



UNIVERSITAT DE BARCELONA

U

B

**UNIVERSIDAD DE BARCELONA**

**FACULTAD DE FARMACIA**

**DEPARTAMENTO  
BIOQUIMICA Y BIOLOGIA MOLECULAR**

**PREDICCIÓN DE RESPUESTA AL  
TRATAMIENTO EN PACIENTES CON  
CARCINOMA ESCAMOSO DE CABEZA Y  
CUELLO**

**MIGUEL ANGEL PAVÓN RIBAS  
2009**

## **II. HIPÓTESIS PRELIMINARES Y** **OBJETIVO GENERAL**



### **II HIPÓTESIS PRELIMINARES Y OBJETIVO GENERAL**

Los protocolos de tratamiento del carcinoma de cabeza y cuello de estadio localmente avanzado que permiten preservar la función del órgano, como son la quimioterapia de inducción seguida de RT/QRT o cirugía, o la quimioradioterapia concomitante, han ido desplazando progresivamente de la práctica clínica a la cirugía radical como único tratamiento desde el inicio.<sup>17</sup>

El objetivo general de este proyecto de tesis es identificar, en biopsias de CECC, marcadores moleculares que posibiliten establecer, con anterioridad al inicio del tratamiento genotóxico, la evolución clínica del paciente. Estos marcadores permitirían identificar qué pacientes podrían beneficiarse clínicamente del tratamiento genotóxico y distinguirlos de los pacientes que no obtendrían beneficio clínico por dicho tratamiento y que serían, por tanto, candidatos a un tratamiento quirúrgico o bien podrían ser tratados con una terapia alternativa.

Este trabajo incluye dos estudios independientes, ambos dirigidos a definir nuevos marcadores moleculares capaces de predecir la evolución clínica de los pacientes con CECC de estadio avanzado, tratados con quimioterapia de inducción, seguida de RT/QRT o cirugía, o quimioradioterapia concomitante.

En el primer estudio se evalúan los niveles de expresión de los genes Ku70, Ku80 y DNA-PKcs, del sistema de reparación del DNA no homóloga por unión de extremos (NHEJ), como posibles predictores de la evolución clínica en pacientes con CECC localmente avanzado tratados con quimioterapia de inducción seguida de RT/QRT o cirugía. Esta hipótesis se basa en las siguientes premisas:

- La elevada frecuencia de alteraciones cromosómicas que presenta el carcinoma de cabeza y cuello sugiere que los mecanismos encargados de mantener la estabilidad genómica se encuentran alterados.<sup>142</sup>
- Los mecanismos de reparación de roturas de doble cadena (RDC) se encuentran alterados en el carcinoma escamoso de cabeza y cuello.<sup>143, 144</sup>
- La citotoxicidad de los antineoplásicos clásicos (5-Fluorouracilo, cisplatino y radioterapia) depende de su capacidad de producir roturas en el DNA de las células tumorales. Todos estos agentes antitumorales producen de manera directa o indirecta RDC en el DNA.<sup>59, 63, 145-147</sup>

## **Hipótesis preliminares y objetivo general**

- En mamíferos, el sistema de reparación no homóloga por unión de extremos (NHEJ) juega un papel predominante en la reparación de RDC.<sup>71</sup>
- El sistema de reparación no homóloga por unión de extremos (NHEJ) suprime la aparición de translocaciones cromosómicas y juega un papel importante en el mantenimiento de la estabilidad genómica.<sup>88, 148</sup>
- En estudios “in vitro” la inactivación del sistema de reparación NHEJ aumenta la sensibilidad de las células tumorales a la radiación ionizante.<sup>149</sup>
- Los genes del complejo DNA-PK (DNA-PKcs, Ku70 y Ku80) participan en el proceso de señalización apoptótica inducida por roturas de doble cadena al DNA producidas por la radioterapia y la quimioterapia.<sup>99, 150, 151</sup>

En el segundo estudio, se utilizan microarrays de expresión para identificar los genes y procesos biológicos implicados en la evolución clínica de los pacientes tratados con quimioterapia de inducción o quimioradioterapia concomitante. Su finalidad es definir un perfil de expresión génico, en biopsias pre-tratamiento, que permita identificar aquellos pacientes que obtienen un beneficio clínico del tratamiento con quimioterapia de inducción, seguida de RT/QRT, o con quimioradioterapia concomitante y distinguirlos de los pacientes a los que el tratamiento genotóxico no les reporta ningún beneficio clínico. Esta es un área escasamente explorada, ya que la mayoría de estudios de microarrays descritos en CECC identifican marcadores pronóstico o genes implicados en el proceso de patogénesis del CECC. Además, en un mismo estudio se incluyen pacientes con tumores en diferentes estadios que, en consecuencia, reciben protocolos terapéuticos distintos. Los estudios dirigidos a localizar genes predictores de la evolución clínica de los pacientes sometidos a tratamiento genotóxico son escasos. En el presente estudio de microarrays de expresión, incluimos exclusivamente pacientes con CECC de estadio localmente avanzado tratados con quimioterapia de inducción, seguida de RT/QRT o cirugía, o quimioradioterapia concomitante. Su propósito es, por tanto, identificar marcadores que permitan predecir la evolución clínica de los pacientes con CECC tratados con 5-Fluorouracilo, cisplatino y radioterapia. En este estudio queremos obtener una visión global de las vías de señalización y de los procesos biológicos implicados en la respuesta tumoral al tratamiento y de la evolución clínica del paciente sin partir de ninguna hipótesis previamente definida. Finalmente, y aunque no es un objetivo prioritario del estudio, evaluaremos la expresión diferencial entre tumores y

## **Hipótesis preliminares y objetivo general**

mucosas normales, para identificar aquellos genes y procesos biológicos implicados en el proceso de patogénesis del carcinoma escamoso de cabeza y cuello.



### **III. OBJETIVOS**





### III OBJETIVOS

#### **ESTUDIO DEL NIVEL DE EXPRESIÓN DE LOS GENES DEL SISTEMA NHEJ EN BIOPSIAS PRE-TRATAMIENTO DE PACIENTES CON CARCINOMA ESCAMOSO DE CABEZA Y CUELLO LOCALMENTE AVANZADO TRATADOS CON QUIMIOTERAPIA DE INDUCCIÓN SEGUIDA DE RT/QRT O CIRUGÍA:**

- Analizar si los niveles de mRNA y/o proteína del complejo DNA-PK en biopsias tumorales pre-tratamiento de estos pacientes predicen la respuesta tumoral a la quimioterapia de inducción.
- Analizar si los niveles de mRNA y/o proteína del complejo DNA-PK en biopsias tumorales pre-tratamiento están asociados con la supervivencia libre de recidiva local y/o con la supervivencia global de estos pacientes.
- Analizar si los niveles de mRNA y/o proteína del complejo DNA-PK en biopsias tumorales pre-tratamiento pueden identificar aquellos pacientes que obtendrán un beneficio clínico de la terapia genotóxica.

#### **ESTUDIO DE EXPRESIÓN CON MICROARRAYS EN BIOPSIAS PRE-TRATAMIENTO DE PACIENTES CON CARCINOMA ESCAMOSO DE CABEZA Y CUELLO LOCALMENTE AVANZADOS TRATADOS CON QUIMIOTERAPIA DE INDUCCIÓN, SEGUIDA DE RT/QRT O CIRUGÍA, O QUIMIORADIOTERAPIA CONCOMITANTE:**

- Identificar posibles subtipos tumorales relacionados con diferencias en la respuesta tumoral al tratamiento genotóxico, la supervivencia libre de recidiva local y la supervivencia global a partir del perfil de expresión de biopsias pre-tratamiento de estos pacientes.
- Identificar genes asociados con supervivencia libre de recidiva local y/o supervivencia global en estos pacientes.
- Identificar genes asociados con la respuesta tumoral tras el tratamiento con quimioterapia de inducción seguida de RT/QRT o cirugía.

## **Objetivos**

---

- Determinar los procesos biológicos implicados en la respuesta tumoral, en la recidiva tumoral y en la supervivencia global de los pacientes con carcinoma escamoso de cabeza y cuello localmente avanzados tras el tratamiento genotóxico.
- Determinar los genes y procesos biológicos implicados en el proceso de patogénesis del carcinoma escamoso de cabeza y cuello, analizando las diferencias entre los perfiles de expresión de los tumores y las mucosas normales.

## **IV. MÉTODOS**



## **IV MÉTODOS**

### **1. CARACTERÍSTICAS DE LOS PACIENTES, RÉGIMEN DE TRATAMIENTO, EVALUACIÓN DE LA RESPUESTA TUMORAL, SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL (SLRL), SUPERVIVENCIA GLOBAL (SG) Y BENEFICIO CLÍNICO**

#### **1.1. CARACTERÍSTICAS CLÍNICAS DE LOS PACIENTES**

El estudio de los niveles de expresión de los genes Ku70, Ku80 y DNA-PKcs se llevó a cabo en dos cohortes independientes de pacientes con carcinoma escamoso de cabeza y cuello (CECC) de estadios localmente avanzados tratados en el “Hospital de la Santa Creu i Sant Pau”. En la primera parte del estudio analizamos los niveles de expresión de Ku70, Ku80 y DNA-PKcs por PCR cuantitativa en un grupo de pacientes seleccionados de modo prospectivo a partir del año 2002. La segunda parte del estudio, en la que se determinó la expresión de las proteínas Ku70 y Ku80, se realizó en una cohorte de pacientes seleccionados de modo retrospectivo que fueron tratados en nuestro centro durante el periodo comprendido entre 1995 y 2003. Los pacientes incluidos en ambos estudios presentaban un diagnóstico confirmado mediante análisis histopatológico, no habían recibido ningún tratamiento con anterioridad y eran candidatos a un protocolo de quimioterapia de inducción seguida de RT/QRT o cirugía.

El estudio de expresión de microarrays se realizó en una cohorte de pacientes con CECC en estadios avanzados seleccionados de modo prospectivo a partir del año 2004. Los pacientes incluidos en este estudio presentaban un diagnóstico confirmado mediante análisis histopatológico y no habían recibido ningún tratamiento con anterioridad. Todos los pacientes incluidos en el estudio eran candidatos a recibir quimioterapia de inducción, seguida de RT/QRT o cirugía, o quimioradioterapia concomitante desde un inicio.

Ambos estudios fueron aprobados por el Comité Ético del Hospital de la Santa Creu i Sant Pau.

#### **1.2. RÉGIMEN DE TRATAMIENTO**

Los pacientes que en el momento del diagnóstico se definen como operables son candidatos a recibir un protocolo de inducción que consiste en la administración de tres ciclos de quimioterapia, seguida de cirugía o de un tratamiento conservador según la respuesta que presente el tumor. Los pacientes con tumores que responden a la

## Métodos

quimioterapia de inducción continúan su tratamiento con un protocolo conservador que consiste en la administración de radioterapia hiperfraccionada o de quimiorradioterapia concomitante (Figura 14). Dicho tratamiento permite preservar la función del órgano. La cirugía radical se utiliza en aquellos pacientes con tumores que no responden a la quimioterapia de inducción (Figura 14). Los pacientes que desde el diagnóstico se definen como no operables son candidatos a recibir quimiorradioterapia concomitante desde un inicio (Figura 14).

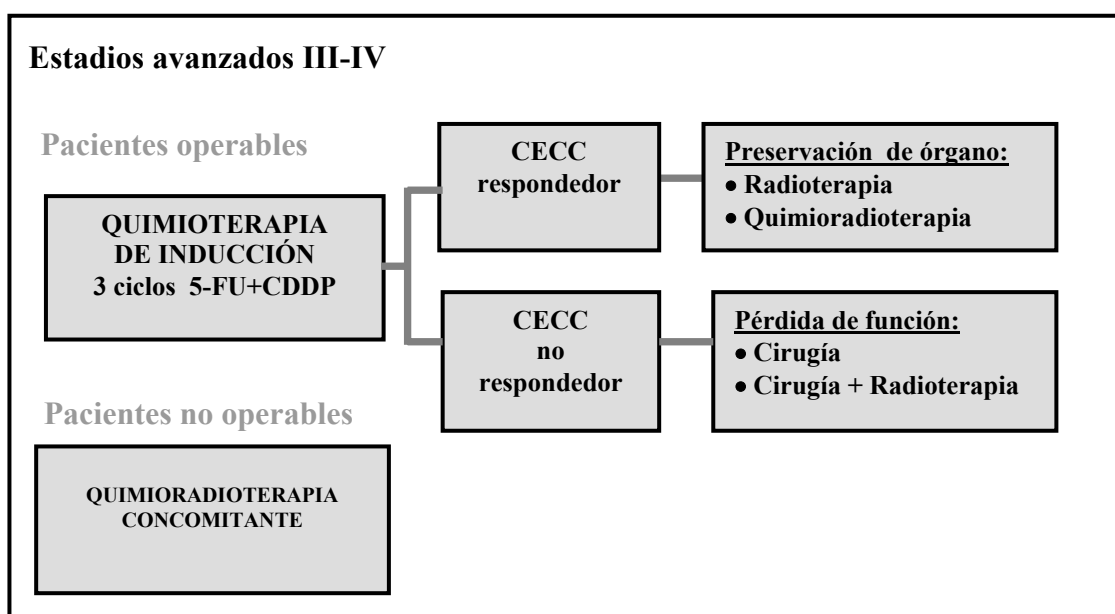


Figura 14. Protocolo de tratamiento utilizado en el Hospital de la Santa Creu i Sant Pau en pacientes con carcinoma escamoso de cabeza y cuello en estadio localmente avanzado.

El tratamiento de quimioterapia de inducción consistió en la administración de 3 ciclos de quimioterapia combinada con 5-Fluoruracilo (5-FU) y cisplatino (CDDP). En cada ciclo de quimioterapia de una semana de duración, se administró CDDP por vía intravenosa el día 1, a una dosis única de  $100 \text{ mg/m}^2$ , y 5-FU a una dosis de  $1000 \text{ mg/m}^2/\text{día}$ , por infusión intravenosa continua de 24 horas desde el día 2 hasta el día 6. El tratamiento se repitió cada tres semanas hasta completar tres ciclos. Los pacientes tratados entre 1995 y 2002, cuyos tumores respondieron a la quimioterapia de inducción, siguieron un protocolo de tratamiento con radioterapia (RT). En 2003, se incorporó la quimiorradioterapia (QRT) como alternativa a la RT en pacientes respondedores. Progresivamente, la QRT fue desplazando a la RT como tratamiento de elección para estos mismos pacientes.

La radioterapia se administró localmente, al tumor primario y a aquellos ganglios regionales positivos, en 35 fracciones de 2 Gy durante una semana hasta alcanzar una dosis total de 70 Gy. Las áreas ganglionares no afectadas recibieron RT hasta una dosis total de 50 Gy.

La quimioradioterapia concomitante, administrada ya sea desde un inicio a los pacientes no operables, o bien, una vez finalizada la quimioterapia de inducción, consistió en la administración de RT a las dosis descritas anteriormente, seguidas de tres ciclos de cisplatino a una dosis de 100 mg/m<sup>2</sup> cada tres semanas.

### **1.3. EVALUACIÓN DE LA RESPUESTA TUMORAL, SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL (SLRL), SUPERVIVENCIA GLOBAL (SG) Y BENEFICIO CLÍNICO**

Para determinar la capacidad de los marcadores evaluados de predecir la evolución clínica de los pacientes tratados con quimioterapia de inducción, seguida de RT/QRT o cirugía, o quimioradioterapia, analizamos la respuesta tumoral inmediata una vez finalizada la quimioterapia de inducción (a los 3 meses de su inicio), la supervivencia libre de recidiva local (SLRL) y la SG (supervivencia global). La respuesta tumoral inmediata sólo pudo ser evaluada en aquellos pacientes tratados con quimioterapia de inducción, ya que la quimioradioterapia concomitante induce una respuesta tumoral completa en la mayoría de pacientes. A continuación, se definen los tres parámetros utilizados para identificar marcadores capaces de identificar a los pacientes con mayor probabilidad de obtener beneficio clínico del tratamiento con quimioterapia de inducción o quimioradioterapia:

La respuesta tumoral a la QTI se evaluó comparando el volumen tumoral antes de iniciar el tratamiento y una vez finalizado el tercer ciclo. La respuesta se definió como una reducción del volumen del tumor primario valorada por examen físico, endoscopia, tomografía computerizada y/o resonancia magnética nuclear, siguiendo los criterios RECIST.<sup>152</sup> Identificamos dos grupos de pacientes según la respuesta. El grupo de respondedores incluyó aquellos pacientes cuyos tumores presentaron una respuesta completa o parcial superior al 50%. El grupo de no respondedores incluyó todos aquellos pacientes con enfermedad estable (<50% de reducción del volumen tumoral) o progresiva (>25% de incremento del volumen tumoral).

Se realizó un seguimiento de todos los pacientes a fin de obtener su supervivencia libre de recidiva local (SLRL) y su supervivencia global (SG). La SLRL se definió como el tiempo transcurrido desde el diagnóstico hasta la recidiva del tumor primario o hasta el



último punto de control del paciente. La supervivencia global (SG) se definió como el tiempo transcurrido desde el diagnóstico hasta la muerte del paciente, o hasta el último punto de control del paciente. La mediana de seguimiento en el estudio de los niveles de expresión de los genes del sistema de reparación NHEJ fue de 2 años y 1 mes (25 meses) para el estudio prospectivo y de 4 años (48 meses) para el retrospectivo. En el estudio de expresión por microarrays la mediana de seguimiento fue de 1 año y 7 meses (19 meses).

## **2. RECOGIDA DE BIOPSIAS DE TEJIDO FRESCO, EXTRACCIÓN DE RNA Y SÍNTESIS CDNA**

### **2.1 RECOGIDA DE BIOPSIAS DE TEJIDO FRESCO**

Se obtuvieron biopsias tumorales, mediante endoscopia, de pacientes con CECC candidatos a quimioterapia de inducción o quimiorradioterapia concomitante antes de que iniciaran tratamiento alguno. Para asegurar al máximo la integridad del RNA, las biopsias se congelaron inmediatamente después de su extracción, sumergiéndolas en un baño de isopentano frío y posteriormente se almacenaron en nitrógeno líquido hasta el momento de ser procesadas. Una parte de la biopsia se utilizó para confirmar, mediante diagnóstico histopatológico, la presencia de células tumorales, excluyendo del estudio aquellas muestras que no presentaban un elevado porcentaje de tejido tumoral. Las biopsias congeladas fueron utilizadas para obtener RNA para los experimentos de PCR y microarrays. Las biopsias de mucosa normal utilizadas en el estudio de microarrays se obtuvieron de localizaciones alejadas del tumor primario, donde no se apreciaba clínicamente ninguna alteración. Mediante análisis histopatológico se confirmó que las 5 mucosas sólo contenían tejido normal.

### **2.2 EXTRACCIÓN DE RNA Y SÍNTESIS CDNA**

La extracción de RNA de tejido tumoral se realizó con Trizol reagent, siguiendo las indicaciones del fabricante (Invitrogen, Paisley, Reino Unido).

El primer paso de la extracción consiste en la disgregación del tejido y lisis de las células utilizando Trizol reagent. Seguidamente, mediante una extracción con fenol-cloroformo, se separa el RNA en la fase acuosa. Posteriormente se realizan 2 pasos de precipitación y lavado del RNA en isopropanol y etanol al 75%. Una vez finalizada la extracción con Trizol, el RNA se limpió mediante el sistema de purificación por columnas RNAeasy (Qiagen, Valencia, USA) con objeto de asegurar un mayor grado de

pureza del RNA. El método de extracción en columnas se basa en la retención del RNA en una columna de afinidad y tras una serie de pasos de limpieza, termina con la elución del RNA en H<sub>2</sub>O libre de RNAasas.

### **Extracción de RNA con TRIZOL**

- 1- Homogenizar 15-20 µg de tejido en 1ml de TRIZOL Reagent.
- 2- Centrifugar 2 min a 7.000 g a 4°C. Pasar el sobrenadante a un tubo nuevo.
- 3- Añadir 200 µl de cloroformo. Homogeneizar por inversión durante 15 s.
- 4- Centrifugar 15 min a 12.000 g a 4°C. Recuperar la fase acuosa (400-450 µl)
- 5- Añadir 450 µl de isopropanol frío. Homogeneizar por inversión durante 15 s.
- 6- Incubar durante 10 min a temperatura ambiente (T.A.)
- 7- Centrifugar 10 min a 12.000 g a 4°C. Eliminar el sobrenadante.
- 10- Añadir 1 ml de etanol 75% frío. Agitar suavemente.
- 11- Centrifugar 5 min a 7.500 g a 4°C. Eliminar el sobrenadante.
- 13- Dejar secar el pellet a T.A. durante 10 minutos.
- 14- Resuspender el RNA en 100µl de H<sub>2</sub>O libre de RNAasas.

### **Purificación de RNA con columnas RNAeasy**

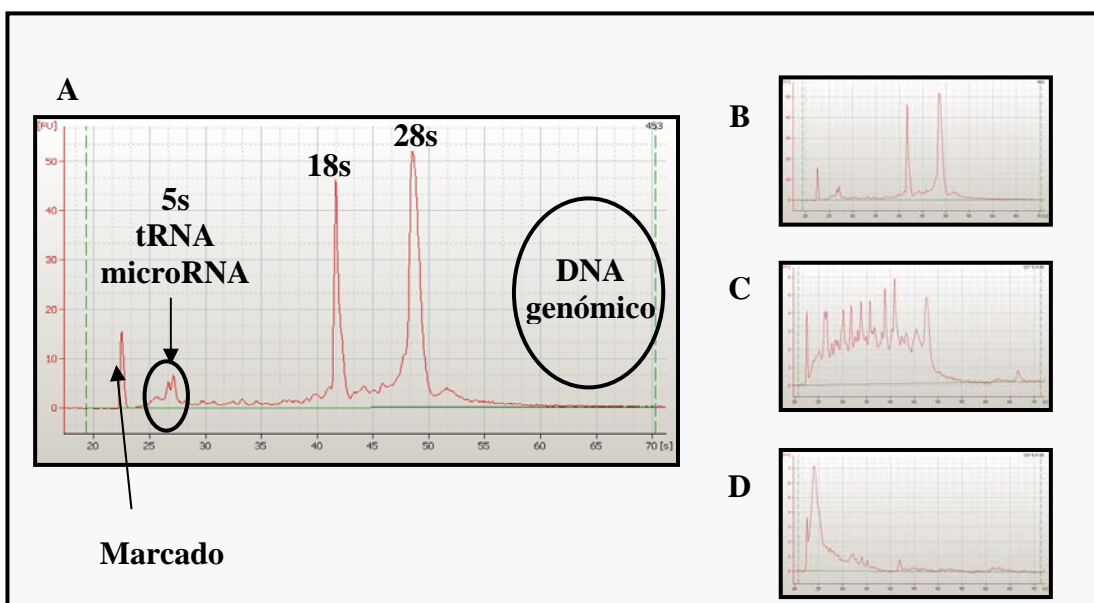
- 1- Añadir 350 µl de buffer RLT al RNA obtenido. Homogeneizar por inversión.
- 2- Añadir 250 µl de etanol 100%. Homogeneizar por inversión durante 15s.
- 3- Pasar la solución (700 µl) a una columna RNAeasy y centrifugar durante 15s a 8.000 g a temperatura ambiente (T.A.)
- 4- Pasar la columna a un tubo nuevo de recogida y añadir 500 µl de buffer RPE. Centrifugar 15s a 8.000 g a T.A.
- 5- Eliminar el eluido y añadir 500 µl de buffer de RPE. Centrifugar durante 2 min a 8.000g, a T.A.
- 6- Eliminar el eluido y centrifugar durante 1min a 12.000 g.
- 7- Pasar la columna a un tubo nuevo de recogida. Añadir 40 µl de H<sub>2</sub>O destilada libre de RNAasas.
- 8- Incubar durante 3min a T.A. y centrifugar 3 min. a 10.000g para eluir el RNA.

La concentración y la pureza del RNA total obtenido se determinaron por medición de la absorbancia a 260 nm y 280 nm. Una unidad de A260 es equivalente a 40 µg/ml de RNA. El ratio A260/A280 nos indica la pureza del RNA. Un ratio de 2 indica una pureza de RNA elevada. Cuando el RNA está contaminado con proteínas este ratio disminuye. Para los estudios de PCR o microarrays utilizamos RNA con un ratio superior a 1,6 o a 1,9, respectivamente.

Para determinar la integridad del RNA y descartar posibles contaminaciones, se utilizó el analizador “Bionalyser” y los chips de RNA de Agilent (Agilent Technologies, Santa Clara, USA). El uso del “Bioanalyser” permite obtener un perfil para las diferentes secuencias de RNA/DNA que conforman una misma muestra separadas por peso

## Métodos

molecular (Figura 15). En una muestra de RNA no degradado, se observa un primer pico que corresponde al marcador, seguido de dos picos muy aparentes que corresponden a las subunidades mayores del RNA ribosómico (18s y 28s). El ratio entre el rRNA 28s y el 18s se aproxima a 2 cuando el RNA no presenta ningún tipo de degradación. A medida que el RNA se degrada el ratio disminuye. Se trata de una medida indirecta, que permite establecer la integridad del mRNA midiendo la integridad del RNA ribosómico que es mucho más abundante que el RNA mensajero. Para el análisis de microarrays utilizamos RNA con ratios  $> 1.8$ .



**Figura 15.** Análisis de la integridad del RNA con el “Bioanalyser”. A) Espectrofenograma obtenido al cuantificar RNA total en el Bioanalyser. Con sólo 1-2  $\mu$ l de RNA purificado podemos evaluar la integridad del RNA y la posible presencia de contaminación por DNA genómico. El primer pico del espectrofenograma corresponde al marcador. Detrás del marcador aparece un pico, que es más o menos aparente según el método de purificación de RNA que se utilice y, que corresponde a los RNAs de bajo peso molecular (RNAr 5s; RNAt; microRNA). Los dos picos que muestran mayor intensidad corresponden a los RNA ribosómicos 18s y 28s. En un RNA integro el ratio 28s/18s debe acercarse a 2. La aparición de alguna señal a un tiempo de elución mayor que el del RNAr 28s es un indicador de contaminación con DNA genómico. En los espectrofenogramas de la derecha podemos observar las diferencias entre un RNA integro (B), un RNA parcialmente degradado (C) y un RNA completamente degradado (D).

La síntesis de cDNA se realizó mediante una reacción de retrotranscripción inversa, utilizando el “High Capacity cDNA Archive Kit” (AppliedBiosystems, Foster City, USA) y 1,5 $\mu$ g de RNA total. Los componentes de la reacción incluyeron 5 $\mu$ l de 10X “RT Buffer”, 2  $\mu$ l de 25X de dNTPs, 5  $\mu$ l de 10X “Random Hexamers Primers”, 125 U de Retrotranscriptasa Reversa “Multiscribe” y 40 U de “RNase Inhibitor”, en un

volumen final de reacción de 50  $\mu$ l. Las condiciones de incubación de la reacción fueron de 10 min a 25°C, 2 horas a 37°C y 3 min. a 95°C.

### 3. DETERMINACIÓN DE LOS NIVELES DE EXPRESIÓN DE KU70, KU80 Y DNA-PKCS POR PCR CUANTITATIVA

Los niveles de expresión de mRNA de los genes del sistema de reparación no homóloga por unión de extremos (NHEJ) fueron analizados mediante PCR cuantitativa a tiempo real, utilizando sondas Taqman (Applied Biosystems, Foster City, USA). La tecnología Taqman se basa en la utilización de una sonda y un juego de dos primers específicos de la región que deseamos analizar. La sonda Taqman está formada por una secuencia de nucleótidos a la que se encuentra unida una molécula de fluorocromo y un amortiguador de fluorescencia. En presencia del amortiguador, el fluorocromo no puede emitir señal. Al inicio de la reacción de PCR, los cebadores y la sonda hibridan específicamente con la región a amplificar. La enzima Taq polimerasa empieza a amplificar el DNA a partir de los dos primers y cuando se encuentra con la sonda Taqman la degrada, utilizando su actividad exonucleasa, liberando de este modo el fluorocromo. Una vez separado del amortiguador, el fluorocromo emite fluorescencia (Figura 16).

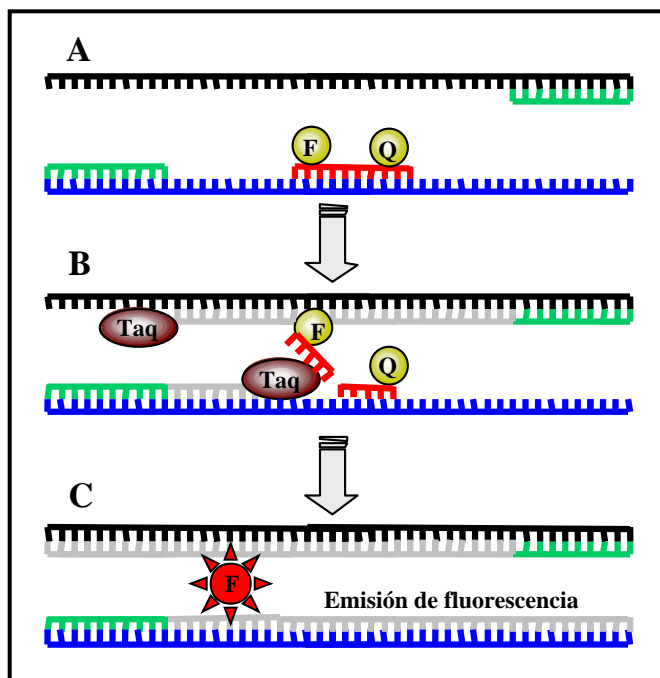


Figura 16. Reacción de la PCR cuantitativa. A) Los primers y la sonda hibridan específicamente con la secuencia a amplificar. B) La DNA polimerasa (Taq) empieza a sintetizar DNA a partir de los primers y cuando encuentra la sonda la degrada, utilizando su actividad exonucleasa, liberando de este modo el fluorocromo (F). C) Una vez separado del amortiguador (Q) el fluorocromo (F) emite fluorescencia que es detectada por un detector de fluorescencia.

La detección de la señal de fluorescencia en cada ciclo de PCR (tiempo real) permite conocer la cantidad de mRNA inicial que había en la reacción. Para realizar su análisis, se fija un umbral de fluorescencia (CT; Cycle Threshold), de modo que una muestra que

## Métodos

alcance el umbral en ciclos iniciales (CT bajos) contiene una cantidad de mRNA mayor que una muestra que alcanza el umbral de fluorescencia en ciclos mas tardíos (CT altos).

### 3.1 PRIMERS Y SONDAS

Para la amplificación de los genes Ku80, Ku70 y  $\beta$ -actina, se utilizaron juegos prediseñados de primers y sondas Taqman, del programa “Assay on Demand” de AppliedBiosystems. Todas las sondas estaban marcadas con el fluorocromo FAM, exepcto la sonda para  $\beta$ -actina que estaba marcada con el fluorocromo VIC. Para la detección de DNA-PKcs se diseñó un juego de dos primers y una sonda Taqman, utilizando el programa “Primer Express” (AppliedBiosystems, Foster City, USA) Primer directo: 5'-TGGGAGCATCACTTGCCTTTAATAA-3'; Primer reverso: 5'-CAAAGTGTCCACAGAGACTCTT-3'; Sonda TaqmanMGB: 5'-CTTCCCTGAATTCCC-3'. La mayoría de sondas fueron escogidas en regiones inter-exónicas, con objeto de descartar señales debidas a la amplificación de DNA genómico. Para comprobar que todos los juegos de primers y sondas presentaban la misma eficiencia de amplificación, se realizaron curvas estándar, a partir de diluciones seriadas del RNA de la línea UM-SCC-22A, y se comprobó, para cada uno de los genes analizados, que las curvas resultantes presentaban la misma pendiente con un valor cercano a -3,32. La Tabla 4 describe el área del gen correspondiente para cada sonda Taqman.

Tabla 4. Descripción de las regiones de amplificación e hibridación de las sonda para el análisis de los niveles de expresión de mRNA por PCR cuantitativa.

Gen	Ensayo <sup>a</sup>	RefSeq <sup>b</sup>	Región Inter-exónica	Tamaño del amplicon (bp)
Ku70	Hs00750856_s1	NM_001469	Exon 1	99
Ku80	Hs00221707_m1	NM_021141	Exon 2 / Exon 3	72
DNA-PKcs	(*)	NM_006904	Exon 27 / Exon 28	72
$\beta$ -actin	Hs99999903_m1	NM_00101	Exon 1	171

<sup>a</sup> Referencia del ensayo “Pre-designed Taqman Gene Expression Assay” utilizado (<http://www.appliedbiosystems.com>)

<sup>b</sup> Secuencia de mRNA de referencia detectado. Disponible en la base de datos RefSeq (NCBI reference Sequence Database).

\*Secuencia diseñada para la detección de DNA-PKcs.

### 3.2 ANÁLISIS POR PCR CUANTITATIVA

El análisis por PCR cuantitativa se llevó a cabo en un termociclador ABI PRISM 7000 (Applied Biosystems, Foster City, USA). Las reacciones se realizaron, por duplicado, en un volumen final de reacción de 20  $\mu$ l, compuesto por 10  $\mu$ l de 2X “Universal Taqman Master Mix”, 1  $\mu$ l de mezcla de primers y sonda, 7  $\mu$ l de H<sub>2</sub>O y 2  $\mu$ l de cDNA. Las condiciones para la reacción de amplificación fueron 2 min a 50°C, 10 min a 95°C para la desnaturalización inicial, y 40 ciclos, cada uno de 15 segundos a 95°C y 1 min. a 60°C.

La expresión génica se calculó aplicando la fórmula  $2^{-\Delta\Delta CT}$  como se ha descrito anteriormente.<sup>153</sup> El resultado de la expresión de un gen problema (Ku70, Ku80 ó DNA-PKcs) se expresa como cambio en el nivel de expresión, respecto al nivel de expresión del mismo gen en una muestra calibradora. Además, para una interpretación correcta de los ensayos de PCR cuantitativa es necesario utilizar un control endógeno que normalice los datos y corrija las diferencias ocasionadas por variaciones en la eficiencia de la reacción de retrotranscripción, la extracción de RNA o la reacción de amplificación. La expresión del control endógeno debe permanecer invariable, independientemente de las diferentes condiciones experimentales o de las diferentes muestras que conforman un mismo experimento. En este estudio se utilizó la  $\beta$ -actina como gen endógeno, gen que codifica para una proteína del citoesqueleto y que se utiliza comúnmente como endógeno en experimentos de PCR cuantitativa.

Como muestra calibradora, se utilizó el RNA extraído de la línea celular UM-SCC-22A de carcinoma escamoso de cabeza y cuello. Esta línea celular fue cultivada en atmósfera húmeda, con 5% de CO<sub>2</sub> y en medio DMEM suplementado con un 10% de FBS, 2mmol/L de glutamina, 50U/ml de penicilina y 50U/ml de estreptomycin.

Para normalizar los datos, calculamos la diferencia entre la expresión del gen problema y la del gen endógeno ( $\Delta CT = CT_{\text{problema (Ku70, Ku80 o DNA-PKcs)}} - CT_{\text{endógeno (\beta-actina)}}$ ) y posteriormente la diferencia entre los niveles de expresión del tumor y los niveles de expresión de la muestra calibradora ( $\Delta\Delta CT = \Delta CT_{\text{Tumor}} - \Delta CT_{\text{UM-SCC22A}}$ ). De este modo, los niveles de expresión de un gen en cada muestra se presentan como aumento o disminución respecto a la expresión del gen de estudio en la línea UM-SCC22A.

### **4. INMUNOHISTOQUÍMICA**

Las técnicas de inmunohistoquímica se basan en la utilización de un anticuerpo específico para la detección de una proteína, en un corte de tejido o muestra biológica, para poder detectar su expresión. Tras la reacción con el anticuerpo primario, se utiliza un anticuerpo secundario ligado a la enzima peroxidasa. Posteriormente se incuba la muestra con una solución que contiene DAB (3,3'-diminobenzidina), la cual se oxida por acción de la peroxidasa. La oxidación de la DAB genera un precipitado de color marrón que se deposita en las zonas positivas para el anticuerpo primario.

La inmunohistoquímica (IHQ) es una técnica utilizada de manera habitual en el diagnóstico histopatológico, que permite analizar el nivel de expresión, la localización celular y el tipo celular que expresa una determinada proteína en un corte histológico, tanto en tejido fresco como parafinado.

El método habitual de cuantificación de la tinción inmunohistoquímica consiste en determinar el porcentaje de células positivas presentes, mediante el conteo de células en diversos campos, y la intensidad de tinción ( + poco teñido; ++ tenido, +++ muy teñido). Este método resulta bastante subjetivo y el resultado final puede variar según la experiencia y criterio del anatomopatólogo que realice el análisis. Por esta razón, desarrollamos un método morfométrico a fin de disminuir la variabilidad en la cuantificación de las tinciones IHQs que utilizamos para determinar los niveles de expresión de las proteínas Ku70 y Ku80.

#### **4.1. PROTOCOLO DE TINCIÓN**

Para el estudio inmunohistoquímico se utilizaron cortes de biopsias tumorales de 4 micras conservadas en parafina y previamente fijadas en formalina. Las muestras que no presentaban un elevado contenido de tejido tumoral se excluyeron del análisis. Una vez desparafinadas en xilol, las secciones tisulares fueron rehidratadas, utilizando un gradiente de etanol y agua destilada. La recuperación antigénica se llevó a cabo sumergiendo los cortes en tampón citrato (Dako Cytomation, Sant Just Desvern, España) y autoclavándolos durante 10 min a 120°C.

Posteriormente, las muestras fueron incubadas en una solución de H<sub>2</sub>O<sub>2</sub> al 3%, para eliminar la actividad peroxidasa endógena y se incubaron con el anticuerpo primario. Se utilizaron los anticuerpos monoclonales de ratón anti-Ku70Ab-4 y el anti-Ku80Ab-7 (LabVision, Freemont, USA), ambos a una dilución 1/200. Para la detección, se utilizó

el sistema “EnVison+Dual link System-HRP Kit” (DakoCytomation, Sant Just Desvern, España). Finalmente, las muestras fueron contrateñidas con hematoxilina.

### **Protocolo de Tinción Inmunohistoquímica**

- 1- Incubación de los cortes a 56-58°C durante toda la noche.
- 2- Desparafinación de la muestra en xilol durante 5 min
- 3- Repetir el paso 2 en una nueva solución de xilol.
- 4- Repetir el paso 2 en una solución 50% xilol + 50% EtOH.
- 5- Rehidratación de la muestra en las siguientes soluciones;
  - a- 2 min en EtOH 100%
  - b- 2 min en EtOH 96%
  - c- 2 min en EtOH 80%
  - d- 2 min en EtOH 70%
  - e- 2 min en EtOH 50%
  - f- 5 min en H<sub>2</sub>O destilada
- 6- Eliminación de la actividad peroxidasa endógena incubando la muestra durante 10 min a una solución de H<sub>2</sub>O<sub>2</sub> al 3%. Lavado con PBS.
- 7- Recuperación antigénica: Autoclavar los cortes durante 10 min a 120°C en el tampón “Target Retrieval pH 6” (DAKO). Lavado con PBS.
- 8- Incubar con anticuerpo primario durante 30 min en cámara húmeda. Lavado con PBS.
- 9- Incubar durante 30 min en EndVision (DAKO) en cámara húmeda. Lavado con PBS.
- 10- Incubar en la solución DAB (DAKO) durante 5 min en cámara húmeda. Lavado con PBS.
- 11- Sumergir las preparaciones durante 2 min en una solución de hematoxilina y lavar con agua.
- 12- Deshidratación en las siguientes soluciones:
  - a- 2 min en EtOH 50%
  - b- 2 min en EtOH 70%
  - c- 2 min en EtOH 80%
  - d- 2 min en EtOH 90-96%
  - e- 2 X 2 min en EtOH 100%
  - f- 2 min en Xilol 50% y EtOH 50%
  - g- 2 X 2 min en Xilol.
- 13- Montaje de la preparación con medio DPX (Sigma Aldrich, Tres Cantos, España)

### **4.2 ANÁLISIS MORFOMÉTRICO**

Para cuantificar las preparaciones inmunohistoquímicas, se utilizaron tres imágenes de cada muestra, que fueron tomadas a 100 aumentos con una cámara DP50 y un microscopio Olympus DX52 (Olympus, Incorp, USA), bajo idénticas condiciones de iluminación y tiempo de exposición. Se utilizaron los programas ViewFinder Lite v1.0 y Studio Lite v1.0 (Olympus, Incorp, USA) para la captura y archivo de imágenes. Posteriormente, con la ayuda del programa AdobePhotoshop v7.0 (AdobePhotoshop systems, Incorp, USA), se eliminaron de la imagen todas aquellas áreas que no presentaban células tumorales.

La tinción inmunohistoquímica fue medida aplicando el modelo de cuantificación de color HSI (Hue-saturation-intensity), utilizando el programa Metamorph (Universal



Imaging Corp, Downnigtown, USA). La aplicación del modelo HSI permite seleccionar una determinada área en función de su color, el cual se define a partir de tres parámetros: HUE (la longitud de onda de transmisión de luz dominante), saturación e intensidad. Para cada pixel presente en una imagen, los valores de HUE, saturación e intensidad son transformados de manera independiente en uno de los 256 valores integrales de 0 a 255.<sup>154, 155</sup>

La herramienta Set Color Threshold del programa Metamorph permite fijar un rango entre 0 y 255 para estos tres parámetros (HUE, intensidad y saturación) seleccionando dentro de una imagen un área de interés, que presenta características de color determinadas.

### **5. MICROARRAYS**

El análisis por microarrays consiste en la hibridación de RNA previamente marcado, obtenido de una muestra problema, con un número elevado de secuencias de DNA conocidas e inmovilizadas de un modo ordenado sobre un soporte físico. En un único experimento, se obtiene información relativa a la expresión de miles de genes de una muestra problema.

Existen dos tipos de microarrays, los de un color y los de dos colores. En los microarrays de un color, generalmente, se sintetiza sobre un soporte físico una serie de oligonucleótidos específicos de los genes a interrogar. Al hibridar el RNA marcado de la muestra a analizar, se produce una emisión de fluorescencia cuya intensidad es mayor o menor dependiendo de la expresión del gen analizado. En los microarrays de dos colores, se disponen sobre un soporte físico, DNA complementario al RNA cuya expresión queremos analizar y se hibridan con dos muestras marcadas con fluorocromos distintos. En cada experimento se hibrida una muestra control y una muestra problema. Se trata de una hibridación por competencia, obteniendo como resultado una señal que corresponde a la fusión de los dos fluorocromos utilizados en el marcaje.

Los experimentos de microarrays descritos en este estudio han sido realizados utilizando la plataforma Affymetrix de microarrays de un color y el chip HG-U133A2. La plataforma Affymetrix está diseñada para la utilización de chips de oligonucleótidos sintetizados directamente sobre un soporte de cristal. Cada gen está representado por 12 pares de sondas de 22 mer; es decir, 12 sondas específicas de la región que se quiere detectar y 12 sondas control adicionales. Los controles son sondas con algún nucleótido

desapareado (mismatch) que permiten descartar hibridaciones inespecíficas. La señal obtenida para un gen determinado es el resultado de integrar la señal del conjunto de sondas que lo representa. El chip HG-U133A2 contiene 20.000 conjuntos de sondas y permite determinar la mayoría de genes humanos que presentan algún tipo de anotación funcional.

### **5.1 MARCAJE, HIBRIDACIÓN, REVELADO Y ESCANEADO**

El proceso completo de marcaje de la muestra, hibridación, lavados y revelado se llevó a cabo en la plataforma Affymetrix, ubicada en el “Hospital de la Santa Creu i Sant Pau”, siguiendo las indicaciones del fabricante (Affymetrix, Santa Clara, USA). La plataforma consta de una estación de fluidos 450, un horno de hibridación y un escaner Affymetrix 3000. Para el marcaje de las muestras, se utilizó el kit “GeneChip One-Cycle Target Labeling” (Affymetrix, Santa Clara, USA) (Figura 17).

El primer paso en la preparación de la muestra consistió en la síntesis de cDNA de doble cadena a partir del RNA total purificado. En el proceso de síntesis se incorporó al cDNA la secuencia del promotor de la polimerasa T7, el cual permitió, mediante una reacción de transcripción in vitro, obtener cRNA marcado con biotina. El cRNA biotinilado se fragmentó dando lugar a secuencias de 35 a 200 bases que posteriormente se hibridaron con el microarray HG-U133A2. A continuación, se realizaron una serie de pasos de lavado y revelado, utilizando un anticuerpo ficoeritrina/etreptavina ligado a un fluorocromo que permite que el cRNA hibridado marcado con biotina emita una señal de fluorescencia detectable por el escaner.

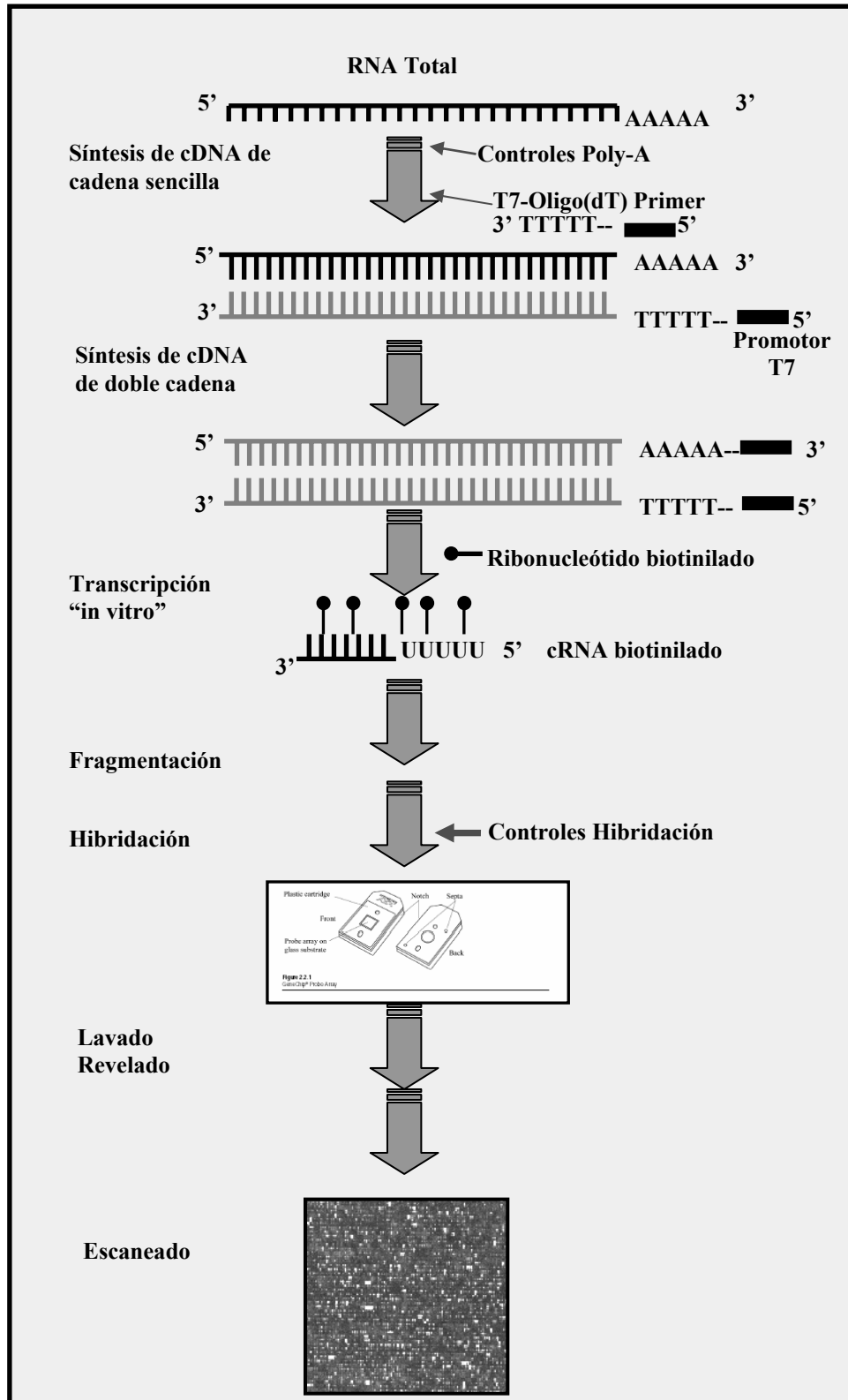


Figura 17. Esquema del proceso de elaboración de los chips de Affymetrix, utilizando el kit de marcaje "GeneChip One-Cycle Target Labelling".

**Síntesis de cDNA:**

**Síntesis de cDNA de cadena sencilla:**

1- Preparar la siguiente reacción (RNA / T7 mix):

2 µg de RNA total -----variable (máximo 8 µl)  
 Dilución de controles Poly A ----- 2 µl  
 T7-Oligo (dt) primers, 50µm ----- 2 µl  
 c.s. de H<sub>2</sub>O libre de RNAsas hasta ----- 12 µl

---

Volumen Total 12 µl

2- Incubar la reacción durante 10 min a 70°C. Enfriar la muestra a 4°C durante un mínimo de 2 min.

3- Añadir 7 µl de la First Strand Master Mix a cada tubo de RNA / T7 Mix e incubar a 42°C durante 2 min.

5 X 1st Strand Reaction Mix ----- 4 µl  
 DTT, 0,1 M ----- 2 µl  
 DNTPs, 10 nM ----- 1 µl

---

Volumen Total 7 µl

4- Añadir 1µl de Superscript II a cada reacción e incubar a 42°C durante 1hora.

**Síntesis de cDNA de doble cadena:**

1- Añadir, a cada reacción 130 µl de la siguiente reacción e incubar durante 2 h a 16°C:

5X 2nd Strand reaction Mix -----30 µl  
 dNTPs, 10nM -----3 µl  
 E.coli DNA ligasa ----- 1 µl  
 E.coli DNA Polimerasa I -----4 µl  
 RNasa H -----1 µl  
 H<sub>2</sub>O libre de RNAsas -----91 µl

---

VOLUMEN TOTAL 130 µl

2- Añadir 2 µl de T4 DNA Polimerasa a cada muestra e incubar durante 5 min a 16°C.

3- Añadir 10 µl de EDTA 0,5 M para parar la reacción y proceder al lavado del cDNA.

**Purificación de cDNA doble cadena:**

1- Añadir 600 µl de cDNA Binding Buffer al cDNA. Agitar durante 3 s.

2- Pasar 500 µl de la mezcla a una columna montada sobre un tubo colector de 2 ml. Centrifugar durante 1 min a 8.000g. Descartar el eluido.

3- Cargar el resto de la mezcla en la columna y centrifugar durante 1 min a 8.000 g. Descartar el eluido.

4- Pasar la columna a un tubo colector limpio. Añadir 750 µl de cDNA Wash Buffer a la columna. Centrifugar durante 1 min a 8.000 g. Descartar el eluido.

5- Centrifugar durante 5 min a máxima velocidad (25.000 g). Descartar el eluido.

6- Pasar la columna a un eppendorf limpio y pipetear 14 µl de cDNA Elution Buffer directamente encima de la membrana. Incubar durante 1 min, a temperatura ambiente, y centrifugar durante 1 min a máxima velocidad (25.000 g) para eluir.

## Métodos

### **Síntesis de cRNA Biotinilado (Genechip IVT Labelling Kit)**

1-Preparar la siguiente reacción e incubar a 37°C durante 16 horas

Template cDNA -----	12 µl
10 X IVT Labeling Buffer -----	4 µl
IVT Labeling NTP MIX -----	12 µl
IVT Labeling Enzyme Mix -----	4 µl
H <sub>2</sub> O libre de RNAsas -----	8 µl

---

Volumen total 40 µl

### **Purificación del cRNA biotilado**

- 1- Añadir 60 µl de H<sub>2</sub>O libre de RNAsas al cRNA biotilado.
- 2- Añadir 350 µl de IVT cRNA Binding Buffer a la muestra.
- 3- Añadir 250 µl de etanol 100%.
- 4- Cargar los 700 µl de muestra en una columna cRNA Cleanup Spin, montada en un tubo colector. Centrifugar durante 15 s a 8000g. Descartar el eluido.
- 5- Pasar la columna a un tubo colector nuevo y añadir 500 µl de IVT cRNA Wash Buffer. Centrifugar 15 s a 8.000 g. Descartar el eluido
- 6- Añadir 500 µl de etanol 80% y centrifugar 15s a 8.000 g. Descartar el eluido.
- 7- Centrifugar durante 5 min a máxima velocidad (25.000 g). Descartar el eluido.
- 8- Pasar la columna a un eppendorf de 1,5 ml. Añadir 11 µl de H<sub>2</sub>O libre de RNAsas sobre la membrana y centrifugar durante 1 min a la máxima velocidad para eluir.
- 9- Añadir 10 µl de H<sub>2</sub>O libre de RNAsas, directamente sobre la membrana. Centrifugar durante 1 min a máxima velocidad para eluir.
- 10- Cuantificar la concentración de cRNA en el espectrofotómetro.

### **Fagmentación del cRNA:**

1- Prepararemos la siguiente reacción de fragmentación.

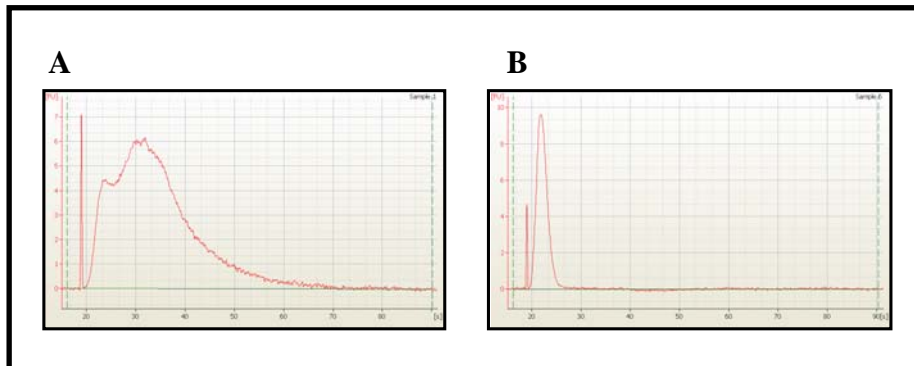
cRNA -----	15 µg (1 to 21µl)
5X Fragmentation Buffer -----	6 µl
c.s. de H <sub>2</sub> O libre de RNAsas hasta--	30 µl

---

VOLUMEN FINAL 30µl.

2- Incubar a 94°C durante 35 min. Colocar en hielo. Guardar una alícuota para el Bioanalyzer y congelar a -70°C).

Una vez finalizada la preparación de la muestra, se determinó la calidad del cRNA y del cRNA fragmentado con el analizador Bioanalyser para comprobar que todo el proceso de marcaje y fragmentación del cRNA había sido eficiente (Figura 18).



**Figura 18. Control de calidad del cRNA marcado. A) Espectrofenograma característico de una muestra de cRNA marcada sin fragmentar. B) Espectrofenograma del mismo cRNA una vez fragmentado. Una fragmentación óptima del cRNA genera fragmentos de 35 a 200 bases.**

La hibridación se llevó a cabo utilizando 10  $\mu\text{g}$  de cRNA fragmentado que, junto al cóctel de hibridación, se introdujeron en el chip HG-U133A2. Cada chip fue incubado durante 16 horas en el horno de hibridación a 45°C, con una rotación de 60 rpm.

## Métodos

### Hibridación:

1- Preparación del el Coctel de hibridación:

		<u>Concentración final</u>
cRNA FRAGMENTADO	10 µg	0,05µg / µl
Oligo B2 controls	3,3 µl	50pM
20 X Eukaryotic Hybridization		
Controls (bioB, bioC, bioD, cre)	10 µl	1,5; 5; 25 y 100 pM
Herring Sperm DNA(10mg / ml)	2 µl	0,1 mg / ml
BSA(50mg/ml)	2 µl	0,5 mg / ml
2 X Hybridization buffer	100 µl	1X
DMSO	20 µl	10%
c.s. de H <sub>2</sub> O hasta	200 µl	
<hr/>		
VOLUMEN FINAL	200 µl	

2- Equilibrar los microarray que vamos a utilizar 10 min a temperatura ambiente.

3. Calentar el cóctel de hibridación a 99°C durante 5 min. Cargar el microarray con 130 µl de 1X Hybridization Buffer.

4. Incubar el microarray a 45°C durante 10 min con una rotación de 60 rpm.

5. Incubar el cóctel de hibridación a 45°C durante 5 min.

6. Centrifugar el coctel de hybridación durante 5 min a máxima velocidad.

7. Vaciar el microarray y cargarlo con 130 µl del coctel de hibridación, evitando coger el material insoluble situado en el fondo del tubo.

8. Incubar el chip en el horno de hibridación a 45°C, con una rotación de 60 rpm durante 16 horas.

### 12 X MES Stock Buffer(1 litro):

64,61 g de MES hydrate  
193,3 g de MES Sodium Salt  
800 ml de H<sub>2</sub>O

### 2 X Hybridization Buffer (50ml):

8,3 ml de 12X MES Stock Buffer  
17,7 ml de 5M NaCl.  
4 ml de EDTA  
0,1 ml de 10% Tween-20  
19,9 ml de H<sub>2</sub>O  
\*Guardar entre 2°C y 8°C, apartado de la luz.

Los lavados y el revelado del chip fueron realizados de manera automática en la estación de fluidos 450, aplicando el protocolo midieuk2v3 y siguiendo las indicaciones del fabricante (Affymetrix, Santa Clara, USA).

### Soluciones y reactivos utilizados en los lavados y el revelado del chip

#### Wash Buffer A: Non-Stringent Wash Buffer (1litro)

300 ml de 20 X SSPE  
1 ml de 10% Tween-20  
699 ml de H<sub>2</sub>O destilada  
Filtrar con filtro de 0,2  $\mu$ m

#### Wash Buffer B: Stringent Wash Buffer (1 litro)

83,3 ml de 12 X MES Stock Buffer  
5,2 ml de 5M NaCl  
1 ml de 10% Tween 20.  
910.5 ml de H<sub>2</sub>O destilada.  
Filtrar con filtro de 0,2  $\mu$ m

#### 2 X Stain buffer (250ml)

41,7 ml de 12 X MES Stock Buffer.  
92,5 ml de 5M NaCl.  
2,5 ml de 10% Tween-20  
113,3 ml de H<sub>2</sub>O destilada.  
Filtrar con filtro de 0,2  $\mu$ m  
Guardar a una temperatura entre 2°C y 8°C, apartado de la luz.

#### SAPE Stain Solution

600  $\mu$ l de 2 X Stain Buffer  
48  $\mu$ l de 50 mg/ml BSA  
12  $\mu$ l de 1mg/ml SAPE  
540  $\mu$ l de H<sub>2</sub>O destilada

#### Antibody Solution:

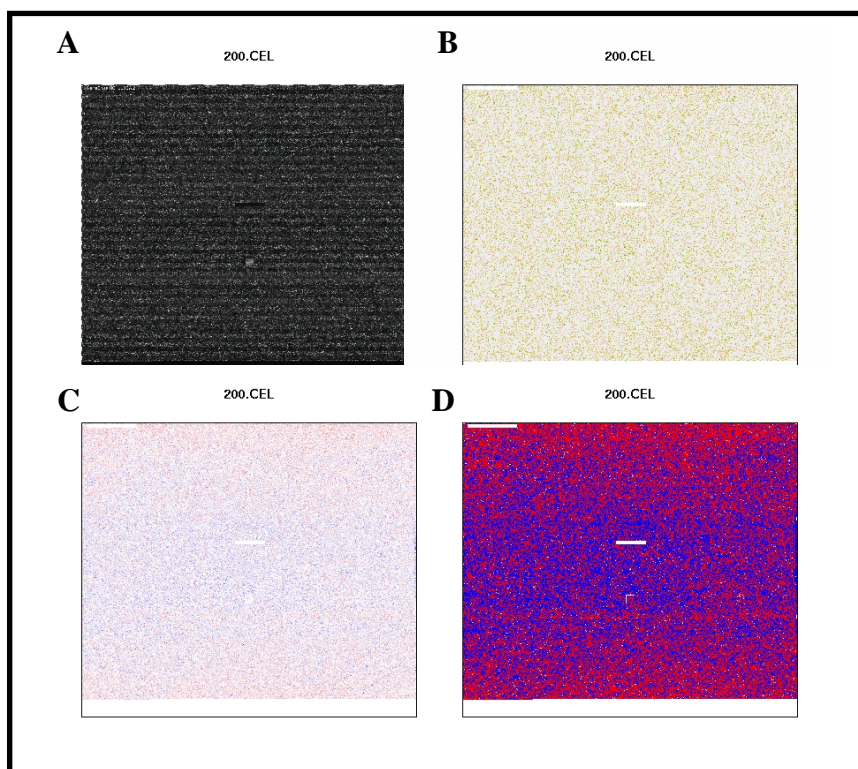
300  $\mu$ l de 2 X Stain Buffer  
24  $\mu$ l de 50 mg /ml BSA  
6  $\mu$ l de 10 mg /ml goat IgG Stock  
3,6  $\mu$ l de 0,5 mg /ml biotinylated antibody  
266,4  $\mu$ l de H<sub>2</sub>O destilada

Una vez finalizado el revelado, se sustituyó el cóctel de hibridación por Wash buffer A y se procedió inmediatamente al escaneado del chip. Como resultado del escaneado obtuvimos, para cada experimento, un archivo .CEL que contiene la imagen, la intensidad de señal y la posición de cada una de las sondas dispuestas en el chip HG-U133A2.



### 5.2 RESULTADOS BRUTOS / CONTROL DE CALIDAD

Una vez escaneados el microarray y antes de realizar el análisis estadístico se inspeccionó detalladamente la imagen obtenida para descartar posibles artefactos producidos durante el proceso de marcaje e hibridación de la sonda. El análisis de imagen y el control de calidad de la muestra fue realizado utilizando los paquetes informáticos “AffyPLM” y “simpleaffy” de Bioconductor. En cada experimento, inspeccionamos la imagen resultante de escanear el chip, así como las diferentes pseudoimágenes que genera AffyPLM para detectar artefactos no apreciables en la imagen escaneada (Figura 19).

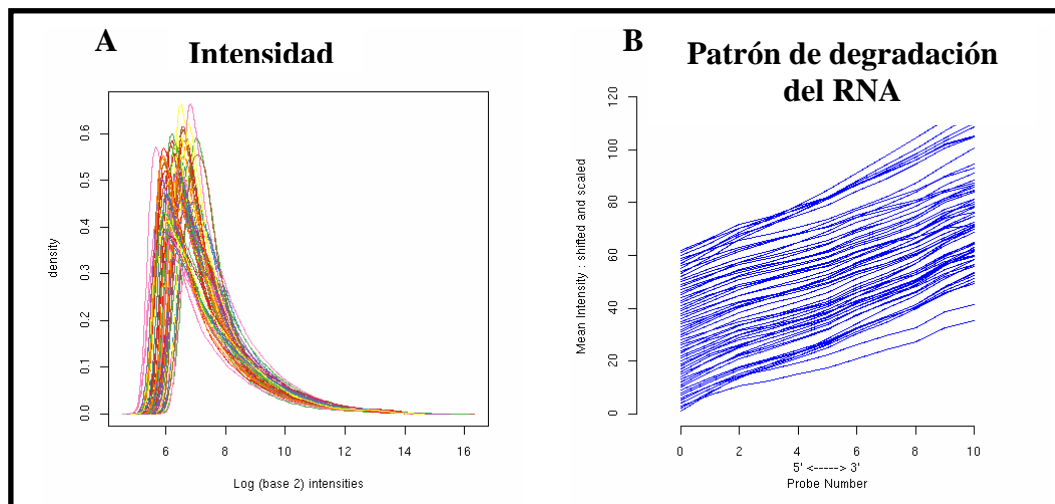


**Figura 19.** Pseudoimágenes obtenidas con el paquete informático "AffyPLM". A) Imagen de escaneado. B) "weight". C) "resids" D) "sign-resids". En este chip no observamos ningún artefacto en la imagen. La hibridación es homogénea.

Todos aquellos chips que presentaban algún artefacto en la imagen fueron eliminados del estudio posterior.

A fin de descartar variaciones en el proceso de marcaje y revelado de la sonda, analizamos las posibles diferencias en la distribución de las intensidades de todas las sondas presentes en cada chip (Figura 20A). La distribución de intensidades debe presentar un patrón similar para todos los experimentos. La calidad del RNA se

determinó calculando la relación entre la intensidad de todas las sondas situadas en el extremo 3' respecto a las sondas situadas en 5' en un mismo microarray. El patrón de degradación del RNA fue similar en todos los chips (Figura 20B).



**Figura 20. Gráficos de intensidad y degradación del RNA.** A) Los histogramas de la densidad relativa de las sondas presentan una determinada señal de intensidad en un mismo experimento de microarrays. No se observan diferencias significativas entre microarrays. B) Representación gráfica del patrón de degradación 3'/5' de las sondas presentes en cada experimento. A pesar de existir diferencias en las medias de intensidad de las diferentes curvas, todas ellas presentan una pendiente similar lo que permite concluir que todos estos experimentos son comparables.

Utilizando el paquete informático “simpleAffy” de Bioconductor, nos aseguramos de que todos los chips analizados superasen los controles de calidad descritos por el fabricante.

## 6. ANÁLISIS ESTADÍSTICO

### 6.1. ESTUDIO DEL NIVEL DE EXPRESIÓN DE LOS GENES DEL SISTEMA DE REPARACIÓN NHEJ EN BIOPSIAS PRE-TRATAMIENTO DE PACIENTES CON CECC LOCALMENTE AVANZADO

Utilizamos el test estadístico no paramétrico de la U de Mann-Whitney para determinar si existían diferencias significativas en los niveles de mRNA y en el porcentaje de células tumorales positivas para la tinción por IHQ al comparar el grupo de tumores con una respuesta tumoral > 50% (respondedores) con el grupo de tumores con una respuesta tumoral < 50% (no respondedores). Este test permite analizar si existen diferencias significativas entre dos grupos de muestras independientes en relación con una variable de estudio.

## **Métodos**

---

Por otra parte, utilizamos el análisis de curvas ROC (Receiver-Operating Characteristics) para determinar la utilidad diagnóstica de los niveles de Ku70, Ku80, DNA-PKcs, o del porcentaje de células tumorales positivas en el caso de Ku70, con objeto de discriminar entre los tumores respondedores y los no respondedores. El análisis de curvas ROC determina el rendimiento global de una prueba diagnóstica, representando sobre una gráfica la sensibilidad frente la inversa de la especificidad que tiene la variable de interés a la hora de predecir un suceso determinado. Cuanto mayor es el área bajo la curva (AUC) mayor sensibilidad y especificidad presenta una prueba. Esta representación determina el punto en el que la prueba alcanza mayor sensibilidad y especificidad. Utilizamos, por tanto, este análisis para establecer los puntos de corte óptimos, en cuanto a niveles de mRNA y de proteína, con mejor balance entre sensibilidad y especificidad con objeto de distinguir los tumores respondedores de los no respondedores.

Usamos el test de regresión logística para evaluar la asociación entre la expresión de Ku70, Ku80 y DNA-PKcs, el tamaño del tumor (T), la afectación ganglionar y la localización, con la respuesta a la quimioterapia de inducción.

Aplicando el método de Kaplan-Meier, generamos curvas de supervivencia libre de recidiva local (SLRL) o de supervivencia global (SG) identificando a los pacientes que presentaban niveles de mRNA o de proteína por encima o por debajo del umbral establecido. Para determinar si existían diferencias estadísticamente significativas en la SLRL o la SG entre el grupo de pacientes que expresaban niveles altos y el grupo de pacientes que expresaban niveles bajos se utilizó el test de Log-Rank.

Con el modelo de Cox analizamos la asociación entre la expresión de Ku70, Ku80 y DNA-PKcs, el tamaño del tumor (T), la afectación ganglionar y la localización con la SLRL y la SG del paciente. Cox construye modelos que permiten estimar el tiempo de aparición de un suceso (recidiva, muerte) en función de los datos obtenidos para las diferentes variables de interés, ya sea en análisis univariante o multivariante. Con el modelo de Cox determinamos la existencia o ausencia de diferencias significativas en el riesgo relativo de muerte o de recidiva tumoral, al comparar el grupo de pacientes con tumores con niveles altos de mRNA, o de proteína, con el grupo de pacientes con tumores con niveles bajos de mRNA, o de proteína. El análisis de Cox multivariante permite establecer si los niveles de mRNA, o de proteína, por encima o por debajo del umbral establecido son factores de riesgo independientes de la recidiva tumoral o muerte del paciente.

En cualquiera de los test estadísticos utilizados se consideraron significativas las diferencias entre grupos cuando la probabilidad estadística fue menor que 0,05 ( $p < 0,05$ ). Todos los tests estadísticos se realizaron utilizando el programa SPSS v 14.01 (SPSS, Inc., Chicago, IL).

## **6.2. ESTUDIO DE EXPRESIÓN POR MICROARRAYS EN CECC DE ESTADIOS LOCALMENTE AVANZADOS**

El pre-procesado, el análisis estadístico y el análisis de anotaciones funcionales, de los experimentos de microarrays, se llevó a cabo utilizando diferentes paquetes informáticos recogidos dentro del proyecto Bioconductor ([www.bioconductor.org](http://www.bioconductor.org)) e implementados para su utilización con lenguaje “R” ([www.r-project.org](http://www.r-project.org)) (Tabla 5).

**Tabla 5. Paquetes informáticos de Bioconductor utilizados en el análisis de datos de los experimentos de microarrays.**

<b>Paquete informático utilizado</b>	<b>Aplicación</b>
<b>Biobase</b>	<b>Aplicaciones básicas Bioconductor</b>
<b>genefilter</b>	<b>Filtrado no específico</b>
<b>affy</b>	<b>Normalización RMA</b>
<b>Multtest</b>	<b>Expresión diferencial</b>
<b>Hgu133a2</b>	<b>Anotaciones microarray HG-U133a2</b>
<b>annaffy</b>	<b>Herramientas análisis anotaciones</b>
<b>annotate</b>	<b>Anotaciones</b>
<b>GStats</b>	<b>Análisis de ontologías</b>

### **6.2.1 Pre-procesado de los experimentos de microarrays**

Todos los microarrays incluidos en el estudio fueron pre-procesados utilizando el método RMA del paquete informático “affy” para descartar variaciones tanto intra-experimentales como inter-experimentales producidas durante el proceso de marcaje e hibridación de los microarrays. El pre-procesado de los chips de Affymetrix incluye tres pasos. En primer lugar se ajusta el “background” de cada experimento, seguidamente se realiza la normalización para que todos los chips incluidos en un mismo estudio sean comparables y finalmente se aplica un método de sumarización que permite obtener un valor de expresión para cada gen, combinando las señales obtenidas de todas las sondas que lo representan. El método de pre-procesado RMA esta implementado para que su funcionamiento sea automático. Utiliza únicamente la señal de las sondas PM para la normalización de los chips, aplica el método de normalización por cuantiles y el método de sumarización basado en el algoritmo “median polish”.<sup>156, 157</sup> La Figura 21 muestra las

diferencias de señal que presentan los microarrays antes y después de aplicar el pre-procesado.

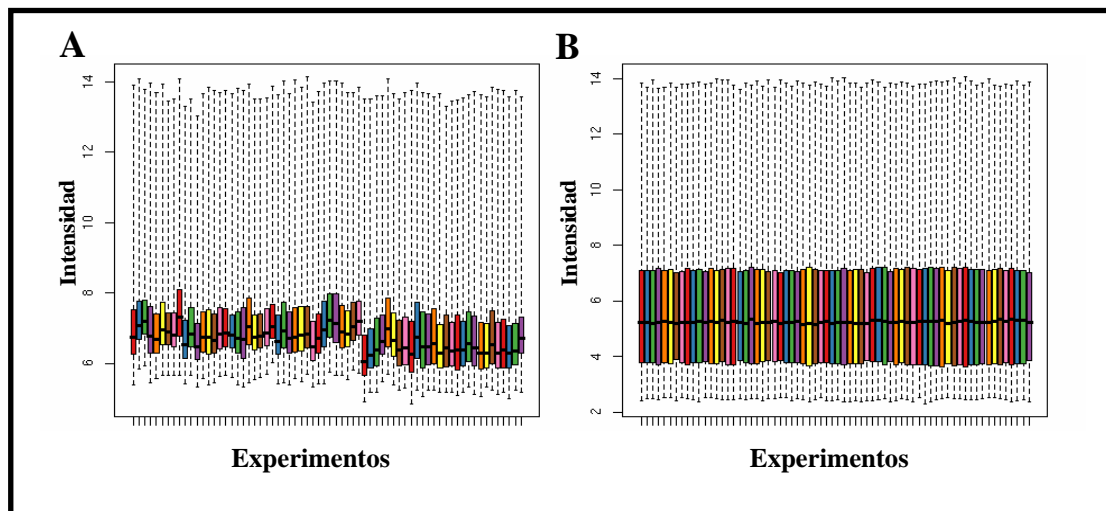


Figura 21. Box plot de los niveles de intensidad de todas las sondas presentes en cada uno de los experimentos de microarrays. Antes de realizar el pre-procesado de los microarrays (A) o bien, una vez finalizado el preprocesado RMA (B).

### 6.2.2 Análisis no supervisado ( “Clustering” )

Antes de generar el análisis de cluster de los microarrays, se realizó un filtrado de genes no específico, utilizando el paquete informático “genefilter” de bioconductor. Descartamos aquellos genes que presentaban un rango inter-cuartílico entre muestras inferior a 0,5 seleccionando de este modo aquellos sets de sondas que presentaban una mayor variabilidad entre muestras. Los otros dos criterios utilizados para el filtrado de los genes fueron que presentasen un nivel de expresión superior al  $\log(100)$  en un 25% de las muestras analizadas y que su nivel de expresión medio en todas las muestras analizadas fuese superior al  $\log(300)$ . Los genes que superaron los tres criterios de filtrado fueron utilizados para realizar la clasificación de todos los tumores y biopsias normales analizadas mediante cluster jerárquico no supervisado. Realizamos el análisis de cluster no supervisado utilizando el programa MEV v4.0 (TIGR Multiexperiment Viewer; [www.tm4.org](http://www.tm4.org). Dana-Farber Cancer Institute).<sup>158</sup> Utilizamos la distancia euclidiana para generar la matriz de similitud y seleccionamos el método de cluster jerárquico no supervisado de ligamiento completo. El cluster agrupó todas las muestras y a todos los genes seleccionados según su perfil de expresión.

Uno de los pacientes incluidos en el análisis de cluster jerárquico presentaba metástasis a distancia en el momento del diagnóstico y fue excluido de todos los análisis

posteriores ya que su tratamiento difiere de aquellos en estadios localmente avanzados, y depende en gran medida del alcance de la lesión producida por el tumor primario, así como del tipo de metástasis que presenta el paciente.

### **6.2.3. Análisis de supervivencia libre de recidiva local (SLRL) y supervivencia global (SG) entre subtipos tumorales**

Aplicando el análisis de Kaplan-Meier y el test de Log rank determinamos si existían diferencias significativas en la supervivencia libre de recidiva local (SLRL) y supervivencia global (SG) entre los pacientes pertenecientes a los diferentes subgrupos de tumores identificados por “clustering”. A continuación, utilizamos el análisis de Cox univariante y multivariante para analizar la asociación entre la clasificación de los tumores en clusters, el tamaño tumoral, la afectación ganglionar y la localización tumoral con la SLRL y la SG del paciente. Las diferencias en SLRL o SG se consideraron significativas cuando la probabilidad estadística fue menor que 0,05 ( $p < 0,05$ ).

### **6.2.4. Análisis de expresión diferencial**

El análisis de expresión diferencial permitió identificar los genes asociados con cada uno de los subtipos tumorales identificados mediante cluster, con la recidiva tumoral a los 2 años, con la supervivencia de los pacientes a los 3 años, con la respuesta tumoral a la quimioterapia de inducción y con el proceso de patogénesis del CECC.

Para realizar todos los análisis de expresión diferencial descritos en el estudio de microarrays se utilizó el paquete informático “Multtest” del proyecto Bioconductor.

En los experimento de microarrays se realizan de manera simultánea múltiples hipótesis nulas que favorecen el incremento del error estadístico de tipo I. Si no se aplica el test de corrección de error para pruebas múltiples es posible que algunos de los genes identificados como expresados diferencialmente sean consecuencia del aumento del error estadístico y que en realidad no presenten diferencias significativas en sus niveles de expresión. Todos los análisis de expresión diferencial se realizaron utilizando la función `mt.maxt` del paquete “Multtest” que aplica el test estadístico de Welch seguido del test de corrección de error para pruebas múltiples. Consideramos que un gen se expresaba diferencialmente cuando su probabilidad estadística ajustada fue menor que 0,05 ( $p < 0,05$ ).

Además, realizamos un análisis de curvas ROC para seleccionar aquellos genes que presentaban mayor sensibilidad y especificidad a la hora de predecir la recidiva tumoral a los 2 años de seguimiento y la supervivencia del paciente a los 3 años de seguimiento.

### **6.2.5. Análisis de anotaciones funcionales**

Llevamos a cabo un análisis de ontologías utilizando el paquete informático GOstats.<sup>159</sup>

GOstats está basado en la aplicación de tests hipergeométricos para determinar, dentro de una lista de genes seleccionados, aquellos términos biológicos que están sobre-representados en comparación con el total de genes presentes en el microarray. De este modo, es posible determinar la activación o inactivación de procesos biológicos o vías de señalización asociados a una condición experimental. GOstats permite analizar los tres niveles de términos biológicos (proceso biológico, componente celular y función molecular) descritos en la base de datos de Gene Ontology (GO; <http://www.geneontology.org/>) y las vías metabólicas o de señalización descritas en la base de datos “Kyoto Encyclopedia of Genes and Genomes” (KEGG; <http://www.genome.jp/kegg/>). Utilizamos GOstats para determinar en qué procesos biológicos estaban implicados los genes sobre-expresados o infra-expresados identificados previamente con el análisis de expresión diferencial. El análisis con GOstats se realizó seleccionando aquellos genes que presentaban una  $p < 0,001$  no ajustada en el análisis de expresión diferencial. De este modo, se aumenta el número de genes incluidos en el análisis y la posibilidad de identificar un grupo de genes implicados en el mismo proceso biológico o vía de señalización. Para todos los análisis de anotaciones funcionales consideramos diferencias significativas con una probabilidad menor a  $< 0,001$  ( $p < 0,001$ ).

Todas las listas de genes incluídas en este trabajo han sido generadas utilizando los paquetes informáticos de Bioconductor “annotate”, “Hgu133a2” y “annaffy”.<sup>159</sup>