

PREMISAS E HIPÓTESIS

PREMISAS

1. La **inestabilidad genómica** es un mecanismo que actúa desde las primeras etapas de la formación del tumor, permitiendo la acumulación de múltiples alteraciones genéticas, potenciando la creación de heterogeneidad y, en consecuencia, favoreciendo la progresión tumoral.
2. En las células existen **múltiples mecanismos** de control de la estabilidad del genoma y la alteración de cualquiera de ellos puede conducir a un tipo concreto de inestabilidad genómica y a un proceso independiente de tumorigénesis.
3. En el cáncer colorrectal se hipotetizan **diferentes vías de progresión tumoral**, algunas bien conocidas, como la inestabilidad de microsatélites, y otras desconocidas y presumiblemente asociadas a inestabilidad cromosómica.
4. El **daño genómico** es la consecuencia directa de la inestabilidad genómica y, por tanto, su medida se considera una estima de ésta. El análisis de las diversas formas de daño genómico presente en los tumores puede ayudar a **determinar los diferentes tipos de inestabilidad genómica** que actúan en la progresión.
5. Existen evidencias de una relación entre el daño genómico y la agresividad tumoral. Se ha sugerido, por tanto, que la **medida del daño genómico** podría tener **valor pronóstico**.

HIPÓTESIS

La inestabilidad genómica presente en el tumor determina su evolución. Esta evolución puede ocurrir por diferentes vías de progresión tumoral que comportan unas características moleculares, cromosómicas y clínico-patológicas concretas. El estudio del daño genómico, consecuencia de la inestabilidad genómica, en un número elevado de tumores puede ayudar a caracterizar las vías de progresión tumoral y puede permitir la identificación de los grupos de tumores con peor pronóstico. La comparación con el daño genómico presente en las metástasis puede ayudar a definir el papel de la inestabilidad genómica a lo largo de la progresión tumoral.

OBJETIVO

OBJETIVO

Caracterización de distintas formas de daño genómico presentes en el cáncer colorrectal y determinación de su relación con el comportamiento biológico del tumor.

ABORDAJE

1. Análisis del daño genómico en una serie amplia de tumores colorrectales esporádicos.
 - 1a. Cuantificación del daño genómico mediante las siguientes técnicas: citometría de flujo, para medir aneuploidía y AP-PCR para cuantificar ganancias y pérdidas alélicas.
 - 1b. Comparación de los dos tipos de daño genómico entre ellos y con las variables clínico-patológicas y moleculares de los tumores.
2. Determinación del valor pronóstico de las medidas de daño genómico en el cáncer colorrectal.
3. Determinación del papel de la aneuploidía en la diseminación metastática.

MATERIAL Y MÉTODOS

1. PACIENTES Y MUESTRAS

Este estudio incluye 151 pacientes que fueron diagnosticados preoperativamente de cáncer colorrectal en el Hospital de la Santa Cruz y San Pablo de Barcelona entre Julio de 1991 y Junio de 1994. Los criterios de inclusión fueron: a) adenocarcinomas primarios intervenidos de forma electiva (cirugía no urgente); b) obtención de muestras frescas de mucosa tumoral y su correspondiente normal en un intervalo de dos horas tras la extirpación del tumor; y c) pacientes sin muerte postoperatoria. La inclusión en el estudio no influyó el tratamiento adyuvante suministrado. El protocolo del estudio fue aprobado por el Comité Ético del centro. Ninguno de los pacientes recibió quimioterapia o radioterapia previamente a la operación.

De los 151 tumores, 131 fueron analizados por citometría de flujo. Dos tumores resultaron ser hipodiploides y se excluyeron del estudio debido a que era un número insuficiente de casos para constituir un grupo específico. Por tanto **el número definitivo de casos incluidos en el estudio fue 129**. Las distribuciones para todas las variables clínico-patológicas y genéticas consideradas fueron indistinguibles entre los pacientes incluidos y excluidos. Las características más importantes son: 73 hombres y 56 mujeres, edad media 66 ± 12 años (rango: 33-96). Cuarenta tumores se localizaron en el colon derecho y 89 en el colon izquierdo incluyendo el recto. La distribución de los carcinomas según la clasificación de Dukes fue 18 A-B1, 54 B2-B3, 40 C y 17 D. A efectos del estudio el seguimiento se finalizó en Enero de 1999 y en ese momento la situación de los pacientes era: 72 pacientes vivos sin enfermedad, 6 muertos sin enfermedad, 1 vivo con enfermedad y 50 muertos por la enfermedad. El seguimiento medio fue de 65 ± 14 meses (rango: 19-85). 108 pacientes sufrieron una resección quirúrgica radical (definida por la ausencia de enfermedad residual macroscópica o microscópica y no Dukes D). Los tumores procedentes de estos pacientes se indicaron con la terminología R0.

De los 129 tumores incluidos en el estudio sólo **113 produjeron resultados satisfactorios en la AP-PCR**. Los 16 tumores descartados presentaron *fingerprints* no válidos en más de un *primer*, lo cual es indicativo de una calidad deficiente del DNA. Todos los análisis en los que intervenían parámetros derivados de la AP-PCR se realizaron con los 113 tumores, excepto los modelos de Cox, en los cuales se utilizaron los datos de los 129 casos tras realizar una imputación de los valores perdidos (o *missings*) de las variables de AP-PCR (ver apartado 5.2.2 de Material y Métodos).

Para el análisis de citometría de flujo se partió de bloques de parafina mientras que para la AP-PCR se utilizó DNA que había sido extraído previamente utilizando métodos convencionales. Para estudiar la asociación entre aneuploidía y metástasis **también se incluyeron en el análisis de citometría de flujo 23 metástasis** parafinadas complementarias a otros tantos tumores primarios incluidos en la serie inicial.

2. CITOMETRIA DE FLUJO

2.1 INTRODUCCIÓN A LA TÉCNICA

La citometría de flujo es una técnica que permite el análisis por medio de un láser de células o partículas que se mueven a lo largo de un líquido que fluye. El funcionamiento del sistema se muestra en la figura 4. Las células se encuentran en el compartimento interno de la cámara de flujo suspendidas en un fluido isotónico que es un líquido envolvente que permite el desarrollo de un flujo laminar. A medida que las células salen, de una en una, de la cámara de flujo, pasan a través del punto de detección, en el que la luz del láser incide directamente sobre ellas. Cuando esto ocurre, la luz se dispersa en todas las direcciones. La luz dispersada hacia delante (forward scatter (FS)) es proporcional al tamaño de la célula o partícula y la luz dispersada 90° respecto al haz del láser (side scatter (SS)) se considera proporcional a la granularidad. Los sistemas sensores recogen estas señales de luz. El citómetro de flujo también puede analizar los fotones de diferente longitud de onda que emiten las moléculas fluorescentes al ser excitadas por una longitud de onda apropiada. Un ejemplo de sustancia fluorescente es el yoduro de propidio (IP) que, como agente intercalante de DNA, permite analizar de forma estequiométrica el contenido de DNA celular. Sus longitudes de onda óptimas de excitación y de emisión son 493 y 639 respectivamente. Los fotones procedentes de la emisión de moléculas fluorescentes son seleccionados mediante un sistema de filtros adecuado. Después se recogen en detectores que los convierten en impulsos eléctricos y luego éstos se procesan a señales digitales por medio de un conversor analógico-digital. Un ordenador capta esta información, la procesa y la presenta en forma de histogramas (Macey, 1994). (ver apartado 1 de Resultados).

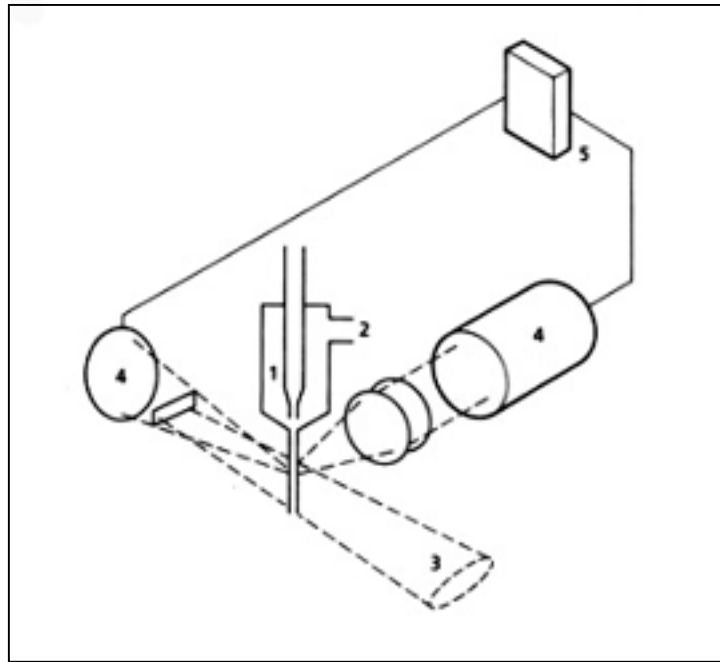


Figura 4. Representación esquemática de un citómetro de flujo. 1, cámara de flujo; 2, lugar de entrada del líquido envolvente; 3, láser; 4, sistemas sensores; 5, ordenador (extraído de Macey, 1994).

En este estudio la citometría de flujo ha sido utilizada para medir el contenido de DNA de las células tumorales y detectar su nivel de aneuploidía mediante comparación con células control diploides. El límite estimado de detección de la técnica se encuentra en aneuploidías que supongan un alteración del contenido de DNA de más del 10%, es decir, de unos dos o tres cromosomas.

2.2 METODOLOGÍA

Microdissección

Consiste en seleccionar la zona del tumor en la que hay un mayor porcentaje de células tumorales con el fin de eliminar la contaminación por células diploides. Para ello se hicieron cortes de 5 μm a partir de los bloques de parafina y se tiñeron con hematoxilina-eosina. Un patólogo experto revisó todos los cortes y señaló en ellos la zona del frente de infiltración del tumor, en la que se estima una presencia aproximada de un 75% de células tumorales. Se tuvo especial cuidado en descartar las zonas con tejido necrótico o con un alto componente inflamatorio. La zona señalada en el portaobjetos

servió de plantilla para ser recortada con máxima precisión a partir de los bloques originales (figura 5). Esta vez se realizaron entre uno y tres cortes (según el tamaño de la zona a recortar) de 50 μm . En algunos casos después se hicieron cortes de 5 μm para confirmar el trozo recortado.

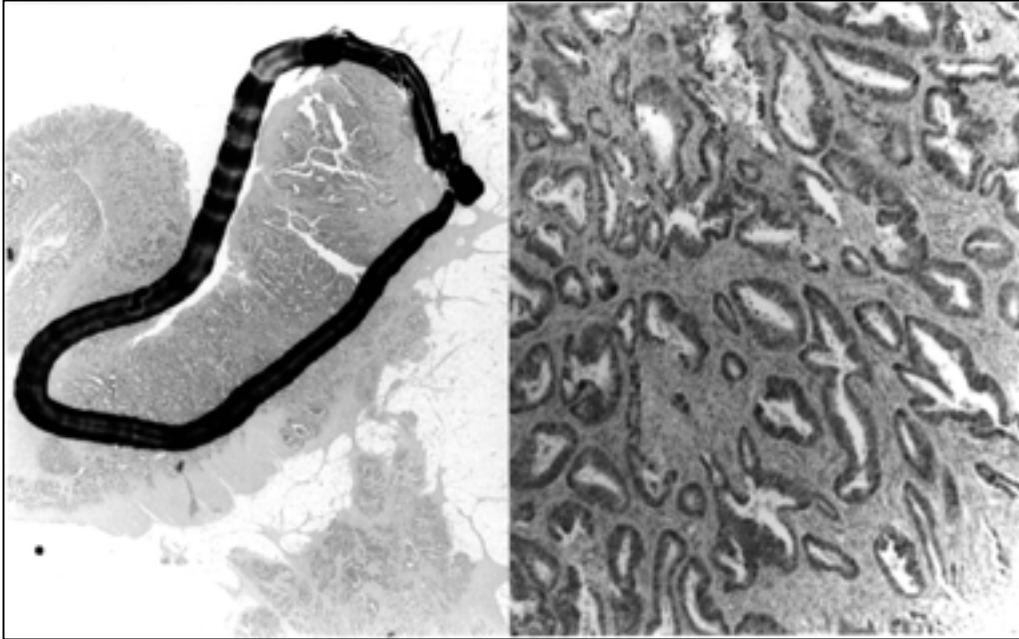


Figura 5. Fotografía ilustrativa del proceso de microdissección. (Izquierda) Carcinoma infiltrante en el que se señala el área correspondiente al frente invasivo, la cual fue microdisseccionada para el análisis de citometría de flujo. (Derecha) Detalle de la zona microdisseccionada.

Procesamiento de tumores parafinados para el análisis del contenido de DNA por citometría de flujo

El protocolo consta de las siguientes fases: desparafinación, hidratación, disgregación mecánica, disgregación enzimática, tinción del DNA, filtración de la suspensión nuclear y análisis.

1. **Desparafinación:**

- Colocar los cortes de 50 μm en tubos de vidrio.
- Lavados por decantación cada 15 min con xileno (5-7 ml) durante 2 horas.

2. Hidratación:

- Eliminación del xileno mediante dos lavados con etanol absoluto (3-5 ml) y decantación.
- Hidratación de los cortes mediante inmersión secuencial en etanol 95%, 75% y 50% por adición de H₂O bidestilada e incubar 15 min a temperatura ambiente.
- Decantar y lavar dos veces con H₂O bidestilada (3-5 ml).

3. Disgregación mecánica:

- Disgregación mediante bisturí y sobre una placa de Petri pequeña hasta obtener de cada corte una pasta totalmente homogénea.
- El tejido disgregado se coloca en un tubo eppendorf de 1.5 ml.

4. Disgregación enzimática:

- Añadir a cada tubo 1ml de una solución de NaCl al 0.9% en H₂O bidestilada, pH=1.5 con un 0.5% de pepsina.
NOTA: Es fundamental que el pH sea de 1.5 para que la pepsina tenga actividad. Tras este paso la estructura celular habrá desaparecido y trabajaremos con núcleos.
- Incubar los tubos a 37°C en agitación durante 1h.

5. Tinción del DNA:

- Centrifugar 8 seg a 14000 rpm (2 veces, cambiando la posición del tubo la segunda vez para facilitar que las células se depositen en el fondo y no en las paredes).
- Eliminar el sobrenadante. Resuspender el botón celular en 1ml de solución al 0.5% de ribonucleasa A en PBS y IP a una concentración de 20µg/ml.
NOTA: El IP es un fluorocromo, por tanto es muy importante preservarlo de la luz para que no pierda actividad. Esta solución lleva ribonucleasa A para degradar el RNA, ya que el IP se une tanto a DNA como a RNA.
- Incubar los tubos de 12-20h a 4°C.

6. Filtración:

- Filtración en malla de nylon de 50 µm de diámetro con el fin de eliminar restos de tejido que puedan obstruir la cámara del citómetro de flujo. Recoger el filtrado en tubos eppendorfs nuevos.
NOTA: Utilizamos jeringas de insulina de aguja extraíble. Colocamos el filtro entre la jeringa y la aguja y después cortamos la aguja por la parte de plástico, de manera que el filtro quede sujeto entre la jeringa y la anilla de plástico de la aguja. La filtración permite la eliminación de los agregados celulares.

7. Análisis:

- Vaciar el contenido en tubos de plástico apropiados para el citómetro de flujo. Se utilizó el citómetro de flujo EPICS-XL-MCL (Coulter Electronics, Hialeah, Florida). Como fuente de excitación se usó un láser de argón sintonizado a 488nm. La lectura de emisión de fluorescencia se hizo a 675 nm.

2.3 PROCESAMIENTO DE DATOS

Controles

Como se ha dicho anteriormente, las aneuploidías se detectan mediante la comparación con células control diploides. En nuestras muestras utilizamos **dos tipos de control: externo e interno**. El mejor control externo se considera que es el **tejido normal** correspondiente al tipo de tumor examinado (Shankey et al, 1993). En nuestro caso utilizamos mucosa colónica normal, la cual se procesó como una muestra más en todas las tandas de análisis de tumores que se llevaron a cabo. El coeficiente de variación medio (de los picos diploides G0/G1) fue de 3.43 ± 0.39 . El control interno lo constituyó la **población de linfocitos infiltrantes presentes en las muestras**. A pesar del enriquecimiento de células tumorales por microdissección, es muy difícil la eliminación de todas las células contaminantes, especialmente los linfocitos, los cuales son muy abundantes en la mucosa colorrectal. Sin embargo, su presencia no constituye un problema, sino una ventaja ya que gracias a las diferentes características de dispersión de luz que muestran sus núcleos, éstos pueden ser fácilmente identificados y seleccionados y se puede analizar por separado su contenido de DNA para usarlo como referencia del contenido diploide.

Procesamiento de datos. 1er nivel: Evaluación del contenido de DNA

Las muestras se analizaron en el citómetro de flujo usando el programa EPICS XL System II. Este programa elabora los datos y los almacena en forma de unos archivos llamados *list mode* que están compuestos de histogramas. Estos histogramas se definen en función de la información que se desea analizar y se describen detalladamente en el apartado 1 de Resultados. El análisis de la ploidía se llevó a cabo siguiendo el método de Petriz et al (1996) y las recomendaciones de la 1992 DNA Flow Cytometry Consensus Conference (Shankey et al, 1993). **La aneuploidía se definió como la presencia de cualquier población de células o núcleos no apoptóticos que produjese un perfil distinto del patrón diploide**. Para cuantificarla se utilizó el parámetro estándar en citometría de flujo que es el **DNA index (DI)**. Éste se define como el contenido de DNA del pico G0/G1 de las células aneuploides dividido entre el contenido de DNA del pico G0/G1 de las células diploides. De esta manera los tumores diploides tienen $DI=1$, lo que equivale a un contenido de DNA $2n$, y los aneuploides $DI \neq 1$. Este puede ser $DI < 1$ si se trata de una hipodiploidía (el pico aneuploide aparece a la izquierda del pico $2n$) o $DI > 1$ si es una hiperploidad (el pico aneuploide aparece a la derecha). En algunos tumores el

grado de aneuploidía es muy pequeño y el pico aneuploide no es fácilmente distinguible porque está parcialmente solapado con el pico diploide. Estos tumores tienen valores de DI entre 1 y 1.1 y se llaman *near* diploides. Los tumores que presentan más de un pico aneuploide se llaman multiploides.

Procesamiento de datos. 2º nivel: Mejora de los porcentajes y análisis de la proliferación

En los histogramas de contenido de DNA, además de la presencia de aneuploidía, también se puede determinar el porcentaje de células que se encuentran en las diferentes fases del ciclo celular. Pero esta determinación conlleva problemas ya que es muy fácil confundir las células en G2/M con picos tetraploides. Para solucionar este problema existen unos programas que permiten aplicar a los histogramas de DNA unos modelos matemáticos que facilitan la interpretación de los gráficos y disminuyen los errores. En nuestro caso todos los archivos *list mode* fueron procesados usando el software Mod Fit LT para Macintosh versión 2.0 (Verity Software House, Inc., Topsham, Maine) y todos los datos utilizados en cálculos posteriores, incluido el DI, fueron los proporcionados por este programa. Se usó el modelo automático para muestras parafinadas con posible debris, en el cual no se tuvo en cuenta la presencia de apoptosis y se consideró la aparición de picos aneuploides (las muestras con aneuploidía se habían determinado previamente mediante el estudio de los archivos *list mode*). Aplicando estas correcciones los análisis de proliferación celular son más fiables, pero, aún así, en el caso de los tumores aneuploides se produce un solapamiento de poblaciones que es difícilmente valorable sin cometer errores. Por esta razón **en nuestras muestras los estudios de proliferación sólo se llevaron a cabo en los casos diploides**. Se utilizaron dos parámetros: la fracción de células en fase S (%S) y el índice de proliferación (PI). Este último es más preciso porque también incluye las células en G2/M y se calcula con la siguiente fórmula: $PI = (\% \text{cel en S} + \% \text{cel en G2M}) / \% \text{cel en G0/G1}$.

Procesamiento de datos. 3er nivel: Limitaciones de los parámetros convencionales y nuevas aproximaciones al análisis de la aneuploidía

La observación de que muchos tumores aneuploides poseían un elevado DNA index pero bajo porcentaje de células aneuploides nos llevó a la hipótesis de que la malignidad del tumor podría depender no tan solo de la existencia de células aneuploides, sino también de su cantidad (apartado 1.1 de Discusión). Para probar esta hipótesis decidimos crear dos nuevos parámetros: el 'cell index' (CI) y el 'aneuploidy index' (AI). El CI mide el *ratio* entre células diploides y aneuploides (ambas en fase G0/G1) independientemente

del contenido de DNA del pico aneuploide. Un tumor con CI=1 tiene la misma proporción de células diploides que de aneuploides. El AI integra los dos conceptos: magnitud de la alteración (DI) y cantidad de células que la presentan (CI). Es decir, mide el contenido de DNA de las células, como el DI, pero lo corrige por la proporción de células aneuploides. Por tanto el AI se puede considerar como la media del contenido de DNA de toda la muestra. A continuación se muestran las fórmulas para el cálculo de cada parámetro y en el apartado 1 de resultados se expone un ejemplo de su aplicación.

$$DI = \frac{\text{contenido DNA cel aneuploides}}{\text{contenido DNA cel diploides}}$$

$$CI = \frac{\% \text{ cel aneuploides}}{\% \text{ cel aneuploides} + \% \text{ cel diploides}}$$

$$AI = \frac{\% \text{ cel diploides} \times 1 + \% \text{ cel aneuploides} \times DI}{\% \text{ cel diploides} + \% \text{ cel aneuploides}}$$

3. AP-PCR

3.1 INTRODUCCIÓN A LA TÉCNICA

La AP-PCR (Arbitrarily Primed PCR) es una técnica de *fingerprint* de DNA que fue descrita por Welsh y McClelland en el 1990 (Welsh y McClelland, 1990). Se basa en la amplificación de DNA genómico con *primers* arbitrarios y unos ciclos iniciales de baja astringencia (que habitualmente se consigue usando baja temperatura de *annealing*). En estas condiciones el *primer* arbitrario se une a múltiples sitios del genoma con los que comparte cierta homología. Después de unas cuantas rondas de amplificación se produce cierto número de productos que han incorporado la secuencia original del *primer* en sus extremos. Después se aumenta la temperatura de *annealing* (alta astringencia) y entonces estos productos se amplifican mucho más, generando un *fingerprint* de bandas discretas cuando se resuelve en gel de electroforesis. La amplificación sólo tiene lugar si existen dos lugares posibles de unión del *primer* en las cadenas opuestas del DNA y a

una distancia moderada, normalmente 1Kb o menos. En las condiciones de rutina se obtienen *fingerprints* que constan de entre 40 y 80 bandas suficientemente intensas como para poder ser detectadas. El patrón de bandas que se obtiene es característico y representativo del genoma usado como molde. Además, **la competición que se establece entre los posibles lugares de *annealing* del *primer* hace que los patrones sean reproducibles y que reflejen las diferencias en la abundancia relativa de las secuencias de DNA** (Figura 6). Otra característica interesante de la AP-PCR es que las secuencias amplificadas proceden de regiones del genoma seleccionadas al azar. Como consecuencia, al estudiar el origen cromosómico de las bandas se encuentra que los cromosomas más grandes están representados por más bandas que los cromosomas pequeños, lo que confirma la arbitrariedad de la técnica (Arribas et al, 1997; Arribas et al, 1999a).

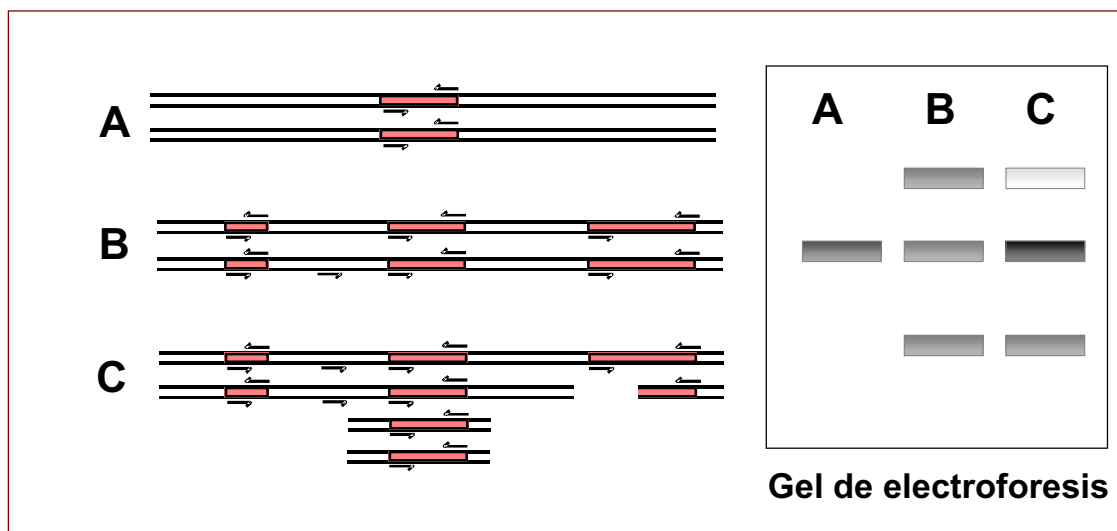


Figura 6. Principio de la AP-PCR. El aumento o disminución del número de copias de una secuencia se refleja en un aumento o disminución de la intensidad de la banda correspondiente.

La AP-PCR permite la detección de la mayoría de alteraciones genéticas que ocurren en el cáncer mediante la comparación de los patrones de bandas del tejido normal y tumoral. Una disminución de la intensidad de una banda se interpreta como una pérdida de la secuencia de DNA molde, la cual puede ser debida a una deleción, a una monosomía o a una translocación que implique parte de la banda. Un aumento de la intensidad de una banda indica una ganancia de la secuencia de DNA, que puede ser debida a una duplicación, a amplificaciones génicas o a trisomía. La aparición o desaparición de bandas puede deberse a translocaciones o a mutaciones puntuales en el lugar de unión del *primer*. Las variaciones de tamaño son una indicación de inserciones y deleciones en el DNA molde. Además, si las inserciones o las deleciones son muy grandes pueden impedir la amplificación de la banda en cuestión. Las triploidías y tetraploidías perfectas son indetectables por AP-PCR (Navarro y Jorcano, 1999).

El análisis de una serie amplia de tumores mediante varios experimentos (usando diferentes *primers*) de AP-PCR permite una visión general del grado de alteración de sus genomas. La información obtenida se puede analizar en dos sentidos: (1) estimación del daño genómico de los tumores, con la intención de estudiar la relación de esta variable con otros factores clínico-patológicos y moleculares y de evaluar su valor pronóstico: a esta aproximación le llamaremos **estudio de casos**; y (2) identificación de secuencias que sufren alteraciones recurrentes y que podrían pertenecer a oncogenes o genes supresores de tumores o bien podrían dar pistas sobre los mecanismos de producción de daño genómico que actúan en los tumores: a esta aproximación le llamaremos **estudio de bandas**. Para llevar a cabo este segundo estudio hay que tener en cuenta otra de las grandes ventajas que ofrece la AP-PCR: la posibilidad de aislar y clonar cualquier banda arbitraria. Esto permite, por un lado, la identificación de la secuencia de la banda y su localización cromosómica y, por otra, la confirmación de los cambios observados con otras técnicas (Peinado et al, 1992; Yasuda et al, 1996; Arribas et al, 1999a).

3.2 METODOLOGÍA

DNA

El DNA se había extraído previamente siguiendo métodos convencionales. Se prepararon diluciones de 20 ng/μl de las cuales se usaron 2.5 μl para cada reacción de AP-PCR. Todos los DNAs se corrieron en geles de agarosa al 1% para **comprobar las concentraciones y la calidad del DNA**. Esto es extremadamente importante en los experimentos de AP-PCR ya que, por un lado, concentraciones diferentes entre el DNA tumoral y su correspondiente normal invalidan la comparación de sus patrones de bandas y, por otro, la degradación del DNA se traduce en patrones anómalos, normalmente con una marcada disminución de la intensidad de las bandas en la zona superior o inferior del carril, lo cual, lógicamente, también invalida la comparación.

Reacción

Se realizaron tres experimentos diferentes, dos de ellos llevados a cabo con un único *primer* (331DB2 y Blue) y un tercero llevado a cabo con una pareja de *primers* (D12S77a y D12S77b). Los *primers*, que en los tres casos habían sido descritos en experimentos previos (331DB2 y D12S77: Arribas et al, 1997; Blue: Arribas et al, 1999a), se seleccionaron en base a su reproducibilidad y a la calidad del *fingerprint*. En la tabla 4 se muestran las condiciones de PCR para cada uno de los experimentos. Todas las reacciones se llevaron a cabo en termocicladores Programmable Thermal Controller-100 (MJ Research, Inc., Watertown, Massachusetts). En el caso de los *primers* 331DB2 y Blue, diseñados específicamente para realizar AP-PCR, la amplificación consta, como suele ser habitual, de un paso inicial de desnaturalización (94°C), 5 ciclos de **baja astringencia**, 35 ciclos de **alta astringencia** y un paso final de extensión (72°C). El experimento D12S77 fue diseñado para amplificar una secuencia microsatélite y por eso se utilizan dos *primers* y no se hacen ciclos iniciales de baja astringencia. En las condiciones descritas, además de la amplificación de la secuencia microsatélite se produce un patrón de bandas de origen cromosómico múltiple y altamente reproducible.

Tabla 4. Condiciones de los experimentos de AP-PCR.

331DB2	BLUE	D12S77
ACAGATCTGAAGGGTGAATATTCTCC	CCGAATTCGCAAAGCTCTGA	GAAGGGCAACAACAGTGAA (D12S77a) CTTTTTTTTCTCCCCACTC (D12S77b)
1X tampón de PCR ¹	1X tampón de PCR ¹	1X tampón de PCR ¹
120 µM de cada dNTP ²	100 µM de cada dNTP ²	200 µM de cada dNTP ²
1 mM MgCl ₂	2.16 mM MgCl ₂	1 mM MgCl ₂
2 µM primer	1.08 µM primer	1 µM primer
1.5 µCi α-33P-dATP ³	1.5 µCi α-33P-dATP ³	1.5 µCi α-33P-dATP ³
1 unidad de Taq Polimerasa ¹	1.1 unidades de Taq Polimerasa ¹	1.25 unidades de Taq Polimerasa ¹
50 ng de DNA genómico	50 ng de DNA genómico	50 ng de DNA genómico
H ₂ O hasta volumen final de 25 µl	H ₂ O hasta volumen final de 25 µl	H ₂ O hasta volumen final de 25 µl
Reacción de PCR:	Reacción de PCR:	Reacción de PCR:
95°C 45 sec	94°C 4 min	94°C 30 sec
94°C 45 sec	94°C 30 sec	94°C 20 sec
40°C 30 sec	50°C 60 sec	55°C 30 sec
72°C 75 sec	72°C 90 sec	72°C 45 sec
94°C 45 sec	94°C 15 sec	72°C 5 min
60°C 30 sec	60°C 15 sec	
72°C 60 sec	72°C 60 sec	
72°C 5 min	72°C 7 min	

¹ DNA polimerasa de *Thermus aquaticus* (Taq) (Boehringer Mannheim) a 5 u/µl. Composición del tampón de PCR en la reacción final (10 mM Tris-HCl pH 8.0, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatina). El tampón viene suministrado junto con la Taq DNA Polimerasa en una concentración 10 veces superior a la necesaria para la reacción. ² dNTP: desoxinucleótidos trifosfato (Pharmacia Biotech). ³ α-33P-dATP: desoxiadenosina trifosfato marcado radioactivamente (10 mCi/ml) (Amersham).

Electroforesis y exposición

Los productos de PCR se diluyeron (1:4) en tampón de carga desnaturante DLB (93% formamida, 0.01M EDTA, 0.01M NaOH, 0.01% azul de bromofenol, 0.01% xilencianol), se desnaturaron a 95°C durante 3 minutos e inmediatamente se enfriaron en hielo. Después se analizaron 3 µl de cada muestra en **geles de secuenciación desnaturantes** 6% acrilamida, 8M urea (40 cm de largo, 30 cm de ancho, 0.4 mm de grueso). Antes de cargar las muestras los geles se precorrieron 10 min a 55w para equilibrar los tampones del gel y de la cubeta. Las electroforesis se llevaron a cabo a 55w durante 6h en el caso de los *primers* 331DB2 y Blue y 3.5h para el *primer* D12S77. Debido a la gran variabilidad que factores externos poco controlables, como la temperatura ambiental, introducen en la velocidad de la carrera, el tiempo de electroforesis se intentó regular no tanto por el tiempo real como por el desplazamiento

de los colorantes del tampón de carga. De esta manera se pretendía estandarizar la longitud de los patrones de bandas con el fin de facilitar su posterior análisis. Los geles se secaron al vacío a 85°C y se expusieron a una película de rayos X a temperatura ambiente y sin pantalla intensificadora durante 3-5 días.

3.3 ANÁLISIS DE LOS PATRONES DE BANDAS

Aunque la observación a simple vista de los patrones obtenidos permite la identificación de bandas que sufren alteraciones, es evidente que para que los resultados obtenidos sean fiables y reproducibles es necesario utilizar un método eficiente y estandarizado que permita llevar a cabo la comparación de los *fingerprints* tumorales y normales. El método más objetivo es la **densitometría**. Para ello las películas se escanearon (EPSON GT-8500 con unidad de transparencias; Seiko Epson Corporation, Nagano, Japón) utilizando el programa de captación y análisis de imágenes Adobe Photoshop y se guardaron en formato TIFF. Estas imágenes se transfirieron al programa de densitometría Phoretix 1D Advanced Version 3.0 (Nonlinear Dynamics Ltd, Newcastle upon Tyne, Reino Unido), el cual permite la cuantificación de la intensidad de las bandas. La figura 7 muestra esquemáticamente el funcionamiento de este programa. Los carriles con intensidades no comparables entre tejido normal y tumoral o con patrones de bandas anómalos se descartaron. Sin embargo, cuando las irregularidades se limitaban a una pequeña zona del carril ésta se descartaba, pero se valoraba el resto de bandas del carril. Se consideraron como bandas todas las marcas bien definidas y claramente distinguibles del fondo. Incluso cuando se identificaban dos bandas cercanas que podían corresponder a las dos cadenas de una única secuencia de DNA debido a su comportamiento similar, éstas se seguían considerando como bandas independientes, ya que constituían controles internos de la técnica. Aunque el programa de densitometría permite la medida de la intensidad de la banda considerando su volumen, debido a la dificultad de establecer los límites de las bandas y al solapamiento de unas con otras, se consideró que el método más apropiado era la determinación de la intensidad de pico. Una vez que se hubo determinado la intensidad de todas las bandas analizadas, los datos obtenidos en Phoretix se pasaron a una hoja de cálculo Excel con el fin de comparar las intensidades de las bandas de los tumores con sus correspondientes normales y así determinar aquellas que habían sufrido alteración.

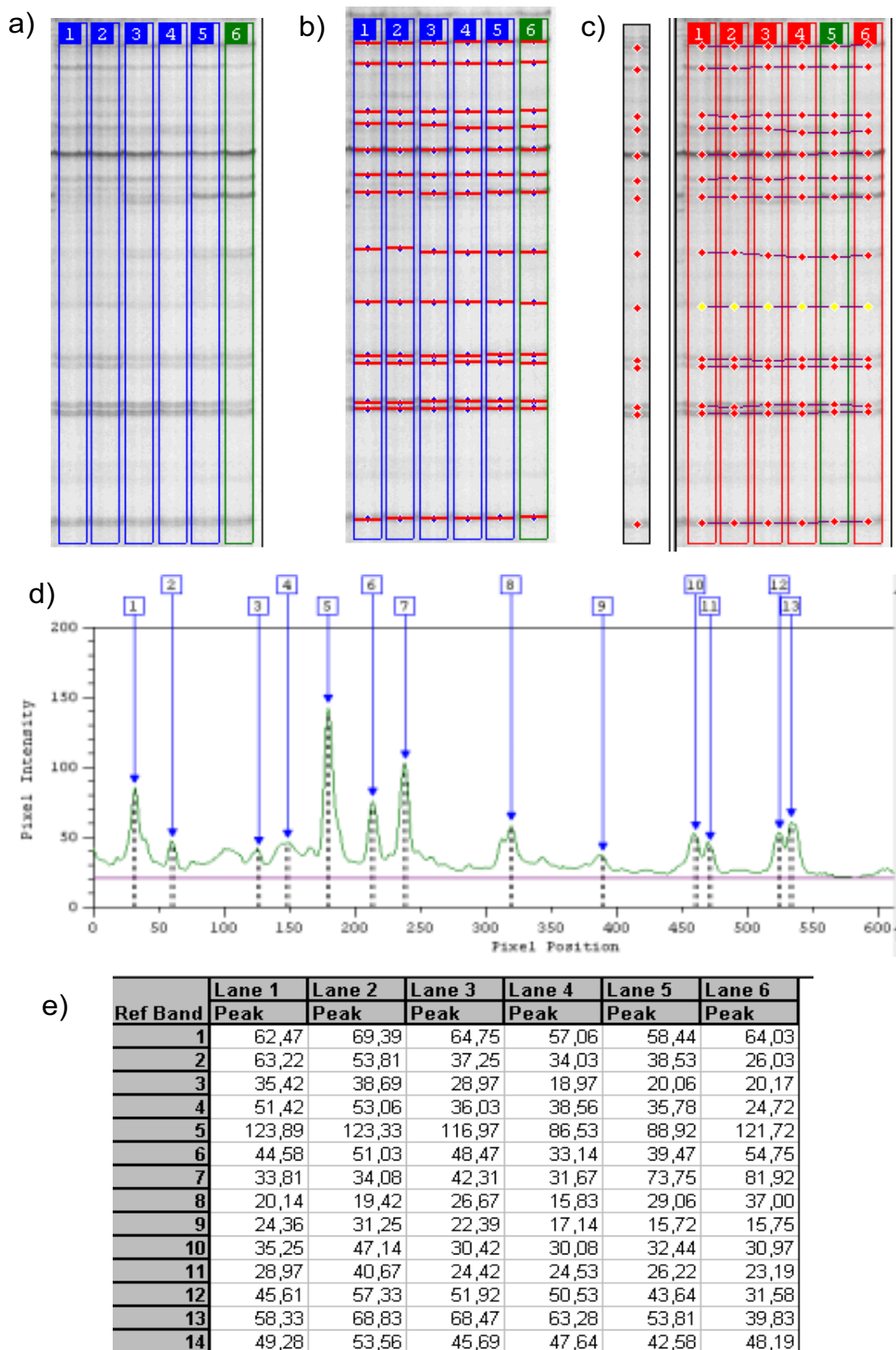


Figura 7. Esquema del funcionamiento del programa de densitometría Phoretix 1D Advanced. a) Definición de los carriles. b) Detección de las bandas. c) Alineación de las bandas respecto al patrón de referencia. d) Representación del perfil de intensidad de las bandas de un carril. e) Cuantificación de la intensidad de las bandas de todos los carriles.

3.4 CUANTIFICACIÓN DEL DAÑO GENÓMICO Y DETECCIÓN DEL LÍMITE DE VARIABILIDAD

Para reducir la posibilidad de error causada por pequeñas variaciones en la intensidad de los carriles, para cada banda **la medida absoluta de la intensidad se modificó a una medida relativa** con la siguiente fórmula:

$$\text{intensidad relativa} = \frac{\text{intensidad de la banda}}{\text{intensidad total del carril}} \times \text{n}^\circ \text{ bandas del carril} \times \text{int. media del experim.}$$

Después se calculó, para cada banda, la **diferencia de intensidad entre tejido tumoral y normal**. Debido a que la valoración de la intensidad se hizo de forma cuantitativa ninguna banda mostró valor 0 para esta diferencia. Pero esto no quiere decir que todas las bandas han sufrido alteraciones, sino que existe cierta variabilidad dentro de la técnica que se pone de manifiesto al cuantificar el valor de la intensidad de cada banda. Esto implica que hay que **analizar la variabilidad de la técnica** para determinar cuál es el límite a partir del cual los cambios que encontramos se consideran alteraciones reales. Para ello se llevó a cabo un **estudio de reproducibilidad intraensayo** que consistió en la realización de cuatro experimentos de AP-PCR, cada uno con una muestra diferente de DNA repetida cuatro veces. A partir de los datos obtenidos se dedujo cuál era la recta que se podía considerar como límite de variabilidad de la técnica y que ofrecía un 5% de probabilidad de error en la determinación de los cambios (ver figura 10 en el apartado 2 de Resultados). Para cada banda se calculó el límite de variabilidad que le correspondía en función de la recta. Para hallar este valor se extrapoló en la recta la intensidad de la banda correspondiente a la media entre el tejido tumoral y el normal. **Sólo cuando la diferencia de intensidad tumor menos normal superaba en valor absoluto el límite de variabilidad correspondiente se consideraba que correspondía a una alteración real en el DNA del tumor**. Para simplificar el análisis **no se consideró la magnitud del cambio** (es decir, su valor numérico), **sino únicamente su presencia o su ausencia y la dirección del cambio** (es decir, ganancias o pérdidas de intensidad) (ver figura 11 en el apartado 2 de Resultados).

Los resultados de los tres experimentos se analizaron conjuntamente y sólo se consideraron válidos los casos que habían producido un patrón de bandas aceptable como mínimo en dos de los tres *primers*. Para cada tumor se cuantificó el daño genómico usando tres variables distintas: la fracción de daño genómico o **GDF (Genomic Damage**

Fraction), la fracción de ganancias o **GF (Gains Fraction)** y la fracción de pérdidas o **LF (Losses Fraction)**, que se definieron de la siguiente manera:

$$\text{GDF} = \frac{\text{n}^\circ \text{ bandas con cambios}}{\text{n}^\circ \text{ bandas totales}} \quad \text{GF} = \frac{\text{n}^\circ \text{ bandas con ganancias}}{\text{n}^\circ \text{ bandas totales}} \quad \text{LF} = \frac{\text{n}^\circ \text{ bandas con pérdidas}}{\text{n}^\circ \text{ bandas totales}}$$

Obviamente, GDF es la suma de GF y LF.

3.5 ANÁLISIS DE BANDAS CON CAMBIOS RECURRENTES

Paralelamente al estudio del daño genómico de los casos, también se llevó a cabo un estudio de las bandas con cambios recurrentes en el tumor. En primer lugar, se realizó la estadística univariante (frecuencia de casos que presentan ganancias o pérdidas de dicha banda), multivariante (asociación con variables clínico-patológicas y moleculares) y de supervivencia de cada una de ellas. En segundo lugar, en base a esos datos y a la localización cromosómica (ver más adelante) se seleccionaron las bandas más interesantes desde el punto de vista de su posible relación con la tumorigénesis. Estas bandas fueron objeto de un estudio más exhaustivo, que es el que se detalla a continuación.

Localización cromosómica y subcromosómica

Las bandas de los *fingerprints* de AP-PCR se localizaron en sus cromosomas correspondientes mediante el uso de un **panel de híbridos monosómicos humano/roedor** (NIGMS mapping panel 2, Coriell Institute for Medical Research, Camden, New Jersey). Este sistema consiste en un conjunto de 24 híbridos celulares, cada uno de los cuales contiene un cromosoma humano distinto en un *background* de DNA de roedor. Se trata de amplificar con los *primers* y las condiciones de la AP-PCR estos híbridos y un control de DNA humano y otro de ratón. Las AP-PCRs de los híbridos producen el patrón de bandas característico del ratón más las bandas correspondientes al cromosoma humano presente en ese híbrido. Por comparación con los controles de humano y ratón se puede identificar la localización cromosómica de todas aquellas bandas humanas que no se solapan con bandas del *fingerprint* de ratón. De esta manera se pudo determinar, de forma fácil y rápida y sin necesidad de clonaje y secuenciación, la localización cromosómica de un gran número de bandas procedentes de los tres *fingerprints* analizados. Este experimento aportó información muy valiosa para acotar el

estudio de bandas interesantes y centrarlo únicamente en aquellas bandas localizadas en regiones cromosómicas de especial relevancia.

Una aproximación más detallada (localización subcromosómica) se llevó a cabo mediante el uso del **panel de híbridos irradiados** Stanford G3 (Research Genetics, Inc., Huntsville, Alabama). Este panel está compuesto por 83 líneas celulares de roedor, cada una de las cuales contiene trozos aleatorios de DNA humano derivado de DNA total fragmentado por irradiación. La amplificación de estos híbridos produce un resultado positivo o negativo según si la banda en cuestión se encuentra o no en alguno de los fragmentos presentes en el híbrido. Los resultados se ordenan en forma de código binario de 83 dígitos en el que 1 indica amplificación y 0 no amplificación. Como un gran número de marcadores de posición conocida han sido mapados previamente usando estos paneles, la comparación del código binario obtenido para la banda en cuestión con los códigos almacenados de la colección de marcadores conocidos (llevado a cabo vía internet por el servidor RH, con dirección <http://shgc-www.stanford.edu>) permite la localización de la banda relativa a la posición de otros marcadores con una elevada precisión (la resolución del panel es de 500 Kb). Aunque este panel, al igual que el anterior, se puede amplificar con las condiciones de AP-PCR, al ser más complejo la identificación de las bandas resulta mucho más difícil. Por ello, además de usar AP-PCR, también se utilizaron condiciones de PCR normal, en las que se usaban *primers* específicos de bandas que previamente habían sido clonadas y secuenciadas (ver más adelante) y cuya localización subcromosómica resultaba de especial interés. Gracias a la reciente publicación de la secuencia del genoma humano hoy en día es posible un mapaje mucho más preciso de estas bandas, por lo que la aproximación utilizada aquí ya es obsoleta.

Aislamiento y clonaje

Las bandas seleccionadas se obtuvieron a partir de los geles de AP-PCR mediante la **escisión de la banda** con una cuchilla y su posterior elución en tampón TE 10:1 (10 mM Tris pH 7.5, 1 mM EDTA) a 65°C durante 2h. Es importante recordar que para que la identificación de la banda en el gel sea posible es necesario, previamente a la exposición, haber marcado el gel con etiquetas fluorescentes que luego permitan el solapamiento de la película, en la que señalaremos la banda a recortar, sobre el gel en cuestión. Una vez hecho el corte, la posterior reexposición del gel nos indicará si éste es válido (es decir, si incluye realmente la banda de interés).

Las bandas eluidas se reamplificaron con sus correspondientes *primers* de AP-PCR, pero usando condiciones estándar de PCR. Los productos se corrieron en geles pequeños de acrilamida (17 cm de largo, 15 cm de ancho, 0.8 mm de grueso) junto con un marcador de peso molecular (ϕ X174-HaeIII) que permitió la cuantificación de los productos. Tras ello se procedió al **clonaje** de las bandas utilizando el sistema TA Cloning Kit (Invitrogen). Una vez obtenidos los clones se seleccionó un número elevado de ellos (10-15) para realizar un lisado de las bacterias (centrifugar 50-100 μ l de cultivo 5 min a velocidad máxima, resuspender el sobrenadante en el mismo volumen de agua y hervir 5 min para lisar las bacterias. Volver a centrifugar y recoger el sobrenadante, en el que se encuentra el plásmido). El lisado se amplificó por PCR y se corrió en gel pequeño de acrilamida junto con una muestra de la banda original que sirvió como referencia para identificar los clones que habían incorporado el inserto correcto (o, en el peor de los casos, un inserto que tenía el mismo tamaño que la banda en cuestión). Una vez identificados estos clones se procedió a la **extracción y purificación del plásmido** mediante el GFX Micro Plasmid Prep Kit (Pharmacia). La razón por la que no se lleva a cabo directamente la extracción del plásmido de todos los clones obtenidos es porque muchos de ellos no contienen el inserto correcto y, ya que la realización de un lisado es mucho más rápida que la extracción del plásmido, resulta conveniente hacerla primero para descartar todos los clones no válidos.

Secuenciación y determinación de homología

Una vez que se obtuvo el plásmido purificado se procedió a la **secuenciación del inserto** mediante el AmpliCycle Sequencing Kit (Perkin Elmer) utilizando los *primers* M13 Forward y M13 Reverse suministrados en el kit. Los tiempos de electroforesis fueron 2h, 4h y 6h, de manera que las secuencias se solapaban de un gel a otro y así se permitía la lectura completa. Los lugares de inicio y finalización de la banda se identificaban por la presencia del *primer* de AP-PCR. En algunas de las bandas más grandes, a pesar de secuenciar por los dos extremos y de hacer carreras muy largas, fue imposible solapar la secuencia, por lo cual se utilizaron los dos trozos para estudiar la homología. Se secuenciaron entre 2 y 4 clones de cada banda y en la mayoría de los casos las secuencias fueron coincidentes y por tanto se dieron como válidas. En los casos en los que esto no ocurrió se realizó un Southern Blot a partir del gel de AP-PCR para determinar cual de las secuencias correspondía a la banda en cuestión.

La **homología** de las bandas aisladas con secuencias descritas previamente e introducidas en las bases de datos se analizó mediante el software BLAST 2.1 (Basic BLAST search) (<http://www.ncbi.nlm.nih.gov/BLAST>).

4. OTROS ANÁLISIS MOLECULARES

Estos casos se habían analizado previamente en nuestro laboratorio para tres variables moleculares de crucial importancia en el cáncer colorrectal: la inestabilidad de microsatélites (Gonzalez-Garcia et al, 2000) y las mutaciones en los genes p53 y k-ras (Tortola et al, 1999a). En el análisis de la **inestabilidad de microsatélites** se habían incluido los 129 casos de este estudio, de manera que se disponía de los datos de todos ellos. La distribución fue la siguiente: 7 casos positivos para inestabilidad de microsatélites y 122 casos negativos. Sin embargo los datos de las mutaciones de p53 y k-ras estaban incompletos: faltaban por analizar 27 casos de p53 y 34 casos de k-ras. Se creyó necesario disponer de esta información y, por tanto, se procedió al análisis de estas muestras siguiendo los protocolos descritos previamente por Tortola et al (1999a). Brevemente, para **p53** se analizaron los exones del 4 al 8 utilizando la técnica de la SSCP (Single-Strand Conformation Polymorphism) y los casos positivos se secuenciaron para confirmar las mutaciones. Las mutaciones de **k-ras** también se determinaron por SSCP, amplificando un fragmento de 241 pb que comprende los codones 12 y 13 del exón 1. Para evitar la secuenciación de todas las muestras positivas se utilizó DNA de líneas celulares con mutaciones de k-ras bien caracterizadas como controles de mutación. De esta manera cuando el patrón de bandas de la SSCP coincidía con el de una línea celular determinada se identificaba inmediatamente la mutación sin necesidad de secuenciar. Cuando el patrón mutado no coincidía con ninguna línea se llevaba a cabo la secuenciación del fragmento. Los resultados de estos experimentos no se describen en la sección de Resultados ya que no son el objeto de esta tesis. Brevemente indicamos aquí los resultados finales. Para p53: 9 casos positivos, 14 negativos y 4 dudosos (SSCP positiva pero mutación no identificada). Para k-ras: 11 casos positivos, 22 negativos y 1 dudoso (incompatibilidad entre el resultado de SSCP y secuencia). Sumando estos datos a los que ya existían, la distribución de mutaciones de estos dos genes en nuestra muestra de tumores quedó de la siguiente manera. Para p53: 52 casos positivos, 66 negativos y 11 dudosos. Para k-ras: 47 casos positivos, 80 casos negativos y 2 casos dudosos.

5. ANÁLISIS ESTADÍSTICO

5.1 ESTADÍSTICA UNI Y BIVARIABLE

La comparación de dos variables cuantitativas se realizó mediante tablas de contingencia sobre las que se calculó, cuando fue posible, el test de Fisher exacto y, en su defecto, el test de Chi-cuadrado. La comparación de las medias de una variable cuantitativa en los grupos definidos por una variable cualitativa se llevó a cabo con el test de la *t* de Student para grupos independientes (si se trataba de dos grupos) o con análisis de la varianza o ANOVA (si se trataba de más de dos grupos). La existencia de una relación lineal entre dos variables cuantitativas se determinó, previo dibujo del diagrama de puntos, mediante el test de la pendiente de un modelo de regresión lineal, el cual es equivalente al coeficiente de correlación de Pearson. Todos los valores de *P* se calcularon a partir de tests estadísticos de dos colas. El criterio de significación se estableció de forma general para $P \leq 0.05$, pero su aplicación fue flexible, de manera que en los análisis más relevantes valores de *P* superiores a 0.05 e inferiores a 0.10 también se consideraron significativos.

5.2 ANÁLISIS DE LA SUPERVIVENCIA

El análisis de la supervivencia se realizó considerando la supervivencia total, es decir, el tiempo transcurrido desde la cirugía hasta la muerte o hasta la última consulta. Dependiendo del subobjetivo considerado se utilizó la serie completa de tumores o sólo aquellos con resección total, como se detalla a continuación:

- **Valoración de las medidas de daño genómico como factores pronósticos.** Se realizó el análisis de la supervivencia total sólo en los pacientes con tumores R0 ($n=108$). Estos pacientes sufrieron una resección quirúrgica radical (definida por la ausencia de enfermedad residual macroscópica o microscópica y no Dukes D) y, por tanto, constituyen la población de referencia para analizar el valor pronóstico de una variable.

- **Análisis multivariante mediante modelos de Cox.** Se realizó el análisis de la supervivencia total con todos los pacientes (n=129). En este caso las variables se estudian con el objetivo de extraer la máxima información que contienen y de averiguar cómo se influyen mutuamente en la determinación de la agresividad de los tumores, la cual se analiza en función del riesgo de muerte. Esta aproximación pretende un estudio más amplio del papel del daño genómico en la progresión tumoral y, por tanto, todos los casos aportan información útil.

5.2.1 VALORACIÓN DE LAS MEDIDAS DE DAÑO GENÓMICO COMO FACTORES PRONÓSTICOS

La valoración del valor pronóstico de una variable se suele llevar a cabo mediante la **estimación de la función de supervivencia por el método de Kaplan-Meier**. Éste permite la creación de una curva que indica la disminución de la supervivencia a medida que pasa el tiempo. La comparación de las curvas de diferentes grupos se realiza mediante el test log rank. Este método se puede utilizar directamente para evaluar el valor pronóstico de las variables cualitativas. En el caso de las variables cuantitativas también se puede usar si previamente éstas se convierten a cualitativas aplicando un punto de corte. Los dos sistemas más usados para determinar puntos de corte son la mediana, que divide la muestra en dos grupos de igual tamaño, y el análisis de las curvas ROC (Receiver Operating Characteristic). Estas curvas proporcionan una imagen gráfica que ayuda a escoger el punto de corte óptimo, que maximiza la suma de la sensibilidad y la especificidad (Domenech, 1993). En nuestro caso **se obtuvo el punto de corte óptimo mediante curvas ROC** para todas las variables cuantitativas incluidas en el estudio usando el programa Medcalc versión 3.0 (Medcalc software, Belgium). El uso de puntos de corte para evaluar el valor pronóstico de variables cuantitativas no es muy recomendable desde el punto de vista estadístico debido a que supone pérdida de información de la variable (ver apartado 2.1 de Discusión). Sin embargo, desde el punto de vista clínico resulta muy conveniente, ya que permite dividir los pacientes en grupos de alto y bajo riesgo. Por esta razón y por la simplicidad de su cálculo, su uso es muy frecuente.

5.2.2 ANÁLISIS MULTIVARIANTE MEDIANTE MODELOS DE COX

Con el doble objetivo de obtener la máxima información incluida en las variables de cuantificación del daño genómico y de explorar como se influyen mutuamente las variables que se asocian a la malignidad de los tumores, se decidió llevar a cabo un estudio más completo y exhaustivo de la supervivencia basado en la utilización de los **modelos de Cox** e incluyendo todos los casos del estudio. A continuación se exponen brevemente los aspectos específicos relacionados con la aplicación de estos modelos a nuestro estudio y en el Apéndice 1 se proporciona una explicación más detallada de su funcionamiento.

En primer lugar hay que señalar que **el modelo de Cox permite analizar la supervivencia**, expresada como 'riesgo' o probabilidad de morir en un tiempo determinado (el riesgo relativo se indica como HR (Hazard Ratio)), **en función de variables tanto cualitativas como cuantitativas**. Esto es especialmente útil para este estudio, ya que las dos variables principales a analizar son cuantitativas (GDF y AI). Además el modelo de Cox ofrece la posibilidad de estudiar el tipo de relación existente (lineal o no) entre la variable y el riesgo. Tanto para el GDF como para el AI se llevaron a cabo tests de desviación de la linealidad usando modelos con el término cuadrático y se representaron gráficamente las funciones de riesgo (evolución del riesgo en función del aumento de la variable cuantitativa). Todos los gráficos se realizaron utilizando métodos de suavizado del tipo *splines* que permitían la curvatura de la función (para más información ver Apéndice 1). Para las variables cualitativas la significación que se obtiene en los modelos de Cox es muy parecida a la que proporciona el test log rank.

El modelo de Cox permite estudiar el efecto que tiene una variable (análisis univariante) o un conjunto de variables (**análisis multivariante**) sobre el riesgo de muerte. En nuestro caso se priorizó la elaboración de modelos sencillos y fácilmente interpretables, por lo que los modelos multivariantes nunca incluyeron más de dos variables. Todos los modelos de dos variables se compararon con sus respectivos modelos de una variable mediante el test de razón de verosimilitud para ver si el efecto sumatorio de las dos variables mejoraba la predicción del riesgo obtenida en los modelos univariados. De esta **manera el modelo de Cox permite determinar si existe una relación de independencia en cuanto a la asociación con supervivencia de las dos variables**. Dicho de otra manera, **proporciona información acerca del valor pronóstico neto** de cada una de las variables incluidas en el modelo. Además, el modelo de Cox **también permite la detección de interacciones entre variables, es decir, de modificaciones**

en el riesgo que produce una variable en función de las categorías de otra. Para detectarlas todos los modelos de dos variables se compararon con sus respectivos modelos de interacción por el test de razón de verosimilitud. Cuando éste resultó significativo se realizaron los gráficos de la función de riesgo de las variables cuantitativas (GDF y AI) en función de las categorías de la segunda variable con el fin de simplificar la interpretación de la interacción. **Es muy importante determinar estas interacciones porque permiten descubrir la existencia de factores pronósticos que afectan sólo a un grupo de casos y cuya significación queda diluida al considerar todos los casos conjuntamente.**

Es fundamental tener en cuenta que **los modelos de Cox no admiten valores perdidos (*missings*)**. En nuestro caso existían *missings* en las variables derivadas de AP-PCR (GDF, GF y LF) y en p53 y k-ras. En la actualidad se recomienda imputar los *missings* si éstos representan menos del 10% de los casos. Puesto que así era, se decidió imputarlos y para ello se usaron distribuciones aleatorias que mantenían las características básicas de la distribución original. Así, para las variables cuantitativas (GDF, GF y LF) se usó una distribución normal con la media y la desviación estándar de los casos analizados y para las variables cualitativas (p53 y k-ras) se usó una distribución binomial de frecuencia calculada con los casos disponibles.

La estadística se llevó a cabo con el software SPSS (Statistical Package for the Social Sciences) versión 6.1.3 (1995) y el paquete estadístico R, versión 0.90.0 (Noviembre 1999).