



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**

## **V. RESULTADOS**



## **V RESULTADOS**

### **1. EXPRESIÓN DE GENES DEL SISTEMA DE REPARACIÓN NHEJ EN BIOPSIAS PRE-TRATAMIENTO DE PACIENTES CON CECC LOCALMENTE AVANZADO TRATADOS CON QUIMIOTERAPIA DE INDUCCIÓN SEGUIDA DE RT/QRT O CIRUGÍA**

#### **1.1. CARACTERÍSTICAS DE LOS PACIENTES INCLUIDOS EN LOS ESTUDIO PROSPECTIVO Y RETROSPECTIVO.**

La expresión de los genes del sistema de reparación no homóloga por unión de extremos (NHEJ) se evaluó en dos cohortes de pacientes independientes, una obtenida de manera prospectiva y otra de manera retrospectiva.

En el primer estudio se analizaron los niveles de mRNA de los genes Ku70, Ku80 y DNA-PKcs mediante PCR cuantitativa en 50 biopsias tumorales obtenidas antes de que los pacientes iniciasen el tratamiento de quimioterapia de inducción seguida de RT/QRT o cirugía. La recogida de biopsias se realizó de manera prospectiva. Veintiocho pacientes presentaron una respuesta tumoral a la quimioterapia de inducción superior al 50% (grupo respondedor), mientras que 22 pacientes presentaron una respuesta tumoral inferior al 50% (grupo no respondedor).

En el segundo estudio, se evaluó la expresión de los genes del sistema NHEJ en 75 biopsias tumorales de pacientes que recibieron quimioterapia de inducción seleccionadas de manera retrospectiva. El análisis de expresión de las proteínas Ku70 y Ku80 se realizó mediante tinción inmunohistoquímica (IHQ). De los 75 pacientes que formaron parte de este estudio, 39 presentaron una respuesta tumoral a la QTI superior al 50% (grupo respondedor) mientras que 36 pacientes tuvieron una respuesta tumoral inferior al 50% (grupo no respondedor).

Las características de los cincuenta pacientes incluidos en los estudios prospectivo y retrospectivo se encuentran recogidas en la Tabla 6.

## Resultados

Tabla 6. Características de los pacientes incluidos en los estudios prospectivo y retrospectivo.

Características de los pacientes	Estudio prospectivo mRNA (n=50)	Estudio retrospectivo IHQ (n=75)
<b>Sexo</b>		
Hombre	45 (90%)	71 (95%)
Mujer	5 (10%)	4 (5%)
<b>Localización tumoral</b>		
Cavidad oral	11 (22%)	7 (9%)
Orofaringe	7 (14%)	18 (24%)
Hipofaringe	14 (28%)	12 (16%)
Laringe	18 (36%)	38 (51%)
<b>Afectación ganglionar</b>		
Presente (N <sup>+</sup> )	40 (80%)	41 (55%)
Ausente (N <sup>-</sup> )	10 (20%)	34 (45%)
<b>Tamaño tumoral (T)</b>		
T2	8 (16%)	9 (12%)
T3	24 (48%)	50 (67%)
T4	18 (36%)	16 (21%)
<b>Estadío tumoral (TMN)</b>		
III	12 (24%)	40 (53%)
IV	38 (76%)	35 (47%)
<b>Tratamiento tras la QTI</b>		
Radioterapia	16 (32%)	38 (51%)
Quimioradioterapia	15 (30%)	4 (5%)
Cirugía	19 (38%)	33 (44%)

### 1.2. EXPRESIÓN DE MRNA DE KU70, KU80 Y DNA-PKcs EN FUNCIÓN DE LA RESPUESTA TUMORAL AL TRATAMIENTO DE QTI (ESTUDIO PROSPECTIVO)

La Tabla 7 recoge la mediana y el rango de expresión de Ku70, Ku80 y DNA-PKcs para las 50 biopsias tumorales incluidas en el estudio prospectivo.

Tabla 7. Mediana y rango de expresión de mRNA de Ku70, Ku80 y DNA-PKcs en biopsias tumorales pre-tratamiento de pacientes candidatos a recibir QTI.

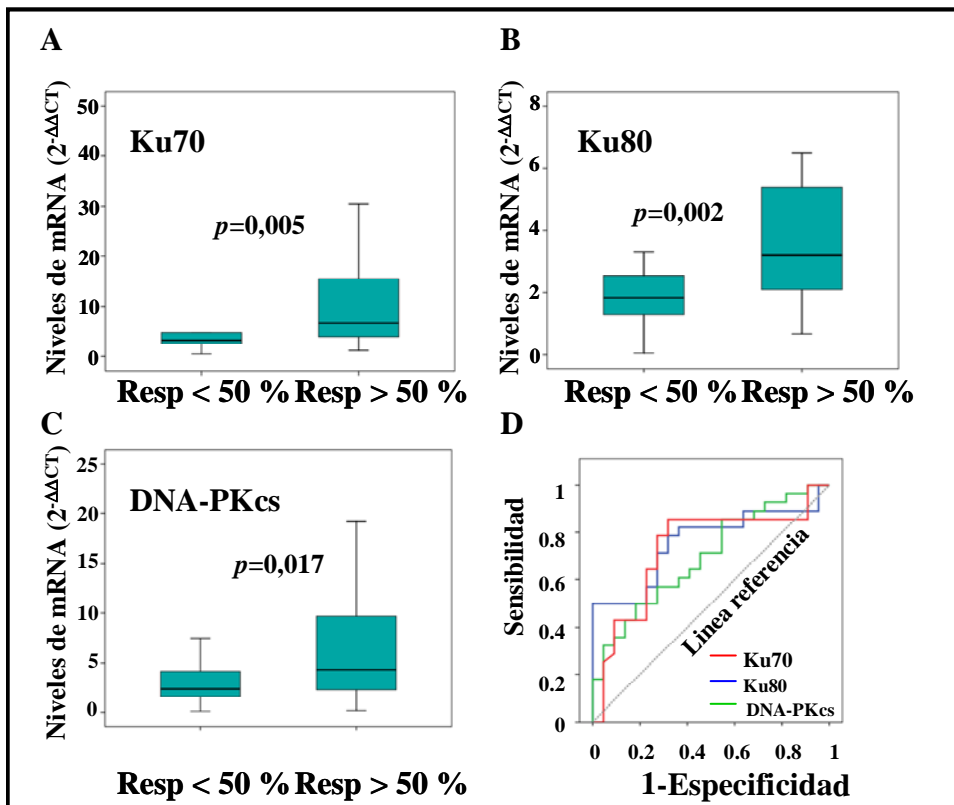
	Todos los tumores (n=50)	Respondedores (n=28)	No respondedores (n=22)
<b>Ku70</b>	4,4 <sup>a</sup> (0,5-309,8) <sup>b</sup>	6,6 <sup>a</sup> (1,2-236,4) <sup>b</sup>	3,3 <sup>a</sup> (0,5-309,0) <sup>b</sup>
<b>Ku80</b>	2,3 <sup>a</sup> (0,1-36,8) <sup>b</sup>	3,2 <sup>a</sup> (0,7-36,8) <sup>b</sup>	1,8 <sup>a</sup> (0,1-3,3) <sup>b</sup>
<b>DNA-PKcs</b>	3,3 <sup>a</sup> (0,1-59,9) <sup>b</sup>	4,2 <sup>a</sup> (0,2-59,9) <sup>b</sup>	2,4 <sup>a</sup> (0,1-11,5) <sup>b</sup>

<sup>a</sup> mediana de expresión de mRNA

<sup>b</sup> rango de expresión de mRNA (mínimo-máximo)

Los niveles de mRNA de Ku70 ( $p=0,005$ ), Ku80 ( $p=0,002$ ) y DNA-PKcs ( $p=0,017$ ) en los tumores de pacientes respondedores fueron significativamente más elevados que en los tumores de pacientes no respondedores (Figura 22A-C).

El análisis de curvas ROC (Receiver-Operating-Characteristics) determinó, para cada uno de los genes estudiados, la sensibilidad y especificidad del nivel de mRNA de predecir la respuesta tumoral a la QTI (Figura 22D). El área bajo la curva (AUC) fue de 0,73 [IC (95%)=0,59-0,88;  $p=0,005$ ] para Ku70; 0,76 [IC (95%)=0,63-0,90;  $p=0,002$ ] para Ku80 y 0,70 [IC(95%)= 0,55-0,88;  $p=0,017$ ] para DNA-PKcs.



**Figura 22.** Diferencias significativas en los niveles de mRNA de Ku70 (A), Ku80 (B) y DNA-PKcs (C) en biopsias pre-tratamiento entre los tumores con una respuesta a la quimioterapia de inducción superior al 50% (Resp >50 %) y los tumores con una respuesta inferior al 50% (Resp<50%) (Test U de Mann-Whitney). D) Curvas obtenidas mediante análisis ROC (Receiver-Operating-Characteristics) para Ku70, Ku80 y DNA-PKcs.

Utilizamos el punto de la curva ROC con mejor equilibrio entre sensibilidad y especificidad como valor de corte para distinguir los tumores con niveles altos de mRNA de los tumores con niveles bajos de mRNA. Se fijaron los siguientes valores de corte para cada gen: 3,6 para Ku70; 2,6 para Ku80 y 5,0 para DNA-PKcs. En los análisis sucesivos se distribuyó a los pacientes en dos grupos según si presentaban niveles de mRNA por encima o por debajo del umbral establecido. La Tabla 8 muestra la sensibilidad, la especificidad y el índice de exactitud (“accuracy”) de los niveles de expresión de los genes del sistema NHEJ (Ku70, Ku80 y DNA-PKcs) en relación con su

## Resultados

capacidad de predecir la respuesta tumoral a la QTI, utilizando los umbrales de expresión fijados para cada gen.

Tabla 8. Sensibilidad, especificidad e índice de exactitud de los niveles de mRNA en la predicción de la respuesta a la QTI utilizando los umbrales de expresión fijados para cada gen.

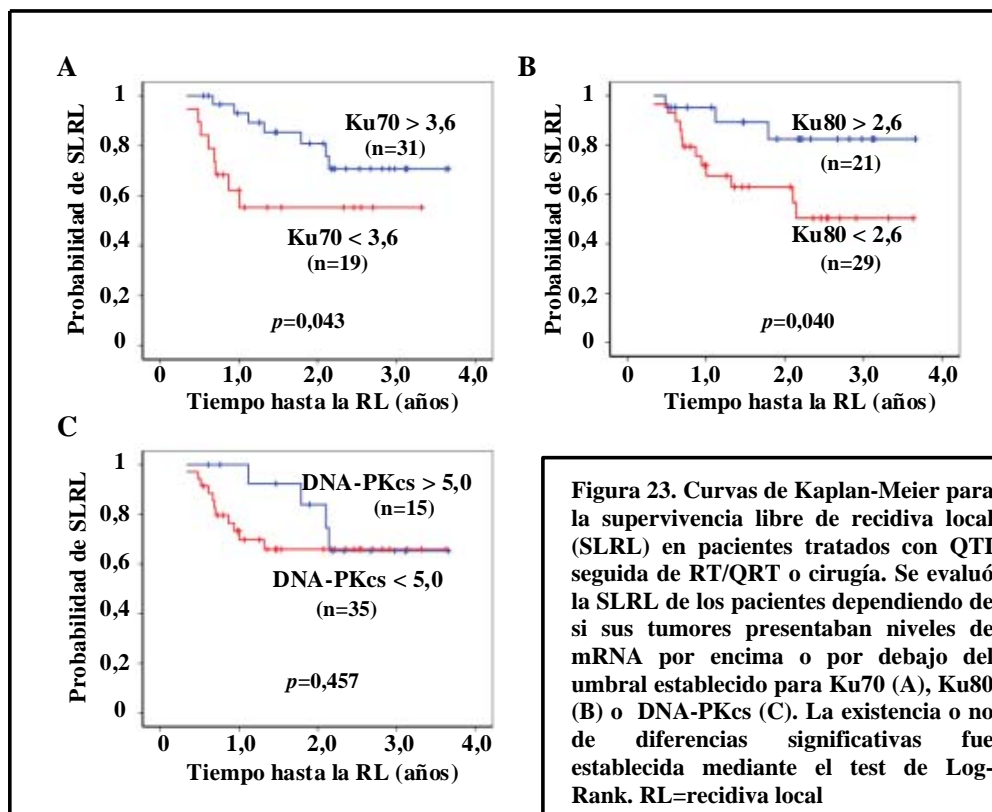
mRNA	Sensibilidad	Especificidad	Índice de exactitud
Ku70	86 %	68 %	78 %
Ku80	57 %	77 %	66 %
DNA-PKcs	43 %	86 %	62 %

Aplicamos un análisis de regresión logística para determinar la relación entre los niveles de mRNA, la localización del tumor, la afectación ganglionar (N) o el tamaño tumoral (T) y la respuesta tumoral a la QTI. Únicamente los niveles de mRNA de Ku70 presentaron diferencias significativas [ $p=0,03$ ; odds ratio 5,9; IC(95%)=1,28-29,6], de modo que niveles altos de mRNA de Ku70 ( $>3,6$ ) se asociaban con una respuesta tumoral a la QTI superiores al 50%.

En resumen, el análisis de ROC y el de regresión logística demostraron que de los tres genes estudiados, la expresión de Ku70 es el mejor marcador para discriminar entre los tumores con una respuesta superior al 50% y los tumores con una respuesta inferior al 50%.

### 1.3. NIVELES DE EXPRESIÓN DE KU70, KU80 O DNA-PKcs Y SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL (ESTUDIO PROSPECTIVO)

La supervivencia libre de recidiva local (SLRL) fue mayor en aquellos pacientes con tumores que expresaban niveles elevados de mRNA de Ku70 ( $>3,6$ ) que en los pacientes cuyos tumores expresaban niveles de Ku70 por debajo del umbral establecido ( $<3,6$ ) ( $p=0,043$ ) (Figura 23A). De manera similar, los pacientes cuyos tumores expresaban niveles de mRNA de Ku80 por encima del umbral ( $>2,6$ ) presentaban mayor SLRL que los pacientes cuyos tumores expresaban niveles de Ku80 por debajo del umbral establecido ( $<2,6$ ) ( $p=0,040$ ) (Figura 23B). La SLRL no mostró diferencias significativas entre pacientes portadores de tumores con niveles de mRNA elevados de DNA-PKcs ( $>5,0$ ) y pacientes con tumores que expresaban niveles bajos de DNA-PKcs ( $<5,0$ ) ( $p=0,457$ ) (Figura 23C).



Aplicando el modelo de Cox se estudió la asociación de los niveles de mRNA, el tamaño tumoral, la localización tumoral y la afectación ganglionar, con la SLRL (Tabla 9A). El análisis de Cox univariante indicó que los niveles de mRNA de Ku70 ( $p=0,05$ ) y de Ku80 ( $p=0,05$ ) están asociados de manera significativa con la SLRL. A su vez, el análisis de Cox multivariante identificó los niveles de mRNA de Ku70 ( $p=0,04$ ) y el tamaño tumoral ( $p=0,03$ ) como los únicos factores de riesgo independiente de padecer una recidiva local (Tabla 9A).



## Resultados

Tabla 9. Tasas de riesgo (TR) para la supervivencia libre de recidiva local (SLRL).

	Supervivencia libre de recidiva local (SLRL)			
	Cox Univariante		Cox Multivariante	
	TR (IC 95%) <sup>a</sup>	p	TR (IC 95%) <sup>a</sup>	p
<b>A. Todos los Pacientes</b>				
Ku70 mRNA (<3,6 vs >3,6)	2,8 ( 1,0-7,7)	<b>0,05*</b>	4,7 (1,1-19,8)	<b>0,04*</b>
Ku80 mRNA (<2,6 vs >2,6)	3,5 (1,0-12,4)	<b>0,05*</b>	2,1 (0,5-9,3)	0,34
DNA-PKcs (< 5,0 vs >5,0)	1,5 (0,5-4,9)	0,46	1,7 (0,2-12,6)	0,59
Tamaño tumoral (T3-T4 vs T1-T2)	2,3 (0,8-6,6)	0,14	7,4 (1,3-42,9)	<b>0,03*</b>
Afectación ganglionar (N+ vs N0)	2,1 (0,5-9,3)	0,33	1,0 (0,2-6,0)	0,97
Localización tumoral	1,4 (0,4-4,3)	0,60	2,1 (0,5-8,6)	0,32
<b>B. Trat. genotóxico conservador</b>				
Ku70 mRNA (<3,6 vs >3,6)	6,9 (1,9-26,5)	<b>&lt;0,01*</b>	28,2 (1,7-47,0)	<b>0,02*</b>
Ku80 mRNA (<2,6 vs >2,6)	2,9 (0,7-11,8)	0,13	0,9 (0,0-8,6)	0,95
DNA-PKcs (< 5,0 vs >5,0)	1,7 (0,4-6,9)	0,45	0,4 (0,02-7,3)	0,53
Tamaño tumoral (T4 vs T1-T2)	1,8 (0,4-7,2)	0,42	4,9 (0,7-36,1)	0,12
Afectación ganglionar (N+ vs N0)	2,6 (0,3-20,6)	0,37	1,3 (0,1-16,5)	0,85
Localización tumoral	1,1 (0,2-5,1)	0,94	0,5 (0,1-3,2)	0,45

<sup>a</sup>TR= Tasa de riesgo de recidiva; IC 95%= Intervalo de confianza al 95%.

\*Existencia de diferencias significativas tras aplicar el modelo de Cox univariante y multivariante.

En resumen, los pacientes con tumores que expresan niveles de Ku70 y Ku80 por encima del umbral presentan un riesgo menor de padecer recidiva local del tumor primario que los pacientes con tumores que expresan niveles de Ku70 y Ku80 por debajo del umbral establecido, tras una media de seguimiento de 25 meses (2 años y un mes). La expresión de Ku70 y el tamaño tumoral son factores de riesgo independientes de recidiva del tumor primario.

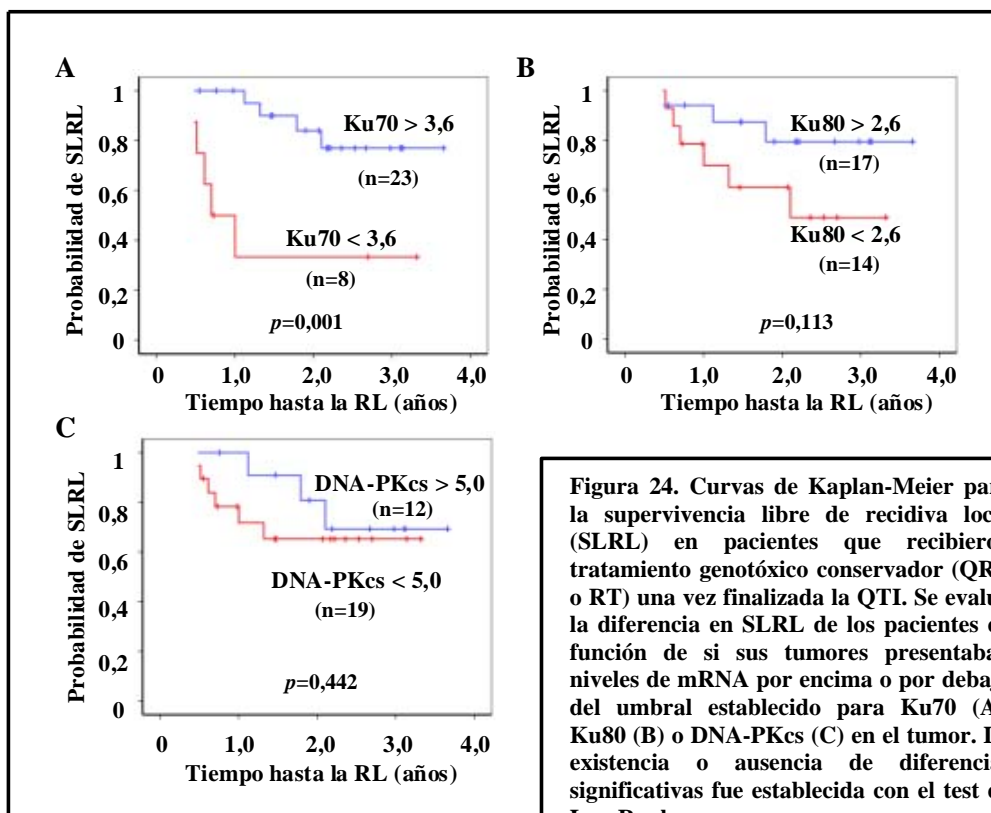
### 1.4. NIVELES DE EXPRESIÓN DE KU70, KU80 O DNA-PKcs Y SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL EN PACIENTES QUE SIGUIERON UN TRATAMIENTO GENOTÓXICO CONSERVADOR (ESTUDIO PROSPECTIVO)

El objetivo de este estudio es determinar la utilidad de Ku70, Ku80 y DNA-PKcs como posibles marcadores de la respuesta y la recidiva tumoral tras el tratamiento genotóxico. Una vez finalizada la QTI, los pacientes incluidos en este estudio siguieron un tratamiento genotóxico conservador de la función del órgano (QRT o RT) o recibieron cirugía dependiendo de la respuesta tumoral. La cirugía se aplica a pacientes con tumores que no responden a la QTI y su inclusión en el análisis de SLRL podría alterar los valores que se obtendrían por efecto directo del tratamiento genotóxico. Por esta razón, analizamos las posibles diferencias en SLRL en pacientes que recibieron QRT o RT tras la QTI, excluyendo del análisis los pacientes tratados con cirugía. Todos los pacientes del grupo respondedor (n=28) siguieron el tratamiento genotóxico

conservador (QRT o RT). A diecinueve pacientes del grupo no respondedor se les aplicó cirugía tras la QTI. Los tres pacientes restantes del grupo no respondedor, a pesar de las indicaciones médicas, rechazaron la cirugía por su efecto mutilante y siguieron el protocolo de tratamiento conservador (QRT o RT).

El análisis de los 31 pacientes que siguieron el protocolo conservador demostró que los pacientes con tumores que expresaban niveles elevados de mRNA de Ku70 tenían una mayor SLRL que los pacientes con tumores que expresaban niveles bajos de Ku70, siendo estas diferencias significativas ( $p=0,001$ ) (Figura 24A). Para Ku80 se observó la misma tendencia pero las diferencias no alcanzaron significación estadística ( $p=0,113$ ) (Figura 24B). DNA-PKcs no mostró diferencias en relación a la SLRL ( $p=0,442$ ) (Figura 24C).

El análisis de Cox univariante confirmó la asociación de los niveles de mRNA para Ku70 con la SLRL ( $p<0,01$ ) (Tabla 9B). El análisis multivariante identificó Ku70 como el único factor de riesgo (o variable) independiente con capacidad de predecir la SLRL en pacientes que siguen un tratamiento conservador después de la QTI ( $p=0,02$ ) (Tabla 9B).



## Resultados

En resumen, en nuestra cohorte de pacientes sometidos a un tratamiento genotóxico una vez finalizada la quimioterapia de inducción y con un seguimiento medio de 25 meses (2 años y un mes), la expresión de Ku70 es el único factor de riesgo independiente de recidiva tumoral. En el análisis multivariante, se observó un riesgo de recidiva 6 veces mayor entre tumores con niveles altos y niveles bajos de mRNA de Ku70 (TR: 28,2 vs 4,7. Tabla 9) cuando sólo se incluyeron en el análisis únicamente los pacientes sometidos a tratamiento genotóxico después de la quimioterapia de inducción.

### 1.5. NIVELES DE MRNA DEL COMPLEJO DNA-PK (KU70, KU80 Y DNA-PKcs) Y SUPERVIVENCIA GLOBAL (ESTUDIO PROSPECTIVO)

Los pacientes con tumores que expresaban niveles elevados de mRNA de Ku70 o de Ku80 tenían una probabilidad de supervivencia mayor que los pacientes cuyos tumores expresaban niveles bajos de mRNA (Figura 25), tras un seguimiento medio de 25 meses (2 años y 1 mes). No obstante, en ningún caso las diferencias alcanzaron la significación estadística.

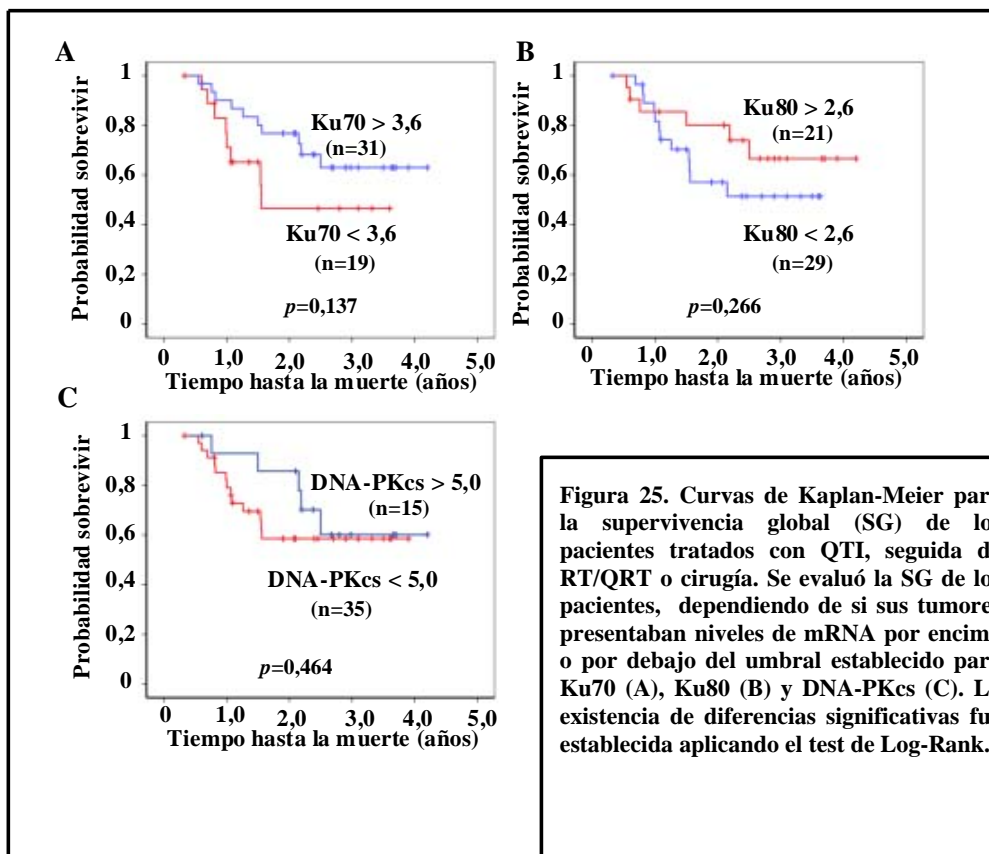


Figura 25. Curvas de Kaplan-Meier para la supervivencia global (SG) de los pacientes tratados con QTI, seguida de RT/QRT o cirugía. Se evaluó la SG de los pacientes, dependiendo de si sus tumores presentaban niveles de mRNA por encima o por debajo del umbral establecido para Ku70 (A), Ku80 (B) y DNA-PKcs (C). La existencia de diferencias significativas fue establecida aplicando el test de Log-Rank.

**1.6. EXPRESIÓN DE LAS PROTEÍNAS KU70 Y KU80, Y RESPUESTA TUMORAL (ESTUDIO RETROSPECTIVO)**

La tinción inmunohistoquímica de Ku70 y Ku80 mostró un patrón de expresión nuclear. Para el análisis cuantitativo de expresión utilizamos la herramienta “Set color threshold” del software Metamorph v.5.0 y aplicamos el modelo de cuantificación de color HSI (Figura 26A). Se estableció un rango de HUE de 180-255 para seleccionar todas las áreas (núcleos) teñidas de marrón, positivas para el anticuerpo primario, asegurándonos que no se seleccionaba ningún área en los controles negativos (Figura 26B-C). Además, se estableció un rango de HUE de 160-255 para seleccionar los núcleos (positivos y negativos) de todas las células presentes en cada imagen (marrones y azules) excluyendo el citoplasma, la membrana celular y los depósitos de queratina (Figura 26D).

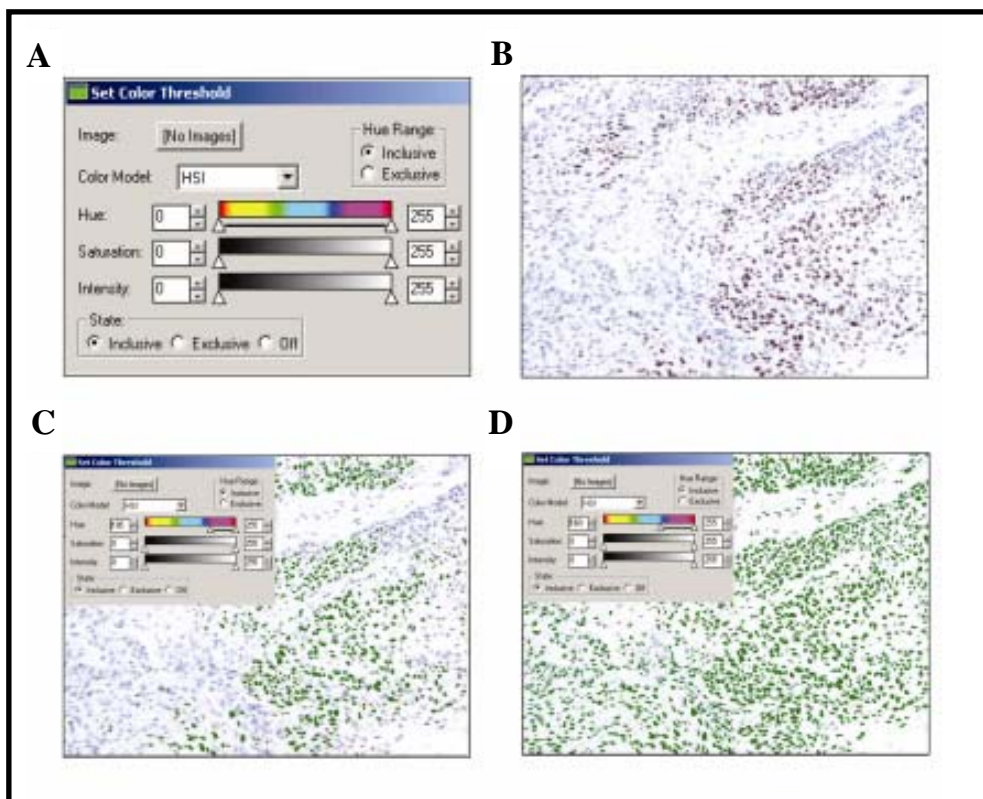
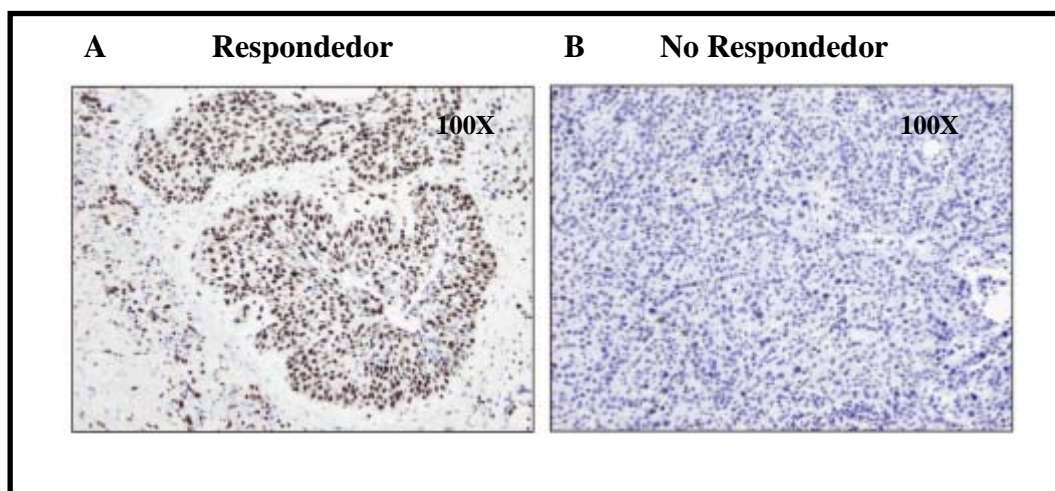


Figura 26. Cuantificación inmunohistoquímica de Ku70 a partir de imágenes tomadas a 100X. A) Utilizamos la herramienta “Set Color Threshold” del programa Metamorph v 5.0 para establecer un rango de HUE que seleccionase las áreas que presentan características inmunohistoquímicas concretas. B) Ku70 tiene un patrón de tinción exclusivamente nuclear. Los núcleos positivos para Ku70 se tiñen de color marrón. C) Usando un rango de HUE de 180-255 se seleccionan (áreas en verde) todos los núcleos positivos para Ku70 presentes en la imagen. D) Utilizando un rango de HUE de 165-255 se seleccionan (áreas en verde) todos los núcleos presentes (positivos y negativos) en la imagen excluyendo el citoplasma, la membrana celular y los depósitos de queratina. A partir de las dos áreas (el área de núcleos positivos y el área de núcleos negativos) determinamos el porcentaje de células tumorales positivas presentes en la imagen. La cuantificación inmunohistoquímica de Ku80 se llevó a cabo siguiendo los mismos criterios establecidos para Ku70.

## Resultados

Puesto que todas las células tumorales presentes en una misma imagen contenían núcleos con un tamaño similar, pudimos calcular el porcentaje de núcleos tumorales positivos dividiendo el área ocupada por los núcleos positivos por el área ocupada por el total de núcleos presentes en cada imagen. Mediante este método de cuantificación inmunohistoquímica, pudimos establecer el porcentaje de células tumorales positivas para las proteínas Ku70 y Ku80 presentes en los cortes histológicos de cada biopsia tumoral.

La Figura 27 muestra la tinción inmunohistoquímica de Ku70 en dos biopsias tumorales obtenidas de pacientes con respuestas dispares a la QTI. Los tumores incluidos en el estudio presentaron una mediana de células tumorales positivas del 63,3% (rango: 7,4-92,8%) para Ku70 y del 60,8% para Ku80 (rango: 2,3-93,2%). En el grupo respondedor, la mediana de células tumorales positivas fue de 70,7% (rango: 19,0-92,8%) para Ku70 y 65,2% (rango: 8,4-93,8%) para Ku80. En el grupo no respondedor, la mediana de células tumorales positivas fue de 57,7% (rango: 7,4-88,3%) para Ku70 y 53,2% (rango: 2,1-82,7%) para Ku80.



**Figura 27. Diferencias en la tinción inmunohistoquímica de la proteína Ku70 en tumores representativos del grupo respondedor y del grupo no-respondedor. A) IHQ de Ku70 en el tumor de un paciente con una respuesta completa. La mayoría de células tumorales presentan tinción nuclear para Ku70. B) IHQ de Ku70 en un tumor de un paciente con una respuesta a la QTI inferior al 50%. La mayoría de células tumorales son negativas para Ku70.**

Se observó una mayor expresión de la proteína Ku70 en el grupo de pacientes con tumores respondedores en comparación con el grupo de pacientes con tumores no respondedores ( $p=0,036$ ) (Figura 28A). Se observó la misma tendencia al cuantificar la expresión de la proteína Ku80, sin embargo las diferencias no alcanzaron significación estadística ( $p= 0,114$ ; U de Mann-Whitney).

El análisis de curvas ROC determinó un área bajo la curva (AUC) de 0,64 [IC(95%)= 0,51-0,77;  $p=0,036$ ] para el porcentaje de células Ku70 positivas (Figura 28B). El valor umbral para diferenciar entre los tumores que presentaban un porcentaje elevado de células positivas para Ku70 y los que presentaban un porcentaje bajo se estableció en el 74%. Con este valor de corte, la sensibilidad, especificidad e índice de exactitud de dicha prueba en cuanto a su capacidad de predecir la respuesta tumoral a la QTI fue de un 47, 69 y 59%, respectivamente.

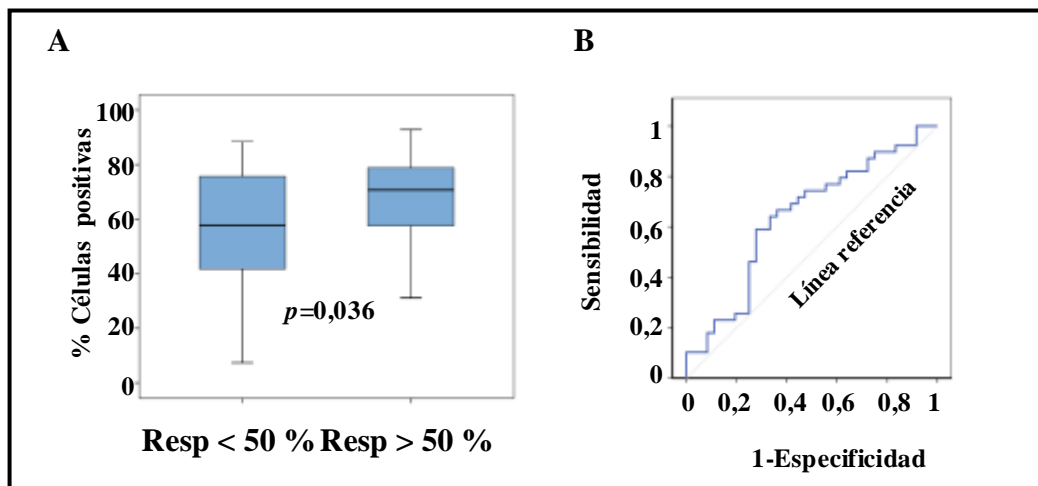


Figura 28. A) Diferencias significativas en el porcentaje de células tumorales positivas para la proteína Ku70 en función de la respuesta a la quimioterapia de inducción (U de Mann-Whitney). Resp >50%: respuesta superior al 50%; resp <50%: Respuesta inferior al 50%. B) Curva ROC para Ku70.

En resumen, en biopsias tumorales pre-tratamiento, el porcentaje de células tumorales positivas para Ku70 está asociado de modo significativo con la respuesta tumoral a la quimioterapia de inducción.

## Resultados

### 1.7. PORCENTAJE DE CÉLULAS TUMORALES POSITIVAS PARA KU70 Y SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL (ESTUDIO RETROSPECTIVO)

Los pacientes cuyos tumores tenían un elevado porcentaje de células tumorales positivas para Ku70 (>74%) mostraron una mayor SLRL que aquellos cuyos tumores presentaron un porcentaje bajo de células tumorales positivas (<74%) ( $p=0,013$ ) (Figura 29A).

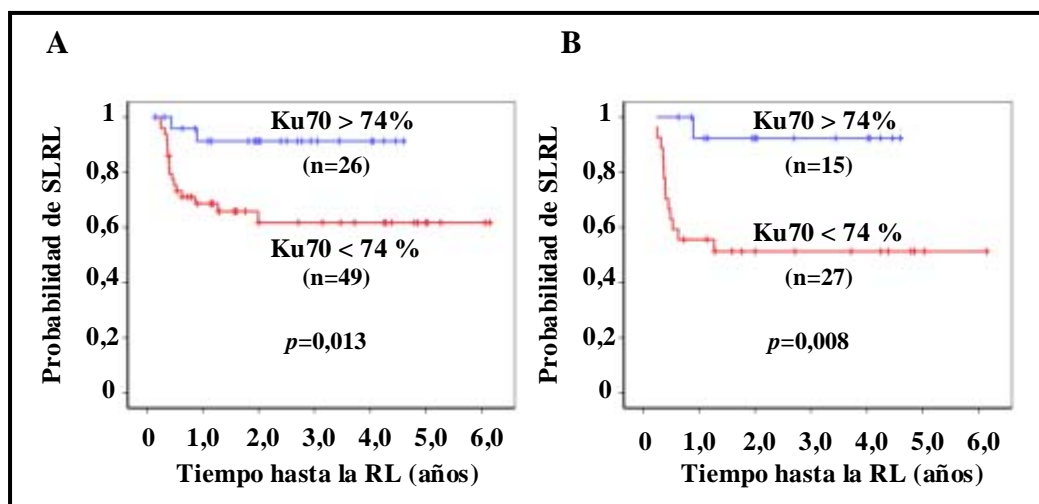


Figura 29. Curvas de Kaplan-Meier para la supervivencia libre de recidiva local (SLRL) de los pacientes en función de si sus tumores presentaban un porcentaje de células tumorales positivas para la proteína Ku70 mayor o menor que el 74 %. El análisis se realizó incluyendo todos los pacientes (A) o únicamente aquellos pacientes que habían recibido el tratamiento conservador (QRT o RT) tras la QTI (B). Diferencias significativas establecidas con el test de Log-Rank.

El análisis de Cox univariante ( $p=0,03$ ) y multivariante ( $p=0,02$ ) mostró que el porcentaje de células tumorales positivas para Ku70 es un factor de riesgo independiente asociado con la SLRL (Tabla 10A).

Tabla 10. Tasas de riesgo (TR) para la supervivencia libre de recidiva local (SLRL) en el estudio retrospectivo.

	Supervivencia libre de recidiva local (SLRL)			
	Cox Univariante		Cox Multivariante	
	TR (IC 95%) <sup>a</sup>	p	TR (IC 95%) <sup>a</sup>	p
<b>A. Todos los pacientes</b>				
% Ku70 positivas (<74 % vs >74%)	5,2 (1,2-22,9)	<b>0,03*</b>	5,6 (1,3-24,3)	<b>0,02*</b>
Tamaño tumoral (T3-T4 vs T1-T2)	2,7 (0,4-20,1)	0,34	2,7 (0,4-21,4)	0,34
Afectación ganglionar (N+ vs N0)	1,1 (0,5-2,8)	0,78	1,0 (0,4-3,0)	0,97
Localización tumoral	1,9 (0,8-4,9)	0,17	2,3 (0,8-6,7)	0,12
<b>B. Trat. genotóxico conservador</b>				
% Ku70 positivas (<74 % vs >74%)	9,5 (1,3-73,1)	<b>0,03*</b>	12,3 (1,6-96,6)	<b>0,02*</b>
Tamaño tumoral (T3-T4 vs T1-T2)	3,3 (0,4-25,0)	0,26	4,6 (0,6-38,2)	0,16
Afectación ganglionar (N+ vs N0)	1,0 (0,4-2,9)	0,99	1,7 (0,5-5,2)	0,37
Localización tumoral	2,0 (0,7-6,0)	0,22	2,8 (0,9-9,0)	0,08

<sup>a</sup>TR= Tasa de riesgo de recidiva; IC 95%= Intervalo de confianza del 95%

\*Existencia de diferencias significativas tras la aplicación del modelo de Cox univariante y multivariante.

Del mismo modo que en el estudio de expresión de mRNA por PCR, analizamos la relación entre el porcentaje de células positivas para la proteína Ku70 y la SLRL, excluyendo a los pacientes que una vez finalizada la QTI fueron tratados mediante cirugía. Todos los pacientes del grupo respondedor (n=39) fueron tratados con QRT o RT tras la QTI. Treinta y tres pacientes del grupo no respondedor fueron tratados con cirugía al finalizar la QTI. Tres pacientes del grupo de no respondedor rehusaron la cirugía radical y, a pesar de la recomendación médica, fueron tratados con tratamiento conservador (QRT o RT). Una vez excluidos los pacientes quirúrgicos, 42 pacientes siguieron un protocolo conservador y fueron los incluidos en el análisis de SLRL. Los pacientes cuyos tumores presentaban un porcentaje de células positivas para Ku70 por encima del umbral fijado (>74%) tenían menor riesgo de padecer una recidiva que aquellos cuyos tumores presentaban un porcentaje de células positivas por debajo del umbral establecido (<74%) ( $p=0,008$ ) (Figura 29B). Los análisis de Cox univariante ( $p=0,03$ ) y multivariante ( $p=0,02$ ) confirmaron la asociación entre el porcentaje de células tumorales positivas para Ku70 y la SLRL de los pacientes que siguieron un tratamiento conservador después de la QTI (Tabla 10B).

En resumen, la expresión de Ku70 es un factor de riesgo independiente de la recidiva del tumor primario en pacientes sometidos a tratamiento genotóxico con RT o QRT una vez finalizada la quimioterapia de inducción. Un elevado porcentaje de células



## Resultados

tumorales positivas para la proteína Ku70 se asocia, de modo significativo, con un aumento de la supervivencia libre de recidiva local. Las diferencias son significativamente mayores cuando se incluye en el análisis únicamente el grupo de pacientes que han seguido el tratamiento conservador (Figura 29B) que cuando se analizan todos los pacientes tratados con QTI, incluyendo los que recibieron tratamiento conservador y los quirúrgicos (Figura 29A).

### 1.8. PORCENTAJE DE CÉLULAS TUMORALES POSITIVAS PARA KU70 Y SUPERVIVENCIA GLOBAL (ESTUDIO RETROSPECTIVO)

En el análisis de Cox univariante, se observó que el porcentaje de células tumorales positivas para la proteína Ku70 ( $p=0,03$ ), la afectación ganglionar ( $p=0,03$ ) y la localización tumoral ( $p<0,01$ ) estaban asociados significativamente con la supervivencia global (SG) del paciente (Tabla 11). El análisis multivariante mostró que el porcentaje de células positivas para Ku70 ( $p<0,01$ ) y la localización tumoral ( $p<0,01$ ) son factores de riesgo independiente asociados con la SG del paciente (Tabla 11A). El análisis de Cox multivariante, realizado sólo con los pacientes que siguieron un protocolo conservador después de la QTI, mostró resultados similares (Tabla 11B).

Tabla 11. Tasas de riesgo (TR) para la supervivencia global (SG) (estudio retrospectivo).

	Supervivencia Global (SG)			
	Cox Univariante		Cox Multivariante	
	TR (IC 95%) <sup>a</sup>	p	TR (IC 95%) <sup>a</sup>	p
<b>A. Todos los pacientes</b>				
% Ku70 positivas (<74 % vs >74%)	3,9 (1,2-13,3)	<b>0,03*</b>	5,1(1,5-17,4)	<b>&lt;0,01*</b>
Tamaño tumoral (T3-T4 vs T1-T2)	1,4 (0,3-6,0)	0,65	2,2 (0,5-9,6)	0,30
Afectación ganglionar (N+ vs N0)	3,1 (1,1-8,3)	<b>0,03*</b>	1,9 (0,6-5,8)	0,26
Localización tumoral	5,9 (2,0-17,4)	<b>&lt;0,01*</b>	5,6 (1,7-18,4)	<b>&lt;0,01*</b>
<b>B. Trat. genotóxico conservador</b>				
% Ku70 positivas (<74 % vs >74%)	2,0 (0,7-13,9)	0,16	5,5 (1,1-26,5)	<b>0,03*</b>
Tamaño tumoral (T3-T4 vs T1-T2)	2,4 (0,3-18,4)	0,42	5,6 (0,7-46,9)	0,11
Afectación ganglionar (N+ vs N0)	2,2 (0,6-8,1)	0,26	2,5 (0,6-10,2)	0,19
Localización tumoral	12,3 (1,6-96,6)	<b>0,02*</b>	16,4 (2,0-134,3)	<b>&lt;0,01*</b>

<sup>a</sup>TR= Riesgo de muerte"; IC 95%= Intervalo de confianza del 95%

\* Existencia de diferencias significativas tras la aplicación del modelo de Cox univariante y multivariante

## **2. ESTUDIO DE EXPRESIÓN POR MICROARRAYS EN BIOPSIAS PRE-TRATAMIENTO DE PACIENTES CON CECC EN ESTADIOS LOCALMENTE AVANZADOS**

### **2.1. CARACTERÍSTICAS DE LOS PACIENTES INCLUIDOS EN EL ESTUDIO**

El estudio de microarrays de expresión se llevó a cabo utilizando 63 biopsias tumorales obtenidas de modo prospectivo de pacientes con carcinoma escamoso de cabeza y cuello (CECC) en estadio localmente avanzado candidatos a recibir tratamiento de quimiorradioterapia concomitante o de quimioterapia de inducción seguida de RT/QRT o cirugía. También se obtuvieron biopsias de mucosa normal de 5 pacientes de localizaciones alejadas del tumor primario donde no se apreciaba clínicamente ninguna alteración. El análisis histológico confirmó que las 5 mucosas contenían únicamente tejido normal. De los 63 pacientes incluidos en el estudio, 17 siguieron un protocolo de quimiorradioterapia concomitante mientras que 46 pacientes siguieron un protocolo de quimioterapia de inducción seguida de RT/QRT o cirugía. Las características de los pacientes incluidos en el estudio de expresión por microarrays se encuentran recogidas en la Tabla 12.

**Tabla 12. Características de los pacientes incluidos en el estudio de microarrays de expresión en CECC.**

<b>Características</b>	<b>Número (%)</b>
<b>Sexo</b>	
Hombre	56 (89%)
Mujer	7 (11%)
<b>Localización tumoral</b>	
Cavidad oral	11 (18%)
Orofaringe	21 (33%)
Hipofaringe	8 (13%)
Laringe	23 (36%)
<b>Afectación ganglionar</b>	
Presente (N <sup>+</sup> )	53 (84%)
Ausente (N <sup>-</sup> )	10 (16%)
<b>Tamaño tumoral (T)</b>	
T2	13 (21%)
T3	30 (48%)
T4	20 (31%)
<b>Estadio tumoral (TMN)</b>	
III	15 (24%)
IV	48 (76%)
<b>Tratamiento:</b>	
QTI seguida de RT	14 (22%)
QTI seguida de QRT	15 (24%)
QTI seguida de cirugía	17 (27%)
QRT concomitante	17 (27%)

## Resultados

### 2.2. ANÁLISIS NO SUPERVISADO (“CLUSTERING”) EN TUMORES Y MUCOSAS NORMALES

Antes de realizar el análisis de cluster se efectuó un proceso de filtrado para seleccionar aquellos genes que presentaban una intensidad de señal más elevada y una mayor variación en su niveles de expresión entre las diferentes muestras (ver métodos). El filtrado seleccionó 814 genes.

El cluster jerárquico no supervisado agrupó a los 63 tumores y las 5 biopsias según su perfil de expresión. Las muestras se agruparon en tres subclusters distintos dependiendo de la similitud entre perfiles (Figura 30). El cluster 1 agrupó 19 biopsias tumorales, el cluster 2 agrupó 23 biopsias tumorales y el cluster 3 agrupó las 21 biopsias tumorales y las 5 mucosas normales (Figura 30).

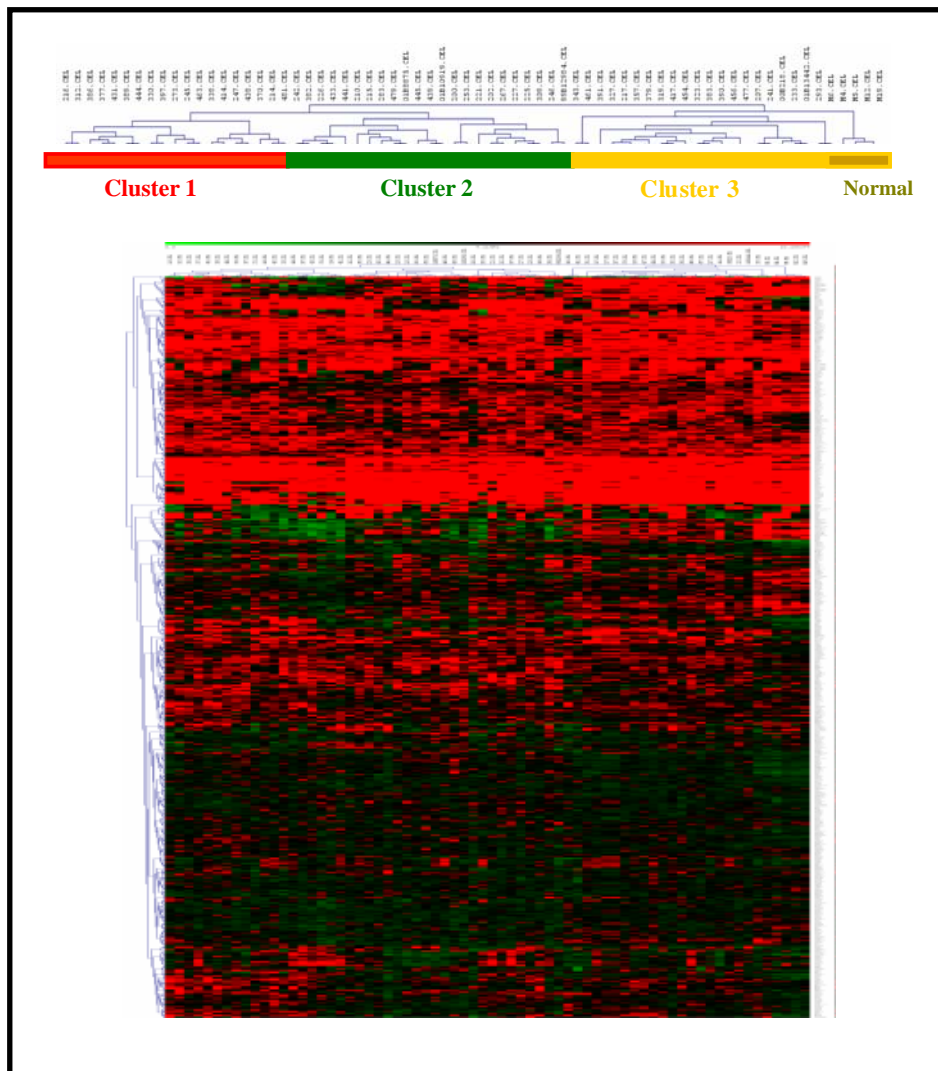


Figura 30. Cluster no supervisado realizado con las 68 muestras (63 biopsias tumorales y 5 biopsias de mucosa normal) y los 814 genes seleccionados. Las muestras se agrupan en tres clusters distintos (cluster 1, cluster 2 y cluster 3). El cluster 3 agrupa un subgrupo de tumores, conjuntamente con 5 muestras de mucosa normal. En este análisis se realizó un cluster jerárquico de ligamiento completo.

### 2.3. IDENTIFICACIÓN DE SUBTIPOS TUMORALES RELACIONADOS CON LA SUPERVIVENCIA LIBRE DE RECIDIVA LOCAL (SLRL) Y LA SUPERVIVENCIA GLOBAL (SG) DE LOS PACIENTES TRAS SU TRATAMIENTO

El análisis de supervivencia libre de recidiva local (SLRL) determinó la existencia de diferencias significativas al comparar los tumores del cluster 1 con los tumores del cluster 2 ( $p=0,008$ ), así como los tumores del cluster 1 con los del cluster 3 ( $p=0,047$ ) (Figura 31A). Los pacientes del cluster 1 mostraron una mayor probabilidad de padecer una recidiva tumoral (menor SLRL) que los pacientes incluidos en los dos clusters restantes. No se observaron diferencias significativas en SLRL entre los pacientes del cluster 2 y los del cluster 3 ( $p=0,484$ ).

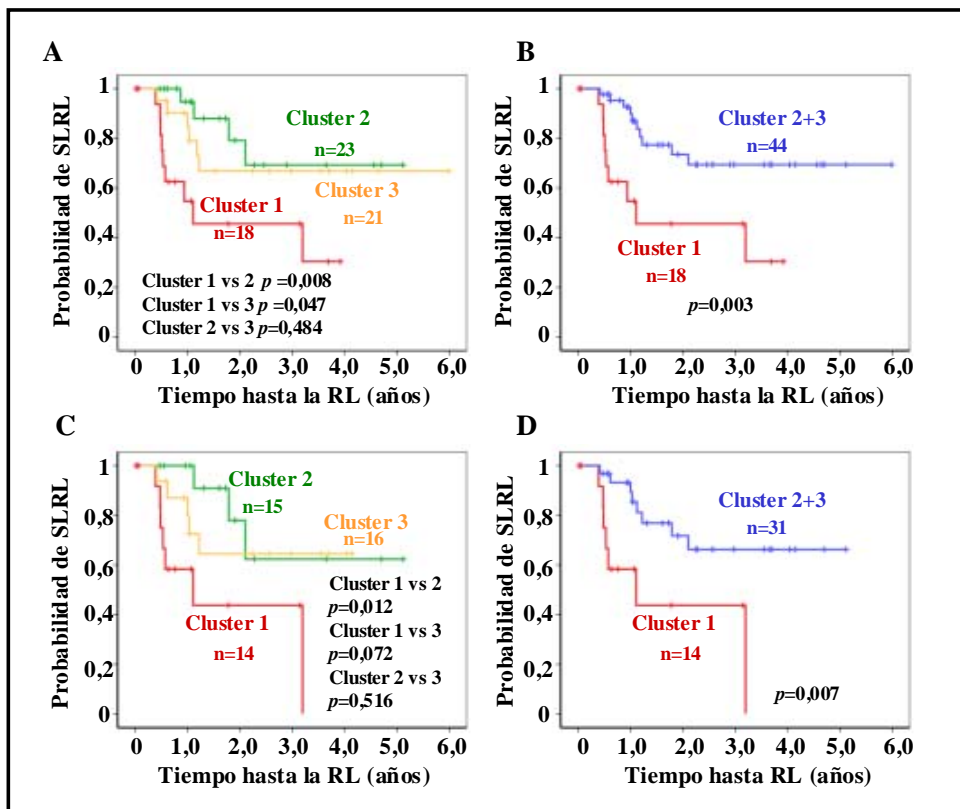


Figura 31. Curvas de Kaplan-Meier para la supervivencia libre de recidiva local (SLRL) en pacientes con CECC en función de su clasificación en clusters según su patrón de expresión A) Curvas de SLRL para los pacientes de los tres clusters. B) Curvas de SLRL para los pacientes del cluster 1 y para el resto de pacientes (cluster 2 y cluster 3). C y D) Curvas de SLRL dentro del subgrupo de pacientes que recibieron tratamiento genotóxico (excluyendo los quirúrgicos). La existencia de diferencias significativas fue establecida aplicando el test de Log-Rank.

La ausencia de diferencias significativas en la SLRL al comparar los pacientes del cluster 2 con los del 3, nos llevó a agrupar a ambos clusters en un mismo grupo para determinar si existían diferencias cuando se comparaban ambos grupos conjuntamente con los pacientes incluidos en el cluster 1. De este modo, se observó que los pacientes

## Resultados

del cluster 1 presentaban una mayor probabilidad de padecer una recidiva local que los pacientes del cluster 2 y 3 agrupados ( $p=0,003$ ) (Figura 31B).

Los análisis de Cox univariante ( $p=0,06$ ) y multivariante ( $p=0,03$ ) mostraron que la clasificación dentro de un cluster es un factor de riesgo independiente de recidiva local del tumor primario (Tabla 13A). No encontramos una asociación significativa entre la SLRL y la afectación ganglionar, la localización o el tamaño tumoral.

**Tabla 13. Tasa de riesgo (TR) de supervivencia libre de recidiva local (SLRL) para los pacientes con CECC en función del cluster en el que se clasifican sus tumores.**

	Supervivencia libre de recidiva local (SLRL)			
	Cox Univariante		Cox Multivariante	
	TR (IC 95%) <sup>a</sup>	p	TR (IC 95%) <sup>a</sup>	p
<b>A. Todos los pacientes</b>				
<b>Clasificación según cluster (cluster 1 vs cluster 2 y 3)</b>	3,6 (1,4-8,8)	0,06	4,2 (1,6-11,0)	<b>0,03*</b>
<b>Tamaño tumoral (T3-T4 vs T1-T2)</b>	1,2 (0,4-3,3)	0,76	1,1 (0,4-3,1)	0,90
<b>Afectación ganglionar (N+ vs N0)</b>	4,6 (0,6-34,8)	0,14	4,9 (0,6-42,6)	0,15
<b>Localización tumoral</b>	1,4 (0,5-3,6)	0,54	1,3 (0,43-3,7)	0,69
<b>B. Tratamiento genotóxico</b>				
<b>Clasificación según cluster (cluster 1 vs cluster 2 y 3)</b>	3,8 (1,4-10,6)	<b>0,01*</b>	4,7 (1,4-15,9)	<b>0,01</b>
<b>Tamaño tumoral (T3-T4 vs T1-T2)</b>	1,7 (0,4-3,7)	0,80	0,9 (0,3-3,2)	0,89
<b>Afectación ganglionar (N+ vs N0)</b>	3,2 (0,4-24,52)	0,26	3,25 (0,33-31,8)	0,31
<b>Localización tumoral</b>	1,03 (0,33-3,3)	0,95	1,4 (0,3-5,8)	0,65

<sup>a</sup> TR= Riesgo de recidiva; IC 95%= Intervalo de confianza del 95%.

\* Existencia de diferencias significativas tras aplicar el modelo de Cox univariante o multivariante.

Con objeto de determinar la capacidad de la clasificación en clusters de predecir la respuesta tumoral al tratamiento genotóxico repetimos el análisis de SLRL incluyendo en el estudio sólo a aquellos pacientes que habían recibido tratamiento genotóxico (quimioradioterapia concomitante o quimioterapia de inducción seguida de RT o QRT), excluyendo a los pacientes quirúrgicos. De los 62 pacientes con biopsias disponibles, 45 siguieron un tratamiento genotóxico con QRT concomitante o QTI seguida de RT o QRT. El análisis de SLRL mostró que, dentro de los pacientes sometidos a un tratamiento genotóxico, aquellos que presentaban tumores en el cluster 1 tenían una mayor probabilidad de padecer una recidiva tumoral que los pacientes cuyos tumores se agrupaban en los clusters 2 y 3 (Figura 31 C-D).

El análisis de Cox univariante ( $p=0,011$ ) y multivariante ( $p=0,013$ ) mostró que la clasificación dentro de un cluster es un factor de riesgo independiente de recidiva local del tumor primario para los pacientes que recibieron un tratamiento genotóxico (Tabla 13B). En cambio, el tamaño tumoral, la afectación ganglionar o la localización no se asociaron de manera significativa con la SLRL.

El análisis de Kaplan-Meier mostró que los pacientes del cluster 3 presentaban mayor supervivencia global (SG) que los pacientes del cluster 1 o del cluster 2. No obstante, dichas diferencias no alcanzaron significación estadística (cluster 1 vs 3,  $p=0,06$ ; cluster 2 vs 3,  $p=0,150$ ; cluster 1 vs 2,  $p=0,52$ ) (Log-Rank test) (Figura 32A).

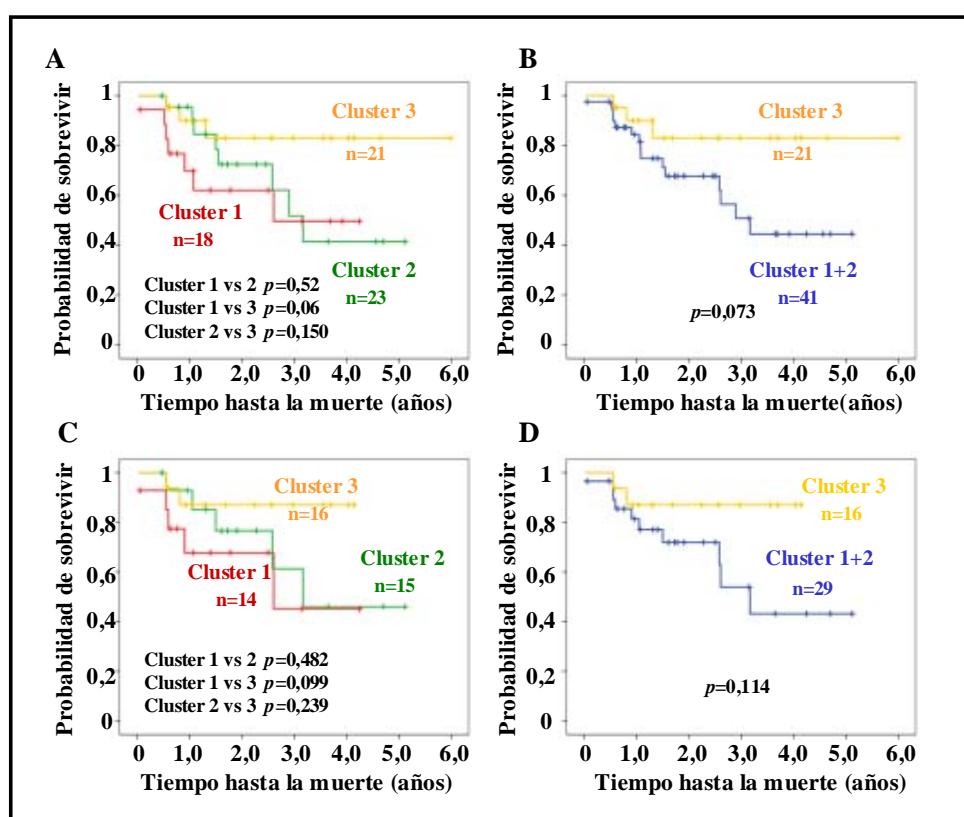


Figura 32. Curvas de Kaplan-Meier para la supervivencia global (SG) en pacientes con CECC en función de su clasificación en clusters según su patrón de expresión A) Curvas de SG para los pacientes de los tres clusters identificados. B) Curvas de SG para los pacientes del cluster 1 respecto al resto de pacientes (cluster 2 y cluster 3). C y D) Curvas de SG de los pacientes que recibieron tratamiento genotóxico (QRT concomitante o QTI seguida de RT/QRT), excluyendo a los quirúrgicos. La existencia de diferencias significativas fue establecida con el test de Log-Rank.

Al agrupar los pacientes del cluster 1 conjuntamente con los del cluster 2 observamos que su supervivencia global era menor que la de los pacientes del cluster 3 (Figura 32B); no obstante, estas diferencias no llegaron a ser estadísticamente significativas ( $p=0,073$ ) (Log-Rank test).

## Resultados

El análisis de Cox multivariante ( $p=0,05$ ) demostró que la clasificación de los pacientes en clusters es un factor de riesgo independiente de supervivencia global, siendo los pacientes del cluster 3 los que registraron mayor supervivencia global (Tabla 14A).

**Tabla 14. Tasas de riesgo (TR) de supervivencia global (SG) de los pacientes con CECC en función de la clasificación de sus tumores mediante cluster jerárquico.**

	Supervivencia Global (SG)			
	Cox Univariante		Cox Multivariante	
	TR (IC 95%) <sup>a</sup>	p	TR (IC 95%) <sup>a</sup>	p
<b>A. Todos los pacientes</b>				
<b>Clasificación según cluster (cluster 3 vs cluster 1 y 2)</b>	3,0 (0,9-10,2)	0,09	3,2 (0,9-11,2)	<b>0,05*</b>
<b>Tamaño tumoral (T3-T4 vs T1-T2)</b>	1,3 (0,5-3,7)	0,58	1,0 (0,4-2,9)	0,95
<b>Afectación ganglionar (N+ vs N0)</b>	27,5 (0,2-4889)	0,21	-----	0,97
<b>Localización tumoral</b>	1,6 (0,6-4,6)	0,34	1,2 (0,4-3,4)	0,76
<b>B. Tratamiento genotóxico</b>				
<b>Clasificación según cluster (cluster 3 vs cluster 1 y 2)</b>	3,2 (0,7-14,6)	0,16	3,2 (0,9-11,2)	0,08
<b>Tamaño tumoral (T3-T4 vs T1-T2)</b>	0,9 (0,2-3,3)	0,85	1,0 (0,4-3)	0,95
<b>Afectación ganglionar (N+ vs N0)</b>	26,6 (0,0-19867)	0,31	-----	0,97
<b>Localización tumoral</b>	1,9 (0,4-8,5)	0,42	1,2 (0,4-3,4)	0,80

<sup>a</sup>TR= Tasa de riesgo de muerte; IC 95%= Intervalo de confianza al 95%

\*Existencia de diferencias significativas tras aplicar el modelo de Cox univariante o multivariante.

De modo similar al procedimiento para evaluar la SLRL, analizamos la supervivencia global de sólo aquellos pacientes que habían recibido un tratamiento genotóxico, excluyendo los quirúrgicos. De nuevo, observamos que los pacientes del cluster 3 presentaban una mayor supervivencia global (SG) que los pacientes del cluster 2 o que los del cluster 1 (Figura 32C-D). Sin embargo, dichas diferencias no alcanzaron significación estadística (cluster 1 vs 3,  $p=0,10$ ; cluster 2 vs 3,  $p=0,24$ ; cluster 1 vs 2,  $p=0,48$ ) (Log-Rank test). La comparación de los pacientes del cluster 2 con el resto de pacientes, demostró que la supervivencia de los pacientes del cluster 3 era mayor que la de los pacientes del cluster 1 y 2 agrupados, siendo esta diferencia próxima a la significación estadística ( $p=0,114$ ) (Log Rank test, Figura 32D)). Los análisis de Cox univariante y multivariante mostraron que ninguna de las variables analizadas se asociaba significativamente con la supervivencia global de los pacientes. La clasificación en clusters fue la única variable que mostró una tendencia cercana a la significación estadística ( $p=0,08$ ) (Tabla 14B).

De modo consistente con los datos obtenidos en el análisis de SLRL, los pacientes del cluster 3 registraron mayor supervivencia global que el resto de los pacientes. A su vez, los pacientes del cluster 1 presentaron menor supervivencia global que los pacientes del cluster 3. Aunque en los pacientes del cluster 2 se observó una SLRL superior a los del cluster 1, ésta no se tradujo en un mayor supervivencia global, ya que el análisis de supervivencia global mostró que los pacientes del cluster 1 y del cluster 2 tenían unos valores de supervivencia similares e inferiores a los del cluster 3.

No se observó asociación significativa entre la respuesta tumoral a la quimioterapia de inducción ( $p=0,73$ ), la localización tumoral ( $p=0,306$ ), el tamaño tumoral ( $p=0,740$ ) o la afectación ganglionar ( $p=0,960$ ) con la clasificación de los tumores en clusters (test U de Mann-Whitney).

#### **2.4. GENES EXPRESADOS DIFERENCIALMENTE ENTRE LOS TUMORES DEL CLUSTER 1 Y LOS TUMORES DE LOS CLUSTERS 2 Y 3 AGRUPADOS**

Una vez identificado el cluster 1 como el grupo de tumores de pacientes con menor supervivencia libre de recidiva local analizamos los genes diferencialmente expresados en los tumores de este cluster, en comparación con los tumores incluidos en los clusters 2 y 3. El resultado de este análisis, que obtuvimos utilizando el paquete de software “Multtest” de Bioconductor, identificó 35 genes expresados diferencialmente en el cluster 1 respecto al cluster 2 y 3, con una probabilidad ajustada menor a 0,05 (Tabla 15 y Anexo 1). Por otro lado, utilizamos la herramienta GOstats de Bioconductor para identificar los términos biológicos sobre-representados a partir de los genes expresados diferencialmente. 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, aumentamos el número de genes incluidos en el análisis y la probabilidad de identificar un grupo de genes implicados en el mismo proceso biológico o vía de señalización. Este análisis detectó 525 genes expresados diferencialmente con una probabilidad estadística menor que 0,001, de los cuales 272 estaban sobre-expresados y 255 infra-expresados, con los que realizamos el análisis de ontologías.



## Resultados

Tabla 15. Genes expresados diferencialmente al comparar los tumores del cluster 1 con los tumores de los cluster 2 y 3 agrupados ( $p$  ajustada  $<0,05$ ).

Genes sobre-expresados	
<b>SERPINE1</b>	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
<b>THBS1</b>	thrombospondin 1
<b>INHBA</b>	inhibin, beta A
<b>MICAL2</b>	microtubule associated monoxygenase, calponin and LIM domain containing 2
<b>SERINC3</b>	serine incorporator 3
<b>PLOD2</b>	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2
<b>LGALS1</b>	lectin, galactoside-binding, soluble, 1 (galectin 1)
<b>FHL2</b>	four and a half LIM domains 2
<b>ITGB1</b>	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)
<b>SCG5</b>	secretogranin V (7B2 protein)
<b>P4HA1</b>	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I
<b>C7orf10</b>	chromosome 7 open reading frame 10
<b>PPFIBP1</b>	PTPRF interacting protein, binding protein 1 (liprin beta 1)
<b>GALNACT-2</b>	chondroitin sulfate GalNAcT-2
<b>TPST1</b>	tyrosylprotein sulfotransferase 1
<b>TNFRSF12A</b>	tumor necrosis factor receptor superfamily, member 12A
<b>CALU</b>	calumenin
<b>TNFAIP6</b>	tumor necrosis factor, alpha-induced protein 6
<b>CAV1</b>	caveolin 1, caveolae protein, 22kDa
<b>FNDC3B</b>	fibronectin type III domain containing 3B
<b>ATP2C1</b>	ATPase, Ca <sup>++</sup> transporting, type 2C, member 1
Genes infra-expresados	
<b>ATP10B</b>	ATPase, Class V, type 10B
<b>PPFIBP2</b>	PTPRF interacting protein, binding protein 2 (liprin beta 2)
<b>CYP2C18</b>	cytochrome P450, family 2, subfamily C, polypeptide 18
<b>CBR3</b>	carbonyl reductase 3
<b>KRT13</b>	keratin 13
<b>TGM3</b>	transglutaminase 3 (E polypeptide, protein-glutamine-gamma-glutamyltransferase)
<b>CYP2C9</b>	cytochrome P450, family 2, subfamily C, polypeptide 9
<b>RAB25</b>	RAB25, member RAS oncogene family
<b>DBI</b>	diazepam binding inhibitor (GABA receptor modulator, acyl-Coenzyme A binding protein)
<b>DUOX2</b>	dual oxidase 2
<b>ITIH4</b>	inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)
<b>SASH1</b>	SAM and SH3 domain containing 1
<b>HPGD</b>	hydroxyprostaglandin dehydrogenase 15-(NAD)
<b>NSF</b>	N-ethylmaleimide-sensitive factor

Los genes sobre-expresados en los tumores del cluster 1 participan en los siguientes procesos biológicos de la base de datos GO: señalización a través de integrinas, adhesión celular, desarrollo de tejidos y motilidad celular (Tabla 16). Además, la mayoría de proteínas para las que codifican estos genes forman parte de componentes celulares como el aparato de Golgi, el retículo endoplasmático, la matriz extracelular y el complejo laminina (Tabla 16). Por otra parte, el análisis utilizando la base de datos KEGG muestra sobre-representadas las rutas de señalización que incluyen los receptores de la matriz extracelular, las adhesiones focales, la patogénesis del carcinoma

de pulmón de células pequeñas (microcítico), la comunicación celular y la regulación del citoesqueleto de actina.

Por otra parte, los genes infra-expresados en los tumores del cluster 1 participan en procesos biológicos como la queratinización, la morfogénesis de la epidermis y la diferenciación de los queratinocitos. El análisis de los genes infra-expresados en el cluster 1 también muestra sobre-representación de las vías de biosíntesis de esteroides y el metabolismo de la  $\beta$ -alanina descritas en la base de datos KEGG (Tabla 16).

**Tabla 16. Términos biológicos sobre-representados en los tumores del cluster 1 respecto a los tumores de los cluster 2 y 3 agrupados.**

<b>Genes sobre-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Adhesión celular
Vía de señalización mediada por integrinas
Desarrollo tisular
Procesos de desarrollo
Motilidad celular
<b><u>Componente celular (<math>p &lt; 0,001</math>)</u></b>
Matriz extracelular
Retículo endoplasmático
Complejo laminina
Membrana basal
Aparato de Golgi
<b><u>Vías KEGG (<math>p &lt; 0,001</math>)</u></b>
Adhesiones focales
Interacción con receptores de la membrana extracelular
Carcinoma de pulmón de células pequeñas
Comunicación celular
Regulación del citoesqueleto de actina
<b>Genes infra-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Queratinización
Morfogénesis de la epidermis
Diferenciación queratinocitos
<b><u>Vías KEGG (<math>p &lt; 0,001</math>)</u></b>
Metabolismo $\beta$ -alanina
Biosíntesis de esteroides

## Resultados

### 2.5. GENES EXPRESADOS DIFERENCIALMENTE ENTRE LOS TUMORES DEL CLUSTER 3 Y LOS TUMORES DE LOS CLUSTERS 1 Y 2 AGRUPADOS

El análisis de cluster identificó al cluster 3 como el grupo de pacientes que tenían mayor supervivencia global. Utilizamos “Multtest” para identificar los genes diferencialmente expresados en los tumores del cluster 3, comparándolos con los tumores de los clusters 1 y 2 agrupados.

Como resultado del análisis de expresión diferencial obtuvimos 51 genes sobre-expresados en el cluster 3, en comparación con los de los clusters 1 y 2, con una probabilidad ajustada menor que 0,05 (Tabla 17 y Anexo 2).

Tabla 17. Genes sobre-expresados en los tumores del cluster 3 respecto a los de los clusters 1 y 2 agrupados (*p* ajustada < 0,05 ).

Genes sobre-expresados	
<b>HOP</b>	homeodomain-only protein
<b>LYPD3</b>	LY6/PLAUR domain containing 3
<b>SLURP1</b>	secreted LY6/PLAUR domain containing 1
<b>SPINK5</b>	serine peptidase inhibitor, Kazal type 5
<b>CRCT1</b>	cysteine-rich C-terminal 1
<b>KLK8</b>	kallikrein-related peptidase 8
<b>CLIC3</b>	chloride intracellular channel 3
<b>SLC39A2</b>	solute carrier family 39 (zinc transporter), member 2
<b>CSTA</b>	cystatin A (stefin A)
<b>RHCG</b>	Rh family, C glycoprotein
<b>GDPD3</b>	glycerophosphodiester phosphodiesterase domain containing 3
<b>IL1RN</b>	interleukin 1 receptor antagonist
<b>KLK7</b>	kallikrein-related peptidase 7
<b>AQP3</b>	aquaporin 3 (Gill blood group)
<b>SULT2B1</b>	sulfotransferase family, cytosolic, 2B, member 1
<b>DUOX1</b>	dual oxidase 1
<b>SCEL</b>	sciellin
<b>TMPRSS11D</b>	transmembrane protease, serine 11D
<b>IL1F5</b>	interleukin 1 family, member 5 (delta)
<b>SLPI</b>	secretory leukocyte peptidase inhibitor
<b>PPL</b>	periplakin
<b>IVL</b>	involucrin
<b>TGM1</b>	transglutaminase 1 (K polypeptide epidermal type I, protein-glutamine-gamma-glutamyltransferase)
<b>CSTB</b>	cystatin B (stefin B)
<b>SFT2D2</b>	SFT2 domain containing 2
<b>EPS8L1</b>	EPS8-like 1
<b>DEFB1</b>	defensin, beta 1
<b>KLK11</b>	kallikrein-related peptidase 11
<b>CD24</b>	CD24 molecule
<b>HSPB8</b>	heat shock 22kDa protein 8
<b>LGALS7</b>	lectin, galactoside-binding, soluble, 7 (galectin 7)
<b>PLK2</b>	polo-like kinase 2 (Drosophila)
<b>MOSPD1</b>	motile sperm domain containing 1
<b>CDA</b>	cytidine deaminase
<b>KLK10</b>	kallikrein-related peptidase 10
<b>LTB4R</b>	leukotriene B4 receptor
<b>KLK6</b>	kallikrein-related peptidase 6
<b>KLK12</b>	kallikrein-related peptidase 12
<b>S100A12</b>	S100 calcium binding protein A12
<b>SPRR1A</b>	small proline-rich protein 1 <sup>a</sup>
<b>KLK5</b>	kallikrein-related peptidase 5
<b>KRT16</b>	keratin 16 (focal non-epidermolytic palmoplantar keratoderma)

<b>KCNK7</b>	potassium channel, subfamily K, member 7
<b>NPEPPS</b>	aminopeptidase puromycin sensitive
<b>BBOX1</b>	butyrobetaine (gamma), 2-oxoglutarate dioxygenase (gamma-butyrobetaine hydroxylase) 1
<b>P11</b>	26 serine protease
<b>S100A9</b>	S100 calcium binding protein A9
<b>CEACAM7</b>	carcinoembryonic antigen-related cell adhesion molecule 7
<b>KRT6C</b>	keratin 6C
<b>RAB25</b>	RAB25, member RAS oncogene family

Utilizando la herramienta GOstats, identificamos los términos biológicos sobre-representados a partir de los genes expresados diferencialmente en los tumores del cluster 3. En el análisis de anotaciones incluimos aquellos genes que presentaban una probabilidad estadística no ajustada menor que 0,001. Detectamos 512 genes expresados diferencialmente con una p no ajustada  $<0,001$ , de los cuales 311 estaban sobre-expresados y 201 infra-expresados, a partir de los que realizamos el análisis de ontologías.

Los genes sobre-expresados en los tumores del cluster 3 participan en los siguientes procesos biológicos de la base de datos Gene Ontology: diferenciación de los queratinocitos, queratinización, morfogénesis de la epidermis y desarrollo de órganos (Tabla 18). Además, los genes sobre-expresados en el cluster 3 se encuentran localizados en el citoplasma o forman parte de estructuras celulares como los filamentos intermedios o los desmosomas.

Por otro lado, los genes infra-expresados en los tumores del cluster 3 participan en los siguientes procesos biológicos: reparación del DNA por “mismatch”, replicación del DNA, catabolismo de la leucina y regulación negativa de la recombinación del DNA.

No encontramos ninguna ruta de señalización de KEGG que alcanzase diferencias significativas tanto para los genes sobre-expresados como para los infra-expresados.

## Resultados

Tabla 18. Términos biológicos sobre-representados en los tumores del cluster 3 respecto a los tumores de los clusters 1 y 2 agrupados.

<b>Genes sobre-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Diferenciación queratinocitos
Queratinización
“Cross-linking” de péptidos
Morfogénesis de la epidermis
Desarrollo epidermis
Desarrollo de los órganos
Desarrollo y biogénesis de la envuelta celular
<b><u>Componente celular (<math>p &lt; 0,001</math>)</u></b>
Citoplasma
Desmosoma
Filamento intermedio
Cubierta cornificada
<b>Genes infra-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Reparación del DNA por mismatch
Replicación del DNA, DNA dependiente
Catabolismo leucina
Determinación de la vida media adulta
Regulación negativa de la recombinación del DNA
Hipermutación somática de los genes de inmunoglobulinas
Metabolismo de la guanosina

### 2.6. GENES EXPRESADOS DIFERENCIALMENTE ENTRE LOS TUMORES DEL CLUSTER 2 Y LOS TUMORES DE LOS CLUSTERS 1 Y 3 AGRUPADOS

El análisis de cluster muestra que los pacientes con tumores del cluster 2 presentan una SLRL similar a la de los pacientes con tumores del cluster 3 y superior a la del cluster 1; sin embargo, la supervivencia global de estos pacientes se aproxima más a la de los pacientes portadores de tumores del cluster 1 y es claramente menor que la de los del cluster 3.

Como resultado del análisis de expresión diferencial obtuvimos 21 genes expresados diferencialmente en el cluster 2 respecto al los de los clusters 1 y 3 combinados, con una probabilidad ajustada menor que 0,05 (Tabla 19 y Anexo 3).

**Tabla 19. Genes expresados diferencialmente en los tumores del cluster 2 respecto a los de los clusters 1 y 3 agrupados ( $p$  ajustada  $< 0,05$ ).**

Genes sobre-expresados	
<b>KRT19</b>	keratin 19
<b>MCCC1</b>	methylcrotonoyl-Coenzyme A carboxylase 1 (alpha)
<b>LMO4</b>	LIM domain only 4
<b>C1orf115</b>	chromosome 1 open reading frame 115
<b>DVL3</b>	dishevelled, dsh homolog 3 (Drosophila)
<b>KRT15</b>	keratin 15
Genes infra-expresados	
<b>PLEK2</b>	pleckstrin 2
<b>PLK2</b>	polo-like kinase 2 (Drosophila)
<b>RGS20</b>	regulator of G-protein signaling 20
<b>FHL2</b>	four and a half LIM domains 2
<b>VEGFC</b>	vascular endothelial growth factor C
<b>EREG</b>	epiregulin
<b>CAB39</b>	calcium binding protein 39
<b>PTGS1</b>	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
<b>PDLIM2</b>	PDZ and LIM domain 2 (mystique)
<b>CDA</b>	cytidine deaminase
<b>CD59</b>	CD59 molecule, complement regulatory protein
<b>UPP1</b>	uridine phosphorylase 1
<b>STK17A</b>	serine/threonine kinase 17a
<b>ATP6V1D</b>	ATPase, H <sup>+</sup> transporting, lysosomal 34kDa, V1 subunit D
<b>ACTN1</b>	actinin, alpha 1

Utilizando la herramienta GOstats determinamos los términos biológicos que estaban sobre-representados a partir de los genes expresados diferencialmente en los tumores del cluster 2 (Tabla 20). En el análisis de anotaciones incluimos aquellos genes que presentaban una probabilidad estadística no ajustada menor a 0,001. Detectamos 479 genes expresados diferencialmente con una  $p < 0,001$ , de los cuales 197 estaban sobre-expresados y 282 infra-expresados, a partir de los que realizamos el análisis de ontologías.

Los genes sobre-expresados en los tumores del cluster 2 participan en los siguientes procesos biológicos de la base de datos Gene Ontology: el catabolismo de la leucina, el metabolismo de los aminoácidos ramificados, la respuesta al daño al DNA y la morfogénesis de la ramificación de estructuras tubulares. Entre los genes sobre-expresados también encontramos sobre-representado el proceso de patogénesis del carcinoma colorectal descrita en KEGG.

Los genes infra-expresados en los tumores del cluster 2 participan en los siguientes procesos biológicos: la angiogénesis, la cascada de señalización proteína quinasa, el

## Resultados

desarrollo de estructuras anatómicas, la organización y biogénesis del citoesqueleto y la formación de adhesiones focales.

**Tabla 20. Términos biológicos sobre-representados en los tumores del cluster 2 respecto a los tumores del cluster 1 y 3 agrupados**

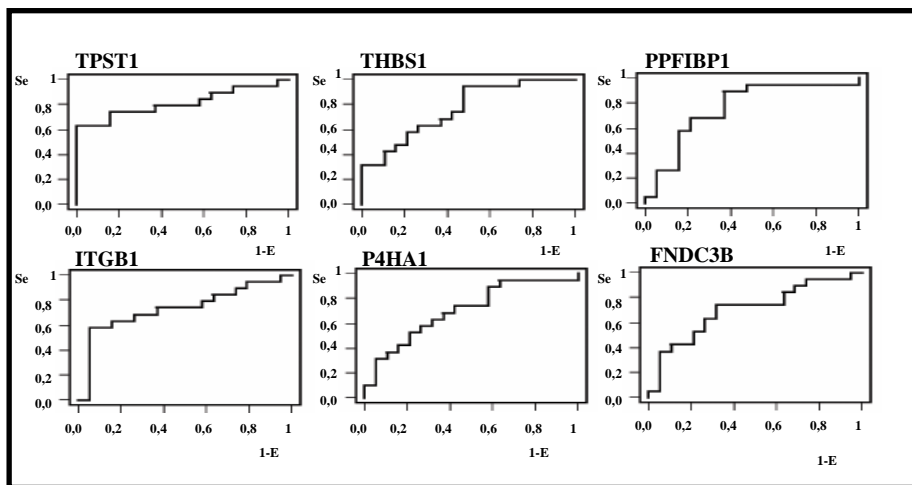
<b>Genes sobre-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Catabolismo de la leucina
Metabolismo de los aminoácidos ramificados
Respuesta al daño al DNA
Morfogénesis de la ramificación de estructuras tubulares
<b><u>Vías KEGG (<math>p &lt; 0,001</math>)</u></b>
Cancer colorectal
<b>Genes infra-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Angiogénesis
Cascada de señalización protein quinasa
Desarrollo de las estructuras anatómicas
Organización y biogénesis del citoesqueleto
Formación adhesiones focales
<b><u>Componente celular (<math>p &lt; 0,001</math>)</u></b>
Citoesqueleto de actina
Citoesqueleto
Componentes de las fibras contráctiles

### **2.7. GENES EXPRESADOS DIFERENCIALMENTE EN EL CLUSTER 1 Y EN EL CLUSTER 3 RELACIONADOS CON LA RECIDIVA LOCAL A 2 AÑOS**

Con objeto determinar los genes con mayor capacidad predictiva de la recidiva tumoral de entre el conjunto de genes identificados en los apartados 2.4. y 2.5. (35 cluster 1 y 51 cluster 3), realizamos un análisis de curvas ROC para todos los genes identificados previamente con una  $p$  ajustada  $< 0,05$ . Las curvas ROC representan la sensibilidad de un marcador determinado frente a los falsos positivos (la inversa de la especificidad) a la hora de predecir el resultado de una prueba, que en nuestro caso es la aparición de una recidiva tumoral. Los genes que obtienen mayor rendimiento global para predecir la recidiva tumoral son los que presentan mayor área bajo la curva (AUC).

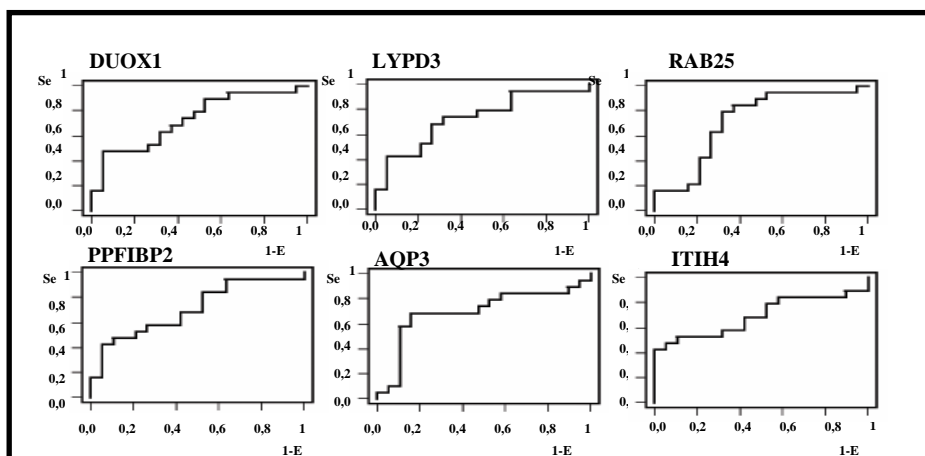
El análisis de recidiva tumoral se llevó a cabo incluyendo aquellos pacientes con un seguimiento mínimo de dos años, excluyendo a aquellos pacientes con un tiempo de

seguimiento menor. Treinta y ocho pacientes tenían un seguimiento mayor a 2 años, de los cuales 19 habían padecido una recidiva del tumor primario mientras que otros 19 permanecían libres de recidiva durante el mismo período de tiempo. Analizamos la sensibilidad y especificidad de cada gen para predecir la recidiva tumoral, durante el período de 2 años de seguimiento, generando curvas ROC para cada uno de los genes seleccionados. Los genes que obtuvieron mayores valores de AUC y una asociación directa con la aparición de recidivas tumorales a los 2 años (en el sentido de una correlación entre mayor expresión y, mayor probabilidad de recidiva tumoral) fueron: TPST1, THBS1, PPFIBP1, ITGB1, P4HA1 y FNDC3B (Figura 33).



**Figura 33. Genes cuya expresión está directamente relacionada con la recidiva tumoral. Mayores niveles de expresión de estos genes aumentan la probabilidad de padecer una recidiva tumoral durante un período de dos años. 1-E: 1-Especificidad. Se: Sensibilidad**

Los genes que presentaron una mayor AUC y una asociación inversa con la recidiva tumoral a los 2 años (en el sentido de una correlación entre mayor expresión y menor probabilidad de recidiva tumoral ) fueron: DUOX1, LYPD3, RAB25, PPFIBP2, AQP3 y ITIH4 (Figura 34).



**Figura 34. Genes cuya expresión está inversamente relacionada con la recidiva tumoral. Mayores niveles de expresión de estos genes disminuyen la probabilidad de padecer una recidiva tumoral durante un período de dos años. 1-E: 1-Especificidad. Se: Sensibilidad.**



## Resultados

### 2.8. GENES EXPRESADOS DIFERENCIALMENTE EN EL CLUSTER 1 Ó EN EL CLUSTER 3 RELACIONADOS CON LA MUERTE DEL PACIENTE A LOS 3 AÑOS

El análisis de supervivencia incluye a los pacientes que tenían un seguimiento mínimo de tres años, excluyendo a aquellos con un tiempo de seguimiento menor. De los treinta pacientes con un seguimiento mayor a 3 años, 18 fallecieron como consecuencia de la enfermedad, mientras que 12 permanecían vivos después de los 3 años de seguimiento. Analizamos la sensibilidad y especificidad de cada gen para predecir la supervivencia de los pacientes a los 3 años, generando curvas ROC para todos aquellos genes que en el análisis de expresión diferencial habían presentado una  $p$  ajustada  $<0,05$ . Los genes con mayor valor de área bajo la curva (AUC) y una asociación directa con la muerte de los pacientes a los 3 años (en el sentido de mayor expresión, más probabilidad de muerte) fueron: THBS1, FNDC3B, ATP2C1, P4HA1 y PPFIBP1 (Figura 35). Los genes que presentaron un AUC mayor y una asociación inversa con la muerte de los pacientes a los 3 años (en el sentido de mayor expresión, menor probabilidad de muerte) fueron: LYPD3, KRT6C, EPS8L1, RAB25, DUOX1 y GDPD3 (Figura 36).

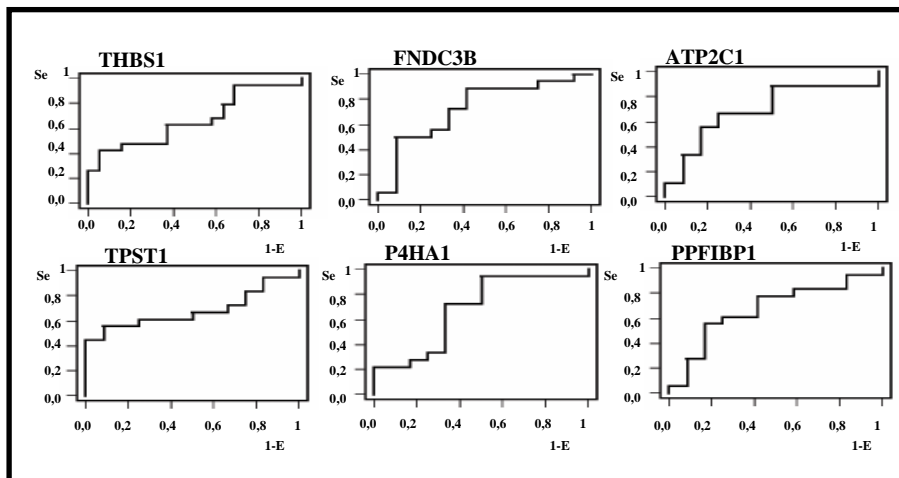
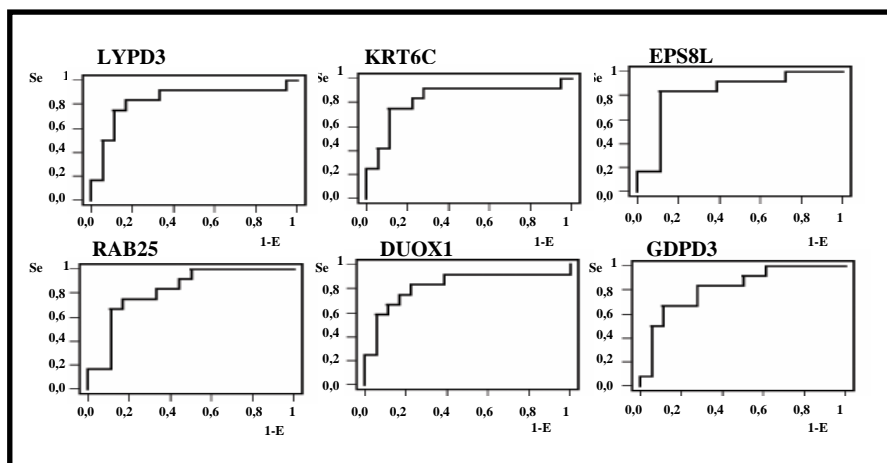


Figura 35. Genes cuya expresión se asocia con la muerte del paciente. Mayores niveles de expresión aumentan la probabilidad de que el paciente muera durante un período de seguimiento de 3 años. 1-E: 1-Especificidad. Se: Sensibilidad



**Figura 36. Genes cuya expresión está inversamente relacionada con la muerte del paciente. Mayores niveles de expresión disminuyen la probabilidad de que el paciente muera durante un período de seguimiento de 3 años. 1-E: 1-Especificidad. Se: Sensibilidad**

## 2.9. GENES EXPRESADOS DIFERENCIALMENTE EN FUNCIÓN DE LA RECIDIVA LOCAL A LOS 2 AÑOS

Con objeto de identificar aquellos genes más directamente relacionados con la recidiva tumoral, independientemente de la clasificación de los tumores mediante análisis de cluster jerárquico, realizamos un estudio de expresión diferencial comparando los pacientes que habían sufrido recidivas con aquellos que permanecían libres de recidiva a los 2 años. Únicamente se incluyeron en el estudio aquellos pacientes que tenían un seguimiento mínimo de dos años. De los treinta y ocho pacientes con un seguimiento mayor a 2 años, 19 habían padecido una recidiva del tumor primario, mientras que otros 19 permanecían libres de recidiva durante el mismo período de tiempo.

En el análisis de expresión diferencial, aplicando “Multtest”, no se identificó gen alguno con una  $p$  ajustada menor que 0,05. No obstante, se identificaron 57 genes (33 sobre-expresados y 24 infra-expresados) con una probabilidad no ajustada menor que 0,001 (Tabla 21 y Anexo 3). El número relativamente bajo de genes identificados fue insuficiente para determinar la posible participación de procesos biológicos o de vías de señalización implicadas en la recidiva tumoral a 2 años.

## Resultados

Tabla 21. Genes expresados diferencialmente al comparar los tumores de pacientes que sufrieron una recidiva tumoral con los tumores de pacientes libres de recidiva a los 2 años ( $p < 0,001$ ).

### Genes sobre-expresados

<b>MAT2B</b>	methionine adenosyltransferase II, beta
<b>FTH1</b>	ferritin, heavy polypeptide 1
<b>RNF130</b>	ring finger protein 130
<b>CD209</b>	CD209 molecule
<b>MED31</b>	mediator complex subunit 31
<b>ARSB</b>	arylsulfatase B
<b>CDC42EP3</b>	CDC42 effector protein (Rho GTPase binding) 3
<b>EPOR</b>	erythropoietin receptor
<b>FADS1</b>	fatty acid desaturase 1
<b>SLC12A3</b>	solute carrier family 12 (sodium/chloride transporters), member 3
<b>PPP2R3C</b>	protein phosphatase 2 (formerly 2A), regulatory subunit B", gamma
<b>EAPP</b>	E2F-associated phosphoprotein
<b>PPID</b>	peptidylprolyl isomerase D (cyclophilin D)
<b>CDC23</b>	cell division cycle 23 homolog (S. cerevisiae)
<b>TPST1</b>	tyrosylprotein sulfotransferase 1
<b>RPL24</b>	ribosomal protein L24
<b>RARS</b>	arginyl-tRNA synthetase
<b>MED7</b>	mediator complex subunit 7
<b>KIR2DS1</b>	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 1
<b>GABARAPL2</b>	GABA(A) receptor-associated protein-like 2
<b>PTP4A1</b>	protein tyrosine phosphatase type IVA, member 1
<b>G3BP1</b>	GTPase activating protein (SH3 domain) binding protein 1
<b>FOXD1</b>	forkhead box D1
<b>HSPC171</b>	HSPC171 protein
<b>MRPL22</b>	mitochondrial ribosomal protein L22
<b>ATP8B4</b>	ATPase, Class I, type 8B, member 4
<b>AKR1B1</b>	aldo-keto reductase family 1, member B1 (aldose reductase)
<b>RABGGTB</b>	Rab geranylgeranyltransferase, beta subunit
<b>AP4S1</b>	adaptor-related protein complex 4, sigma 1 subunit
<b>YIPF5</b>	Yip1 domain family, member 5
<b>NR3C1</b>	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
<b>CLDN16</b>	claudin 16
<b>IFNAR1</b>	interferon (alpha, beta and omega) receptor 1

### Genes infra-expresados

<b>RRP1</b>	ribosomal RNA processing 1 homolog (S. cerevisiae)
<b>STAP2</b>	signal transducing adaptor family member 2
<b>POLR1B</b>	polymerase (RNA) I polypeptide B, 128kDa
<b>INTS6</b>	integrator complex subunit 6
<b>STK11</b>	serine/threonine kinase 11
<b>MAP2K2</b>	mitogen-activated protein kinase kinase 2
<b>UBR4</b>	ubiquitin protein ligase E3 component n-recognin 4
<b>PLXNA2</b>	plexin A2
<b>ANAPC5</b>	anaphase promoting complex subunit 5
<b>WIPF2</b>	WAS/WASL interacting protein family, member 2
<b>TIMELESS</b>	timeless homolog (Drosophila)
<b>C7orf43</b>	chromosome 7 open reading frame 43
<b>ATG4B</b>	ATG4 autophagy related 4 homolog B (S. cerevisiae)
<b>TNPO2</b>	transportin 2 (importin 3, karyopherin beta 2b)
<b>FAM128B</b>	family with sequence similarity 128, member B
<b>SLC24A1</b>	solute carrier family 24 (sodium/potassium/calcium exchanger), member 1
<b>IPO13</b>	importin 13
<b>ATP8B1</b>	ATPase, Class I, type 8B, member 1
<b>POLR2A</b>	polymerase (RNA) II (DNA directed) polypeptide A, 220kDa
<b>PYHIN1</b>	pyrin and HIN domain family, member 1
<b>FGFR1</b>	fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)
<b>LRRC40</b>	leucine rich repeat containing 40
<b>TNPO2</b>	transportin 2 (importin 3, karyopherin beta 2b)
<b>NEBL</b>	nebullette

**2.10. GENES EXPRESADOS DIFERENCIALMENTE EN TUMORES EN FUNCIÓN DE LA SUPERVIVENCIA A 3 AÑOS**

Con objeto de identificar aquellos genes más directamente relacionados con la supervivencia de los pacientes, realizamos un estudio de expresión diferencial comparando los pacientes fallecidos con aquellos que permanecían vivos a los 3 años de seguimiento. Incluimos en el estudio aquellos pacientes que contaban con un período de seguimiento mínimo de tres años. De los treinta pacientes con un tiempo de seguimiento mayor a 3 años, 18 fallecieron como consecuencia de la enfermedad, mientras que 12 permanecían vivos a los 3 años.

Ningún gen alcanzó una *p* ajustada menor que 0,05. Se identificaron 83 genes con una *p* no ajustada <0,001 (26 sobre-expresados y 57 infra-expresados) (Tabla 22 y Anexo 4). Sin embargo, este número resultó excesivamente pequeño para realizar el análisis de ontologías con el paquete de software GOstats y determinar los procesos biológicos y vías de señalización que se encontraban alteradas.

**Tabla 22. Genes sobre-expresados al comparar los tumores de pacientes fallecidos con los tumores de pacientes vivos a los 3 años de seguimiento (*p*<0,001).**

<b>Genes sobre-expresados</b>	
<b>TNIK</b>	TRAF2 and NCK interacting kinase
<b>MAT2B</b>	methionine adenosyltransferase II, beta
<b>MMP8</b>	matrix metalloproteinase 8 (neutrophil collagenase)
<b>ENPP4</b>	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)
<b>NRGN</b>	neurogranin (protein kinase C substrate, RC3)
<b>AKR1B1</b>	aldo-keto reductase family 1, member B1 (aldose reductase)
<b>ABCG2</b>	ATP-binding cassette, sub-family G (WHITE), member 2
<b>FAM49A</b>	family with sequence similarity 49, member A
<b>FUT8</b>	fucosyltransferase 8 (alpha (1,6) fucosyltransferase)
<b>ADEF</b>	adipose differentiation-related protein
<b>WDR48</b>	WD repeat domain 48
<b>ELOVL2</b>	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2
<b>OLR1</b>	oxidized low density lipoprotein (lectin-like) receptor 1
<b>CCR2</b>	chemokine (C-C motif) receptor 2
<b>HSD11B1</b>	hydroxysteroid (11-beta) dehydrogenase 1
<b>HBG2</b>	hemoglobin, gamma G
<b>PCLO</b>	piccolo (presynaptic cytomatrix protein)
<b>ZXDB</b>	zinc finger, X-linked, duplicated B
<b>HBG1</b>	hemoglobin, gamma A
<b>KLRG1</b>	killer cell lectin-like receptor subfamily G, member 1
<b>TNIK</b>	TRAF2 and NCK interacting kinase
<b>ZMYM6</b>	zinc finger, MYM-type 6
<b>HBD</b>	hemoglobin, delta
<b>HBQ1</b>	hemoglobin, theta 1
<b>ADCY1</b>	adenylate cyclase 1 (brain)
<b>RAB6B</b>	RAB6B, member RAS oncogene family

## Resultados

### Genes infra-expresados

<b>STRN4</b>	striatin, calmodulin binding protein 4
<b>ACTN4</b>	actinin, alpha 4
<b>UBR4</b>	ubiquitin protein ligase E3 component n-recognin 4
<b>TNPO2</b>	transportin 2 (importin 3, karyopherin beta 2b)
<b>LMNA</b>	lamin A/C
<b>GNA11</b>	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
<b>DNAJC11</b>	DnaJ (Hsp40) homolog, subfamily C, member 11
<b>CC2D1A</b>	coiled-coil and C2 domain containing 1A
<b>MAP3K6</b>	mitogen-activated protein kinase kinase kinase 6
<b>KIAA1026</b>	kazrin
<b>ATG4B</b>	ATG4 autophagy related 4 homolog B (S. cerevisiae)
<b>LAD1</b>	ladinin 1
<b>ZNF668</b>	zinc finger protein 668
<b>NLRP1</b>	NLR family, pyrin domain containing 1
<b>CNOT3</b>	CCR4-NOT transcription complex, subunit 3
<b>SMARCA4</b>	SWI/SNF related, matrix associated, actin depend regulator of chromatin, subfam a, m4
<b>C1orf116</b>	chromosome 1 open reading frame 116
<b>PDLIM5</b>	PDZ and LIM domain 5
<b>NAPA</b>	N-ethylmaleimide-sensitive factor attachment protein, alpha
<b>MAP2K2</b>	mitogen-activated protein kinase kinase 2
<b>SLC24A6</b>	solute carrier family 24 (sodium/potassium/calcium exchanger), member 6
<b>ARHGEF4</b>	Rho guanine nucleotide exchange factor (GEF) 4
<b>ZBTB7A</b>	zinc finger and BTB domain containing 7A
<b>KRT5</b>	keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne)
<b>CIDEB</b>	cell death-inducing DFFA-like effector b
<b>MINK1</b>	misshapen-like kinase 1 (zebrafish)
<b>PLA2G4A</b>	phospholipase A2, group IVA (cytosolic, calcium-dependent)
<b>SGTA</b>	small glutamine-rich tetratricopeptide repeat (TPR)-containing, alpha
<b>AKAP8L</b>	A kinase (PRKA) anchor protein 8-like
<b>SH3GL1</b>	SH3-domain GRB2-like 1
<b>LETM1</b>	leucine zipper-EF-hand containing transmembrane protein 1
<b>SUPT5H</b>	suppressor of Ty 5 homolog (S. cerevisiae)
<b>LRRFIP1</b>	leucine rich repeat (in FLII) interacting protein 1
<b>PLEKHM2</b>	pleckstrin homology domain containing, family M (with RUN domain) member 2
<b>RHOD</b>	ras homolog gene family, member D
<b>FBXL11</b>	F-box and leucine-rich repeat protein 11
<b>NDUFA4L2</b>	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2
<b>RICS</b>	Rho GTPase-activating protein
<b>TRIM29</b>	tripartite motif-containing 29
<b>TUBB2A</b>	tubulin, beta 2A
<b>PPP2R1A</b>	protein phosphatase 2 (formerly 2A), regulatory subunit A , alpha isoform
<b>PPP5C</b>	protein phosphatase 5, catalytic subunit
<b>CAMKK2</b>	calcium/calmodulin-dependent protein kinase kinase 2, beta
<b>PRKD2</b>	protein kinase D2
<b>TNFRSF25</b>	tumor necrosis factor receptor superfamily, member 25
<b>PPP2R4</b>	protein phosphatase 2A activator, regulatory subunit 4
<b>MINK1</b>	misshapen-like kinase 1 (zebrafish)
<b>UBAP2L</b>	ubiquitin associated protein 2-like
<b>NOSIP</b>	nitric oxide synthase interacting protein
<b>P53AIP1</b>	p53-regulated apoptosis-inducing protein 1
<b>DOCK6</b>	dedicator of cytokinesis 6
<b>TIAF1</b>	TGFB1-induced anti-apoptotic factor 1
<b>SFRS8</b>	splicing factor, arginine/serine-rich 8
<b>NOC2L</b>	nucleolar complex associated 2 homolog (S. cerevisiae)
<b>CEP164</b>	centrosomal protein 164kDa
<b>CSRP1</b>	cysteine and glycine-rich protein 1
<b>CCNL2</b>	cyclin L2

**2.11. GENES EXPRESADOS DIFERENCIALMENTE EN FUNCIÓN DE LA RESPUESTA TUMORAL A LA QUIMIOTERAPIA DE INDUCCIÓN**

En este análisis, se incluyeron únicamente aquellos pacientes que recibieron quimioterapia de inducción, excluyendo a los tratados desde un inicio con quimioradioterapia concomitante.

De los 62 pacientes incluidos en el estudio, 45 recibieron quimioterapia de inducción. De estos 45, 25 presentaron una respuesta tumoral a la quimioterapia de inducción superior al 50% (grupo repondedor), mientras que 20 pacientes tuvieron una respuesta tumoral al tratamiento inferior al 50% (grupo no respondedor).

En el análisis de expresión diferencial, no se identificó ningún gen con una *p* ajustada <0,05. No obstante, identificamos 6 genes con una *p* no ajustada <0,001 (Tabla 23 y Anexo 6).

**Tabla 23. Genes expresados diferencialmente en función de la respuesta a la quimioterapia de inducción (*p*<0,001).**

<b>Genes sobre-expresados</b>	
<b>PPFIA1</b>	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF),interacting protein (liprin), $\alpha$ 1
<b>CBARA1</b>	calcium binding atopy-related autoantigen 1
<b>PCSK6</b>	proprotein convertase subtilisin/kexin type 6
<b>KIDINS220</b>	kinase D-interacting substrate of 220 kDa
<b>MAN1A2</b>	mannosidase, alpha, class 1A, member 2
<b>Genes infra-expresados</b>	
<b>CTDSPL</b>	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase-Like

**2.12. EXPRESIÓN DIFERENCIAL ENTRE TUMOR Y MUCOSA NORMAL**

Para identificar los genes relacionados con el proceso de patogénesis del CECC, aplicamos “Multtest” y llevamos a cabo un estudio de expresión diferencial comparando el perfil de expresión de las 62 biopsias tumorales con el perfil de expresión de las 5 mucosas normales. Se identificaron 48 genes con una probabilidad estadística ajustada < 0,05 (Tabla 24 y Anexo 7).

Utilizando la herramienta GOstats de Bioconductor, determinamos los términos biológicos que estaban sobre-representados a partir de los genes expresados diferencialmente. Este análisis fue realizado seleccionando aquellos genes que presentaban una probabilidad estadística no ajustada <0,001. La comparación de los

## Resultados

niveles de expresión de las 62 biopsias tumorales con la de las 5 mucosas normales identificó 961 genes con una  $p$  no ajustada  $<0,001$  (463 sobre-expresados y 498 infra-expresados).

El análisis de los genes sobre-expresados en los tumores muestra sobre-representación de los siguientes procesos biológicos descritos en la base de datos “Gene Ontology”: regulación del ciclo celular, organización del huso mitótico, reparación del DNA, reparación de roturas de doble cadena por recombinación homóloga, organización de la matriz extracelular y “checkpoint” mitótico (Tabla 25).

Del mismo modo, utilizando la base de datos KEGG, encontramos sobre-representadas vías relacionadas con el ciclo celular y el proceso de patogénesis del carcinoma de pulmón de células pequeñas (microcítico).

**Tabla 24. Genes expresados diferencialmente tras comparar los niveles de expresión en tumores respecto a mucosas normales con una  $p$  ajustada  $<0,05$ .**

### Genes sobre-expresados

<b>DCC1</b>	defective in sister chromatid cohesion homolog 1 (S. cerevisiae)
<b>HOXC6</b>	homeobox C6
<b>MMP1</b>	matrix metalloproteinase 1 (interstitial collagenase)
<b>TCF3</b>	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
<b>PARN</b>	poly(A)-specific ribonuclease (deadenylation nuclease)
<b>LAPTM4B</b>	lysosomal associated protein transmembrane 4 beta
<b>SHCBP1</b>	SHC SH2-domain binding protein 1
<b>INHBA</b>	inhibin, beta A
<b>HEATR6</b>	HEAT repeat containing 6
<b>ISG15</b>	ISG15 ubiquitin-like modifier
<b>TRIP13</b>	thyroid hormone receptor interactor 13
<b>IFI6</b>	interferon, alpha-inducible protein 6
<b>RUNX3</b>	runt-related transcription factor 3
<b>PAIP1</b>	poly(A) binding protein interacting protein 1
<b>HIST1H2BH</b>	histone cluster 1, H2bh
<b>GNB5</b>	guanine nucleotide binding protein (G protein), beta 5
<b>CCDC59</b>	coiled-coil domain containing 59

### Genes infra-expresados

<b>C4orf31</b>	chromosome 4 open reading frame 31
<b>CLEC3B</b>	C-type lectin domain family 3, member B
<b>RBPM5</b>	RNA binding protein with multiple splicing
<b>QARS</b>	glutaminyl-tRNA synthetase
<b>LRP10</b>	low density lipoprotein receptor-related protein 10
<b>CXCL12</b>	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)
<b>DARC</b>	Duffy blood group, chemokine receptor
<b>FOLR2</b>	folate receptor 2 (fetal)
<b>CHRD1</b>	chordin-like 1
<b>TTC9</b>	tetratricopeptide repeat domain 9
<b>CLCA4</b>	chloride channel, calcium activated, family member 4
<b>CYP3A5P2</b>	cytochrome P450, family 3, subfamily A, polypeptide 5 pseudogene 2
<b>SLC25A20</b>	solute carrier family 25 (carnitine/acylcarnitine translocase), member 20
<b>FLJ20273</b>	RNA-binding protein
<b>KIAA1026</b>	kazrin

<b>MAN1A1</b>	mannosidase, alpha, class 1A, member 1
<b>CYP2C9</b>	cytochrome P450, family 2, subfamily C, polypeptide 9
<b>PPAP2B</b>	phosphatidic acid phosphatase type 2B
<b>CYP2C18</b>	cytochrome P450, family 2, subfamily C, polypeptide 18
<b>CEACAM1</b>	carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)
<b>MYOC</b>	myocilin, trabecular meshwork inducible glucocorticoid response
<b>EMP1</b>	epithelial membrane protein 1
<b>TP53I3</b>	tumor protein p53 inducible protein 3
<b>LOC650465</b>	similar to Short transient receptor potential channel 2 (TrpC2) (mTrp2)
<b>TNXB</b>	tenascin XB
<b>PIK3C2B</b>	phosphoinositide-3-kinase, class 2, beta polypeptide
<b>PPP2R2A</b>	protein phosphatase 2 (formerly 2A), regulatory subunit B, alpha isoform
<b>KRT13</b>	keratin 13
<b>FUT6</b>	fucosyltransferase 6 (alpha (1,3) fucosyltransferase)
<b>DUOX1</b>	dual oxidase 1
<b>MAP3K3</b>	mitogen-activated protein kinase kinase kinase 3

Entre los genes infra-expresados en los tumores, respecto a la mucosa normal, encontramos una sobre-representación de los siguientes procesos biológicos descritos en la base de datos GO: La diferenciación de los queratinocitos, la morfogénesis tisular y la queratinización (Tabla 25). Entre los genes infra-expresados no encontramos ningún término de KEGG sobre-representado.

**Tabla 25. Términos biológicos sobre-representados obtenidos por la comparación de los tumores con las mucosas normales.**

<b>Genes sobre-expresados</b>
<b><u>Procesos biológicos de Gene Ontology (<math>p &lt; 0,001</math>)</u></b>
Ciclo celular
Regulación del ciclo celular
Interfase mitótica
Organización del huso mitótico
Reparación del DNA
Reparación RDC por recombinación homóloga
Organización de la matriz extracelular
“Checkpoint” mitótico
“Checkpoint” mitótico del ciclo celular
“Checkpoint” de la replicación del DNA
Respuesta a estímulos endógenos
Condensación mitótica de los cromosomas
Regulación de la progresión a través del ciclo celular
<b><u>Vías KEGG (<math>p &lt; 0,001</math>)</u></b>
Ciclo celular
Carcinoma de pulmón de células pequeñas



## Resultados

---

### Genes infraexpresados

#### Procesos biológicos de Gene Ontology ( $p < 0,001$ )

Diferenciación de queratinocitos

Organización y biogénesis del citoesqueleto de actina

Morfogénesis tisular

Queratinización