

# CAPÍTULO VI

## Los Mapas Genéticos

---

### 6.1 Introducción

El ratón, después del hombre, es el mamífero que tiene el mapa genético más completo y fue el segundo genoma mamífero completamente secuenciado (ver más adelante). Por otro lado, las técnicas de mapeo, los programas de computación y las bases de datos están enormemente desarrolladas en esta especie. Las cualidades que hacen del ratón la especie modelo para el desarrollo de mapas genéticos son: (i) Tienen un tiempo generacional muy corto y son muy prolíficos, lo que permite criar miles de ratones híbridos F1, F2 o N2. (ii) Es una especie en la cual se pueden llevar a cabo cruces inter-específicas. (iii) Existe una cantidad enorme de líneas genéticamente definidas, como las consanguíneas y congénicas, además de miles de mutaciones y un gran número de rearreglos cromosómicos disponibles.

Por todo lo mencionado, tomaremos al mapa genético del ratón y sus métodos de desarrollo como guía (con reservas) para todos los animales de laboratorio.

Como ya se vio, el ratón de laboratorio "estándar" tiene un cariotipo de 40 cromosomas (19 autosómicos más el par sexual). Los cromosomas murinos son difíciles de diferenciar en los preparados citogenéticos porque, contrariamente a los humanos, son todos acrocéntricos y muestran una graduación continua en el tamaño. Los cromosomas autosómicos se numeran del 1 al 19, en orden decreciente de longitud.

Según las estimaciones más recientes, el genoma murino contendría entre 30.000 y 35.000 genes, de los cuales un tercio ya ha sido identificado, gracias a la existencia de mutaciones y a los trabajos de genética molecular. Es interesante resaltar que la mayoría de las mutaciones descubiertas por "criadores de ratones" identifican, en general, un locus nuevo más que nuevos alelos de un locus ya conocido. En este contexto, si consideramos los recientes resultados provenientes de las experiencias en noqueo de genes, es probable que la mayoría de las mutaciones que ocurren en el genoma del ratón sean indetectables, por ser incompatibles con el desarrollo normal del embrión, o por no causar efectos fenotípicamente obvios. Considerando que un gran número de genes del ratón espera ser descubierto, la existencia de un mapa genético de alta resolución y alta densidad facilitaría enormemente esa tarea. Como veremos en detalle más adelante, un mapa de **alta resolución** se logra analizando un gran número de ratones (meiosis) y un mapa de **alta densidad** se obtiene por medio del uso de muchos marcadores polimórficos.

La forma más racional de desarrollar un mapa genético del genoma del ratón es hacerlo en forma gradual, primero ubicando una serie de **loci de referencia** (del inglés *anchor loci*), o de marcadores, e ir aumentando progresivamente el número de los mismos en los espacios desiertos hasta lograr un "andamio" sólido. Con este enfoque escalonado y, teniendo en cuenta que el número de marcadores a disposición es suficientemente grande, es posible hoy en día realizar mapas de alta densidad en el ratón. De hecho, ésta es la estrategia que vienen utili-

zando los genetistas desde comienzos del siglo veinte. En 1915, John B. Haldane y colaboradores informaron en el *Journal of Genetics* que dos mutaciones del color del pelaje, *albino* (*Tyr<sup>c</sup>*) y *pink-eyed dilution* (*p*), estaban ligadas en el cromosoma 7 del ratón<sup>1</sup>. Este trabajo es considerado hoy como el "puntapié inicial" del mapa genético del ratón. Durante las décadas de 1970 y 1980 se produjeron grandes cambios que llevaron a un gran aumento del flujo informativo referente al mapeo de genes. Estos cambios incluyen el descubrimiento de nuevos marcadores polimórficos (primero bioquímicos y luego de ADN), la puesta a punto de la tecnología de ADN recombinante, el desarrollo de las cruces inter-específicas de ratones y, finalmente, la elección del ratón como modelo de enfermedades humanas.

Desde un punto de vista físico (y muy reduccionista!) los genes no son más que segmentos de ADN alineados a lo largo de los cromosomas como cuentas de un rosario. Una manera de establecer un mapa "final" sería clonando el genoma en pequeños segmentos superpuestos y secuenciando estos fragmentos uno tras otro. Esto llevaría, finalmente, a la secuencia completa del genoma. Este planteo fue considerado poco realista en el pasado (teniendo en cuenta las técnicas disponibles) pero, debido a la gran disminución en los costos y al aumento en la eficiencia de la secuenciación automática, es hoy una realidad.

Se sabe que los genes en los mamíferos no son tan compactos como la mayoría de los genes procariotas y están salpicados con secuencias repetitivas de distinto tipo, como copias de provirus, minisatélites, microsátélites y pseudogenes, cuya función, si existe, es aún desconocida (ver Capítulo I). Es por todo esto que se criticó, al principio, la idea de secuenciar completamente un genoma haciendo experimentos muy tediosos con gastos enormes e injustificados debido a las secuencias de ADN irrelevante (lo que se conoció en una época como ADN de descarte, del inglés *junk DNA*). Sin embargo, como veremos en detalle más adelante, la disponibilidad de las secuencias de los genomas humano y del ratón está trayendo valiosa información sobre la estructura del genoma de los mamíferos.

## 6.2 Tipos de mapa del genoma

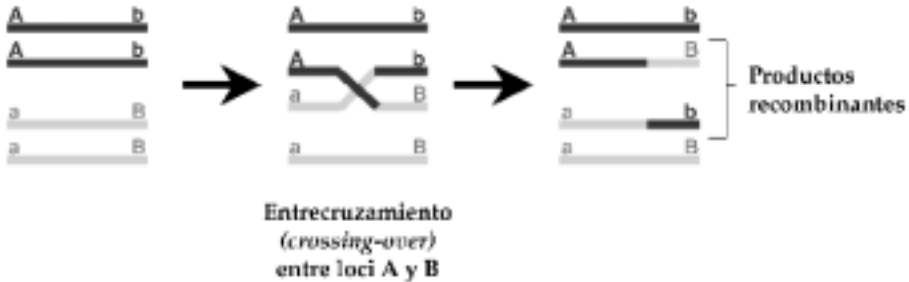
Hay varias clases de mapas que interesan a los genetistas, aunque los tres más importantes son los mapas de ligamiento, los mapas cromosómicos y los mapas físicos.

### 6.2.1 Mapas de Ligamiento

El establecimiento de **mapas de ligamiento** (*linkage maps*), también llamados **mapas meióticos**, se basa en el hecho de que, durante la meiosis, los loci que se encuentran en diferentes cromosomas se separan al azar en las gametas, mientras que los que se encuentran en un mismo cro-

<sup>1</sup> El mapa actual de esa misma región del cromosoma 7 posee alrededor de 2.300 loci!

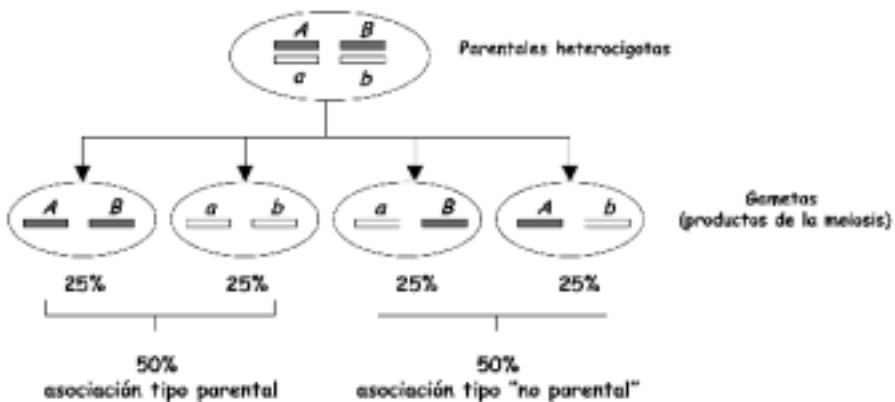
mosoma tienden a co-segregar, al menos que un evento de **recombinación** rompa esa asociación de tipo "parental" (**Figura 6.1**). La presencia de este fenómeno de **entrecruzamiento** (del inglés *crossing-over*) se visualiza en los cromosomas como una estructura llamada **quiasma**. Por lo tanto, para desarrollar este tipo de mapas de ligamiento hace falta realizar cruza de animales.



**Figura 6.1. Recombinación del ADN.** El esquema muestra la recombinación del ADN en cromátidas no hermanas de un par de cromosomas homólogos durante el entrecruzamiento. Como consecuencia, se generan gametas recombinantes entre los loci A y B (AB y ab no se encontraban en el mismo cromosoma en los padres). Este fenómeno de recombinación es la base de los estudios de ligamiento (del inglés *linkage analysis*).

Básicamente, la probabilidad de que dos genes sean separados por un evento de recombinación dependerá de la distancia que haya entre ellos. En el caso de dos loci no ligados, la frecuencia de gametas esperada será 50% tipo parental y 50% tipo no parental (recombinante), mientras que en aquellos loci que estén ligados la frecuencia de gametas tipo parentales será siempre mayor al 50% (**Figura 6.2**). Esto se ve reflejado en la elección de la unidad de mapeo, el **centiMorgan (cM)**, que corresponde al 1% de probabilidad de producir una gameta recombinante luego de una meiosis.

**Figura 6.2 A. Segregación independiente de loci no ligados**



En otras palabras, un mapa de ligamiento es un diagrama de los rearrreglos lineales de los genes localizados en un cromosoma dado. Estos mapas fueron diseñados a principios del siglo XX por el científico americano (premio Nobel de Medicina y Fisiología en 1933) Thomas H. Morgan y sus discípulos (en particular Alfred H. Sturtevant), usando cruces de la mosca de la fruta (*Drosophila melanogaster*) y fueron aplicados posteriormente a los mamíferos. Como se verá más adelante, para desarrollar este tipo de mapas es esencial cruzar animales (genera-

Figura 6.2 B. Segregación no independiente de loci ligados

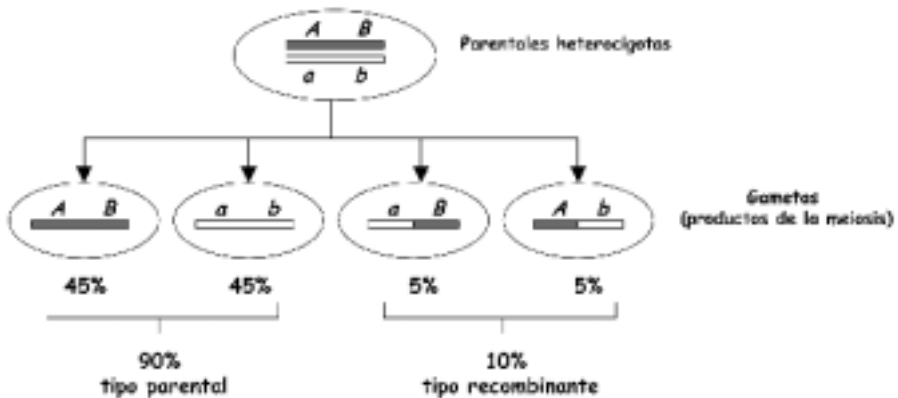
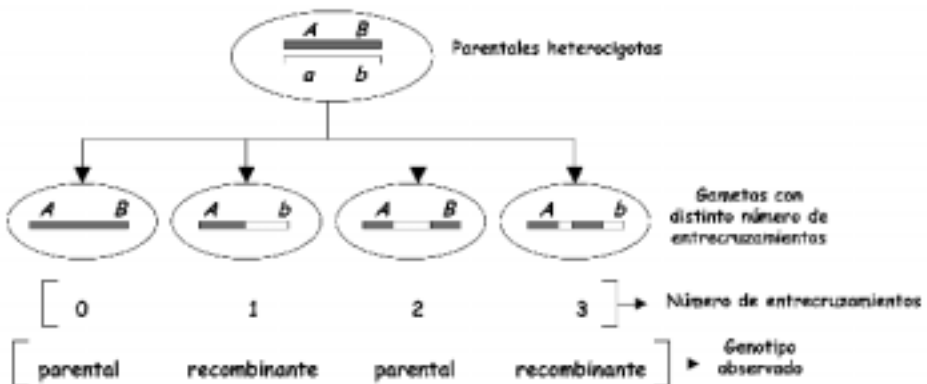
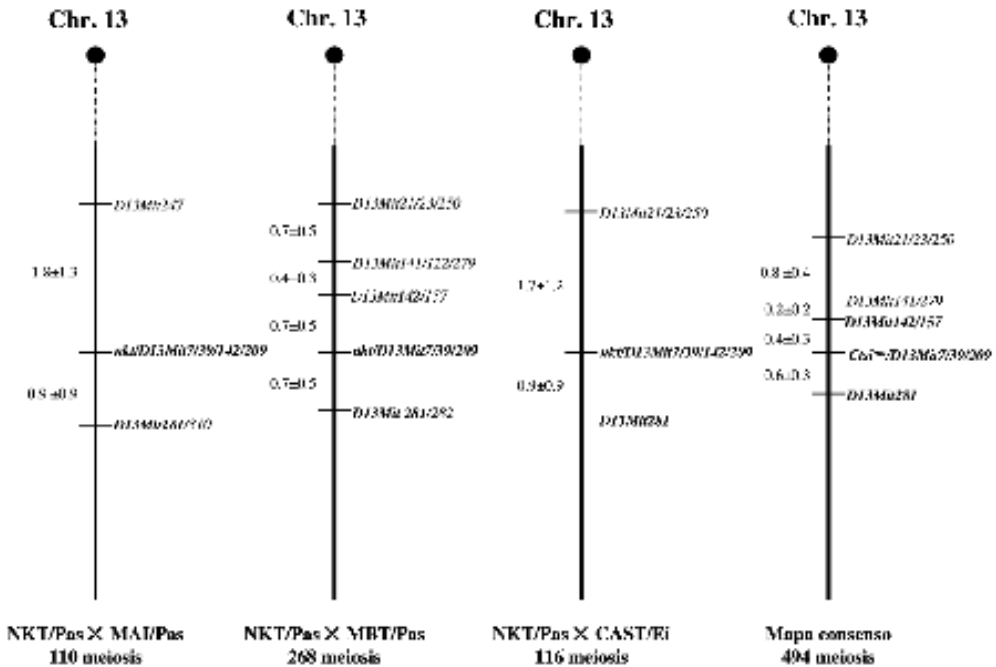


Figura 6.2 C. Efecto de los entrecruzamientos múltiples en la detección de cromosomas recombinantes



**Figura 6.2. Tipos de segregación.** **A.** Segregación independiente de alelos en el caso de loci no ligados. Las cuatro clases de productos de la meiosis (gametos) tienen frecuencias similares. **B.** Segregación no independiente de alelos pertenecientes a loci ligados. Como consecuencia del ligamiento, el porcentaje de gametos del tipo parental es mayor del 50%. En el ejemplo, el porcentaje de gametos recombinantes es de sólo el 10%. **C.** Efecto de la presencia de múltiples entrecruzamiento en la detección de cromosomas recombinantes. Aunque este tipo de eventos es muy infrecuente, cuando el número de recombinaciones múltiples es par, la recombinación no se detecta. Adaptado de Silver L. M. (ed) *Mouse Genetics. Concepts and applications.* Oxford University Press, Oxford, 1995.

ción parental) que tengan alelos diferentes en dos o más loci, para poder estudiar el patrón de segregación en la descendencia. Es por esta razón que los mapas de ligamiento dependen enormemente de la disponibilidad de **polimorfismos** (la existencia de 2 ó más alelos en un locus dado). La **Figura 6.3** muestra, a modo de ejemplo, un mapa (parcial) del cromosoma 13 del ratón donde se localizó la mutación inmunodeficiente *nackt* (*Cts<sup>lnkt</sup>*).



**Figura 6.3. Mapas de ligamiento.** Este ejemplo de mapa de ligamiento muestra la localización de la mutación recesiva *nackt* (*Cts<sup>lnkt</sup>*) en el cromosoma 13 del ratón. Se muestran tres intercruzas sub-específicas, empleando las líneas MAI y MBT (*M. m. musculus*) y CAST/Ei (*M. m. castaneus*), y un mapa consenso (derecha) incluyendo el total de meiosis analizadas. Todos los marcadores indicados son microsatélites. A la izquierda de los mapas se observa la distancia en cM (± desvío estándar) entre loci, calculadas al 5% de nivel de riesgo. Tomado de Benavides et al., *Immunogenetics* 53: 233-242, 2001.

Existen un par de puntos a tener en cuenta con respecto a estos mapas:

- (i) Para establecer el orden de los loci se requiere, por definición, que al menos un evento de recombinación rompa el orden lineal de origen parental. Esto es poco frecuente si los dos marcadores están fuertemente ligados.
- (ii) La densidad de un mapa genético estará correlacionada con el número de polimorfismos que segregan en una cruce particular, mientras que su resolución depende del número de gametas (= número total de meiosis) analizadas en la progenie.

La longitud total del mapa de ligamiento del ratón ha sido estimada por diversos investigadores, usando diferentes enfoques, en el orden de los 1500 a 1600 cM. Esto quiere decir que, en promedio, 1 cM en el genoma del ratón corresponde a 1700 kb, mientras que en el humano esta cifra es equivalente a 1000 kb. Estos datos deben considerarse estrictamente como promedios porque no hay una correlación absoluta entre escalas en cM y kb, debido a la distribución despareja de los eventos de recombinación a lo largo de las cromátides (en el ratón se calcula un promedio de 28 quiasmas por genoma). Por ello, dos genes pueden aparecer completamente ligados en un mapa de alta resolución y resultar estar muy distantes en términos moleculares mientras que, contrariamente, dos genes pueden parecer relativamente alejados uno del otro si existe un "hot spot" de recombinación en la región intergénica. También hay que considerar que existe un fenómeno conocido como **interferencia de entrecruzamientos** (del inglés *crossover interference*) que impide la formación de quiasmas cercanos en un mismo cromosoma (la localización de los quiasmas en las cromátides no es al azar). En particular, el genoma del ratón muestra un fenómeno de interferencia mucho más fuerte que el descrito en el genoma humano.

Varios programas de computación han sido desarrollados para ayudar a los genetistas a procesar los datos obtenidos de las cruza experimentales y la evaluación de los marcadores en un análisis de ligamiento. Entre ellos: MAPMAKER/EXP 3.0 (1992) de Eric Lander (*Whitehead Institute*, Estados Unidos), MapManager QT (1999) y MapManager QTX (2001), de Kenneth Manly y colaboradores (*Roswell Park Cancer Institute*, Estados Unidos) y Gene-Link de Xavier Montagutelli (*Institut Pasteur*, Francia).

En el momento de este escrito, el mapa de ligamiento del ratón cuenta con alrededor de 18.000 genes y 15.000 marcadores genéticos (la mayoría marcadores moleculares). La distancia promedio entre marcadores es menor a 0,01 cM y la densidad promedio del mapa murino es de 12,5 loci por cM. El mapa de la rata, el tercero más avanzado después de los mapas del hombre y el ratón, posee localizados alrededor de 1.600 genes, 6.500 secuencias que podrían representar genes (en inglés, *predicted genes*) y 11.000 marcadores de ADN. Ningún otro roedor de laboratorio posee un mapa genético avanzado como los del ratón y la rata. La versión actualizada de los mapas genéticos de estos roedores de laboratorio se encuentra a disposición a través de la base de datos de *The Jackson Laboratory* (MGD) (<http://www.informatics.jax.org>) y en la base de datos del genoma de la rata (<http://rgd.mcgw.edu>).

## 6.2.2. Mapas Cromosómicos

Mientras que los mapas de ligamiento requieren de la realización de protocolos con cruza de animales, los **mapas cromosómicos** se logran usando técnicas que no incluyen la reproducción sexual y, por lo tanto, asignan una localización según las regiones citogenéticas (además, estos mapas presentan la ventaja de no requerir del uso de marcadores polimórficos). Esas técnicas son: (i) hibridación *in situ*, (ii) híbridos de células somáticas y (iii) híbridos de radiación.

### 6.2.2.1 Hibridación In Situ.

Cuando una secuencia de ADN, en este caso referida como una “sonda” de ADN (en inglés, *probe*), es marcada con isótopos radioactivos o fluorescencia, podemos hibridarla con su secuencia complementaria en el ADN genómico de un cromosoma específico y observar la marca directamente sobre el extendido cromosómico con un microscopio óptico. La hibridación *in situ* fluorescente (**FISH**) es una técnica muy sensible siempre que se cuente con sondas lo suficientemente grandes (y homólogas) y que se suprima la fluorescencia de fondo. La sensibilidad de esta variante de la técnica de hibridación *in situ* permite localizar una secuencia con una precisión de hasta 30 Mb y de detectar anomalías cariotípicas en cromosomas metafásicos tanto como interfásicos. Los **fluorocromos** más empleados son la **fluoresceína** (FITC), la **rodamina** (XRITC) y el **Texas Red**.

Aunque se trata de una técnica muy confiable, el FISH es poco utilizado en la construcción de mapas en el ratón por varias razones: (i) en muchos casos es más fácil y práctico realizar esquemas reproductivos, (ii) la hibridación *in situ* nos provee de una posición regional, no una localización precisa y (iii) los cromosomas murinos, como ya se dijo, son muy difíciles de distinguir entre sí, en un preparado citogenético. De todas maneras, la hibridación *in situ* es la técnica de elección para la localización rápida de los transgenes (ver Capítulo VIII). Es también muy útil para evaluar si los grandes insertos en los cromosomas artificiales de levadura (YAC's) derivan todos del mismo cromosoma o, por el contrario, son quiméricos. Por último, el **pintado de cromosomas** (del inglés *chromosome painting*) es una variante del FISH que emplea una mezcla de sondas de ADN derivada de un cromosoma entero o de una región cromosómica de interés. Se utiliza para obtener patrones de regiones homólogas entre diferentes especies y es una forma muy directa de evaluar la cantidad de rearrreglos que se produjeron entre dos especies en el curso de la evolución.

### 6.2.2.2 Híbridos de Células Somáticas

Para esta técnica se utilizan **células híbridas** (viables) derivadas de la fusión *in vitro* de células somáticas de especies diferentes de mamíferos. Aunque se desconoce la razón, se observa que los cromosomas que se van perdiendo en las células híbridas pertenecen exclusivamente a uno de los conjuntos parentales. Por ejemplo, los híbridos humano/roedor tienden a perder (a través de los sucesivos cultivos y en forma preferencial) los cromosomas humanos y quedarse con una mayoría de cromosomas de roedor. Debido a esto, es posible derivar un panel de células híbridas que representen cada cromosoma humano en forma única, sobre un conjunto de cromosomas de roedor. Por lo tanto, todos los genes de origen roedor son retenidos (y se expresan normalmente), mientras que sólo los genes presentes en el único cromosoma humano que quedó retenido en ese híbrido podrán ser localizados por medio del uso de sondas moleculares o el análisis de expresión. De esta manera, detectando patrones de presencia o ausencia de marcadores (o expresión de genes) y correlacionando con los cromosomas presentes en el híbrido, se puede asignar un gen humano a un cromosoma en particular. Los híbridos de células somáticas humano/ ratón, con un complemento limitado de cromosomas humanos (intactos), han sido muy usados para localizar genes humanos en los

últimos 20 años, a pesar de ser muy inestables. También existen híbridos similares segregando cromosomas de ratón, pero no han sido muy utilizados porque es en general más fácil y rápido usar el ADN derivado de cruza de animales. Además, contrariamente a los cromosomas humanos, son raros los híbridos de células somáticas del ratón con deleciones o translocaciones, lo que complica la asignación sub-cromosómica de genes por este método.

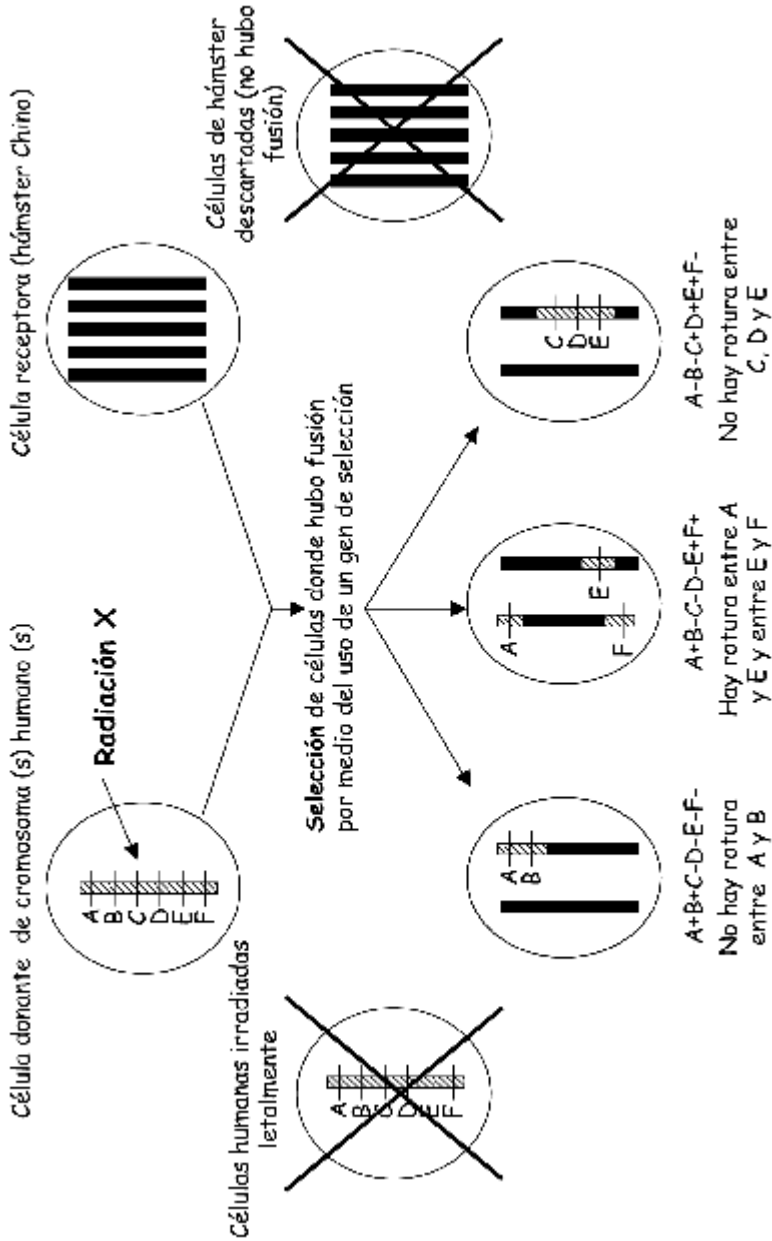
### 6.2.2.3 Híbridos de Radiación (HR)

Otra técnica desarrollada para obtener mapas de alta resolución en humanos son los paneles de **híbridos de radiación** (*Radiation Hybrid* –RH), enfoque descrito por Goss y Harris en 1975. Las dos grandes ventajas de este sistema son que pueden mapearse marcadores o genes no polimórficos (ya que se evalúa sólo la presencia o ausencia de un marcador) y que se logra una resolución mayor que con los mapas de ligamiento, constituyendo un puente entre éstos y los mapas físicos. Estos paneles de células híbridas se construyen a partir de células donantes (de la especie a ser mapeada) que han sido irradiadas letalmente (ya sea con rayos  $\gamma$  o X) para causar la fragmentación de sus cromosomas (por roturas de doble cadena). Las células así irradiadas son fusionadas con líneas celulares receptoras deficientes para un **marcador de selección**, como ser la **timidina quinasa** (TK<sup>-</sup>) o la **hipoxantina fosforibosil transferasa** (HPRT<sup>-</sup>). Usando condiciones de selección con medios apropiados, sobrevivirán sólo aquellas células que contengan, por lo menos, algún fragmento cromosómico proveniente de las células dadoras. Por ejemplo, para el mapeo de genes humanos se utilizó siempre un fondo de células receptoras de roedor, en particular de ratón o hámster chino (*Cricetulus griseus*) (**Figura 6.4**). Muchas de las líneas de células híbridas obtenidas por este método retienen cantidades detectables del ADN donante (normalmente entre 20 y 30% del genoma). Por lo tanto, estas células híbridas, obtenidas por las técnicas clásicas de fusión celular, contienen sólo un pequeño segmento del genoma donante (irradiado) incorporado en los cromosomas de la especie receptora (no irradiada).

El concepto teórico es que aquellos loci que se encuentran muy próximos unos de otros serán retenidos en el mismo fragmento cromosómico después de la irradiación. Esta característica es similar al principio de ligamiento genético que hemos visto en los mapas meióticos. Como en los paneles de ADN obtenidos con cruza de animales, la información de los paneles HR es acumulativa y puede proveer datos para el ordenamiento (de alta resolución) de regiones no polimórficas o no separables por los mapas de ligamiento.

En 1998 se puso a disponibilidad de la comunidad científica un panel HR del ratón, el primero en su momento para un organismo modelo. Su nombre es T31 y fue creado por Linda McCarthy y sus colaboradores usando fragmentos de cromosomas de ratón (de una línea de células embrionarias de la cepa I29) retenidos en cromosomas de hámster chino (provenientes de una línea de fibroblastos TK<sup>-</sup>), totalizando 104 líneas celulares. La información más actualizada sobre este panel y las secuencias que se encuentran localizadas se puede obtener en *The Jackson Laboratory T31 Mouse Radiation Hybrid Database* (<http://www.jax.org/resources/documents/cmdata/rhmap/>). Desde el año 1999, la rata posee también su panel HR (T55), en el cual se han localizado más de 5.000 marcadores genéticos. Los paneles HR ratón/hámster y rata/hámster se encuentran





**Figura 6.4. Híbridos de radiación.** Diagrama esquemático mostrando la construcción de clones de híbridos de radiación (HR) a partir de células donantes con un único cromosoma humano (híbrido mono-cromosómico) y células receptoras de hámster Chino. La selección de células donde hubo fusión se realiza por medio del uso de genes de selección como la timidina kinasa. Cada clon (abajo) presenta una colección diferente (al azar) de fragmentos del cromosoma humano original. Los 6 loci hipotéticos (A a F) se marcan como + (presencia) o - (ausencia), dando una idea del ordenamiento en el cromosoma original.

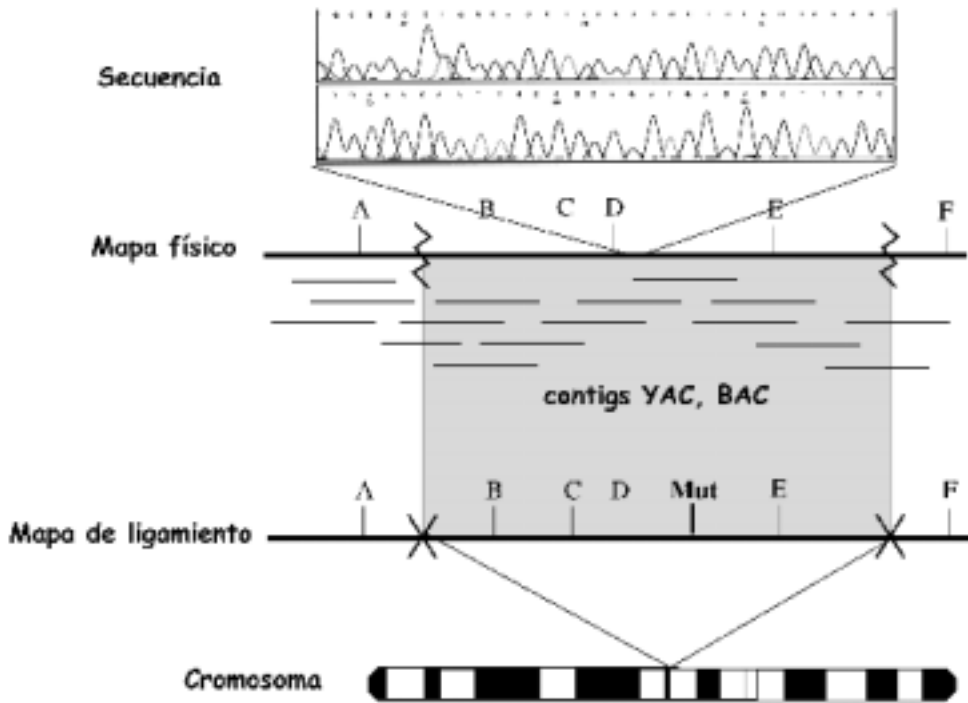
disponibles comercialmente a través de la empresa Research Genetics Inc. (<http://www.res-gen.com/products/MRH.php3#info>).

Cuando se evalúa un marcador (por Southern blot, FISH o PCR), el patrón de presencia (+) o ausencia (-) a través del panel, define el emplazamiento del mismo: aquellos marcadores con el mismo patrón de + y - estarán localizados en el mismo "bin" o posición. Estos estudios se conocen como "**patrones de retención de marcadores**" y nos ayudan a determinar la ubicación de un gen o marcador. En estos casos, la unidad utilizada para medir la distancia en el mapa es el **Ray** o el **centiRay (cR)**. Debido a que la dosis de rayos X tiene gran influencia sobre esta medida, es necesario aclararlo de la siguiente manera: un  $cR_{3000}$  indica una frecuencia de rotura del 1% entre dos loci dados luego de una exposición a 3.000 rads de radiación X. Como esta distancia entre marcadores puede tener poca significación para un mapa genético, será necesario convertirla a distancias de mapas físicos (en kb o Mb). Por ejemplo, en el panel T55 de la rata un  $cR_{3000}$  es equivalente a una distancia física de 106 kb.

La evaluación estadística de los resultados debe hacerse con programas de computación especiales, como RHMAPPER (<http://www-genome.wi.mit.edu/ftp/distribution/software/rhmapper/>). El mapa más reciente del ratón basado en híbridos de radiación (panel T31) fue publicado en la revista *Nature Genetics* en el año 2001. Este mapa contiene 11.109 genes (incluyendo genes expresados sólo en embriones y neonatos) que fueron emplazados con respecto a un mapa de referencia conteniendo alrededor de 2000 marcadores genéticos, e incluye 3.658 **genes homólogos** con secuencias humanas (**genes ortólogos**). Uno de los aspectos más notables que surge del análisis de este mapa es la distribución irregular de genes a lo largo de los cromosomas; por ejemplo, existe un exceso de genes en los cromosomas 11 y 19 y un déficit en los cromosomas 18 y X.

### 6.2.3 Mapas Físicos

Un **mapa físico** (*physical map*) es la representación real del alineamiento de los genes en un cromosoma, en la misma forma que un mapa de ruta indica la localización de las ciudades a lo largo de una autopista. El orden de los genes es el mismo que el dado por los mapas de ligamiento, pero las distancias son medidas en kb o Mb. Estos mapas constituyen un paso crucial en la caracterización estructural y funcional del genoma. Pueden lograrse por diversos métodos pero el más conveniente es el ordenamiento de grupos de clones superpuestos (**contigs**), ya sean estos clonados en fagos (PI en general), cósmidos, cromosomas artificiales de bacterias (BAC's) o cromosomas artificiales de levaduras (YAC's) (ver Capítulo I). Aunque en teoría es posible desarrollar un mapa físico *de novo*, por ejemplo uniendo extremos de insertos de ADN clonados en YAC's (y detectando otros clones usando esos extremos como sondas hasta tener grandes segmentos ordenados), el establecimiento de dicho mapa se vería muy facilitado si existiese previamente un mapa de ligamiento de alta densidad (200-400 marcadores por cromosoma). Siguiendo esta estrategia, la realización de un mapa físico es un "se-



**Figura 6.5. Clonaje posicional.** El diagrama muestra el sistema escalonado que nos lleva del mapa de ligamiento donde se ha localizado una mutación (Mut) a la secuencia de un gen candidato por el proceso denominado clonaje posicional. El mapa físico es la etapa siguiente al mapa genético y se elabora gracias a la ayuda de clones contiguos (contigs) de vectores de gran capacidad como los YAC y BAC. La secuencia de un gen o región del genoma es la etapa final de este proceso.

gundo paso” una vez establecido el mapa de ligamiento (Figura 6.5). Como en los mapas de ligamiento, existen varios puntos a tener en cuenta:

- (i) Antes de establecer contigs de YAC's, BAC's o PI es importante confirmar que cada molécula de ADN clonada representa realmente una secuencia inalterada y derivada sólo de un cromosoma. Dado que todas las **bibliotecas genómicas** (o **genotecas**) son construidas con la misma estrategia (fragmentación del ADN de alto peso molecular con enzimas de restricción y el posterior ligado de los productos a un vector), es muy común que, por azar, dos segmentos completamente independientes sean empaquetados juntos en el mismo vector. El vector resultante será un mosaico y su uso puede llevar a resultados incongruentes. Una situación similar, pero aún más grave, puede ocurrir si un segmento de ADN queda deleciónado en el proceso de clonación.
- (ii) Para establecer contigs superpuestos de clones de YAC's, BAC's, cósmidos o PI es necesario construir varias genotecas independientes con diversas enzimas de restricción, con digestión completa o incompleta.

Se ha publicado recientemente (*Nature*, 2002) un mapa físico del ratón que cuenta con 296 contigs superpuestos de clones BAC y casi 17.000 marcadores genéticos. Este mapa fue creado básicamente en dos etapas. La primera etapa fue identificar la superposición de los 300.000 clones de BAC disponibles por medio del uso de enzimas de restricción (*fingerprints*), para poder construir 7.500 contigs de clones superpuestos. Usando la información proveniente de las secuencias de los extremos de los fragmentos clonados en los BAC's, se pudo alinear estos contigs con la secuencia del genoma humano, lo que generó un mapa de homología humano-ratón. En la segunda etapa los investigadores usaron miles de marcadores genéticos para posicionar los contigs en el mapa genético del ratón. Asumiendo que el tamaño del genoma del ratón es de alrededor de 2.500 Mb (2,5 Gb), este mapa cubre la totalidad del genoma ya que el tamaño promedio de los contigs es de 9,3 Mb ( $296 \times 9,3 = 2.752$  Mb). Este mapa físico permitirá terminar la secuencia completa del genoma del ratón (con 99,9% de precisión) antes del 2005.

#### 6.2.4 Secuenciando Genomas

La información más detallada que podemos obtener de una región del genoma (o de un genoma completo) es su secuencia nucleotídica. Secuenciar fragmentos pequeños de ADN (obtenidos por técnicas de sub-clonación o por PCR) es hoy en día relativamente accesible, aunque se trata de un emprendimiento de gran envergadura cuando se quiere secuenciar más allá de un gen y sus alrededores. Como veremos, existen ambiciosos proyectos conjuntos, como el de la secuenciación completa del genoma humano (lanzado en 1988 y terminado en el año 2001), el ratón, la rata, y otros vertebrados como el **pez cebra** (*Danio rerio*) ([http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/)) y el **pez globo** japonés (*Fugu rubripes*) (<http://fugu.hgmp.mrc.ac.uk/>).

En la actualidad, se encuentran terminadas las secuencias genómicas completas correspondientes a varios microorganismos, entre ellos, más de 150 virus, alrededor de 10 eubacterias (*Haemophilus influenzae*, *Mycoplasma genitalium*, *Escherichia coli*, *Bacillus subtilis* y *Mycobacterium tuberculosis*, entre otras), arcobacterias (*Methanococcus jannaschii*) y eucariotas (la levadura *Saccharomyces cerevisiae* fue el primero). El primer organismo metazoario fue el nematode *Caenorhabditis elegans*, terminado después de 15 años de trabajo en el año 1998. Hasta el año 2000, el genoma más grande secuenciado en forma completa era el de la mosca de la fruta (*Drosophila melanogaster*) con 180 Mb y alrededor de 13000 genes, secuenciado por la empresa *Celera Genomics* (Rockville, Estados Unidos). Recientemente se sumó el mosquito (*Anopheles gambiae*) como segundo insecto con su genoma completamente secuenciado (2002); el mismo tiene un tamaño aproximado de 278 Mb y alrededor de 14000 genes. Con un genoma de apenas 120 Mb, se encuentra *Arabidopsis thaliana*, la primer planta secuenciada (2000). El genoma humano, el primer vertebrado secuenciado, tiene un tamaño de 2.900 Mb, conteniendo alrededor de 35.000 genes, un número sorpresivamente inferior a los 100.000 calculados con anterioridad. Recientemente se sumaron a la lista de vertebrados secuenciados

el ratón de laboratorio (*Mus musculus*), con 2.500 Mb, la rata (*Rattus norvegicus*), con un tamaño estimado de 2.800 Mb y el pez globo, con su genoma llamativamente compacto de 365 Mb (Tabla 6.1).

Organismo	Fecha de terminación	Tamaño del Genoma	Nº estimado de genes
<i>H. influenzae</i>	1995	1,8 Mb	1.700
<i>S. cerevisiae</i>	1996	12 Mb	6.000
<i>C. elegans</i>	1998	97 Mb	19.000
<i>A. thaliana</i>	2000	100 Mb	25.000
<i>D. melanogaster</i>	2000	180 Mb	13.000
<i>H. sapiens</i>	2001	2.900 Mb	30.000-40.000
<i>M. musculus</i>	2002	2.500 Mb	30.000-40.000
<i>R. norvegicus</i>	2003	2.800 Mb	30.000-40.000
<i>F. rubripes</i>	2002	365 Mb	30.000-35.000
<i>A. gambiae</i>	2002	278 Mb	14.000

**Tabla 6.1. Secuencias completas de genomas.** La tabla muestra algunos organismos cuyos genomas fueron secuenciados en forma completa, el tamaño comparativo de sus genomas y el número estimado de genes. Se incluye un procarionta, la bacteria *Haemophilus influenzae*, y nueve de los eucariotas secuenciados hasta la actualidad (algunos en una versión borrador). Dentro de los eucariotas secuenciados encontramos plantas (*Arabidopsis thaliana*), hongos (*Saccharomyces cerevisiae*) y animales pertenecientes a las clases de los insectos, nematodos, peces y mamíferos (ver texto).

A diez años del lanzamiento del **Proyecto Genoma Humano** por James Watson, la presentación simultánea de dos versiones del primer borrador de la secuencia casi completa del genoma humano fue un acontecimiento sin precedente y de indudable importancia. El 15 de febrero de 2001 la revista *Nature* publicaba el artículo titulado *Initial Sequencing and Analysis of The Human Genome* con los resultados de una colaboración entre 20 centros de Estados Unidos, Inglaterra, Japón, Francia, Alemania y China, involucrando una enorme cantidad de participantes. Se trató de un emprendimiento de gran envergadura con fondos públicos originados en los Institutos Nacionales de la Salud de los Estados Unidos (*National Institutes of Health*, NIH) y liderado por Francis Collins. Un día después, aparecía en la revista *Science* con el título de *The Sequence of the Human Genome* el trabajo de Craig Venter de la empresa *Celera Genomics*, y sus múltiples colaboradores, esta vez totalmente con fondos privados. Se llegaba así a

un final feliz de lo que representó una costosa y a veces muy agresiva disputa entre los dos grupos, una verdadera carrera contra el tiempo ya que inicialmente la meta estaba fijada para el año 2005.

En el caso del ratón, el consorcio público (*Mouse Genome Sequencing Consortium*) que realiza su secuenciación (utilizando ADN de una hembra de la línea C57BL/6)<sup>2</sup>, ha hecho disponible el borrador inicial del genoma de este roedor en el año 2002 con un volumen especial en la revista *Nature* [volumen 420 (6915), 2002]. Este consorcio es un grupo internacional de investigadores pertenecientes al *Whitehead Institute* y la *Washington University* en Estados Unidos, y el *Wellcome Trust Sanger Institute* y el *European Bioinformatics Institute* en Inglaterra. Los fondos provienen del *National Human Genome Research Institute* (NIH) en Estados Unidos y el *Wellcome Trust* en el Reino Unido. El genoma del ratón fue secuenciado por un sistema mixto que combina el **escopetazo jerarquizado** (*BAC-based shotgun sequencing*, sistema que también eligió el Proyecto Genoma Humano) y la **secuencia completa por escopetazo** (*whole-genome shotgun sequencing*, método que usó *Celera Genomics* para secuenciar la mosca de la fruta y parte del genoma humano) (ver Capítulo I). Brevemente, el conjunto de secuencias clonadas en *contigs*, fue organizado (por análisis de superposición de secuencias) en fragmentos más grandes llamados **supercontigs** (de alrededor de 17 Mb) y éstos, a su vez, fueron anclados en el mapa físico del ratón (basado en clones de BAC's). Luego, los *supercontigs* adyacentes fueron unidos para formar **ultracontigs** (de alrededor de 50 Mb), llegándose a una situación donde la secuencia casi completa de todo un cromosoma estaba contenida en unas decenas de *ultracontigs*.

Para realizar el ensamble de secuencias, los científicos usaron la información obtenida de más de 33 millones de experimentos individuales de secuenciación y dos sistemas de computación (ARACHNE y PHUSION), conocidos en inglés como *genome assemblers*, desarrollados en el *Whitehead Institute* y el *Sanger Institute*, respectivamente. La secuencia borrador obtenida incluye más del 96% del genoma del ratón y representa una cobertura equivalente a siete veces el genoma. Esto quiere decir que cada nucleótido en el genoma fue determinado, en promedio, hasta siete veces, lo que asegura la alta precisión de la secuencia.

Los resultados de este esfuerzo indican que el genoma del ratón tendría un tamaño de alrededor de 2.500 Mb (2,5 gigabases), 14% más chico que el genoma humano (2,9 gigabases); esta reducción en el tamaño estaría reflejando una mayor proporción de deleciones en el genoma de los ratones. No obstante, el análisis preliminar de las secuencias permite prever un número de genes (codificantes de proteínas) cercano al del hombre (alrededor de 35.000). Alrededor del 80% de los genes del ratón posee un gen ortólogo (ascendencia común) reconocible en el genoma humano. Al nivel de sus secuencias nucleotídicas, alrededor del 40% del genoma humano puede ser alineado con la secuencia del ratón.

Los resultados de estos análisis pueden encontrarse (con acceso público y gratuito) en varios sitios de Internet, incluyendo <http://mouse.ensembl.org> (*European Bioinformatics Institute*),

<sup>2</sup> El cromosoma Y del ratón está siendo secuenciado en forma independiente por "escopetazo jerarquizado".

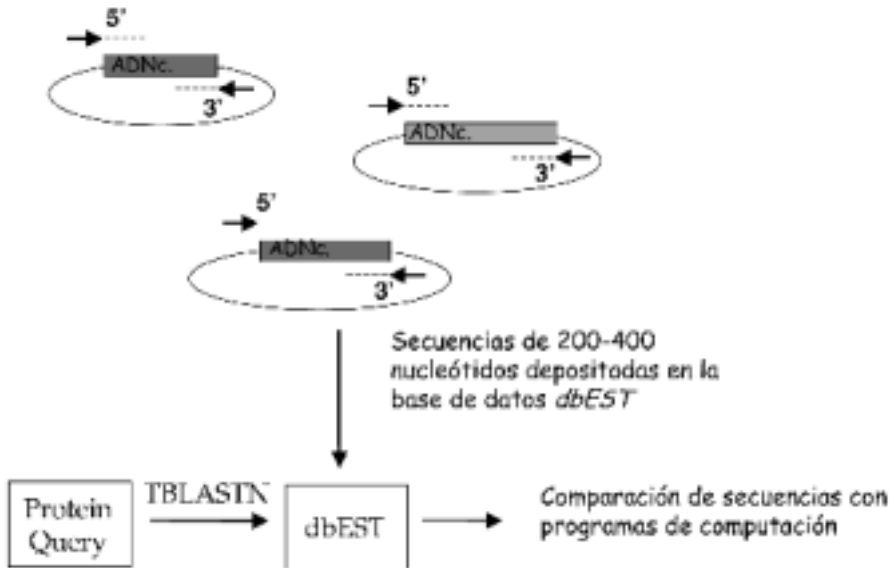
<http://www.ncbi.nlm.nih.gov/genome/guide/mouse> (National Center for Biotechnology Information at the National Library of Medicine) y <http://genome.ucsc.edu> (University of California, Santa Cruz). Los tres sitios incluyen comparaciones de las secuencias del ratón con la del genoma humano. La próxima fase del proyecto internacional es producir una versión "terminada", llenando todos los espacios y resolviendo los errores e incongruencias. Esta etapa será realizada exclusivamente por escopetazo jerarquizado, haciendo uso del mapa físico del ratón. La disponibilidad de la secuencia final del genoma tendrá grandes implicancias para la genética del ratón de laboratorio y ayudará al avance de las ciencias biomédicas. Las áreas más beneficiadas serán el clonaje posicional de genes, la identificación de loci responsables de rasgos cuantitativos (QTL's), la creación de mutaciones dirigidas (ratones *knock-out* y *knock-in*) y la genómica funcional (del inglés *functional genomics*), entre otras. Por otra parte, la empresa privada *Celera Genomics* ya ofreció su versión del borrador del genoma murino el año 2001, realizado con las líneas DBA/2, AJ y 129X1/Sv (<http://www.celera.com/products/mouse.cfm>), pero el acceso a la información no es gratuito.

Al momento de este escrito se acaba de hacer público el primer ensamble (versión 2.0) del genoma de la rata *Rattus norvegicus* (línea BN/SsMCMW), haciendo uso del mismo sistema mixto que el utilizado para el ratón. Este borrador inicial cubre más del 90% de las estimadas 2.800 Mb (2,8 gigabases) que tendría este genoma (ver <http://www.hgsc.bcm.tmc.edu/projects/rat/> y <http://rgd.mcw.edu>). El consorcio encargado del proyecto (*Rat Genome Sequencing Consortium*) está liderado por el *Baylor College of Medicine Human Genome Sequencing Center*, más colaboraciones de *Celera Genomics*, *Genome Therapeutics*, *The Institute for Genome Research*, *The University of Utah*, *Children's Hospital Oakland Research Institute*, *Medical College of Wisconsin*, y *BC Genome Sciences Center*.

Aún cuando estén mapeados y secuenciados todos los genes de estas especies, estará por delante la tarea más ardua: aquella de analizar su función, su regulación y su interacción. La denominada era "**pos-genómica**" (en inglés *post-genomics*) ya está entre nosotros y será esencial contar con modelos animales para el estudio funcional de las secuencias obtenidas de los proyectos del genoma humano y murino. La secuenciación del genoma del ratón tiene dos razones fundamentales: el conocimiento científico propiamente dicho y la aceleración en la detección de las regiones codificantes de los genes en las dos especies, ya que son las únicas regiones que conservan una alta homología.

### 6.2.5 Mapas de Transcripción

Más allá del tremendo valor que tendrán las versiones completas (y analizadas) de las secuencia genómicas para la biomedicina, no es razonable esperar hasta ese momento para diseñar **mapas de transcripción** (del inglés *transcript maps* o *gene maps*). Es por eso que se está trabajando actualmente en construir mapas sobre la base de fragmentos cortos de ADNc, amplificados por PCR a partir de extremos no-traducidos de ARNm, denominados **EST's** (*Expressed Sequence Tags*) (**Figura 6.6**). Se ha logrado un progreso enorme en la identificación de



**Figura 6.6. Las secuencias EST.** Los EST (*Expressed Sequence Tags*) son secuencias parciales provenientes de los extremos (tanto 3' como 5') de clones de ADNc, obtenidas por medio de una secuenciación de pasaje simple. La estrategia de los EST se desarrolló para poder lograr una identificación rápida de los genes que se expresan gracias a la posibilidad de contar con secuencias únicas que identifiquen cada clon de ADNc. Todas las secuencias EST se depositan en la bases de datos *dbEST* y están disponibles para comparaciones con programas de computación.

secuencias parciales obtenidas de clones de ADNc (EST's) en el hombre y en el ratón, conociéndose incluso su posición en el mapa genético, lo que ha generado los denominados mapas de transcripción. La colección pública de EST's humanos es de cerca de cuatro millones de secuencias, entre las cuales se encuentran representados el 90% de los genes humanos. Para el ratón, el número de EST's es alrededor de tres millones y para la rata en el orden de los 800.000 (dbEST: <http://www.ncbi.nlm.nih.gov/dbEST/index.html>).

El primer mapa de transcripción basado en híbridos de radiación (panel T31) para el ratón fue presentado por Philip Avner y colaboradores en el año 2001 (*Nature Genetics*). Este mapa de primera generación emplaza alrededor de 6.000 marcadores genéticos de los cuales 4.000 son EST's; éstos serán útiles en la identificación de muchos de los mutantes generados por mutagénesis química por ENU (ver Capítulo VII), por medio del enfoque "gen candidato". Mapas de este tipo existen también para la rata, el perro, el gato, el pez cebra y el hombre. Los EST's sirven como sondas moleculares para la identificación de secuencias del genoma que están siendo expresadas y que seguramente corresponden a genes. Así mismo, se está trabajando en la localización de EST's humanos en el genoma de la rata y el ratón (utilizando iniciadores humanos), lo que permitirá descubrir nuevas homologías entre estas especies. Aunque su uso generó controversias en un comienzo, en la actualidad los EST's se han con-



vertido en una herramienta de rutina, junto con los híbridos de radiación, los YAC's y los BAC's (estos últimos han sido esenciales para preparar el terreno a la secuenciación completa del genoma humano, la rata y el ratón).

En Japón, los laboratorios del *RIKEN Genomic Sciences Center* (<http://genome.gsc.riken.go.jp/home.html>) se encuentran trabajando en una enciclopedia de ADNc completos del ratón (*full-length cDNA encyclopedia*), lo que sin dudas acelerará el descubrimiento de nuevos genes. Es, al día de hoy, la colección más completa de unidades de transcripción (**transcriptoma**) en los mamíferos. Actualmente, este centro cuentan con alrededor de 60.000 ADNc completos (ver *RIKEN Mouse Full-length cDNA Encyclopedia* <http://genome.gsc.riken.go.jp/activity.html#3>), lo que implica que muchos genes del ratón (más del 30%) están utilizando el mecanismo de *splicing* alternativo (ver Capítulo I). Para mayor información sobre los genes que se expresan durante el desarrollo embrionario del ratón consultar el *Edinburgh Mouse Atlas of Gene Expression* (EMAGE) <http://genex.hgu.mrc.ac.uk/Emage/database/intro.html>.

### 6.2.6 Integración de Mapas (Mapas Consenso)

Para una cruce de animales determinada, se puede establecer un mapa de ligamiento exclusivamente para los marcadores que segregan en la misma pero, lamentablemente, el número de esos marcadores es muchas veces limitado. Esta limitación se extiende también al polimorfismo entre las líneas parentales (o las técnicas utilizadas para detectarlo). Por eso, es muy importante combinar los resultados obtenidos en distintas cruces independientes en un **mapa consenso**. Esto es posible solamente cuando los datos obtenidos de cruces independientes poseen marcadores en común.

Para ayudar a la integración de los diversos mapas genéticos, los genetistas han definido un grupo de loci de referencia que se encuentran uniformemente distribuidos en todos los cromosomas y son también muy polimórficos. Los mapas consenso actualizados para el genoma del ratón y la rata pueden obtenerse vía Internet: (<http://www.informatics.jax.org>), (<http://rgd.mcw.edu>) y (<http://www-genome.wi.mit.edu/>).

### 6.2.7 Programas de Mapeo a Gran Escala

Varios programas de mapeo a gran escala se han venido desarrollando en los últimos 12 años en el genoma del ratón. Históricamente, los dos más importantes son los llevados a cabo en el *Institut Pasteur*, París, Francia (Guénet y colaboradores) y en el *Frederick Cancer Research and Development Center*, Maryland, USA (Copeland y colaboradores), sumándose en la actualidad los programas del *Jackson Laboratory*, el *Mouse Genome Project* del *National Center for Human Genome Research* (NIH) y el programa en colaboración entre el MRC Británico y el *Institut Pasteur* (*European Collaborative Interspecific Backcross* o EUCIB). El panel de retrocruzas

del EUCIB (*Mus spretus* y C57BL/6) es en la actualidad el más grande del mundo con un número de 982 muestras de ADN evaluadas para loci de referencia a lo largo del genoma. Con este panel del EUCIB, dos marcadores genéticos que aparezcan ligados en las casi 1000 muestras de ADN (porque no logran recombinar) estarán a una distancia menor a 0,3 cM con un nivel de confianza de 95% (equivalente a 500 kb en el genoma del ratón). Estos programas proveen los medios necesarios para la realización de mapas de alta resolución, y los consecuentes mapeos físicos en el ratón, los cuales aportarán información muy útil para el mapeo comparativo entre especies.

### 6.2.8 Mapeos Comparativos

El mapeo comparativo entre los vertebrados (**genoma comparativo**), y en particular entre los mamíferos, está aportando una nueva dimensión a los proyectos de mapeo del genoma humano y de otras especies de interés científico o zootécnico. Estos trabajos se están realizando fundamentalmente con la técnica de pintado de cromosomas o **Zoo-FISH** y son coordinados por la *Comparative Genome Organization*. Teniendo en cuenta que el genoma de los mamíferos está muy conservado, la combinación de la información genética proveniente del hombre y del resto de los mamíferos está siendo analizada en beneficio de varios proyectos conjuntos.

Los mamíferos evolucionaron hace aproximadamente 200 millones de años a partir de un grupo de reptiles (**sinápsidos**), y son sus únicos descendientes vivos. Existen tres grandes grupos de mamíferos: **Euteria** (placentados), **Metateria** (marsupiales) y **Prototeria** (monotremas). Los marsupiales divergieron de los placentados hace 120-150 millones de años, mientras que los monotremas lo hicieron hace 200 millones de años. Los mamíferos placentados poseen un genoma extremadamente conservado y probablemente compartan la gran mayoría de su conjunto de 30.000 a 40.000 genes.

Curiosamente, este genoma común lo podemos encontrar dividido desde tres pares de cromosomas grandes ( $2n=6$ ) en el **Muntjac Indio** (*Muntiacus muntjak*), hasta 46 pares pequeños ( $2n=92$ ), en el **rinoceronte negro** (*Diceros bicornis*). Esta variación cariotípica hizo que, durante muchos años, los investigadores imaginaran que el orden de los genes en las diferentes especies estaba muy mezclado. Además, por mucho tiempo se asumió que, como el ratón y el hombre mostraban tantos cambios provenientes de rearrreglos cromosómicos, los genomas de los mamíferos placentados, los marsupiales y los monotremas estarían "revueltos" hasta el nivel de no poder reconocer los segmentos de origen común. El mapeo comparativo está probando que esto no es así y que el grado de conservación entre especies es sorpresivamente grande. Por ejemplo, se han identificado grupos de genes ligados en el humano que también forman grupos (del inglés *clusters*) en el pollo y hasta en los peces. Otro dato particularmente llamativo es la conservación de cromosomas enteros entre los primates y otros ordenes de mamíferos. Por ejemplo, el cromosoma X se encuentra casi completamente conservado en todos los mamíferos placentados y los cromosomas 4 y 17 del humano aparecen casi intactos en los carnívoros, artiodáctilos, cetáceos e insectívoros.

Por otra parte, el genoma de los vertebrados parece haber permanecido estable por algo así como 400 millones de años (la conservación parece ser la regla en los genomas de las aves y los peces) y las diferencias que se observan entre los genomas de los placentados serían más bien perturbaciones recientes en la evolución. En particular, el mapa del ratón se presenta especialmente "roto" con respecto al humano, con alrededor de 180 rearrreglos que los separan. El tamaño promedio de los segmentos conservados entre estas dos especies es de sólo 7 cM, aunque debemos tener en cuenta que sólo se conocen aquellos segmentos de fácil detección. En forma similar, se pudo identificar (por medio del uso de zoo-FISH) la presencia de 49 segmentos cromosómicos que se encuentran conservados entre la rata y el ratón. Analizando las regiones con sondas fluorescentes ("pintadas"), prácticamente la totalidad del genoma de la rata está cubierto de regiones homólogas del ratón, con la excepción de pequeños segmentos cerca de los centrómeros y de los brazos cortos de los cromosomas 3, 11, 12, y 13. Podríamos decir que los placentados y en especial el ratón, con su gran número de rearrreglos, son los animales que rompieron la regla de la evolución de cromosomas. Al día de hoy, se desconocen los mecanismos de rearrreglos de los cromosomas durante la especiación y también su finalidad, si es que la hay.

La disponibilidad de un mapa detallado del genoma humano y el ratón ayudará a diseñar los mapas de las otras especies de animales de laboratorio, incluidas aquellas especies no tradicionales. La información genética puede fluir también en la dirección opuesta: la información de los mapas en las diferentes especies (notar que, después del mapa humano, los cuatro mapas de ligamiento más completos son los del ratón, la rata, el bovino y el cerdo, en ese orden) puede ser transferida al mapa humano en el marco de las regiones homólogas.

### 6.2.9 Localización de rasgos cuantitativos (*Quantitative Trait Loci*)

Hasta ahora hemos visto los mapas realizados en base a rasgos de herencia simple (**rasgos monogénicos**), como es el caso de muchas mutaciones del ratón y la rata. Sin embargo, muchos rasgos y enfermedades de interés médico, como la presión arterial, la obesidad y la resistencia a infecciones son naturalmente de carácter cuantitativo; por lo tanto se ven influenciados por muchos genes simultáneamente (**rasgos multigénicos**) y por factores ambientales. Debido a que este tema excede los propósitos de este libro, daremos sólo una breve reseña de los pasos a seguir para localizar loci responsables de **rasgos cuantitativos** (del inglés *quantitative trait loci* o **QTL**). Se considera que el proceso que lleva del descubrimiento de un QTL a la identificación del gen responsable involucra cinco pasos:

- (i) localizar el QTL en un segmento cromosómico. Esto implica la generación de cientos de animales de retrocruza (N2) o intercruza (F2) que serán genotipados con marcadores moleculares espaciados entre 15 y 20 cM; normalmente se utilizan entre 75 y 100 marcadores microsatélites para cubrir todo el genoma (*whole genome scan*). Este grupo de animales debe ser analizado, a su vez, para el rasgo de interés, de manera de poder clasificar a los animales siguiendo una graduación de fenotipos. El objetivo de este mapeo es identificar aquellas re-

giones del genoma que están contribuyendo a la variación en el rasgo de interés. En el caso de una retrocruza teórica AA × AB, habrá que asociar el fenotipo de un animal en particular con el hecho de haber heredado (del parental F1) el alelo A o el alelo B. Las asociaciones entre los marcadores y el fenotipo se evalúan de manera de calcular la probabilidad estadística de que un QTL se encuentre cerca de un marcador. Estos valores son presentados normalmente como **LOD score** (del inglés, *log of the odds*), considerándose ligamiento cuando este valor es igual o mayor a 3. Para tal propósito existen programas de computación, entre los cuales se encuentra la familia de programas Map Manager (<http://mapmgr.roswellpark.org/mapmgr.html>). Esta primer etapa lleva mucho tiempo, cientos de análisis de fenotipo y decenas de miles de genotipados, y además se considera que el QTL queda localizado con baja precisión.

- (ii) Construcción de una línea congénica (ver Capítulo IV) con el objetivo de aislar genéticamente el QTL (localizado en la etapa anterior) de otros QTL's (normalmente se realiza un grupo de líneas congénicas abarcando los distintos QTL's). De esta manera, manipulamos un rasgo poligénico para estudiarlo como monogénico. La única limitante es que aquellos QTL's que estén localizados muy cerca en la misma región cromosómica, seguramente no serán separados.
- (iii) Reducir la región de interés a un tamaño menor a 1 cM por medio de nuevas cruza entre la línea congénica y la línea parental, con el objetivo de hacer una localización más precisa del QTL. Una vez más, esta etapa incluye análisis de fenotipo y genotipados.
- (iv) Búsqueda de genes candidato. Como veremos más adelante, la identificación y evaluación de genes candidato es un trabajo muy arduo ya que, en promedio, un segmento de 1 cM contendrá más de 30 genes para analizar.
- (v) El paso final será la prueba de identidad del gen candidato, lo que se suele hacer por técnicas de manipulación genética de embriones (transgénesis o mutagénesis dirigida).

Es necesario aclarar que estos métodos convencionales de mapeo de QTL pierden poder a medida que la complejidad genética del rasgo aumenta. El principal inconveniente es que la habilidad para detectar QTL's que tienen poca influencia en el fenotipo (débiles) es función del número de animales analizados. Normalmente, el análisis de cientos de animales permite detectar aquellos QTL que aportan más del 10% de la varianza del fenotipo. Como hemos visto, en el caso del ratón, las líneas recombinantes congénicas y las líneas consómicas (ver Capítulo IV) son una alternativa interesante para el mapeo de QTL's, otras alternativas son el uso de **grupos heterogéneos** (del inglés *heterogeneous stocks*) o de las **líneas de intercruza avanzada** (del inglés *advanced intercross lines*). A pesar de disponer de estas estrategias, hasta el momento de este escrito sólo un puñado de genes QTL han sido identificados (ocho en el ratón y tres en la rata), situación que seguramente se revertirá con la disponibilidad de las secuencias de los genomas de estos roedores. Aquellos lectores interesados en los rasgos cuantitativos pueden consultar la página de Internet del *Complex Trait Consortium* (<http://www.complextrait.org/>).

## 6.3 Marcadores utilizados en la construcción de mapas de ligamiento

Cualquier tipo de cambio en el ADN que haga a un individuo o grupo de individuos diferente de los otros miembros de la misma especie puede ser considerado como un potencial marcador genético (suponiendo que contamos con los medios, directos o indirectos, de seguirlo a través de las generaciones). Los marcadores genéticos que se vienen utilizando han sido seleccionados por tres características fundamentales: abundancia, fácil evaluación y técnicas de detección de bajo costo.

### 6.3.1 Marcadores Fenotípicos

Como fue mencionado anteriormente, el mapa de ligamiento del ratón empezó a formarse en 1915, con la observación de un ligamiento entre el locus albino (*Tyr<sup>r</sup>*) y el locus *pink eye dilution* (*p*). Desde esta observación inicial, el mapa genético se ha enriquecido y crecido en densidad, gracias al descubrimiento de nuevos alelos mutantes en diferentes loci del genoma murino. La acumulación de estas mutaciones fue debida, en parte, al uso del ratón como modelo experimental en estudios de radiaciones y mutágenos químicos; y en parte también a la elección de la endocría como esquema reproductivo para mantener las líneas consanguíneas (la consanguinidad facilita la detección de mutaciones recesivas).

Más de 200 alelos mutantes con efectos sobre el color del pelaje, los ojos, el esqueleto, el comportamiento etc., han contribuido al establecimiento de la "columna vertebral" del actual mapa de ligamiento del ratón (algunos todavía son considerados marcadores útiles por algunos genetistas). Desafortunadamente, tienen dos grandes inconvenientes y ya no representan una fuente importante de nuevos marcadores:

- (i) Frecuentemente alteran la viabilidad y/o fertilidad de los animales afectados y resulta muy difícil establecer cruces con más de tres o cuatro marcadores de esta naturaleza segregando al mismo tiempo.
- (ii) Si bien la mayoría de las mutaciones que ocurren *de novo* en el ratón afecta a nuevos loci, el número total de este tipo de mutaciones se mantiene, en términos de marcadores genéticos, relativamente bajo. Finalmente, los alelos mutantes detectados por la inspección grosera del fenotipo externo de los animales pueden ser interesantes por su efecto sobre el desarrollo o como modelos de enfermedades genéticas, pero ya no representan una fuente importante de marcadores genéticos.

### 6.3.2 Marcadores Proteicos

El desarrollo de los geles de electroforesis en la década de 1960 permitió la identificación de nuevos loci, atribuibles a polimorfismos resultantes de variaciones en las cargas eléctricas de algunas proteínas. Estas proteínas tienen funciones similares pero diferentes estructuras primarias (ej: **isoenzimas**). Este tipo de marcadores moleculares conocidos como variantes electroforéticas (ej: **aloenzimas**), tienen tres ventajas sobre los clásicos marcadores fenotípicos: (i) se expresan en forma co-dominante, por lo tanto pueden ser evaluados en animales heterocigotas, (ii) son en general compatibles con el normal funcionamiento de la proteína (enzimas) y no alteran la viabilidad ni la fertilidad de los portadores y (iii) son relativamente abundantes, probablemente por ser neutros en términos de selección, y tienen una distribución amplia en el genoma (existen actualmente más de 100 marcadores de este tipo y siguen apareciendo nuevas variantes).

Estos polimorfismos han contribuido al rápido desarrollo del mapa genético del ratón, pero tienen el inconveniente de requerir de técnicas tediosas y específicas para cada enzima o proteína, lo que resulta muy caro y dificulta los proyectos de mapeos de gran envergadura. Además, cabe recordar que no más del 3 % del ADN genómico es finalmente traducido a proteínas y que sólo unas pocas de éstas pueden mutar sin alterar la viabilidad de los animales. Por lo tanto, es obvio que si los marcadores proteicos fuesen los únicos disponibles, serían insuficientes para el desarrollo de un mapa de alta densidad.

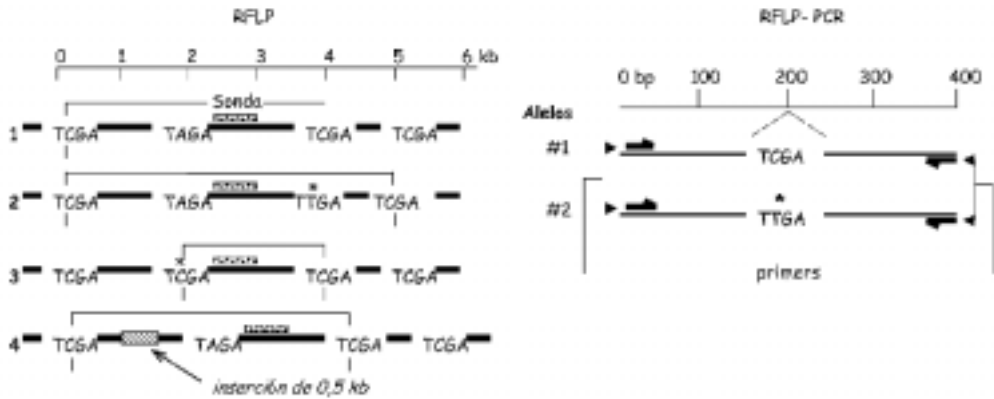
### 6.3.3. Marcadores de ADN

Todos los marcadores de ADN usados para construir mapas de ligamiento pertenecen a los llamados **marcadores de tipo I**, secuencias no codificantes y por lo tanto muy variables. Sin embargo, hay que tener en cuenta que los **marcadores de tipo II** (genes codificantes) son muy útiles para la genética comparativa.

#### 6.3.3.1 RFLP's (*Restriction Fragment Length Polymorphisms*)

Con la llegada de las técnicas de genética molecular (*Southern blot*, sondas moleculares etc.), surgieron los primeros polimorfismos en las secuencias de ADN que son los **RFLP's** o polimorfismos de longitud en fragmentos de restricción. Los fragmentos generados por la digestión de una muestra de ADN genómico, cuando son separados en un gel de agarosa e identificados por una sonda molecular marcada, exhiben polimorfismos determinados por la presencia o ausencia de un sitio de enzima de restricción específico. Al comportarse como caracteres mendelianos co-dominantes, los RFLP's son muy útiles como marcadores genéticos (**Figura 6.7**).

Sin embargo, la variabilidad genética de estos marcadores es relativamente baja en el hombre y los animales y la búsqueda de nuevos RFLP's es un trabajo caro, largo y tedioso. Este traba-



**Figura 6.7. Eventos moleculares que generan polimorfismos RFLP.** El cromosoma 1 es la secuencia ancestral con tres sitios de restricción para la enzima *TaqI* (TCGA). La banda en el gel será de 4 kb. El cromosoma 2 sufrió una mutación puntual C por T (asterisco) que elimina el sitio de restricción (la banda será de 5 kb). El cromosoma 3 sufrió una mutación A por C (asterisco) lo que creó un nuevo sitio de restricción (banda de 2 kb). El cromosoma 4 sufrió una inserción de 0,5 kb (banda de 4,5 kb). A la derecha, un polimorfismo **RFLP-PCR**. Los iniciadores (*primers*) amplifican un segmento que contiene el sitio de restricción para *TaqI*. El producto inicial de PCR será el mismo para ambos alelos (400 pb) pero luego de la restricción con *TaqI*, el alelo #1 dará 2 bandas de 200 pb y el alelo #2 (que perdió el sitio de restricción) será de 400 pb. Adaptado de Silver L. M. (ed) *Mouse Genetics. Concepts and applications*. Oxford University Press, Oxford, 1995.

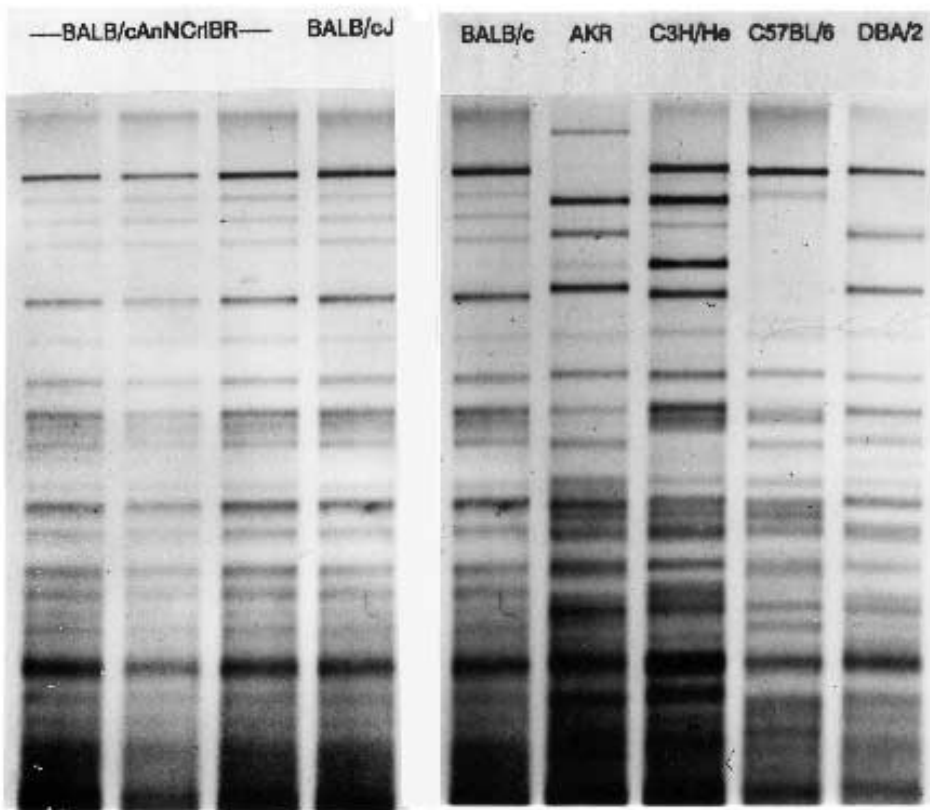
jo incluye técnicas de *Southern blot* con el uso de isótopos radioactivos (como  $^{32}\text{P}$  o  $^{33}\text{P}$ ), grandes cantidades de ADN genómico (típicamente 5 a 10 mg) y varias enzimas de restricción (aunque existen miles de enzimas de restricción sólo se comercializan algunos cientos). Además, en el caso del estudio de pedigrís humanos, el bajo número de alelos que presenta un locus de RFLP (la mayoría son di-alélicos) produce una disminución en el porcentaje de heterocigosis con la consecuente pérdida de valor como marcador en los estudios de ligamiento.

### 6.3.3.2 Minisatélites

En 1980 A. R. Wyman y R. White descubrieron, por azar, la primera **región hipervariable** del ADN humano, y subsecuentemente se fueron hallando en el genoma humano otras regiones de ese tipo cerca de los genes de la  $\alpha$  globina, la insulina, la mioglobina y otros. Estas secuencias no codificantes e hipervariables del ADN denominadas "**minisatélites**", están compuestas por "motivos" (del inglés *motifs*) cortos (de una longitud aproximada de entre 10 y 50 pb) repetidos en *tandem* un número variable de veces, formando secuencias de entre 5 y 50 kb de longitud. Esto genera una gran cantidad de alelos diferenciados por el número de repeticiones que posee cada uno. Esta última característica se ve reflejada en el alto porcentaje de heterocigosis (generalmente mayor al 80 %) alcanzado para estos loci, lo que los convierte en marcadores altamente informativos en estudios de análisis de ligamiento y pruebas de identificación. A diferencia de los RFLP's, estos polimorfismos no surgen por la presencia o ausencia de un sitio de restricción sino por el número variable de repeticiones en *tandem* en cada alelo.

En estas técnicas, los sitios de restricción de las enzimas se encuentran flanqueando las secuencias repetidas en tandem.

Muchos minisatélites comparten entre sí una secuencia más corta (del orden de 15 pb) llamada "core", formando así familias de minisatélites con un core en común. Las sondas moleculares desarrolladas por Jeffreys y colaboradores en el año 1985 a partir de secuencias repetitivas de un intrón de la mioglobina humana, dieron origen a la técnica conocida como "**fingerprinting de ADN**". La misma ha revolucionado el campo de la medicina forense y los estudios de paternidad en todo el mundo. El fingerprinting de ADN es el patrón de bandas único y específico de individuo (excepto para los gemelos monocigóticos) obtenido por la hibridación de muestras de ADN digerido con sondas basadas en la secuencia core, capaces de detectar múltiples minisatélites a lo largo del genoma (**fingerprinting multi-locus**) (Figura 6.8).



**Figura 6.8. Fingerprinting de ADN.** La foto muestra los patrones de bandas generados por la técnica de *Fingerprinting de ADN* utilizando la sonda 33.6 desarrollada por Jeffreys y colaboradores. Como se puede apreciar, cada cepa consanguínea presenta un patrón de bandas individual (*fingerprint*).



En el caso del ratón y la rata de laboratorio, los fingerprinting de ADN generados por las sondas de Jeffreys 33.6 y 33.15 evidencian una complejidad e intensidad de bandas comparables a las obtenidas por las mismas sondas en humanos. Esto nos ofrece la ventaja de generar *fingerprints* específicos de línea que sirven para el control de la calidad genética de las líneas consanguíneas (ver Capítulo IV). Al día de hoy ya no se utilizan más en el desarrollo de mapas genéticos.

En el año 1987 Nakamura y sus colegas definieron los VNTR's (*Variable Number of Tandem Repeats*) secuencias polimórficas usadas como marcadores en los mapeos de genes. Estos investigadores introdujeron un nuevo enfoque en el manejo de los minisatélites al desarrollar una serie de sondas locus-específicas que ya no generan un patrón de bandas o fingerprinting sino una o dos bandas por locus de VNTR, según sea homocigota o heterocigota, técnica denominada **fingerprinting locus-específico**. Hasta el momento, se han descrito y mapeado cientos de loci de VNTR en el genoma humano y de algunos animales, aunque se calcula que su número podría ser de varios miles por genoma.

### 6.3.3.3 Microsatélites

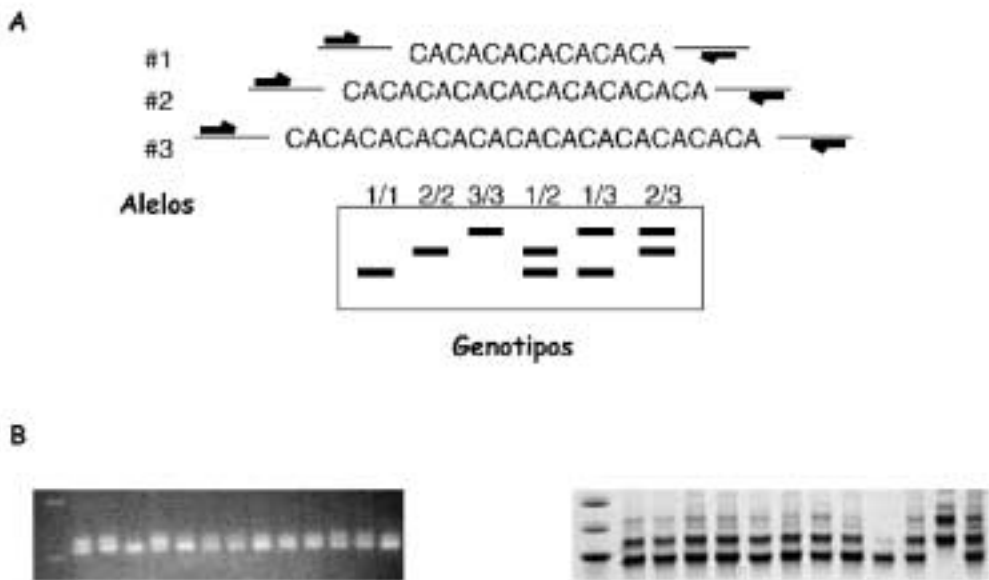
A fines de la década de 1980, se da a conocer la existencia de polimorfismos de ADN conteniendo repeticiones en tandem de dos nucleótidos [en particular  $(CA)_n$  y  $(GA)_n$ ] además de otros formados por secuencias de entre 1 y 5 nucleótidos. A estas secuencias se las conoce como **microsatélites**, **SSLP** (*Simple Sequence Length Polymorphisms*), **STR** (*Simple Tandem Repeats*) o **SSR** (*Simple Sequence Repeats*) y podemos considerarlas como otro tipo de VNTR. Los microsatélites se encuentran en los genomas de todos los eucariotas, desde las plantas hasta los mamíferos, aunque existen diferencias en cuanto a su densidad. Por ejemplo, los estudios comparativos de los genomas humano y del ratón mostraron que éste tiene entre dos y tres veces más microsatélites que el humano.

El número de repeticiones, por ejemplo del motivo  $(CA)_n$ , comprende un rango de entre 15 y 40, lo que facilita la posibilidad de amplificación de estos pequeños segmentos por medio de la técnica de PCR para el análisis de los polimorfismos. El uso masivo de estos marcadores en la década de 1990 los convirtió en la "vedette" entre los marcadores moleculares, siendo su único inconveniente el hecho de que los iniciadores son, en general, específicos de cada especie. Su frecuencia es mayor que la de los minisatélites, calculándose que su número podría superar los 200.000 por genoma. Por ejemplo, se calcula que existen cerca de 100.000 copias del dinucleótido  $(CA)_n$  (con un n mayor a 10) en el genoma del ratón, dando un promedio de un locus cada 30 kb. Además, presentan la ventaja de estar distribuidos al azar, en intrones, regiones codificantes o intergénicas; aunque con una leve tendencia a presentar una mayor densidad en la región distal (telomérica) de los cromosomas.

El uso de los microsatélites en el ratón y la rata aceleró los trabajos de mapeo del genoma gracias a su abundancia (hay más de 8.000 descriptos para ambas especies, lo que asegura, en promedio, un marcador cada 400 kb) y estabilidad, características que los convierten en excelentes marcadores genéticos. Aproximadamente, el 50 % de los microsatélites descriptos has-

ta el momento en el ratón muestra polimorfismo entre las distintas líneas consanguíneas de laboratorio, aunque ese porcentaje se eleva al 77% entre la línea C57BL/6 y *Mus musculus castaneus*, y al 90% si se trata de *Mus spretus* (ver más adelante el punto 6.4.5.1). Muchas de estas variaciones, en el rango de 2 a 40 pb, son detectables por la observación de los productos de PCR en geles de agarosa o poliácridamida (Figura 6.9). La lista completa de SSLP's del ratón y la rata, la secuencia de sus primers, el tamaño de sus productos y el polimorfismo entre las líneas de laboratorio puede encontrarse en Internet (<http://www.informatics.jax.org> y <http://www.genome.wi.mit.edu>). En la rata, se encuentra disponible información sobre tamaños de alelos de 4.500 SSLP en 48 líneas consanguíneas en el *Rat Genome Database* (ver <http://rgd.mcw.edu/GENOMESCANNER>). Los iniciadores para los SSLP del ratón (*Mouse MapPairs® Primers*), la rata (*Rat MapPairs® Primers*) y otras especies pueden adquirirse a precios accesibles en *Research Genetics Inc.* (<http://www.resgen.com>).

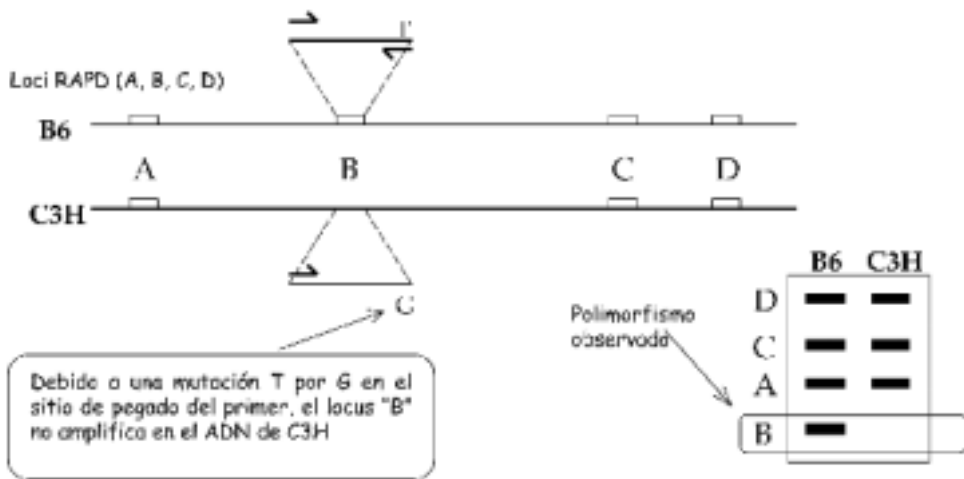
Aquellos lectores interesados en otros aspectos de los microsatélites, como ser sus posibles funciones en la regulación de genes, sus usos en genética poblacional y estudios de evolución, pueden consultar el libro *Microsatellites: Evolution and Applications* de DV Goldstein y C Schlotterer C (2000).



**Figura 6.9. Los marcadores microsatélite. A.** Detección de alelos de microsatélite por PCR. Se muestran tres alelos diferentes de un locus compuesto de repeticiones CA y los respectivos genotipos en la electroforesis. **B.** Fotografías de electroforesis de productos de microsatélites del ratón por PCR. A la izquierda un gel de agarosa teñido con bromuro de etidio, a la derecha un gel de poliácridamida teñido con sales de plata. Fotos: F. Benavides, *The University of Texas - M.D. Anderson Cancer Center, Department of Carcinogenesis - Science Park, Texas, Estados Unidos.*

### 6.3.3.4 RAPD's (Random Amplified Polymorphic DNA).

Técnica basada en el uso de iniciadores únicos (o combinados de a pares) cuyas secuencias cortas (generalmente 10 nucleótidos) fueron diseñadas al azar, en reacciones de PCR de muy bajo requerimiento. Este es actualmente el método de elección en las especies menos estudiadas (por ejemplo es muy usado en peces) debido a que (i) no requiere información previa sobre secuencias de ADN en el organismo a estudiar, (ii) se identifican muchos marcadores moleculares simultáneamente, y (iii) los costos son relativamente bajos y no se requiere laboratorio ni mano de obra muy sofisticada. Pero debido a las altas concentraciones de  $Cl_2Mg$  y las bajas temperaturas de pegado de los iniciadores presenta la desventaja de tener baja "repetibilidad". Además, las bandas amplificadas no presentan variantes de tamaño (alelos), simplemente el producto de PCR está presente o ausente, lo que no permite diferenciar animales heterocigotas de homocigotas para un locus determinado (Figura 6.10).



**Figura 6.10. Otros marcadores moleculares.** Detección de loci polimórficos por la técnica de RAPD (AP-PCR). Las líneas horizontales representan los cromosomas de las líneas consanguíneas C3H y C57BL/6 (B6). Las cajas blancas representan fragmentos genómicos (loci RAPD) amplificables por PCR con un iniciador diseñado al azar. Cada iniciador genera un *fingerprinting* individual. Como puede observarse, el locus B es polimórfico entre estas líneas. Adaptado de Silver L. M. (ed) *Mouse Genetics. Concepts and applications*. Oxford University Press, Oxford, 1995.

### 6.3.3.5 SNP's (Single Nucleotide Polymorphisms)

Los SNP's, o "snips", como se lo suele pronunciar en inglés, son el resultado de la presencia de una base variante en cualquier sitio del genoma y constituyen la forma más común de variación genética. Los encontramos tanto dentro de secuencias anónimas de ADN como en secuencias codificantes (genes) y presentan en general dos alelos. Aunque la fuerza informativa de estos marcadores -en los estudios humanos- es menor a la de los microsatélites, se ven

compensados por su enorme frecuencia, calculada empíricamente (promedio) como 1 SNP cada 600 a 1.200 pb. El número total (teórico) de sitios polimórficos SNP en el genoma humano estaría cerca de los 3-4 millones (esto coincide con los 3,4 millones de SNP's hallados por la empresa *Celera* comparando cuatro líneas consanguíneas de ratones). Por consiguiente, estos marcadores bi-alélicos son la última gran revolución en la creación de mapas de alta resolución y alta densidad de marcadores. En particular, esto se vio favorecido por la aparición de métodos de "genotipado" automático, basados en PCR y microchips de ADN, que permiten analizar miles de muestras a la vez.

Si bien se desarrollaron primero en los mapas humanos (ver <http://www-genome.wi.mit.edu/SNP/human/>), hoy en día están tomando mucha importancia en el ratón y la rata. Este tipo de marcadores se está desarrollando velozmente en el ratón debido a la ventaja que ofrecería contar con nuevos marcadores polimórficos, aún entre líneas muy emparentadas. Actualmente, gracias a la secuencia inicial del genoma del ratón, se encuentran identificados alrededor de 80.000 SNP's. Esta colección inicial de SNP's identificó alrededor de 47.000 candidatos a SNP's entre las líneas I29S1/SvImj y C57BL/6j; 20.000 entre C3H/HeJ y C57BL/6j; y 12.000 entre BALB/cByJ y C57BL/6j. La frecuencia de polimorfismos (promedio) entre las líneas clásicas (*M. m. domesticus*) analizadas fue de aproximadamente 1 SNP cada 500-700 pb. Por otra parte, las comparaciones entre cualquier línea clásica y la línea CAST/Ei (derivada de *M. m. castaneus*) muestra un promedio de un SNP cada 200 pb, promedio que se eleva a un SNP cada 80-100 pb en el caso de comparar con líneas consanguíneas derivadas de *Mus spretus*.

Más allá de estos promedios, es interesante ver que la distribución de los SNP's no es en absoluto uniforme. Existen regiones muy grandes que se encuentran desiertas de SNP's (en un estudio de secuencias de C57BL/6 se detectó una región de este tipo de un tamaño de 26 Mb) y otras regiones son muy ricas en SNP's. Estos datos surgieron a partir de los primeros estudios de la secuencia del genoma del ratón, donde se pudo comprobar la presencia de grandes segmentos de ADN con niveles de polimorfismo extremadamente alto (alrededor de 40 SNP's cada 10 kb) y otros bloques de muy baja densidad (alrededor de 0,5 SNP's cada 10 kb). Para obtener información acerca de SNP's en 15 líneas consanguíneas (clásicas y derivadas de ratones salvajes) se puede consultar la base de datos del *Whitehead Institute for Biomedical Research/MIT Center for Genome Research* (<http://www-genome.wi.mit.edu/SNP/mouse/>), *La Roche* (*The mouse SNP database*, <http://mouseSNP.roche.com>) y *Celera* (*Mouse Reference SNP database*, <http://www.celera.com/genomics/academic/home.cfm?ppage=cds&cpage=mousesnps>).

Finalmente, si bien no es un tipo de marcador *per se*, mencionaremos la técnica **SSCP** (del inglés *Single Strand Conformation Polymorphism*) como una herramienta muy útil a la hora de detectar nuevos polimorfismos. Esta técnica detecta polimorfismos en las secuencias propiamente dichas, y no en su longitud, según las estructuras secundarias que adoptan las cadenas simples. Esto se logra comparando la movilidad electroforética de los productos de PCR en geles de poliacrilamida después de su desnaturalización. El método es tan sensible que es capaz de reconocer una mutación puntual en secuencias de entre 80 y 250 pb de longitud. El SSCP es muy utilizado para detectar diferencias alélicas en marcadores microsátélites y SNP.

## 6.4 Cruzas para análisis de ligamiento

La utilización de líneas de ratones de laboratorio favorece los trabajos de mapeo genético porque: (i) es posible establecer cruzamientos a voluntad, (ii) las líneas consanguíneas producen un solo tipo de gameta, y (iii) es posible criar un número ilimitado de animales de un cruzamiento dado. Es importante recordar que, en el ratón como en el hombre, la presencia de formas alternativas (alelos) de los marcadores genéticos en la generación parental es una condición indispensable para los estudios de ligamiento.

Aunque se han usado ocasionalmente técnicas no sexuales en mapeos genéticos del ratón, la mayoría de la información ha provenído, sin ninguna duda, del uso de cruzas informativas de ratones. Estas cruzas son: las retrocruzas y las intercruzas (líneas clásicas entre sí o líneas clásicas con líneas de origen salvaje), de un lado, y las líneas consanguíneas recombinantes (RIS) por el otro.

### 6.4.1 Retrocruzas

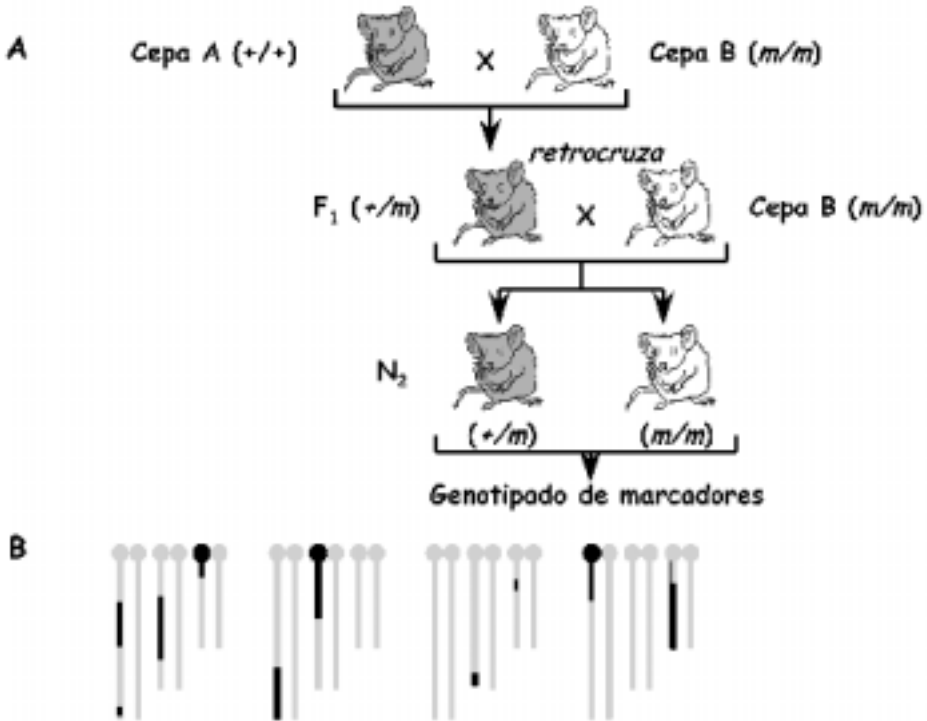
La **retrocruza** (del inglés *backcross*, BC) es la cruce de un animal F1 por un representante de cualquiera de las líneas parentales, lo que genera animales N2 (**Figura 6.11**). Los mismos se anotan de la siguiente manera: (C57BL/6xDBA/2)F1xDBA/2, donde C57BL/6 y DBA/2 son las líneas parentales (la hembra, en este caso el híbrido F1, se anota primero). Aun si el entrecruzamiento se produce en ambos reproductores (en el ratón F1 y en la línea parental) sólo las recombinaciones entre los cromosomas del animal F1 darán recombinantes reconocibles. El uso de las retrocruzas para localizar mutaciones está recomendado en los siguientes casos: (i) si el alelo mutante (*Mu*) es dominante, (ii) si el alelo mutante es recesivo y los homocigotas (*mu/mu*) son fértiles.

Veamos un ejemplo teórico del uso del análisis de ligamiento en una retrocruza: A y B son dos marcadores que se quieren analizar para determinar ligamiento, con dos alelos cada locus (A y a; B y b), la retrocruza AaBb X AA BB producirá cuatro genotipos: AaBb, AABb, AaBB y AABb en iguales proporciones, si los marcadores no están ligados. Si, por el contrario, están ligados, la retrocruza producirá una proporción de genotipos recombinantes de acuerdo a la fuerza de ese ligamiento. Si **f** es el número de genotipos recombinantes observado y **n** el número total de animales, el valor **r** (porcentaje de recombinación) será dado por la fórmula:

$$r = \frac{f}{n}$$

El desvío estándar estará dado por la siguiente fórmula:

$$SE_r = \sqrt{\frac{r(1-r)}{n}}$$

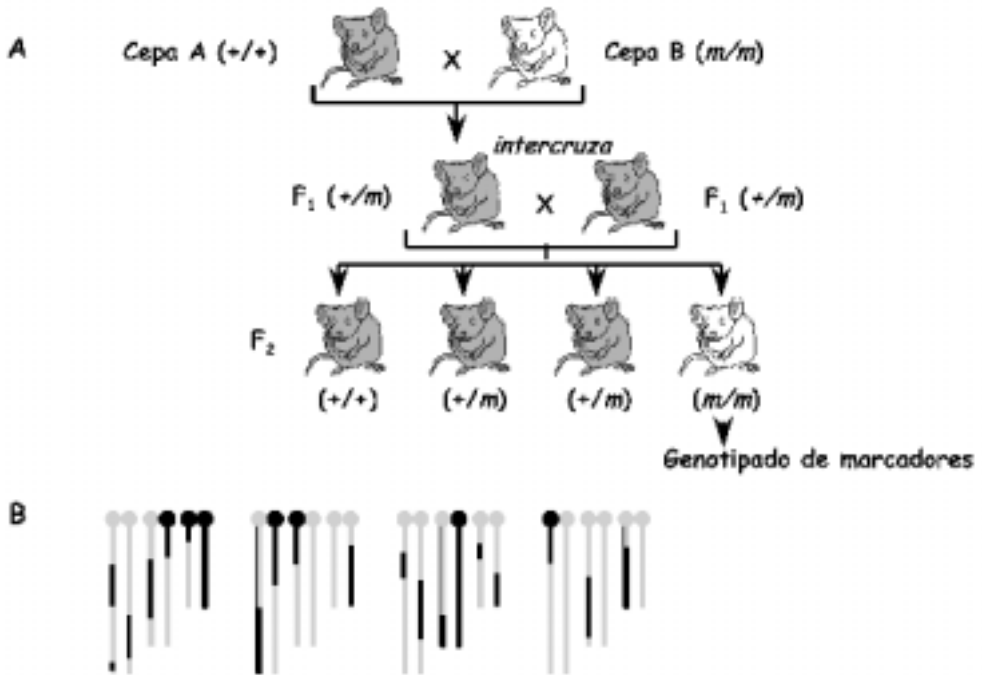


**Figura 6.11. Las retrocruzas. A.** Esquema de una retrocruza. En el ejemplo, la cepa B es homocigota para una mutación recesiva ( $m$ ) que no altera la viabilidad ni la fertilidad. Todos los ratones de la generación  $N_2$  pueden ser genotipados sabiendo que, en promedio, 50% serán heterocigotas (normales) y 50% homocigotas con fenotipo mutante. **B.** Tres pares de cromosomas [representando grande (100 cM), mediano (75 cM) y chico (50cM)] de cuatro progenies independientes de ratones  $N_2$ . Notar que los ratones  $N_2$  reciben siempre un cromosoma entero (no recombinante) de la cepa B.

Con esta fórmula uno puede calcular que un panel de retrocruzas de 300 animales provee un 95 % de probabilidad de detectar ligamiento para dos genes que se encuentren a 1 cM de distancia. Esta probabilidad se eleva a 99 % si el panel de progenie es de 460 animales. De todas formas, 100 a 200 animales  $N_2$  es un número adecuado para comenzar un trabajo de localización de una mutación. (Una vez localizada una mutación entre dos marcadores, para afinar el mapa sólo hace falta estudiar los recombinantes entre estos marcadores, lo que reduce considerablemente el trabajo de genotipado.)

#### 6.4.2 Intercruzas

La **intercruza** (del inglés *intercross*) es el resultado del acoplamiento de dos híbridos  $F_1$ , lo que genera animales  $F_2$  (Figura 6.12), los cuales se anotan de la siguiente manera:



**Figura 6.12. Las intercruzas. A.** Esquema de una intercruza. En el ejemplo, la cepa B es homocigota para una mutación recesiva (m) que no altera la viabilidad ni la fertilidad. Los ratones de la generación F<sub>2</sub> serán, en promedio, 75% normales y 25% con fenotipo mutante. Estos últimos son los ratones que elegimos para el genotipado con marcadores. **B.** Tres pares de cromosomas [representando grande (100 cM), mediano (75 cM) y chico (50cM)] de cuatro progenies independientes de ratones F<sub>2</sub>. Notar que los ratones F<sub>2</sub> pueden recibir cromosomas recombinantes de ambos progenitores.

(C57BL/6xDBA/2)F<sub>2</sub>. Presentan dos ventajas sobre las retrocruzas: (i) cada cría F<sub>2</sub> que analizamos nos proporciona el doble de información recombinogénica (proveniente de las gametas paterna y materna), ya que ocurren meiosis informativas en ambos progenitores y (ii) pueden ser usadas para mutaciones recesivas deletéreas, ya que ambos padres (heterocigotas para la mutación) serán viables, y el ADN de las crías F<sub>2</sub> puede ser obtenido en cualquier momento del desarrollo (inclusive prenatal) para su tipificación. Entonces, el uso de las intercruzas para localizar mutaciones está recomendado en los siguientes casos: (i) si el alelo mutante es recesivo y los homocigotas (*mu/mu*) son infértiles y (ii) si el alelo mutante es co-dominante. Entre 50 y 100 animales F<sub>2</sub> es un número adecuado para comenzar un trabajo de localización de una mutación.

¿Cuál es la mejor elección entonces: intercruza o retrocruza?

Las ventajas de la intercruza son:

- (i) Se analizan dos meiosis por animal (por lo tanto se usa la mitad de animales).
- (ii) Permite estudiar los tres genotipos:  $mu/mu$ ,  $mu/+$  y  $+/+$  ( $mu$  = mutante,  $+$  = alelo salvaje).

Las desventajas de la intercrusa son:

- (i) En el caso de mutaciones recesivas, sólo un animal de cada cuatro (25%) será mutante (y por lo tanto directamente informativo para el análisis de ligamiento), los animales de fenotipo normal deberán ser genotipados para discernir entre  $mu/+$  y  $+/+$ .
- (ii) Exige tener animales F1 fértiles en los dos sexos (por lo tanto no sirve para las cruza inter-específicas).

Las ventajas de la retrocrusa son:

- (i) En el caso de las mutaciones recesivas, el 100% de los animales será informativo.
- (ii) Permite tener animales F1 fértiles en un sólo sexo (ideal para las cruza inter-específicas).

Las desventajas de la retrocrusa son:

- (i) Se analiza sólo una meiosis por animal.

### 6.4.3 Líneas Consanguíneas Recombinantes

Las **líneas consanguíneas recombinantes** (*Recombinant Inbred Strains*, RIS) son paneles (del inglés *set*) de líneas que se producen por 20 ó más generaciones de endocria sistemática de las crías sucesivas de una crusa entre dos animales F1, donde las parejas en la generación F2 son tomadas al azar (**Figura 6.13**). Hasta la actualidad se crearon 30 paneles de líneas RIS (con un total de 441 líneas) derivadas de 23 líneas consanguíneas parentales, pero sólo 13 grupos (162 líneas) se comercializan. Uno de los grupos más usado es el BXD, cuyas líneas parentales son C57BL/6 y DBA/2, junto con los grupos AXB y BXA, muy útiles debido a su alto nivel de polimorfismo. El *Jackson Laboratory* es el laboratorio que dispone de mayor número de líneas RIS, las mismas pueden consultarse en Internet: <http://jaxmice.jax.org/html/infosearch/pricelistframeset.html>. Al optar por estos animales hay que tener en cuenta dos características:

- (i) Siendo líneas consanguíneas, cada línea de un panel de ratones RIS representa una colección de individuos con genomas idénticos, homocigotas para todos sus genes, y que pueden ser criados en número ilimitado, lo que significa una gran ventaja. Estas lí-



Figura 6.13 A. Construcción de un panel de cepas consanguíneas recombinantes (RIS)

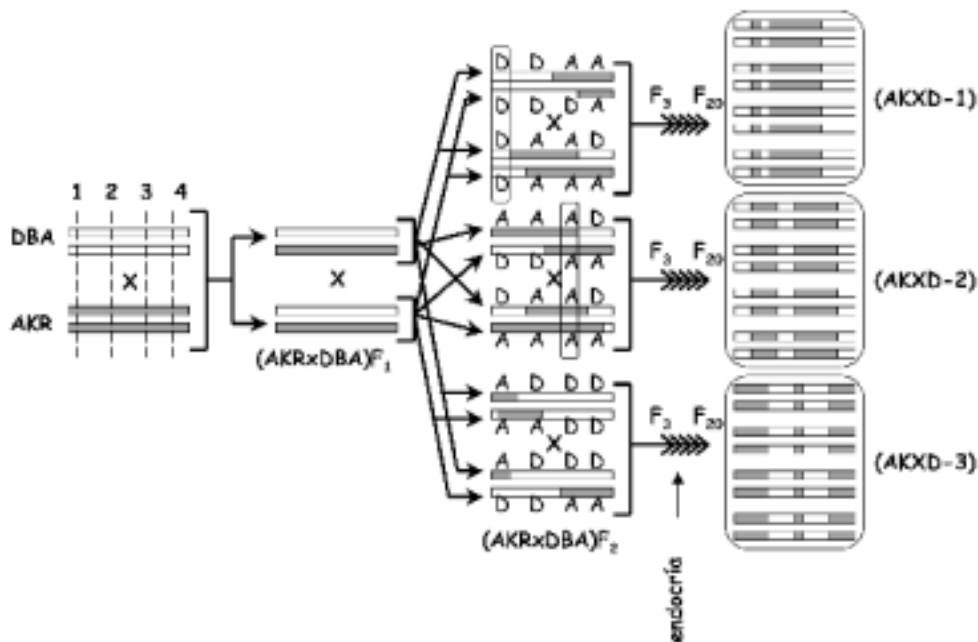


Figura 6.13 B. Paneles de RIS disponibles

Línea	Líneas parentales	Número de líneas disponibles		
		Total	ADN	Ratones vivos
AKXD	AKR/J, DBA/2J	30	24	21
AKXL	AKR/J, C57L/J	21	23	18
AXB, BXA	A/J, C57BL/6J	51	41	31
BXD	C57BL/6J, DBA/2J	26	27	26
BXH	C57BL/6J, C3H/HeJ	13	12	12
BXJ	C57BL/6J, SJL/J	2	2	2
CXB	BALB/cBy, C57BL/6By	13	13	13
CXJ	BALB/cKe, SJL/J	12	7	4
LXB	C57L/J, C57BL/6J	5	3	1
LXPL	C57L/J, PL/J	11	7	1
NX129	NZB/BINJ, 129J	11	8	0
NXSM	NZB/BINJ, SM/J	18	15	15
SWXJ	SWR/Bm, SJL/Bm	14	14	14
SWXL	SWR/J, C57L/J	7	7	7

Figura 6.13. Las líneas recombinantes consanguíneas. A. Esquema de la construcción de un grupo de líneas RIS usando AKR (AK) y DBA (D) como progenitoras. Con motivos didácticos, el diagrama muestra un único par de cromosomas a lo largo del proceso. La posición de 4 loci hipotéticos se indica en los cromosomas parentales. Los recuadros son alelos que ya han sido fijados en la generación F2. Luego de 20 generaciones de endocria obtenemos cepas consanguíneas (isogénicas) que portan en promedio 50% de cada progenitor (se muestran el mismo par de cromosomas para cuatro individuos en cada línea). B. La tabla muestra los paneles de líneas RIS disponibles en el Jackson Laboratory, Maine, Estados Unidos.

neas también permanecen estables a lo largo de las generaciones, con la sola excepción de la posibilidad de mutaciones.

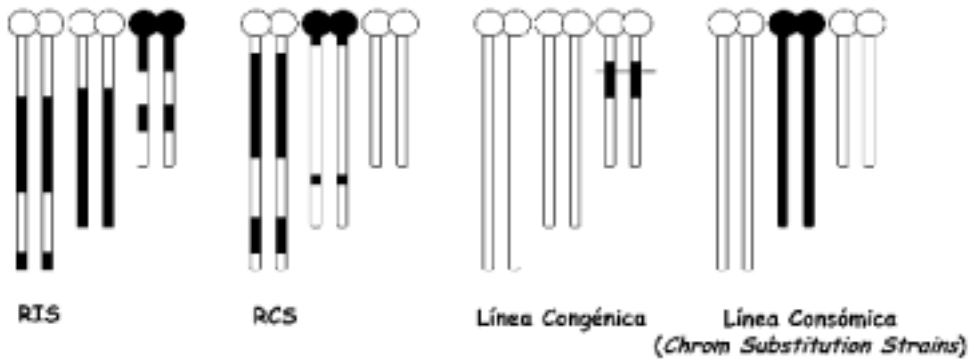
- (ii) Habiéndose originado de dos líneas progenitoras no relacionadas, heredan cada componente de su estructura genética de una u otra línea parental. Cuando un panel de RIS está siendo creado, pueden ocurrir eventos de recombinación meiótica en cada generación, en ambos miembros de la pareja reproductora. De esta manera, los cromosomas son “fragmentados” progresivamente en pequeñas regiones que se fijarán, al azar, en cada línea RIS. En otras palabras, a nivel genético, cada línea es como un “tapiz” hecho de “trozos” de cromosomas derivados, probablemente al azar, de las dos líneas parentales (50% y 50% en promedio). La distribución de los marcadores o genes en estas líneas se conoce como **patrón de distribución por cepa** (del inglés *strain distribution pattern*, **SDP**). Aquellos genes que se encuentren ligados en el mismo cromosoma tendrán una tendencia a permanecer juntos a través de las sucesivas generaciones, excepto cuando una recombinación anule esa asociación. La gran ventaja que presentan estas líneas para el análisis de ligamiento es la posibilidad de acumular información en forma aditiva a lo largo de los años sobre el mismo panel de ratones RIS. El programa MapManager también puede ser usado para identificar ligamientos en estas cruzas.

#### 6.4.4 Líneas Congénicas Recombinantes

Las **líneas congénicas recombinantes** (del inglés *recombinant congenic strains*, RCS) constituyen una suerte de variación de la filosofía de las RIS y se forman, en primer lugar, cruzando dos líneas consanguíneas. Luego, a partir de los animales F1, se realizan algunas (generalmente dos) generaciones de retrocruza contra una de las líneas parentales (la línea de fondo) para finalmente proseguir con la endocría -sin selección- de estos animales. De esta manera, la proporción de la línea de fondo y la donante no será, como en el caso de las RIS, 50% y 50%. Por ejemplo, para el caso de dos retrocruzas, será 87,5% de la línea de fondo y 12,5% de la línea donante. Las RCS, junto con las líneas consómicas (en las cuales se reemplaza un cromosoma completo), son especialmente útiles para el estudio de caracteres hereditarios complejos (QTL's), como son los genes de susceptibilidad al cáncer, entre otros (**Figura 6.14**). Una variante de las líneas congénicas recombinantes son las **líneas congénicas recombinantes inter-específicas** (*inter-specific recombinant congenic strains*, IRCS) donde C57BL/6 es la línea de fondo y SEG/Pas (*Mus spretus*) la línea donante (ver Capítulo IV).

#### 6.4.5 Cruzas inter-específicas (la contribución de los ratones salvajes)

Hablamos de **cruzas inter-específicas** cuando apareamos ratones de distintas especies (siempre dentro del mismo género), como por ejemplo *Mus musculus* y *Mus spretus*. En cambio, las **cruzas inter-subespecíficas** se logran cuando apareamos ratones de distinta subespecie (el



**Figura 6.14. Cromosomas recombinantes.** El esquema muestra la imagen hipotética de las recombinaciones en tres pares de cromosomas (representando grande, mediano y chico) para las líneas RIS, RCS, congénicas y consómicas. Para las RIS el promedio de cada línea parental es de 50%. Para las líneas RCS (empezando la endocria en la segunda generación de retrocruza), la línea donante (color negro) contribuye con un 12,5%. Para las líneas congénicas, el segmento seleccionado es en promedio de 20 cM en la generación N10. Las líneas consómicas poseen un cromosoma entero que fue transferido a la línea receptora.

tercer componente del trinomio en latín), como por ejemplo *Mus musculus domesticus* y *Mus musculus castaneus* (para detalles sobre la sistemática del ratón remitirse al Capítulo III).

#### 6.4.5.1 Los ratones salvajes como fuente de polimorfismo genético

La relativa homogeneidad genética de las líneas clásicas de laboratorio ayudó a los primeros genetistas simplificando, en cierta medida, el análisis del determinismo genético de algunos genotipos complejos, como aquellos implicados en la histocompatibilidad o la predisposición a ciertos tipos de cáncer. Por otro lado, esto se convirtió en un gran inconveniente a la hora de realizar mapas genéticos de alta densidad (indispensables por ejemplo para el clonaje posicional de genes), estudiar la **impronta genética** (del inglés *genomic imprinting*), los efectos de la **epistasia** o la predisposición a ciertas enfermedades infecciosas. Haciendo una analogía con la pintura podríamos decir que al trabajar con las líneas clásicas de laboratorio los genetistas están usando una paleta a la cual le faltan varios colores fundamentales. Es en parte por esta razón que fueron creadas nuevas líneas consanguíneas a partir de ratones capturados en la naturaleza, en estado salvaje. Estas nuevas líneas han introducido en los laboratorios un polimorfismo considerable a partir del cual fueron posibles, por ejemplo, los estudios de genética cuantitativa (QTL's), entre otros.

La desventaja de las líneas clásicas de poseer un conjunto (del inglés *pool*) genético restringido, con una variación alélica relativamente baja comparada con el hombre, fue entonces superada cuando los genetistas del ratón decidieron tomar ventaja de la diversidad existente en los especímenes salvajes del género *Mus*. Esto sirvió para desarrollar nuevas líneas consanguíneas que permitieron cruza muy polimórficas entre líneas de origen salvaje y líneas clásicas de laboratorio (cruzas inter-específicas e inter-subespecíficas). Salvo algunas excepciones, el

establecimiento de líneas consanguíneas a partir de ratones salvajes del complejo *Mus musculus* fue siempre exitoso. En cambio, la creación de nuevas líneas consanguíneas a partir de otras especies del género *Mus*, tales como *Mus spretus* o *Mus spicilegus*, fue más difícil, aunque se logró en varias ocasiones.

*Mus spretus* es la especie más distante, dentro del género *Mus*, que puede ser cruzada con las líneas de laboratorio y producir híbridos fértiles (solamente las hembras, debido al denominado efecto Haldane). La cruce con esta especie se ha convertido, en los últimos años, en la cruce inter-específica de elección para la realización de mapas genéticos de alta densidad/alta resolución. Esta estrategia de mapeo de genes ha demostrado ser confiable y probablemente no introduce ningún desvío en la localización de los genes en el genoma del ratón.

Haciendo un análisis del polimorfismo en estas nuevas líneas consanguíneas se ha encontrado que el 81% de los clones de ADN, tomados al azar, entre *Mus spretus* y C57BL/6 muestran un polimorfismo de restricción (RFLP), utilizando sólo cuatro enzimas. Al evaluar el polimorfismo por PCR (para un panel de marcadores microsatélites) se observó un nivel de 77% entre C57BL/6 y *Mus m. castaneus* y un 90% entre C57BL/6 y *Mus spretus*. Estos hallazgos fueron confirmados y expandidos recientemente por el grupo dirigido por Carlos Penha-Gonçalves usando 10 líneas consanguíneas derivadas de ratones salvajes y el análisis de polimorfismo de 254 microsatélites. Las líneas salvajes estudiadas fueron *M. m. domesticus* (WMP, WLA, BIK/g y 38CH), *M. m. musculus* (MBT, PWK y MAI), *Mus spretus* (STF y SEG) y *Mus spicilegus* (ZYD). Es interesante notar que estos investigadores hallaron 632 alelos derivados de las líneas salvajes que no se encontraban presentes en las líneas clásicas de laboratorio y que, en promedio, el nivel de polimorfismo de estas líneas, comparadas con cualquier línea clásica, era de 80%. Es necesario hacer notar que algunos iniciadores de PCR, al haber sido diseñados para las líneas clásicas de laboratorio, pueden fallar en la amplificación del segmento de ADN "salvaje" homólogo, resultando en la ausencia del producto de PCR proveniente de *Mus spretus* (este resultado puede conducir a una interpretación errónea, por ejemplo clasificando a un heterocigota como homocigota).

El uso de las nuevas técnicas para detectar polimorfismo ha hecho que cobren más valor las cruces inter-subespecíficas (con las líneas consanguíneas derivadas de *Mus m. musculus* o *Mus m. castaneus*), menos distantes de las líneas de laboratorio que *Mus spretus*, pero mucho más fáciles de reproducir. Además, como ya se vio, el polimorfismo de los marcadores microsatélites entre las líneas clásicas y cualquier línea derivada del grupo *Mus musculus* tiene un nivel aceptable para cualquier trabajo de cartografía genética. Este mismo tipo de comparaciones fue realizado entre líneas consanguíneas de rata y ratas capturadas en la naturaleza, donde alrededor del 80% de los microsatélites estudiados fueron polimórficos. Este alto nivel de polimorfismo debería ser aprovechado para generar líneas consanguíneas derivadas de ratas salvajes; lo que aportaría nuevos alelos de microsatélites para la construcción de mapas genéticos, como lo hemos visto para el ratón.

Además de su gran valor como fuente de polimorfismo, los ratones salvajes representan también una fuente valiosa de variaciones morfológicas a nivel del cariotipo. En las líneas de labo-

ratorio existe una gran variedad de translocaciones recíprocas e inversiones que son el resultado de todos los ensayos de radiación realizados a lo largo de los años. En cambio, las translocaciones Robertsonianas o las fusiones céntricas, por ejemplo, son poco comunes en estas líneas de laboratorio. Sin embargo, este tipo de variaciones cromosómicas son muy comunes en las poblaciones naturales de la especie *Mus musculus domesticus*. Se han hallado en todos los cromosomas acrocéntricos del cariotipo normal (con la excepción del cromosoma Y) diversos tipos de translocaciones Robertsonianas y es muy probable que existan muchas más por descubrir. Estos rearrreglos espontáneos (de ocurrencia natural) han sido muy usados para generar una gran variedad de monosomías y trisomías y también para la demostración del fenómeno de impronta genética. Son además herramientas útiles para la localización de genes por hibridación *in situ* ya que ayudan en la discriminación entre pares cromosómicos, lo que suele ser difícil usando el complemento cromosómico normal del ratón.

#### 6.4.5.2 El valor de los ratones salvajes para el clonaje posicional de genes

Por todo lo expresado, queda claro que las líneas consanguíneas salvajes son una herramienta fundamental para el **clonaje posicional** (del inglés *positional cloning*) de genes. Como veremos más adelante el clonaje de un gen que es conocido sólo por su fenotipo deletéreo implica que el gen en cuestión sea aislado en un vector con la menor cantidad de ADN flanqueante posible, para evitar trabajo innecesario en los pasos finales. Para cumplir este objetivo se requiere de la realización de un mapa de alta densidad de marcadores —de las regiones que flanquean al gen— y la identificación de los dos marcadores más cercanos al gen. La situación ideal es cuando el locus mutado puede ser flanqueado por dos marcadores que se encuentren a una distancia menor de 0,3 cM, lo que representaría, en promedio, entre 500 y 600 kb de ADN genómico. Lamentablemente esto es muy difícil de lograr en la práctica con las líneas clásicas, debido a su bajo nivel de polimorfismo, y por eso es muy apropiado pensar en diseñar cruza con líneas consanguíneas de origen salvaje para el clonaje posicional de genes en el ratón.

No parece arriesgado asegurar que, en el futuro, el uso de estos ratones salvajes va a expandirse hacia otras áreas, mas allá del mapeo de genes y el estudio de la impronta genética. Seguramente se convertirá en una herramienta útil en el estudio de las interacciones epistáticas y como una fuente de “genes a clonar”, así como lo es la hierba *Arabidopsis thaliana* para los que trabajan en genética de plantas. Teniendo en cuenta que las líneas de ratones de laboratorio presentan un espectro de reacciones muy pobre cuando son infectados con organismos patógenos o son expuestos a una droga, es muy probable que estudiando la reacción de los ratones salvajes —de diferentes especies— se llegue a la identificación de genes nuevos potencialmente interesantes.

## 6.5 Aplicaciones de los mapas genéticos

Veremos muy brevemente las aplicaciones más importantes de los mapas que hemos delineado en este capítulo:

### 6.5.1 Identificación de Genotipos Específicos

Criar ratones mutantes para investigación es muchas veces complicado por el hecho de que los genotipos homocigotas, los cuales son buscados para muchas experiencias, generalmente mueren *in utero*, en el período perinatal o el defecto les impide reproducirse. Esto último es lo que le ocurre, entre otras, a las mutaciones *obese (ob)* y *dwarf (dw)*. Es aquí donde la información de un mapa genético puede ser de gran utilidad. Por ejemplo, en la mutación autosómica recesiva *progressive motor neuronopathy* (ahora *Tbce<sup>pmn</sup>*) los ratones homocigotas pueden ser identificados entre los días 14 y 20 de nacidos (cuando comienzan a desarrollar una parálisis de los miembros anteriores), pero estos animales mueren rápidamente, entre los días 50 y 55 de vida. Por trabajos de mapeo, se localizó esta mutación muy cerca ( $< 0,2$  cM) del locus *Xt* (*Extra toes*, ahora *Gli3<sup>Xt</sup>*) en el cromosoma 13. Esta localización ha permitido aumentar la eficiencia de los programas de cría de mutantes *pmn/pm*n utilizando al dedo extra de la mutación *Xt* como marcador fenotípico de fácil identificación. Esto se logró estableciendo un grupo de ratones con la mutación *Xt* en repulsión, es decir sobre el cromosoma que no porta el alelo mutante *pmn*. Haciendo uso de esta ayuda, se intercrucan animales con genotipos *Xt +/+ pmn* (dedo suplementario): los animales *Xt +/Xt +*, que no portan más el alelo *pmn*, mueren *in utero*, mientras que los homocigotas *+ pmn/+ pmn* (dedos normales) pueden ser reconocidos desde recién nacidos, antes de que enfermen (ver Anexo I).

### 6.5.2 Estudios de Evolución del Genoma

Existen múltiples ejemplos para ilustrar la importancia de los mapas genéticos en los estudios de evolución del genoma, entre ellos citaremos sólo dos. Dentro del genoma del ratón existen muchas secuencias que se encuentran repetidas un número determinado de veces. Algunas de ellas son genes funcionales y sus diferentes copias pueden estar en el mismo cromosoma o en cromosomas distintos. Algunas, inclusive, presentan una expresión diferencial según los tejidos. Este es el caso de la duplicación de un gen ancestral que codifica para la amilasa en el cromosoma 3 del ratón. Como resultado de esa duplicación existen dos copias ligadas del gen, *Amy1* y *Amy2*, siendo el primer gen activo en las glándulas salivales y el segundo en el páncreas. Esta distribución responde a cambios producidos durante la evolución del genoma en diferentes momentos de la historia de la especie. Estos mecanismos aún no son bien conocidos pero, con el desarrollo de buenos mapas genéticos, se podrá comprender mejor lo ocurrido durante la evolución de estas secuencias.

El otro ejemplo lo constituyen los estudios realizados sobre el origen del Complejo Mayor de Histocompatibilidad (ver capítulo V). Los enigmas sobre su origen parecen haber sido resueltos recientemente con el descubrimiento de regiones que evolucionaron paralelamente -por duplicación- de un gen ancestral común. Es decir que estas regiones emergieron como resultado de una duplicación cromosómica producida, aparentemente, en un ancestro común a todos los vertebrados con mandíbula, antes de su divergencia de los peces sin mandíbulas (vertebrados primitivos como la lamprea). Estas observaciones señalan la importancia evoluti-

va que tienen los fenómenos de duplicación cromosómica en la creación de sistemas cada vez más complejos en los organismos vivos.

### 6.5.3 Establecimiento de Homologías Cromosómicas entre Especies

De los 18.000 genes mapeados actualmente en el ratón, se conocen alrededor de 3.500 homólogos con genes humanos (incluyendo cerca de 400 loci homólogos con enfermedades genéticas humanas), de los cuales alrededor de 2.500 están mapeados en ambas especies en distintos segmentos homólogos. (Además, la información vertida recientemente por las comparaciones de los genomas completos del ratón y el hombre –con un ancestro común que se remonta a 75 millones de años– estaría indicando que aproximadamente el 99% de los genes del ratón tiene un homólogo en el genoma humano). Cada grupo de genes homólogos constituye un segmento cromosómico conservado, o sea que se trata de grupos de genes que muestran el mismo orden lineal en ambas especies, lo que se denomina conservación de la **sintenia** (del griego “sobre el mismo hilo”), lo que hace referencia a los genes que se encuentran en el mismo cromosoma. Algunos autores diferencian entre “**segmentos conservados**” (también “ligadura conservada” o, en inglés, *syntenic segment*), aquellas regiones donde el orden de varios genes es el mismo en ambas especies, y “**sintenia conservada**”, donde se conserva la localización cromosómica de los genes, pero no necesariamente el mismo orden (también llamados –en inglés– *syntenic blocks*).

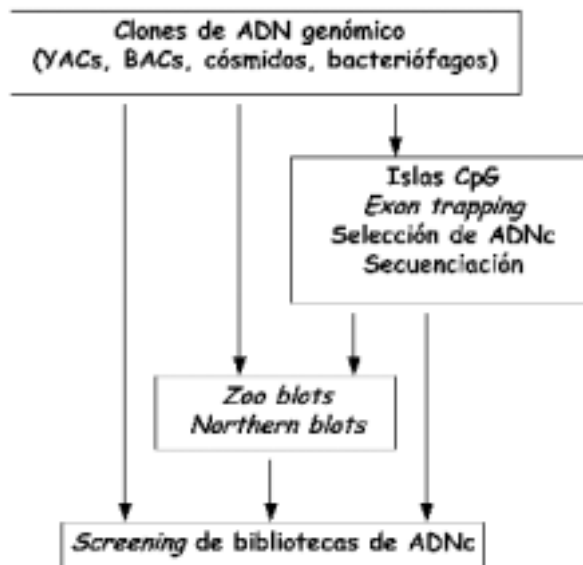
Hasta la fecha de este escrito (incluyendo la información proveniente del primer borrador del genoma del ratón), se identificaron 342 segmentos conservados entre los genoma del ratón y el humano, con tamaños que van desde 300 kb hasta 65 Mb (un promedio de 7 Mb). Alrededor del 90% del genoma humano y 93% del genoma del ratón reside en estos segmentos conservados. Otros hallazgos recientes son que el cromosoma X está representado por un único bloque sinténico en ambas especies y que el cromosoma 20 humano corresponde enteramente a una porción del cromosoma 2 del ratón, con una conservación del orden de genes casi perfecta.

Estos mapas comparativos pueden usarse para predecir ligamientos entre genes y para identificar “genes candidato” para enfermedades genéticas en ambas especies, como sucedió con los genes *Pax6*, *Kit* y *Lyst* (mutación *beige*) cuyas enfermedades homólogas humanas son la aniridia, el piebaldismo y el síndrome de Chediak-Higashi, respectivamente. Por ejemplo, el mapeo de un gen murino (*aku*) homólogo de la enfermedad metabólica humana alkaptonuria, en el cromosoma 16 del ratón, sirvió de base para la localización del gen humano en la región homóloga correspondiente del cromosoma 3. Otro ejemplo es el desarrollo de un modelo murino para el síndrome de Down: se han identificado por lo menos 25 genes conservados entre la región de homología del cromosoma 16 del ratón y el brazo largo del cromosoma 21 humano (recordar que la imagen citogenética del síndrome es la trisomía del cromosoma 21). Esto demuestra la importancia que puede tener este tipo de estudios comparativos en la disección genética de las enfermedades complejas, como la diabetes, la hiper-

tensión y el cáncer. Finalmente, el conocimiento preciso del orden de los genes dentro de estos segmentos homólogos puede incluso ser usado para transferir información desde el ratón y el hombre a otras especies de animales de laboratorio y domésticos.

#### 6.5.4 Clonaje Posicional de Mutaciones

El proceso de clonaje de un gen que es conocido sólo por su fenotipo, pero no por la proteína para la cual codifica, se conoce como clonaje posicional. Esta estrategia, antes llamada “genética reversa”, consiste en que el gen en cuestión sea aislado en un vector con la menor cantidad posible de ADN flanqueante. Esto requiere, en primer lugar, el establecimiento de un mapa de alta densidad en las regiones flanqueantes del gen, junto con el hallazgo de los dos marcadores genéticos más cercanos, para evitar trabajo innecesario en los pasos finales de la identificación del gen. El clonaje posicional implica un “caminar” a través de esos marcadores moleculares cercanos a la mutación utilizando clones de ADN (YAC's, BAC's o fagos PI). Una vez que se identifica un intervalo que (con gran certitud) posee el gen de interés, se procede a identificarlo por medio de varias técnicas, como ser la **trampa de exones** (del inglés *exon trapping*), la **selección de ADNc** o la **secuenciación**, entre las más utilizadas (Figura 6.15). La estrategia de clonaje posicional ha sido utilizada en medicina humana para aislar los genes responsables de varias enfermedades genéticas como la fibrosis quística y la enfermedad de Huntington.



**Figura 6.15. Detectando genes.** Identificación de genes a partir de ADN clonado. El diagrama muestra las técnicas más utilizadas para detectar genes en una secuencia de ADN clonado. Esta situación se presenta en las etapas finales del clonado posicional de una mutación, donde no se ha tenido éxito con la búsqueda de genes candidato en la región.



Otras mutaciones han sido identificadas en términos moleculares por el método del **gen candidato**: si un gen, que es conocido bajo la forma de un alelo mutante, es ubicado en el mapa genético y, al mismo tiempo, un clon de ADNc es posicionado en la misma región del genoma se analiza la posible conexión entre ellos. Esto es lo que ocurrió con el gen *Pax1*, que fue localizado en el cromosoma 2 del ratón en una región donde había sido mapeada una vieja mutación llamada *undulated* (*un*). De esta manera, *un* resultó ser una mutación del gen *Pax1*. De la misma forma, *balding* (*bal*) resultó ser una mutación del gen *desmoglein 3* (*Dsg3*) y *nackt* (*nkt*) del gen catepsina L (*Ctsl*), entre muchos ejemplos más. Hasta la fecha, casi tres cuartos de todos los genes clonados en el ratón han sido identificados usando este enfoque del gen candidato.

Existen muchos genes en el ratón que son de interés por ser modelos de enfermedades genéticas humanas o por alterar alguna función esencial del desarrollo embrionario, muchos de ellos se encuentran actualmente en proceso de clonaje. Hay que tener en cuenta que en el ratón existen más de 1.000 mutaciones espontáneas que causan enfermedades genéticas, lo cual constituye una fuente de nuevos modelos animales interesantes para la biomedicina (ver Capítulo IX) y que pueden, a su vez, aportar ideas sobre la función de los genes. La identificación del gen *obese* (*Lep<sup>ob</sup>*) responsable de fenómenos de obesidad en el ratón y el hombre es un ejemplo reciente de este tipo de trabajos. Por otro lado, muchos genes responsables de caracteres de variación cuantitativa (QTL's) han sido mapeados pero, como hemos visto, muy pocos han sido clonados.

Con la disponibilidad de varias genotecas de BAC's y de las secuencias completas del genoma del ratón y la rata, es posible actualmente llevar a cabo el clonaje posicional de cualquier mutación. El rápido aumento en el número de EST's y ADNc humanos y murinos más el refinamiento y el aumento de la densidad de ambos mapas genéticos a mejorado mucho las posibilidades de realizar clonaje posicional. Además, el uso de información cruzada entre estas dos especies aumenta las posibilidades de sugerir, descartar o confirmar genes candidato para la mutación.

## Bibliografía General

- AHMADIAN A, LUNDEBERG J. *A brief history of genetic variation analysis*. Biotechniques 32: 1122-1136, 2002.
- AITMAN T, HEARNE C, TODD J. *Mononucleotide repeats are an abundant source of length variants in mouse genomic DNA*. Mammalian Genome 1: 206-210, 1991.
- APARICIO S, CHAPMAN J, STUPKA E, PUTNAM N, CHIA JM, DEHAL P, CHRISTOFFELS A, et al. *Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes**. Science 297: 1301-1310, 2002.
- AVNER P, AMAR L, DANDOLO L, GUÉNET J-L. *Genetic analysis of the mouse using interspecific crosses*. Trends in Genetics 4: 18-23, 1988.
- AVNER P, BRULS T, PORAS I, ELEY L, GAS S, RUIZ P, WILES MV, SOUSA-NUNES R, KETTLEBOROUGH R, RANA A, MORISSETTE J, BENTLEY L, GOLDSWORTHY M, HAYNES A, HERBERT E, SOUTHAM L, LEHRACH H, WEISENBACH J, MANENTI G, RODRIGUEZ-TOME P, BEDDINGTON R, DUNWOODIE S, COX RD. *A radiation hybrid transcript map of the mouse genome*. Nature Genetics 29: 194-200, 2001.

- BAILEY D. Recombinant inbred strains. An aid to finding identity, linkage and function of histocompatibility and other genes. *Transplantation* 11: 325-327, 1971.
- BAXEVANIS AD. *The Molecular Biology Database Collection: 2003 update*. *Nucleic Acids Research* 31: 1-12, 2003.
- BEDDELL M, JENKINS N, COPELAND N. Mouse models of human disease. Part I: Techniques and resources for genetic analysis in mice. *Genes and Development* 11: 1-10, 1997.
- BEIER D. Single-strand conformation polymorphism (SSCP) analysis as a tool for genetic mapping. *Mammalian Genome* 4: 627-631, 1993.
- BENAVIDES F, GUÉNET J-L. Mapeo de genes en el ratón. *Medicina (Buenos Aires)* 57: 507-510, 1997.
- BIRREN B, GREEN ED, HIETER P, KLAPHOLZ S, MYERS R, RETHMAN H, ROSKAMS J. (Eds.) *Mapping Genomes*. En: *Genome Analysis: A Laboratory Manual Series 4*, Cold Spring Harbor Laboratory Press, New York, 1999.
- BONHOMME F, MARTIN S, THALER L. Hybridization between *Mus musculus domesticus* and *Mus spretus* under laboratory conditions. *Experientia* 34: 1140-1141, 1978.
- BOTSTEIN D, WHITE R, SKOLNICK M, DAVIS R. Constitution of genetic linkage map in man using RFLP's. *American Journal of Human Genetics* 32: 314-331, 1980.
- BRADLEY A. Mining the mouse genome. *Nature* 420: 512-514, 2002.
- BREEN M, DEAKIN L, MACDONALD B, MILLER S, SIBSON R, TARTTELIN E, AVNER P., et al. Towards high resolution maps of the mouse and human genomes-A facility for ordering markers to 0.1 cM resolution. *Human Molecular Genetics* 3: 621-627, 1994.
- BROWN SD. Whither the mouse genome? *Mammalian Genome* 13: 481-482, 2002.
- BROMAN KW. Review of statistical methods for QTL mapping in experimental crosses. *Lab Animal (NY)* 30: 44-52, 2001.
- BROMAN KW, ROWE LB, CHURCHILL GA, PAIGEN K. Crossover interference in the mouse. *Genetics* 160: 1123-1131, 2002.
- CAMPINO S, BEHRSCHEMIDT C, BAGOT S, GUENET JL, CAZENAVE PA, HOLMBERG D, PENHA GONCALVES C. Unique genetic variation revealed by a microsatellite polymorphism survey in ten wild-derived inbred strains. *Genomics* 79: 618-620, 2002.
- COMPARATIVE GENOME ORGANIZATION OF VERTEBRATES. *The First International Workshop on Comparative Genome Organization*. *Mammalian Genome* 7: 717-734, 1996.
- COPELAND N, JENKINS N, GILBERT D, EPPIG J, MALTAIS L, MILLER J, DIETRICH W, WEAVER A, LINCOLN S, STEEN R, STEIN L, NADEAU J, LANDER E. A genetic linkage map of the mouse: current applications and future prospects. *Science* 262: 57-66, 1993.
- DAVISSON M, BRADT D, MERRIAM J, ROCKWOOD S, EPPIG J. The mouse gene map. *ILAR Journal* 39: 96-131, 1998.
- DEMANT P, HART A. Recombinant congenic strains-a new tool for analyzing genetic traits determined by more than one gene. *Immunogenetics* 24: 416-422, 1986.
- DIETRICH W, MILLER J, STEEN R, LINCOLN S, STEIN L, PAGE D, LANDER E. A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature Genetics* 7: 220-245, 1994.
- DIETRICH W, MILLER J, STEEN R, MERCHANT M, DAMRON-BOLES D, HUSAIN Z, DREDGE R, et al. A comprehensive genetic map of the mouse genome. *Nature* 380: 149-152, 1996.
- DOOLITTLE D, DAVISSON M, GUIDI J, GREEN M. Catalog of mutant genes and polymorphic loci. En: "Genetic Variants and Strains of the Laboratory Mouse" (M.F. Lyon, S. Rastan, y S. Brown, Editores), pp. 17-854. Oxford University Press, Oxford, 1996.
- FLICEK P, KEIBLER E, HU P, KORF I, BRENT MR. Leveraging the mouse genome for gene prediction in human: from whole-genome shotgun reads to a global synteny map. *Genome Research* 13: 46-54, 2003.
- GLAZIER AM, NADEAU JH, AITMAN TJ. Finding genes that underlie complex traits. *Science* 298: 2345-2349, 2002.
- GOFFEAU A, BARRELL B, BUSSEY H, DAVIS R, DUJON B, FELDMANN H, GALBERT F, et al. Life with 6,000 genes. *Science* 274: 546-567, 1996.
- GOLDSTEIN DV, SCHLOTTERER C (Eds). *Microsatellites: Evolution and Applications*. Oxford University Press, Oxford, 2000.
- GRAVES JA. Background and overview of comparative genomics. *ILAR Journal* 39: 48-65, 1998.
- GREGORY SG, SEKHON M, SCHEIN J, ZHAO S, OSOEGAWA K, SCOTT CE, EVANS RS, BURRIDGE PW, COX TV, FOX CA, HUTTON RD, MULLENGER IR, PHILLIPS KJ, SMITH J, et al. A physical map of the mouse genome. *Nature* 418: 743-750, 2002.

- GUÉNET J-L, SIMON D, AVNER P. *The use of interspecific mouse crosses for gene localization: present status and future perspectives*. En: "Current Topics in Microbiology and Immunology 137" Mock y Potter, Editores, pp 13-17, 1988.
- GUÉNET J-L. *The Mouse Genome*. En: "Genomes, Molecular Biology and Drug Discovery", pp. 27-51. Academic Press Ltd., New York, 1996.
- GUÉNET J-L, BONHOMME F. *Wild mice: an ever-increasing contribution to a popular mammalian model*. Trends in Genetics 19: 24-31, 2003.
- HEARNE C, GHOSH S, TODD J. *Microsatellites for linkage analysis of genetic traits*. Trends in Genetics 8: 288-294, 1992.
- HELOU K, VALENTINSSON A, LEVAN G, STAHL F. *Between rat and mouse zoo-FISH reveals 49 chromosomal segments that have been conserved in evolution*. Mammalian Genome 12: 765-771, 2001.
- HOLT RA, SUBRAMANIAN GM, HALPERN A, SUTTON GG, CHARLAB R, NUSSKERN et al. *The genome sequence of the malaria mosquito Anopheles gambiae*. Science 298: 129-149, 2002.
- HUDSON TJ, CHURCH DM, GREENAWAY S, NGUYEN H, COOK A, STEEN RG, VAN ETTEN WJ, CASTLE AB, STRIVENS MA, TRICKETT P, HEUSTON C, DAVISON C, SOUTHWELL A, HARDISTY R, VARELA-CARVER A, HAYNES AR, RODRIGUEZ-TOME P, DOI H, KO MS, PONTIUS J, SCHRIML L, WAGNER L, MAGLOTT D, BROWN SD, LANDER ES, SCHULER G, DENNY P. *A radiation hybrid map of mouse genes*. Nature Genetics 29: 201-205, 2001.
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. *Initial sequencing and analysis of the human genome*. Nature 409: 860-921, 2001.
- JACKSON I, ABBOTT C (Eds). *Mouse Genetics and Transgenics. A Practical Approach*. Oxford University Press, New York, 2000.
- JACOB HJ, KWITEK AE. *Rat genetics: attaching physiology and pharmacology to the genome*. Nature Reviews Genetics 3: 33-42, 2002.
- JAFFE DB, BUTLER J, GNERRE S, MAUCELI E, LINDBLAD-TOH K, MESIROV JP, ZODY MC, LANDER ES. *Whole-genome sequence assembly for mammalian genomes: Arachne 2*. Genome Research 13: 91-96, 2003.
- JEFFREYS A, WILSON V, KELLY R, TAYLOR B, GRAHAME B. *Mouse DNA fingerprints: analysis of chromosome localization and germ-line stability of hypervariable loci in recombinant inbred strains*. Nucleic Acids Research 15: 2823-2837, 1987.
- JULIER C, GOUYON B DE, GEORGES M, GUÉNET J-L, NAKAMURA Y, ARNER P, LATHROP M. *Minisatellite linkage maps in the mouse by cross-hybridization with human probes containing tandem repeats*. Proceedings of the National Academy of Science USA 87: 4585-4589, 1990.
- KAWAI J, SHINAGAWA A, SHIBATA K, YOSHINO M, ITOH M, ISHII Y, ARAKAWA T, HARA A, FUKUNISHI Y, KONNO H, ADACHI J, FUKUDA S, AIZAWA K, IZAWA M, NISHI K, et al. *Functional annotation of a full-length mouse cDNA collection*. Nature 409: 685-690, 2001.
- KIKKAWA Y, MIURA I, TAKAHAMA S, WAKANA S, YAMAZAKI Y, MORIWAKI K, SHIROISHI T, YONEKAWA H. *Microsatellite database for MSM/IMs and JF1/IMs, molossinus-derived inbred strains*. Mammalian Genome 2: 750-752, 2001.
- KORSTANJE R, PAIGEN B. *From QTL to gene: the harvest begins*. Nature Genetics 31: 235-236, 2002.
- LENNON G, AUFRAY C, POLYMEROPOULOS M, SOARES M. *The I.M.A.G.E. Consortium: An integrated molecular analysis of genomes and their expression*. Genomics 33: 151-152, 1996.
- LINDBLAD-TOH K, WINCHESTER E, DALY MJ, WANG DG, HIRSCHHORN JN, LAVIOLETTE JP, ARDLIE K, REICH DE, ROBINSON E, SKLAR P, SHAH N, THOMAS D, FAN JB, GINGERAS T, WARRINGTON J, PATIL N, HUDSON TJ, LANDER ES. *Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse*. Nature Genetics 24: 381-386, 2000.
- LINDBLAD-TOH K, LANDER ES, MCPHERSON JD, WATERSTON RH, RODGERS J, BIRNEY E. *Progress in sequencing the mouse genome*. Genesis 31: 137-141, 2001.
- LOVE J, KNIGHT A, MC ALEER M, TODD J. *Towards construction of a high resolution map of the mouse genome using PCR-analyzed microsatellites*. Nucleic Acids Research 18: 4123-4130, 1990.
- LYON L, LAUGHLIN T, COPELAND N, JENKINS N, WOMACK J, O'BRIEN S. *Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes*. Nature Genetics 15: 47-56, 1997.
- MANLY KF, OLSON JM. *Overview of QTL mapping software and introduction to map manager QT*. Mammalian Genome 10: 327-334, 1999.
- MANLY KF, CUDMORE RH JR, MEER JM. *Map Manager QTX, cross-platform software for genetic mapping*. Mammalian Genome 12: 930-932, 2001.

- MCCARTHY L, TERRETT J, DAVIS M, KNIGHTS C, SMITH A, CRITCHER R, et al. A first-generation whole genome-radiation hybrid map spanning the mouse genome. *Genome Research* 7: 1153-1161, 1997.
- MONTAGUTELLI X, SERIKAWA T, GUÉNET J-L. PCR analyzed microsatellites: data concerning laboratory and wild-derived mouse inbred strains. *Mammalian Genome* 1: 255-259, 1991.
- NADEAU J, DAVISSON M, DOOLITTLE D, GRANT P, HILLYARD A, KOSOWSKY M, RODERICK H. Comparative map for mice and humans. *Mammalian Genome* 3: 480-536, 1992.
- NADEAU JH, FRANKEL WN. The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nature Genetics* 25: 381-384, 2000.
- NADEAU JH, BALLING R, BARSH G, BEIER D, BROWN SD, BUCAN M, CAMPER S, CARLSON G, COPELAND N, EPPIG J, FLETCHER C, FRANKEL WN, GANTEN D, GOLDDOWITZ D, GOODNOW C, GUÉNET J-L et al. Sequence interpretation. Functional annotation of mouse genome sequences. *Science* 291: 1251-1255, 2001.
- NILSSON S, HELOU K, WALENTINSSON A, SZPIRER C, NERMAN O, STAHL F. Rat-mouse and rat-human comparative maps based on gene homology and high-resolution zoo-FISH. *Genomics* 74: 287-298, 2001.
- OKAZAKI Y, FURUNO M, KASUKAWA T, ADACHI J, BONO H, KONDO S, NIKAIKIDO I, OSATO N, SAITO R, SUZUKI H, YAMANAKA I, KIYOSAWA H, YAGI K, TOMARU Y, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420: 563-573, 2002.
- PARK YG, CLIFFORD R, BUETOW KH, HUNTER KW. Multiple cross and inbred strain haplotype mapping of complex-trait candidate genes. *Genome Research* 13: 118-121, 2003.
- PARRA G, AGARWAL P, ABRIL JF, WIEHE T, FICKETT JW, GUIGO R. Comparative gene prediction in human and mouse. *Genome Research* 13: 108-117, 2003.
- PEROU C, MOORE K, NAGLE D, MISUMI E, WOOLF E, MCGRAIL S, et al. Identification of the murine beige gene by YAC complementation and positional cloning. *Nature Genetics* 13: 303-308, 1996.
- PEVZNER P, TESLER G. Genome rearrangements in mammalian evolution: lessons from human and mouse genomes. *Genome Research* 13: 37-45, 2003.
- RHODES M, STRAW R, FERNANDO S, EVANS A, LACEY T, DEARLOVE A, GREYSTRONG J, WALKER J, et al. A high-resolution microsatellite map of the mouse genome. *Genome Research* 8: 531-542, 1998.
- ROWE LB, BARTER ME, KELMENSEN JA, EPPIG JT. The comprehensive mouse radiation hybrid map densely cross-referenced to the recombination map: a tool to support the sequence assemblies. *Genome Research* 13: 122-133, 2003.
- SCHULER G, BOGUSKI M, STEWART E, STEIN L, GYAPAY G, RICE K, WHITE R. A gene map of the human genome. *Science* 274: 540-546, 1996.
- SERIKAWA T, MONTAGUTELLI X, SIMON D, GUÉNET J-L. Polymorphisms generated by PCR with short sized arbitrary primers can be used as reliable markers for mouse and rat gene mapping. *Mammalian Genome* 3: 65-72, 1992.
- SILVER L M. (ed). *Mouse Genetics. Concepts and applications*. Oxford University Press, Oxford, 1995. (Versión online disponible en <http://www.informatics.jax.org/silver/>)
- STASSEN A, GROOT P, EPPIG J, DEMANT P. Genetic composition of the recombinant congenic strains. *Mammalian Genome* 7: 55-58, 1996.
- STEEN RG, KWITEK-BLACK AE, GLENN C, GULLINGS-HANDLEY J, VAN ETTEN W, ATKINSON OS, APPEL D, TWIGGER S, MUIR M, MULL T, GRANADOS M, KISSEBAH M, RUSSO K, CRANE R, POPP M, PEDEN M, MATISE T, BROWN DM, LU J, KINGSMORE S, TONELLATO PJ, ROZEN S, SLONIM D, YOUNG P, JACOB HJ, et al. A high-density integrated genetic linkage and radiation hybrid map of the laboratory rat. *Genome Research* 9: AP1-AP8, 1999.
- THOMAS JW, TOUCHMAN JW. Vertebrate genome sequencing: building a backbone for comparative genomics. *Trends in Genetics* 18: 104-108, 2002.
- THOMAS JW, SCHUELER MG, SUMMERS TJ, BLAKESLEY RW, McDOWELL JC, THOMAS PJ, IDOL JR, MADURO WV, LEE-LIN SQ, TOUCHMAN JW, BOUFFARD GG, BECKSTROM-STERBERG SM, PROGRAM NC, GREEN ED. Pericentromeric duplications in the laboratory mouse. *Genome Research* 13: 55-63, 2003.
- THREADGILL DW, HUNTER KW, WILLIAMS RW. Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort. *Mammalian Genome* 13: 175-178, 2002.
- TWIGGER S, LU J, SHIMOYAMA M, CHEN D, PASKO D, LONG H, GINSTER J, CHEN CF, NIGAM R, KWITEK A, EPPIG J, MALTAS L, MAGLOTT D, SCHULER G, JACOB H, TONELLATO PJ. *Rat Genome Database (RGD): mapping disease*

- onto the genome. *Nucleic Acids Research* 30: 125-128, 2002.
- VENTER JC, ADAMS MD, MYERS EW, LI PW, MURAL RJ, SUTTON GG, SMITH HO, YANDELL M, EVANS CA, HOLT RA, GOCAYNE JD, AMANATIDES P, BALLEW RM, HUSON DH, WORTMAN JR, ZHANG Q, KODIRA CD, ZHENG XH, CHEN L, SKUPSKI et al., *The sequence of the human genome*. *Science* 291: 1304-1351, 2001.
- WADE CM, KULBOKAS EJ, KIRBY AW, ZODY MC, MULLIKIN JC, LANDER ES, LINDBLAD-TOH K, DALY MJ. *The mosaic structure of variation in the laboratory mouse genome*. *Nature* 420: 574-578, 2002.
- WATANABE TK, BIHOREAU MT, MCCARTHY LC, KIGUWA SL, HISHIGAKI H, TSUJI A, BROWNE J, et al. *A radiation hybrid map of the rat genome containing 5,255 markers*. *Nature Genetics* 22: 27-36, 1999.
- WATERSTON RH, LINDBLAD-TOH K, BIRNEY E, ROGERS J, ABRIL JF, AGARWAL P, AGARWALA R, AINSCOUGH R, ALEXANDERSSON M, AN P, ANTONARAKIS SE, et al. *Initial sequencing and comparative analysis of the mouse genome*. *Nature* 420: 520-562, 2002.
- WELSH J, PETERSEN C, MCCLELLAND M. *Polymorphisms generated by arbitrary primed PCR in the mouse: application to strain identification and genetic mapping*. *Nucleic Acids Research* 19: 303-306, 1991.
- WHEELER DL, CHURCH DM, LASH AE, LEIPE DD, MADDEN TL, PONTIUS JU, SCHULER GD, SCHRIML LM, TATUSOVA TA, WAGNER L, RAPP BA. *Database resources of the National Center for Biotechnology Information: 2002 update*. *Nucleic Acids Research* 30: 13-16, 2002.
- ZHANG Y, PROENÇA R, MAFFEI M, BARONE M, FRIEDMAN J. *Positional cloning of the mouse obese gene and its human homologue*. *Nature* 372: 425-432, 1994.