

UNIVERSIDAD DE SEVILLA



RADIOBIOLOGÍA MOLECULAR EN CÁNCER DE
PULMÓN DE CÉLULAS NO PEQUEÑAS:
FACTORES PRONÓSTICOS Y PREDICTIVOS DE
RESPUESTA.

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

Javier Peinado Serrano

Director:

Amancio Carnero Moya

Co-directores:

María José Ortiz Gordillo

Sevilla, 2021

RESUMEN

RESUMEN

El cáncer de pulmón de célula no pequeña (CPCNP) es la principal causa de muerte por cáncer a nivel mundial, generando un enorme impacto económico y social que no ha dejado de crecer en los últimos años. El tratamiento oncológico para esta neoplasia engloba, de forma habitual, a la cirugía, quimioterapia, tratamientos sobre dianas moleculares y radiación ionizante. El pronóstico en términos de supervivencia global (SG) y las respuestas terapéuticas dispares entre pacientes, pueden explicarse, en gran medida, por la existencia de perfiles moleculares ampliamente heterogéneos.

El objetivo principal de este estudio fue la identificación de firmas génicas pronósticas y predictivas de respuesta al tratamiento oncológico (radioterápico), que pudieran ayudar en la toma de decisiones terapéuticas en paciente afectos de CPCNP.

Para ello, tomando como referencia los patrones de expresión génicos diferenciales entre líneas celulares comerciales, diferenciadas por su perfil de respuesta a radiación ionizante (líneas radiosensibles vs. radorresistentes) y extrapolando dichos resultados a una cohorte de 107 pacientes afectos de CPCNP que habían recibido tratamiento radioterápico (entre otras terapias), obtuvimos una firma génica con capacidad para pronosticar supervivencia y capaz de predecir la supervivencia libre de progresión (SLP), lo cual podría traducirse en una predicción de la respuesta al tratamiento oncológico recibido.

Los pacientes que se asociaban con la firma pronóstica desfavorable tenían una mediana de SG de 24,13 meses contra 71,47 meses de la firma favorable y una mediana de SLP de 12,65 meses contra 47,11 meses de los pacientes con firma favorable.

ÍNDICE

ÍNDICE

1. INTRODUCCIÓN

- 1.1. La radiación ionizante como tratamiento oncológico
- 1.2. Radiobiología clásica: las 5 “R” que tratan de explicar la respuesta a radiación ionizante.
- 1.3. Epidemiología y clasificación del cáncer de pulmón.
- 1.4. Abordaje terapéutico del CPCNP.
- 1.5. Caracterización molecular del CPCNP y su impacto en el abordaje terapéutico.
 - 1.5.1. Papel del receptor del factor de crecimiento epidérmico (EGFR).
 - 1.5.2. Familia de oncogenes RAS: papel de KRAS en CPCNP y su respuesta a radiación ionizante
 - 1.5.3. EML4-ALK.
- 1.6. Firmas génicas con valor pronóstico y predictivo de respuesta en CPCNP.

2. HIPÓTESIS

3. OBJETIVOS DEL ESTUDIO

4. MATERIALES Y MÉTODOS

- 4.0. Resumen del proceso de elaboración del proyecto de tesis
- 4.1. Líneas celulares.
- 4.2. Extracción de RNA de las líneas celulares.
- 4.3. Ensayo de clonogenicidad.
- 4.4. Procesamiento de las muestras: cuantificación del ARN y evaluación de su calidad.
- 4.5. Microarrays de DNAC.
- 4.6. Análisis transcriptómico.
- 4.7. Cohorte de pacientes del estudio.
- 4.8. Paquetes estadísticos empleados para la identificación de biomarcadores génicos pronósticos/predictivos y obtención de firmas génicas.
- 4.9. Identificación de genes como potenciales elementos pronósticos y/o predictivos de respuesta de forma individual.
- 4.10. Obtención de firmas génicas.
- 4.11. Ensayo de las firmas génicas en cohorte adicional.

5. RESULTADOS

5.1. Valores de supervivencia a 2 Gy y curvas de supervivencia de todas las líneas celulares.

5.2. Evaluación de la síntesis y purificación del ARNc previa a la hibridación en el microarray.

5.3. Evaluación de la calidad de los datos de los microarrays.

5.4. Análisis bioinformático. Identificación de genes con expresión diferencial entre líneas radiosensibles y radiorresistentes.

5.4.1. Análisis bioinformático incluyendo todas las líneas celulares.

5.4.2. Resultados del análisis mediante enriquecimiento en funciones moleculares, componentes celulares y vías de señalización.

5.4.3. Validación individual de genes con expresión diferencial entre líneas radiosensibles y radiorresistentes como biomarcador pronóstico en la cohorte de interés del TCGA.

5.4.4. Validación conjunta de genes como firma pronóstica.

5.4.5. Validación individual de genes con expresión diferencial entre líneas radiosensibles y radiorresistentes como biomarcador predictivo de respuesta en la cohorte de interés del TCGA.

5.4.6. Validación conjunta de genes como firma predictiva de respuesta al tratamiento.

5.5. Generación de firma génica común para predicción de supervivencia global y supervivencia libre de progresión.

5.5.1. Aplicación de la firma génica para predicción de SG y SLP en función del subtipo histológico.

5.6. Ensayo de la firma génica en cohorte adicional

5.7. Cohorte de pacientes del TCGA. Análisis univariante y multivariante de variables clínicas, patológicas y moleculares como factores pronósticos y predictivos de respuesta a tratamiento.

5.7.1. Análisis estadístico de las variables clínicas y patológicas y su impacto en supervivencia global.

5.7.2. Análisis estadístico de las variables clínicas y patológicas y su impacto en supervivencia libre de progresión.

6. DISCUSIÓN.

7. CONCLUSIONES.

8. BIBLIOGRAFIA.

* ANEXO 1. Ensayo de clonogenicidad

* ANEXO 2. Tabla resumen de los valores de SF2 publicados en la literatura

- * ANEXO 3. Listado de genes diferencialmente expresados entre líneas celulares radiosensibles y radiorresistentes.
- * ANEXO 4. Identificadores de pacientes de la cohorte del TCGA.
- * ABREVIATURAS
- * PUBLICACIONES

INTRODUCCIÓN

1. INTRODUCCIÓN

1.1. LA RADIACIÓN IONIZANTE COMO TRATAMIENTO ONCOLÓGICO

Desde el descubrimiento de los Rayos X por Roentgen y los fenómenos de radiactividad por el matrimonio Curie hace ya más de 110 años, la incorporación de la radiación ionizante en la medicina no se hizo esperar. El impacto en términos de beneficios clínicos para el paciente solo tiene su homólogo, hoy en día, en el tratamiento quirúrgico, siendo estas dos armas terapéuticas las responsables de poder hablar de curación en cáncer. La radioterapia se administra en al menos el 60% de los pacientes afectados por alguna neoplasia, con unos perfiles de tolerancia y seguridad en constante mejora y unos resultados clínicos excelentes en el escenario clínico radical, complementario o paliativo.

En la actualidad, hablar de tratamiento antineoplásico es hablar de medicina personalizada. La personalización pretende ofrecer al paciente un tratamiento individualizado adaptado a sus características físicas y a las características de la enfermedad neoplásica que padece. No obstante, los tratamientos oncológicos empleados, tanto en estadio localizado como en el metastásico, distan bastante de lo que podríamos considerar como personalización óptima en términos de individualización. Aunque ya son varias las neoplasias a las que ofrecer alternativas terapéuticas más precisas, seguimos empleando de manera estándar, tratamientos generalistas como la cirugía, radiación ionizante y quimioterápicos. Hablar de tratamiento individualizado en terapias sistémicas es hacerlo centrando el abordaje terapéutico sobre dianas moleculares específicas. Sin embargo, la personalización en el tratamiento radioterápico se centra, principalmente, en la aplicación de elementos de inmovilización adaptados a la anatomía de cada paciente, que permiten una ganancia de precisión, así como de avances tecnológicos en diagnóstico por la imagen con el empleo de la Tomografía Computarizada (TC) para la planificación de tratamientos y su fusión con la Tomografía por Emisión de Positrones (PET) y la resonancia magnética

(RM), entre otros. Lo que queda claro es que el tratamiento radioterápico en sus diferentes modalidades era eficaz mucho antes de que se descubrieran las bases moleculares que podían explicar su efecto biológico. Dicho efecto molecular, centrado directa y/o indirectamente en el daño sobre el DNA nuclear y/o mitocondrial, supone solo la punta del iceberg del efecto biológico de esta potente arma terapéutica.

El tratamiento radioterápico ha soportado con éxito la aparición de numerosas oleadas de alternativas terapéuticas que amenazaban con su sustitución como segunda terapia más importante tras cirugía. Lejos de ello, ha sabido evolucionar hacia la personalización terapéutica, ofreciendo ahora gran precisión y perfil de tolerancia sin precedentes. La radioterapia ofrece a algunos pacientes la posibilidad de curación a largo plazo con un perfil de toxicidad excelente.

A pesar de lo anterior, seguimos precisando de una profundización mayor en el conocimiento sobre las bases moleculares que gobiernan sobre los diferentes efectos terapéuticos tras radiación ionizante.

1.2. RADIOBIOLOGIA CLÁSICA: LAS 5 “R” QUE TRATAN DE EXPLICAR LA RESPUESTA A RADIACIÓN IONIZANTE.

Antes de que los avances en biología molecular del cáncer nos permitieran profundizar en las bases moleculares relacionadas con la respuesta a radiación, se habían descrito varios fenómenos que condicionan las respuestas dispares apreciadas al tratamiento con radiación ionizante. Esto sentó las bases para el uso de diferentes esquemas de tratamiento en función del subtipo histológico y el estadio de la enfermedad. Dichos factores, conocidos comúnmente como las 5 “R” de la radiobiología, son aplicables tanto sobre tejidos sanos como tumorales y son:

- **Radiosensibilidad intrínseca**: dos tumores de diferente estirpe histológica muestran una respuesta terapéutica muy diferente. Un buen ejemplo de ello es el caso de un melanoma, considerado como radiorresistente a fraccionamientos estándares y el linfoma folicular de bajo grado, cuyo tratamiento curativo pasa por dosis muy bajas de radiación en monoterapia. Así mismo, dos tumores considerados de la misma estirpe histológica también muestran, en ocasiones, respuestas terapéuticas muy dispares. La profundización en el estudio de esta variable abre el campo de las alteraciones moleculares propias de cada célula tumoral. Las alteraciones en las rutas de señalización que median en proliferación, reparación del daño o muerte celular programada, entre otras, condicionan el amplio abanico de respuestas objetivadas en la clínica. El estudio de la radiosensibilidad intrínseca *in-vitro* se lleva a cabo mediante el ensayo de clonogenicidad, obteniéndose el parámetro fracción superviviente a 2 Gy (SF2) que permite comparar resultados de diferentes subtipos histológicos y permite predecir la potencial respuesta clínica de dicho tumor. El estudio de los factores moleculares implicados en esta respuesta es uno de los pilares de la radiobiología molecular moderna, en cuyo marco se encuadra la presente tesis.
- **Reoxigenación**: Los primeros estudios que sugerían un importante papel de la oxigenación tisular en la respuesta al tratamiento radioterápico datan de 1909, cuando Schwarz demostró que el efecto de la radiación ionizante sobre la piel era menor si ésta se comprimía externamente, reduciendo así el flujo sanguíneo en la zona irradiada¹. En 1910, Müller publica su estudio en el que obtiene mayores respuestas al tratamiento mediante el aumento del flujo sanguíneo con diatermia². En 1953, Grey y colaboradores³, como pioneros en radiobiología, publican sus estudios sobre el efecto de la hipoxia en la respuesta a radiación ionizante, haciendo uso de un modelo *in-vitro* sobre fibroblastos y otro *in-vivo* sobre ratón. Postulan que la concentración de oxígeno condiciona la respuesta

química inducida en las células tras la radiación. En ese mismo año, se describe la estructura de doble hélice del ADN por Watson y Crick⁴. En base a esto se postula que el daño directo e indirecto (por radicales libres de oxígeno-nitrógeno) al ADN, es uno de los principales mecanismos de acción de la radiación ionizante. Este daño se ve influido por los fenómenos de hipoxia-reoxigenación, siendo la piedra angular en la radiobiología física y molecular actual. Ya en 1955 se describe la presencia de áreas hipóxicas en algunos carcinomas escamosos de pulmón y su potencial relación con la respuesta a radiación⁵. Desde entonces se han estudiado muchos subtipos histológicos neoplásicos, siendo la hipoxia un factor pronóstico independiente en cáncer de cervix^{6,7}, carcinomas de cabeza y cuello⁸ y sarcomas de partes blandas⁹, entre otros.

En los últimos 20 años se ha profundizado de forma significativa en el conocimiento sobre los mecanismos moleculares que median en respuesta a los niveles de oxígeno y su papel en los fenómenos de iniciación, mantenimiento, proliferación, metástasis y respuesta a tratamientos oncológicos¹⁰.

- **Redistribución**: Se trata de las diferencias de respuesta a la radiación ionizante en función de la fase del ciclo celular en la que se encuentre la célula. Los estudios pioneros, hacia finales de los años 60, sitúan a la fase tardía de síntesis del ciclo celular (fase S) como la de menor sensibilidad a la radiación¹¹. Por el contrario, la fase tardía de G2 y la fase de Mitosis, es la de mayor sensibilidad a la misma. Así mismo, se establece que, mediante el fraccionamiento de la dosis, se consigue una redistribución hacia fases más sensibles del ciclo celular. Esto podría explicar el mayor beneficio de dichos esquemas fraccionados. Los factores responsables de las variaciones de respuesta dependientes del ciclo celular son múltiples y aún no del todo determinados, suponiéndose a la capacidad de reparación del daño mediante

recombinación homóloga (RH), uno de los principales motivos detrás de esas diferencias.

- **Reparación**: dado que el daño, directo y/o indirecto, sobre el ADN, es el responsable último del efecto inducido por la radiación ionizante, la reparación de dicho daño inducido es un elemento clave para entender la respuesta terapéutica. Después de la exposición a una fracción de tratamiento, existen 3 subtipos de daños potencialmente inducibles: 1) daño letal, que no es reparable y que supone generalmente la muerte celular (tanto tumoral como de células normales); 2) daño subletal, que supone la generación de alteraciones moleculares que sí son reconocidas de forma adecuada por los sensores de daño y pueden ser reparados y 3) daño potencialmente letal, que de no repararse de forma adecuada llevará a un evento letal. Sabemos que las células neoplásicas poseen en ocasiones mutaciones, como la inhibición de TP53 o de otros sensores, como *ATM* y *BRCA*, que supone un reconocimiento anómalo o fallido de los daños inducidos sobre el ADN. Esto, sumado a una maquinaria imperfecta de reparación, lleva finalmente a la aparición de clones sucesivos con nula o escasa respuesta al tratamiento.

- **Repoblación**: cuando sobre una lesión tumoral dada se avanza en el tratamiento radioterápico fraccionado administrado, existe una subpoblación celular con propiedades pluripotenciales, capaces de regenerar al propio tumor dañado¹². Dichos fenómenos celulares son descritos en la mayoría de las entidades neoplásicas y justificarían en parte, las recaídas locales más o menos tempranas que se objetivan en la clínica. Dicho fenómeno es consecuencia de una mayor resistencia al tratamiento de esos clones pluripotenciales (célula madre cancerígena (CSC)). Se supera de forma habitual alcanzando una dosis

total de radiación absorbida suficiente y no deteniendo el tratamiento fraccionado antes de su teórica finalización. La radiación es altamente efectiva en neoplasias del área de cabeza y cuello o ano, entre otras^{13,14}. Por el contrario, y a pesar de alcanzar altas dosis de radiación, encontramos neoplasias como los gliomas de alto grado o la mayoría de CPCNP localmente avanzados, donde se obtienen pobres resultados clínicos. Esta mala respuesta terapéutica encuentra, en parte, justificación por el perfil extremadamente radio-resistente de dichas CSC, pudiendo aunar dicha explicación al hecho de que, como en el caso del cáncer de pulmón, a mayor tamaño de la lesión, mayores tasas de fracaso terapéutico¹⁵⁻²⁰. Se ha sugerido que a mayor tamaño del tumor habría, potencialmente, mas CSC, posibilitando la aparición de mayor numero de clones resistentes y por tanto mayor posibilidad de fracaso terapéutico^{21,22}.

En un intento por demostrar que, a mayor dosis total, mejor control de la enfermedad, el ensayo clínico fase III RTOG 0617, planteó dos brazos de tratamiento sobre pacientes afectos por CPCNP localmente avanzado, en el que el brazo de 74 Gy (versus 60 Gy) obtuvo peores resultados clínicos, hecho que sorprendió a la comunidad oncológica y fue justificado por un efecto perjudicial sobre los tejidos sanos circundantes, lo que supuso un efecto deletéreo en supervivencia de los pacientes tratados con dosis más altas²³.

Los mencionados factores que pueden determinar las diferencias en respuesta a la radiación son fenómenos generales que durante mucho tiempo han sido considerados como elementos predictivos de respuesta al tratamiento. Hoy en día, aún distan bastante de ser elementos personalizables y difieren sustancialmente entre diferentes subtipos histológicos, localizaciones anatómicas e incluso entre dos tumores de similares características. Si bien dichos factores son tenidos en cuenta a la hora de plantear el abordaje terapéutico con radiación ionizante, no son suficientes, tomados de

forma individual o conjunta, para predecir respuestas terapéuticas ni explican por completo la falta de eficacia del tratamiento en algunos pacientes.

1.3. EPIDEMIOLOGÍA Y CLASIFICACIÓN DEL CÁNCER DE PULMÓN.

El cáncer de pulmón, debido a su frecuencia, la repercusión en la vida de los pacientes y en los sistemas sanitarios, constituye uno de los problemas sanitarios más importantes en nuestra sociedad. Supone la mayor causa de mortalidad por cáncer a nivel mundial²⁴. La incidencia en la Unión Europea es de 288,100 nuevos casos al año con una mortalidad de 252,500 (181,900 hombres y 70,600 mujeres). En España, la incidencia es de 26,715 nuevos casos al año (21,780 hombres y 4,935 mujeres) con 21,180 fallecimientos (17,430 hombres y 3,688 mujeres). El principal factor de riesgo para su desarrollo es el consumo de tabaco²⁵⁻²⁸. Atendiendo a sus características histológicas, se divide en 2 grandes grupos: cáncer de pulmón de células pequeñas (microcítico) y cáncer de pulmón de células no pequeñas (CPCNP). Este último supone aproximadamente el 85-90% del total. A su vez, este grupo se subdivide, atendiendo a características histológicas y moleculares en: adenocarcinomas (la mayoría), escamoso, células grandes, neuroendocrino y NOS (acrónimo del inglés non otherwise specified) (figura 1).

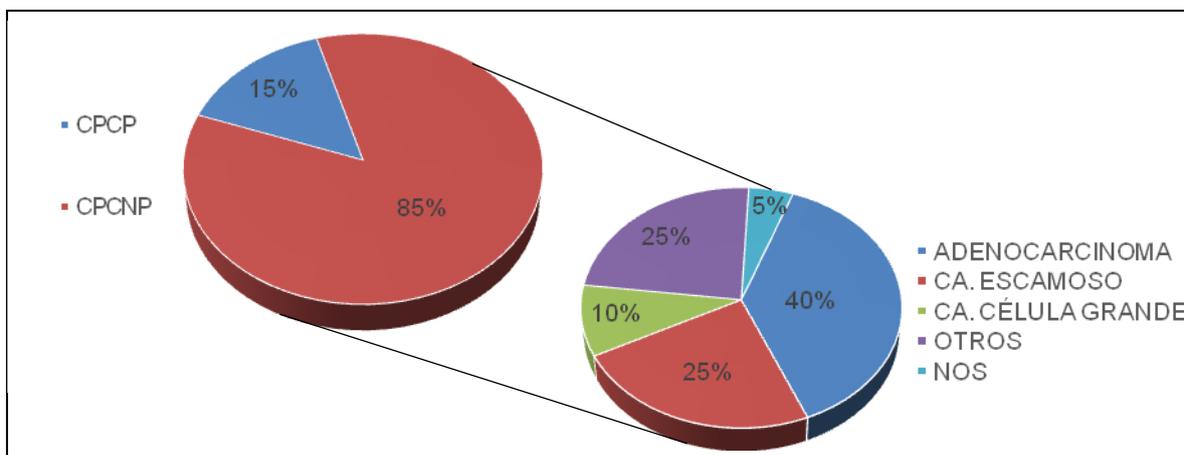


Figura 1. Clasificación general del cáncer de pulmón y clasificación del subgrupo de células no pequeñas

1.4. ABORDAJE TERAPÉUTICO DEL CPCNP

En la actualidad se consideran 3 escenarios clínicos a la hora de plantear el tratamiento oncológico en pacientes recién diagnosticados:

- 1- Resecable
- 2- Localmente avanzado (no resecable)
- 3- Metastásico

Tal y como se muestra en las principales guías internacionales terapéuticas²⁹, el tratamiento radioterápico juega un papel importante en todos los escenarios.

- Es una alternativa al tratamiento quirúrgico en estadios tempranos (Estadio I y II sin carga nodal) mediante radioterapia estereotáxica corporal (SBRT, de sus siglas en inglés) y complementario tras resecciones quirúrgicas con bordes afectados y/o carga nodal positiva (N2).
- En estadios localmente avanzados (Estadio III), donde no es posible una resección quirúrgica de entrada, el tratamiento radioterápico normofraccionado junto al tratamiento quimioterápico (concomitante o secuencial) son el estándar terapéutico.

- En los pacientes con 1 a 5 lesiones torácicas y/o extratorácicas (enfermedad oligometastásica), la opción del tratamiento local con intención radical es considerada una alternativa eficaz.
- Aquellos pacientes con importante carga de enfermedad sistémica se benefician de tratamiento radioterápico paliativo con intención antiálgica, hemostática o descompresiva.
- Por último, la mayoría de los pacientes con diseminación metastásica a nivel del sistema nervioso central, reciben, si su estado de salud lo permite, de forma protocolizada tratamiento radioterápico.

Centrados sobre el escenario localmente avanzado irresecable (Estadios IIIA, IIIB y IIIC), en el momento actual no existe un régimen de radio-quimioterapia estándar, aunque la combinación de un esquema basado en platinos y radioterapia torácica ha mejorado significativamente la supervivencia de estos pacientes. Éstos son tratados con dosis estándar de radiación de 60-66 Gy en concomitancia, o de manera secuencial, con tratamiento quimioterápico combinado³⁰⁻³². Los fraccionamientos pueden variar, pero generalmente se hace uso de 1,8 a 2 Gy/fracción/día (normofraccionados). A pesar de la aplicación de un tratamiento combinado citotóxico, seguimos observando tasas de recaídas locales del 30-50% en este grupo de pacientes^{33,34}. Este hecho justifica, por sí solo, la necesidad de seguir profundizando en las claves biológicas que gobiernan los pobres resultados clínicos obtenidos hasta la fecha.

1.5. CARACTERIZACIÓN MOLECULAR DEL CPCNP Y SU IMPACTO EN EL ABORDAJE TERAPÉUTICO.

La investigación centrada sobre la mejora de las terapias anti-neoplásicas en cáncer de pulmón se ha basado en el estudio genómico y proteómico de tumores con una base genética específica conocida, como la mutación de EGFR y KRAS. Esta clasificación molecular ha tenido una influencia en la respuesta a terapias biológicas basadas en anticuerpos monoclonales e inhibidores de tirosina quinasas en pacientes con cáncer de pulmón³⁵⁻³⁷. Por otro lado, se ha objetivado una respuesta no uniforme de los pacientes a estas terapias, proponiéndose un modelo de resistencia que podría estar mediado por otras mutaciones en algunos genes relevantes (inserciones en EGFR³⁸, KRAS³⁹, amplificación de MET^{40,41} o mutaciones en el dominio quinasa de HER-2⁴²).

Conocemos, además, que el perfil genético tumoral tiene un impacto relevante en la respuesta a quimioterapia⁴³ y radioterapia^{44,45}. Existen en la actualidad múltiples estudios que relacionan algunas de las mutaciones mencionadas con mecanismos de radioresistencia o radiosensibilidad en CPCNP^{45,46}. En la tabla 1 se exponen algunos biomarcadores relacionados con los fenómenos radiobiológicos que pudieran estar implicados en la respuesta dispar a radiación ionizante. No obstante, se conoce poco sobre el papel que pueden jugar los distintos esquemas de fraccionamiento radioterápico y su combinación con otros tratamientos sistémicos o locales, en la respuesta tumoral dependiente de los perfiles moleculares descritos^{47,48}.

Tabla 1. Algunos biomarcadores relacionados con los fenómenos radiobiológicos.

FENÓMENOS RADIOBIOLÓGICOS	BIOMARCADORES
Radiosensibilidad Intrínseca	EGFR, HER2, ALK
Reparación del ADN	ATM, XRCC1, RAD21, BRCA
Repoblación	CD44, Ki67
Redistribución del ciclo celular	CDKN1A
Reoxigenación	HIF1 α , HIF2 α , VEGF

Hoy en día seguimos sin considerar ningún biomarcador como condicionante del tratamiento radioterápico a administrar, a pesar de lo cual se sigue ampliando el conocimiento sobre el efecto que algunas mutaciones podrían tener sobre la respuesta a la radiación ionizante. Así pues, la clasificación del CPCNP atendiendo exclusivamente a las características clínico-patológicas era el único determinante de la terapia administrada. En los últimos años ha habido un cambio radical en la forma de clasificar estos tumores. La disponibilidad de material biológico y los avances en técnicas de análisis transcriptómicos, han permitido mejorar la subclasificación del grupo de neoplasias englobadas dentro de los CPCNP^{49,50}. Así mismo, los avances en biología molecular y genética han mostrado que algunas moléculas específicas contribuyen a la aparición esporádica de cáncer de pulmón y son útiles como dianas terapéuticas y/o como biomarcadores predictivos de respuesta⁵¹. En 2011 se publicaron los resultados del estudio mutacional de genes responsables de cáncer de pulmón (*EGFR*, *KRAS*, *BRAF*, *HER2*, *AKT1*, *PIK3CA*, *MEK1*, *EML4-ALK*, *MET*), identificando en más del 60% de las muestras estudiadas, la existencia de al menos una mutación en los genes mencionados y en más del 90% de los casos, dichas mutaciones fueron consideradas excluyentes, entendiendo esto como la existencia de una única mutación en una muestra tumoral⁵².

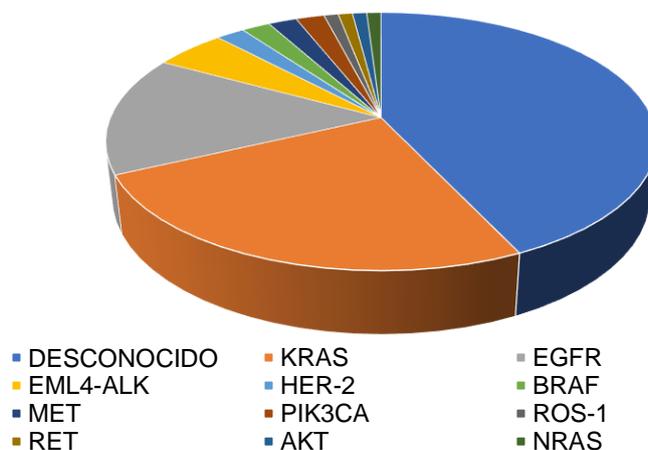


Figura 2. Mutaciones relevantes en Adenocarcinoma y su incidencia

La identificación de dichas mutaciones llevó a crear terapias moleculares dirigidas a mejorar la supervivencia en subgrupos de pacientes con enfermedad metastásica⁵³⁻⁶⁰. Así mismo, se ha profundizado en el papel que juegan dichas mutaciones en la respuesta a radiación ionizante, con el objetivo de poder ofrecer tratamientos combinados que potencialmente puedan mejorar los resultados terapéuticos en aquellos pacientes con CPCNP en estadios tempranos y localmente avanzados. No es objetivo de la presente tesis profundizar en las alteraciones moleculares abordadas mediante inmunoterapia ni las potenciales sinergias derivadas de su combinación con el tratamiento radioterápico.

1.5.1: PAPEL DEL RECEPTOR DEL FACTOR DE CRECIMIENTO EPIDÉRMICO (EGFR)

EGFR es un importante regulador del proceso tumorigénico, mediando en los procesos de proliferación, apoptosis, angiogénesis e invasión tumoral. Se encuentra amplificado y/o sobre-expresado hasta en el 6% y mutado entre el 10-15% de los CPCNP^{52,61}. Así mismo, junto con sus ligandos, se activa constitutivamente de forma habitual durante la iniciación y la progresión de la neoplasia.

Las líneas celulares de CPCNP que albergan mutaciones en el dominio tirosina quinasa de EGFR muestran mayor radiosensibilidad en comparación con las líneas celulares con EGFR nativo. La radiosensibilidad, tanto de las líneas celulares de CPCNP con EGFR mutante como de las células epiteliales bronquiales humanas que de forma estable expresan formas mutantes de EGFR, se ha atribuido a diversos aspectos: 1) al retraso en la cinética de reparación del ADN, 2) a defectos en los STOP's inducidos por la radiación durante la síntesis de ADN o en la mitosis y 3) al aumento tanto del fenómeno apoptótico como de la aparición de micronúcleos. Aparentemente, EGFR mutado es incapaz de trasladarse al núcleo, lo que dificulta su interacción con la proteína

quinasa dependiente de ADN (DNA-PK), que es una enzima fundamental en los procesos de reparación de roturas de cadena doble inducidas por la radiación⁶².

La inhibición de EGFR mediante inhibidores de tirosina quinasa (TKI, de sus siglas en inglés) o por anticuerpos monoclonales (mAb, de sus siglas en inglés), ha demostrado limitaciones a la hora de radiosensibilizar líneas celulares de CPCNP *in vitro* e *in vivo*^{62,64,65}.

1.5.2: FAMILIA DE ONCOGENES RAS: PAPEL DE KRAS EN CPCNP Y SU RESPUESTA A RADIACIÓN IONIZANTE

La familia de oncogenes RAS (*HRAS*, *KRAS* y *NRAS*) codifican para proteínas de transducción de señal que se relacionan con la transmisión de señales procedentes de receptores extracelulares de crecimiento, como EGFR. Se trata de una proteína pequeña que une GTP, localizada en la cara interna de la membrana plasmática, que tiene actividad GTPasa. Tras la activación de RAS mediante el cambio de GDP por GTP, se activan múltiples efectores aguas abajo en la cascada de señalización, tales como MAPK, STAT y PI3K, que regulan, en última instancia, los fenómenos de proliferación, motilidad y apoptosis⁶⁶. Las mutaciones activadoras de RAS evitan la hidrólisis de GTP a GDP, llevando a un estado de activación constitutiva activando las cascadas de señalización citadas.

Aproximadamente el 30% de los adenocarcinomas y un 5% de los carcinomas escamosos, presentan mutaciones en KRAS y aunque en la actualidad sigue siendo una cuestión algo controvertida, la activación de mutaciones en KRAS constituye un marcador de mal pronóstico en CPCNP^{67,68}. En el 2011, Sun y colaboradores publican los resultados de su estudio en el que evalúan el papel de dichas mutaciones en la radorresistencia tumoral⁶⁹. Los resultados mostraron que las líneas celulares de cáncer de pulmón denominadas HCC2429, a las que se había transfectado KRAS mutado en posición 12V, presentaron una disminución en la respuesta apoptótica tras radiación. El

mismo equipo ha mostrado, así mismo, que la inhibición específica de JAK2 mediante la molécula TG101209 induce un efecto radio-sensibilizante mediante la inhibición de la fosforilación de STAT3 y la consiguiente reducción en la expresión de Survivina en la línea HCC2429 y H460. Además, en experimentos *in vivo*, la inhibición de Survivina se relacionó con un aumento en la apoptosis, una reducción de la proliferación tumoral y de la densidad vascular asociada. Una vez superado el efecto protector de la Survivina, las diferencias observadas en la inducción de apoptosis entre las dos líneas celulares usadas en ese estudio (H460 y HCC2429) parecen explicarse por la situación mutacional de KRAS.

Wang y colaboradores han propuesto a la osteopontina (SPP1) y su relación con la condensación cromatínica de tipo mitótico (MLCC) dependiente de EGFR, con un perfil de radioresistencia mayor en algunas subpoblaciones celulares de CPCNP con KRAS mutado⁷⁰. La condensación cromatínica se ha relacionado con una mayor protección contra roturas de doble hélice de DNA y, de acuerdo con el trabajo citado, potencialmente con una regulación negativa de los inhibidores de las propiedades de tipo stem de algunas células, como la invasión y la metástasis. De esta forma, se propone un modelo en el que se conectan el fenotipo stem con la vía de EGFR y SPP1, cooperando para modular la condensación cromatínica y la inducción de roturas de doble hélice. Las células de CPCNP KRAS mutadas que presentan activación de la vía de MLCC y niveles bajos de BIM, son más proclives a alteraciones genómicas en los supresores tumorales TP53 y CDKN2A. Así mismo, parece existir una relación positiva entre niveles de SPP1 y mutación de TP53. La identificación de co-mutaciones en KRAS y TP53 se ha relacionado con niveles más altos de radio-resistencia en modelos in-vitro e in-vivo basados en líneas celulares y xénografts de CPCNP, proponiendo a la escalada de dosis radioterápica y/o a la radiosensibilización mediante inhibidores de tirosina quinasas, como rutas para vencer dicha resistencia⁷¹.

1.5.3: EML4-ALK

La fusión del gen de la proteína 4 asociada a microtúbulo de equinodermo (EML4) con el gen de la quinasa del linfoma anaplásico (ALK) fue identificada por primera vez en células neoplásicas de cáncer de pulmón no microcítico en 2007⁷². La fusión lleva asociada la generación de un transcrito traducido a una proteína (EML4-ALK) con dominio tirosina-quinasa, que promueve y mantiene las características de malignidad de la célula neoplásica. Esta fusión, resultado de una inversión en el brazo corto del cromosoma 2, sucede en aproximadamente el 5% de los CPCNP, generalmente pacientes jóvenes, no fumadores y subtipo adenocarcinoma. Su existencia es mutuamente excluyente con mutaciones en RAS y EGFR. Su función como oncogén es dependiente de su actividad tirosina-quinasa, lo que ha permitido emplear tratamientos dirigidos contra esa diana (por ejemplo, crizotinib y ceritinib, entre otros). Los resultados clínicos derivados de ensayos en fase 3, han demostrado la superioridad de crizotinib en comparación con quimioterápicos, en pacientes metastásicos que poseen el reordenamiento de ALK, tanto en primera como en segunda línea terapéutica^{73,74}. En la actualidad se sabe que la mayoría de los pacientes tratados con crizotinib adquieren resistencia al fármaco en los 12 primeros meses de tratamiento⁷⁵. En este sentido, ceritinib, inhibidor tirosina-quinasa de ALK de segunda generación, ha mostrado una actividad hasta 20 veces mayor que crizotinib, demostrando respuestas clínicas en pacientes previamente tratados con crizotinib⁷⁴, así como mejores resultados que quimioterapia estándar en primera y segunda línea de tratamiento^{77,78}. Por otro lado, estos tratamientos dirigidos han demostrado actividad frente a la función tirosina-quinasa de MET y ROS1, cuya amplificación y reordenamiento-fusión, respectivamente, han sido descritas en CPCNP. MET y ROS1 se han descrito como responsables de resistencia en el caso de tumores con mutaciones en EGFR que han adquirido resistencia a los inhibidores tirosina-quinasa (erlotinib, gefitinib)⁷⁹.

Con respecto al papel del reordenamiento de ALK y la inhibición de su dominio funcional, en la modulación de la respuesta a radiación ionizante, en 2013, Sun y colaboradores demuestran las propiedades radiosensibilizadoras de crizotinib en un modelo *in vivo* e *in vitro* en líneas ALK⁺⁸⁰. Analizando la vía de señalización de ALK aguas abajo, demuestran una activación de AKT, ERK y STAT3 tras irradiación y cómo al combinar crizotinib y radioterapia se inhibe a ALK por completo y se disminuye la activación de los efectores. Similares resultados se han obtenido con posterioridad tanto con la combinación de crizotinib como con las moléculas de segunda generación de la misma familia, confirmándose su efecto radiosensibilizador tanto a fotones como a iones de carbono⁸¹. Hasta la fecha no existen ensayos clínicos en fase de aplicación humana que testen la combinación de ambas terapias en cáncer de pulmón. Algunos grupos de trabajo han publicado su experiencia en pacientes tratados con crizotinib y técnicas de radioterapia ablativa en pacientes con enfermedad oligometastásica que progresaban al tratamiento dirigido. Estos estudios proponen la irradiación como arma para superar los fenómenos de resistencia que aparecen tarde o temprano, sugiriendo un beneficio en términos de supervivencia para los pacientes que continuaban con el tratamiento dirigido⁸².

A pesar de la evidencia científica acumulada sobre el impacto de mutaciones relevantes en CPCNP, hoy en día ninguna de ellas condiciona el tratamiento radioterápico, obteniéndose respuestas clínicas en tumores donde la terapia sistémica no ha resultado efectiva, lo que confirma el efecto multi-diana de la radiación y su condición como arma terapéutica fundamental en todos los estadios de la enfermedad.

1.6. FIRMAS GÉNICAS CON VALOR PRONÓSTICO Y PREDICTIVO DE RESPUESTA EN CPCNP.

Una firma de expresión génica puede definirse como el patrón específico de expresión de uno o varios genes con especificidad validada en términos de diagnóstico, pronóstico o predicción de respuesta a algún tratamiento. Desde hace más de 20 años,

con el desarrollo y mejora de la tecnología de microarrays, se han generado numerosos estudios, basados principalmente en el análisis transcriptómico, que proponen diferentes firmas génicas como biomarcadores pronósticos tanto en adenocarcinoma⁸³⁻⁹¹, carcinoma epidermoide⁹²⁻⁹⁴ como en CPCNP en general⁹⁵⁻¹⁰⁶. Algunos de esos estudios tratan de identificar biomarcadores pronósticos y predictivos de respuesta a tratamientos sistémicos. La mayoría se centra en la identificación de marcadores que ayuden en la toma de decisiones clínicas sobre la idoneidad de administrar tratamiento sistémico adyuvante en estadios tempranos de CPCNP tras cirugía^{107,108}. Esto tienen como referencia el manejo de las pacientes afectas por cáncer de mama, en los que ya existen varias plataformas génicas que ayudan a predecir supervivencia y el potencial beneficio de administrar quimioterapia adyuvante¹⁰⁹.

Por el contrario, son pocos los estudios publicados que plantean firmas predictivas de respuesta a radiación ionizante en CPCNP. Así, en base al trabajo de investigación realizado por el Instituto Nacional del Cáncer Americano (NCI, de sus siglas en inglés) sobre análisis *in-vitro* de los efectos biológicos de diferentes fármacos antineoplásicos en líneas celulares¹¹⁰, Torres-Roca y colaboradores, publican en el año 2005 los resultados de su trabajo, centrado en la identificación de genes correlacionados con la respuesta a radiación ionizante en 35 de las líneas celulares del panel NCI-60, representativas de 9 tipos de cáncer. En él evaluaron las contribuciones genéticas a la radiosensibilidad mediante la cuantificación de la fracción sobreviviente después de la exposición de las líneas celulares a una dosis de radiación estándar de 2 Gy (el ensayo de clonogenicidad). Concluyen que su modelo predictor de respuesta en base a la expresión génica podría ser de gran utilidad para mejorar el abordaje terapéutico de los pacientes, asumiendo que dicho modelo necesitaría de una correcta validación *in-vivo*¹¹¹.

Con posterioridad, Eschrich y colaboradores ampliaron el modelo a 48 líneas celulares del panel del NCI60 e incluyeron otras variables biológicas como el estado

mutacional de KRAS, TP53, así como el tejido de origen¹¹². Combinando estos datos, crearon un algoritmo lineal basado en rangos para calcular un índice de radiosensibilidad (RSI, de sus siglas en inglés). El RSI ha sido validado desde entonces en múltiples cohortes de pacientes con diferentes entidades neoplásicas (cáncer de cabeza y cuello, cáncer de mama, gliomas de alto grado, páncreas, metástasis cerebrales y hepáticas)¹¹³⁻¹¹⁵.

En 2017, Scott y colaboradores, publican un trabajo en el que se propone un modelo para adaptar la prescripción radioterápica a la sensibilidad individual del tumor de cada paciente. El modelo, denominado GARD (de sus siglas en inglés, *genome-based model for adjusting radiotherapy dose*), combina la información derivada del índice de radiosensibilidad (RSI) y el modelo lineal cuadrático (modelo LQ). Este extenso trabajo emplea múltiples cohortes de diferentes entidades neoplásicas (cáncer de mama, esófago, cabeza y cuello, estómago, cervix, gliomas, páncreas, pulmón, cáncer de piel no melanoma y melanoma) y establece un valor numérico a GARD (normalmente en rangos de 1 a 200), relacionándose un mayor nivel de GARD con un mayor efecto terapéutico del tratamiento radioterápico y viceversa¹¹⁶.

Casi de manera sincrónica a estos trabajos, en 2008, Amundson y colaboradores publican un artículo en el que analizan la respuesta génica ante el estrés producido por la radiación ionizante en un panel de 60 líneas celulares del NCI y en 3 líneas celulares propias. Dicho trabajo, en contra de los realizados con anterioridad, analiza los cambios en expresión génica inducidos por radiación y no en función al perfil génico basal. A diferencia de lo objetivado al analizar los niveles de expresión basales, en su aproximación no se observan diferencias significativas en los niveles de expresión génica con dependencia del tejido de origen (ovario, pulmón, mama, etc), excepto en el caso de las líneas celulares derivadas de tejido linfoide/mieloide. Esto podría apoyar la hipótesis de la existencia de unos determinantes moleculares de sensibilidad/resistencia a radiación que serían comunes a todos los subtipos tumorales (sin incluir estirpes hematológicas). Concluyen que, a pesar de que existen cambios en la expresión génica

inducidos por radiación, sobre todo en la vía de p53, los niveles de expresión génica basales pueden ser mejores predictores de respuesta¹¹⁷.

Por otro lado, aunque no es el objetivo principal de nuestro estudio, es preciso comentar alguno de los trabajos publicados en relación con el potencial papel predictor y pronóstico de determinados microRNA's (tumORAles y circulantes). Existe un amplio campo de investigación centrado sobre la función de este RNA no codificante que potencialmente actúan como reguladores de expresión génica. Algunos de los más de 2000 miRNA identificados hasta la fecha, pueden ser considerados como biomarcadores diagnósticos y pronósticos en numerosas entidades neoplásicas, incluido el CPCNP¹¹⁸⁻¹¹⁹. En 2010, Hu y colaboradores, publican un artículo en el que identifican una firma pronóstica en CPCNP compuesta por 4 miRNA. Incluyeron para su estudio muestras de pacientes en estadios tempranos y localmente avanzados (estadio I a III) que habían sido tratados con cirugía y quimioterapia. Aparentemente ningún paciente recibió tratamiento radioterápico¹²⁰.

Uno de los trabajos más interesantes es el publicado por Sun y colaboradores en 2018, en el que plantean el papel de determinados miRNA circulantes (c-miRNA) en asociación con otros factores clínicos como determinantes de respuesta a dosis radicales de radiación ionizante en CPCNP. Para ello, proponen la generación de una escala o puntuación en función de la dosis de radiación y la respuesta objetivada, lo que denominan "DRS" (Dose Response Score, de sus siglas en inglés). La cohorte empleada fueron pacientes incluidos en varios ensayos clínicos en los que se planteaba la escalada de dosis de radiación en estadios II y III (más del 90% eran estadio III). Proponen un total de 11 c-miRNA, que, junto con variables como el estadio, la edad, la dosis de radiación administrada, el tratamiento sistémico y la escala de estado general de Karnofsky, servían para la determinación del DRS de cada paciente. La fórmula que resume la obtención del DRS es la siguiente:

$$\begin{aligned} \text{Dose Response Score} = & 3.25 * 10^{-4} * [\text{Age}] + 4.72 * 10^{-3} * [\text{Stage III}] - 6.19 * 10^{-4} * \\ & [\text{KPS}] + 36.028 * [\text{miR} - 10b - 5p] + 0.188 * [\text{miR} - 125b - 5p] - 9.89 * 10^{-4} * \\ & [\text{miR} - 126 - 3p] - 1.518 * [\text{miR} - 134] - 4.855 * [\text{miR} - 155 - 5p] + 4.975 * \\ & [\text{miR} - 200b - 3p] - 0.413 * [\text{miR} - 205 - 5p] + 53.658 * [\text{miR} - 34a - 5p] - 4.21 * 10^{-2} * \\ & [\text{miR} - 92a - 3p] - 2.986 * [\text{miR} - 145 - 5p] - 4.59 * 10^{-3} * [\text{miR} - 22 - 3p]. \end{aligned}$$

Los resultados mostraron que aquellos pacientes que tenían un DRS bajo, se beneficiaban de dosis altas de radiación, con impacto en supervivencia, representado mediante curvas Kaplan Meier. Por el contrario, pacientes con un DRS alto no mostraban diferencias en supervivencia en función de la dosis administrada. Así mismo, demostraron que un valor de DRS bajo en pacientes tratados con dosis altas permitía predecir un menor riesgo de metastatización. Por el contrario, no se consiguió demostrar de forma estadísticamente significativa su capacidad predictora sobre el control local de la enfermedad¹²¹.

Por último, existen otras publicaciones donde se plantean la generación de firmas pronósticas mediante la consideración conjunta de múltiples biomarcadores no específicos de cáncer de pulmón y variables clínicas múltiples. En ese contexto, en 2011, Dehing-Oberije y colaboradores publican los resultados de su estudio, en el que desarrollan y validan una firma pronóstica en paciente afectos de CPCNP tratados con radioterapia +/- quimioterapia. En ella se combinan la determinación de biomarcadores en sangre periférica (antígeno carcinoembrionario (CEA) e interleuquina 6 (IL-6)), con otros factores clínicos, como el sexo, el estado general del paciente, el volumen espiratorio forzado (FEV₁), número de ganglios afectos y volumen del tumor primario (GTV, de sus siglas en inglés)¹²².

Dada la escasa evidencia científica publicada, consideramos que la búsqueda de firmas génicas pronósticas y/o predictivas de respuesta al tratamiento radioterápico, puede ayudar de forma importante a la toma de decisiones clínicas. Estas firmas pueden condicionar los fraccionamientos, la dosis total a administrar y/o la combinación de tratamientos sistémicos en concomitancia con radiación. El fin último es alcanzar

mejores resultados clínicos, minimizando los efectos adversos asociados con las terapias oncológicas actuales.

HIPÓTESIS

2. HIPÓTESIS

El CPCNP, que engloba a los subtipos histológicos más frecuentes de cáncer de pulmón, son entidades neoplásicas cuyo comportamiento biológico es altamente heterogéneo. La heterogeneidad molecular podría explicar las diferentes respuestas y comportamientos antes los tratamientos oncológicos, así como al propio impacto de la neoplasia sobre la supervivencia global de los pacientes.

Considerando dicha heterogeneidad molecular, creemos que la identificación de firmas génicas con capacidad pronóstica y de predicción de respuesta al tratamiento oncológico, en este caso, centrado sobre el tratamiento radioterápico, podría ser de gran ayuda a la hora de optimizar el abordaje terapéutico por parte de los profesionales de la salud.

OBJETIVOS DEL ESTUDIO

3. OBJETIVOS DEL ESTUDIO

El objetivo final es la identificación de una firma de expresión génica pronóstica y predictora de resistencia o sensibilidad a tratamiento radioterápico en CPCNP. Para ello podemos proponer los siguientes objetivos individuales:

- 1- Análisis de sensibilidad innata a radiación ionizante en 15 líneas celulares comerciales de CPCNP.
- 2- Realización de análisis transcriptómico completo de las líneas celulares y su exploración bioinformática, llevándose a cabo un subanálisis mediante enriquecimiento en funciones moleculares, componentes celulares y vías de señalización.
- 3- Explotación de cohorte del “TCGA lung cancer”: selección de pacientes tratados con radioterapia. Análisis de supervivencia en función de diferentes parámetros clínicos-patológicos (análisis uni y multivariante).
- 4- Obtención de genes con expresión diferencial significativa entre líneas radiosensibles y radiorresistentes y validación individual como biomarcador pronóstico/predictivo en cohorte de pacientes afectados de CPCNP.
- 5- Obtención de firmas génicas pronósticas y/o predictivas de respuesta en la misma cohorte utilizada en el punto anterior.
- 6- Ensayo de las firmas en otra cohorte pública compuesta por pacientes afectados de CPCNP.

MATERIALES Y MÉTODOS

4. MATERIALES Y MÉTODOS

4.0. RESUMEN DE LOS PASOS FUNDAMENTALES SEGUIDOS EN LA ELABORACIÓN DE LA PRESENTE TESIS DOCTORAL.

En la figura 3 representamos un diagrama resumen de los pasos seguidos para la consecución de nuestros objetivos.

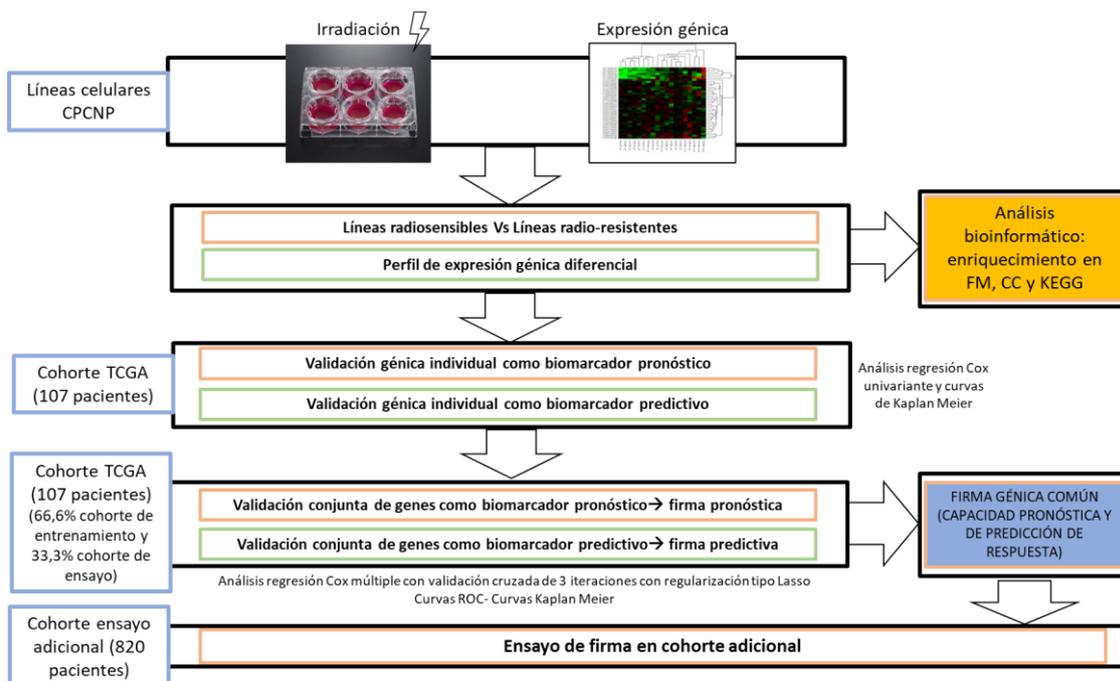


Figura 3. Diagrama en el que se representan los pasos seguidos durante el proyecto de tesis en cuestión.

4.1. LÍNEAS CELULARES

Se emplearon un total de 15 líneas celulares tumorales de CPCNP obtenidas de la American Type Culture Collection (ATCC). De las líneas celulares empleadas, 3 corresponden a carcinoma escamoso (H226, H520 y Calu 1), 2 son carcinomas bronquioloalveolares (H1437 y H358), 1 corresponde a carcinoma de células grandes (H460) y las restantes 9 son catalogadas como adenocarcinomas (A549, Calu-3, H1650, H1781, H1975, H2009, H2228, H3122, HCC827). Todas las líneas celulares crecían en adherencia en placas de cultivo. Se empleó medio de cultivo RPMI con suplemento de suero fetal bovino y antibiótico en todas las líneas celulares, excepto para A549 (DMEM

+ piruvato de sodio, HEPES y aminoácidos no esenciales), Calu 1 (McCay + suplemento de glutamina) y Calu 3 (DMEM), de acuerdo a las recomendaciones de la ATCC.

La tabla 2 resume la información histológica y el perfil molecular de algunas mutaciones relevantes en CPCNP de las líneas celulares empleadas.

Tabla 2. Resumen de las líneas celulares tumorales comerciales empleadas en el estudio

LÍNEA CELULAR	HISTOLGIA	EGFR	KRAS	OTROS
A549	Adenocarcinoma	WT	WT	
H1650	CBA	MUT (exon 19)	WT	
H1975	Adenocarcinoma	MUT (790M)	WT	CDKN2A/Pi3KCA/TP53
H2009	Adenocarcinoma	WT	MT	
H358	CBA	WT	MT	
H460	CCG	WT	MT	
H520	Ca. Escamoso	WT	desconocido	
HCC827	Adenocarcinoma	MUT	WT	
H1437	Adenocarcinoma			CDKN2A/TP53
H1781	Adenocarcinoma			
H2228	Adenocarcinoma			ALK
H226	Ca. Escamoso			
Calu 1	Ca. Escamoso		MT	
Calu 3	Adenocarcinoma	WT	MT	Amplificación ERBB2 /p53/CDKN2A
H3122	Adenocarcinoma			ALK

Se representa información sobre la histología y el perfil molecular de genes relevantes en CPCNP. CBA: carcinoma bronquioloalveolar, CCG: carcinoma de células grandes. WT: wild type (nativo). MUT: mutado.

4.2. EXTRACCIÓN DE RNA DE LAS LÍNEAS CELULARES

Mediante el Kit comercial de extracción de miRNA (miRNeasy® Mini Kit (Qiagen, California, USA)) siguiendo las instrucciones del fabricante.

4.3. ENSAYO DE CLONOGENICIDAD

Empleamos el ensayo de clonogenicidad para la valoración de respuesta a radiación mediante la determinación del parámetro SF2, o fracción superviviente a 2 Gy, contabilizando el número de unidades formadoras de colonias (UFC) existentes a esa dosis. Dicho ensayo se compone de los siguientes pasos:

1- Siembra de la línea celular de interés: en placas de 6 pocillos, con el medio de cultivo recomendado por la ATCC. Dejamos crecer hasta alcanzar un 70-90% de

confluencia (dependiendo de la tasa de proliferación de cada línea celular), manteniendo a las células en condiciones estándares de incubación (37°C de temperatura, 5% de CO₂ y 21% de O₂).

2- Irradiación de la línea celular: tras alcanzar el nivel de confluencia deseado, procedimos a la irradiación en acelerador de electrones modelo Synergy® (Elekta, Estocolmo, Suecia) usando fotones de 6 Mv de energía, con dosis de 2, 4, 6, y 8 Gy, administradas en una sola fracción, con una tasa de dosis de 400 unidades monitor (UM). Dicha unidad se ubicaba en el S^o de Oncología Radioterápica del Virgen del Rocío (Sevilla) (imagen 1). Se empleó una fantoma (imagen 2) diseñado para albergar las placas de cultivo en cuestión, rellenando los pocillos con medio de cultivo, de tal forma que no existiera interfase aire- líquido entre la fantoma y el medio de cultivo, evitando así interferencias electrónicas con el haz de radiación.

3- Cambio del medio en campana de cultivo e incubación durante 36-48 horas.

4- Siembra de células irradiadas: procedimos a la retirada del medio de cultivo, tripsinizando con tripsina al 0,25% durante 3-5 minutos, previo lavado con solución salina (PBS). Centrifugamos las células en 5 ml de medio durante 5 minutos a 10000 r.p.m. Retirada del sobrenadante y resuspensión del pellet celular en 1 ml de medio de cultivo. Con la ayuda de hematocitómetro (cámara de Neubauer) contabilizamos las células, procediendo a la siembra en placas de cultivo de 10 cm de diámetro, de 300-2000 células por triplicado por cada dosis de radiación empleada (control, 2,4,6 y 8 Gy), siendo menor el número de células sembradas en las placas controles y las sometidas a 2 Gy.

5- Incubación: entre 7 y 25 días en atmósfera estándar, para permitir la formación de colonias.

6- Fijación y tinción: Tras el periodo de incubación apropiado para cada línea celular, se llevó a cabo la fijación mediante glutaraldehído al 1% durante 25-30 minutos.

Lavamos de forma adecuada con PBS y teñimos las placas con cristal violeta al 0,05% durante 30-60 minutos. Dejamos secar.

7. Escaneado y contabilización de UFC: escaneamos las placas de cultivo y procedemos a la contabilización empleando el programa informático Image J®.

8. Cálculo de los parámetros radiobiológicos:

a) Eficiencia de crecimiento en placa (plating efficiency o PE)

PE = número de colonias contabilizadas / número de células sembradas

b) Fracción superviviente a 2 Gy (SF2)

SF2 = (número de colonias contabilizadas / número de células sembradas) / PE

9. Elaboración de curvas de supervivencia: en base a los valores de fracción superviviente en las placas control, 2, 4, 6 y 8 Gy.

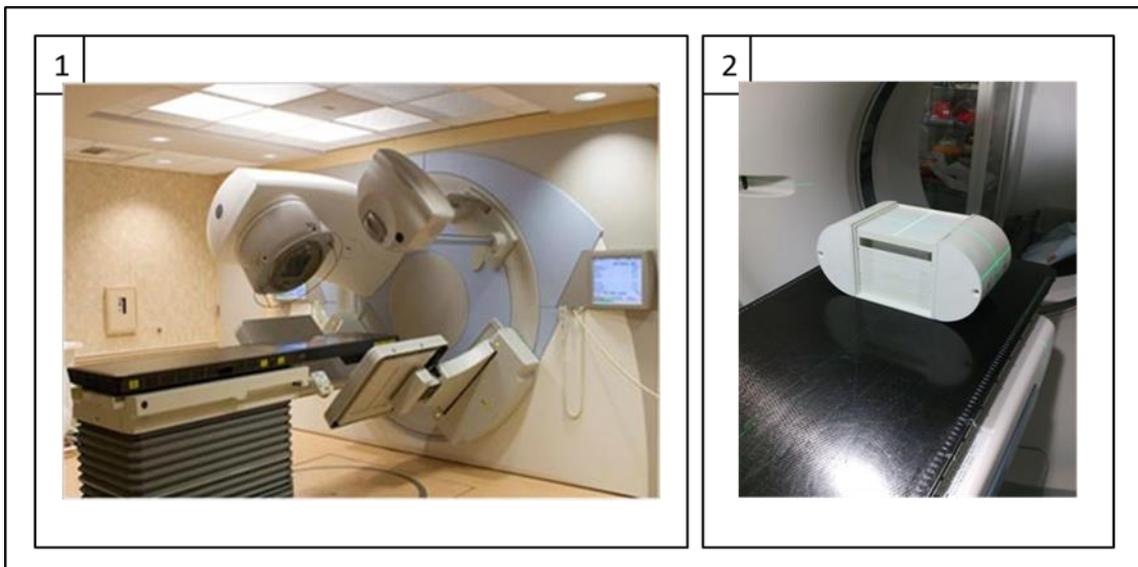


Figura 4. 1) Acelerador lineal de electrones. 2) Fantoma empleado para la irradiación de placas de cultivos.

4.4. PROCESAMIENTO DE LAS MUESTRAS: CUANTIFICACIÓN DE ARN Y EVALUACIÓN DE SU CALIDAD

La cantidad y pureza del ARN obtenido se evaluó mediante la medición de la absorbancia con el espectrofotómetro Nanodrop 2000 (Thermo Scientific; Waltham, MA, USA). La concentración del ARN se midió en ng/ μ L. Para comprobar la calidad del ARN, se empleó el método microfluídico de electroforesis del RNA 6000 Nano kit (Agilent® Technologies; Santa Clara, CA, USA). El análisis se realizó en el Bionalyzer 2100 de Agilent®.

4.5. MICROARRAYS DE ADNc

El ARN obtenido se sometió a una reacción de retrotranscripción en presencia de cebadores aleatorios de 6 nucleótidos (Invitrogen/Life Technologies, Carlsbad, CA), usando dCTPs marcados con Cy3 o Cy5 (Amersham Biosciences) y la retrotranscriptasa MultiScribe™ (Invitrogen/Life Technologies). Como referencia se utilizó el ARN extraído de células normales epiteliales de pulmón. Una vez obtenidas todas las muestras, se utilizaron cantidades equivalentes de ADNc para realizar el microarray. La prehibridación se realizó en una cámara humidificada a 42°C durante 16 a 20 horas, y la hibridación a 65°C en una GeneTac Hybridization Station (Genomics Solutions). Las muestras hibridadas se analizaron con un escáner láser confocal (ArrayExpress, Packard Bioscience) y los datos fueron cuantificados utilizando el software QuantArray (Packard Bioscience). El nivel de significación de la expresión de cada gen se determinó utilizando el test estadístico t-Student (10000 permutaciones), utilizando, como corrección, el test FDR (False Discovery Rate) para eliminar los falsos positivos.

4.6. ANÁLISIS TRANSCRIPTÓMICO

Después de la adquisición de imágenes y la cuantificación, por intensidad de señal, la intensidad media de señales entre las réplicas se determinó para cada secuencia. En el QT -02 (replicados 3 a 6 veces representan el mismo gen) valores con intensidad de la señal baja y baja reproducibilidad entre repeticiones fueron excluidos (media \pm 2 desviaciones estándar de corte). Las señales cuantificadas fueron sometidas

a transformación logarítmica y estandarizada. Utilizamos la prueba de T de student's de permutación (10000 permutaciones) para determinar el nivel de significación de la expresión de cada gen individual, y la tasa de falso positivo es utilizado como corrección para el análisis múltiple. El análisis de conglomerados jerárquico y la fiabilidad de la agrupación se evaluó usando técnicas Bootstrap, empleando el software TMEV. Identificamos los mRNAs alterados en células resistentes vs no resistentes a radioterapia mediante agrupamientos dirigidos, con posterior enriquecimiento de pathways. Este proceso fue llevado a cabo con la ayuda del servicio de bioinformática del Instituto de Biomedicina de Sevilla (IBIS).

4.7. COHORTE DE PACIENTES DEL ESTUDIO

Se empleó la cohorte pública del Atlas del genoma del cáncer (TCGA) centrada sobre pacientes afectos de CPCNP (versión del año 2018). De dicha cohorte se seleccionaron aquellos pacientes en estadio I al III en base a la clasificación TNM (tumor, ganglios, metástasis), cuya edición variaba en función del año del diagnóstico de cada paciente incluido. El requisito principal para la selección de la cohorte era haber recibido tratamiento radioterápico. Para el análisis del impacto de variables clínicas /patológicas sobre la SG y SLP, se llevó a cabo un análisis univariante por regresión de Cox. Posteriormente llevamos a cabo un análisis bivariante con la variable "firma génica" como elemento principal y considerando el resto de las variables como independientes, con el objetivo de identificar alguna relación modificadora o variables de confusión, que condicionaran el impacto de la firma génica en la SG o SLP. Empleamos el paquete estadístico SPSS® versión 20 (IBM, Nueva York, USA). Para la obtención de las firmas génicas se emplearon los datos de la secuenciación de RNA de las muestras tumorales de los pacientes del TCGA, cuyos valores de expresión se obtuvieron como transcritos por millón (TPM) y posteriormente fueron transformados a escala logarítmica.

4.8. PAQUETES ESTADÍSTICOS EMPLEADOS PARA LA IDENTIFICACIÓN DE BIOMARCADORES GÉNICOS PRONÓSTICOS/PREDICTIVOS Y OBTENCIÓN DE FIRMAS GÉNICAS.

El análisis de supervivencia se ha realizado usando el software R versión 3.6.3 (2020-02-29) con los paquetes *Survival* (en versión 3.2.7) para la regresión de Cox, *Survminer* (en versión 0.4.6) para las curvas Kaplan-Meier, *Glmnet* (versión 3.0.2) para la validación cruzada y obtención de las firmas y *SurvivalROC* (versión 1.0.3) para las curvas ROC.

4.9. IDENTIFICACIÓN DE GENES COMO POTENCIALES ELEMENTOS PRONÓSTICOS Y/O PREDICTIVOS DE RESPUESTA DE FORMA INDIVIDUAL

Tras la identificación de los genes diferencialmente expresados entre líneas celulares radiosensibles y radiorresistentes, hicimos uso de los datos del RNAseq de la cohorte del TCGA seleccionada. A continuación, se llevó a cabo un análisis de supervivencia por regresión de Cox univariante. Después de seleccionar aquellos genes cuyo impacto en supervivencia era estadísticamente significativo ($HR > 0 < 1$; $p < 0,05$), se llevó a cabo un análisis de estimación de supervivencia mediante Kaplan Meier, estableciendo como punto de corte el valor de la mediana de expresión de cada gen. Valoramos las diferencias en SG/SLP con la prueba de los rangos logarítmicos. Seleccionamos como elementos potencialmente relevantes para la SG o SLP a aquellos genes que resultaron estadísticamente significativos en ambos análisis ($p < 0,05$)

4.10. OBTENCIÓN DE FIRMAS GÉNICAS

Para la obtención de la firma, a partir de los genes seleccionados como elementos individuales con impacto pronóstico y/o predictivo (HR con $p < 0,05$ y $\log Rank$ Test con $p < 0,05$), se procedió con un análisis de regresión de Cox multivariante con regularización de tipo L1 o Lasso. Para ello, separamos a la cohorte de 107 pacientes del TCGA en 1) cohorte de entrenamiento y 2) cohorte de ensayo. Procedimos a realizar un análisis con validación cruzada de 3 iteraciones, separando aleatoriamente, por 1000 veces como máximo, el 66.6 % de la cohorte para la de entrenamiento y el 33.3%

restante para el set de ensayo. La firma pronóstica y predictiva se construyeron en base a la combinación lineal del coeficiente de regresión obtenido a partir de los coeficientes derivados del análisis de regresión de Cox con regularización de tipo L1 (β) multiplicados por el nivel de expresión de cada gen (en unidades de transcritos por millón). En este caso, tanto la cohorte de entrenamiento como de ensayo procedía de la misma cohorte seleccionada del TCGA. Para el cálculo de la predicción de supervivencia, llevamos a cabo un análisis por curvas ROC (a 12, 24 y 60 meses) y Kaplan Meier. Por último, llevamos a cabo una prueba U de Mann-Whitney para conocer si las diferencias de expresión de los genes de cada firma en cada grupo de riesgo eran estadísticamente significativas. A continuación, separamos a la cohorte en función del tipo histológico (adenocarcinoma o epidermoide) y ensayamos las firmas propuestas.

4.11. ENSAYO DE LAS FIRMAS GÉNICAS EN COHORTE ADICIONAL.

Empleamos la información del TCGA lung cancer para ensayar las firmas propuestas. Se estableció como requisito para la inclusión en la cohorte de ensayo el no ser metastásicos. No se consideraron otros criterios clínicos ni terapéuticos para la selección de esta cohorte adicional. Se excluyeron los pacientes considerados en la cohorte utilizada en el presente trabajo para la generación de las firmas propuestas (que sí habían sido seleccionadas por haber recibido tratamiento radioterápico y no ser metastásicos al diagnóstico). En un primer paso, ensayamos las firmas en la cohorte completa de pacientes (adenocarcinoma + epidermoide) y, seguidamente, probamos cada firma en cada cohorte con dependencia del subtipo histológico. Para el ensayo de la firma, multiplicamos los coeficientes del modelo de regresión multivariante de Cox (para las firmas propuestas) por los respectivos valores de expresión (en la cohorte de ensayo) para obtener las puntuaciones de riesgo. A continuación, valoramos la capacidad de predicción, sobre los mismos datos, de las firmas mediante curvas ROC con el cálculo del área bajo la curva (AUC) a 12, 24 y 60 meses. Por último, considerando como punto de corte al valor de la mediana de las puntuaciones de riesgo,

dividimos a los pacientes en 2 grupos (alto o bajo riesgo) y analizamos la estimación de SG y SLP por Kaplan Meier.

RESULTADOS

5. RESULTADOS

5.1. VALORES DE SUPERVIVENCIA A 2 Gy Y CURVAS DE SUPERVIVENCIA DE TODAS LAS LÍNEAS CELULARES.

Empleamos el ensayo de clonogenicidad para la valoración de la respuesta a radiación ionizante mediante la determinación del parámetro SF2 (fracción superviviente a 2 Gy), contabilizando el número de unidades formadoras de colonias (UFC) existentes a esa dosis. Para ello, sembramos las células en placas de 6 pocillos y dejamos crecer hasta alcanzar un 70-90% de confluencia. Tras ello, irradiamos la línea celular en un acelerador lineal de fotones de 6 Mv de energía con una tasa por minuto de 400 unidades monitor (UM). Tras un cambio de medio e incubación durante 36-48 horas, contabilizamos las células, procediendo a su siembra en placas de cultivo de 10 cm de diámetro, sembrando entre 300 y 2000 células por placa, en triplicado por cada dosis de radiación empleada (0, 2, 4,6 y 8 Gy), siendo menor el número de células sembradas en las placas control y las sometidas a 2 Gy. Dejamos incubar un periodo comprendido entre 7 y 25 días en atmósfera estándar para permitir la formación de colonias. Tras dicho periodo de incubación (variable entre líneas celulares), llevamos a cabo la fijación mediante glutaraldehído al 1% durante 25-30 minutos y teñimos las placas con cristal violeta al 0,05% durante 30-60 minutos. Escaneamos las placas y, por último, procedemos a la contabilización de las UFC mediante el programa informático Image J®. El número de UFC se normalizó al número de células sembradas en cada caso. Se utilizó el valor medio del sumatorio de todos los valores de SF_{2Gy} de cada una de las líneas celulares para obtener un valor medio (0,54), que usamos como punto de corte sobre el cual establecer una clasificación como líneas sensibles (valor de SF2 < 0,54) o resistentes (SF a 2 Gy ≥ 0,54) a radiación. Dicho valor, medido mediante el ensayo de clonogenicidad descrito previamente, así como el resto de los valores de supervivencia a las distintas dosis planteadas (control, 4,6 y 8 Gy), nos permitieron generar las curvas dosis-respuesta mostradas en figura 3. En la tabla 3 se muestra la información sobre los valores de SF a 2 Gy de cada línea celular. El valor de la línea H520 fue extraído de la

literatura. En el anexo 1 se incluye la figura con el ensayo de clonogenicidad de cada línea celular estudiada. En el anexo 2 se incluye la tabla con algunos de los valores de SF2 publicados en la literatura.

Tabla 3. Valores de la fracción superviviente a 2 Gy (SF2) de todas las líneas celulares

ENSAYO DE CLONOGENICIDAD			
	LÍNEA CELULAR	SF2	SD
RADIO-RESISTENTES	H1975	0,891	0,01040341
	A549	0,832	0,16252891
	HC827	0,745	0,2
	H358	0,676	0,157
	H2228	0,646	-
	H3122	0,622	0,05628992
	H460	0,617	0,12920136
	H1650	0,570	0,15699525
RADIO-SENSIBLES	CALU 1	0,454	0,079754773
	H520	0,49	
	H1437	0,448	-
	H226	0,430	0,09
	CALU 3	0,325	0,131673504
	H2009	0,228	0,02413002
	H1781	0,180	0,067

SF2: fracción superviviente a 2 Gy; SD: desviación estándar.

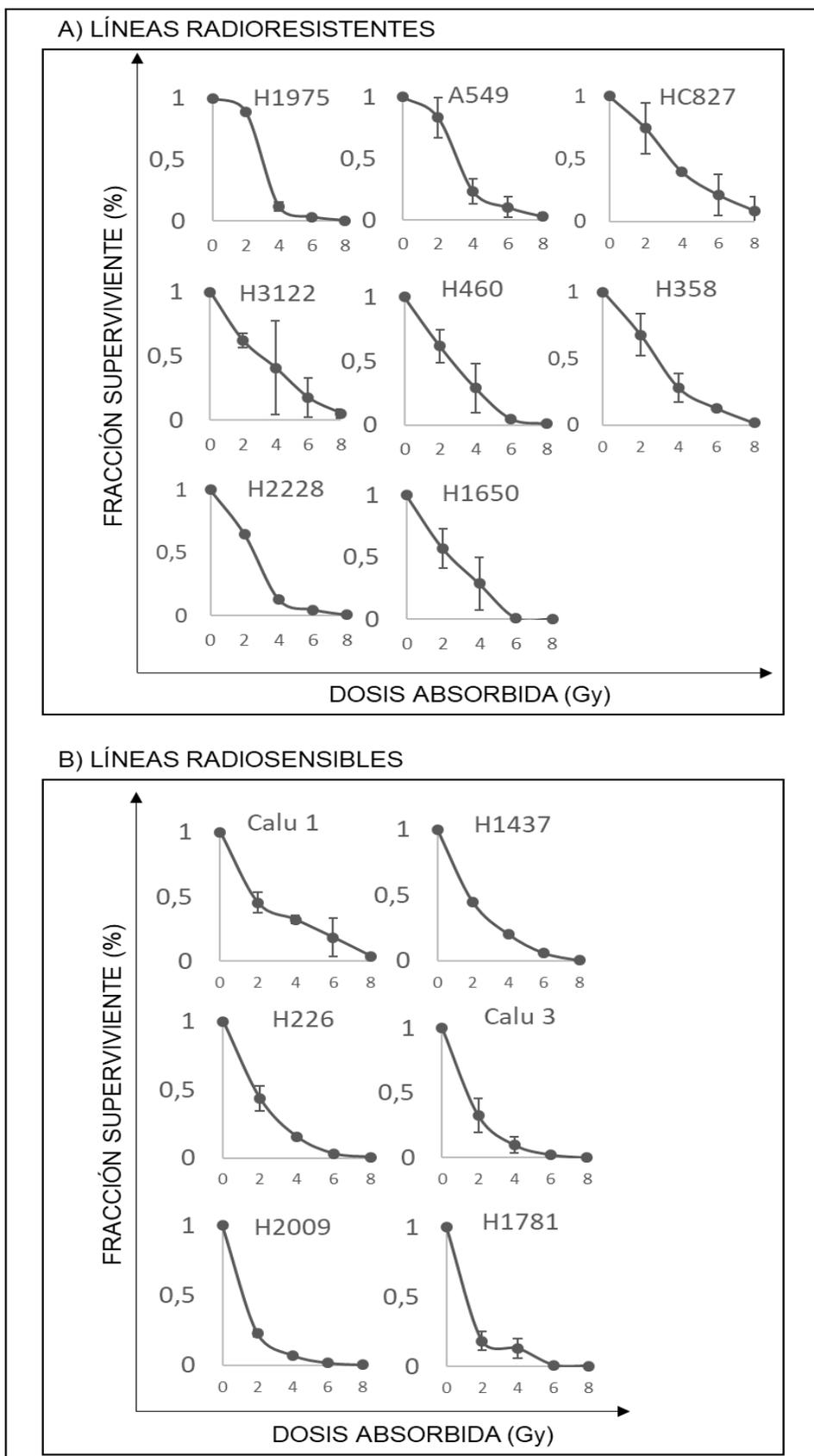


Figura 5. Representación gráfica de las curvas de supervivencia celular tras radiación, medido mediante el ensayo de clonogenicidad. En el eje de ordenadas se representa la fracción superviviente en porcentaje y en el eje de abscisas la dosis de radiación administrada (0- 8 Gy). A) líneas celulares radioresistentes. B) líneas celulares radiosensibles.

5.2. EVALUACIÓN DE LA SÍNTESIS Y PURIFICACIÓN DEL ARNc PREVIA A LA HIBRIDACIÓN EN EL MICROARRAY

En el apartado de materiales y métodos se describen los pasos dados para la obtención del ARNc, y la medición de su concentración y pureza. En la tabla 4 se detallan los valores obtenidos en cada una de las muestras, apreciándose una calidad idónea de las mismas. En la figura 4 se muestra de forma gráfica el análisis practicado mediante el Bioanalyzer 2100® que confirma la calidad del RNA.

Tabla 4. Datos sobre la concentración, pureza y calidad del RNA de todas las muestras empleadas en el estudio.

Nombre Muestra	Concentración (ng/μl)	Abs 260/280	Volumen	RIN
H3122	1948.5	2.1	9 ul	9.6
H460	362.6	2.08	“	9.8
H2228	424.6	2.09	“	9.9
H1975	335.8	2.07	“	10
H1437	403.2	2.1	“	9.8
H1650	339.3	2.06	“	9.9
HC827	269.7	2.09	“	9.9
H1781	1507.2	2.11	“	8.3
Calu 1	111.1	2.05	“	9.9
Calu 3	359.5	2.05	“	9.6
H520	638.4	2.15	“	9.5
Nuli-1	81.5	1.88	“	10
NL 20	785.7	1.84	“	9.9
H226	371	2.09	“	9.8
H2009	135.3	2.03	“	10
H358	531.8	2.12	“	9.8
A549	613.6	2.12	“	9.3
Mix hibridación tejido sano(13 muestras)	50		26	8.3

RIN: número de integridad de RNA. Abs: absorbancia.

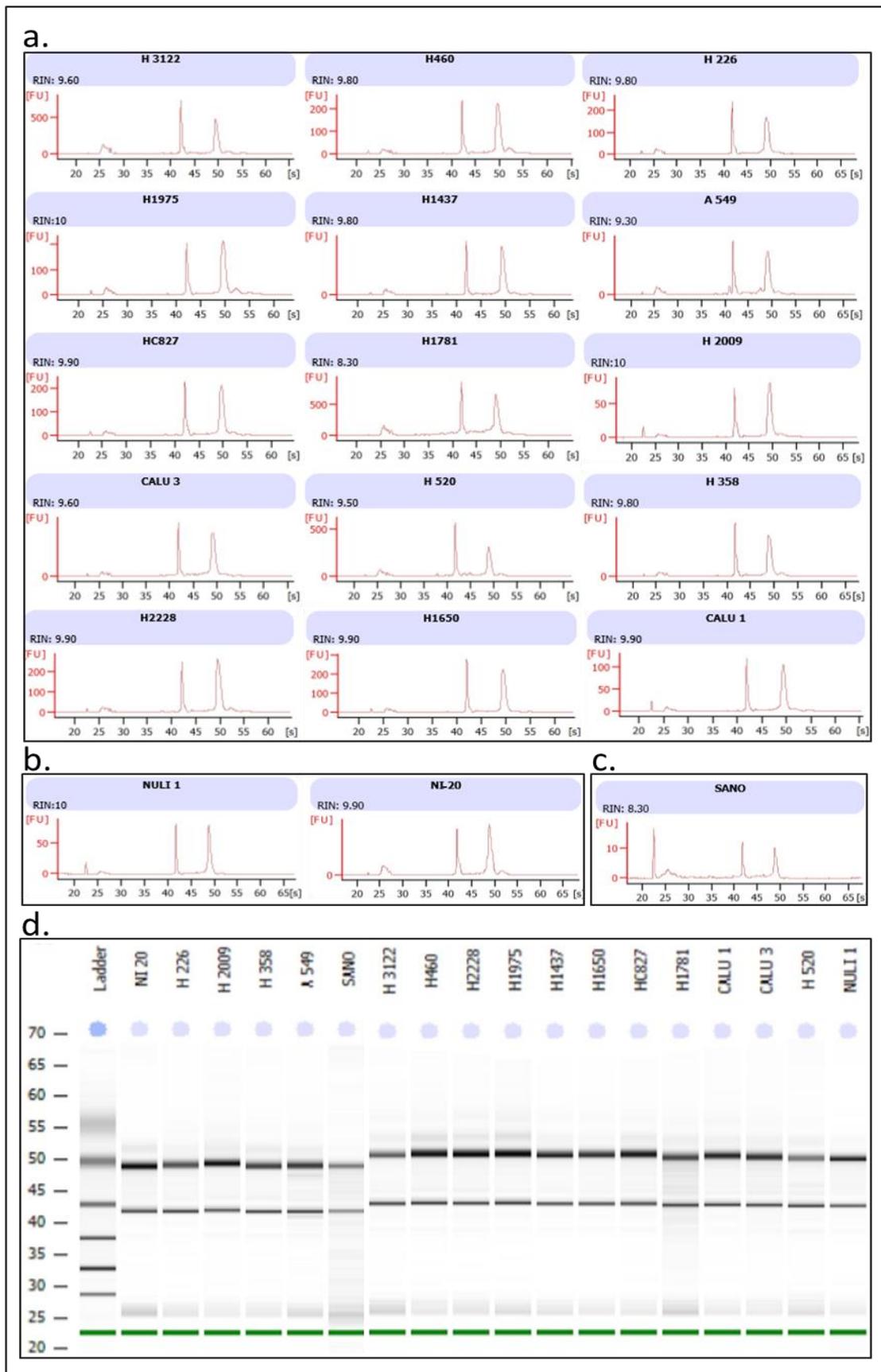


Figura 6. Electroferograma de las líneas celulares empleadas. a) líneas celulares tumorales; b) líneas celulares inmortalizadas no tumorales; c) mix celular tejido sano pulmonar; d) gel de electroforesis de todas las muestras.

5.3. EVALUACIÓN DE LA CALIDAD DE LOS DATOS DE LOS MICROARRAYS.

Para estudiar la calidad de los microarrays llevados a cabo, se emplearon los parámetros NUSE (valores de error estándar sin escala normalizados) y RLE (expresión logarítmica relativa).

Los valores NUSE reflejan el error estándar sin escala normalizado para cada set de sondas a partir de los valores de intensidad de cada sonda en el set. En el estudio, todas las muestras presentaron unos valores NUSE dentro del rango adecuado, cuyo valor medio ronda el 1.

Con respecto a los valores de RLE, en el supuesto biológico de que la mayoría de los genes no están diferencialmente expresados y que hay una misma cantidad de genes sobre o infrarregulados, el comportamiento colectivo de los datos de RLE pueden emplearse como un control de calidad adecuado. Nuestro array cumple con los criterios de calidad de acuerdo con el valor de RLE, cuyo mediana debe estar próxima a 0 y con un rango intercuartílico pequeño. Los datos en relación con dichos parámetros se representan en la figura 7.

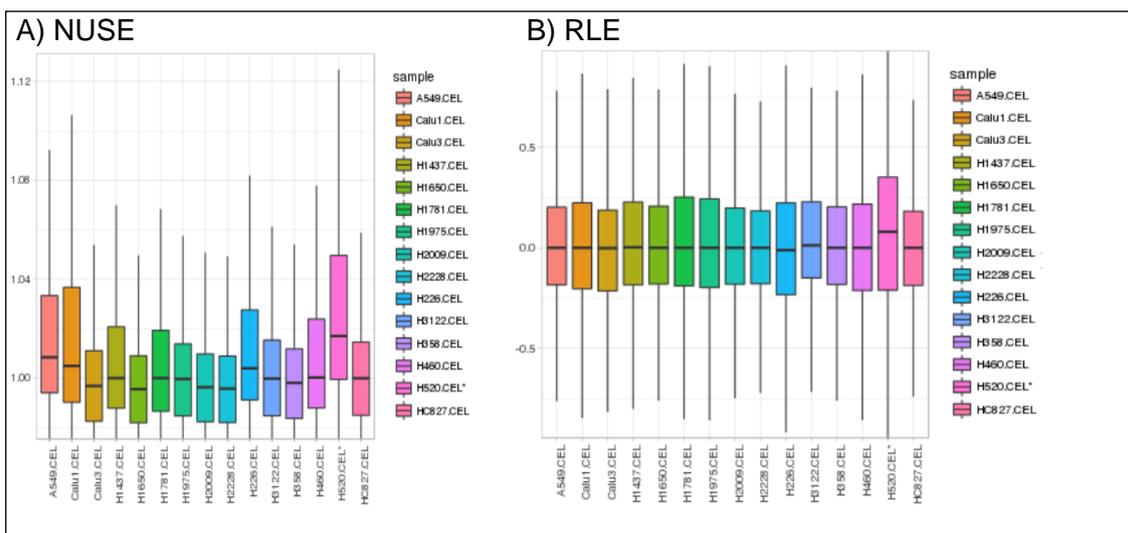


Figura 7. Representación gráfica mediante diagrama de cajas de los valores de NUSE (a) y RLE (b).

5.4. ANÁLISIS BIOINFORMÁTICO. IDENTIFICACIÓN DE GENES CON EXPRESIÓN DIFERENCIAL ENTRE LÍNEAS RADIOSENSIBLES Y RADIORRESISTENTES.

Con objeto de identificar genes diferencialmente expresados entre líneas sensibles y resistentes a la radiación, empleamos un análisis de microarray basado en cDNA. Nuestro panel de líneas celulares estaba compuesto mayoritariamente por el subtipo histológico de adenocarcinoma, pero también incluimos a líneas de estirpe escamosa y de células grandes, por lo que decidimos llevar a cabo un análisis bioinformático en el que se incluían todas las líneas celulares, con el objetivo de no perder poder estadístico.

5.4.1. Análisis bioinformático incluyendo todas las líneas celulares.

- **Agrupación no supervisada:** se representa en la figura 8. Permite observar si la expresión transcripcional es similar dando libertad a la agrupación en función de la transcripción diferencial de todos los genes. Observamos que no existe una diferenciación estadística de ambos subgrupos fisiológicos, aunque si cierta tendencia en la agrupación.

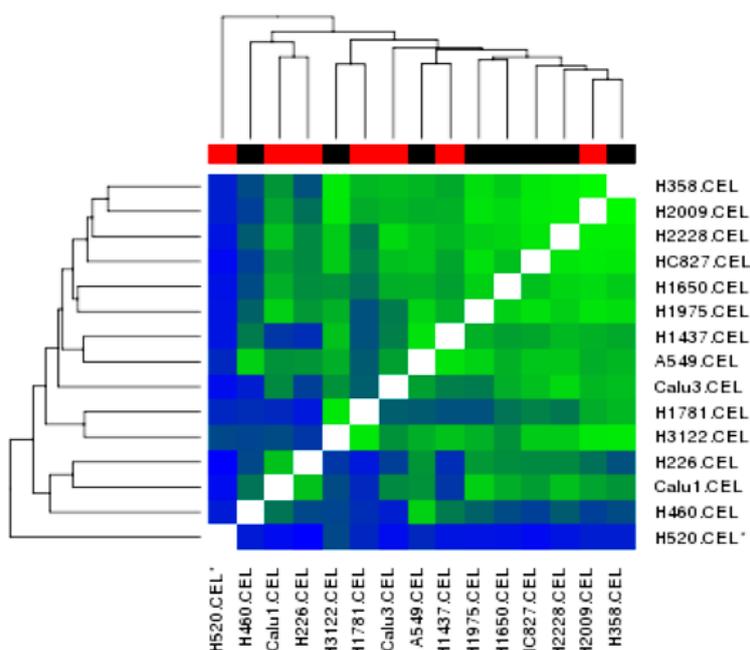


Figura 8. Representación gráfica de la agrupación no supervisada. En la parte superior se representa, mediante color rojo, a las líneas celulares clasificadas como radiosensibles y en negro las radiorresistentes.

- **Agrupación supervisada:** Permite observar la expresión transcripcional diferencial en el subconjunto de radioresistentes y diferente al de radiosensibles. Mediante la aproximación FDR (o false discovery rate) no se obtuvieron valores de $p < 0,05$, por lo que establecimos unos límites más relajados para la expresión diferencial (DE). Consideramos como punto de corte aquellos con un $|\log_2FC| > 1$ y $p < 0,05$. Obtuvimos un total de 127 genes con expresión diferencial (figura 9), de los cuales 76 se encontraban sobreexpresados y 51 infraexpresados (ver listado de genes en ANEXO 3).

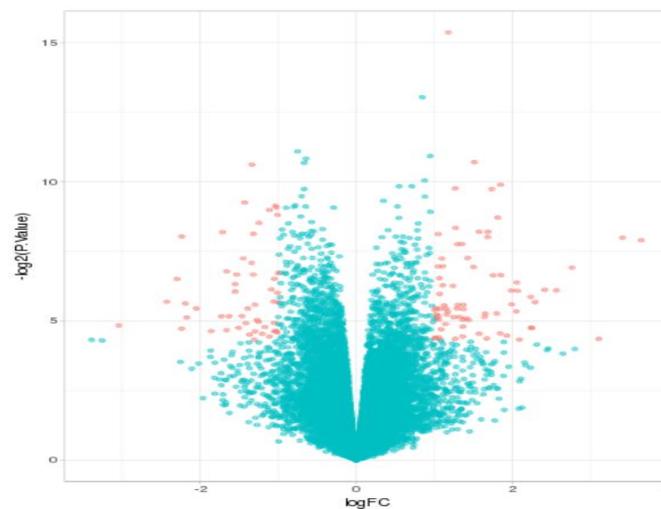


Figura 9. Diagrama tipo volcano. Se representa el \log_2FC en el eje de abscisas y el logaritmo en base 2 del valor p en eje de ordenadas. En color rosa se representan los 127 genes de interés. En verde el resto de los genes que no cumplen las condiciones preestablecidas en el análisis.

Para representar los resultados de esta aproximación, se llevó a cabo un mapa de calor (heatmap) en el que se establece una combinación de escala colorimétrica (azul-rojo) y agrupación por similitud en los cambios de expresión génica por parte de los genes de interés, representado mediante dendogramas. En las columnas se representan las líneas celulares (en rojo, las radiosensibles y en negro las radioresistentes), ordenadas de acuerdo con el dendograma situado en la parte superior de la figura. En las filas se representan los genes de interés, agrupados de acuerdo con el dendograma mostrado en la parte izquierda de la figura. Se representa en azul oscuro un nivel de cambio de 2 y en rojo oscuro, un nivel de cambio de 14.

A pesar de no existir un marcado patrón diferencial de expresión génica que nos permita hacer una agrupación perfecta entre líneas radiosensibles y radiorresistentes, podemos apreciar una clara tendencia de expresión y algunos datos interesantes. Así, podemos establecer 4 subgrupos de genes con una tendencia de expresión diferencial entre líneas sensibles y resistentes (figura 10). En los grupos de genes 1, 2 y 4, se aprecia una tendencia hacia niveles de expresión mayores en las líneas consideradas radiorresistentes versus las sensibles. Por el contrario, el grupo número 3 muestra una tendencia hacia mayores niveles de expresión en las líneas radiosensibles.

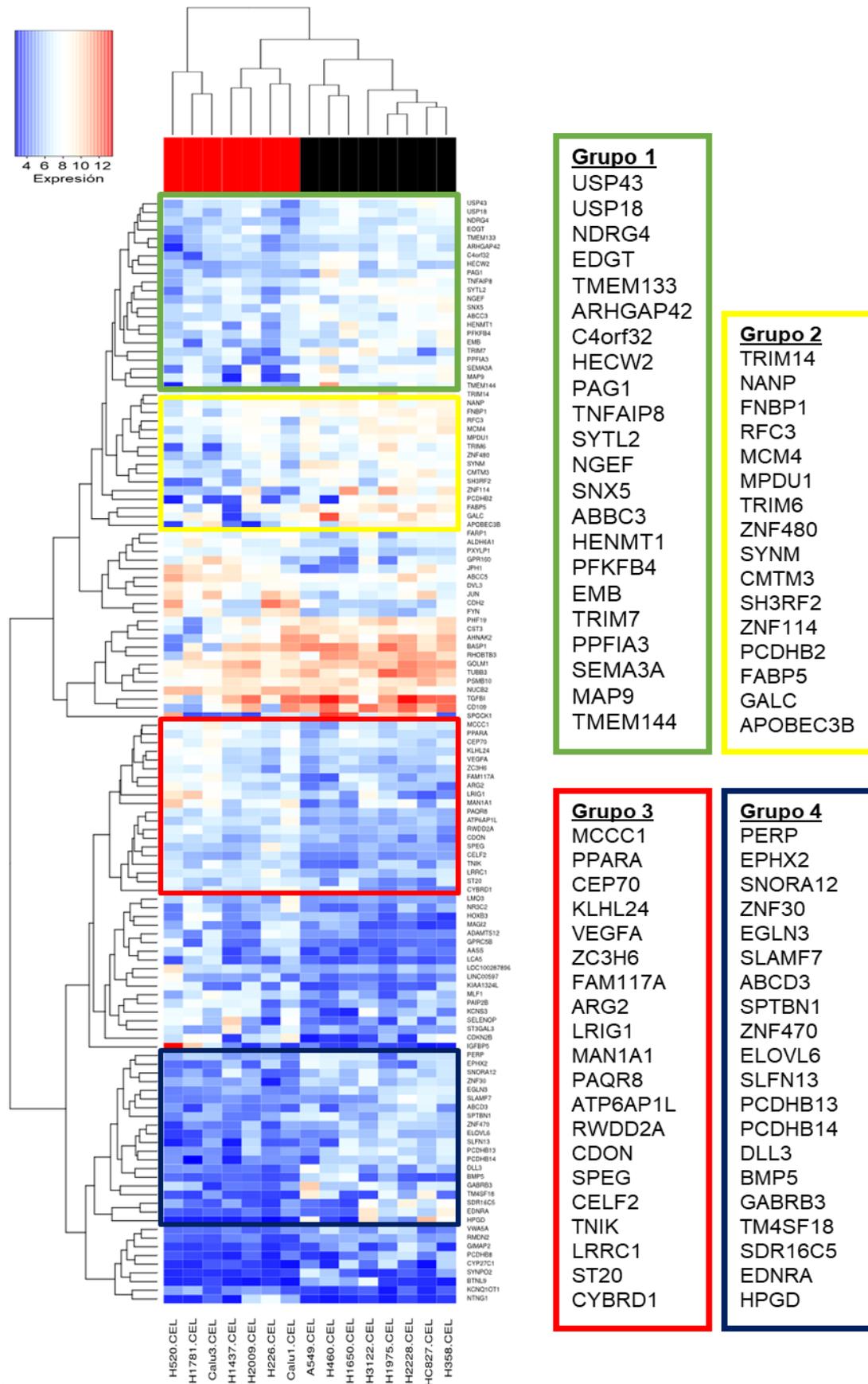


Figura 10. Mapa de calor representando el resultado de la agrupación supervisada de los genes de interés.

5.4.2. Resultados del análisis mediante enriquecimiento en funciones moleculares, componentes celulares y vías de señalización.

Enriquecimiento en funciones moleculares:

De los 127 genes de interés obtenidos dentro del análisis de todas las líneas celulares, con un nivel de cambio (\log_2FC) ≤ -1 y ≥ 1 y un valor de $p < 0,05$, se identificaron por parte del software libre de análisis de ontologías (panther.org®) un total de 125. De esos 125 genes, el 48% (60 genes) fueron relacionados con un total de 7 ontologías principales de función molecular (figura 11). A su vez, de esos 60 genes, 36 estaban relacionados con una ontología principal única (17 genes relacionados con “actividad catalítica”, 15 genes en relación con “unión” y 4 genes en relación con “actividad transportadora”). Los restantes 24 genes se podían relacionar con dos o más ontologías principales. La tabla 5 muestra la relación de genes con información ontológica relacionada con funciones moleculares principales. En dicha tabla se muestran los genes que solo se relacionan con una ontología principal, así como aquellos que comparten al menos 2 ontologías.

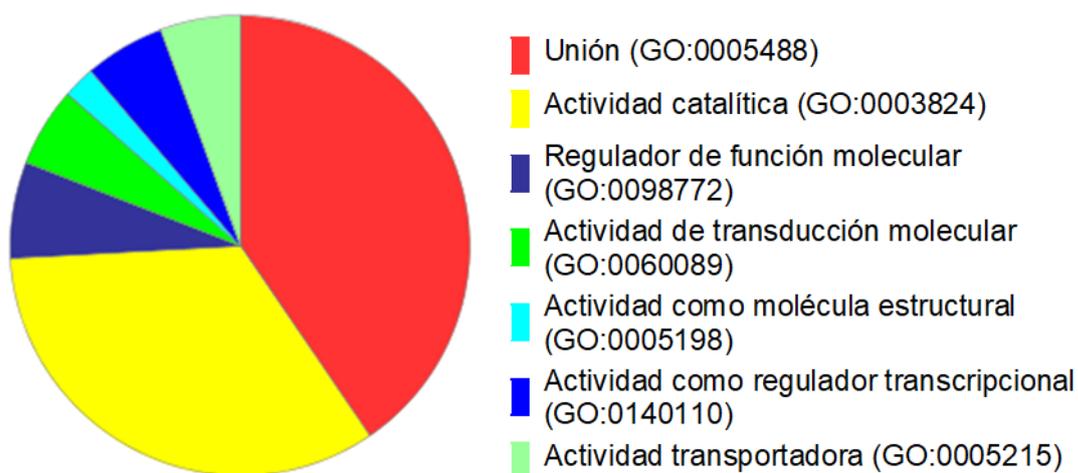


Figura 11. Diagrama circular en el que se representan las ontologías principales de función molecular a las que se relacionan los genes de interés. GO: ontología génica (identificador).

Tabla 5. Listado de genes vinculados con ontologías principales relacionadas con funciones moleculares.

ONTOLOGÍAS PRINCIPALES	Nº GENES	GENES
ACTIVIDAD CATALÍTICA	17	ELOVL6, SPEG, TNIK, PFKFB4, CYP27C1, EOGT, MCCC1, ABCC5, HENMT1, CYBRD1, ST3GAL3, AASS, USP18, ALDH6A1, PSMB10, MAN1A1, ABCC3
UNIÓN	15	LRRC1, AHNAK2, CELF2, EMB, SYMPO2, DVL3, NUCB2, SELENOP, SLFN13, MLF1, TGFB1, IGFBP5, SPOCK1, CDH2, BTNL9
UNIÓN ACTIVIDAD CATALÍTICA	9	RFC3, RHOBTB3, SH3RF2, FYN, ARG2, SLAMF7, MCM4, SYNM, APOBEC3B
UNIÓN ACTIVIDAD COMO REGULADOR TRANSCRIPCIONAL	4	NR3C2, PHF19, JUN, HOXB3
ACTIVIDAD TRANSPORTADORA	4	GABRB3, ABCD3, JPH1, KCNS3
UNIÓN ACTIVIDAD CATALÍTICA REGULADOR DE FUNCIÓN MOLECULAR	2	ADAMTS12, GPRC5B
UNIÓN REGULADOR DE FUNCIÓN MOLECULAR ACTIVIDAD DE TRANSDUCCIÓN MOLECULAR	2	BMP5, VEGFA
UNIÓN ACTIVIDAD COMO MOLÉCULA ESTRUCTURAL	2	TUBB3, MAGI2
UNIÓN ACTIVIDAD DE TRANSDUCCIÓN MOLECULAR ACTIVIDAD COMO REGULADOR TRANSCRIPCIONAL	1	PPARA
UNIÓN ACTIVIDAD DE TRANSDUCCIÓN MOLECULAR	1	PAQR8
ACTIVIDAD CATALÍTICA REGULADOR DE FUNCIÓN MOLECULAR	1	ARHGAP42
ACTIVIDAD CATALÍTICA ACTIVIDAD TRANSPORTADORA	1	HECW2
REGULADOR DE FUNCIÓN MOLECULAR ACTIVIDAD DE TRANSDUCCIÓN MOLECULAR	1	SEMA3A

Haciendo un breve análisis por cada función molecular principal, encontramos que de los 36 genes relacionados con “unión”, el 43,8 y 29,2% de ellos se relacionan con las ontologías “unión a proteínas” y la “unión de compuesto heterocíclico”, respectivamente.

Con respecto a los 30 genes relacionados con “actividad catalítica”, 13 (39.4%) y 12 genes (36.4%) se relacionan con “actividad hidrolasa” y “actividad transferasa”. Considerando los 6 genes relacionados con la ontología principal “regulador de función molecular”, el 50% está relacionado con “actividad reguladora enzimática” y el otro 50% en relación con “actividad reguladora de receptor”.

Los 5 genes relacionados con la ontología principal “actividad de transducción molecular” se asocian a la ontología secundaria “actividad de señalización de receptor”.

La ontología principal “actividad reguladora de transcripción” se encuentra asociada a 5 genes. Esos están asociados a la ontología secundaria “actividad como factor de transcripción unido a DNA”. Así mismo, 2 de esos 5 genes también se relacionan con la “actividad correguladora de la transcripción”.

Los 5 genes relacionados con la ontología principal “actividad transportadora” se relacionan con la ontología secundaria “actividad transportadora transmembrana”.

Por último, los 2 genes relacionados con la ontología “actividad de molécula estructural” se relación, así mismo, con la ontología secundaria “actividad relacionada con complejos proteínicos de soporte”

En la figura 12 se resumen de forma gráfica las ontologías secundarias relacionadas con las funciones moleculares principales a las que se vinculan 60 de los 127 genes de interés del presente análisis.

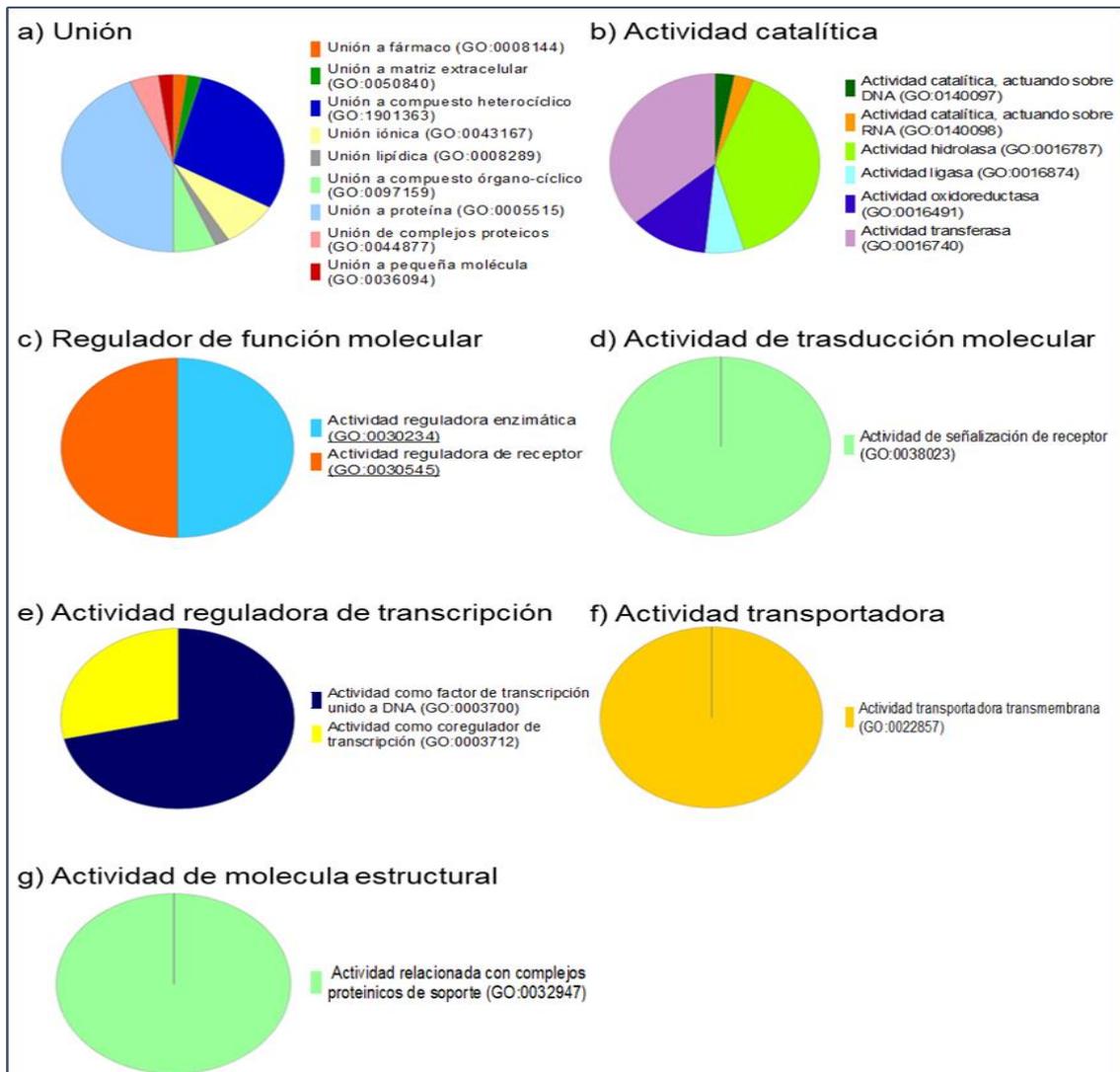


Figura 12. Se representan mediante diagrama circular, las principales ontologías secundarias asociadas a cada función molecular principal. GO: ontología génica (identificador)

Enriquecimiento en componente celular (CC):

Sesenta y uno de los 125 genes identificados por el software de análisis de ontologías génicas se vincularon con 7 ontologías principales relacionadas con componente celular. En la figura 13 se representan dichas ontologías mediante gráfico circular. La ontología “célula” aglomera al 36,6% de los genes, seguido por “organela” (21,5%), “región extracelular” y “membrana” con un 14% cada una, “complejo que contiene proteínas” (9,7%), “unión celular” (3,2%) y por último “complejo supramolecular” con un 1,1% de los genes.

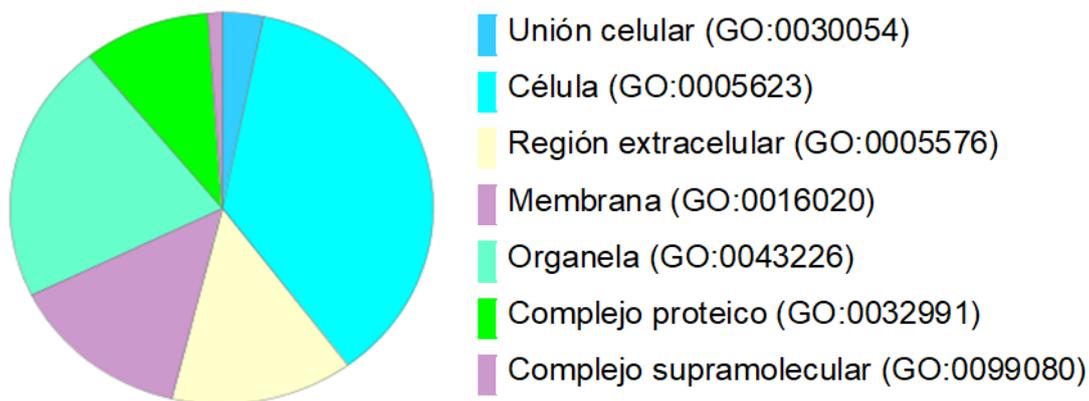


Figura 13. Gráfico circular donde se representan las ontologías principales de componente celular a las que se relacionan los genes de interés. GO: identificador de ontología génica.

La ontología principal “célula” resulta ser muy genérica, y comprende hasta un total de 98 ontologías secundarias y terciarias. El 50% aproximadamente de dichos genes se relacionan con las ontologías secundarias “parte intracelular” e “intracelular”, que a su vez comprende a elementos del citoplasma, orgánulos asociados o no a membrana, etc.

La ontología principal “organela” (a la que se asocian el 21.5% de los genes), se sigue de ontologías secundarias relacionadas con orgánulos asociados o no asociados a membrana, así como a vesículas intracelulares.

Dentro de la ontología principal “región extracelular” (14% de los genes), podemos encontrar las ontologías secundarias “matriz extracelular” y “espacio extracelular”.

También el 14% de los genes se relaciona con “membrana”, que a su vez se relacionan con hasta 35 ontologías secundarias y terciarias distintas, todas ellas principalmente relacionadas con las distintas partes de la membrana plasmática.

EL 9,7% de los genes se relacionan con “complejo proteico” que, al igual que sucede con los previos, engloba hasta un total de 10 ontologías secundarias relacionadas con complejos proteicos asociados a receptor, a proteasoma, complejos ribonucleoproteicos, complejos catalíticos, etc.

El 3,2% de los genes asociados a la ontología principal “unión celular” se relaciona con la ontología secundaria “unión adherente célula-célula”

Por último, el 1,1% de los genes se relaciona con la ontología “complejo supramolecular”, que se asocia como ontología secundaria y terciaria a “sarcómero” y “disco z” respectivamente.

Por otro lado, en la tabla 6 se identifican los 61 genes de interés y las ontologías principales a las que se relacionan, de ellos 37 se asocian de forma exclusiva a una ontología principal (15 genes a “célula”; 9 a “región extracelular”; 6 a “membrana” y “organela” y 1 gen a “complejos proteicos”. Los 24 genes restantes se encuentran asociados a 2 o más ontologías.

Tabla 6. Genes asociados a ontologías principales relacionadas con componente celular.

ONTOLOGÍA	Nº GENES	GENES
CELULA	15	NDRG4, AHNAK2, RHOBTB3, HECW2, TNIK, MPDU1, PFKFB4, TNFAIP8, EMB, DVL3, GIMAP2, PAQR8, CYBRD1, AASS, ALDH6A1
REGION EXTRACELULAR	9	VEGFA, CST3, NTNG1, ADAMTS12, LRIG1, SELENOP, TGFBI, IGFBP5, SPOCK1
MEMBRANA	6	PCDHB2, PCDHB13, TM4SF18, CMTM3, PCDHB14, PCDHB8
ORGANELA	6	SH3RF2, PPARA, HOXB3, SDR16C5, MLF1, APOBEC3B
CELULA ORGANELA	6	TUBB3, ARG2, JUN, HENMT1, SYTL2, PSMB10
CELULA MEMBRANA	4	ELOVL6, GABRB3, FYN, BTNL9
CELULA COMPLEJOS PROTEICOS ORGANELA	3	CELF2, JPH1, LCA5
COMPLEJOS PROTEICOS ORGANELA	2	RFC3, PHF19
CELULA MEMBRANA UNION CELULAR	1	LRRC1
COMPLEJO SUPRAMOLECULAR ORGANELA	1	SYNPO2
COMPLEJOS PROTEICOS	1	ATP6AP1L
CELULA COMPLEJOS PROTEICOS UNION CELULAR	1	CDH2
CELULA ORGANELA REGION EXTRACELULAR	1	MAN1A1
CELULA UNION CELULAR	1	MAGI2
CELULA REGION EXTRACELULAR	1	BMP5
MEMBRANA REGION EXTRACELULAR	1	SEMA3A
COMPLEJOS PROTEICOS MEMBRANA	1	KCNS3
CELULA COMPLEJOS PROTEICOS ORGANELA REGION EXTRACELULAR	1	GPRC5B

Enriquecimiento en vías de señalización (pathways):

En el análisis de enriquecimiento en pathways (mediante el software PANTHER®), obtuvimos información sobre 24 genes, los cuales se relacionaban con 31 pathways diferentes (figura 12). Destaca la “vía de wnt”, la “vía de señalización de

cadherina” y “angiogénesis”, con 6, 6 y 4 genes, respectivamente. La vía de wnt y cadherina comparten 5 de los 6 genes (de la familia de las cadherinas y proto-cadherinas). Ambas vías y sus relaciones se han asociado con el proceso tumorigénico en CPCNP, su pronóstico, así como con la eficacia de los tratamientos oncológicos empleados comúnmente (radio y quimioterapia).

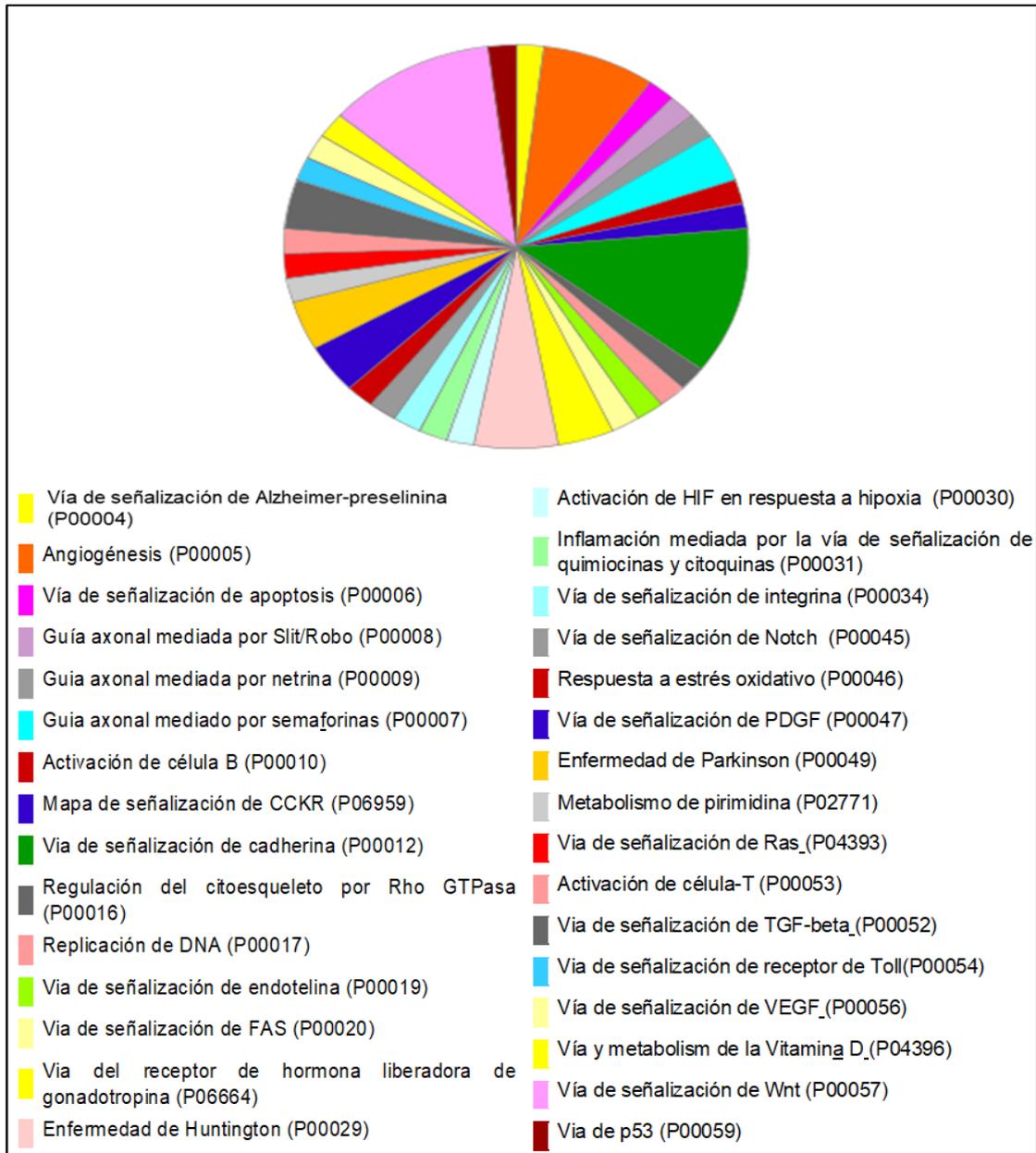


Figura 14. Representación mediante gráfico circular de los pathways relacionados con los genes de interés del análisis. (P00000): identificador de pathway Panther

Haciendo el mismo análisis por enriquecimiento en KEGG, obtuvimos información asociada a 21 de los 127 genes. En la tabla 7 se muestran las ontologías significativas asociadas a cada uno de los genes. Hay que destacar la presencia de rutas de señalización relacionadas con cáncer (transportadores ABC (ABCC3, ABCD3 y ABCC5), que median en fenómenos de resistencia a fármacos antineoplásicos); vías de señalización en carcinoma renal (EGLN3, JUN y VEGFA) y en replicación del DNA (MCM4 y RFC3); cascadas moleculares relacionadas con la vía de Notch (DVL3 y DLL3) comúnmente desregulado en cáncer y vías de señalización en cáncer en general (EDNRA y CDKN2B).

Tabla 7. Resumen del análisis por enriquecimiento en KEGG considerando todas las líneas celulares. Se representan los genes asociados a las ontologías más significativas resultantes del análisis.

IDENTIFICADOR	NOMBRE PATHWAY	Nº GENES	GENES
path:hsa04360	Guía axonal	4	FYN, SEMA3A, NGEF, NTNG1
path:hsa02010	Transportadores ABC	3	ABCC3, ABCD3, ABCC5
path:hsa05211	Carcinoma Renal	3	EGLN3, JUN, VEGFA
path:hsa03030	Replicación de DNA	2	MCM4, RFC3
path:hsa04330	Vía señalización de Notch	2	DVL3, DLL3
path:hsa00514	Otros tipos de biosíntesis de O-glucanos	2	EOGT, ST3GAL3
path:hsa05200	Vías en cáncer	2	EDNRA, CDKN2B
path:hsa00280	Degradación de valina, leucina e isoleucina	2	ALDH6A1, MCCC1
path:hsa05130	Infección patogénica por E.Coli	1	TUBB3

5.4.3. Validación individual de genes con expresión diferencial como biomarcador pronóstico en la cohorte de interés del TCGA.

Empleamos la cohorte pública del TCGA de pacientes afectados de CPCNP (Pan Cancer atlas) (actualización 2018) para la validación de nuestros resultados. Dicha cohorte se componía de CPCNP tipo adenocarcinoma y tipo escamoso. Los criterios clínicos principales para la selección de los pacientes fueron 1) haber recibido tratamiento radioterápico y 2) ser pacientes clasificados como no metastásicos al diagnóstico. Con dichos criterios de inclusión seleccionamos un total de 107 pacientes.

De la explotación bioinformática de los resultados obtenidos considerando todas las líneas celulares, se obtuvieron un total de 127 genes con expresión diferencial entre las líneas radiosensibles y radorresistentes (valor absoluto de $\log_2FC > 1$ y valor de p ajustado < 0.05).

Al analizar los datos de la cohorte seleccionada del TCGA sobre el RNAseq de los 127 genes reflejados más arriba, identificamos la falta de información sobre 6 de ellos (C4orf32, KIAA1324L, LINC00597, LOC100287896, SELENOP y TMEM133). Así pues, considerando los niveles de expresión de los 121 genes restantes, se llevó a cabo un análisis de supervivencia mediante regresión de Cox univariante, en el que obtuvimos un total de 21 genes, cuyo nivel de expresión se relacionaba de forma estadísticamente significativa con la SG en la cohorte de pacientes seleccionada (tabla 8).

Tabla 8. Genes con un impacto significativo ($p < 0.05$) en la supervivencia global medido mediante la prueba de regresión de Cox univariante. Se muestran ordenados de forma descendente por el valor de p .

Gen	Beta	HR (IC 95% para HR)	wald.test	Valor p
APOBEC3B	-0.303	0.738 (0.623-0.875)	12.30	0.000456
GOLM1	0.352	1.42 (1.15-1.76)	10.20	0.001430
FAM117A	-0.473	0.623 (0.461-0.843)	9.43	0.002130
CDH2	0.283	1.33 (1.09-1.61)	8.26	0.004060
KCNQ1OT1	1.19	3.27 (1.44-7.43)	8.03	0.004600
SDR16C5	0.197	1.22 (1.05-1.41)	7.10	0.007730
KLHL24	-0.373	0.688 (0.512-0.926)	6.10	0.013500
PCDHB13	0.315	1.37 (1.06-1.78)	5.67	0.017300
PCDHB8	0.244	1.28 (1.04-1.56)	5.66	0.017400
PCDHB2	0.222	1.25 (1.04-1.5)	5.61	0.017900
RHOBTB3	0.24	1.27 (1.04-1.55)	5.49	0.019100
NUCB2	0.468	1.6 (1.08-2.37)	5.44	0.019700
BASP1	0.194	1.21 (1.03-1.43)	5.27	0.021800
TRIM6	0.284	1.33 (1.02-1.72)	4.58	0.032400
ABCC5	-0.166	0.847 (0.726-0.987)	4.52	0.033600
EPHX2	-0.272	0.762 (0.587-0.989)	4.18	0.040900
LRIG1	-0.283	0.754 (0.574-0.989)	4.16	0.041400
TMEM144	0.341	1.41 (1.01-1.95)	4.11	0.042500
PXYLP1	-0.373	0.688 (0.479-0.989)	4.08	0.043300
TUBB3	0.194	1.21 (1.01-1.47)	4.05	0.044100
USP43	0.252	1.29 (1-1.65)	3.92	0.047600

Beta= coeficientes de regresión, HR= relación de riesgo (HR <1 significa reducción del riesgo y viceversa), IC =intervalo de confianza.

Centrados sobre los genes anteriormente expuestos, llevamos a cabo un análisis de estimación de supervivencia mediante curvas de Kaplan Meier, tomando la mediana de la expresión del gen como punto de corte para dividir las muestras en alto y bajo nivel de expresión. Un total de 10 genes de esos 21 identificados más arriba, arrojan un valor de $p < 0,05$ por la prueba de los rangos logarítmicos (figura 15).

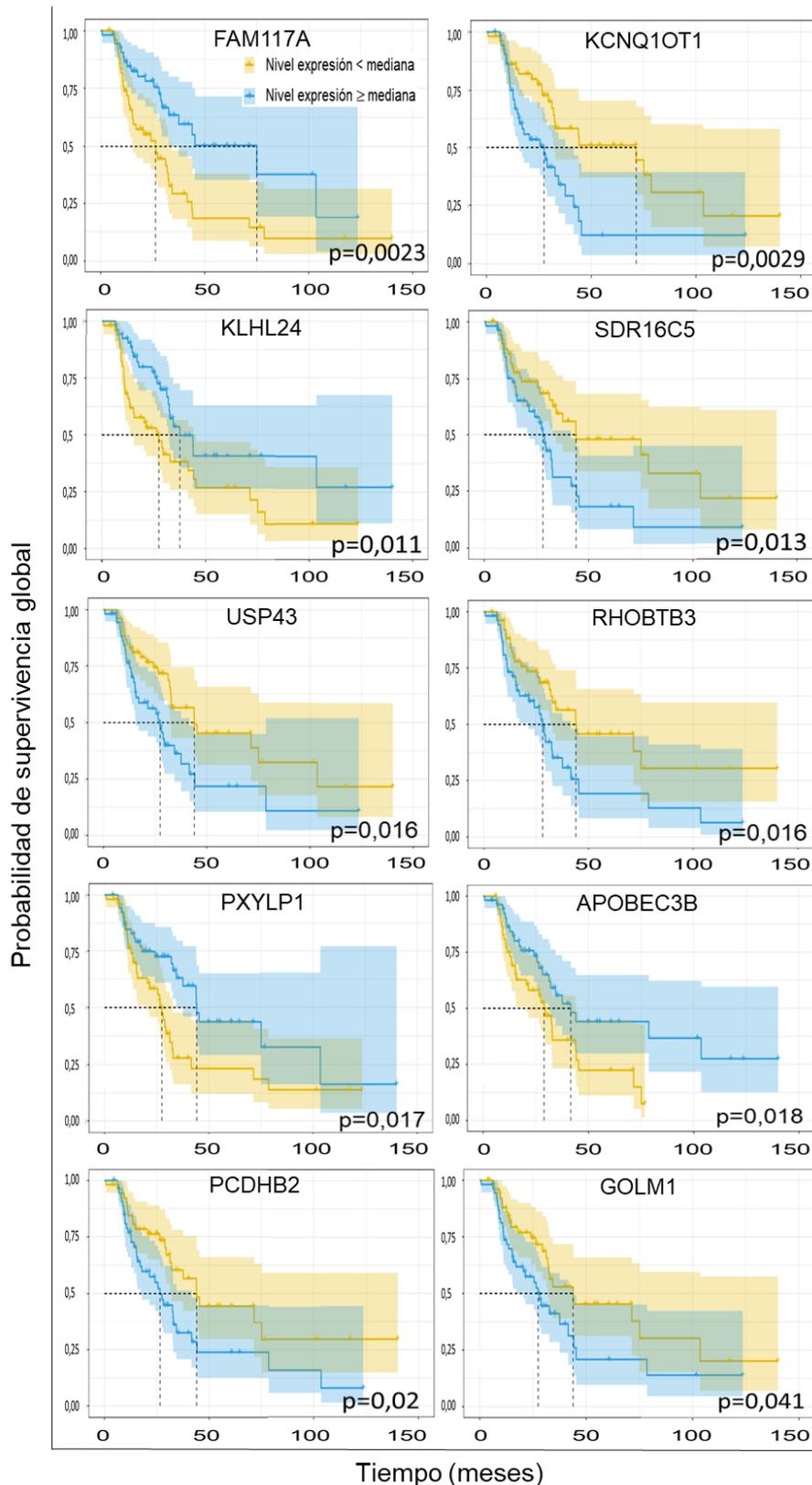


Figura 15. Curvas de Kaplan Meier con la información referente a los 10 genes de interés ordenados por nivel de significación estadística (valor p). En el eje de ordenadas se representa la probabilidad de supervivencia (0-1) y en el eje de abscisas, la supervivencia global en meses. En amarillo se representan los pacientes cuyo nivel de expresión del gen en cuestión era menor a la mediana de expresión del gen y en azul los pacientes con niveles de expresión iguales o superiores a la mediana de expresión del gen en cuestión.

Así pues, los genes FAM117A, KCNQ1OT1, KLHL24, SDR16C5, USP43, RHOBTB3, PXYLP1, APOBEC3B, PCDHB2 y GOLM1, son los que mantuvieron la significación estadística en el análisis de regresión de Cox univariante y en la estimación de SG por Kaplan Meier y, además de ser potenciales biomarcadores pronósticos de forma individual, son los seleccionados para la elaboración de la firma pronóstica.

5.4.4. Obtención de la firma pronóstica.

Seleccionamos los 10 genes identificados en el apartado previo como potenciales candidatos a formar parte de la firma pronóstica. A continuación, realizamos una regresión múltiple de Cox con una validación cruzada de 3 iteraciones, en la que 2/3 de los datos (71 pacientes) se usan como cohorte de entrenamiento y 1/3 (36 pacientes) como cohorte de ensayo. Aplicamos, además, una regularización Lasso (tipo L1) a la regresión, para llevar a cabo una selección de los genes que influyen en la supervivencia de forma más significativa. De acuerdo con el método descrito en el apartado de “materiales y métodos”, obtuvimos unos coeficientes (betas en la regresión de Cox) para los genes APOBEC3B, GOLM1, FAM117A, KCNQ1OT1, SDR16C5, PCDHB2, RHOBTB3 y USP43.

A continuación, obtuvimos un valor de firma para cada uno de los pacientes, partiendo de los valores de expresión (logTPM) de cada uno de los genes de interés multiplicados por sus respectivos coeficientes. Por último, llevamos a cabo un análisis de regresión de Cox univariante para todos los valores de la firma (puntuaciones de riesgo).

La puntuación de riesgo es el resultado de la combinación lineal del coeficiente de regresión multiplicado por el nivel de expresión de cada gen en cuestión, resumido en la siguiente fórmula:

$$\text{Puntuación riesgo} = (-0.1353493) * \text{Expresión APOBEC3B} + 0.0671513 * \text{Expresión GOLM1} + (-0.4002509) * \text{Expresión FAM117A} + 0.6455878 * \text{Expresión KCNQ1OT1} + 0.0613337 * \text{Expresión SDR16C5} + 0.0542689 * \text{Expresión PCDHB2} + 0.0309122 * \text{Expresión RHOBTB3} + 0.0834290 * \text{Expresión USP43}$$

Mediante un análisis por regresión de Cox univariante de la firma pronóstica, los pacientes asociados al grupo pronóstico adverso muestran una SG significativamente más pobre que aquellos asociados al grupo favorable, con una HR de 3.9 (IC95% 2.39-6.37) y una $p < 0,0001$. El índice de concordancia fue de 0,713 +/- 0.0374.

El análisis de predicción de la supervivencia en función de la firma génica propuesta mediante curvas ROC a 12, 24 y 60 meses, arrojó un valor de AUC de 0.73, 0.72 y 0.79, respectivamente (figura 16). Para el análisis mediante Kaplan Meier (figura 17), consideramos como punto de corte para separar a los dos grupos pronósticos, a la mediana de los valores de las puntuaciones de riesgo. Dicho valor es de -0.8. En amarillo se identifican los pacientes cuya firma génica muestra unos niveles de expresión menores al punto de corte y que se asocian a mejor supervivencia (mediana de SG de 71.47 meses (IC 95% 36,7-106,2) y en azul los pacientes en los que los valores de la firma quedan igual o por encima de dicho punto y que se relaciona con peor pronóstico (mediana de SG de 24,13 meses (IC al 95% 11.98-36,27). Los valores de expresión de cada gen en cada uno de los dos grupos diferenciados se exponen en la figura 18 mediante diagrama de cajas. Todos los genes presentan un nivel de expresión diferencial estadísticamente significativo en función de su relación con cada grupo de riesgo.

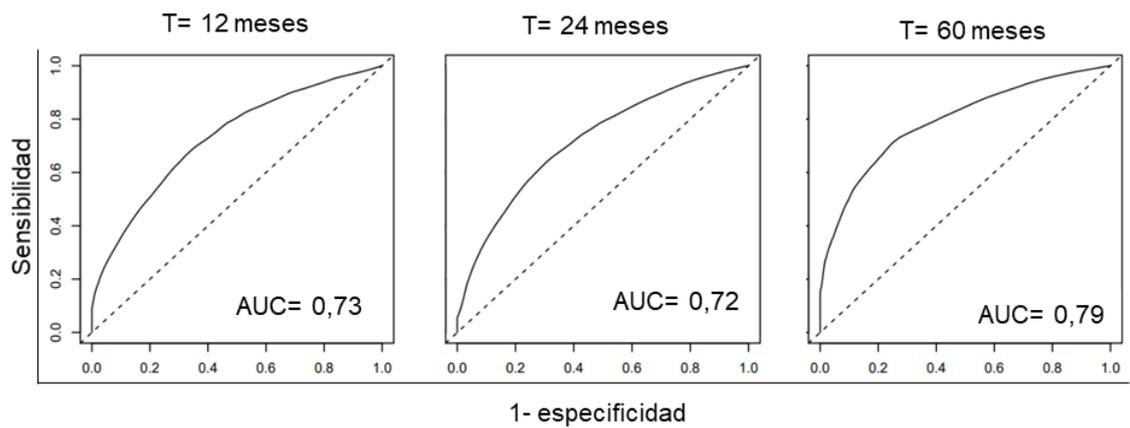


Figura 16. Curvas ROC en función del tiempo (T). AUC: área bajo la curva.

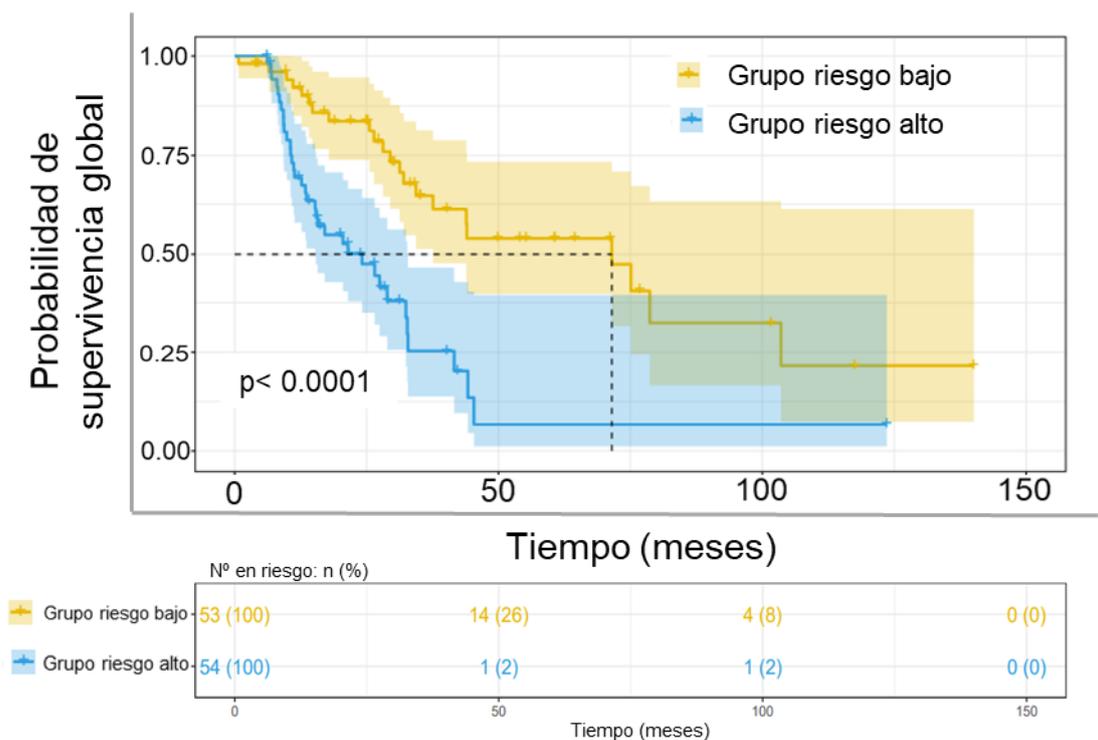


Figura 17. Curvas de Kaplan-Meier en la que se representa la probabilidad de supervivencia global de los pacientes en función de su agrupación de acuerdo con el punto de corte establecido, que corresponde a la mediana de las puntuaciones de riesgo de cada paciente en la cohorte, separando a ésta en 2 grupos (alto o bajo riesgo, en función de si su puntuación de riesgo individual es mayor/igual o menor a dicho punto de corte, respectivamente). Se representa, así mismo, el número de pacientes en riesgo en cada grupo en función de los intervalos de tiempo establecidos.

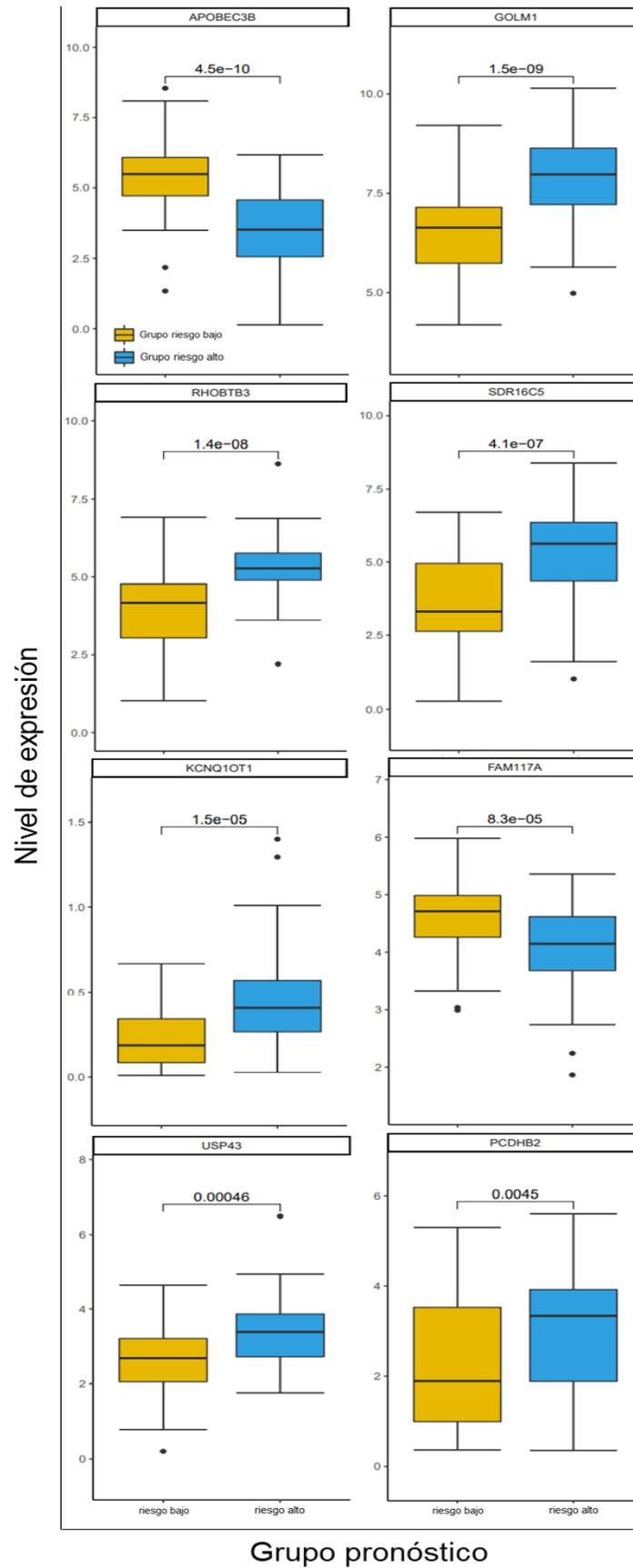


Figura 18. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) (normalizados) en eje de ordenadas, de cada gen que conforma la firma en función del grupo pronóstico (bajo y alto riesgo)

5.4.5. Validación individual de genes con expresión diferencial entre líneas radiosensibles y radiorresistentes como biomarcador predictivo de respuesta en la cohorte de interés del TCGA.

Hemos considerado a la SLP como variable para la valoración del efecto de los tratamientos oncológicos en nuestra cohorte de interés. Al igual que lo realizado en el apartado de SG, consideramos los 127 genes diferencialmente expresados e identificamos 121 de ellos en los datos de RNAseq de nuestra cohorte del TCGA. Procedimos a hacer un análisis de impacto en la SLP mediante la prueba de regresión de Cox univariante. En la tabla 9 se representan los 29 genes cuyo impacto en SLP resultó estadísticamente significativa.

Tabla 9. Genes con un impacto significativo ($p < 0.05$) en la SLP medido mediante la prueba de regresión de Cox univariante. Ordenados de forma descendente por el valor p.

Gen	Beta	HR (IC 95% para HR)	wald test	Valor p
GOLM1	0.467	1.6 (1.3-1.96)	20.20	7.00e-06
APOBEC3B	-0.323	0.724 (0.621-0.844)	17.00	3.75e-05
KCNQ1OT1	1.41	4.08 (1.82-9.15)	11.60	6.58e-04
PCDHB2	0.239	1.27 (1.08-1.5)	8.29	4.00e-03
TUBB3	0.257	1.29 (1.08-1.54)	8.13	4.37e-03
USP43	0.333	1.4 (1.11-1.76)	7.99	4.72e-03
NTNG1	0.369	1.45 (1.12-1.87)	7.78	5.29e-03
PCDHB8	0.259	1.3 (1.08-1.56)	7.76	5.35e-03
BASP1	0.209	1.23 (1.06-1.43)	7.55	6.01e-03
FARP1	0.429	1.54 (1.12-2.1)	7.29	6.95e-03
MPDU1	0.636	1.89 (1.19-3.01)	7.19	7.31e-03
ATP6AP1L	0.526	1.69 (1.14-2.51)	6.87	8.74e-03
PXYLP1	-0.413	0.662 (0.473-0.926)	5.80	1.60e-02
CYP27C1	0.229	1.26 (1.04-1.52)	5.77	1.63e-02
ABCC5	-0.17	0.844 (0.733-0.971)	5.59	1.81e-02
PAIP2B	0.319	1.38 (1.05-1.79)	5.54	1.86e-02
PCDHB13	0.279	1.32 (1.05-1.67)	5.52	1.88e-02
KLHL24	-0.33	0.719 (0.545-0.947)	5.51	1.90e-02
SDR16C5	0.159	1.17 (1.03-1.34)	5.47	1.94e-02
CDH2	0.208	1.23 (1.03-1.47)	5.44	1.97e-02
RHOBTB3	0.205	1.23 (1.03-1.47)	5.04	2.47e-02
HECW2	0.312	1.37 (1.04-1.8)	4.88	2.72e-02
FAM117A	-0.328	0.721 (0.538-0.966)	4.81	2.83e-02
SPTBN1	0.36	1.43 (1.03-1.99)	4.68	3.06e-02
CELF2	0.189	1.21 (1.01-1.45)	4.28	3.85e-02
JPH1	-0.212	0.809 (0.66-0.992)	4.13	4.21e-02
SLFN13	0.205	1.23 (1.01-1.5)	4.11	4.27e-02
ELOVL6	-0.212	0.809 (0.655-0.998)	3.93	4.75e-02
MLF1	-0.187	0.829 (0.688-0.998)	3.90	4.82e-02

Beta= coeficiente de regresión, HR= relación de riesgo (HR <1 significa reducción del riesgo y viceversa), IC =intervalo de confianza.

Centrados sobre los genes anteriormente expuestos, llevamos a cabo un análisis predictivo de SLP, mediante curvas de Kaplan Meier, tomando la mediana de la expresión del gen como punto de corte para dividir las muestras en alto nivel de expresión y bajo nivel de expresión. Las curvas de supervivencia con valor de $p < 0,05$ se muestran en la figura 19.

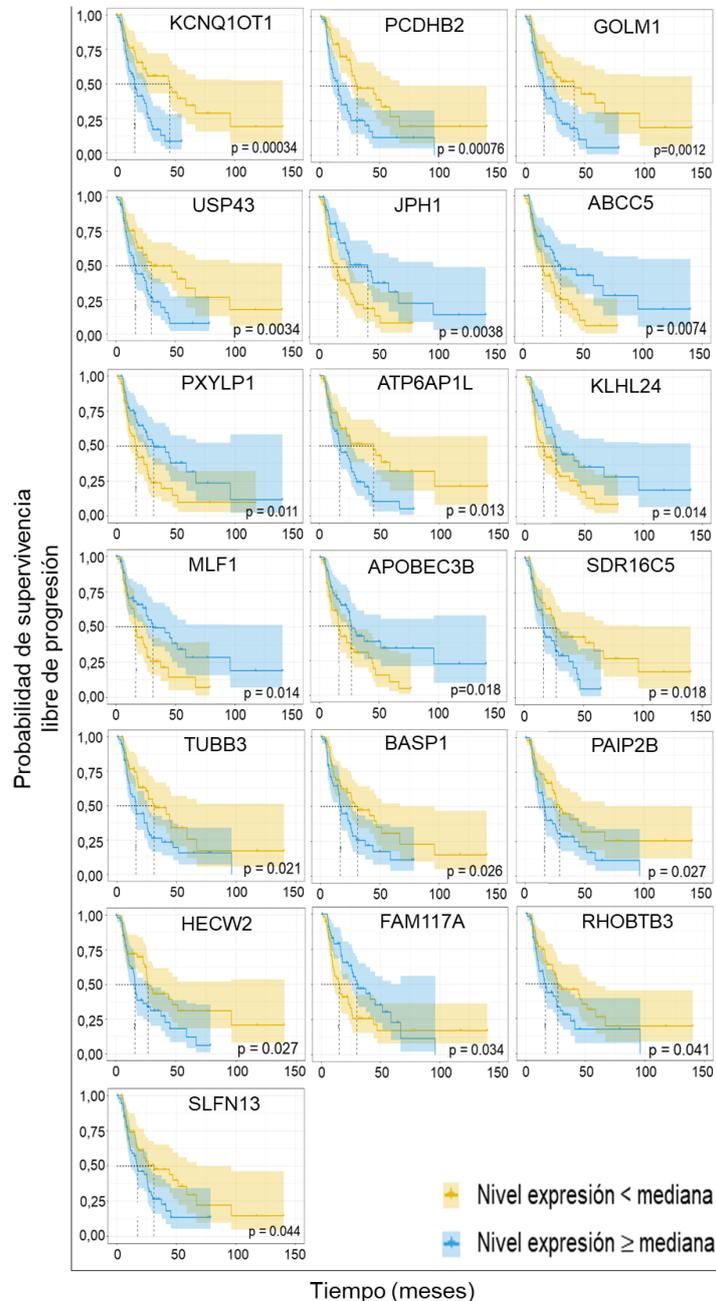


Figura 19. Curvas de Kaplan Meier con la información referente a los 19 genes de interés (valor $p < 0,05$) ordenados por nivel de significación estadística. En el eje de ordenadas se representa la probabilidad de SLP (0-1) y en el eje de abscisas, la SLP en meses. En amarillo se representan los pacientes cuyo nivel de expresión del gen en cuestión era menor a la mediana de expresión del gen y en azul los pacientes con niveles de expresión iguales o superiores a la mediana de expresión del gen en cuestión.

Así pues, los genes KCNQ1OT1, PCDHB2, GOLM1, USP43, JPH1, ABCC5, PXYLP1, ATP6AP1L, KLHL24, MLF1, APOBEC3B, SDR16C5, TUBB3, BASP1, PAIP2B, HECW2, FAM117A, RHOBTB3 Y SLFN13 son los que mostraron de forma individual un impacto en SLP de forma estadísticamente significativa por análisis de regresión de Cox y, además, mantuvieron la significación estadística en el análisis de estimación de SLP por Kaplan Meier. Seleccionamos dichos genes para desarrollar la firma predictiva de respuesta.

5.4.6. Validación conjunta de genes como firma predictiva de respuesta al tratamiento.

Considerando los niveles de expresión de los genes identificados en el apartado previo, llevamos a cabo un análisis por regresión de Cox multivariante con regularización de tipo L1 con validación cruzada de 3 iteraciones. Solo 7 de los 19 genes se mantenían en el modelo.

La puntuación de riesgo obtenida de igual forma a la descrita en el apartado previo queda resumida en la siguiente fórmula:

$$\text{Puntuación riesgo} = (-0.1462772) * \text{Expresión APOBEC3B} + 0.1824914 * \text{Expresión GOLM1} + (-0.1364778) * \text{Expresión FAM117A} + 0.5059698 * \text{Expresión KCNQ1OT1} + 0.0545114 * \text{Expresión PAIP2B} + 0.0624953 * \text{Expresión PCDHB2} + 0.0317391 * \text{Expresión USP43}$$

Calculamos la puntuación de riesgo en base a los 7 genes de interés sobre cada paciente y a continuación, mediante el paquete estadístico Survminer R, identificamos el punto de corte para dicha puntuación de riesgo, clasificando a todos los pacientes en cohorte de alto o bajo riesgo en función del punto de corte establecido. A continuación, se llevó a cabo un análisis por regresión de Cox univariante de la propia firma predictiva, obteniendo un valor de HR de 5,04 (IC95% 3-8,47) y $p < 0.001$. El índice de concordancia fue de 0.703.

Por último, llevamos a cabo un análisis de predicción de SLP en función de la firma génica propuesta. Para ello, realizamos un análisis mediante curvas ROC a 12, 24 y 60 meses (figura 20) y posteriormente un análisis mediante Kaplan Meier, considerando como punto de corte la mediana de las puntuaciones de riesgo, cuyo valor resultó ser + 0.72, para separar a los dos grupos pronósticos (figura 21). Los pacientes cuya firma génica muestran unos niveles de expresión menores o mayores/iguales a dicho punto de corte, obtuvieron una mediana de SLP de 51,55 (IC95% 20,1-83) y 15,45 meses (IC95% 10,1-20,8), respectivamente. Los valores de expresión de cada gen en cada uno de los dos grupos diferenciados se exponen en la figura 22 mediante diagrama de cajas. Todos los genes de la firma excepto PAIP2B, presentan un nivel de expresión diferencial estadísticamente significativo en función de su relación con cada grupo pronóstico. (test U de Mann-Whitney con $p < 0.05$).

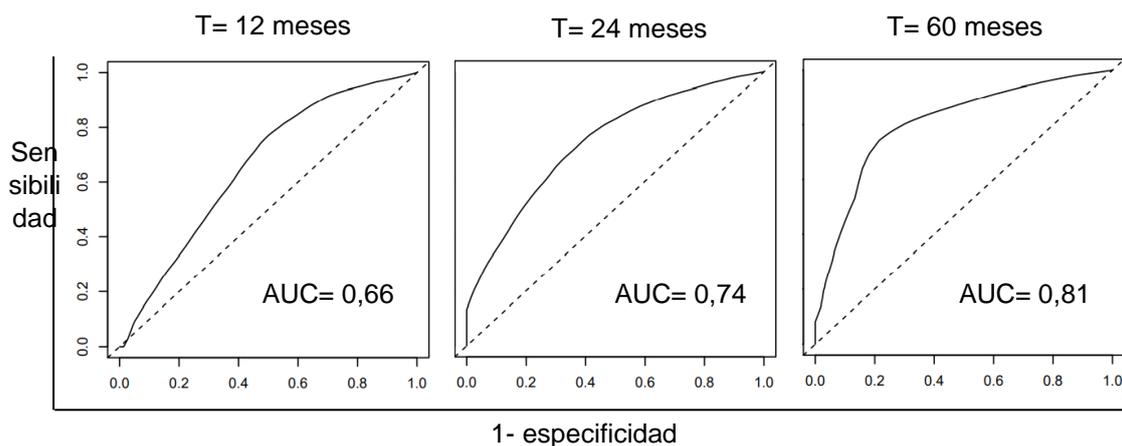


Figura 20. Curvas ROC de predicción de SLP en función del tiempo (T). AUC: área bajo la curva

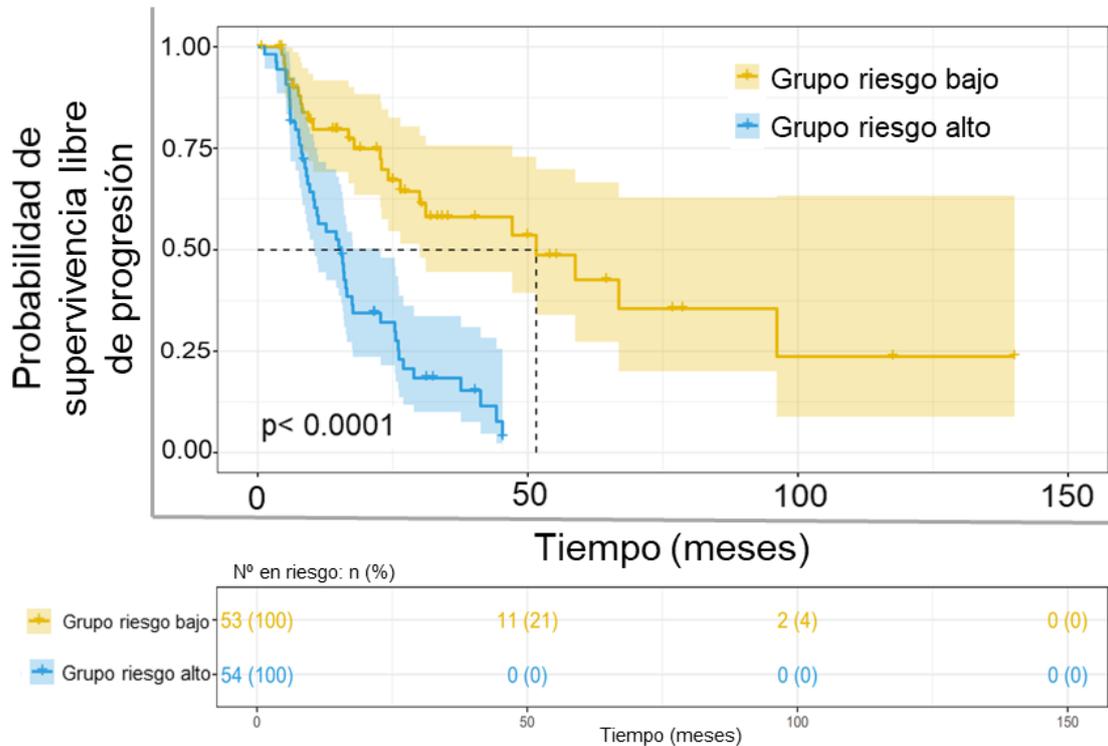


Figura 21. Curvas de Kaplan-Meier en la que se representa la probabilidad de SLP de los pacientes en función de su agrupación de acuerdo con el punto de corte establecido, que corresponde a la mediana de las puntuaciones de riesgo de cada paciente en la cohorte, separando a ésta en 2 grupos (alto o bajo riesgo, en función de si su puntuación de riesgo individual es mayor/igual o menor a dicho punto de corte, respectivamente). Se representa, así mismo, el número de pacientes en riesgo en cada grupo en función de los intervalos de tiempo establecidos.

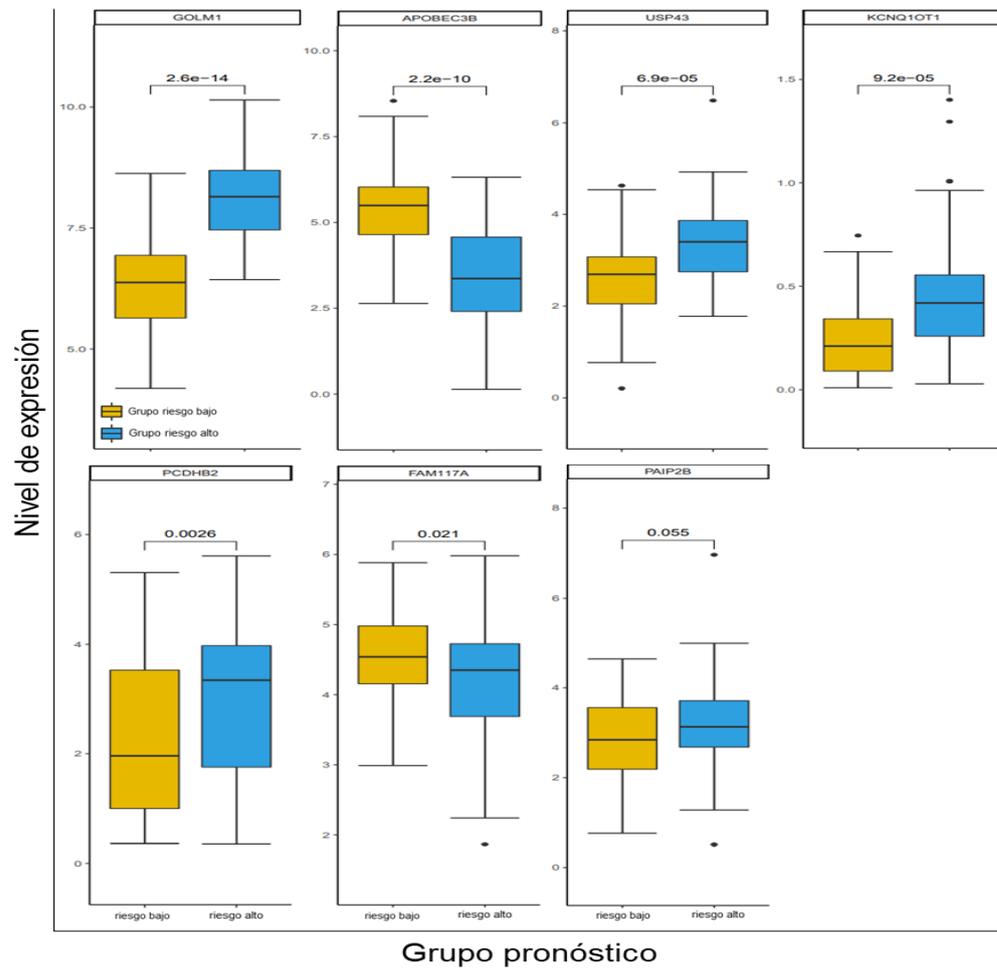


Figura 22. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) (normalizados) en eje de ordenadas de cada gen en función de cada grupo pronóstico (bajo o alto riesgo).

5.5. GENERACIÓN DE FIRMA GÉNICA COMÚN PARA PREDICCIÓN DE SUPERVIVENCIA GLOBAL Y SUPERVIVENCIA LIBRE DE PROGRESIÓN.

Tras la obtención de las firmas génicas con valor pronóstico y predictivo de respuesta, objetivamos que entre ambas se compartían 6 elementos génicos (APOBEC3B, GOLM1, FAM117A, KCNQ1OT1, PCDHB2 y USP43), lo que nos llevó a la generación de una firma única con capacidad pronóstica de SG y predictiva de respuesta. Para ello, siguiendo igual metodología a la empleada en la obtención de las firmas previas, pero en este caso empleando los datos de SG de la cohorte, obtuvimos una puntuación de riesgo resumida en la siguiente fórmula:

$$\text{Puntuación riesgo} = (-0.1783969) * \text{Expresión APOBEC3B} + 0.0877709 * \text{Expresión GOLM1} + (-0.4162868) * \text{Expresión FAM117A} + 0.8842986 * \text{Expresión KCNQ1OT1} + 0.0859888 * \text{Expresión PCDHB2} + 0.1117924 * \text{Expresión USP43}$$

A continuación, llevamos a cabo un análisis de regresión de Cox univariante de la firma para la SG y la SLP. Tal y como se refleja en la tabla 10, la HR para SG y SLP asociado a la firma génica de alto riesgo fue de 3,23 y 3,46, con un índice de concordancia de 0.71 y 0.72, respectivamente.

Tabla 10. Resultados del análisis de la firma génica común por regresión de Cox univariante para SG y SLP

	Beta	HR (IC95% para HR)	Test Wald	Valor p	IC
Firma (SG)	1,17	3,23 (2,11-4,95)	20,2	6,66e-08	0,71±0,0373
Firma (SLP)	1,24	3,46 (2,3-5,2)	35,4	2,62e-09	0,72±0,032

Beta= coeficiente de regresión; HR= razón de riesgo; IC =intervalo de confianza.

Llevamos a cabo en análisis de capacidad de predicción de la firma común mediante curvas ROC, con un AUC a 12, 24 y 60 meses para la SG de 0.69, 0.70 y 0.81 y para la SLP de 0.67, 0.78 y 0.79, respectivamente. Para las curvas de Kaplan Meier, establecimos un punto de corte en función de la mediana de los valores de la firma

(puntuación de riesgo) cuyo valor resultó ser -1.059, obteniendo dos curvas comparadas mediante la prueba de los rangos logarítmicos. Las curvas ROC y Kaplan Meier se muestran en la figura 23.

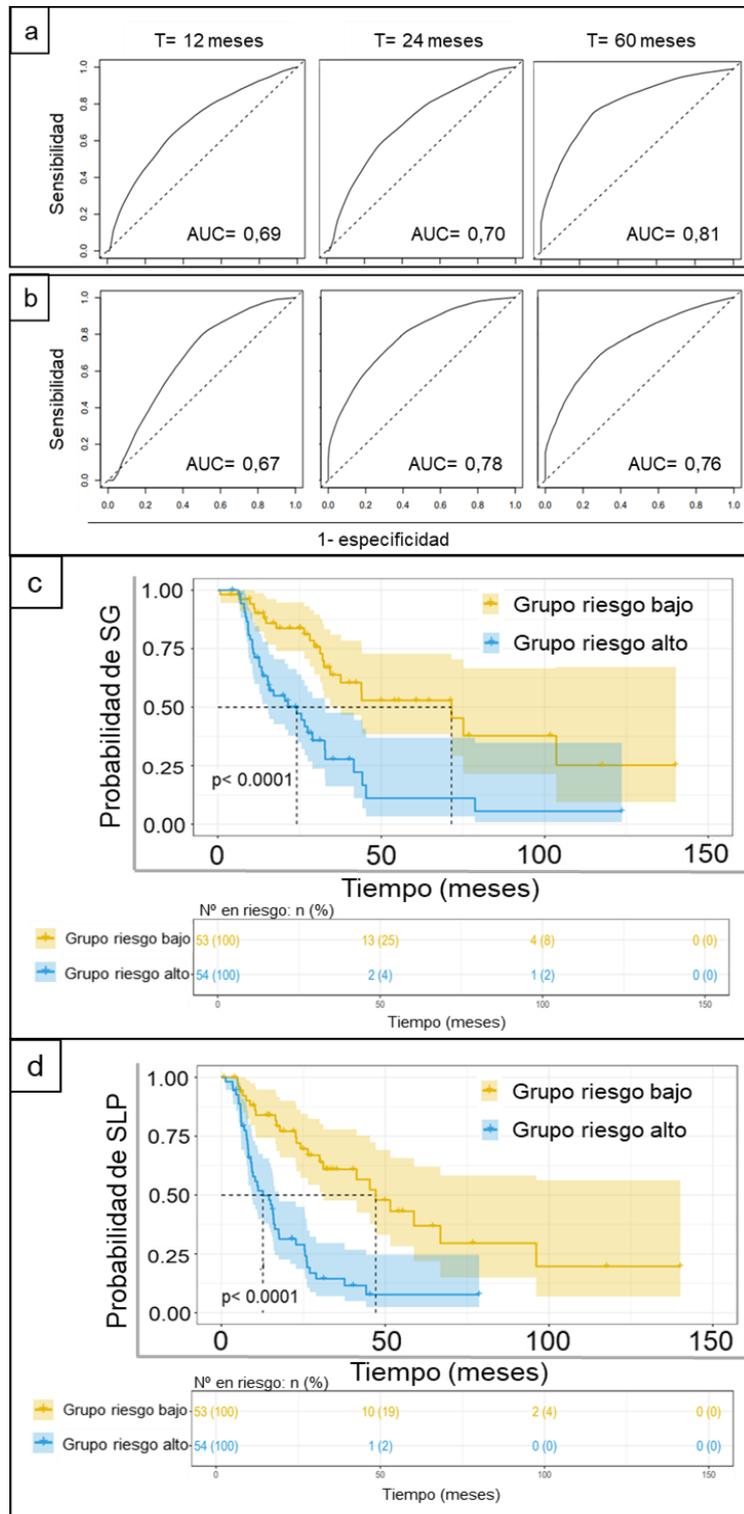


Figura 23. Curvas ROC a 12, 24 y 60 meses de la firma génica común para SG (a) y SLP (b). Curvas de Kaplan Meier en la que se representa la probabilidad de SG (c) y SLP (d) en función del grupo de riesgo de la firma propuesta (bajo o alto riesgo). AUC: área bajo la curva.

Por último, determinamos los niveles de expresión de cada uno de los genes en función del grupo de riesgo considerado en la firma (bajo o alto riesgo). Comparamos dichos niveles mediante la prueba de U de Mann-Whitney y lo representamos mediante diagrama de cajas (figura 24).

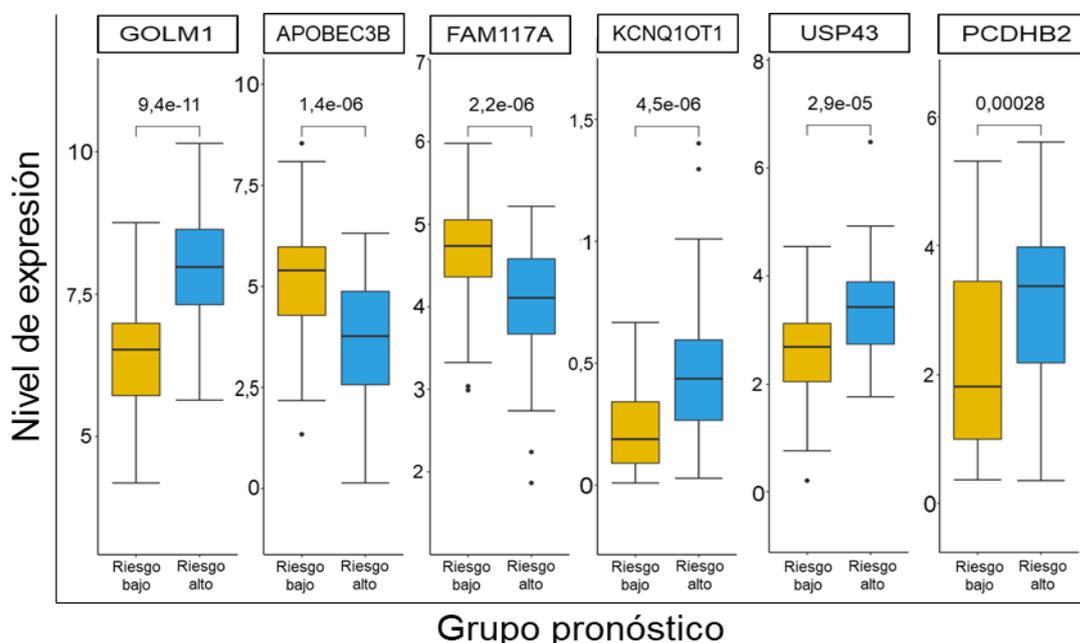


Figura 24. Diagrama de cajas en el que se muestran los niveles de expresión (logTPM) en eje de ordenadas de cada gen en función de cada grupo pronóstico (bajo o alto riesgo).

5.5.1. Aplicación de la firma génica para predicción de SG y SLP en función del subtipo histológico.

Dado que la firma génica propuesta se generó sin distinción por subgrupo histológico principal (adenocarcinoma contra escamoso) planteamos comprobar si la firma era válida para cada subtipo de forma individual. Procedimos metodológicamente igual a lo mostrado en apartados previos para la aplicación de la firma. En la tabla 11 se representan los valores del análisis de regresión de Cox univariante de la firma génica común, distinguiendo entre adenocarcinoma y escamoso. Al comparar los valores de la firma con los obtenidos en el análisis previo sin distinción por subtipo histológico, objetivamos valores de HR similares entre ellos, sin perder poder estadístico.

Tabla 11. Resultados del análisis de la firma génica común por regresión de Cox univariante para SG y SLP en función del tipo histológico principal (adenocarcinoma /escamoso)

Histología		Beta	HR (IC95% para HR)	Test Wald	Valor p	IC
ADC	Firma (SG)	1.26	3.52 (1.85-6.69)	14.8	0.00012	0.69 ±0.055
	Firma (SLP)	1,32	3.76 (2,03-6.97)	17.6	2,68e-05	0,72±0,047
CE	Firma (SG)	1.35	3.87 (1.91-7.83)	14.2	0.00016	0.73±0.05
	Firma (SLP)	1.33	3.78 (1.93-7.42)	15	0.00011	0.73±0.048

Beta= coeficiente de regresión; HR= razón de riesgo; IC= intervalo de confianza; ADC=adenocarcinoma; CE= carcinoma escamoso

Las curvas ROC con los valores de AUC a 12, 24 y 60 meses se muestran en la figura 25. Apreciamos como la capacidad predictiva de la firma se mantiene tanto en adenocarcinoma como en escamoso (valores de AUC iguales o mayores a 0,6 en todos los casos).

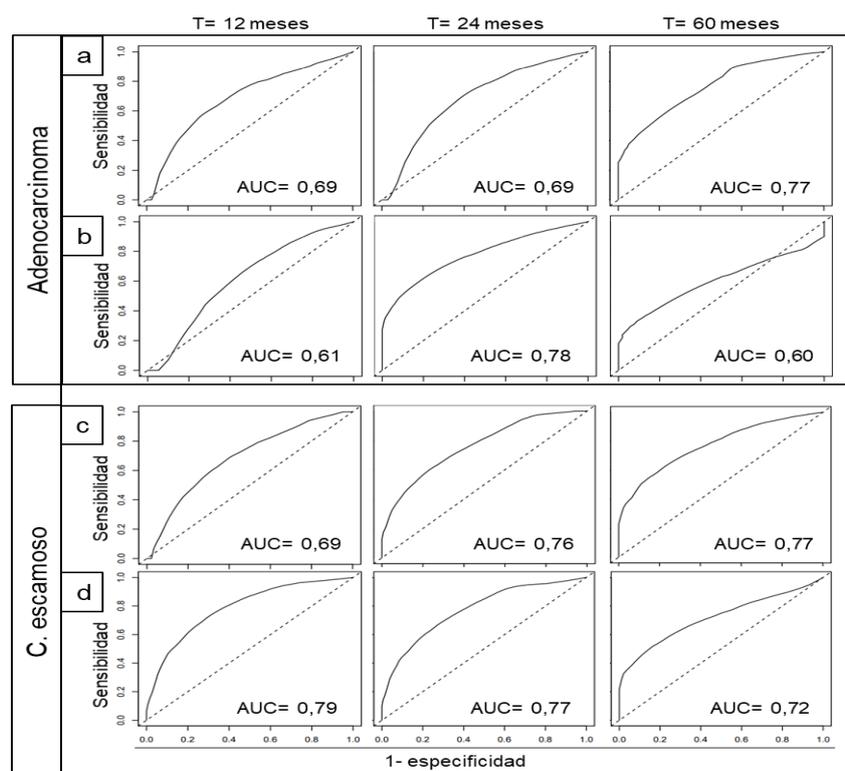


Figura 25. Curvas ROC a 12, 24 y 60 meses de la firma génica común para SG (a y c) y SLP (b y d) en función del grupo histológico (adenocarcinoma y ca. escamoso)

Por último, en la figura 26 se muestran las curvas de Kaplan Meier, con separación estadísticamente significativa (logRank test $p < 0,05$) entre los pacientes asignados al grupo de bajo o alto riesgo, separados de acuerdo con el punto de corte empleado en la cohorte global (-1.05).

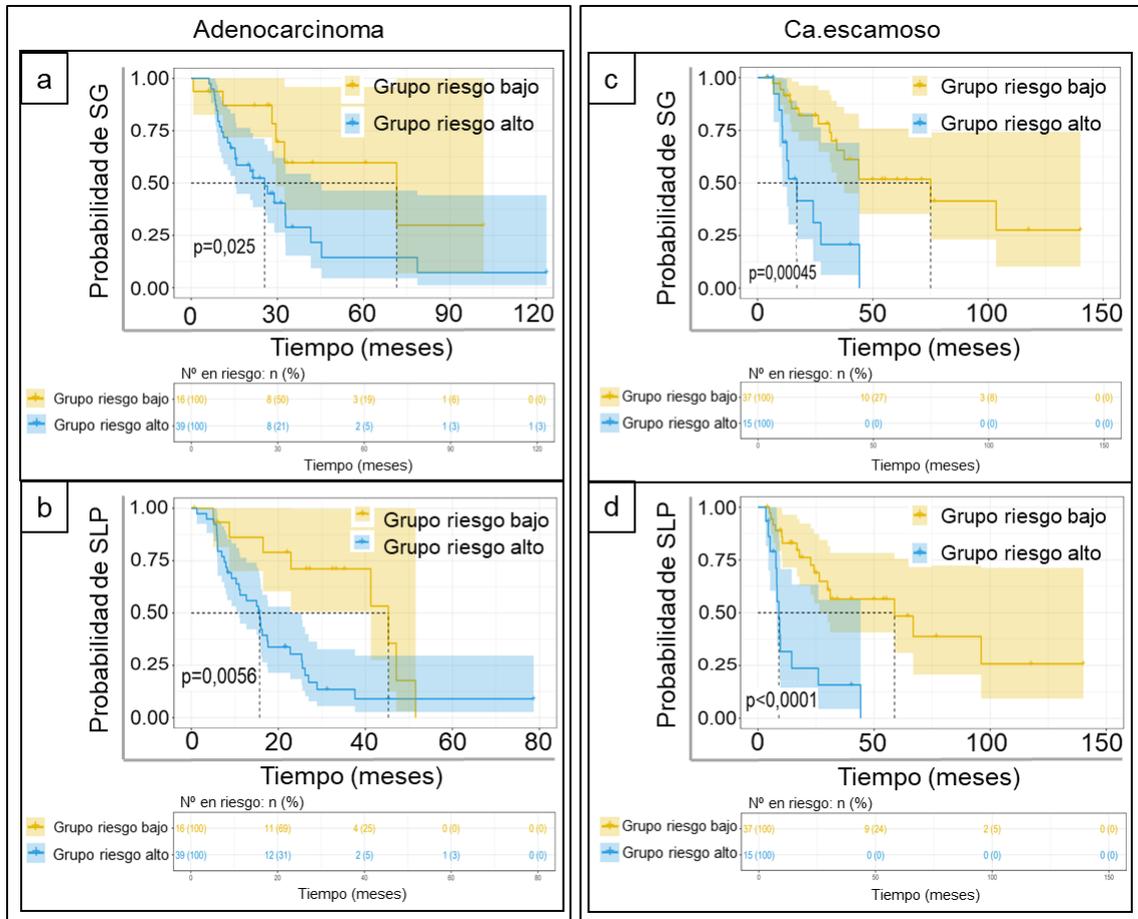


Figura 26. Curvas de Kaplan Meier en las que se representan las probabilidades de SG y SLP en función del grupo de riesgo asociado a la firma génica para el subgrupo de pacientes afectados de adenocarcinoma (imagen a y b) y para el subtipo escamoso (imagen c y d). En cada gráfica se muestra el número de pacientes en riesgo para diferentes tiempos. Se empleó como punto de corte para separar a ambos grupos, el valor obtenido del análisis de la cohorte global.

Decidimos, así mismo, llevar a cabo un análisis adicional por Kaplan Meier, considerando como punto de corte al valor de la mediana de las puntuaciones de riesgo obtenido en el subconjunto de adenocarcinoma y ca. escamoso de forma individual. Así, el punto de corte, al considerar la mediana, separa a las cohortes en 2 grupos con igual número de pacientes. El punto de corte para la cohorte de adenocarcinoma y ca. escamoso fue de -0.81 y -1.48, respectivamente.

Las curvas de KM se muestran en la figura 27. Tal y como se aprecia, las diferencias en estimación de supervivencia para los pacientes de alto y bajo riesgo siguen siendo estadísticamente significativas.

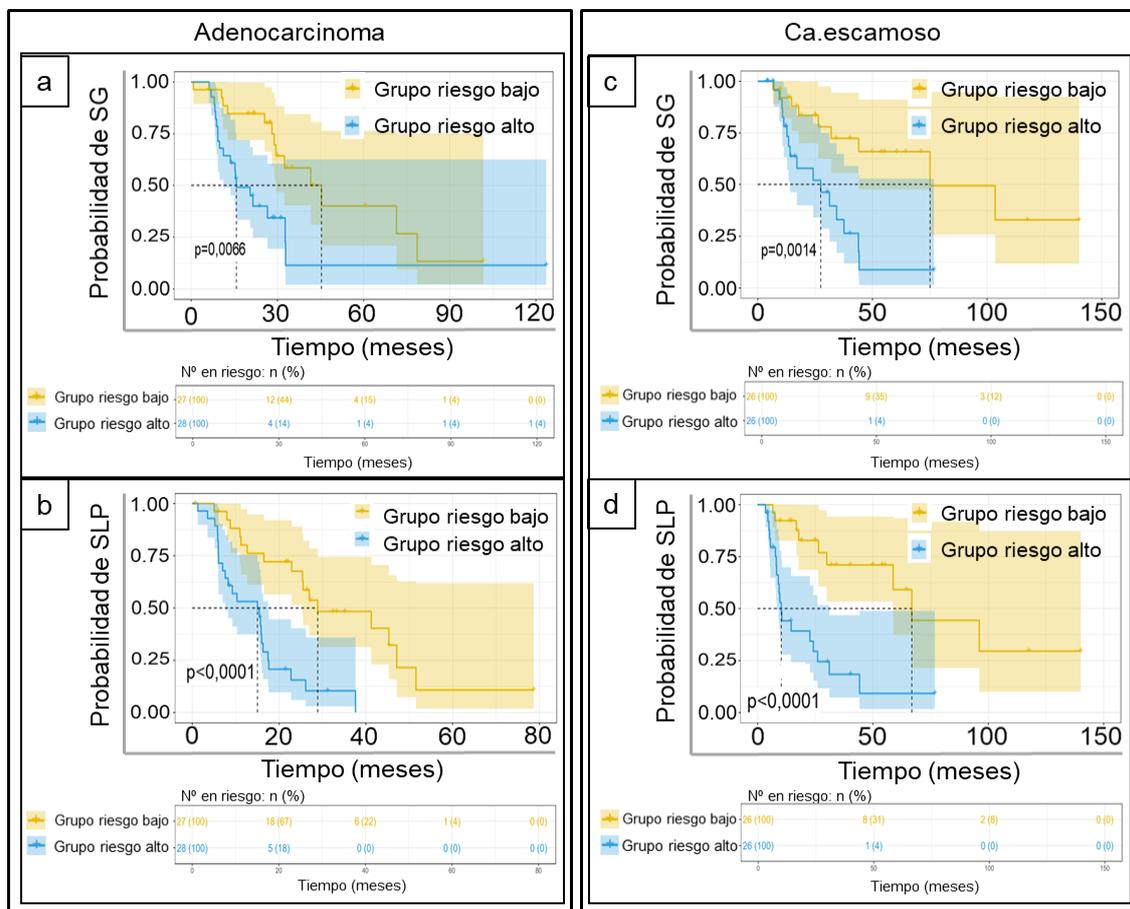


Figura 27. Curvas de Kaplan Meier en las que se representan las probabilidades de SG y SLP en función del grupo de riesgo asociado a la firma génica para el subgrupo de pacientes afectados de adenocarcinoma (imagen a y b) y para el subtipo escamoso (imagen c y d). En cada gráfica se muestra el número de pacientes en riesgo para diferentes tiempos. Se empleó como punto de corte para separar a ambos grupos, la mediana de los valores de la puntuación de riesgo de cada cohorte de forma independiente.

Por último, determinamos los niveles de expresión de cada uno de los genes en función del grupo de riesgo considerado en la firma (bajo o alto riesgo). Comparamos dichos niveles mediante la prueba de U de Mann-Whitney y lo representamos mediante diagrama de cajas (figura 28).

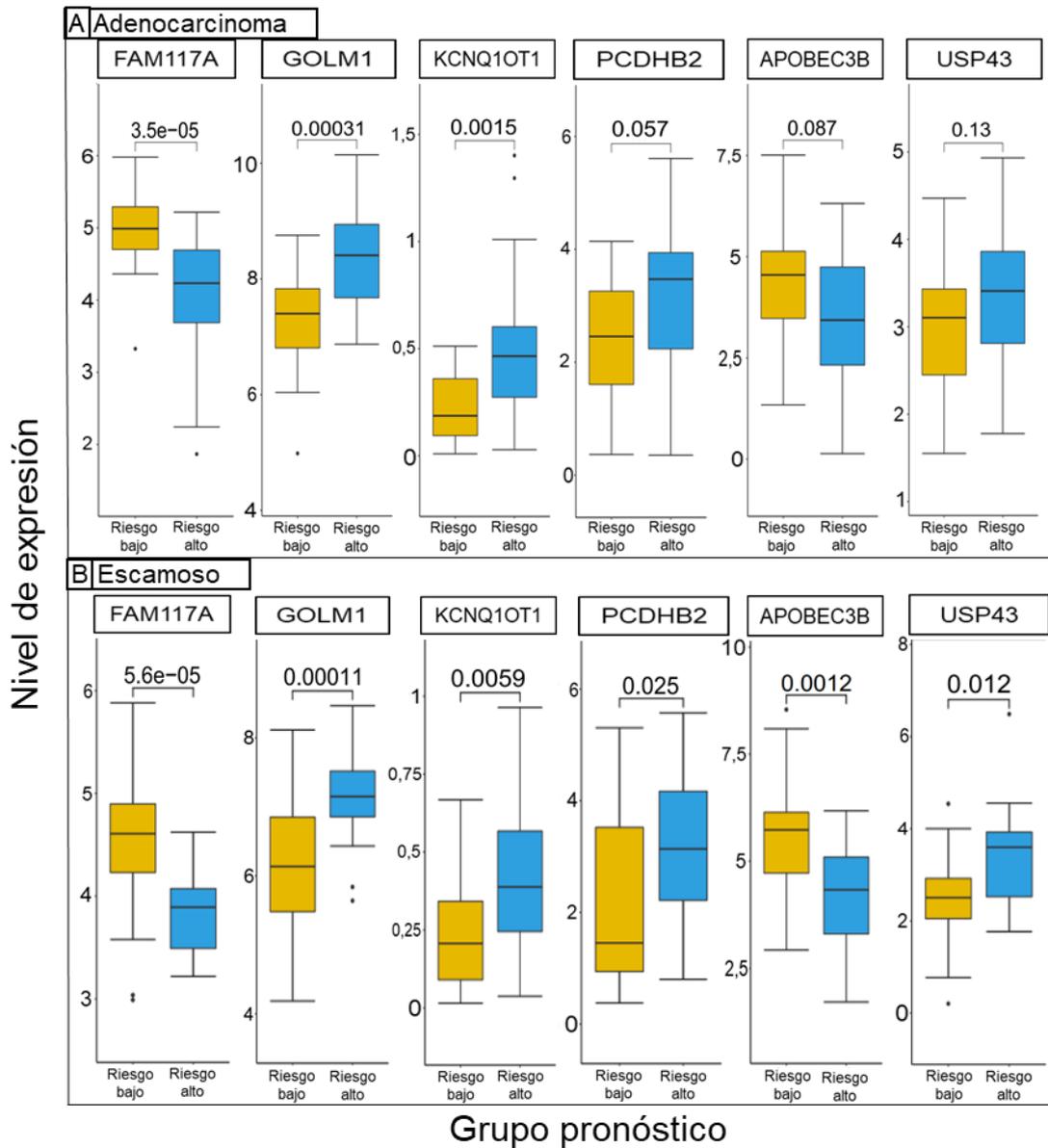


Figura 28. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) en eje de ordenadas de cada gen que conforma la firma génica en función del grupo pronóstico (bajo o alto riesgo). Se muestran los resultados para el análisis de la cohorte de adenocarcinoma (A) y ca. escamoso (B) y se ordenan los elementos génicos de izquierda a derecha en función del valor de p (test de U de Mann-Whitney).

5.6. ENSAYO DE LA FIRMA GÉNICA EN COHORTE ADICIONAL DE PACIENTES AFECTOS DE CPCNP.

A pesar de no disponer de una cohorte adicional en la que los pacientes hubieran recibido tratamiento radioterápico, decidimos ensayar la firma génica propuesta en las cohortes del TCGA de cáncer de pulmón (adenocarcinoma y escamoso) (versión 2018), excluyendo a los pacientes con estadio IV al diagnóstico y los pacientes que habían recibido tratamiento radioterápico, por haber sido esos los considerados en nuestra cohorte para el desarrollo de la firma génica. Se incluyeron un total de 820 pacientes (410 pacientes de cada grupo histológico principal), los cuales habían recibido tratamiento oncológico quirúrgico +/- sistémico. Desde este momento denominaremos al conjunto completo de pacientes de esta cohorte adicional como “cohorte completa”.

5.6.1. Ensayo de la firma génica como factor pronóstico de SG.

Al igual que en lo descrito en los apartados previos, calculamos el valor de la firma génica para cada paciente en función del nivel de expresión de cada uno de los 6 genes propuestos multiplicados por sus respectivos coeficientes de regresión. Llevamos a cabo un análisis por regresión de Cox univariante de la firma para dicha cohorte adicional, obteniendo un valor de HR para la SG de 0.956 (IC al 95% de 0.79-1.15) con valor $p=0.638$ y un IC= 0.503. Las curvas ROC con el cálculo de la predicción de supervivencia a 12, 24 y 60 meses arrojó unos valores de AUC de 0.54, 0.51 y 0.47, respectivamente. Al analizar el impacto de la firma en función del subtipo histológico, apreciamos resultados similares a los obtenidos en la cohorte global, siendo la predicción de SG discretamente mejor para el subgrupo de adenocarcinoma. Las curvas ROC se muestran en la figura 29.

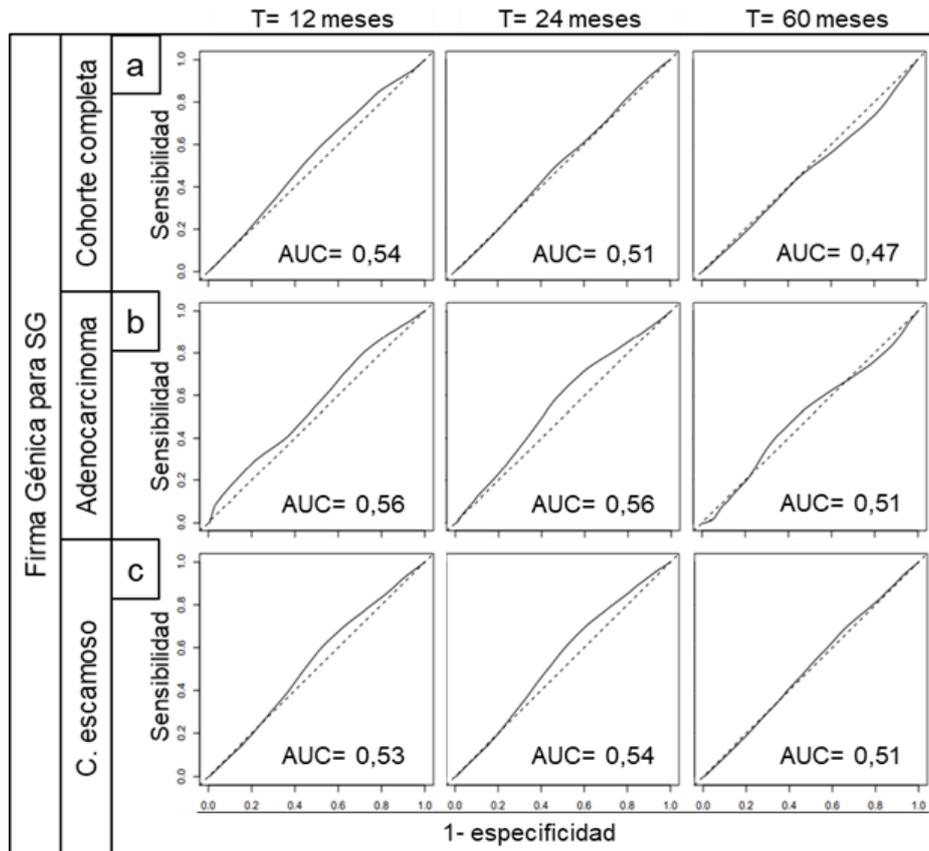


Figura 29. Curvas ROC con el valor del AUC a 12, 24 y 60 meses de la firma génica pronóstica para la cohorte completa de ensayo del TCGA (imagen a), para adenocarcinoma (b) y para c.escamoso (c). AUC= área bajo la curva.

Decidimos llevar a cabo un análisis adicional por Kaplan Meier, considerando como punto de corte al valor de la mediana de las puntuaciones de riesgo obtenido al analizar la cohorte adicional (820 pacientes), y al subconjunto de adenocarcinoma y al de ca. escamoso de forma individual. Así, el punto de corte, al considerar la mediana, separa a las cohortes en 2 grupos con igual número de pacientes. El punto de corte para la cohorte completa, para adenocarcinoma y ca. escamoso, fue de -1.05, -0.81 y -1.48, respectivamente. Las curvas de Kaplan Meier se muestran en la figura 30. Tal y como se aprecia en esta, en el caso del análisis del subgrupo de adenocarcinoma, las diferencias en la estimación de la SG se sitúan en el límite de la significación estadística ($p=0,05$).

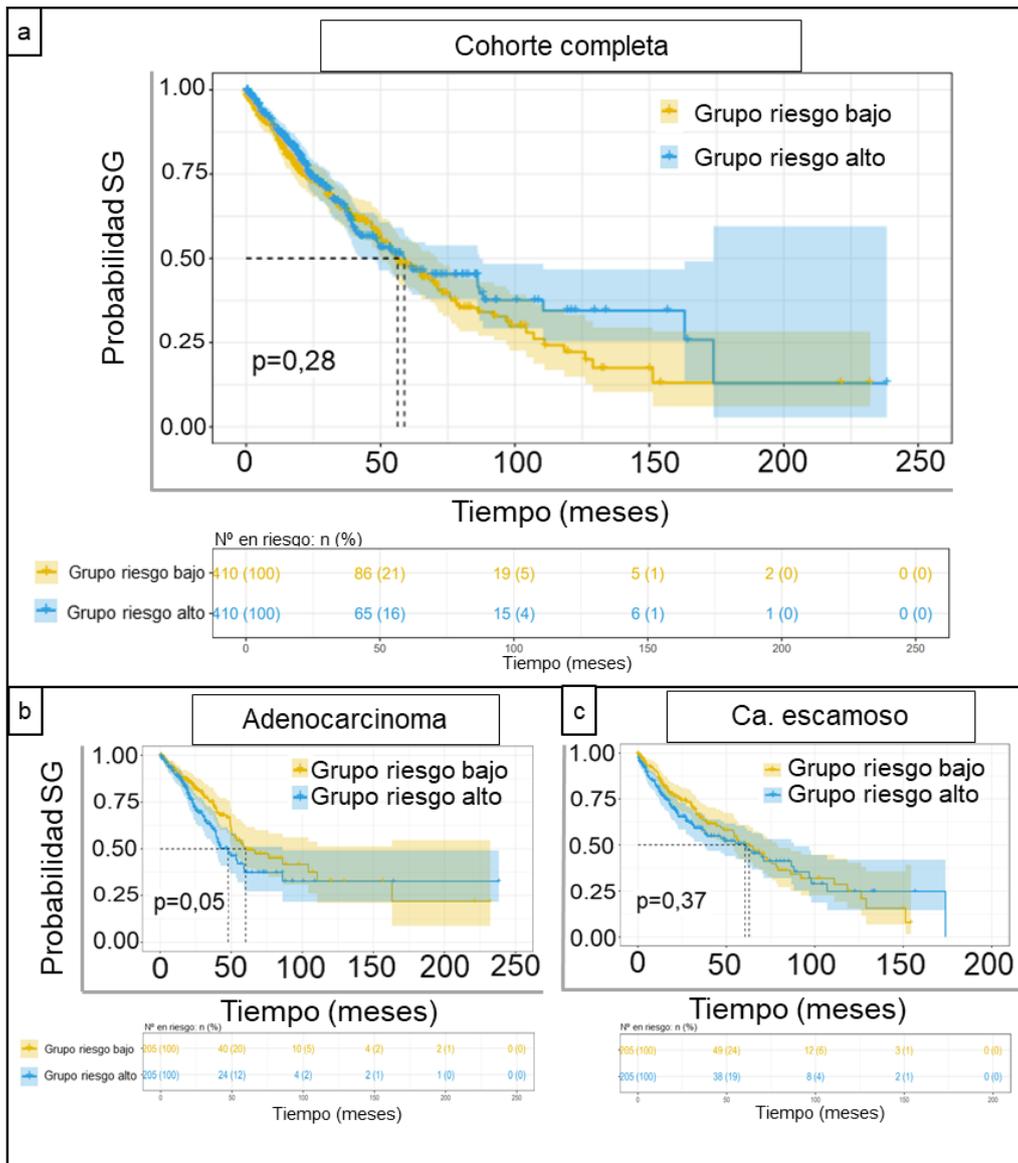


Figura 30. Curvas de Kaplan Meier en las que se representan la probabilidad de SG en función del grupo de riesgo asociado al valor de la firma génica propuesta. Información de la cohorte adicional completa del TCGA (imagen a), del subtipo adenocarcinoma (imagen b) y del ca. escamoso (imagen c). Se empleó como punto de corte para dividir a los pacientes en grupo de alto y bajo riesgo a la mediana del valor de las puntuaciones de riesgo para cada cohorte (-1.15 para cohorte completa, -0.88 para adenocarcinoma y -1.48 para ca. escamoso).

5.6.2. Ensayo de la firma génica como factor predictor de respuesta terapéutica (impacto en SLP).

A continuación, se muestran los resultados del análisis de la firma génica común como elemento predictor de respuesta terapéutica (impacto en SLP) sobre la misma cohorte de pacientes del TCGA empleada en el apartado previo, analizando primero la cohorte completa y posteriormente su impacto en función del subtipo histológico. El valor de la HR para la firma génica predictiva de respuesta fue de 1.48 (IC al 95% de 1.21-

1.81) con un valor $p=0.00015$ y un IC $=0.592 \pm 0.019$. Las curvas ROC a 12, 24 y 60 meses mostraron un AUC de 0.58, 0.56 y 0.52, respectivamente. EL valor de la HR para adenocarcinoma fue de 1.45 (IC al 95% 1.04-2.01) con un valor $p=0.0263$ y un IC de 0.575 ± 0.025 . El valor de AUC a 12, 24 y 60 meses fue de 0.58, 0.58 y 0.53, respectivamente. Por último, en el caso del subtipo escamoso, la HR para la firma resultó ser de 1.28 (IC al 95% de 0.91-1.78) con un valor $p=0.146$ y un IC $=0.578 \pm 0.031$. El valor del AUC a 12, 24 y 60 meses resultó ser de 0.58, 0.54 y 0.51, respectivamente. Las curvas ROC de los 3 supuestos se representan en la figura 31.

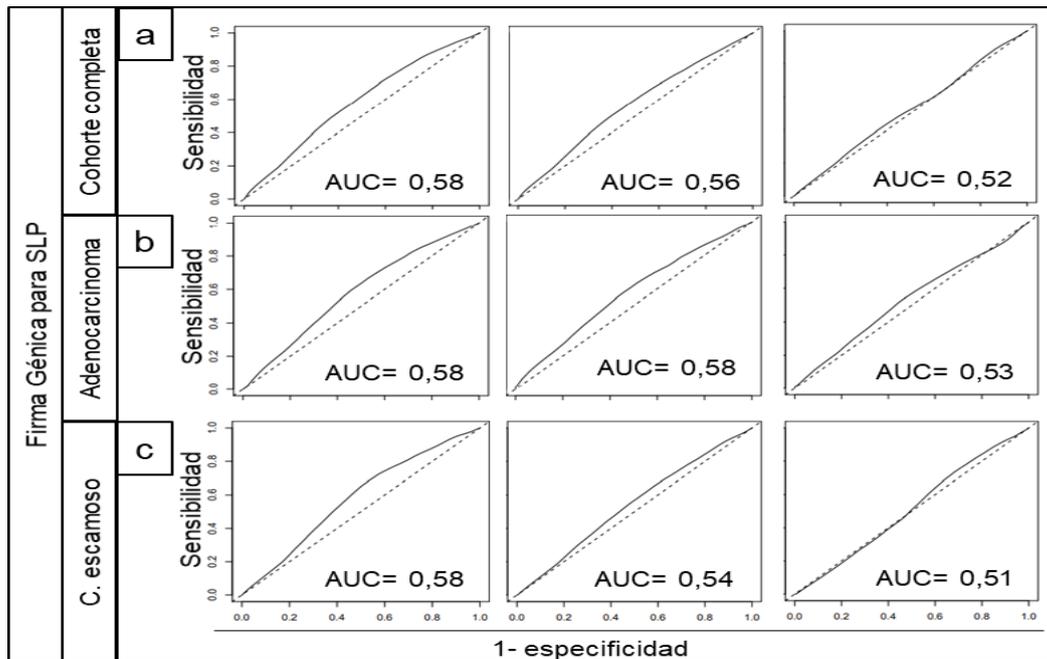


Figura 31. Curvas ROC con el valor del AUC a 12, 24 y 60 meses de la firma génica con intención predictiva de respuesta para la cohorte completa de ensayo del TCGA (imagen a), para adenocarcinoma (b) y para c.escamoso (c). AUC= área bajo la curva.

Al igual que en los previos, llevamos a cabo un análisis adicional por Kaplan Meier, considerando como punto de corte al valor de la mediana de las puntuaciones de riesgo obtenido al analizar la cohorte adicional (820 pacientes; punto de corte=-1.15), el subconjunto de adenocarcinoma (410 pacientes; punto de corte=-0.88) y el de ca. escamoso (410 pacientes; punto de corte=-1.48) de forma individual. Así, el punto de corte, al considerar la mediana, separa a las cohortes en 2 grupos con igual número de pacientes. Las curvas de Kaplan Meier se muestran en la figura 32. Tal y como se

aprecia en la figura, se mantiene y mejora el valor de p para la cohorte y en el caso del análisis del subgrupo de adenocarcinoma, las diferencias en la estimación de la SLP alcanzan la significación estadística ($p=0,012$). Las diferencias en SLP en el caso del ca.escamoso no consiguen alcanzar la significación estadística ($p=0.13$).

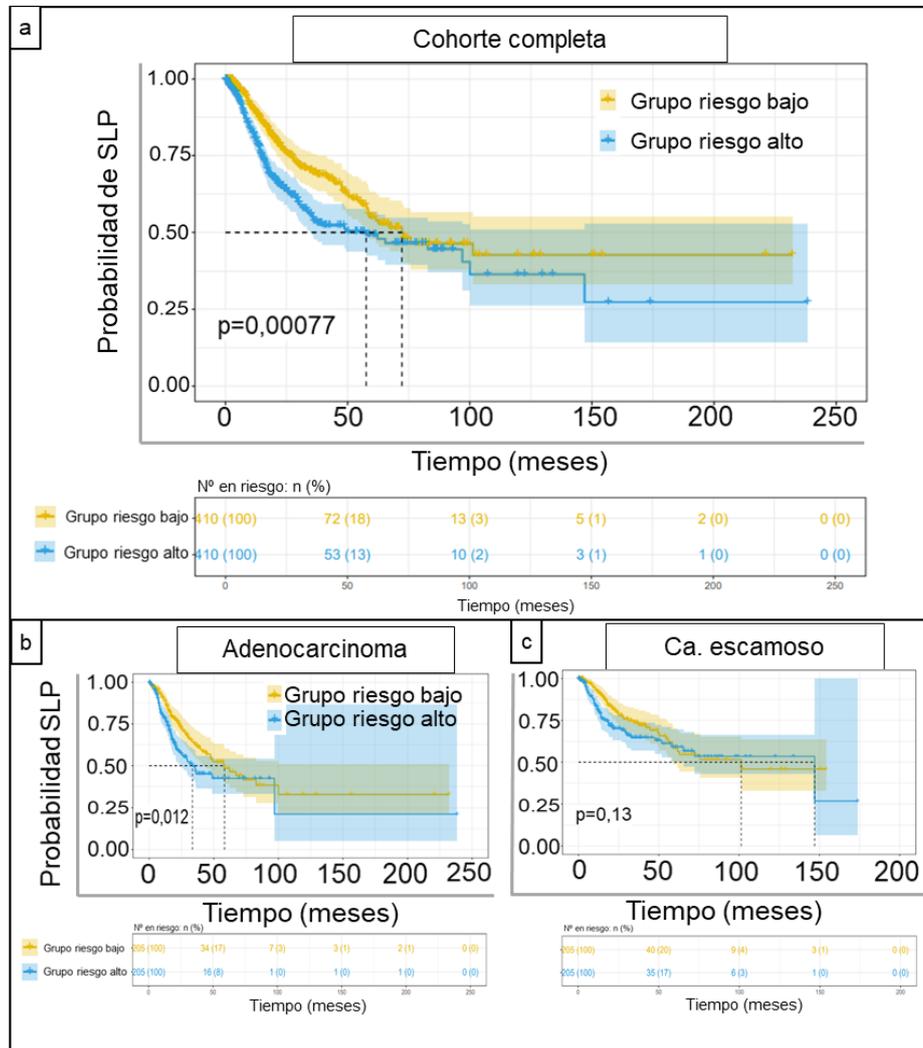


Figura 32. Curvas de Kaplan Meier en las que se representan la probabilidad de SLP en función del grupo de riesgo asociado al valor de la firma génica propuesta. Información de la cohorte adicional completa del TCGA (imagen a), del subtipo adenocarcinoma (imagen b) y del ca. escamoso (imagen c). Se empleó como punto de corte para dividir a los pacientes en grupo de alto y bajo riesgo, la mediana del valor de las puntuaciones de riesgo para cada cohorte (-1.15 para cohorte completa, -0.88 para adenocarcinoma y -1.48 para ca. escamoso).

5.7. COHORTE DE PACIENTES DEL TCGA. CARACTERÍSTICAS Y ANÁLISIS DE LAS VARIABLES CLÍNICAS, PATOLÓGICAS Y MOLECULARES COMO FACTORES PRONÓSTICOS Y PREDICTIVOS DE RESPUESTA A TRATAMIENTO.

En la tabla 12 se muestran los datos clínicos de la cohorte de pacientes seleccionada. En el anexo 4 se encuentran los identificadores del TCGA.

La edad media al diagnóstico era de 63 años (rango 39-86 años). 59 de los pacientes eran varones y 48 mujeres. El diagnóstico inicial de los pacientes y por tanto el tratamiento, se llevó a cabo entre los años 1992 y 2013. Con respecto a la clasificación patológica por estadios, basado en la valoración TNM (de la 3ª a la 7ª edición, dependiendo del año del diagnóstico), 24/107 pacientes se clasificaron como estadio I; 33/107 pacientes como estadio II y 50/107 pacientes como estadio III. Entre este último grupo, 42/107 pacientes se clasificaron como estadio IIIa y 7 pacientes al estadio IIIb.

Con respecto al abordaje terapéutico, todos los pacientes recibieron tratamiento radioterápico, sin poder especificar si se administró de forma adyuvante o en contexto prequirúrgico o como modalidad radical sola o en combinación con tratamiento sistémico.

En el momento de la recolección de los datos de la cohorte (actualización 2018 del TCGA), 57/107 pacientes ya habían fallecido y 68/107 pacientes ya habían sufrido una progresión de su enfermedad. La mediana de SLP, supervivencia cancer-específica (SCE) y SG fue de 22,9 meses (rango 0,72-140), 41,6 meses (rango 0,72- 140) y 32,7 meses (rango 0.72-140), respectivamente (figura 35). La SG, SLP y SCE a 1, 2 y 5 años fue del 80, 66 y 33%, 67,5, 50 y 21% y 82,5, 72,5 y 44%, respectivamente.

Tabla 12. Resumen de las características de los pacientes de la cohorte principal del estudio.

Variables	N	%
Edad al diagnóstico		
< 40 años	1	0,9
[40-49]	8	7,5
[50-59]	28	26,2
[60-69]	37	34,6
≥70	32	29,9
Valor perdido	1	0,9
Año del diagnóstico		
<2000	5	4,7
[2000-2009]	51	47,7
≥ 2010	36	33,6
Valor perdido	15	14
Género		
Mujer	48	44,9
Hombre	59	51,4
Estadio tumoral (AJCC)		
Estadio I	24	22,4
Estadio II	33	30,8
Estadio III	50	46,7
Subtipo histológico		
Adenocarcinoma	55	51,4
Epidermoide	52	48,6
Estado vital		
Vivo	50	46,7
Fallecido	57	53,3

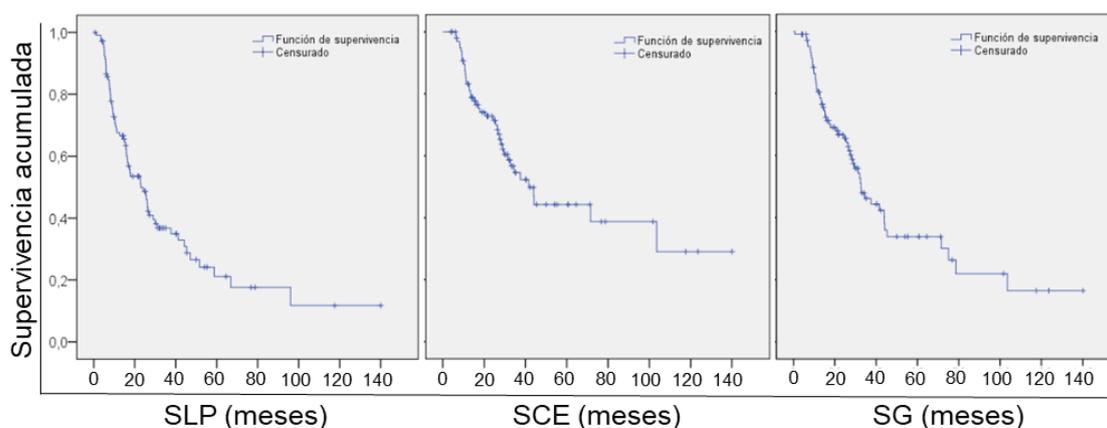


Figura 33. Curvas de Kaplan Meier en las que se representan la supervivencia libre de progresión (SLP), la supervivencia cáncer específica (SCE) y la supervivencia global (SG) de la cohorte de interés.

Para conocer si existen factores que condicionen la SG y la SLP, se consideraron una serie de variables clínicas y patológicas para analizar su influencia en dichos parámetros. Para facilitar el análisis estadístico, las variables se dicotomizaron. Dichas variables son: sexo (varón contra mujer); edad al diagnóstico (menor o mayor-igual a 63 años); año del diagnóstico (antes o después del 2008); el subtipo histológico (adenocarcinoma contra epidermoide); el componente T (T1 y T2 contra T3 y T4); la carga nodal (n0 y n1 contra n2 y n3) y el estadio tumoral (estadio I y II contra estadio III). Así mismo, se dicotomizó la firma génica en pacientes de bajo y alto riesgo en función de si el valor de la firma individual era menor o mayor/igual a la mediana de los valores de la firma establecida.

La información referente a la estimación de SG y SLP en función de las variables descritas más arriba se muestra en la tabla 13. Se empleó un análisis por Kaplan Meier con comparación de las curvas mediante la prueba de los rangos logarítmicos, considerando al valor de $p < 0,05$ como nivel de significación estadística. Tal y como se muestra, tan solo los pacientes con firma génica de alto riesgo contra la de bajo riesgo muestran unas diferencias en SG estadísticamente significativas. Con respecto a la SLP, el subtipo histológico adenocarcinoma y la firma génica de alto riesgo propuesta se relacionaban con peor SLP de forma estadísticamente significativa.

5.7.1. Análisis uni y multivariante de las variables clínicas y patológicas, su impacto en supervivencia global y su influencia sobre la firma génica propuesta.

Con el objetivo de comprobar el potencial impacto de las variables consideradas sobre la SG de los pacientes de la cohorte de interés, se llevó a cabo en primer lugar un análisis por regresión de Cox univariante, seguido de un análisis bivariante con el objetivo de identificar la influencia de dichas variables sobre el impacto de la firma génica en la SG. En el análisis por regresión de Cox univariante, tan solo la firma génica propuesta en nuestro trabajo mostraba un impacto en la SG estadísticamente significativo, con una HR para el grupo de alto riesgo de 2,99 (IC al 95% para la HR de

1,72-5,19) y $p < 0.0001$ (tabla 14). La variable “estadio tumoral” y dentro de ella, el estadio I y II, se relacionó con una disminución del riesgo de muerte (HR= 0,60 con IC al 95% para HR de 0.35-1.01) sin alcanzar el nivel de significación estadística planteado ($p=0.051$).

Tabla 13. Resultados del análisis univariante (test de rangos logarítmicos) para potenciales factores pronósticos y predictivos de respuesta.

Variable	Categorías	n	Mediana SG (meses)	p (Log Rank)	Mediana SLP (meses)	p (Log Rank)
Edad	< 63 años	47	37,6	0,346	25,38	0.457
	≥ 63 años	59	32		22,68	
	desconocido	1	-		-	
Año diagnóstico	< 2008	39	28,9	0,128	22,91	0.831
	≥ 2008	53	44,19		26,14	
	desconocido	15	-		-	
Género	Hombre	59	37,58	0,42	26	0,612
	Mujer	48	32,45		22,75	
Subtipo histológico	Adenocarcinoma	55	29,46	0.158	17,46	0.028
	Escamoso	52	43,89		29,98	
Componente T	T1 y T2	78	32,84	0,188	26	0.1
	T3 y T4	29	24,13		12,66	
Carga nodal	n0-1	65	37,58	0,37	22,92	0.814
	n2-n3	40	27,45		17,62	
	desconocido	2	----		-	
Estadio	I y II	57	43,89	0.051	26,96	0,183
	III	50	26,53		15,88	
Firma pronóstica	Bajo riesgo	53	71,473	<0,001	47,11	<0,001
	Alto riesgo	54	24,131		12,65	
Supervivencia cohorte						
Mediana SG (meses): 32,7 (IC al 95% 25-40,4)						
Mediana SLP (meses): 22,9 (IC al 95% 15-30,8)						
Mediana de SCE (meses): 41,6 (IC al 95% 30,1-53,1)						

SG: supervivencia global; SLP: supervivencia libre de progresión; SCE: supervivencia cáncer específica. IC: intervalo de confianza. N: número de pacientes

Antes de descartar la necesidad de un análisis multivariante, quisimos plantear un análisis bivariante, donde la variable “firma génica pronóstica” se establecía como la principal. En la tabla 15 se muestran los resultados de este análisis. Ninguna variable considerada resultó impactar de forma estadísticamente significativa en la supervivencia ni modificaba el impacto de la firma génica pronóstica en la cohorte seleccionada. Como

ejemplo, en el caso de análisis bivalente con la variable “estadio tumoral”, la HR asociado a la variable “firma pronóstica” (2,81), indica que los pacientes con firma pronóstica de alto riesgo tienen un riesgo de éxitus 2,81 veces mayor que aquellos con firma génica de bajo riesgo, una vez que la posible influencia del estadio tumoral en la SG está controlada. Además, el efecto deletéreo que la firma pronóstica desfavorable ejerce sobre la SG es estadísticamente significativo (valor $p < 0.0001$), cuando el efecto del estadio tumoral en la SG está controlado. Como la HR de la variable “firma génica” del presente modelo (2.81) es muy similar al obtenido en el modelo univariante (2.99), el estadio tumoral no es un factor confusor de la relación entre la firma génica pronóstica y la SG y tampoco es un factor de riesgo de la SG ya que su HR no es estadísticamente significativo (valor $p = 0,24$). Esta interpretación es extensible al resto de variables incluidas y nos permite omitir un posterior análisis multivariante.

Tabla 14. Análisis de regresión de Cox univariante. Se analiza el impacto de las variables categóricas independientes sobre la SG.

Variables	Beta	ET	Wald	HR	95% IC para HR		Valor p
					Inferior	Superior	
Edad al diagnóstico	-0,25	0,27	0,88	0,78	0,46	1,32	0,35
Año diagnóstico	0,46	0,31	2,28	1,59	0,87	2,91	0,13
Género	-0,21	0,27	0,65	0,81	0,48	1,36	0,42
Subtipo histológico	0,38	0,27	1,97	1,46	0,86	2,48	0,16
Componente T	-0,38	0,29	1,71	0,68	0,39	1,21	0,19
Carga Nodal	-0,24	0,27	0,80	0,78	0,46	1,34	0,37
Estadio tumoral (AJCC)	-0,52	0,27	3,73	0,60	0,35	1,01	0,05
Firma génica SG	1,10	0,28	15,16	2,99	1,72	5,19	<0,0001

IC= Intervalo de confianza; HR= razón de riesgo

Tabla 15. Análisis por regresión de Cox bivariante. Valoración del efecto de las variables categóricas independientes sobre la HR de la firma génica pronóstica de SG.

Variables	Beta	ET	Wald	HR	95,0% IC para HR		Valor p
					Inferior	Superior	
Firma génica SG	1,09	0,28	15,04	2,99	1,72	5,20	0,00
Edad al diagnóstico	-0,34	0,27	1,58	0,71	0,42	1,21	0,21
Firma génica SG	1,29	0,31	16,93	3,65	1,97	6,76	0,00
Año diagnóstico	0,52	0,31	2,88	1,69	0,92	3,09	0,09
Firma génica SG	1,11	0,29	14,58	3,04	1,72	5,39	0,00
Género	0,06	0,28	0,05	1,07	0,62	1,83	0,82
Firma génica SG	1,20	0,32	13,69	3,32	1,76	6,26	0,00
Subtipo histológico	-0,20	0,31	0,42	0,82	0,45	1,50	0,52
Firma génica SG	1,11	0,28	15,50	3,03	1,74	5,26	0,00
Componente T	-0,42	0,29	2,10	0,65	0,37	1,16	0,15
Firma génica SG	1,09	0,29	13,99	2,98	1,68	5,28	0,00
Carga Nodal	-0,02	0,28	0,01	0,98	0,57	1,68	0,93
Firma génica SG	1,03	0,29	12,95	2,81	1,60	4,93	0,00
Estadio tumoral (AJCC)	-0,32	0,27	1,36	0,73	0,43	1,24	0,24

IC= Intervalo de confianza; HR= razón de riesgo

5.7.2. Análisis uni y multivariante de las variables clínicas y patológicas, su impacto en supervivencia libre de progresión y su influencia sobre la firma génica propuesta.

Al igual que en el apartado de SG, llevamos a cabo un análisis por regresión de Cox univariante de cada una de las variables clínicas y patológicas consideradas para conocer su impacto en SLP. Tal y como se muestra en la tabla 16, el subtipo histológico adenocarcinoma y la firma génica de alto riesgo resultaron impactar en la SLP con una HR de 1.74 y 3.66 y un valor de p de 0.03 y <0.0001, respectivamente.

En el análisis bivariante por regresión de Cox, tomando como variable principal a la firma génica predictiva y enfrentándola con el resto de las variables, ninguna resultó ser un factor modificador ni confusor del impacto de la firma génica sobre la SLP de forma estadísticamente significativa. En la tabla 17 se representan los resultados de dicho análisis.

Tabla 16. Análisis de regresión de Cox univariante. Se analiza el impacto de las variables categóricas independientes sobre la SLP

Variables	Beta	ET	Wald	HR	95% IC para HR		Valor p
					Inferior	Superior	
Edad al diagnóstico	-0,18	0,25	0,55	0,83	0,51	1,35	0,46
Año diagnóstico	0,06	0,27	0,05	1,06	0,62	1,81	0,83
Género	-0,12	0,24	0,25	0,88	0,55	1,43	0,62
Subtipo histológico	0,55	0,25	4,74	1,74	1,06	2,86	0,03
Componente T	-0,44	0,27	2,63	0,64	0,38	1,10	0,10
Carga Nodal	0,06	0,25	0,06	1,06	0,65	1,74	0,81
Estadio tumoral (AJCC)	-0,32	0,24	1,76	0,72	0,45	1,17	0,18
Firma génica SLP	1,30	0,27	23,45	3,66	2,16	6,18	<0,0001

IC= Intervalo de confianza; HR= razón de riesgo

Tabla 17. Análisis por regresión de Cox bivariante. Valoración del efecto de las variables categóricas independientes sobre la HR de la firma génica pronóstica de SLP

Variables	Beta	ET	Wald	HR	95,0% IC para HR		Valor p
					Inferior	Superior	
Firma génica SLP	1,28	0,27	23,07	3,60	2,14	6,08	0,00
Edad al diagnóstico	-0,27	0,25	1,21	0,76	0,47	1,24	0,27
Firma génica SLP	1,46	0,30	24,14	4,33	2,41	7,76	0,00
Año diagnóstico	0,09	0,28	0,12	1,10	0,64	1,89	0,73
Firma génica SLP	1,31	0,27	23,50	3,72	2,19	6,33	0,00
Género	0,10	0,25	0,16	1,11	0,68	1,81	0,69
Firma génica SLP	1,33	0,30	19,13	3,79	2,09	6,89	0,00
Subtipo histológico	-0,07	0,29	0,06	0,93	0,53	1,64	0,80
Firma génica SLP	1,31	0,27	24,19	3,69	2,19	6,21	0,00
Componente T	-0,51	0,28	3,37	0,60	0,35	1,04	0,07
Firma génica SLP	1,32	0,27	22,90	3,73	2,17	6,39	0,00
Carga Nodal	0,32	0,26	1,50	1,37	0,83	2,28	0,22
Firma génica SLP	1,28	0,27	21,99	3,59	2,10	6,11	0,00
Estadio tumoral (AJCC)	-0,10	0,25	0,15	0,91	0,56	1,48	0,70

IC= Intervalo de confianza; HR= razón de riesgo

DISCUSIÓN

6. DISCUSIÓN

El CPCNP constituye la principal causa de muerte por cáncer en el varón y la segunda en la mujer a nivel mundial²⁴. Un porcentaje elevado de pacientes se diagnostica en estadios localmente avanzados e irresecables y la mayoría (inclusive aquellos diagnosticados en estadios tempranos operables) terminarán sucumbiendo a fenómenos de diseminación metastásica. En dicha situación, el pronóstico a 5 años sigue siendo sombrío.

Hoy en día, el tratamiento estándar en estadios iniciales es cirugía más/ menos tratamiento sistémico y/o radioterápico. En estadios más avanzados, no resecables, pero no metastásicos, el tratamiento estándar sigue siendo regímenes combinados de radio-quimioterapia (basada en platinos). En estadios metastásicos al diagnóstico, el tratamiento consiste en quimioterapia o tratamientos dirigidos sobre dianas moleculares específicas, más/menos la adición de tratamiento radioterápico en algunos casos, todo ello con fines paliativos²⁹.

Actualmente, el sistema de estadificación tumoral por la clasificación TNM sigue siendo el instrumento más potente para predecir la supervivencia de los pacientes y es el eje en el que se centra la comunidad oncológica para plantear el abordaje terapéutico en el caso del CPCNP y en la mayoría de las neoplasias ^{29,123,124}.

A pesar de los esfuerzos por obtener información clínica, patológica y/o molecular que pudiera servir para predecir respuesta a los tratamientos y mejorar la capacidad pronóstica, no existen en el momento actual biomarcadores validados en el CPCNP que faciliten a la comunidad oncológica la toma de decisiones con respecto a la selección individualizada de tratamientos en el contexto no metastásico. Existen múltiples propuestas de firmas génicas que intentan pronosticar la supervivencia o predecir la respuesta a tratamiento (no radioterápico), pero principalmente centrados en estadios tempranos o bien en los metastásicos al diagnóstico. Por el contrario, son escasos los estudios que plantean dichos elementos predictivos y/o pronósticos en

estadios localmente avanzados no metastásicos, donde el tratamiento con radiación ionizante juega un papel importante. El oncólogo radioterápico, en particular, carece de marcadores moleculares que le sirvan para condicionar el tratamiento radioterápico a emplear, más allá de las recomendaciones generales que se derivan, por ejemplo, del informe anatomopatológico, considerando la situación de márgenes quirúrgicos o la carga nodal positiva, entre otros ^{29,123}.

Una de las dificultades a las que nos enfrentamos cuando planteamos la identificación de firmas predictivas y pronósticas en CPCNP, es la incapacidad para identificar si las variables clínicas, terapéuticas, histológicas o moleculares, poseen el mismo peso a la hora de condicionar la respuesta terapéutica mantenida y la supervivencia global, tal y como refleja Subramanian y Simon, en la publicación de 2010, donde hacen un análisis profundo de los sesgos que presentan las diferentes firmas génicas propuestas para esta entidad patológica, llegando a la conclusión de que son esos sesgos y el mal diseño de los estudios pronósticos y predictivos lo que limita la inclusión de los resultados en la práctica clínica diaria¹²⁵.

El objetivo principal de nuestro estudio fue el desarrollo y validación de firmas génicas con capacidad pronóstica y predictiva de respuesta a radioterapia. Para ello realizamos el análisis de expresión basal del genoma completo en muestras tumorales de líneas celulares comerciales de CPCNP, que fueron clasificadas por su sensibilidad al tratamiento radioterápico (líneas radiosensibles contra líneas radorresistentes). Obtuvimos así un total de 127 genes cuyos niveles de expresión diferían entre ambos grupos.

Con el objetivo potencial de identificar ontologías génicas o rutas de señalización que pudieran estar relacionados de forma directa o indirecta con el perfil de respuesta al tratamiento radioterápico, procedimos al enriquecimiento en funciones moleculares, componente celular y *pathways* de los 127 genes diferencialmente expresados. De forma similar a lo publicado con anterioridad, el enriquecimiento en funciones

moleculares y componente celular resultó demasiado genérico y no nos ha permitido definir de forma clara una ontología predominante sobre otras. No obstante, la información obtenida en nuestro estudio sobre un potencial enriquecimiento en actividad oxidorreductasa, tal y como se muestra en el apartado de resultados, nos llevó a testar un fármaco (amitriptilina) con conocido efecto sobre estrés oxidativo¹²⁶, para intentar radiosensibilizar las líneas celulares comerciales incluidas en nuestro trabajo. Los resultados (no mostrados en la presente tesis) no confirmaron la potencial radiosensibilización mediante este modificador de radicales libres.

El enriquecimiento en vías de señalización mediante el software PANTHER o por KEGG, muestra la existencia de enriquecimiento en múltiples pathways de relevancia en cáncer, como la vía de p53, la vía de Wnt, la de PDGF, vías de señalización relacionadas con receptor tipo Toll, vías de señalización de cadherinas, angiogénesis, etc. Estos resultados son observados en otras publicaciones relacionadas con CPCNP, y tal y como se refleja en éstas, el papel de las potenciales interacciones entre estas vías y su desregulación, podrían ayudar a explicar los fenómenos de generación, mantenimiento, progresión y respuesta terapéutica del CPCNP, entre otros tumores¹²⁷⁻
129.

Los genes diferencialmente expresados entre las líneas sensibles y resistentes a radiación fueron considerados para su posterior análisis en la cohorte de interés del estudio (107 pacientes del TCGA (conjunto de adenocarcinoma y ca. escamoso) que cumplían con los requisitos de no ser metastásicos al diagnóstico y haber recibido tratamiento radioterápico). Se llevó a cabo un análisis (regresión de Cox univariante) del impacto de la expresión de dichos genes en la SG de los pacientes que conformaban dicha cohorte, obteniendo un total de 21 genes con impacto pronóstico estadísticamente significativo. A continuación, llevamos a cabo un análisis de predicción de supervivencia mediante curvas de Kaplan Meier, considerando a la mediana del nivel de expresión de cada uno de los 21 genes anteriores como punto de corte para establecer de forma arbitraria una clasificación por altos o bajos niveles de expresión y poder relacionarlo

con la variable SG. De esos 21 genes, solo 10 genes mantenían la significación estadística (log Rank test $p < 0,05$). A continuación, considerando esos 10 genes y mediante una aproximación estadística por regresión de Cox múltiple con regularización de tipo L1 y una validación cruzada de 3 iteraciones (similar a lo empleado en otras publicaciones)¹³⁰, se obtuvo una firma génica pronóstica óptima compuesta por 8 genes: APOBEC3B, GOLM, FAM117A, KCNQ1OT1, SDR16C5, PCDHB2, RHOBTB3 y USP43. La obtención de los coeficientes de regresión de cada elemento multiplicados por el nivel de expresión de cada gen, eran capaz de separar de forma óptima a la cohorte de interés en “pacientes con mal pronóstico” y “pacientes con mejor pronóstico”. La mediana de SG fue de 71,47 y 24,13 meses para el grupo relacionado con la firma favorable y desfavorable, respectivamente. La HR para la firma pronóstica desfavorable fue de 3.9 (IC al 95% 2,39-6,37) ($p < 0,001$). Ninguna variable clínica ni patológica considerada resultó modificadora de dicho impacto en SG.

Por otro lado, y con la misma aproximación, se utilizaron los 127 genes descritos más arriba con intención de obtener una firma predictora de respuesta a los tratamientos. Para ello, hicimos uso de la misma cohorte de pacientes y empleamos la información sobre SLP para obtener dicha firma. Mediante un primer análisis univariante de impacto en SLP, se obtuvieron un total de 29 genes con impacto estadísticamente significativo, de los que solo 19 mantuvieron la significación estadística en Kaplan Meier. Al igual que en el análisis previo, dichos genes se analizaron mediante la prueba de regresión de Cox múltiple con regularización tipo Lasso, con el objetivo de poder obtener los coeficientes de regresión, que, multiplicados por el valor de expresión de cada gen en cada paciente, nos permitió obtener una puntuación de riesgo. El resultado es la obtención de una firma génica predictiva compuesta por 7 genes (APOBEC3B, GOLM1, KCNQ1OT1, PCDHB2, USP43, FAM117A y PAIP2B). La mediana de SLP fue de 47,11 y 12,65 meses para los pacientes relacionados con la firma favorable y desfavorable, respectivamente. En el análisis por regresión de Cox univariante, la firma predictiva desfavorable muestra una HR de 5,04 (IC al 95% de 3-8.47) ($p < 0,001$). En el análisis por

regresión de Cox bivariante, considerando a la firma predictiva como variable principal, ninguno de las variables clínicas o patológicas consideradas, modificaba el valor de la HR asociado a la firma génica predictora propuesta.

Dado que ambas firmas compartían 6 elementos génicos (APOBEC3B, GOLM1, FAM117A, KCNQ1OT1, PCDHB2 y USP43), decidimos llevar a cabo una única firma que fuera válida como elemento pronóstico y como predictor de respuesta terapéutica. Usamos la misma metodología bioinformática que la llevada a cabo en las 2 firmas previas, con un análisis por regresión de Cox con regularización de tipo L1 y validación cruzada de 3 iteraciones. Decidimos emplear los datos de SG de la cohorte para la obtención de los coeficientes de regresión que conformaban la firma génica común.

Los valores de HR para SG y SLP asociado a la firma génica de alto riesgo fue de 3,23 y 3,46, con un índice de concordancia de 0.71 y 0.72, respectivamente ($p < 0.0001$). Apreciamos diferencias con respecto a los valores obtenidos en las firmas génicas diferenciadas expuestas más arriba, pero manteniendo un significativo impacto en dichas variables. Dado que nuestra cohorte se componía de pacientes afectados del subtipo adenocarcinoma y ca. escamoso, decidimos valorar la capacidad predictiva de la firma en función de la histología. Así, en el caso del subgrupo de adenocarcinoma, la HR de la firma génica de alto riesgo sobre la SG y SLP fue de 3,52 y 3.76, respectivamente (valor $p < 0.001$). En el caso del ca. epidermoide, en valor de la HR de la firma sobre la SG y SLP fue de 3,87 y 3,78, respectivamente. Dichos resultados nos indican el valor de la firma génica propuesta con independencia del subtipo histológico principal en los que se dividía nuestra cohorte de estudio, sin poder asegurar que sea válida para otros subtipos histológicos menos frecuentes y no analizados en nuestro estudio.

Por otro lado, a pesar de no disponer de una cohorte adicional para practicar un ensayo de la firma en la que los pacientes hubieran recibido tratamiento radioterápico como parte de su tratamiento oncológico, decidimos emplear la base de datos del TCGA

de cáncer de pulmón (incluyendo a los subtipos adenocarcinoma y escamoso) y descartando a los pacientes en estadio IV al diagnóstico, así como la exclusión de aquellos que sí habían recibido tratamiento radioterápico y que habían sido incluidos en nuestra cohorte de entrenamiento y validación. Con dichos criterios obtuvimos un total de 820 pacientes (50% de adenocarcinoma y 50% de ca. escamoso). Los datos de expresión génica se normalizaron de igual forma a la empleada para la cohorte de entrenamiento y validación del estudio (normalización a logTPM). La firma génica compuesta por 6 genes perdía su capacidad pronóstica en la cohorte adicional (HR para SG de 0,956; $p=0,638$ y un IC=0,53), manteniendo la capacidad de predicción de respuesta sobre la cohorte con una HR=1.48 (IC al 95% de 1.21-1.81), valor $p=0.00015$ y un IC =0.592+/- 0.019. Al realizar el análisis de estimación de supervivencia por Kaplan Meier, se consideró el punto de corte el valor de la mediana de las puntuaciones de riesgo de cada paciente para la cohorte completa adicional, y para la de adenocarcinoma y en el conjunto de ca. escamoso de forma independiente, obteniendo unos valores de -1.15, -0.88 y -1.489, respectivamente.

En nuestro caso, decidimos emplear los datos del TCGA de cáncer de pulmón como cohorte adicional para ensayar la firma obtenida, porque podíamos garantizar, al menos, la normalización de los valores de expresión génica tal y como se había empleado para los datos de la cohorte de desarrollo de la firma propuesta.

Encontramos que nuestra firma no era buena predictora de supervivencia global en pacientes tratados sin radioterapia, a diferencia de cómo sí lo era en radioterapia. Solo en la predicción de supervivencia libre de progresión, especialmente en adenocarcinoma, podíamos observar una estratificación de pacientes debido a nuestra firma molecular.

Con estos resultados, no podemos afirmar que la firma sea predictora específica de la respuesta al tratamiento únicamente radioterápico, pues los resultados de la aplicación de ésta en la cohorte completa adicional empleada para el ensayo de la firma (que no habían recibido tratamiento RT), también predice de forma adecuada la SLP,

aunque con unos valores subóptimos del índice de concordancia y AUC en curvas ROC. Esto podría indicar que la firma puede ser útil como predictora a los tratamientos oncológicos en general en estadios localizados y localmente avanzados manejados con múltiples tratamientos oncológicos entre los que se incluye RT (que es como se ha manejado a la cohorte de 107 pacientes que hemos empleado). Sería realmente difícil la obtención de una cohorte de pacientes exclusivamente tratados con radioterapia de la que dispusiéramos de muestra tumoral para poder llevar a cabo el análisis transcriptómico necesario para la validación adicional de nuestra firma génica. Este es un punto débil de la búsqueda de firmas en patologías donde se emplean tratamientos combinados y/o donde la toma de muestras biológicas, hoy en día, es insuficiente o inexistente, tal y como sucede en algunos casos. La mejora en el diagnóstico mediante biopsia líquida o la radiogenómica, por ejemplo, podría facilitar el desarrollo y/o validación de ésta u otras firmas génicas.

A continuación, mostramos un resumen de la información sobre los genes que se incluyen en la firma génica propuesta.

APOBEC3B (apolipoprotein B mRNA editing enzyme catalytic subunit 3B): Este gen es un miembro de la familia de citidina desaminasas. Se ubica en el brazo largo del cromosoma 22 (22q13.1). Se piensa que estas proteínas pueden ser enzimas de edición de RNA y jugar un papel importante en el crecimiento celular o en el control del ciclo celular, entre otras funciones. Se ha sugerido que la delección de este gen podría relacionarse con un aumento de susceptibilidad para desarrollar neoplasias¹³¹.

GOLM1 (golgi membrane protein 1): codifica para una proteína transmembrana de tipo 2 del aparato de Golgi. El gen se ubica en el brazo largo del cromosoma 9 (9q21.33). Numerosas publicaciones ponen de manifiesto la relación de los niveles de expresión alterados de GOLM1 con proliferación, migración, invasión e inhibición de apoptosis en cáncer de próstata¹³², mama¹³³, páncreas¹³⁴, carcinoma hepatocelular¹³⁵ y

carcinoma escamoso de cavidad oral¹³⁶. Así mismo, y relacionado con CPCNP, Liu y colaboradores publican un artículo en el que relacionan la sobreexpresión de GOLM1 con detrimento en supervivencia global y supervivencia libre de recurrencia en pacientes con el subtipo adenocarcinoma. Se basaron en los datos de la cohorte del TCGA de cáncer de pulmón¹³⁷.

FAM117A (family with sequence similarity 117 member A): es un gen que codifica para la proteína FAM117A. El gen se ubica en el brazo largo del cromosoma 17 (17q21.33). No hemos identificado información relevante sobre su implicación en cáncer ni de forma específica en CPCNP.

KCNQ1OT1 (KCNQ1 opposite strand/antisense transcript 1): el gen es regulado de forma epigenética. Se ubica en el brazo corto del cromosoma 11 (11p15.5). El transcrito interacciona con la cromatina y regula la transcripción de múltiples genes mediante modificaciones epigenéticas. Parece jugar un papel importante en la carcinogénesis colorrectal y se ha sugerido un rol como facilitador de la progresión en CPCNP¹³⁸.

PCDHB2 (protocadherin beta 2): es un gen que codifica para una protocadherina que forma parte del subconjunto de beta protocadherinas que se ubican en el brazo largo del cromosoma 5 (5q31.3). Estas proteínas de tipo cadherina de estirpe neural, forman parte de la membrana plasmática, y aunque su función exacta se desconoce, parecen tener un papel importante en la unión célula-célula en melanocitos y melanoma¹³⁹.

USP43 (ubiquitin specific peptidase 43): codifica para una proteína de la familia de proteasas deubiquitinizantes. EL gen se encuentra en el brazo pequeño del cromosoma 17 (17p13.1). Algún estudio ha relacionado los niveles de expresión del gen

con la regulación del ciclo y el proceso de transición epitelio-mesénquima en cáncer de mama¹⁴⁰. No hemos encontrado ninguna referencia en relación con CPCNP.

FIRMAS GÉNICAS

Beer y colaboradores, hacia el año 2002, publican uno de los artículos pioneros en la generación de firmas génicas con capacidad pronóstica¹⁴¹. En dicho artículo, proponen un perfil génico compuesto por 50 genes, que permite la identificación de un subgrupo de pacientes dentro de los englobados como estadio I cuyo comportamiento y supervivencia se asemejan a los englobados como estadio III. Proponen que la modificación del abordaje terapéutico en dichos pacientes podría condicionar una mejora en su pronóstico. Ninguno de los genes de esta firma pronóstica está representado en la nuestra. Existen diferencias notables en la metodología y materiales empleados en dicho estudio con respecto al nuestro. Tomaron exclusivamente muestras de adenocarcinomas, principalmente en estadio iniciales. La mayoría de los pacientes no recibieron tratamiento adyuvante, a diferencia de la cohorte empleada en nuestro estudio.

Hacia el año 2004, Tomida y colaboradores proponen una firma pronóstica compuesta por 25 genes resultado del análisis de un total de 8644 genes en muestras tumorales de 50 pacientes intervenidos quirúrgicamente de CPCNP¹⁴². 30, 16 y 4 muestras correspondían a los subtipos adenocarcinoma, carcinoma escamoso y carcinoma de celular grandes, respectivamente. Con respecto a la clasificación por estadios, 23, 11 y 16 pacientes, se clasificaron como estadio I, estadio II y estadio III, respectivamente. Llevan a cabo varios subanálisis por histologías, proponiendo una firma de 16 genes para una mejor clasificación pronóstica en subtipo escamoso y otra firma de 12 genes para el subtipo adenocarcinoma. Dichos genes procedían de la firma común de 25 propuesta, sin distinción entre subtipos histológicos. Concluyen que su estudio es parcialmente robusto (limitado por un tamaño muestra escaso principalmente) y que dichas firmas son independientes al estadio TNM al diagnóstico. Con respecto a

nuestro trabajo, no apreciamos ningún gen en común con nuestra firma pronóstica. Al igual que en previos estudios, las diferencias en la metodología, con el análisis de menos de 9000 genes (versus el análisis de más de 20000 genes en nuestro estudio), así como las diferencias en las muestras biológicas empleadas y las potenciales diferencias en el abordaje terapéutico adyuvante, podrían justificar la ausencia de coincidencia en los resultados.

En 2006, Raponi y colaboradores publican un trabajo centrado en la identificación de subgrupos pronósticos en el subtipo escamoso de CPCNP⁹². Emplean un total de 129 muestras de carcinoma escamoso (la mayoría procedente de pacientes en estadio I de la enfermedad) y hacen un análisis mediante microarray con validación por RT-PCR e inmunohistoquímica. Obtienen un grupo de 50 genes con capacidad de separar a las 129 muestras de carcinoma escamoso por subgrupos pronósticos (clasificador pronóstico). Posteriormente, se llevó a cabo una validación en otra cohorte independiente con 36 muestras (mayoritariamente de pacientes en estadio I). Así mismo, empleando de forma conjunta a dicho set de genes, los obtenidos en un trabajo centrado en el subtipo adenocarcinoma y publicado por los mismos autores, testan el set de 100 genes en una cohorte con 52 muestras (50% escamoso y 50% adenocarcinoma) con el objetivo de poder establecer una firma pronóstica válida para los dos subtipos histológicos principales en CPCNP, concluyendo que el set génico clasificador mantenía su capacidad para separar subgrupos pronósticos en la muestra de adenocarcinoma pero no así en el subtipo escamoso. Concluyen, como en la mayoría de los estudios publicados, que es necesario nuevas validaciones en cohortes más extensas y homogéneas. Con respecto a nuestro trabajo, no podemos conocer si parte de los genes que componen ese set génico clasificador comparte alguno de los genes propuestos en nuestra firma pronóstica/predictiva. Lo más probable, dado que el estudio se centraba en muestras de carcinoma escamoso en estadios tempranos, es que no haya elementos comunes.

En 2007, Lau y colaboradores publican una firma pronóstica de 3 genes centrados en CPCNP en estadios tempranos¹⁴³. El objetivo de dicho estudio era la validación de una firma con capacidad para identificar, dentro de estadios iniciales de la enfermedad, aquellos pacientes con perfiles pronósticos diferenciales. Los genes propuestos en la firma fueron STX1A, CCR7 y HIF1A. Al igual que en previos, la firma pronóstica no incluye ninguno de los genes de nuestro estudio, lo cual denota las diferencias notables en cuanto a la aproximación metodológica y la muestra empleada para el estudio, entre otras razones.

Chen y colaboradores publican en el 2007 una firma pronóstica compuesta por 5 genes (DUSP6, MMD, STAT1, ERBB3 y LCK), que fueron el resultado de un análisis combinado de microarrays y RT-PCR en muestras de tumor de 101 pacientes intervenidos quirúrgicamente (mezcla de subtipo adenocarcinoma, escamosos y otros)¹⁴⁴. Posteriormente, los resultados fueron validados en una cohorte independiente de 60 pacientes y en un set de microarrays de 86 pacientes. En el análisis multivariante considerando otras variables clínicas como la edad o el estadio, la firma propuesta mantenía su significación estadística. Por otro lado, dicha firma génica predice de forma estadísticamente significativa la SLR en la cohorte inicial del estudio, no mostrando este resultado en las cohortes de validación. Los motivos que podrían explicar la no coincidencia de ninguno de los genes propuestos con nuestra firma pronóstica son varios: 1) las muestras biológicas de la cohorte principal corresponden casi al 50% para subtipo adenocarcinoma y otro 50% para escamoso, lo que contrasta con nuestro modelo, basado en líneas celulares comerciales de adenocarcinoma mayoritariamente y 2) el tratamiento ofrecido a los pacientes en dicho estudio (solo cirugía, sin adyuvancia) difiere sustancialmente del tratamiento empleado en nuestra cohorte (cirugía, radioquimioterapia, tratamiento sobre dianas moleculares) lo cual puede suponer un impacto en términos de supervivencia y condicionar el valor de la firma pronóstica.

Más recientemente, Zuo y colaboradores proponen una firma de 6 genes con capacidad para predecir la supervivencia libre de enfermedad y la supervivencia global

en CPCNP, sin discriminar por subtipos histológicos¹⁴⁵. Se basaron en la combinación de información génica procedente de 3 bases de datos públicas que engloban todos los subtipos histológicos (mayoritariamente adenocarcinomas). Tras obtener los genes candidatos con capacidad pronóstica, se llevó a cabo una validación con la cohorte de TCGA de cáncer de pulmón. La firma propuesta se compone de los genes PLEKHH2, ISCU, CLUL1, CHRDL1, PAIP2B y CDCP1.

El mismo equipo publica en el mismo año otro artículo en el que proponen una firma pronóstica de 8 genes para pacientes con CPCNP en estadio temprano¹⁴⁶. La aproximación estadística y bioinformática para la obtención de dicha firma se asemeja a la empleada en el artículo referido más arriba y que también es similar a lo realizado en nuestro trabajo. Los genes propuestos (CDCP1, HMMR, TPX2, CIRBP, HLF, KBTBD7, SEC24B-AS1, and SH2B1) no coinciden con los propuestos en nuestro estudio. Esto puede deberse también a múltiples causas: 1) dicha firma se basó en la identificación de genes con valor pronóstico ($HR < 1$ y $p < 0,05$) comunes en 4 bases de datos públicas diferentes (nuestro modelo se basaba inicialmente en un *screening* de muestras tumorales de líneas comerciales y posteriormente el empleo de la cohorte seleccionada de TCGA, exclusivamente) y 2) seleccionaron exclusivamente las muestras que correspondían a estadio temprano de la enfermedad.

A diferencia de los trabajos expuestos más arriba, los artículos publicados sobre la identificación de marcadores génicos predictivos o pronósticos centrados sobre cohortes cuyo tratamiento principal era la radiación ionizante son más escasos y prácticamente inexistentes en el caso específico de CPCNP.

Cabe destacar los trabajos realizados por Torres Roca y colaboradores desde el año 2005 hasta la actualidad, donde tratan de identificar elementos génicos comunes a todas las neoplasias, que pudieran explicar las diferencias de radiosensibilidad objetivadas tanto *in-vitro* como en la práctica clínica. La generación del denominado

“índice de radiosensibilidad” (RSI de sus siglas en inglés) conforma una firma predictiva de respuesta al tratamiento radioterápico compuesta por los genes AR, cJUN, STAT1, PKC, RELA, ABCc, SUMO1, CDK1, HDAC1 e IRF1, que ha sido posteriormente validada en cohortes de pacientes afectados de cáncer de mama, cabeza y cuello, esófago, cáncer de recto y glioblastoma multiforme^{111,147,148}. La ecuación que permite obtener el valor de RSI es:

$$RSI = -0.0098009 * AR + 0.0128283 * cJun + 0.0254552 * STAT1 - 0.0017589 * PKC - \beta - 0.0038171 * RelA + 0.1070213 * cABL - 0.0002509 * SUMO1 - 0.0092431 * PAK2 - 0.0204469 * HDAC1 - 0.0441683 * IRF1$$

Uno de los trabajos más importantes para la comunidad de radio-oncólogos y la radiobiología física y molecular, es el publicado en 2017 por Scott y colaboradores¹¹⁶. En dicho artículo se propone un modelo para adaptar la prescripción radioterápica a la sensibilidad individual del tumor de cada paciente. El modelo, denominado GARD (*genome-based model for adjusting radiotherapy dose*, de sus siglas en inglés), combina la información derivada del índice de radiosensibilidad (RSI) (ya comentado en otros artículos del grupo de trabajo de Torres-Roca) y el modelo lineal cuadrático (modelo LQ), que propone la existencia de dos parámetros que impactan en la capacidad citotóxica de la radiación, siendo uno de ellos proporcional a la dosis de radiación administrada (factor α) y otro que es proporcional al cuadrado de la dosis (factor β). Este modelo matemático ha sido y es empleado, desde hace décadas, para el cálculo de la dosis biológica equivalente de diferentes esquemas de tratamiento radioterápico, teniendo en cuenta el índice α/β de cada tumor, lo cual se ha utilizado para plantear fraccionamientos radioterápicos alterados que han permitido alcanzar resultados biológicamente iguales o superiores a los tratamientos basados en normofraccionamiento diario (1.8/2 Gy /fracción)^{149,150}. Este extenso trabajo emplea múltiples cohortes de diferentes entidades neoplásicas (cáncer de mama, esófago,

cabeza y cuello, estómago, cérvix, gliomas, páncreas, pulmón, cáncer de piel no melanoma y melanoma) y establece un valor numérico a GARD (normalmente en rangos de 1 a 200), relacionándose un mayor nivel de GARD con un mayor efecto terapéutico del tratamiento radioterápico y viceversa. Los autores enfatizan la idea de que el modelo GARD no es útil para la predicción de supervivencia. No obstante, defienden que los diferentes valores de GARD predicen de forma adecuada las diferencias en respuesta terapéutica (a radiación ionizante) que se aprecian en la clínica diaria, entre los distintos subtipos tumorales o distintas localizaciones anatómicas afectas tratadas, lo cual podría relacionarse con un impacto en términos de control locorregional de la enfermedad, así como sobre la supervivencia de los pacientes.

Una de las potenciales críticas a este extenso trabajo es la presunción de que la respuesta tumoral tras radiación ionizante depende de forma predominante de la propia biología tumoral, sin considerar otros factores no tumorales *per se* (del entorno tumoral, comorbilidades, homeostáticos, inmunológicos, etc) o a la usual combinación con otros tratamientos oncológicos, que suele ser la norma en el abordaje multidisciplinar de la mayoría de las neoplasias consideradas en el estudio.

En base a los modelos de RSI y GARD, trabajos más recientes siguen planteando su aplicación en la práctica clínica habitual con el objetivo de predecir la respuesta a tratamiento normofraccionado o estereotáxico ablativo a nivel pulmonar u otras localizaciones y la adaptación del fraccionamiento y DBE en aquellas estirpes histológicas que, de forma habitual, muestran peor respuesta a radiación ionizante^{151,152}.

Existen factores no estrictamente génicos/ moleculares, como el estado general del paciente, el tamaño tumoral, la carga nodal, la edad, la existencia de comorbilidades, los tratamientos previos administrados y la clasificación TNM, entre otros, junto a otros elementos identificados en muestras sanguíneas (marcadores de inflamación como interleuquinas y proteína C reactiva; marcadores indirectos de hipoxia como osteopontina, anhidrasa carbónica IX y lactato deshidrogenasa; o bien marcadores

indirectos de carga tumoral como el antígeno carcinoembrionario o fragmentos de citoqueratina 21-1, que se han empleado de forma clásica para predecir la respuesta al tratamiento y el pronóstico vital, tanto en cáncer de pulmón como en otras neoplasias sólidas y hematológicas¹⁵³⁻¹⁶². Esos factores, generalmente validados en estudios de forma univariante y con importantes sesgos, se consideraron en el trabajo publicado por Dehing-Oberije y colaboradores, proponiendo la combinación de factores clínicos (la clasificación del estado de salud de la WHO (WHO-PS), el volumen espiratorio forzado en el primer segundo (FEV1), el volumen tumoral grosero (GTV) equivalente al tamaño del componente tumoral principal, la carga nodal y el género) con biomarcadores obtenidos de sangre periférica (CEA e IL-6), obteniendo así una mejora en la capacidad pronóstica a 2 años de los pacientes afectos y tratados por un CPCNP¹²².

En nuestro estudio, centrado sobre la cohorte de 107 pacientes del TCGA, ninguna variable contemplada resultó impactar en la SG de forma estadísticamente significativa en el análisis univariante por regresión de Cox. Tampoco en el análisis bivariante, considerando a la firma génica como la variable principal, ninguna variable modificaba el valor de la HR asociado a la firma. En el caso del impacto de las variables contempladas en la SLP, tan solo la variable “subtipo histológico adenocarcinoma” resultó estadísticamente significativa en el análisis univariante. Sin embargo, en el análisis bivariante el subtipo histológico perdía la significación estadística y no modificaba de forma sustancial el valor de la HR asociado a la firma génica.

LIMITACIONES DE NUESTRO ESTUDIO

Tal y como se refleja en la mayoría de los trabajos publicados hasta la fecha, centrados en la obtención de determinados biomarcadores, tomados individualmente o en forma de firmas génicas con potencial capacidad pronóstica y/o predictiva, el principal problema es la dificultad para extrapolar los datos obtenidos *in-vitro* a la clínica diaria. La heterogeneidad de las muestras, las diferentes técnicas de extracción del material genético y el desarrollo constante en la aproximación bioestadística y bioinformática hace muy difícil la comparación entre los estudios pioneros y los más recientes.

En nuestro trabajo hemos planteado una aproximación básica con potencial capacidad traslacional, aunque son varias las limitaciones:

1- El tamaño muestral: empleamos un total de 15 líneas comerciales de CPCNP, de las cuales obtuvimos el material biológico (RNA) para plantear nuestro trabajo, lo que limita la potencia estadística del análisis bioinformático.

2- La valoración de la respuesta a radiación ionizante se llevó a cabo con el ensayo de clonogenicidad, por ser éste el estándar para la determinación de dicho parámetro. Tal y como se refleja en la literatura^{112,163-165}, existen no pocas diferencias en los valores publicados de SF2 de las distintas líneas celulares comerciales (anexo 2), lo cual implica que pueda existir un claro sesgo a la hora de establecer de forma arbitraria una clasificación en radiosensibles y radiorresistentes y esto, obviamente, puede condicionar el análisis bioinformático supervisado, que es la base en la que asienta el resto del estudio.

3- En el caso de CPCNP localmente avanzado, en muchas ocasiones no existe suficiente material de biopsia para poder ampliar la batería de determinaciones mutacionales ni poder realizar microarrays tisulares ni técnicas de validación por inmunohistoquímica. En nuestro estudio, la idea inicial era emplear una cohorte de pacientes en estadio localmente avanzado que hubiesen recibido de forma exclusiva

radio y quimioterapia. Ante la ausencia de muestras en biobanco local que cumplieren con dichos objetivos, optamos por la consideración de bases de datos públicas, como es el caso del TCGA.

4- La mayoría de las bases de datos públicas de información bioinformática en CPCNP se nutren de muestras de pacientes principalmente en estadios localizados y/o metastásicos, que generalmente no han recibido o no reflejan información sobre el tratamiento radioterápico, lo que ha limitado mucho el tamaño muestral empleado en nuestro estudio (n=107), así como la obtención de cohortes adicionales para la validación adicional de nuestras firmas.

5- La generación de firmas génicas de carácter pronóstico y/o predictivo de respuesta a algún tratamiento, no valoran, de forma habitual, otros factores biológicos no relacionados de forma directa con la biología del propio tumor. Dichos factores, cuyas bases genéticas y epigenéticas pueden condicionar la respuesta a determinados tratamientos oncológicos e incluso condicionar de forma importante la supervivencia global de los pacientes, son sin duda alguna, los mayores sesgos a la hora de dar valor traslacional a dichas firmas. En nuestro estudio, con intención de conocer si otras variables clínicas y patológicas podían modificar el impacto pronóstico y/o predictivo de nuestra firma génica, llevamos a cabo un análisis univariante de múltiples variables clínicas y patológicas, y un análisis bivariante por regresión de Cox, sin que ningún factor modificara de forma sustancial el valor de la HR de la firma génica propuesta. Es imposible conocer si otra/s variable/s no incluidas en dicho análisis eran más importantes a la hora de condicionar la supervivencia de los pacientes de la cohorte seleccionada.

La generación de nuevas firmas génicas con capacidad pronóstica y/o predictivas en patologías como el CPCNP puede favorecer de forma importante a los pacientes. En nuestro estudio, hemos propuesto una firma génica pronóstica de SG y con capacidad para predecir la SLP. El conocimiento sobre el perfil respondedor o el

pronóstico vital del paciente, previo al inicio del tratamiento, puede hacernos optimizar el abordaje terapéutico y evitar la temida y frecuente iatrogenia.

No obstante, nuestro estudio aún necesita pasos de validación adicionales antes de su entrada en clínica. Por ejemplo, la validación de un corte de referencia con cual comparar cada paciente, como podría ser su sangre, tejido pulmonar no tumoral, pool de biopsias tumorales, recordando que nuestra firma de alto y bajo riesgo para cada gen está hecha por encima de la mediana de una cohorte de muestras tumorales. También, y de manera muy importante, se necesitaría calibrar la contribución relativa de cada uno de los genes a la firma predictiva y organizar un algoritmo de priorización de dichos genes. Todo esto podría formar parte de trabajos posteriores.

Creemos que nuestro trabajo es un buen ejemplo de cómo es posible hacer radio-oncología traslacional y confiamos que, junto a los resultados ya publicados y el esfuerzo de futuros trabajos, podamos sumar para alcanzar el objetivo final de toda la comunidad oncológica, que es mejorar la calidad de vida y el pronóstico de los pacientes con cáncer.

CONCLUSIONES

7. CONCLUSIONES

1- El análisis de respuesta a radiación ionizante de 15 líneas celulares de CPCNP mediante el valor de supervivencia a 2 Gy permite diferenciar entre subgrupos de líneas sensibles y resistentes al tratamiento.

2- Mediante el análisis de microarray del material genético de las líneas celulares, previamente caracterizadas por su sensibilidad basal a radiación ionizante, hemos obtenido un set de 127 genes diferencialmente expresados.

3- A partir de la información obtenida del análisis bioinformático de las líneas celulares y empleando los datos públicos de microarray de la cohorte seleccionada del TCGA de cáncer de pulmón, hemos obtenido un total de 10 genes como factores pronósticos de SG de forma individual y 19 genes como potenciales elementos predictivos de respuesta (medido por su impacto en SLP).

4- Hemos obtenido una firma génica con capacidad pronóstica y predictiva de respuesta terapéutica. Los genes que forman dicha firma son: APOBEC3B, GOLM1, FAM117A, KCNQ10T, PCDHB2 y USP43. La SG y SLP asociada al grupo de alto y bajo riesgo de acuerdo con el valor de la firma propuesta fue de 24,13 contra 71,47 y 12,65 contra 47,11 meses, respectivamente. Se hizo un análisis multivariante en el que se incluían variables clínicas y patológicas, manteniendo su impacto pronóstico de forma estadísticamente significativa.

5- La firma génica propuesta era válida al ser aplicada de forma individual tanto en el conjunto de pacientes afectos de CPCNP subtipo adenocarcinoma, como en el caso del carcinoma escamoso en pacientes tratados con radioterapia.

6- En el ensayo de la firma génica en una cohorte de 820 pacientes procedentes del TCGA (versión del 2018), en estadios I a III y en los que no se había empleado tratamiento radioterápico como parte del tratamiento oncológico, la firma génica perdía la capacidad pronóstica para la SG pero mantenía la capacidad de predicción de

respuesta, con un valor de HR de 1.48 (IC al 95% de 1.21-1.81) con un valor $p < 0.001$ y un índice de concordancia de 0.592. La mediana de SLP para el grupo de alto y bajo riesgo fue de 49 y 73 meses, respectivamente ($p < 0.001$).

BIBLIOGRAFIA

BIBLIOGRAFIA

1. Schwarz G. über Desensibilisierung gegen Röntgen- und Radiumstrahlen. *Munchener Medizinische Wochenschrift* 1909;24:1–2.
2. Müller C. Eine neue Behandlungsmethode bösartiger Geschwülste. *Munchener Medizinische Wochenschrift* 1910;28:1490–3.
3. Gray LH, Conger AD, Ebert M et al. The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br J Radiol* 1953;26:638–48
4. Watosn JD, Crick FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 1953;171:737-738. doi:10.1038/171737a0
5. Thomlinson RH, Gray LH. The histological structure of some human lung cancers and the possible implications for radiotherapy. *Br J Cancer* 1955;9:539–549.
6. Bergsjø P, Kolstad P. Clinical trial with atmospheric oxygen breathing during radiotherapy of cancer of the cervix. *Scand J Clin Lab Invest* 1968;106 (suppl.);167–71.
7. Grogan M, Thomas GM, Melamed I et al. The importance of hemoglobin levels during radiotherapy for carcinoma of the cervix. *Cancer* 1999;86:1528–36.
8. Nordmark M, Bentzen SM, Rudat V et al. Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother Oncol* 2005;77:18–24.
9. Siemann DW, Hill RP, Bush RS. The importance of the pre-irradiation breathing times of oxygen and carbogen (5% CO₂; 95% O₂) on the in vivo radiation response of a murine sarcoma. *Int J Radiat Oncol Biol Phys* 1977;2:903–11.
10. Eales KL, Hollinshead KER, Tennant DA. Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis*. 2016 Jan; 5(1): e190
11. Sinclair WK. Cyclic x-ray responses in mammalian cells in vitro. *Radiat Res*. 1968 Mar;33(3):620-43.
12. Gupta PB, Chaffer CL, Weinberg RA. Cancer stem cells: mirage or reality? *Nat Med*. 2009;15:1010-2
13. Bartelink H, Roelofsen F, Eschwege F, et al. Concomitant radiotherapy and chemotherapy is superior to radiotherapy alone in the treatment of locally advanced anal cancer: results of a phase III randomized trial of the European Organization for Research and Treatment of Cancer Radiotherapy and Gastrointestinal Cooperative Groups. *J Clin Oncol* 1997;15:2040–9.
14. Guckenberger M, Wulf J, Mueller G, et al. Dose-response relationship for image-guided stereotactic body radiotherapy of pulmonary tumors: relevance of 4D dose calculation. *Int J Radiat Oncol Biol Phys* 2009;74:47–54.
15. Dubben HH, Thames HD, Beck-Bornholdt HP. Tumor volume: a basic and specific response predictor in radiotherapy. *Radiother Oncol*. 1998;47(2):167–74. [PubMed]
16. Yaromina A, Krause M, Thames H, et al. Pre-treatment number of clonogenic cells and their radiosensitivity are major determinants of local tumour control after fractionated irradiation. *Radiother Oncol*. 2007;83(3):304–10. [PubMed]
17. Alexander BM, Othus M, Caglar HB, Allen AM. Tumor volume is a prognostic factor in non-small-cell lung cancer treated with chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2011;79(5):1381–7. [PubMed]
18. Soliman M, Yaromina A, Appold S, et al. GTV differentially impacts locoregional control of non-small cell lung cancer (NSCLC) after different fractionation schedules: Subgroup analysis of the prospective randomized CHARTWEL trial. *Radiother Oncol*. 2013 [PubMed]
19. Werner-Wasik M, Swann RS, Bradley J, et al. Increasing tumor volume is predictive of poor overall and progression-free survival: secondary analysis of the Radiation Therapy Oncology Group 93–11 phase I–II radiation dose-escalation study in patients with inoperable non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2008;70(2):385–90. [PubMed]
20. Zhao L, West BT, Hayman JA, Lyons S, Cease K, Kong FM. High radiation dose may reduce the negative effect of large gross tumor volume in patients with medically

- inoperable early-stage non-small cell lung cancer. *Int J Radiat Oncol Biol Phys.* 2007;68(1):103–10. [PubMed].
21. Dubben H.H., Thames H.D., Beck-Bornholdt H.P. Tumor volume: A basic and specific response predictor in radiotherapy. *Radiother. Oncol.* 1998;47:167–174. doi: 10.1016/S0167-8140(97)00215-6.
 22. Linge A., Dubrovskaja A., Baumann M., Krause M. The role of cancer stem cells in tumour radioresponse. In: Anscher M., Valerie K., editors. *Strategies to Enhance the Therapeutic Ratio of Radiation as a Cancer Treatment.* Springer International Publishing; Cham, Switzerland: 2016. pp. 43–74.21.
 23. Cox JD. Are the results of RTOG 0617 mysterious? *Int J Radiat Oncol Biol Phys.* 2012;82(3):1042–4.
 24. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer.* 2019 Apr 15;144(8):1941-1953. doi: 10.1002/ijc.31937.
 25. Doll R and R. Peto. *Epidemiology of Cancer.* Oxford Textbook of Medicine, ed. D. Warrell, et al. 2003, Oxford: OUP.
 26. IARC, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Tobacco smoke and involuntary smoking. . Vol. 83. 2004, Lyon: IARCPress.
 27. Sasco A et al., Tobacco smoking and cancer: a brief review of the epidemiological evidence. *Lung Cancer,* 2004. 45: p. S3-S9.
 28. Powell HA et al. Chronic obstructive pulmonary disease and risk of lung cancer: the importance of smoking and timing of diagnosis. *J Thorac Oncol.* 2013 Jan; 8(1):6-11.
 29. National Comprehensive Cancer Network (NCCN). *NCCN Clinical Practice Guidelines in Oncology. Non-small cell lung cancer. Version 1.2020*
 30. O'Rourke N, Roqué I Figuls M, Farré Bernadó N, Macbeth F..Concurrent chemoradiotherapy in non-small cell lung cancer. *Cochrane Database Syst Rev.* 2010 Jun 6;(6):CD002140. doi: 10.1002/14651858.CD002140.pub3
 31. Aupérin A, Le Péchoux C, Pignon JP, et al.: Concomitant radio-chemotherapy based on platin compounds in patients with locally advanced non-small cell lung cancer (NSCLC): a meta-analysis of individual data from 1764 patients. *Ann Oncol* 17 (3): 473-83, 2006
 32. Curran WJ, Scott CB, Langer CJ, et al.: Long-term benefit is observed in a phase III comparison of sequential vs concurrent chemo-radiation for patients with unresected stage III nsclc: RTOG 9410. [Abstract] *Proceedings of the American Society of Clinical Oncology* 22: A-2499, 2003
 33. Le Chevalier T, Arriagada R, Quoix E, et al. Radiotherapy alone versus combined chemotherapy and radiotherapy in nonresectable non-small-cell lung cancer: first analysis of a randomized trial in 353 patients. *J Natl Cancer Inst.* 1991;83(6):417–23.
 34. Curran WJ, Jr., Paulus R, Langer CJ, et al. Sequential vs. concurrent chemoradiation for stage III non-small cell lung cancer: randomized phase III trial RTOG 9410. *J Natl Cancer Inst.* 2011;103(19):1452–60.
 35. Pao W et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proceedings of the National Academy of Science USA* 2004; 101(36) 13306-13311.
 36. Lynch TJ et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *New England of Journal Medicine* 2004; 350(21) 2129-2139.
 37. Paez JG et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; 304(5676) 1497-500.
 38. Yasuda H, Kobayashi S, Costa DB. EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications. *Lancet Oncology* 2012; 13(1) e23-31.
 39. Pao W et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Medicine* 2005; 2(1) e17.
 40. Bean J et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proceedings of the National Academy of Science USA* 2007; 104(52) 20932-20937.

41. Engelman JA et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007; 316(5827) 1039- 1043.
42. Wang SE et al. HER2 kinase domain mutation results in constitutive phosphorylation and activation of HER2 and EGFR and resistance to EGFR tyrosine kinase inhibitors. *Cancer Cell* 2006; 10 25-38.
43. Eberhard DA et al. Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *Journal of Clinical Oncology* 2005; 23(25) 5900-5909.
44. Das AK, Sato M, Story MD. Non-small-cell lung cancers with kinase domain mutations in the epidermal growth factor receptor are sensitive to ionizing radiation. *Cancer Research* 2006; 66(19) 9601-9608.
45. Raben D, Helfrich B, Bunn PA Jr. Targeted therapies for non-small-cell lung cancer: biology, rationale, and preclinical results from a radiation oncology perspective. *International Journal of Radiation Oncology Biology Physic* 2004; 59(2 Suppl) 27-38.
46. Bergqvist M et al. p53 status and its in vitro relationship to radiosensitivity and chemosensitivity in lung cancer. *Anticancer Res.* 2003 Mar-Apr; 23(2B):1207-12.
47. Kimberly L Johung et al. A clinical model for identifying radiosensitive tumor genotypes in non small cell lung cancer. *Clin Cancer Res* . Volume: 19, Issue: 19, Date: 2013 Oct 1, Pages: 5523-32
48. Bradford A et al. Assessing the Radiation Response of Lung Cancer with Different Gene Mutations Using Genetically Engineered Mice. *Front Oncol.* 2013; 3: 72. Published online Apr 2, 2013. doi: 10.3389/fonc.2013.00072
49. Hayes DN1, Monti S, Parmigiani G, Gilks CB, Naoki K, Bhattacharjee A, Socinski MA, Perou C, Meyerson M. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol.* 2006 Nov 1;24(31):5079-90
50. Travis WD1, Brambilla E, Riely GJ. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. *J Clin Oncol.* 2013 Mar 10;31(8):992-1001. doi: 10.1200/JCO.2012.46.9270. Epub 2013 Feb 11
51. Choong NW, Salgia R, Vokes EE. Key signaling pathways and targets in lung cancer therapy. *Clinical Lung Cancer* 2007; 8 Suppl 2 S52-S60.
52. Kris MG et al. Identification of driver mutations in tumor specimens from 1000 patients with lung adenocarcinoma: the NCI's lung cancer mutation consortium (LCMC) [abstract CRA7506]. *Journal of Clinical Oncology* 2011; 29(Suppl).
53. Maemondo M et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med* 362:2380- 2388, 2010
54. Shepherd FA et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005;353:123-132
55. Park K, Goto K. A review of the benefit-risk profile of gefitinib in Asian patients with advanced non-small-cell lung cancer. *Curr Med Res Opin* 2006;22:561-573
56. Thatcher N et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 2005;366:1527-1537
57. Mitsudomi T et al: Gefitinib versus cisplatin plus docetaxel in patients with non-smallcell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. *Lancet Oncol* 11:121-128, 2010
58. Mok TS et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947-957, 2009.
59. Rosell R et al: Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* 13:239-246, 2012.

60. Zhou C et al: Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. *Lancet Oncol* 12:735-742, 2011.
61. Tang X et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Research* 2005; 65(17) 7568-7572.
62. Das AK, Chen BP, Story MD. Somatic mutations in the tyrosine kinase domain of epidermal growth factor receptor (EGFR) abrogate EGFR-mediated radioprotection in non-small cell lung carcinoma. *Cancer Research* 2007; 67(11) 5267-5274.
63. Schuurbiens OC1, Kaanders JH, van der Heijden HF, Dekhuijzen RP, Oyen WJ, Bussink J. The PI3-K/AKT-pathway and radiation resistance mechanisms in non-small cell lung cancer. *J Thorac Oncol.* 2009 Jun;4(6):761-7. doi: 10.1097/JTO.
64. Chinnaiyan P, Huang S, Vallabhaneni G. Mechanisms of enhanced radiation response following epidermal growth factor receptor signaling inhibition by erlotinib (Tarceva). *Cancer Research* 2005; 65(8) 3328-3335. Biomarkers in Lung Cancer: Integration with Radiogenomics Data 69
65. Shibuya K, Komaki R, Shintani T. Targeted therapy against VEGFR and EGFR with ZD6474 enhances the therapeutic efficacy of irradiation in an orthotopic model of human non-small-cell lung cancer. *International Journal of Radiation Oncology Biology Physics* 2007; 69(5) 1534-1543.
66. Sanclemente M et al. c-RAF Ablation Induces Regression of Advanced Kras/Trp53 Mutant Lung Adenocarcinomas by a Mechanism Independent of MAPK Signaling. *Cancer Cell.* 2018 Feb 12;33(2):217-228.e4. doi: 10.1016/j.ccell.2017.12.014. Epub 2018 Jan 27.
67. Riely GJ et al. Frequency and distinctive spectrum of KRAS mutations in never smokers with lung adenocarcinoma. *Clinical Cancer Research* 2008; 14(18) 5731-5734.
68. Brose MS et al. BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Research* 2002; 62(23) 6997-7000.
69. Sun Y et al. Inhibition of JAK2 signaling by TG101209 enhances radiotherapy in lung cancer models. *Journal of Thoracic Oncology* 2011; 6(4) 699-706.
70. Wang M, Han J, Marcar L, et al. Radiation Resistance in KRAS-Mutated Lung Cancer Is Enabled by Stem-like Properties Mediated by an Osteopontin-EGFR Pathway. *Cancer Res.* 2017;77(8):2018-2028. doi:10.1158/0008-5472.CAN-16-0808
71. Gurtner K, Kryzmien Z, Koi L, et al. Radioresistance of KRAS/TP53-mutated lung cancer can be overcome by radiation dose escalation or EGFR tyrosine kinase inhibition in vivo. *Int J Cancer.* 2020;147(2):472-477. doi:10.1002/ijc.32598
72. Soda M, Choi YL, Enomoto M, et al. (August 2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature.* 448 (7153): 561–6. doi:10.1038/nature05945. PMID 17625570.
73. Shaw AT, Kim DW, Nakagawa K, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N Engl J Med* 2013; 368: 2385–2394. [PubMed]
74. Solomon BJ, Mok T, Kim DW, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med* 2014; 371: 2167–2177. [PubMed]
75. Katayama R, Shaw AT, Khan TM, et al. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung cancers. *Sci Transl Med* 2012; 4: 120ra17
76. Crinò L, Ahn MJ, De Marinis F, et al. Multicenter phase II study of whole-body and intracranial activity with ceritinib in patients with ALK-rearranged non-small-cell lung cancer previously treated with chemotherapy and crizotinib: results from ASCEND-2. *J Clin Oncol* 2016; 34: 2866–2873. [PubMed]
77. Soria JC, Tan DS, Chiari R, et al. First-line ceritinib versus platinum-based chemotherapy in advanced ALK-rearranged non-small-cell lung cancer (ASCEND-4): a randomised, open-label, phase 3 study. *Lancet* 2017; 389: 917–929. [PubMed]
78. Shaw AT, Kim TM, Crinò L, et al. ASCEND-5: a randomised, phase 3 study of ceritinib versus chemotherapy in adult patients with ALK-rearranged non-small cell lung cancer, previously treated with chemotherapy and crizotinib. *Lancet Oncol* 2017; 18: 874–886.

79. Nanjo S, Yamada T, Nishihara H, et al. Ability of the Met kinase inhibitor crizotinib and new generation EGFR inhibitors to overcome resistance to EGFR inhibitors. *PLoS ONE*. 2013;8(12):e84700.
80. Sun Y, Nowak KA, Zaorsky NG, et al. ALK inhibitor PF02341066 (crizotinib) increases sensitivity to radiation in non-small cell lung cancer expressing EML4-ALK. *Molecular cancer therapeutics*. 2013;12(5):696-704. doi:10.1158/1535-7163.MCT-12-0868.
81. Ying Dai , Quanxiang Wei, Christian Schwager, et al. Oncogene addiction and radiation oncology: effect of radiotherapy with photons and carbon ions in ALK-EML4 translocated NSCLC. *Radiat Oncol*. 2018 Jan 5;13(1):1. doi: 10.1186/s13014-017-0947-0.
82. Gan GN, Weickhardt AJ, Scheier B, Doebele RC, Gaspar LE, Kavanagh BD, Camidge DR. Stereotactic radiation therapy can safely and durably control sites of extra-central nervous system oligoprogressive disease in anaplastic lymphoma kinase-positive lung cancer patients receiving crizotinib. *Int J Radiat Oncol Biol Phys*. 2014 Mar 15;88(4):892-8. doi: 10.1016/j.ijrobp.2013.11.010. Epub 2014 Jan 22.
83. Beer D.G., Kardia S.L., Huang C.C., Giordano T.J., Levin A.M., Misek D.E., Lin L., Chen G., Gharib T.G., Thomas D.G., et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med*. 2002;8:816–824.
84. Director's Challenge Consortium for the Molecular Classification of Lung Adenocarcinoma. Shedden K., Taylor J.M., Enkemann S.A., Tsao M.S., Yeatman T.J., Gerald W.L., Eschrich S., Jurisica I., Giordano T.J., et al. Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. *Nat. Med*. 2008;14:822–827. doi: 10.1038/nm.1790.
85. Larsen J.E., Pavey S.J., Passmore L.H., Bowman R.V., Hayward N.K., Fong K.M. Gene expression signature predicts recurrence in lung adenocarcinoma. *Clin. Cancer Res*. 2007;13:2946–2954. doi: 10.1158/1078-0432.CCR-06-2525.
86. Li Y., Tang H., Sun Z., Bungum A.O., Edell E.S., Lingle W.L., Stoddard S.M., Zhang M., Jen J., Yang P., et al. Network-based approach identified cell cycle genes as predictor of overall survival in lung adenocarcinoma patients. *Lung Cancer*. 2013;80:91–98. doi: 10.1016/j.lungcan.2012.12.022.
87. Tomida S., Takeuchi T., Shimada Y., Arima C., Matsuo K., Mitsudomi T., Yatabe Y., Takahashi T. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. *J. Clin. Oncol*. 2009;27:2793–2799. doi: 10.1200/JCO.2008.19.7053
88. Bianchi F., Nuciforo P., Vecchi M., Bernard L., Tizzoni L., Marchetti A., Buttitta F., Felicioni L., Nicassio F., di Fiore P.P. Survival prediction of stage I lung adenocarcinomas by expression of 10 genes. *J. Clin. Investig*. 2007;117:3436–3444. doi: 10.1172/JCI32007.
89. Raz D.J., Ray M.R., Kim J.Y., He B., Taron M., Skrzypski M., Segal M., Gandara D.R., Rosell R., Jablons D.M. A multigene assay is prognostic of survival in patients with early-stage lung adenocarcinoma. *Clin. Cancer Res*. 2008;14:5565–5570. doi: 10.1158/1078-0432.CCR-08-0544.
90. Park Y.Y., Park E.S., Kim S.B., Kim S.C., Sohn B.H., Chu I.S., Jeong W., Mills G.B., Byers L.A., Lee J.S. Development and validation of a prognostic gene-expression signature for lung adenocarcinoma. *PLoS One*. 2012;7 doi: 10.1371/journal.pone.0044225.
91. Akagi I., Okayama H., Schetter A.J., Robles A.I., Kohno T., Bowman E.D., Kazandjian D., Welsh J.A., Oue N., Saito M., et al. Combination of protein coding and noncoding gene expression as a robust prognostic classifier in stage I lung adenocarcinoma. *Cancer Res*. 2013;73:3821–3832. doi: 10.1158/0008-5472.CAN-13-0031.
92. Raponi M., Zhang Y., Yu J., Chen G., Lee G., Taylor J.M., Macdonald J., Thomas D., Moskaluk C., Wang Y., et al. Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res*. 2006;66:7466–7472. doi: 10.1158/0008-5472.CAN-06-1191.

93. Larsen J.E., Pavey S.J., Passmore L.H., Bowman R., Clarke B.E., Hayward N.K., Fong K.M. Expression profiling defines a recurrence signature in lung squamous cell carcinoma. *Carcinogenesis*. 2007;28:760–766.
94. Skrzypski M., Jassem E., Taron M., Sanchez J.J., Mendez P., Rzyman W., Gulida G., Raz D., Jablons D., Provencio M., et al. Three-gene expression signature predicts survival in early-stage squamous cell carcinoma of the lung. *Clin. Cancer Res.* 2008;14:4794–4799. doi: 10.1158/1078-0432.CCR-08-0576.
95. Landi M.T., Zhao Y., Rotunno M., Koshiol J., Liu H., Bergen A.W., Rubagotti M., Goldstein A.M., Linnoila I., Marincola F.M., et al. MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clin. Cancer Res.* 2010;16:430–441. doi: 10.1158/1078-0432.CCR-09-1736.
96. Hou J., Aerts J., den Hamer B., van Ijcken W., den Bakker M., Riegman P., van der Leest C., van der Spek P., Foekens J.A., Hoogsteden H.C., et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS One*. 2010;5 doi: 10.1371/journal.pone.0010312.
97. Mitra R., Lee J., Jo J., Milani M., McClintick J.N., Edenberg H.J., Kesler K.A., Rieger K.M., Badve S., Cummings O.W., et al. Prediction of postoperative recurrence-free survival in non-small cell lung cancer by using an internationally validated gene expression model. *Clin. Cancer Res.* 2011;17:2934–2946. doi: 10.1158/1078-0432.CCR-10-1803.
98. Roepman P., Jassem J., Smit E.F., Muley T., Niklinski J., van de Velde T., Witteveen A.T., Rzyman W., Floore A., Burgers S., et al. An immune response enriched 72-gene prognostic profile for early-stage non-small-cell lung cancer. *Clin. Cancer Res.* 2009;15:284–290. doi: 10.1158/1078-0432.CCR-08-1258.
99. Xie Y., Xiao G., Coombes K.R., Behrens C., Solis L.M., Raso G., Girard L., Erickson H.S., Roth J., Heymach J.V., et al. Robust gene expression signature from formalin-fixed paraffin-embedded samples predicts prognosis of non-small-cell lung cancer patients. *Clin. Cancer Res.* 2011;17:5705–5714. doi: 10.1158/1078-0432.CCR-11-0196.
100. Lee E.S., Son D.S., Kim S.H., Lee J., Jo J., Han J., Kim H., Lee H.J., Choi H.Y., Jung Y., et al. Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin. Cancer Res.* 2008;14:7397–7404. doi: 10.1158/1078-0432.CCR-07-4937.
101. Boutros P.C., Lau S.K., Pintilie M., Liu N., Shepherd F.A., Der S.D., Tsao M.S., Penn L.Z., Jurisica I. Prognostic gene signatures for non-small-cell lung cancer. *Proc. Natl. Acad. Sci. USA*. 2009;106:2824–2828. doi: 10.1073/pnas.0809444106.
102. Chen D.T., Hsu Y.L., Fulp W.J., Coppola D., Haura E.B., Yeatman T.J., Cress W.D. Prognostic and predictive value of a malignancy-risk gene signature in early-stage non-small cell lung cancer. *J. Natl. Cancer Instit.* 2011;103:1859–1870. doi: 10.1093/jnci/djr420.
103. Guo N.L., Wan Y.W., Tosun K., Lin H., Msiska Z., Flynn D.C., Remick S.C., Vallyathan V., Dowlati A., Shi X., et al. Confirmation of gene expression-based prediction of survival in non-small cell lung cancer. *Clin. Cancer Res.* 2008;14:8213–8220. doi: 10.1158/1078-0432.CCR-08-0095.
104. Lu Y., Lemon W., Liu P.Y., Yi Y., Morrison C., Yang P., Sun Z., Szoke J., Gerald W.L., Watson M., et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med.* 2006;3 doi: 10.1371/journal.pmed.0030467.
105. Chen H.Y., Yu S.L., Chen C.H., Chang G.C., Chen C.Y., Yuan A., Cheng C.L., Wang C.H., Terng H.J., Kao S.F., et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N. Engl. J. Med.* 2007;356:11–20. doi: 10.1056/NEJMoa060096.
106. Lau S.K., Boutros P.C., Pintilie M., Blackhall F.H., Zhu C.Q., Strumpf D., Johnston M.R., Darling G., Keshavjee S., Waddell T.K., et al. Three-gene prognostic classifier for early-stage non small-cell lung cancer. *J. Clin. Oncol.* 2007;25:5562–5569. doi: 10.1200/JCO.2007.12.0352.
107. Zhu C-Q, Ding K, Strumpf D, et al. Prognostic and Predictive Gene Signature for Adjuvant Chemotherapy in Resected Non–Small-Cell Lung Cancer. *Journal of Clinical Oncology*. 2010;28(29):4417-4424. doi:10.1200/JCO.2009.26.4325.

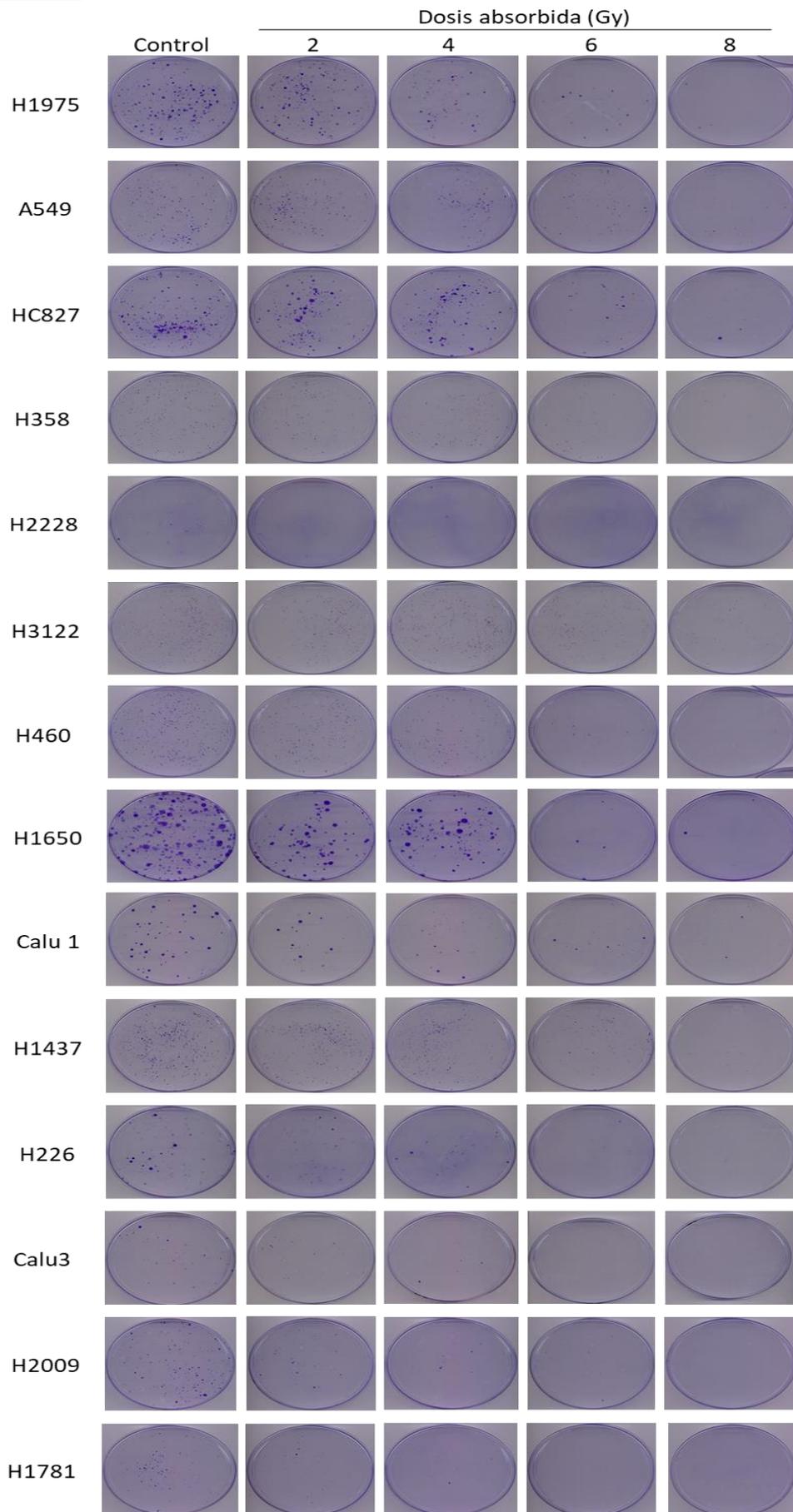
108. Tang H, Xiao G, Behrens C, et al. A 12-Gene Set Predicts Survival Benefits from Adjuvant Chemotherapy in Non-Small-Cell Lung Cancer Patients. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2013;19(6):1577-1586. doi:10.1158/1078-0432.CCR-12-2321.
109. Balázs Györfy, Christos Hatzis, Tara Sanft, Erin Hofstatter, Bilge Aktas, Lajos Pusztai. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res*. 2015; 17(1): 11. doi: 10.1186/s13058-015-0514-2
110. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer*. 2006 Oct;6(10):813-23.
111. Torres-Roca JF, Eschrich S, Zhao H, et al. Prediction of radiation sensitivity using a gene expression classifier. *Cancer Res*. 2005;65:7169-76.
112. Eschrich S, Zhang H, Zhao H, et al. Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. *Int J Radiat Oncol Biol Phys*. 2009;75:497-505.
113. Strom T, Hoffe SE, Fulp W, Frakes J, Coppola D, Springett GM, Malafa MP, Harris CL, Eschrich SA, Torres-Roca JF, Shridhar R. Radiosensitivity index predicts for survival with adjuvant radiation in resectable pancreatic cancer. *Radiother Oncol*. 2015 Oct;117(1):159-64. doi: 10.1016/j.radonc.2015.07.018. Epub 2015 Jul 30.
114. Ahmed KA, Caudell JJ, El-Haddad G, Berglund AE, Welsh EA, Yue B, Hoffe SE, Naghavi AO, Abuodeh YA, Frakes JM, Eschrich SA, Torres-Roca JF. Radiosensitivity Differences Between Liver Metastases Based on Primary Histology Suggest Implications for Clinical Outcomes After Stereotactic Body Radiation Therapy. *Int J Radiat Oncol Biol Phys*. 2016 Aug 1;95(5):1399-1404. doi: 10.1016/j.ijrobp.2016.03.050. Epub 2016 Apr 8.
115. Ahmed KA, Berglund AE, Welsh EA, Naghavi AO, Kim Y, Yu M, Robinson TJ, Eschrich SA, Johnstone PAS, Torres-Roca JF. The radiosensitivity of brain metastases based upon primary histology utilizing a multigene index of tumor radiosensitivity. *Neuro Oncol*. 2017 Aug 1;19(8):1145-1146. doi: 10.1093/neuonc/nox043.
116. Scott JG, Berglund A, Schell MJ, et al. A genome based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. *Lancet Oncol* 2017;18:202-11.
117. Amundson SA, Do KT, Vinikoor LC, Lee RA, Koch-Paiz CA, Ahn J, Reimers M, Chen Y, Scudiero DA, Weinstein JN, Trent JM, Bittner ML, Meltzer PS, Fornace AJ Jr. Integrating global gene expression and radiation survival parameters across the 60 cell lines of the National Cancer Institute Anticancer Drug Screen. *Cancer Res*. 2008 Jan 15;68(2):415-24. doi: 10.1158/0008-5472.CAN-07-2120.
118. Hammond SM. An Overview of MicroRNAs. *Adv Drug Deliv Rev* 2015;87:3-14.
119. Boeri M, Verri C, Conte D, et al. MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. *Proc Nat Acad Sci U S A* 2011;108:3713-8.
120. Hu Z, Chen X, Zhao Y, et al. Serum MicroRNA signatures identified in a genome-wide serum MicroRNA expression profiling predict survival of non-small-cell lung cancer. *J Clin Oncol* 2010;28:1721-6.
121. Sun Y, Schipper MJ, Hawkins PG, et al. Serum MicroRNA signature predicts response to high-dose radiation therapy in locally advanced non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2018;100:107-14.
122. Dehing-Oberije C, Aerts H, Yu S, De Ruyscher D, Menheere P, Hilvo M, van der Weide H, Rao B, Lambin P. Development and validation of a prognostic model using blood biomarker information for prediction of survival of non-small-cell lung cancer patients treated with combined chemotherapy and radiation or radiotherapy alone (NCT00181519, NCT00573040, and NCT00572325). *Int J Radiat Oncol Biol Phys*. 2011 Oct 1; 81(2):360-8.
123. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, Escrú C, Peters S; ESMO Guidelines Committee. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017 Jul 1;28(suppl_4):iv1-iv21. doi: 10.1093/annonc/mdx222.
124. Detterbeck FC, Boff DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest*. 2017;151:193–203. doi: 10.1016/j.chest.2016.10.010.

125. Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? *J Natl Cancer Inst.* 2010 Apr 7;102(7):464-74. doi: 10.1093/jnci/djq025. Epub 2010 Mar 16.
126. Cordero MD, Sánchez-Alcázar JA, Bautista-Ferrufino MR, Carmona-López MI, Illanes M, Ríos MJ, Garrido-Maraver J, Alcudia A, Navas P, de Miguel M. Acute oxidant damage promoted on cancer cells by amitriptyline in comparison with some common chemotherapeutic drugs. *Anticancer Drugs.* 2010 Nov;21(10):932-44. doi: 10.1097/CAD.0b013e32833ed5f7.
127. Chen R, Li WX, Sun Y, Duan Y, Li Q, Zhang AX, Hu JL, Wang YM, Gao YD. Comprehensive Analysis of lncRNA and mRNA Expression Profiles in Lung Cancer. *Clin Lab.* 2017 Feb 1;63(2):313-320. doi: 10.7754/Clin.Lab.2016.160812.
128. Li Z, Sang M, Tian Z, Liu Z, Lv J, Zhang F, Shan B. Identification of key biomarkers and potential molecular mechanisms in lung cancer by bioinformatics analysis. *Oncol Lett.* 2019 Nov;18(5):4429-4440. doi: 10.3892/ol.2019.10796. Epub 2019 Sep 4.
129. Long T, Liu Z, Zhou X, Yu S, Tian H, Bao Y. Identification of differentially expressed genes and enriched pathways in lung cancer using bioinformatics analysis. *Mol Med Rep.* 2019 Mar;19(3):2029-2040. doi: 10.3892/mmr.2019.9878. Epub 2019 Jan 18.
130. Ouyang, G., Yi, B., Pan, G. et al. A robust twelve-gene signature for prognosis prediction of hepatocellular carcinoma. *Cancer Cell Int* 20, 207 (2020).
131. Hashemi M, Moazeni-Roodi A, Taheri M. Association of APOBEC3 deletion with cancer risk: A meta-analysis of 26 225 cases and 37 201 controls. *Asia Pac J Clin Oncol.* 2019 Dec;15(6):275-287. doi: 10.1111/ajco.13107. Epub 2019 Jan 29. PMID: 30693645.
132. Yan G, et al. GOLM1 promotes prostate cancer progression through activating PI3K-AKT-mTOR signaling. *Prostate.* 2018 Feb; 78(3):166-177.
133. Zhang R, et al. Golgi Membrane Protein 1 (GOLM1) Promotes Growth and Metastasis of Breast Cancer Cells via Regulating Matrix Metalloproteinase-13 (MMP13). *Med Sci Monit.* 2019 Jan 29;25:847-855.
134. Song YX, et al. Overexpression of GP73 promotes cell invasion, migration and metastasis by inducing epithelial-mesenchymal transition in pancreatic cancer. *Pancreatology*, 2018 Oct. PMID 30217697
135. Jiao C, et al. Clinical significance and expression of serum Golgi protein 73 in primary hepatocellular carcinoma. *J Cancer Res Ther*, 2018 Oct-Dec. PMID 30488837
136. Li H, et al. Overexpression of Golgi Phosphoprotein 2 Is Associated With Poor Prognosis in Oral Squamous Cell Carcinoma. *Am J Clin Pathol.* 2018 May 31;150(1):74-83. doi: 10.1093/ajcp/aqy029.
137. Liu X, Chen L, Zhang T. Increased GOLM1 Expression Independently Predicts Unfavorable Overall Survival and Recurrence-Free Survival in Lung Adenocarcinoma. *Cancer Control.* 2018;25(1):1073274818778001. doi:10.1177/1073274818778001
138. Dong Z, Yang P, Qiu X et al. KCNQ1OT1 facilitates progression of non-small-cell lung carcinoma via modulating miRNA-27b-3p/HSP90AA1 axis. *J Cell Physiol.* 2019 Jul;234(7):11304-11314. doi: 10.1002/jcp.27788. Epub 2018 Nov 23. PMID: 30471108.
139. Matsuyoshi N, Tanaka T, Toda K, Imamura S. Identification of novel cadherins expressed in human melanoma cells. *J Invest Dermatol.* 1997 Jun;108(6):908-13.
140. Lin F, Xie Z, Chang L, Li W, Wang L, Hou Y, Li L, Zhu J, Xia Y, He W, Li W. USP43 promotes tumorigenesis through regulating cell cycle and EMT in breast cancer. *Int J Clin Exp Pathol.* 2017 Nov 1;10(11):11014-11021. PMID: 31966446; PMCID: PMC6965873.
141. Beer DG, Kardias SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med.* 2002;8(8): 816–824
142. Tomida S, Koshikawa K, Yatabe Y, et al. Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. *Oncogene.* 2004;23(31):5360–5370.
143. Lau SK, Boutros PC, Pintilie M, et al. Three-gene prognostic classifier for early-stage non small-cell lung cancer. *J Clin Oncol.* 2007;25(35): 5562–5569.
144. Chen HY, Yu SL, Chen CH, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med.* 2007;356(1):11–20.
145. Zuo S, Wei M, Zhang H. et al. A robust six-gene prognostic signature for prediction of both disease-free and overall survival in non-small cell lung cancer. *J Transl Med* 17, 152 (2019) doi:10.1186/s12967-019-1899-y.

146. He R, Zuo S. A Robust 8-Gene Prognostic Signature for Early-Stage Non-small Cell Lung Cancer. *Front Oncol*. 2019 Jul 31;9:693. doi: 10.3389/fonc.2019.00693. eCollection 2019.
147. Torres-Roca JF. A molecular assay of tumor radiosensitivity: a roadmap towards biology-based personalized radiation therapy. *Per Med* 2012;9:547-57.
148. Ahmed KA, Chinnaiyan P, Fulp WJ, et al. The radiosensitivity index predicts for overall survival in glioblastoma. *Oncotarget*. 2015; 6: 34414-34422
149. Fowler JF. 21 years of biologically effective dose. *Br J Radiology* 2010; 83: 554–68.
150. Miles EF, Lee WR. Hypofractionation for prostate cancer: a critical review. *Semin Radiat Oncol* 2008; 18: 41–47.
151. Ahmed KA, Scott JG, Arrington JA, et al. Radiosensitivity of lung metastases by primary histology and implications for stereotactic body radiation therapy using the genomically adjusted radiation dose. *J Thorac Oncol* 2018;13:1121-7.
152. Ahmed KA, Grass GD, Orman AG, et al. Personalizing Radiation Treatment Delivery in the Management of Breast Cancer. *Int J Breast Cancer* 2018;2018:6729802.
153. Boldrini L, Donati V, Dell'Omodarme M, et al. Prognostic significance of osteopontin expression in early-stage non-small-cell lung cancer. *Br J Cancer* 2005;93:453–457.
154. Donati V, Boldrini L, Dell'Omodarme M, et al. Osteopontin expression and prognostic significance in non-small cell lung cancer. *Clin Cancer Res* 2005;11:6459–6465.
155. Kaminska J, Kowalska M, Kotowicz B, et al. Pretreatment serum levels of cytokines and cytokine receptors in patients with non-small cell lung cancer, and correlations with clinico-pathological features and prognosis. M-CSF—An independent prognostic factor. *Oncology* 2006;70:115–125.
156. Koukourakis MI, Giatromanolaki A, Sivridis E, et al. Lactate dehydrogenase-5 (LDH-5) overexpression in non-small-cell lung cancer tissues is linked to tumour hypoxia, angiogenic factor production and poor prognosis. *Br J Cancer* 2003;89:877–885.
157. Le QT, Chen E, Salim A, et al. An evaluation of tumor oxygenation and gene expression in patients with early stage non-small cell lung cancers. *Clin Cancer Res* 2006;12:1507–1514.
158. Orditura M, De Vita F, Catalano G, et al. Elevated serum levels of interleukin-8 in advanced non-small cell lung cancer patients: Relationship with prognosis. *J Interferon Cytokine Res* 2002;22:1129–1135.
159. Songur N, Kuru B, Kalkan F, et al. Serum interleukin-6 levels correlate with malnutrition and survival in patients with advanced non-small cell lung cancer. *Tumori* 2004;90:196–200.
160. Muley T, Dienemann H, Ebert W. CYFRA 21-1 and CEA are independent prognostic factors in 153 operated stage I NSCLC patients. *Anticancer Res* 2004;24:1953–1956.
161. Pujol JL, Molinier O, Ebert W, et al. CYFRA 21-1 is a prognostic determinant in non-small-cell lung cancer: Results of a meta-analysis in 2063 patients. *Br J Cancer* 2004;90:2097–2105.
162. Swinson DE, Jones JL, Richardson D, et al. Carbonic anhydrase IX expression, a novel surrogate marker of tumor hypoxia, is associated with a poor prognosis in non-small-cell lung cancer. *J Clin Oncol* 2003;21:473–482.
163. Tumati V, Kumar S, Yu L, Chen B, Choy H, Saha D. Effect of PF-02341066 and radiation on non-small cell lung cancer cells. *Oncol Rep*. 2013 Mar;29(3):1094-100. doi: 10.3892/or.2012.2198. Epub 2012 Dec 18.
164. Jung J, Lee JS, Lee YS, Lee K. Radiosensitivity of Cancer Cells Is Regulated by Translationally Controlled Tumor Protein. *Cancers (Basel)*. 2019 Mar 19;11(3). pii: E386. doi: 10.3390/cancers11030386.
165. Girdhani S, Rodriguez A, Dua R, Parry R. Correlation of radiosensitivity of lung carcinoma to CD137L endogenous expression levels. *Journal of Clinical Oncology*. 37. e14229-e14229. 10.1200/JCO.2019.37.15_suppl.e14229.

ANEXOS

Anexo 1: Ensayo de clonogenicidad



Anexo 2: Valores de SF2 de las líneas celulares en publicaciones previas.

Línea celular	Valor de SF2								Resultado propio
	Carmichael 1989	Eschrich 2009	Tumati 2013	Abazeed 2014	Ahn 2014	Anakura 2019	Jung 2019	Girdhani 2019	
H1975						0,47/0,47			0,89
A549	0,82	0,61	0,73		0,58	0,71/0,53	0,896	0,7	0,83
HCC827						0,35/0,55		0,67	0,74
H358	0,64							0,62	0,67
H2228			0,76						0,65
H3122			0,64						0,62
H460	0,64	0,84	0,54			0,53/0,72	0,518	0,76	0,61
H1650					0,4	0,26/0,23		0,27	0,57
CALU1				0,56					0,45
H520	0,49			0,48		0,49/0,83		0,54	0,49
H1437									0,45
H226	0,61	0,79		0,57					0,43
CALU3									0,32
H2009									0,23
H1781									0,18

Anexo 3. Listado de genes diferencialmente expresados entre líneas celulares radiosensibles y radiorresistentes.

GEN	EXPRESIÓN (log2FC)	GEN	EXPRESIÓN (log2FC)	GEN	EXPRESIÓN (log2FC)
AASS	-1,71	KCNS3	-1,66	TMEM144	1,99
ABCD3	1,15	KIAA1324L	-1,63	TNIK	-1,45
ADAMTS12	-1,30	KLHL24	-1,00	TRIM14	1,10
AHNAK2	2,29	LCA5	-1,02	TRIM6	1,80
ALDH6A1	-1,32	LINC00597	-1,20	TRIM7	1,67
APOBEC3B	2,76	LMO3	-1,25	USP18	1,05
ARG2	-1,66	LOC100287896	-1,27	USP43	1,24
ARHGAP42	1,01	LRIG1	-2,30	VWA5A	1,07
ATP6AP1L	-1,43	LRRC1	-1,01	ZC3H6	-1,06
BASP1	2,25	MAP9	2,07	ZNF114	2,56
BTNL9	1,41	MCM4	1,37	ZNF30	1,35
C4orf32	1,74	MLF1	-1,53	ZNF470	1,30
CD109	3,41	MPDU1	1,05	ZNF480	1,16
CDKN2B	-2,24	NANP	1,00	ABCC3	1,11
CDON	-1,00	NDRG4	1,27	ABCC5	-1,05
CELF2	-1,01	NGEF	1,36	DLL3	1,68
CEP70	-1,05	NTNG1	-1,42	MAN1A1	-2,05
CMTM3	1,43	PAG1	1,36	NUCB2	-1,09
CYBRD1	-1,51	PAIP2B	-1,39	PCDHB13	1,57
CYP27C1	1,33	PERP	1,06	PCDHB14	1,68
DVL3	-1,02	PFKFB4	1,17	PCDHB2	2,09
EDNRA	2,23	PHF19	1,85	PCDHB8	1,14
EGLN3	1,51	PPFIA3	1,06	CST3	1,64
ELOVL6	1,02	PSMB10	1,05	SPOCK1	3,11
EOGT	1,02	PXYLP1	-1,25	TNFAIP8	1,27
EPHX2	1,37	RFC3	1,01	TGFB1	2,24
FAM117A	-1,55	RHOBTB3	1,65	IGFBP5	-3,04
FARP1	-1,17	RMDN2	1,18	VEGFA	-1,05
FNBP1	1,06	RWDD2A	-1,34	FABP5	1,98
FYN	-2,24	SDR16C5	1,84	MCCC1	-1,34
GABRB3	2,06	SELENOP	-2,17	EMB	1,44
GALC	1,93	SEMA3A	2,05	CDH2	-2,43
GIMAP2	1,09	SLAMF7	1,30	MAGI2	-1,46
GOLM1	1,10	SLFN13	1,75	PPARA	-1,01
GPR160	-1,86	SNORA12	1,29	SH3RF2	1,57
GPRC5B	-1,55	SNX5	1,06	SYTL2	1,85
HECW2	1,50	SPEG	-1,32	NR3C2	-1,30
HENMT1	1,36	ST20	-1,11	PAQR8	-1,11
HOXB3	-1,37	ST3GAL3	-1,30	SPTBN1	1,12
HPGD	3,65	SYNPO2	1,07	TUBB3	1,25
JUN	-1,73	TM4SF18	2,41	JPH1	-2,19
KCNQ1OT1	-1,04	TMEM133	1,02	SYNM	1,81
				BMP5	1,27

Anexo 4. Identificadores de los pacientes de la cohorte principal del estudio

Identificadores de pacientes			
TCGA Adenocarcinoma		TCGA Escamoso	
TCGA-05-4382	TCGA-50-6673	TCGA-18-3411	TCGA-63-7023
TCGA-05-4384	TCGA-50-8460	TCGA-18-5592	TCGA-63-A5MN
TCGA-05-4424	TCGA-53-A4EZ	TCGA-21-1081	TCGA-63-A5MP
TCGA-05-5425	TCGA-55-6970	TCGA-21-5787	TCGA-66-2757
TCGA-05-5428	TCGA-55-6981	TCGA-22-4595	TCGA-66-2759
TCGA-44-5643	TCGA-55-7574	TCGA-22-4605	TCGA-66-2778
TCGA-44-6774	TCGA-55-A48Z	TCGA-22-5471	TCGA-66-2783
TCGA-44-6777	TCGA-62-A46Y	TCGA-33-4582	TCGA-66-2794
TCGA-49-4487	TCGA-64-5779	TCGA-33-4586	TCGA-68-7756
TCGA-49-4490	TCGA-69-7974	TCGA-33-AASI	TCGA-68-A59I
TCGA-49-4506	TCGA-71-8520	TCGA-37-A5EL	TCGA-70-6722
TCGA-49-4507	TCGA-73-4676	TCGA-39-5040	TCGA-70-6723
TCGA-49-4510	TCGA-73-A9RS	TCGA-43-2576	TCGA-77-6844
TCGA-49-AAR4	TCGA-78-7166	TCGA-43-A475	TCGA-77-7139
TCGA-49-AARO	TCGA-78-7220	TCGA-46-3768	TCGA-77-8143
TCGA-4B-A93V	TCGA-78-7536	TCGA-46-3769	TCGA-77-8145
TCGA-50-5044	TCGA-78-8660	TCGA-46-6026	TCGA-85-6561
TCGA-50-5045	TCGA-86-7955	TCGA-52-7810	TCGA-85-A4CN
TCGA-50-5049	TCGA-86-8279	TCGA-52-7812	TCGA-90-A59Q
TCGA-50-5051	TCGA-86-A4P8	TCGA-56-7579	TCGA-92-8063
TCGA-50-5072	TCGA-91-6847	TCGA-56-7730	TCGA-LA-A7SW
TCGA-50-5930	TCGA-95-A4VK	TCGA-56-8308	TCGA-NC-A5HE
TCGA-50-5933	TCGA-97-7546	TCGA-56-A4BY	TCGA-NC-A5HG
TCGA-50-5935	TCGA-97-8176	TCGA-56-A62T	TCGA-NC-A5HQ
TCGA-50-5942	TCGA-J2-8194	TCGA-58-8391	TCGA-O2-A52S
TCGA-50-6592	TCGA-L9-A8F4	TCGA-58-A46K	TCGA-O2-A52V
TCGA-50-6593	TCGA-MP-A4T9		
TCGA-50-6595			

ÍNDICE DE FIGURAS

ÍNDICE DE FIGURAS

Figura 1. Clasificación general del cáncer de pulmón y clasificación del subgrupo de células no pequeñas

Figura 2. Mutaciones relevantes en Adenocarcinoma y su incidencia

Figura 3. Diagrama en el que se representan los pasos seguidos durante el proyecto de tesis en cuestión.

Figura 4. 1) Acelerador lineal de electrones. 2) Fantoma empleado para la irradiación de placas de cultivos.

Figura 5. Representación gráfica de las curvas de supervivencia celular tras radiación, medido mediante el ensayo de clonogenicidad. En el eje de ordenadas se representa la fracción superviviente en porcentaje y en el eje de abscisas la dosis de radiación administrada (0- 8 Gy).

Figura 6. Electroferograma de las líneas celulares empleadas. a) líneas celulares tumorales; b) líneas celulares inmortalizadas no tumorales; c) mix celular tejido sano pulmonar; d) gel de electroforesis de todas las muestras.

Figura 7. Representación gráfica mediante diagrama de cajas de los valores de NUSE (a) y RLE (b).

Figura 8. Representación gráfica de la agrupación no supervisada. En la parte superior se representa, mediante color rojo, a las líneas celulares clasificadas como radiosensibles y en negro las radiorresistentes.

Figura 9. Diagrama tipo volcano. Se representa el logFC en el eje de abscisas y el logaritmo en base 2 del valor p en eje de ordenadas. En color rosa se representan los 127 genes de interés. En verde el resto de los genes que no cumplen las condiciones preestablecidas en el análisis.

Figura 10. Mapa de calor representando el resultado de la agrupación supervisada de los genes de interés.

Figura 11. Diagrama circular en el que se representan las ontologías principales de función molecular a las que se relacionan los genes de interés. GO: ontología génica (identificador).

Figura 12. Se representan mediante diagrama circular, las principales ontologías secundarias asociadas a cada función molecular principal. GO: ontología génica (identificador)

Figura 13. Gráfico circular donde se representan las ontologías principales de componente celular a las que se relacionan los genes de interés. GO: identificador de ontología génica.

Figura 14. Representación mediante gráfico circular de los pathways relacionados con los genes de interés del análisis. (P00000): identificador de pathway Panther.

Figura 15. Curvas de Kaplan Meier con la información referente a los 23 genes de interés ordenados por nivel de significación estadística. En el eje de ordenadas se representa la probabilidad de supervivencia (0-1) y en el eje de abscisas, la supervivencia en meses. En amarillo se representan niveles altos de expresión y en azul los niveles bajos.

Figura 16. Curvas ROC en función del tiempo (T). AUC: área bajo la curva.

Figura 17. Curvas de Kaplan-Meier en la que se representa la probabilidad de supervivencia global de los pacientes en función de su agrupación de acuerdo con los niveles de expresión de los genes que conforman la firma génica pronóstica propuesta (bajo y alto riesgo). Se representa, así mismo, el número de pacientes en riesgo en cada grupo en función de los intervalos de tiempo establecidos.

Figura 18. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) en eje de ordenadas, de cada gen que conforma la firma en función del grupo pronóstico (bajo y alto riesgo).

Figura 19. Curvas de Kaplan Meier con la información referente a los 19 genes de interés (valor $p < 0,05$) ordenados por nivel de significación estadística. En el eje de ordenadas se representa la probabilidad de SLP (0-1) y en el eje de abscisas, la SLP en meses. En amarillo se representan los pacientes cuyo nivel de expresión del gen en cuestión era menor a la mediana de expresión

del gen y en azul los pacientes con niveles de expresión iguales o superiores a la mediana de expresión del gen en cuestión.

Figura 20. Curvas ROC de predicción de SLP en función del tiempo (T). AUC: área bajo la curva.

Figura 21. Curvas de Kaplan-Meier en la que se representa la SLP de los pacientes en función de su agrupación de acuerdo con los niveles de expresión de los genes que conforman la firma génica predictiva de respuesta propuesta (bajo y alto riesgo). Se representa, así mismo, el número de pacientes en riesgo en cada grupo en función del tiempo.

Figura 22. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) en eje de ordenadas de cada gen en función de cada grupo pronóstico (bajo o alto riesgo).

Figura 23. Curvas ROC a 12, 24 y 60 meses de la firma génica común para SG (a) y SLP (b). Curvas de Kaplan Meier en la que se representa la probabilidad de SG (c) y SLP (d) en función del grupo de riesgo de la firma propuesta (bajo o alto riesgo). AUC: área bajo la curva.

Figura 24. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) en eje de ordenadas de cada gen en función de cada grupo pronóstico (bajo o alto riesgo).

Figura 25. Curvas ROC a 12, 24 y 60 meses de la firma génica común para SG (a y c) y SLP (b y d) en función del grupo histológico (adenocarcinoma y ca. escamoso).

Figura 26. Curvas de Kaplan Meier en las que se representan las probabilidades de SG y SLP en función del grupo de riesgo asociado a la firma génica para el subgrupo de pacientes afectados de adenocarcinoma (imagen a y b) y para el subtipo escamoso (imagen c y d). En cada gráfica se muestra el número de pacientes en riesgo para diferentes tiempos. Se empleó como punto de corte para separar a ambos grupos, el valor obtenido del análisis de la cohorte global.

Figura 27. Curvas de Kaplan Meier en las que se representan las probabilidades de SG y SLP en función del grupo de riesgo asociado a la firma génica para el subgrupo de pacientes afectados de adenocarcinoma (imagen a y b) y para el subtipo escamoso (imagen c y d). En cada gráfica se muestra el número de pacientes en riesgo para diferentes tiempos. Se empleó como punto de

corte para separar a ambos grupos, la mediana de los valores de la puntuación de riesgo de cada cohorte de forma independiente.

Figura 28. Diagrama de cajas en las que se muestran los niveles de expresión (logTPM) en eje de ordenadas de cada gen que conforma la firma génica en función del grupo pronóstico (bajo o alto riesgo). Se muestran los resultados para el análisis de la cohorte de adenocarcinoma (A) y ca. escamoso (B) y se ordenan los elementos génicos de izquierda a derecha en función del valor de p (test de U de Mann-Whitney).

Figura 29. Curvas ROC con el valor del AUC a 12, 24 y 60 meses de la firma génica pronóstica para la cohorte completa de ensayo del TCGA (imagen a), para adenocarcinoma (b) y para c.escamoso (c). AUC= área bajo la curva.

Figura 30. Curvas de Kaplan Meier en las que se representan la probabilidad de SG en función del grupo de riesgo asociado al valor de la firma génica propuesta. Información de la cohorte completa del TCGA (imagen a), del subtipo adenocarcinoma (imagen b) y del ca. escamoso (imagen c). Se empleó como punto de corte para dividir a los pacientes en grupo de alto y bajo riesgo a la mediana del valor de las puntuaciones de riesgo para cada cohorte (-1.15 para cohorte completa, -0.88 para adenocarcinoma y -1.48 para ca. escamoso).

Figura 31. Curvas ROC con el valor del AUC a 12, 24 y 60 meses de la firma génica con intención predictiva de respuesta para la cohorte completa de ensayo del TCGA (imagen a), para adenocarcinoma (b) y para c.escamoso (c). AUC= área bajo la curva.

Figura 32. Curvas de Kaplan Meier en las que se representan la probabilidad de SLP en función del grupo de riesgo asociado al valor de la firma génica propuesta. Información de la cohorte completa del TCGA (imagen a), del subtipo adenocarcinoma (imagen b) y del ca. escamoso (imagen c). Se empleó como punto de corte para dividir a los pacientes en grupo de alto y bajo riesgo, la mediana del valor de las puntuaciones de riesgo para cada cohorte (-1.15 para cohorte completa, -0.88 para adenocarcinoma y -1.48 para ca. escamoso).

Figura 33. Curvas de Kaplan Meier en las que se representan la supervivencia libre de progresión (SLP), la supervivencia cáncer específica (SCE) y la supervivencia global (SG) de la cohorte de interés.

ÍNDICE DE TABLAS

ÍNDICE DE TABLAS

Tabla 1. Algunos biomarcadores relacionados con los fenómenos radiobiológicos.

Tabla 2. Resumen de las líneas celulares tumorales comerciales empleadas en el estudio.

Tabla 3. Valores de la fracción superviviente a 2 Gy (SF2) de todas las líneas celulares.

Tabla 4. Datos sobre la concentración, pureza y calidad del RNA de todas las muestras empleadas en el estudio.

Tabla 5. Listado de genes y su relación con ontologías principales relacionadas con funciones moleculares.

Tabla 6. Genes asociados a ontologías principales relacionadas con componente celular.

Tabla 7. Resumen del análisis por enriquecimiento en KEGG considerando todas las líneas celulares. Se representan los genes asociados a las ontologías más significativas resultantes del análisis.

Tabla 8. Genes con un impacto significativo ($p < 0.05$) en la supervivencia global medido mediante la prueba de regresión de Cox univariante. Se muestran ordenados de forma descendente por el valor de p .

Tabla 9. Genes con un impacto significativo ($p < 0.05$) en la SLP medido mediante la prueba de regresión de Cox univariante. Ordenados de forma descendente por el valor p .

Tabla 10. Resultados del análisis de la firma génica común por regresión de Cox univariante para SG y SLP

Tabla 11. Resultados del análisis de la firma génica común por regresión de Cox univariante para SG y SLP en función del tipo histológico principal (adenocarcinoma /escamoso).

Tabla 12. Resumen de las características de los pacientes de la cohorte principal del estudio.

Tabla 13. Resultados del análisis univariante (test de rangos logarítmicos) para potenciales factores pronósticos y predictivos de respuesta.

Tabla 14. Análisis de regresión de Cox univariante. Se analiza el impacto de las variables categóricas independientes sobre la SG.

Tabla 15. Análisis por regresión de Cox bivariante. Valoración del efecto de las variables categóricas independientes sobre la HR de la firma génica pronóstica de SG.

Tabla 16. Análisis de regresión de Cox univariante. Se analiza el impacto de las variables categóricas independientes sobre la SLP.

Tabla 17. Análisis por regresión de Cox bivariante. Valoración del efecto de las variables categóricas independientes sobre la HR de la firma génica pronóstica de SLP.

ABREVIATURAS

ABREVIATURAS

CPCNP: Cáncer de pulmón de células no pequeñas

RT: Radioterapia

SF2: Fracción superviviente a 2 Gy

TC: Tomografía computarizada

PET: Tomografía por emisión de positrones

RM: Resonancia magnética

RH: recombinación homóloga.

EGFR: Receptor del factor de crecimiento epidérmico

KRAS: Kirsten rat sarcoma viral ongene

MET: Proto-oncogen met

HER-2: Receptor 2 de factor de crecimiento epidérmico humano.

TKI: Inhibidor de tirosina quinasa.

DNA-PK: proteína quinasa dependiente de ADN

GDP: Guanosín difosfato

GTP: Guanosin trifosfato

MAPK: Proteínas quinasas activadas por mitógenos

STAT: Proteínas transductor de señal y activador de transcripción.

PI3K: fosfoinositol-3 quinasa

JAK2: Janus quinasa 2

EML4-ALK: Proteína 4 asociada a microtúbulo de equinodermo- quinasa de linfoma anaplásico

ROS1: proto-oncogén ROS1

NCI: Instituto nacional del cáncer.

RSI: Índice de radiosensibilidad.

GARD: Modelo basado en genoma para el ajuste de la dosis radioterápica

LQ: modelo lineal cuadrático

miRNA: microRNA

DRS: puntuación dosis-respuesta.

CEA: antígeno carcinoembrionario

IL-6: interleuquina 6

FEV₁: Volumen espiratorio forzado en primer segundo

GTV: Volumen tumoral grueso.

TCGA: Atlas del genoma del cáncer.

ATCC: American Type culture collection

RPMI: medio de cultivo Roswell Park Memorial Institute.

DMEM: medio de cultivo modificado de Dulbecco

HEPES: buffer HEPES

MV: megavoltio

UM: unidad monitor

PBS: tampón fosfato salino.

UFC: unidad formadora de colonias.

dCTP: Deoxicitidina trifosfato

TMeV: TIGR visualizador multiexperimento

SPSS: Paquete estadístico para las ciencias sociales.

NUSE: valores de error estándar normalizados sin escala

RLE: expresión logarítmica relativa.

FDR: False Discovery rate (tasa de falsos descubrimientos)

Log₂: logaritmo en base 2

KEGG: Enciclopedia de Genes y Genomas de Kioto

ECOG: Eastern Cooperative Oncology Group (Grupo de Oncología Cooperativa del Este)

TNM: Tumor, nodal, metástasis.

SLP: supervivencia libre de progresión

SG: supervivencia global

SCE: Supervivencia cáncer específica

IC: intervalo de confianza

HR: Hazard ratio (razón de riesgo)

DBE: dosis biológica equivalente

WHO-PS: World Health Organization- Performance Status (escala de estado de salud de la Organización Mundial de la Salud)

PUBLICACIONES

Publicaciones:

1- Gen signatures in non-small cell lung cancer: past, present and future. (En fase de redacción).

2- A 6-gene prognostic and predictive signature for early and locally advanced stages in non-small cell lung cancer. (En fase de redacción).