283
 Desaminaciones oxidativas espontáneas, 173
 Desarrollo embrionario, 44, 247, 264, 265
 Desnaturalización, 21, 24, 25, 104, 120, 143, 151-154
 Desnucleación, 255
 Desoxirribosa, 19, 20, 22, 86, 172, 173
dhfr, 13
 di-β-D-galactopiranosido fluoresceína, 214
 Diabetes, 209, 252
 Diagnóstico, 107, 108, 111, 117, 161, 163
 Diana, 48, 106, 107, 109, 158-160, 180, 182, 185, 197, 199, 252, 255, 262, 263
 Diana de reconocimiento, 48
 Diana de restricción, 106, 107, 109, 158-160, 180, 199
 Diana génica, 182, 185, 197, 252, 255, 262, 263
 Dianas, 97, 104, 107, 126, 128, 130, 131, 135, 138, 140, 182, 184, 185-187, 196, 215, 251, 269, 279
 Dianas de inserción, 126
 Dianas de restricción, 97, 104, 107, 126, 128, 130, 131, 135, 138, 140, 215
 Dianas génicas, 182, 184, 185-187, 196, 251, 269, 279
 Dicotiledóneas, 270, 274, 279, 281
 Dideóxidos, 116, 117, 133
 Digestión parcial, 105, 119, 137, 144, 145, 200
 Digestión total de ADN humano, 103
 Digestiones dobles, 106
 Digoxigenina, 91, 158
 Disgénesis híbrida P-M, 74
 Distrofia muscular, 161, 163, 209
Dmcl, 59
 Doxiciclina, 216, 255
Drosophila, 31, 32, 35, 36, 39, 41, 43-46, 50, 59, 61, 66, 70, 74, 77, 80, 82, 178, 244, 245
Drosophila melanogaster, 39, 41, 44, 59, 61, 70, 74, 80, 244
Dscam, 35, 36, 45
 Duplicación del ADN, 258
 DYZ1, 163

 E1A, 222-224
 E1B, 222, 224
 E2A, 222, 224
 E2B, 223, 224
 E3, 36, 46, 223
 E4, 36, 46, 223, 224, 241
E. coli, 39, 50, 54, 57, 61, 63, 70, 76, 88-91, 93, 94, 97, 98, 125, 126, 129-131, 133, 137, 138, 146-150, 177, 180, 181, 187-189, 195, 196, 199, 202, 212-214, 216, 219, 270, 272, 289, 290, 293
E. coli lacZ, 76, 213
 Eco R1, 141, 204
 EcoB, 102
 EcoK, 102
 Efecto de posición, 283
 Eficiencia, 63, 76, 87, 89, 90, 97, 101, 123, 129, 149, 151, 158-160, 176, 186, 193, 198, 199, 202, 204, 206, 210, 214, 215, 218, 231, 233, 234, 239, 241, 256, 259, 269, 282
 Eficiencia de clonaje, 199
 Eficiencia de transformación, 198, 202
 Elaboración de un mapa, 104
 Electroforesis, 103, 105, 106, 114, 116, 118, 151, 194, 199, 200.
 Electroforesis en campo pulsante, 194, 199, 200
 Electroporación, 135, 210, 263, 264, 265, 272, 281
 Electrottransferencia, 103
 Elemento de respuesta a la tetraciclina, 255
 Elemento funcional, 18
 Elemento móvil, 74, 82, 177
 Elemento P, 74, 244-246
 Elementos de respuesta, 33
 Elementos móviles, 27, 69, 70, 80, 168, 178
 EMBL, 129-131
 Embrioblastos, 244, 246-248, 251, 252, 255, 264
 Embriones de cerdo, 252
 Empalmes, 22, 27, 42, 45
 Empaquetamiento del ADN, 28, 29, 137
 Empaquetamiento del fago lambda, 128
 Empaquetamiento in vitro, 131, 132, 138, 215
 ENCODE, 39, 41, 48
 Endonucleasas, 80, 85, 88, 96, 99
 Endorribonucleasa, 89
 Enfermedades, 117, 161-163, 187, 209, 218, 221, 222, 231-237, 241-244, 252, 266, 267, 277, 292, 295
 Enfermedades genéticas, 209, 235, 241
 Enfermedades monogénicas, 11, 233
 Enterotoxina de *Escherichia coli*, 295
env, 77, 78, 177, 218-222, 238
 Enzima ACCS, 285
 Enzima arilsulfatasa, 235
 Enzima ARN polimerasa dependiente de ARN, 283
 Enzima de restricción, 27, 48, 98, 105-107, 109, 110, 112, 114, 115, 119, 139-141, 145, 154, 158-160, 179, 186, 187
 Enzima de restricción Fok I, 186, 187
 Enzima fotorreactivante, 176
 Enzima Klenow, 91
 Enzima luciferasa, 178, 214
 Enzima puromicín acetil transferasa, 212
 Enzima uracil-N glicosilasa, 173
 Enzimas de restricción, 85, 94, 96-99, 102-107, 123, 128, 132, 133, 138, 144, 241
 Enzimas de restricción de otros tipos, 102
 Enzimas de restricción de tipo I, 102

- Enzimas de restricción de tipo II, 98, 99, 144
 Enzimas de restricción de tipo III, 102
 Epigenética, 42
 Episoma, 50, 51, 131, 226
 Episoma F, 131
 Episomas citoplasmáticos, 214
 EPO, 252, 253
Epstein-Barr, 226, 227
 Erizo de mar, 40-42
 Escisión-reparación, 173
 Esclerodermia CREST, 230
 Espacio periplasmático, 146, 150, 194, 205
 Espinaca, 280, 289
 Espontáneas, 18, 70, 167, 168, 172, 173, 175, 176, 213
 Esporulación, 194
 Esquizofrenia, 209
 Estable, 17, 18, 25, 32, 44, 66, 78, 87, 119, 124, 135, 173, 201, 209, 210, 212-214, 230, 233, 237
 Esterilidad masculina, 284, 285
Estigmene acrea, 206, 208
 Estrategia de «salto», 144, 145
 Estructura del ADN, 21, 172
 Éticas, 261
 Etileno, 285, 286
 Eucromatina, 30
 Evolución, 49, 82, 167, 252, 260, 278
Ex vivo, 231, 234, 236, 239, 242
 exoma, 18
 Exones, 18, 27, 28, 34-36, 42, 45, 107, 159, 163, 167
 Exonucleasa, 54, 56, 86, 88-91, 93, 152, 154, 173, 178, 179
 Exonucleasa III de *E. coli*, 88
 Exonucleasas, 61, 66, 88
 Exorribonucleasas, 89, 283
 Exportación, 146
 Expresión, 34, 38, 113, 146, 193, 196, 198, 202, 218, 243, 244, 251, 255, 282
 Expresión de los genes clonados, 146
 Expresión de un gen, 34, 38, 113, 146, 193, 196, 243, 255
 Expresión del gen clonado, 218
 Expresión del inserto, 198, 202
 Expresión del transgén, 244, 251, 282
 Expresión en levaduras: sistema del doble híbrido, 202
 Extremos *cos* de λ , 62, 128
eyeless, 43
 F1, 74, 133-135, 203, 204, 216, 245, 249, 250
 Fabáceas y lamiláceas, 289
 Factor de crecimiento epidérmico, 206, 208
 Factor F, 14, 50
 Factor IX antihemofílico humano, 261, 262
 Factor IX humano, 261, 262
 Factores antiangiogénicos, 238
 Factores de fertilidad F, 50, 51, 135
 Factores de transcripción, 28-35, 38, 39, 42-45, 47, 88, 93, 94, 101, 117, 119, 226, 258, 266, 267, 283
 Fago, 49, 50, 61-64, 76, 88, 103, 124, 127, 128, 130, 131, 133, 134, 135, 138, 140, 145, 146, 148, 204
 Fago «auxiliar», 134, 204
 Fago fl, 133
 Fago filamentosos, 131, 133
 Fago λ , 49, 50, 61-66, 76, 88, 103, 124, 128, 130, 131, 134, 135, 138, 140, 145, 146, 148, 180
 extremos cohesivos *cos*, 61
 mapa genético, 127, 128, 134
 Fago lambda, 61, 131, 138
 Fago M13, 135, 180
 Fago P1, 61, 63, 65, 66, 135
 Fago P2, 131
 Fagómido, 131, 133, 134, 135
 Fagómido λ ZAP, 135
 Fagos atemperados, 61
 Fagos de la serie Tpar, 103
 FAIRE-sec, 47
 Falciforme, 107, 108
 Farmacología, 261
 fcIX/adenovirus, 51, 69, 214, 222-224, 226, 229, 232, 233, 240, 241
 fcVIII/retrovirus, 11, 27, 45, 69, 77, 78, 154, 156, 177, 214-216, 218-221, 229, 233-235, 238-240, 242, 248-250, 264, 273
 Fecundación, 49, 162, 163, 246-248, 256
 Fertilidad, 50, 51, 134, 135, 272, 284, 285
 Fibra de cromatina, 28, 29
 Fibroblastos, 173, 210, 224, 261-264, 266, 267
 Fibrosis quística, 143, 144, 209, 232
 FIH, 63-66
 Filamentos sexuales F, 131
 Fitohormonas, 271, 273, 278, 279
 Flag, 215, 216
Flavobacterium okeanokoites, 100
 Fluoresceína, 112, 116, 154, 158, 159, 214
 Fluorescencia, 142, 148, 151, 154, 155, 213, 214
 Forma replicativa de M13, 133
 Formaldehído, 25, 47, 119, 121
 Formamida, 25
 Formas alélicas, 167
 Formas polimórficas, 167
 Fos/fox, 251
 Fosfatasa ácida, 205
 Fosfatasa alcalina, 148, 150
 Fosfatasa alcalina de *E. coli*, 150
 Fosfatasas, 95, 96
 Fosfinotricin acetil transferasa, 275, 286, 287
 Fosfinotricina, 275, 286, 287
 Fosfonometil glicina, 286
 Fosforilaciones, 28, 149, 151, 193

- Fósiles, 165
 Fotorreactivación, 173
 Fragmentación o digestión de genomas, 119, 120, 173
 Fragmento Klenow de la ADN polimerasa I de *E. coli*, 91
 Frecuencia, 20, 37, 50, 70, 74, 77, 78, 90, 94, 98, 99, 129, 144, 150, 152, 163, 167-169, 172, 175, 184, 185, 190, 191, 195, 197, 198, 202, 210, 214, 222, 229, 231, 236, 238, 252, 269, 277, 279, 289
 Frecuencia mutacional espontánea, 175
 Frío, 284, 289
 Fundadores transgénicos, 248, 249
 Fusiones génicas, 148
- G0, 244-246, 248, 249, 251, 258, 259, 264, 265
 G1, 258, 264, 265, 282, 294
 G2, 258, 282
 G418, 126, 195, 197, 212, 231, 244, 261-263
gag, 77, 78, 108, 218-222, 238
 Gal 4, 126
 GAL1, 202, 204
 GAL10, 202, 204
 Galactosa, 202, 208, 262
Gallus gallus, 39, 41
 Ganadería, 261
 Ganciclovir, 183, 185, 212, 213, 231, 238, 239, 263
 Gc/retrovirus, 33, 89, 99, 176, 178, 179
 Gel de agarosa, 103, 104, 106, 112, 138
 Gel de poliacrilamida, 149
 Geles de agarosa, 105, 106, 151, 164
 Gemelos idénticos, 260
 Geminivirus, 274, 275
 Gen, 35, 36, 69, 73, 94, 101, 108, 129, 136, 142, 147, 148, 177, 182, 185, 195, 196, 197, 199, 202, 204-208, 212, 213, 214-217, 219-229, 231, 232, 234, 235, 236, 238, 251, 255, 262, 272, 274, 275, 277, 278, 281, 285, 286, 287
 Gen *acc*, 101, 285, 286
 Gen *arsa*, 235, 236
 Gen β -galactosidasa (lac Z) de *E. coli*, 127
 Gen *bar*, 275, 287
 Gen *cap*, 69, 225
 Gen de interés, 94, 142, 147, 148, 196, 197, 202, 204-208, 212, 214-217, 219-229, 231, 255, 274, 277, 278
 Gen de la β -hemoglobina, 108, 227
 Gen de la miosina, 35, 36
 Gen de la timidina quinasa, 229, 238
 Gen de resistencia, 73, 129, 148, 202, 216, 231, 287
 Gen de selección, 136, 177, 182, 185, 195, 197, 199, 212, 213, 219, 221, 222, 227, 231, 232, 234, 236, 238, 251, 262, 272, 281
 Gen de selección del inserto, 136
 Gen de selección puro, 177, 232, 234
- Gen del factor de crecimiento epidérmico humano, 206, 208
 Gen eucariótico, 65, 85, 138, 147, 149, 150, 158, 193, 202, 204, 228
 Gen *lac*, 124, 126, 129, 130, 133
 Gen *luc*, 179, 214
 Gen *neo*, 183, 185, 195, 197, 212, 244, 261, 262, 287
 Gen *gfp*, 213
 Gen *rep*, 69, 225
 Gen *rpe65*, 237
 Gen *sacB*, 136, 137
 Gen señal, 178, 179, 202, 206, 214, 216, 272, 287
 Gen señal *gus*, 287
 Gen señal *luc*, 179
 Genes, 146, 188, 211-213, 223, 224, 270, 271, 273, 277, 231, 284, 290
 Genes clonados, 146, 188
 Genes de selección, 211, 212, 231
 Genes osmoprotectores, 290
 Genes que producen esterilidad masculina, 284
 Genes que restauran la fertilidad a plantas estériles masculinas, 284
 Genes señal, 211-213, 277
 Genes tempranos, 223, 224
 Genes *vir*, 270, 271, 273
 Genética Molecular, 11
 Genéticamente idénticos, 243, 260
 Genoma, 27, 31, 33, 34, 35, 42, 45, 66, 69, 109, 138, 141, 146, 172, 173, 191, 224, 252, 284
 Genoma de Ad5, 224
 Genoma de mamíferos, 45, 141, 252
 Genoma humano, 27, 31, 33, 34, 35, 42, 45, 66, 69, 109, 138, 146, 172, 173, 191, 284
 Genómica, 11, 18, 24, 26, 45, 79, 138, 142, 218, 241, 242, 255, 260
 Genoteca, 137-139, 142-146
 Germinales, 74, 110, 168, 243-245, 249, 279
gfp, 206, 212, 213, 277
 Girasol, 280, 293, 294
 Glándula mamaria, 254, 255, 258-262
 Glicina-betaína, 289, 290
 Glicosilación, 193, 194, 204, 205, 208, 293, 294
 Glicosilación de *S. cerevisiae*, 205
 Glicosilaciones, 149, 151, 193
 Glifosato, 286
 Globodera, 288
 GMP (o dGMP), 175, 212, 213
 GTP, 20, 175
 Grado de pureza, 151
 Guanina, 19, 21, 23, 31, 48, 59, 172, 175, 176, 183, 213, 238
 Guanosina, 20, 112
 Guisante, 266, 282
 Gus, *gus*, 272
 gusano, 40, 41, 80, 206, 208

- Hansenula polymorpha*, 205
 HAT, 212
 Hemaglutinina del virus de la gripe, 205
 Hemicigóticos, 248, 249
 Hemofilia, 161, 162, 163
 Herbicida, 275, 286, 287
 Herbicida Basta, 275, 286, 287
 Herbicidas, 11, 269, 283, 284, 286, 287, 290
Herpes simplex, 183, 185, 212-214, 216, 217, 226, 238
 Heterocigótico, 248
 Heterocromatina, 18, 27, 30, 45, 82
Heterodera glicines, 288
 Heterodúplices, 51-55, 57-59
 Hfr F+, 50, 51
 Hibridación, 25, 26, 28, 39, 91, 104, 108, 112, 113, 125, 127, 135, 137, 142-144, 150, 151, 157, 161, 165, 146
 Hibridación *in situ*, 143, 157, 246
 Hidrólisis, 19, 85, 86, 102, 137, 173, 214, 292
 Hidrólisis de ésteres fosfóricos, 173
 Higromicina B, 212
 Hind II, 98
 Hind III, 98, 154, 202
 Hiper Cromicidad, 24
 Hipervariabilidad, 109
 Hipoxantina, 173, 176, 212, 213
 HIS3, 195, 202, 203
 Histona H1, 28, 215
 Histonas, 23, 24, 27-32, 42, 43, 142, 188
 Hiwi, 82
 HLA, 221, 223, 234
Homo sapiens, 39, 41
 Homocigosis de factores deletéreos, 284
 Homocigóticos transgénicos, 248
 Hormona del crecimiento, 251, 253
 Hormona del crecimiento de rata, 251
 Hormonas vegetales, 271
 Hospedador, 61-63, 70, 80, 97, 124, 129, 233
 Hospedadores, 97, 124, 222, 227, 274, 288
hprt, 213, 231, 232
 Huella de los fragmentos de restricción, 105, 107, 112, 113
 Huella genética, 109-111
 Huella genética individual, 109
 Huéspedes/hospedadores, 97, 124, 222, 227, 274, 288
 Huevo desnucleado, 259
 Huevos fecundados, 244, 246, 251-253, 255, 257
 Huevos irradiados, 257
 Huntington, 161
 huntintina, 161
 HXGPRT, 213
hygro, 212
 Identificación individual, 110
 Identificación de los productos clonados, 150
 IgG, 318, 294
 Imidazol glicerolfosfato IGP dehidratasa, 195
 IMP, 14
 Implantación, 146, 149, 162, 247, 257
In vivo, 85, 87, 88, 96, 123, 134, 135, 178, 180, 181, 204, 214, 215, 219, 223, 224, 228, 231, 236, 237, 240, 251, 266, 274, 275
 Individuos clónicos, 243, 255
 Individuos genéticamente idénticos, 243, 260
 Individuos transgénicos, 243, 245, 248
 Inducidas, 137, 176, 194, 267
 Infección, 69, 91, 97, 127, 169, 180, 188, 206, 208, 209, 215, 218, 219, 221, 222, 224, 226, 227, 234, 248, 249, 261, 270, 276, 277, 280, 282, 287, 288
 Infección de discos de hoja, 280
 Infiltración al vacío, 276, 289, 281
 Información genética, 17, 19, 20, 31, 173, 243, 260
 Ingeniería genética, 11, 48, 85, 92, 93, 123, 142, 146, 150, 193, 227, 269, 270, 283, 284, 293
 Inhibición de la producción de polen, 284
 Inhibición del glifosato en la vía metabólica, 286
 Iniciación de la transcripción, 31-35, 46
 Inmunización pasiva, 293
 Inmunoglobulina, 206
 Inmunodeficiencia combinada severa, 209, 233, 234, 252
 SCID-ADA, 209, 233, 234, 252
 SCID-X1, 209, 233, 234, 252
 Inmunoprecipitación de proteínas expresadas en bacterias, 149
 Insectos, 103, 193, 206, 208, 283, 284, 287, 288, 289
 Inserciones de elementos móviles, 168
 Inserto, 124, 126, 128, 130, 131, 133-138, 143, 144, 146, 180, 197, 198, 199, 201, 202, 204, 221, 226, 227, 229, 230, 231, 236, 272-274, 277-279, 283, 287
 Insulina, 140, 251
 Integración, 11, 49, 61-67, 69, 71, 77, 78, 80, 128, 129, 148, 183, 185, 195-198, 210, 218, 221, 225-227, 233, 235, 241, 246, 251, 271, 272, 274, 283
 Integración del elemento P, 246
 Interferón, 205, 208, 240
 Interferón a humano, 208
 Interleuquina 4 humana, 154
 Intrón, 35, 74, 107, 202, 225, 253, 254, 262, 263, 287
 Inversiones, 65, 66, 168, 177
 Invertasa, 205
 Investigación forense, 110, 163
 iPS, 243, 257, 266, 267
 iPS-H, 266, 267
 IPTG, 126, 134, 147, 148
 IRES, 219, 221, 222, 262, 263, 277
 Islas CpG, 30-32, 44, 45, 48, 144, 175
 Isoesquizómeras, 99, 100

- Isopropil-tio- β -D-galactosido, 126
 IX anticoagulante humano, 261, 262
- Kanamicina, 73, 125, 126, 136, 182, 183, 212, 244, 272, 279-281, 287, 289
Kluyveromyces lactis, 205, 206
 KO, 186, 187, 234, 251, 252, 255, 262, 264, 284
Knock-out, 251
Knock-out tejido-específico, 255
- λ ZAP, 134, 135, 203, 204
 λ ZAP pEL URA3, 203, 204
 Lac, 124, 126, 127, 129, 130, 147, 148, 177, 212, 272
lac, 124, 126, 127, 129, 130, 147, 148, 177, 212, 272
lacI de *E. coli*, 126
 Lactoferrina humana, 154
lacZ de *E. coli*, 76, 213
 LAP1, LAP2, 227
 Latencia, 226, 227
 LDM, 235, 236
 Lentivirus, 218, 221, 236, 240, 242
 Lepidópteros, 289
 LEU2, 195, 201, 202
 Leucemias, 191, 235, 236, 238, 241
 Leucodistrofia metacromática, 233, 235, 236
 Levaduras, 11, 12, 28, 31, 40, 59, 61, 77, 103, 124, 137, 193-208, 230, 270
 Levansacarosa, 137.
 Lex A, 137
 Librería de ADNc, 202, 203
 Librería genómica, 142
 Librerías, 129, 137, 142, 199
 Librerías de ADNc, 142
 Librerías de genomas muy grandes, 199
 Librerías genómicas, 129, 137
 LCA, 236
 Ligasa, 58, 63, 91, 94-96, 119, 138-141, 180
 LMO2, 235
 Linamarasa, 239
 Linamarasa/linamarina, 239
 Linamarina, 239
 Línea transgénica, 248, 249, 252, 288
 Líneas estables, 248, 249
 Linfocitos T, 221, 233, 235
 Liposomas, 211, 214, 227, 230, 248, 250
 Liposomas de carácter aniónico, 211
 Lisis fágica, 142
 Lisogenia, 61, 62, 63
 Lisogénica, 62, 138
 Lisozima humana, 205
 Locus MAT, 194
 lox-P, 135, 256
 lox P del fago P1, 135
 LT-A, 295
 LT-B, 295
luc, 179, 214
 Luciferasa, 178, 180, 214
 Lugares de iniciación, 32, 35
 Luminómetro, 178, 214
- M, 24, 74, 78, 102, 110, 111, 112, 114, 115, 155, 164, 182, 183, 244, 245, 258, 282
 M13, 95, 124, 131, 133, 134, 135, 146, 147, 180, 181, 203, 204
 MAC, 230
 Maduración, 11, 22, 35, 80, 131, 149, 193, 194, 284-286
 Maduración controlada de frutos y flores, 284, 285
 Maduración de ARN, 149, 194
 Maíz, 39, 41, 165, 206, 269, 270, 274, 275, 278, 281, 284, 286-289
 Mamestra brassicae, 206
 Manduca sexta, 289
 Manipulación, 11, 85, 123, 149, 193, 206, 214, 218, 226, 229, 246, 250, 267, 269, 275, 276, 278, 282, 290
 Manipulación genética, 11, 85, 269, 278, 290
 Mapa de restricción, 104, 106
 Mapa del plásmido natural Ti de *Agrobacterium*, 270
 Mapa del vector CAB-P1, 136
 Mapas de restricción, 106
 Marcaje con fluorocromos, 116
 Mastrevirus, 274, 275
 Material genético, 17, 49, 50, 51, 54, 57, 103, 107, 165, 167, 211, 214, 226, 232, 256-258, 260, 261, 287
 Mcr, 102
 Medio HAT, 212
 Megacebador, 180, 181, 183, 184
 Mejora de plantas con interés agrícola, 284
 Meloidogine, 288
 Mensajero, 19, 23, 24, 28, 35, 37, 38, 39, 45, 78, 82, 113, 137, 140, 149, 150, 154, 201, 202, 221-223, 240, 254, 273
 Metaloproteasas, 188
 Metilación, 23, 24, 29, 30, 32, 33, 42, 44, 48, 79, 80, 98, 99, 102, 188, 191, 249, 283
 Metilación de ADN vía ARN, 283
 Metilasa, 24, 98, 140, 141
 Metionina inicial, 37, 146, 149, 150, 151, 193, 205
 Método de «infiltración al vacío», 279
 Método de sustracción, 142
 Método del fosfato cálcico, 210
 Método químico de secuenciación del ADN, 115
 Métodos de cultivo, 278
 Métodos de secuenciación denominados químico y enzimático, 114
 Métodos inmunológicos, 129, 148, 150

- Métodos para detectar mutágenos, 190
 Metotrexato, 14
 MHCI, 239
 Micro ARN, 23, 33, 38, 39, 42
 miARN, 23, 24, 35, 39, 45, 282, 283
 Microchips de ADN, 112, 113
 Microinyección, 211, 244, 246, 248, 251, 253, 254, 261, 265
 Microinyección de ADN, 244, 246
 Microinyección de ADN a huevos fecundados, 244
 Microproyectiles, 280, 281, 282, 286
 Microsatélites, 159
 Microtúbulos, 37, 38, 222
 Mili, 82
 Mini T, 272, 273, 293
 Minisatélites, 110, 111
 Miosina adulta, 35, 36
 Miosina embrionaria, 35, 36, 45
 Miwi, 82
 MLV, 218, 250
mob, 125, 272
 Modo de acción, 23, 60, 87-90, 92-94, 96, 99, 100
 Monocotiledóneas, 270, 274, 278
 Monogénicas, 11, 209, 233
 Mononucleosoma, 28
 Mórula, 246, 258, 259
 Mosca del vinagre, 39, 41, 80, 244
 Moscas, 74, 243, 244, 245, 274
 Movimiento, 35, 49, 106, 274, 275, 276, 277, 287
 Movimiento y la expansión viral, 275
 Muestras forenses, 163, 165
 Multiloci, 109, 111
 Mus 81, 59
Mus musculus, 41, 244, 246
 Mutación, 18, 43, 44, 73, 79, 103, 106-108, 129, 133, 134, 144-146, 167-171, 174, 178, 180-182, 186, 195, 199, 202, 209, 240, 241, 252
 Mutación *ade* 2, 200
 Mutación somática, 168
 Mutaciones,
 condicionales, 168
 en zonas reguladoras, 169
 letales, 168, 169
 silenciosas, 169
 supresiones, 170, 177
 viables, 168
 Mutaciones al azar, 177
 Mutaciones en el operón «lac», 177
 Mutaciones espontáneas, 18, 172, 175, 176, 213
 Mutaciones hfl, 129
 Mutaciones internas, 179
 Mutaciones puntuales, 18, 167-169, 180, 191
 Mutaciones resistentes, 169
 Mutaciones y cáncer, 188
 Mutagénesis, 11, 74, 82, 91, 133, 134, 149, 167, 177, 178, 181-183, 186, 187, 212, 213, 216, 229, 232, 233, 242
 Mutagénesis al azar, 177, 178
 Mutagénesis dirigida, 91, 133, 134, 149, 178, 213, 216
 Mutagénesis insercional, 74, 82, 177, 229, 232, 233, 242
 Mutagénesis mediante reacción en cadena de la polimerasa, 182
 Mutagénesis por recombinación homóloga, 182
 Mutagénesis por tijeras químicas, 187
 Mutágenos, 176
 Mutante, 129, 138, 167, 168, 178, 190, 195, 286
 Mutantes sensibles a la temperatura, 168
myc, 204, 205, 240
 N, 19, 20, 22, 31, 62, 85, 100, 102, 103, 108, 112, 127-130, 137, 138, 151, 163, 172-177, 290
 N-acetil- β -glucosidasa, 208
 N-glucosílicos, 85
 Nailon, 103, 104, 110, 112, 142, 149
 Naranja de acridina, 176
 Nemátodos, 287, 288
neo, 126, 182, 183, 185, 195-197, 211, 212, 222, 230, 231, 238, 244, 245, 251, 261-263, 272, 273, 287
 neo-centrómeros, 230
 Neomicin acetil fosfotransferasa, 212
 Neomicina, 125, 126, 182, 183, 195, 211, 212, 244, 272
 Neurona, 38
Nicotiana, 270, 272, 278, 279, 284, 288
Nicotiana tabacum, 270, 278, 279, 284
 Nitrosoguanidina, 176, 177
 Niveles de empaquetamiento del ADN, 28, 29
 Niveles de expresión de una proteína en *E. coli*, 150
 No permisivas, 215
 Nomenclatura, 98
 Nopalina sintetasa, 271, 287
nos, 271, 287, 292
 NotI, 99, 136, 137
 Nucleasa BAL31, 89
 Nucleasa S1, 88
 Nucleasas de dedos de zinc, 186, 241
 Núcleo, 19, 21, 23, 35, 79, 82, 183, 188, 202, 205, 206-208, 211, 215, 218, 221-223, 226, 227, 230, 231, 243, 255-265, 271, 272, 274, 275
 Nucleoide, 28
 Núcleos donantes, 256, 258
 Nucleósidos, 20, 183, 213
 Nucleosomas, 28, 29, 31, 34, 47, 210
 Nucleótico, 28, 29, 31, 34, 47, 210
 Nucleótido,
 Obsidad,
 Objetivos, 124, 214, 260

- ocs*, 271
 Octopina, 270
 Octopina sintetasa, 271
 Oligo-dC, 142
 Oligocebador, 155, 157-160, 164, 180-182, 184, 253, 263
 Oligonucleótido, 87, 113, 154, 160
 Oligos, 87, 151, 152, 153, 158, 160, 162, 182, 189
 Oncogén *c-myc*, 204
 Oncogén *fos*, 251
 Oncogén *ras*, 204
 Oncogenes, 188, 191, 270
 Oncosupresores, 188, 191, 238
 Operón, 126, 127, 148, 177, 216, 272
 Operón de la tetraciclina, 216
 Operón *lac*, 177
 Opina sintetasa, 271, 272
 Opinas, 270-273
 ORF, 74
 Órganos artificiales, 266
 Orígenes de replicación, 17, 18, 214, 216, 227, 230
Oryza sativa, 39
 Orotidín 5'fosfato decarboxilasa, 195
 Osmoprotectores, 289, 290
 Ovejas, 258, 262
 Ovejas clónicas, 261, 262
 Ovejas nodrizas, 259, 261, 262
 Ovejas transgénicas, 261, 262
 Ovejas transgénicas y clónicas, 261, 262
 Ovulación, 246, 247
 Ovulación en el ratón, 246
 Óxido de cerio, 87

 p53, 222, 223, 252
 pac, 135-137, 212
 Pacasa, 137
 Pacientes de glioblastoma, 239
 Palpitaciones, 25
 PAM, 277
Pan troglodites, 41
 Pancreática, 89
 Papaya, 282
 Parvovirus, 61, 67
 Paseo cromosómico, 142, 144-146
 Patata, 270, 274, 277, 287-290, 295
 Paternidad, 110-112
 Patrón de restricción de ADN, 111
 Pauta de lectura, 168-170
Pax6, 43, 44
 pBR322, 126, 201, 216
 PC, 275
 PCR, 85, 92, 119, 123, 140, 151-165, 180, 182-184, 248, 249, 253, 254, 263, 264, 279, 281
 PCR arbitraria, 160, 161
 PCR asimétrica, 153, 155, 156, 165
 PCR in situ, 153, 157, 158
 PCR inversa, 153, 154, 156
 PCR múltiple, 159, 160, 163
 PCR-aplicaciones
 diagnóstico precoz, 161, 163
 prediagnóstico de enfermedades, 161
 Pelos sexuales, 50
 Peptidasa, 150, 205
 Péptido señal, 150, 205
 Permeasas, 205
 Permisivas, 168, 215
 Petunia, 279, 280, 286, 288
 Petunias, 269, 275, 281, 283, 284
 Pez cebra, 40, 41, 66, 80, 185
 Pfu, 92, 152
 pH, 19, 21, 25, 26, 85, 86, 89, 91, 114, 115, 172, 222
 PHB, 292, 293
 PHB sintasa, 292
phbA, 292
phbB, 292
phbC, 292
 PHBV, 293, 294
Pichia pastoris, 205
 Pili, 50, 51, 131
 Pinturas, 284
 Pirimidinas, 19, 89, 93, 173, 212
 Pistolas génicas, 280, 281
 Piwi, 11, 13, 80, 82
 Piwi/ARNpi, 82
 Placenta, 246
Plantago, 290
 P. marítima, 290
 P. media, 290
 Planta de la colza, 284
 Planta de tabaco, 276, 294, 295
 Plantas como fábricas de plásticos, anticuerpos o vacunas, 292
 Plantas como fábricas de plásticos, pinturas, anticuerpos, encimas, 284
 Plantas de arroz, 278
 Plantas de maíz, 278, 286
 Plantas resistentes a hongos, 284
 Plantas resistentes a patógenos: virus, bacterias, hongos, nemátodos e insectos, 284, 287
 Plantas resistentes a virus, 287
 Plantas tolerantes, 284, 290
 Plantas tolerantes a las situaciones estresantes, 284
 Plantas transgénicas, 11, 269, 278, 279, 281-287, 289-295
 alfalfa, 287-289
 algodón, 282, 284, 287, 289
 Arabidopsis, 39, 41, 269, 270, 272, 277-279, 281-284, 289-291, 293
 arándanos, 289

- arroz, 39, 41, 269, 270, 278, 281, 282, 284, 289, 291, 292
- brécol, 289
- cacahuete, 282, 288, 289
- caña de azúcar, 282, 289
- colza, 284, 287, 289, 293
- maíz, 39, 41, 165, 206, 269, 270, 274, 275, 278, 281, 284, 286-289
- manzano, 289
- nogal, 289
- patata, 270, 274, 277, 287-290, 295
- peral, 289
- repollo, 289
- soja, 282, 284, 286-289, 293
- tomate, 274, 275, 279, 284, 287, 288, 289, 294, 295
- vid, 289
- Plantas transgénicas resistentes, 287
- Plásmido de expresión de levadura URA3, 203
- Plásmido de replicación en levaduras, 14
- Plásmido de transferencia, 206, 208
- Plásmido etiqueta, 144-146
- Plásmido pBR, 251
- Plásmido Ri, 273
- Plásmido Ti, 270-272, 276, 280
- Plásmidos, 50, 51, 70, 80, 124-126, 129, 131, 134, 144, 166, 159, 214, 216, 221, 226, 227, 270-272, 286, 293
- Plásmidos artificiales de mamíferos, 227
- Plásmidos retrovirales, 220
- Plásmidos Ti, 270-272, 286, 293
- Plásticos, 11, 269, 284, 292
- Plásticos biodegradables, 11, 269, 292
- PM, 62, 275, 276
- P_M, 62
- pMB9, 126
- pMGH, 251
- pol, 54, 77, 78, 89-92, 218-222, 238
- poli A, 23, 218, 225, 253
- poli[R-(–)-3-hidroxibutirato], 292
- Polibreno-DMSO, 210
- Policlonaje, 94, 95, 124, 126, 129-132, 134, 196, 197, 204, 216, 272, 273
- Poliedrina, 206-208, 215
- Polihidroxialcanoatos, 292
- Polimerasa I, 89-91, 93, 104, 125, 140
- Polimerasa II, 200, 226, 283
- Polimerasa III, 240
- Polimerasa Taq, 92
- Polimerasas para uso comercial, 208
- Polimorfismos del tamaño de fragmentos de restricción, 107
- Polimorfismos en secuencias, 107
- Populus*, 282
- Producción de anticuerpos, 222, 293
- Producto, 17, 18, 26, 35, 43, 45, 62, 63, 70, 74, 85, 91, 123, 125, 129, 133, 136, 137, 143, 148, 150, 151, 163, 167-170, 177, 180, 181, 183, 184, 197, 198, 200, 202, 203, 206, 213, 214, 216, 222, 231, 238, 255, 271, 272, 277, 287, 288, 292
- Productoras, 220, 231, 239, 249, 250, 283, 293, 295
- Profago, 61-64, 102, 127, 138
- Proflavina, 176
- Promotores:
 - Estrechos, 32
 - Anchos, 32
- Promotor 35S del virus del mosaico de la coliflor, 285, 287
- Promotor de la β -lactoglobulina, 262
- Promotor de la metalotioneína, 251
- Promotor de la tetraciclina, 216
- Promotor específico de expresión en el floema, 289
- Promotor específico de los sacos polínicos del estambre, 284
- Promotor específico de tejido, 255, 256
- Promotor eucariótico de citomegalovirus, 216, 255
- Promotor sintético de respuesta a la tetraciclina, 216
- Promotor Sp6, 95, 127, 136
- Promotor T3, 95, 134
- Promotor T7, 95, 134, 136
- Promotor/ARN polimerasa del fago T7, 147
- Promotores, 23, 28, 31, 32-35, 45, 62, 70, 90, 94, 95, 126, 127, 135, 137, 146, 177, 202-204, 214, 215, 222, 225, 227, 235, 236, 252, 271, 272, 275, 276, 280, 289, 290
- Promotores confinados al polen, 289
- Promotores de diseño, 215, 221, 222
- Promotores de latencia, 227
- Promotores específicos, 95, 289, 290
- Promotores inducidos por heridas, 289
- Promotores P_D, 62
- Promotores P_I, 62
- Pronúcleo femenino, 248, 253
- Pronúcleo masculino, 246, 248, 253
- Pronúcleos, 246, 251
- Propiedades, 25, 92, 136, 168, 222, 254, 271, 292, 293
- Proteína, 206, 207, 212, 213, 215, 230, 272-277, 287
- Proteína centromérica CENP-B, 230
- Proteína de la cápside, 215, 275
- Proteína de la cubierta, 274, 287
- Proteína de movimiento, 276, 277, 287
- Proteína fusión, 149, 217
- Proteína G, 294
- Proteína recombinante, 202
- Proteína verde, 206, 207, 212, 213, 272, 277
- Proteína verde fluorescente, 212, 213, 277
- Proteínas fusión, 129, 141, 146
- Proteínas mutantes quiméricas por PCR, 184
- Proteínas terapéuticas en la leche de animales domésticos, 261
- Proteínas Mer3, Msh4 y Msh5, 60
- Proteínas SYCP1, SYCP2 y SYCP3, 60
- Proteínas ZMM, 60

- Proteínas Zip: Zip1, Zip2, Zip3, Zip4, 60
 Protoplastos, 278, 279, 281
 Provirus, 78, 218, 219, 221, 249, 264
 pSC101, 126
 Pseudopartículas, 215
 pSP, 95, 127
 pTA-29, 184
 pTi, 270
 PTF, 107
 pUC, 125-127, 133, 134, 148, 204, 221
 Puentes de hidrógeno, 21, 24, 25
 Puntuales, 18, 167-169, 180, 184, 191
 Purinas, 19, 85, 93, 114, 172, 173, 212
 Puro, 54, 177, 195, 212, 219, 222, 231, 232, 234, 236-238, 251
 Puromicina acetil transferasa, 177
 Puromicina N-acetil transferasa, 177
 Puromicina, 177, 195, 212
- Quiasma, 50, 51, 54, 55, 57, 58, 63
 Quimera sintética, 186
 Quimera transactivadora, 216
 Quimeras, 99-102, 184, 186, 187, 199, 244, 246, 248, 250
 Quimeras ADN-ARN de hebra sencilla, 186
 Quimeras transgénicas, 244, 246, 248, 250
 Quinasa de timidina (tk), 228, 229, 238
 Quinasa de timidinas, 183, 212, 213, 226, 229
 Quinasas, 95, 96
- Rad 51, Rad 52, Rad55, Rad 57, 59
 Radiaciones ultravioleta, 103, 173
 Radiactividad, 151, 159
 Radicales libres de la molécula de oxígeno, 173
 Raíces de *Arabidopsis*, 282
 Rana, 255-257
Rana pipiens, 255
ras, 204, 205
 Ratón, 40, 41, 43-45, 60, 66, 80, 82, 93, 140, 206, 210, 230, 244, 246-251, 256-259, 266, 267, 294
 Ratón casero, 244, 246
 Ratón transgénico, 251
 Ratona nodriza, 250
 Ratones, 234, 236, 237, 238, 240, 243, 246, 251, 252, 255, 256, 258, 260, 264-266, 295
 Ratones «desnudos», 238
 Ratones gigantes, 251
 Ratones SCID, 252
 Ratones transgénicos, 236, 243, 251, 252, 256
 Rayos láser, 210
 Rayos X, 28, 43, 177, 178
 RBS, 68, 69
 RCP, 87
- Reacción en cadena de la polimerasa, 48, 85, 92, 123, 140, 151, 152, 157, 161, 181, 182
 Rec A, 54, 56, 57, 59, 76
 Rec BCD, 54, 56, 76
 Receptor de la acetil colina en el *Torpedo California*, 205
 Recombinación al azar: transposición, 69
 Recombinación específica de sitio:
 Cre/loxP, 255
 fago λ , 49, 50, 61-67, 76, 85, 88, 91, 95, 97-99, 103, 124, 127, 128, 130-142, 145-148, 180, 181, 204
 ϕ C31, 66
 Flp-FRT, 66
 AAV-cromosoma 19 humano, 68, 69
 Recombinación homóloga,
 Modelo de Holliday, 51, 52, 54
 Modelo de Meselson-Radding, 51, 54
 Encrucijada de Holliday, 51, 53, 55, 57, 58
 Recombinaciones homólogas, 74, 187, 278
 Recombinasa, 63, 66, 67, 68, 75, 136, 137, 148, 196, 201, 255
 Recombinasa Cre, 63, 137
 Redundancia funcional, 167, 152
 Redundancia génica, 167
 Región centromérica, 198, 230
 Región codificante, 35, 147, 154, 160, 272
 Región de policlonaje, 94, 95, 126, 129-131, 134, 196, 197, 204, 216, 272, 273
 Región hipervariable, 110, 112, 164, 165
 Región promotora, 27, 31, 32, 33, 44, 179, 180, 200, 261, 262
 Región promotora en levaduras, 200
 Región reguladora, 62, 82, 154, 178, 179, 227, 262
 Región subtelomérica, 161, 230
 Región TATA, 178
 Región *vir*, 273
 Regiones hipervariables dispersas, 111
 Reimplantación en madres nodrizas pseudopreñadas, 246
 Relleno, 130, 131, 136, 137
 Remolacha, 280, 284, 287, 288
 Remolachas transgénicas, 284, 287
 Renacuajos, 256, 257
 Renaturalización, 25, 26
 Reordenaciones, 43, 134, 167, 168, 185, 288
 Reordenaciones estructurales, 167
 Reparación por escisión, 173
 Replicación, 17, 18, 27, 28, 43, 50, 55, 59, 60, 66, 69, 72, 74, 76, 80, 98, 109, 124-127, 131, 133, 134, 140, 148, 158, 165, 173, 174, 177, 195, 198, 201, 202, 206, 213-216, 222, 223, 225-227, 229, 230, 238, 270-272, 274, 275, 278, 283
 Replicación del ADN, 18, 43, 50, 55, 98, 125, 173, 222, 225
 Replicón, 125, 136, 137
 Replicones, 50, 124, 125
 Represor, 23, 61, 62, 126, 127, 129, 147, 148, 216, 217, 255
 Represor de la tetraciclina, 255

- Resbalón de la polimerasa, 109
 Rescate del fago, 109
 Resistencia, 19, 73, 76, 125, 126, 131, 134-136, 148, 182, 195, 199, 201, 202, 207, 212, 214-216, 228, 231, 244, 246, 252, 255, 261-263, 272, 275, 281, 284, 286, 287-290
 Resistencia a bacterias y hongos, 284, 287
 Resistencia a *Heterodera*, 288
 Resistencia a insectos, 284, 288
 Resistencia a la ampicilina, 126, 215
 Resistencia a nemátodos, 288
 Resistencia a tetraciclina, 201, 255
 Resistencia a virus, 287
 Resistencia al cloranfenicol, 135, 148, 195
 Resistentes a herbicidas, 11, 269, 283, 284, 286, 287, 290
 Resolvasa, 57, 69, 74-76
 Respuesta inmune, 226, 233, 237, 239, 264, 295
 Respuesta inmunológica, 208, 224, 240, 295
 Restrictasa, 98
 Retro-PCR, 153
 Retrocruzamiento, 248-250
 Retrovirus, 11, 27, 45, 69, 77, 78, 154, 156, 177, 214-216, 218-221, 229, 233-235, 238-240, 242, 248-250, 264, 273
 Retrovirus endógenos, 264
 Retrovirus como vectores, 220
 Retrovirus defectivos, 77, 220, 238, 249
 Retrovirus de la leucemia murina, 218, 249, 250
 Retrovirus terapéutico, 234, 239
 Reversión, 70, 121, 170, 195
 Revertiente, 170
 Revertientes, 190
rho, 147
 Ribonucleasa P humana, 89
 Ribonucleótido reductasa, 226
 Ribosa, 19-21, 116, 172, 175
 Ribozimas, 23, 86, 89, 90
 RNA, 23, 82
 Rojo de Texas, 116
rop, 125, 148
 Rot, 28
Round up, 286
 RRBS, 48
 RT-PCR, 140, 153
 Ruptura de los enlaces glicosílicos, 172
 Rupturas cromosómicas, 177
 Ruv A, Ruv B, 57
 S, 23, 39, 62, 66, 78, 98, 102, 117, 128, 129, 130, 136, 139, 193, 194, 197, 199, 201, 202, 205, 206, 212, 223, 230, 258, 282, 286
 S-adenosil-metionina, 98
 S. pombe, 205, 230
 Saccharomyces cerevisiae, 39, 41, 59, 66, 193, 198, 199
 SalI, 99, 100, 127, 134, 136, 137, 204
 Salinidad, 284, 289, 290
 Salmonella, 94, 190, 270, 286, 295
 Salmonella typhimurium, 94, 190, 286
 Salmones, 253, 254
 SAM, 129
 Sam100, 129
 Sapo africano, 244
 Sapos, 256, 257
 Sau3A, 100, 131, 137-139, 145
 Sau3AI, 98, 99
 Schizosaccharomyces pombe, 198, 205
 Schwanniomyces occidentalis, 205, 206
 Secreción, 37, 150, 190, 193, 204, 205
 Secreción del producto, 150
 Secreción en levaduras, 205.
 Secuenasa, 90-92, 116, 117
 Secuencia *chi*, 54, 56
 Secuencia codificante, 147, 253
 Secuencia consenso de iniciación de empalme, 15
 Secuencia de activación, 202
 Secuencia de replicación autónoma, 198
 Secuencia de ruptura, 102
 Secuencia de unión al ribosoma, 147, 160
 Secuencia repetida, 80, 144, 146
 Secuenciación,
 método enzimático, 114, 117, 119, 133, 151
 método químico, 114-116, 187
 Secuenciación del ADN, 91, 114, 115, 178
 Secuenciación masiva o de nueva generación, 116
 Secuenciación por el método enzimático, 117
 Secuenciación de genomas completos, 119, 120
 Secuenciar, 114, 116, 119, 120, 123, 133, 134, 143, 155, 165, 180
 Secuencias muy repetidas, 26
 Secuencias de repetitividad intermedia, 26, 27
 Secuencias únicas, 26, 27, 80, 126, 218
 Secuencias *chi*, 54, 185
 Secuencias de iniciación de la traducción, 146
 Secuencias reguladoras, 43, 167, 214, 255
 Segregación de transgénicos en ratón, 249
 Selección, 212, 214, 215
 Selección a largo plazo, 212
 Selección de una colonia o calva específicas, 143
 Selección en células de mamíferos, 212
 Selección estable, 212, 214
 Selección transitoria, 214
 Semipermisivas, 215
 Señales para la terminación de la transcripción, 35
 Serotipo Ad 2, 222
 Serotipo Ad 4, 222
 Serotipo Ad 5, 222
 Serotipo Ad 7, 222

- Serotipos, 67, 222
 SfiI, 137, 144
 Shigella, 295
 Shine & Dalgarno, 147
 SIDA, 221, 264, 277
 SIE, 15
 Silenciado, 227, 283
 Silenciamiento, 23, 24, 44, 48, 80, 82, 240, 277, 282, 283, 288
 Silenciamiento génico, 48, 240, 282, 283
 Silenciamiento de transgenes, 288
 Silenciamiento de transposones, 82
 Silenciar, 241
 Síndrome de la inmunodeficiencia combinada severa, 252
 Síndrome de Wiskott-Aldrich, 235, 236
 Síntesis de ácidos nucleicos, 85, 90
 Síntesis de ADN, 54, 57, 59, 62, 69, 125, 127, 128, 222, 258
 Síntesis de ARN, 46, 94, 283
 Síntesis de proteínas, 46, 125, 126, 212
 Síntesis del ácido corísmico, 286
 Síntesis por iniciación al azar, 142
 Sistema binario de transferencia, 273
 Sistema de expresión transitorio, 215
 Sistema de restricción y modificación, 97
 Sistema del doble híbrido o del cebo, 202, 203
 Sistema inmune, 80, 96, 190, 221, 222, 295
 Sistemas asesino-suicidas condicionales, 238
 Sistemas de selección, 195
 SIT, 31, 32, 34, 201
 Situaciones estresantes, 284
 Soja, 282, 284, 286-289, 293
 Soja transgénica, 288
 Somáticas, 57, 74, 161, 168, 243
 Sonda,
 radiactiva, 104, 110, 143, 246
 Sondas
 fluorescentes, 104
 radiactivas, 104, 110
 unilocus, 111, 164
 Sondas alelo-específicas,
 SP6, 94, 95, 126, 127, 135-137, 280
 Spi+, 131
 Spo 11, 59
 Spodoptera frugiperda, 206
 SRA, 198, 199, 204
 SRA/CEN, 198, 199
 STE, 15
 Stem cells, 167
 Streptococcus, 294
 Streptococcus mutans, 294
 Subclonaje, 143, 144, 146
 Superratones, 251, 252
 Supresión por ARNt, 171
 Supresora, 170
 SUR, 147-149, 160
 SV40, 179, 203, 214-216, 229, 251, 253
sxl, 45, 46
 T7, 57, 90, 91, 94, 95, 98, 126, 127, 134-137, 147, 148, 204, 280
 Tabaco, 191, 270, 274, 276, 277, 279, 282, 286, 287-290, 294, 295
 tac, 147
 TALEN, 101, 102, 186, 187, 241
 Tallos de arándano, 282
 Taq, 90, 92, 151, 152
 Tardíos, 206, 224, 258
 tat, 216, 221
 TATA, 31, 32, 34-36, 178-180, 201
 Técnica de infiltración al vacío, 281
 Técnica Southern, 103
 Telomerasa, 230
 Telómeros, 27, 45, 194, 198, 199, 230
 Telómeros cromosómicos, 198
 Temperatura de fusión, 24, 25
 Terapia, 69, 209, 231, 232-235, 237-239, 241, 242
 Terapia génica, 11, 69, 209, 231, 233-235, 237-239, 241, 242
 Terapia génica del cáncer, 232, 237, 238
 Terapias génicas, 187, 208, 214, 219, 223, 224, 231, 236
 Terapias in vivo, 216, 223
 Terminación de la transcripción, 35, 202, 221, 222, 287, 292
 Terminaciones repetidas invertidas, 223
 Terminador de la transcripción, 147
 Terminal repetido largo, 221
 Terminales cohesivos, 62, 98, 124, 128, 138, 139
 Terminales planos o romos, 98
 Terminales repetidos e invertidos, 69
 Terminales repetidos largos, 78, 218, 221
 Terminator, 284
 Termociclador, 152
 Terneros, 260
 Test de Ames, 190
 Tetraciclín-acetil transferasa, 216
 Tetraciclina, 125, 126, 199, 201, 202, 215-217, 222, 255
 Tetrahymena piriformis, 199
 Tetrametilrodamina, 116
 Tf_i, 93
 Thermus aquaticus, 90, 92, 151
 Ti, 62, 270-273, 276, 280, 286, 287, 293
 Ti «desarmado», 272
 Tijeras químicas, 87, 187, 241
 Timidina, 112, 168, 213, 228, 229, 238
 Timidina monofosfato, 213, 229
 Timidina quinasa, 228, 229, 238
 Timina, 19-21, 31, 107, 172-174, 176, 212

- Tipo I, 54, 60, 89, 98, 102, 221, 223
 Tipo silvestre, 128, 131, 167, 177, 180, 195, 212, 228, 229, 275
 Tipos de mutaciones, 167, 191
TK/tk, 183, 185, 212, 123, 228, 229, 238, 251
Tm, 24, 25, 219
TMP, 212, 213, 229
TMV, 276, 277, 288
Tn5, 70, 73, 76, 126, 197, 212
 Tolerancia, 239, 240, 277, 284, 289-291
 Tomate, 274, 275, 279, 284, 287-289, 294, 295
 Tomates que maduran a voluntad, 269
 Totipotentes, 231, 247, 248, 250, 251, 257, 261, 264
 Toxinas, 289, 295
 Toxinas Bt, 289
tra, 45, 46, 50, 125, 270-273
 Traducción, 23, 24, 37-39, 45, 48, 62, 70, 79, 82, 146, 147, 149, 150, 170, 204, 221, 223, 277, 282, 283
 Traducción inhibida por el híbrido, 150
 Trampas de estimuladores o potenciadores, 33-35, 45, 180, 201
 Trampas de exones o del poli-A, 23
 Trampas de poli-A, 23
 Trampas de promotores, 177
 Trans-activador EBNA-1, 227
 Transcripción, 124, 137
 Transcripción del inserto, 124, 137
 Transcriptasa inversa, 77-79, 90, 93, 104, 113, 119, 137, 138, 140, 141, 153, 154, 218, 219, 221, 230, 274, 276
 Transducción, 129, 132, 136, 204, 209, 215, 239, 281
 Transfección, 179, 180, 209-212, 214, 215, 218, 220, 248, 250, 276, 281
 Transfección estable, 210
 Transfección por liposomas, 250
 Transfección transitoria, 209
 Transferasa terminal, 90, 93, 94, 142
 Transferencia a protoplastos, 281
 Transferencia de ADN, 50, 149, 273, 279
 Transferencia génica, 11, 12, 206, 209, 214, 222, 234, 244, 269, 270, 293
 Transferencia génica a células de mamíferos, 209
 Transferencia génica a plantas, 12, 269
 Transferencia génica en *Drosophila*, 244
 Transferencia génica en insectos, 206
 Transferencia génica estable, 214
 Transferencia Southern, 104
 Transformación, 22, 43, 126, 129, 133, 179, 181, 194, 195, 198, 200, 202, 209, 272, 278, 280, 281, 286, 295
 Transformación en levaduras, 194, 200
 Transformación por plásmido Ti, 280
 Transgén, 233, 235, 243, 244, 246, 248-255, 262, 263, 269, 273, 281-283, 287, 288, 290
 Transgenes, 262, 264, 269, 274, 282, 283, 288
 Transgénicas, 11, 244-246, 248, 250, 251, 254, 261, 262, 269, 278, 279, 281-287, 289-295
 Transgénicas de arroz, 282
 Transgénico, 248, 249, 251, 254, 255, 286, 291
 Transgénico y fértil, 286
 Transgénicos, 11, 211, 236, 243-265, 267, 282-286, 292
 Cerdos, 252, 253, 258, 260-264
 Ratones, 234, 236, 237, 238, 240, 243, 246, 251, 252, 255, 256, 258, 260, 264-266, 295
 Salmones, 253, 254
 Vacas, 227, 254, 258, 261
 Transgénicos y clónicos, 11, 243
 Ovejas, 258-262
 Transgénicos en *Drosophila*, 245
 Transgénicos integrales, 244, 246, 248, 250
 Transiciones, 137, 138, 176
 Transitoria, 54, 209, 214, 216, 229, 233, 240, 273
 Translocaciones, 43, 65, 66, 106, 168, 177, 191
 Translocaciones de fragmentos, 168
 Transposasa, 69-74, 76, 80, 277, 244
 Transposición, 49
 Control de la, 32, 33, 46, 79-82, 178, 240, 252, 278
 No replicativa, 74, 76, 77, 79
 Replicativa, 67, 72, 74-77, 79, 85, 124, 131, 133, 180, 181, 218, 225, 249, 275
 Vía ADN, 69
 Vía ARN, 69, 77-79, 283
 Transposón Mu, 74
 Transposón o elemento P, 74, 244-246
 Transposón TnA, 76
 Transposón Tn 10 de *E. coli*, 216
 Transposón Ty, 77, 78
 Transposón copia, 72, 76
 Transposón LI, 77, 78
 Transposones y retrovirus, 11, 27, 45
 Transversiones, 167, 168, 173
 Traslado de la mella, 90, 91, 104
 Trasplante de núcleos, 255, 257, 261
 Trasplante de núcleos en anfibios, 257
 Treonín-desaminasa de *E. coli*, 293
 TRI, 68, 69, 76, 223-226, 237
 Trigo, 269, 270, 281, 282, 284, 286, 288
 Triptófano-sintetasa, 195
Triticum aestivum, 286
 TRL, 78, 218, 219, 221, 222, 235, 236
 Troncales, 80, 231, 234, 236, 238, 239, 243, 246, 247, 248, 250, 251, 261, 264, 266, 267
 TRP5, 195, 198
 TRS, 68, 69
 tTA, 216, 217
 Tth, 93, 154
 Tubulina, 37, 38
 Tumor, 113, 188, 190, 238-241, 264, 270, 271
 Tumor de la corona, 270, 271

- Tumores, 140, 161, 188, 191, 215, 222, 230, 238-240, 251, 252, 264, 265, 270, 271, 273
 Tumores de colon humanos, 190, 191
 Tumores sólidos, 191
 Tunicamicina, 205
- UAA «ocre», 37
 UAG «ámbar», 37, 170, 171
 UAS, 202
 UGA «ópalo», 37
Ultrabithorax (Ubx), 43, 44
 Ultrasonidos, 281
 Un solo lugar en el genoma, igualmente hipervaria, 111
 Unilocus, 109, 111, 112, 164
 Unión al ribosoma, 146, 147, 160
 URA3, 195, 199, 200, 203, 204, 232
 Uracil N-glicosilasa, 20
 Uracilo, 19, 20, 21, 48, 173, 174, 176
 Urea, 25
- VAA, 225, 226, 237
 Vacas adultas, 258
 Vacas transgénicas, 254
 Vacunas, 11, 222, 239, 240, 269, 277, 283, 284, 292, 294, 295
 Vacunas (inmunización activa), 294
 Vacunas en plantas, 295
 Vacuolas, 205, 290
 Variabilidad entre:
 células, 42
 individuos, 43
 Especies, 26, 28, 39, 40, 42, 77, 78, 123, 144, 165, 205, 216, 257, 260, 264, 272-274, 277-279, 281, 282, 284, 288, 290, 292, 293
 Variantes polimórficas, 110
 Varicella-zoster, 226
 Vector, 129, 130, 134, 135, 136, 146, 147, 148, 182, 185, 205, 225, 227, 274
 Vector adenoasociado, 225
 Vector BAC, 134
 Vector CAB, 129, 134, 135, 136, 146, 227
 Vector de expresión en *E. coli*, 147, 148
 Vector de integración en levadura, 195
 Vector de reemplazamiento, 130, 182, 185, 274
 Vector de secreción, 205
 Vector de transferencia, 207, 222, 228, 229
 Vector defectivo, 219
 Vector derivado de SV40, 214, 215, 229
 Vector mixto derivado de SV40, 214, 215
 Vectores,
 basados en el círculo de 2 μ m, 202
 basados en elementos P, 244
 caronte, 129
 con promotores específicos, 95
 de expresión, 148, 195, 199, 215, 274, 277
 de expresión basados en SV40, 215
 de expresión de origen mixto, 195
 de expresión en levaduras de origen mixto, 204
 de inserción, 128, 129
 de integración, 195, 196, 198
 de levaduras, 195
 de sustitución, 128, 129, 130, 131
 derivados del Herpes, 226
 EMBL, 129, 130, 131
 eucarióticos, 214
 mixtos, 195, 214, 215, 218
 mixtos de expresión, 214
 plasmídicos, 125, 126
 retrovirales, 232, 235, 238, 239, 240
 retrovíricos, 249
 SIN, 223
 SRA, 198
 SRA/CEN, 198
 SRA/CEN/TEL, 198
 virales, 229, 236, 237, 280
 YAC, 193
 Vent, 92, 152
 Ventajas e inconvenientes, 220, 229, 233
 Ventajas e inconvenientes de los retrovirus como vectores, 220
 Verifica por PCR, 281
 VIH, 221, 222
 Vir A, 271
 Vir D1/D2, 271
 Vir G, 271
 VirE2, 271, 272
 Virus,
 adenoasociados, 67, 224, 236
 ADN, 273
 animales, 215, 295
 ARN, 80, 273-276, 283, 288
 auxiliar, 69, 203, 219, 223, 224, 226, 227
 con forma de varilla, 276
 de la hoja estriada del maíz, 274
 de la inmunodeficiencia adquirida, 218
 de la leucemia murina, 218, 219, 250
 de la poliedrosis nuclear de Autographa California, 206
 de la poliedrosis nuclear de Bombyx mori, 206
 de la Sharka, 277, 278
 de la vacuna, 227-229
 de la viruela, 227
 de los simios, 178, 215
 de plantas como vectores, 273, 278
 del grabado (jaspeado) del tabaco, 277
 del Herpes, 183, 185, 212-214, 216, 217, 224, 226, 227

- del Herpes simplex, 183,185, 212, 214, 216, 217, 226, 238
- del mosaico de la Cassava africana, 274
- del mosaico de la coliflor, 270, 274, 280, 281, 285, 287, 291, 292
- del mosaico del caupí, 277
- del mosaico del tabaco , 270, 274, 276, 287
- del mosaico dorado del tomate, 274
- del mosaico rayado de la cebada, 198
- del sida, 264, 277
- X de la patata, 270, 274, 277
- VP16, 216, 217, 255
- VPNAc, 206, 207
- VV, 228, 229

- WAS/*was*, 233, 235, 236
- WPRe, 221, 222

- X-gal, 76, 126, 134, 203, 213
- Xa21 del arroz (*Oriza sativa*), 41
- Xantina-guanina fosforribosil transferasa, 15
- Xenopus laevis*, 244, 255, 256
- Xenotransplantes, 261
- Xeroderma pigmentosa, 173
- XGPRT, 15
- XhoI, 99, 100, 135

- YAC, 193, 195, 198-200
- Yen 1, 59

- Zanahorias transgénicas,

- Zea mays*, 41
- ZFN, 101

TÍTULOS PUBLICADOS

ÁLGEBRA LINEAL (vol. 2), *A. Gutiérrez Gómez y F. García Castro.*
ANÁLISIS DE DATOS EN LAS CIENCIAS DE LA ACTIVIDAD FÍSICA Y DEL DEPORTE, *M.^a I. Barriopedro y C. Muniesa.*
CIENCIA DE MATERIALES, *P. Coca Rebollero y J. Rosique Jiménez.*
CURSO DE GENÉTICA MOLECULAR E INGENIERÍA GENÉTICA, *M. Izquierdo Rojo*
ECOLOGÍA, *J. Rodríguez.*
ECUACIONES DIFERENCIALES II, *C. Fernández Pérez y J. M. Vegas Montaner.*
ENZIMOLOGÍA, *I. Núñez de Castro.*
FÍSICA CUÁNTICA, *C. Sánchez del Río (coord.).*
FISIOLOGÍA VEGETAL, *J. Barceló Coll, G. Nicolás Rodrigo, B. Sabater García y R. Sánchez Tamés.*
INTEGRACIÓN DE FUNCIONES DE VARIAS VARIABLES, *J. A. Facenda Aguirre, F. J. Freniche Ibáñez.*
INTRODUCCIÓN A LA ESTADÍSTICA Y SUS APLICACIONES, *R. Cao Abad, M. Francisco Fernández, S. Naya Fernández, M. A. Presedo Quindimil, M. Vázquez Brage, J. A. Vilar Fernández, J. M. Vilar Fernández.*
MÉTODOS NUMÉRICOS. Teoría, problemas y prácticas con MATLAB, *J. A. Infante del Río y J. M.^a Cabezas.*
PROBLEMAS, CONCEPTOS Y MÉTODOS DEL ANÁLISIS MATEMÁTICO. 1. Números reales, sucesiones y series, *M. de Guzmán y B. Rubio.*
PROBLEMAS, CONCEPTOS Y MÉTODOS DEL ANÁLISIS MATEMÁTICO. 2. Funciones, integrales, derivadas, *M. de Guzmán y B. Rubio.*
SERIES DE FOURIER Y APLICACIONES. Un tratado elemental, con notas históricas y ejercicios resueltos, *A. Cañada Villar.*
TABLAS DE COMPOSICIÓN DE ALIMENTOS, *O. Moreiras, A. Carbajal, L. Cabrera y C. Cuadrado.*
TECNOLOGÍA MECÁNICA Y METROTECNIA, *P. Coca Rebollero y J. Rosique Jiménez.*

Si lo desea, en nuestra página web puede consultar el catálogo completo o descargarlo:

www.edicionespiramide.es