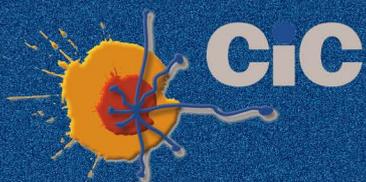
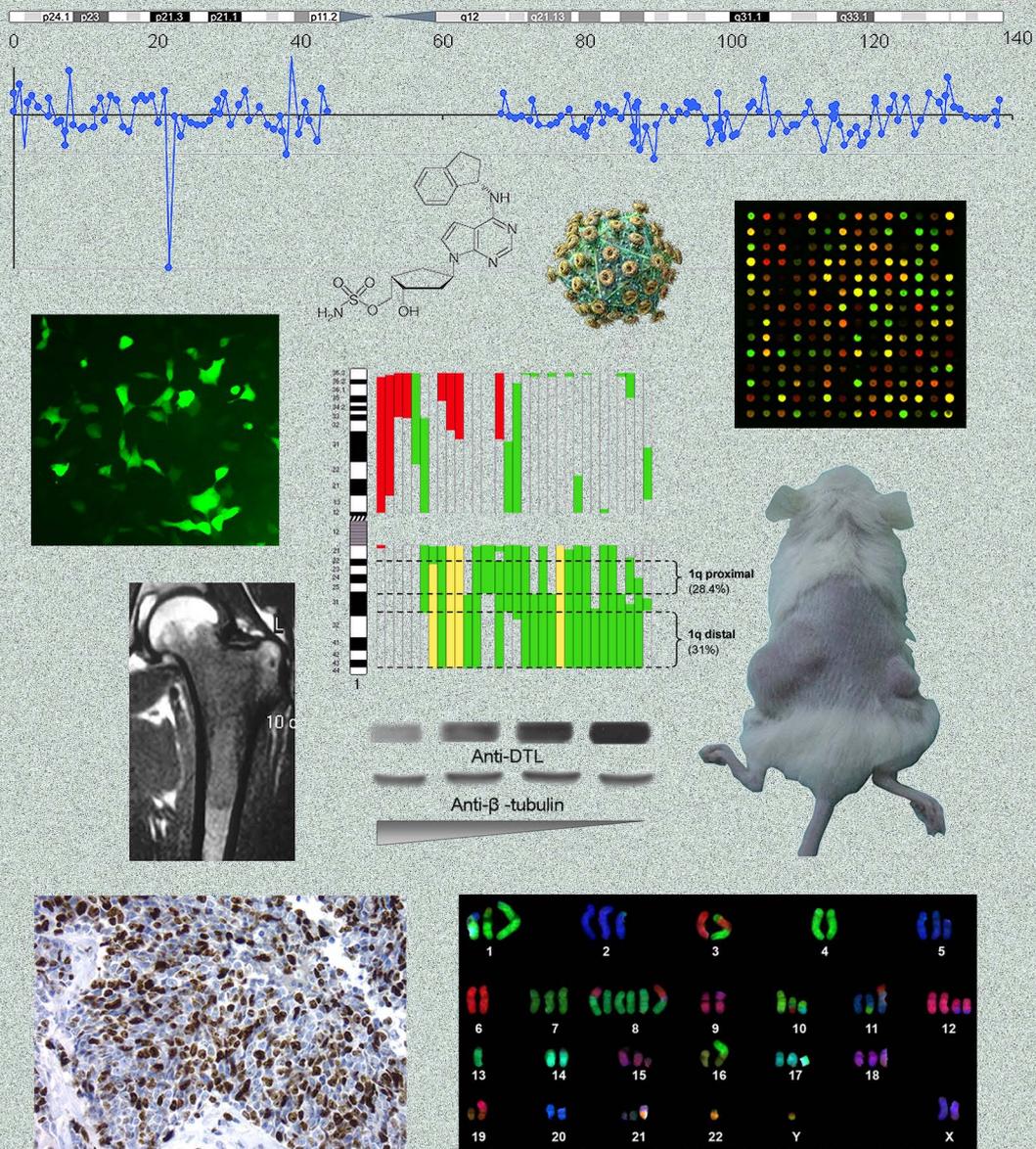


IDENTIFICACIÓN Y CARACTERIZACIÓN DE SUBTIPOS DE ENFERMEDAD EN SARCOMA DE EWING

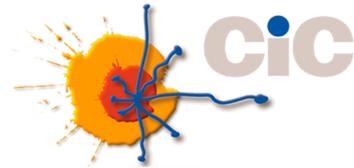
MEDIANTE TÉCNICAS DE CRIBADO MASIVO ENFOCADAS AL DESCUBRIMIENTO DE BIOMARCADORES Y DIANAS TERAPÉUTICAS



CARLOS MACKINTOSH GINEL
TESIS DOCTORAL - 2010



UNIVERSIDAD
DE SALAMANCA



INSTITUTO DE BIOLOGÍA MOLECULAR Y CELULAR DEL CÁNCER /
CENTRO DE INVESTIGACIÓN DEL CÁNCER

Identificación y caracterización de subtipos de enfermedad en Sarcoma de Ewing mediante técnicas moleculares de cribado masivo enfocadas al descubrimiento de biomarcadores y dianas terapéuticas.

TESIS DOCTORAL

Carlos Mackintosh Ginel
Salamanca, 25 de Junio de 2010

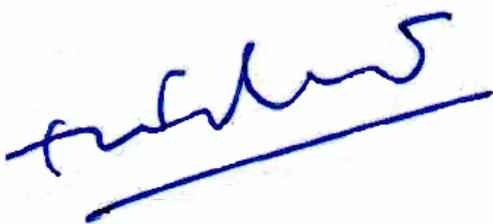
D. ENRIQUE DE ÁLAVA CASADO, Doctor en Medicina y Cirugía y Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Molecular y Celular del Cáncer (IBMCC, CSIC-USAL) de Salamanca,

CERTIFICA:

Que el trabajo realizado bajo mi dirección por el licenciado en Biología **CARLOS MACKINTOSH GINEL** titulado “**Identificación y caracterización de subtipos de enfermedad en Sarcoma de Ewing mediante técnicas moleculares de cribado masivo enfocadas al descubrimiento de biomarcadores y dianas terapéuticas**”, reúne las condiciones de originalidad y contenido requeridas para ser presentado ante el tribunal correspondiente y optar al grado de doctor por la Universidad de Salamanca.

Y para que conste a los efectos oportunos, expido y firmo el presente certificado mediante el cual autorizo la presentación de dicha tesis.

Salamanca, a 25 de Junio de 2010.



Fdo. Dr. Enrique de Álava Casado

Este proyecto de tesis ha sido financiado por:

- Una beca FPU del Ministerio de Educación (2004-2008)
- La Red de Excelencia de la Comunidad Europea EUROBONET, con el Proyecto Europeo “Molecular Pathology of Bone Tumors” (FP6-2004-Lifescihealth-5, 018814).
- El Ministerio de Ciencia e Innovación, a través del Instituto de Salud Carlos III (RD06/0020/0059, PI081828, PI052524, PI020828)

AGRADECIMIENTOS

Delante de esta tesis, antes de ella, está mi familia: tíos y tías, primos y primas y abuelos. Éstos muy especialmente: Pedro, Sacramento, Juan Carlos y Socorro. Pero sobre todo están mis madres. No tengo que decir aquí cuánto quiero a mis dos madres, a su vez la una madre de la otra. Con su cuerpo me tejieron este traje con el que asisto al mundo desde hace casi 30 años. Sobrevalorado ADN nuclear... son las mitocondrias las que conocen el camino a casa. Y además de ser el mismo barro, el mismo maravilloso y milenario barro de billones de citoplasmas maternos, soy barro amasado por las mejores manos, manos de madre que hacen la obra, más o menos imperfecta (para ellas por supuesto perfecta), y luego la cuidan todos y cada uno de sus días hasta morir/seguir viviendo en ella, porque en mi viviréis siempre Adela y Sacramento, porque soy lo que vosotras habéis hecho de mí.

Dentro de esta tesis está la otra familia. ¡Oiga!, con todas las de la ley que compartimos apellido: somos los Eleveintez (o L20ez). Cualquiera puede verlo en los libros de reservas de equipos o en la lista de los fines de semana (a la que damos demasiado uso), registros estos tan válidos para atestiguar el parentesco como cualquier anodino papelajo administrativo. Hay muchos parientes, algunos más lejanos y otros de primera sangre, aunque todos de la familia. De primera sangre son Jose L20 (también llamado José Luis Ordóñez), Vicky L20 (Victoria Sevillano pone en su DNI) y Ana Pastora L20 (en Bucaramanga es Ana Pastora Otero-Motta). Ellos hacen de este laboratorio un campo eólico cuyas aspas las mueve la risa y el buen rollo, la ayuda sincera y el apoyo incondicional. Son la materia misma que compone a este laboratorio, sin la cual no sería nada. Si la vida es una tesis ellos llevan tiempo graduados *cum laude, honoris causa* y tienen la cátedra de la materia de la Cojonudez Absoluta. Son las páginas en que descansan todas las letras de aquí a la última, son el buffer de reacción de esta tesis, una y otra vez el andamio sobre el que he levantando esta estructura que tantas veces se ha derrumbado (y yo con ella), son el abono pero no la mierda (esa son otros), son el cálido recibimiento por la mañana y el descojone en el café, son sencillamente lo mejor. Mucho más allá de las pipetas, las moléculas, las fórmulas y las cepas bacterianas, OS QUIERO. Aunque su paso por aquí ha sido más breve, no se escapan de la familia María, Juan, Alfredo, Carolina, Nilka y Cuca, todos ellos gente sencillamente excelente. Y por supuesto Dani, que se incorporó hace poco pero que ya es también de la familia. Compañero y amigo, qué más se puede pedir. El equipo perfecto de gitanos sureños.

Destacar especialmente a una Eleveinte que se nos fue a trabajar fuera de casa y de la que no pudimos disfrutar más, o al menos no tan a menudo. La pillina de pelo rizado y pecas, la niña más buena y alegre del planeta, la Pasión Vega de laboratorio que te canta y te hace olvidar las penas. Gracias por pasar por aquí Virginia Delgado Maroto, diamante puro, más que brisa fresca, aire acondicionado permanente.

Mención especial también para otro ilustre miembro de esta familia afincado ahora en Suiza: el “crack” de la risa y el compadreo, esa fuerza de la naturaleza, ese terremoto que dejó a Rychter sin escala y a Pfizer le dio de su sangre el principio activo de la viagra. Ese chorro vital con voz de camionero atrapado en el fondo de un pozo, el único, el inigualable... Deivid L20 (David Herrero, miembro fundacional de la familia, con carné desde el 2003).

Mis oráculos, mis wikipedias, mis Maniatis, mis Aristóteles de Alejandro Magno, mis Sanchos en la lucha diaria con los molinos, mis Yodas sin hiperbatos sintácticos: Antonella Chiechi, Alicia Ginel (la prima), y de nuevo Jose L20. Mis “Brainstorming” suelen dejarlos “Brainstormed” pero les da igual, si cito a su cerebro a un vis a vis con mis problemas, acudirán y me dejarán hasta su última sinapsis. A ellos debo la oposición dialéctica que conduce la chispa de la idea hasta la solución del problema científico, así como les debo rescates de las peligrosas parras a las que a veces me subo. En este apartado a Jose le corresponde mi inmensa gratitud por su vigilancia paternal, el interés diario por mi trabajo y sus dudas cartesianas que tantas veces me han hecho redibujar el esquema mal trazado. A Alicia Ginel, compañera de desgracias, por sus brillantes ideas, sus más que chispas fogonazos de miles de vatios que alumbran todo con la luz del talento puro. Si alguien conoce esta tesis aparte de mí es ella, si alguien conoce la suya aparte de ella soy yo.

El cerebro de Antonella Chiechi es digno de ser estudiado. Enciclopédica y llena de recursos, es capaz de ver la aguja en el pajar antes del primer parpadeo a la vez que habrá retenido la perfecta imagen del conjunto y de sus elementos uno por uno. Mi referente más asiduo. Ella es además un alto en el camino de esta tesis, un monumento en mi vida, uno de esos escasos y valiosísimos descubrimientos personales de los que me enorgullezco de llamar amiga.

Mención como no al jefe, Enrique de Álava, él ha puesto los medios a mi disposición y le agradezco sinceramente el haber hecho todo lo posible para que finalizara esta tesis más allá de los 4 años de mi beca.

Hay mucha más gente aquí. Mucha gente que ha contribuido con su trabajo, poniendo su pequeño granito de arena. Citarlos a todos es imposible pero su presencia anónima está por todo el texto de esta tesis: en las esquinas inferiores ayudando a pasar de página, en cada párrafo sujetando cada línea o en las pausas, plegados sobre sí mismos haciendo de comas. Son tantos que vista más allá de las palabras esto más que una tesis es un libro de “Buscando a Wally” en el que soy difícil de encontrar. La mayoría dibujados en plazas iluminadas o sorprendidos en un bello patio andaluz que se abre en el recodo del camino sofocante. Los menos en callejones oscuros a la espera del ataque por la espalda, con el filo listo para regalar un disgusto. Si hubiera una mención de “Desagradecimientos” habría que citar aquí a unas pocas personas, muy pocas por suerte, que han intentado torpedear esta investigación y mi propia vida. Como no existe dicho apartado y no deseo yo inventarlo los dejaré al resguardo de esa oscurísima sombra en la que tan bien se desenvuelven, en la que urden sus ridículos planes entonando su descerebrada e inagotable melopea de rumores y cuchicheos, hablando en esa lengua que tiene a la envidia y a la injuria como principales elementos articuladores. Esta tesis es a pesar suya.

De entre los que aportaron sus conocimientos y experiencia a este modesto trabajo que tanto carece de ambas cosas, me gustaría agradecer especialmente a Karoly Szuhai, Felipe Pimentel, Alberto de Luis, María Sacristán, Agustín Mayo Iscar, Mariano Igual, Martín Pérez de Andrés y Cristina Teodosio.

A los de “fuera del área”, no puedo citarlos a todos pero Pache, Álex y David no pueden faltar. Por ser ellos ya tienen que figurar, que para eso es mi tesis.

Y tras la tesis Carmen. “Tú eres todo, ¿lo sabes qué eres todo?, you are everything, everything. Eres la primera mujer en el primer día de la creación, tú eres la madre, la hermana, la amante, la amiga, el ángel, el demonio, la tierra, la casa. ¡Eso es lo que eres!, la casa...”. Carmen esperando tras la última página con el motor en marcha y la ruta al fin del mundo marcada en el mapa.

Con todos he aprendido, con la mayoría he disfrutado, muchos hemos sufrido juntos. Gracias a todos por vuestro aporte a mi vida. Todo el que pasó por aquí ha firmado con fuego en mi libro de visitas y nunca será olvidado. Glosaros a todos y todo lo que pasó sería imposible y daría para una historia mucho más extensa e interesante que esta, que quién sabe si algún día se escriba. Aquí quedan como hitos en el camino los que más han contribuido a lo que viene a continuación. Ha sido este un camino largo y duro, que probablemente no mereció la pena y del que he aprendido demasiado. Sin vosotros habría sido estéril pues sois su único fruto, su principal descubrimiento, el mejor resultado.

Finalmente, esta tesis está dedicada a Mercedes MM, quién más la ha sufrido después de mí. De este pedazo de mi vida tú te has llevado la peor parte. Este “tochazo”, que nunca leerás, te habría aburrido, te habrías burlado de él. Más tarde, sin mí delante, lo habrías admirado en silencio, impresionada en secreto por su lenguaje críptico, por el estilo añejo y a la vez atemporal de esta porción de nuestra tan respetada pero descuidada ciencia, museo de cera tan ajeno a ti que eres vida inmediata, sin teoremas. Tras ese lapso de reverencias habrías pensado “tantas palabrejas que no importan a nadie”. Y como siempre tendrías razón. Ojalá algún día lo leas hasta aquí y te burles del resto. Cómo nos lo pasaremos pintarrajeándolo todo.

TABLA DE CONTENIDOS

1. INTRODUCCIÓN	1
1.1 El sarcoma de Ewing	2
1.2 Presentación clínica	3
1.3 Histología	4
1.4 Factores con valor pronóstico	5
1.5 Tratamiento	6
1.6 La fusión quimérica EWS-FLI1	9
1.6.1 EWS	9
1.6.2 FLI1	10
1.6.3 EWS-FLI1	10
1.7 Origen celular y requerimiento de alteraciones secundarias	12
1.8 Estudios previos de alteraciones moleculares secundarias en SE	15
1.8.1 CDKN2A y TP53	15
1.8.2 Alteraciones de número de copia (ANC)	16
2. HIPÓTESIS Y OBJETIVOS	22
2.1 Hipótesis	23
2.2 Objetivos	23
2.2.1 Objetivo general	23
2.2.2 Objetivos específicos	24
3. MATERIAL Y MÉTODOS	25
3.1 Muestras y pacientes	26
3.2 Líneas celulares	27
3.3 Extracción de DNA	27
3.4 Hibridación genómica comparativa, BAC-array	28
3.4.1 Construcción del BAC-array	28
3.4.2 DOP-PCR y Aminolinking PCR	28
3.4.3 Microdeposición del producto de PCR	32
3.4.4 Marcaje del DNA	34
3.4.5 Hibridación	36
3.5 Análisis de los datos de los aCGH BAC-array	39
3.6 Hibridación de los SNP-array, modelo “GeneChip Human Mapping 500K Array Set” (Affymetrix)	41
3.6.1 Preparación de los DNA genómicos	41
3.6.2 Amplificación del DNA mediante PCR y procesamiento del producto de PCR	42
3.6.3 Marcaje del producto de PCR	44
3.6.4 Hibridación del array, lavados y escaneo	44
3.7 Análisis de los datos de los aCGH SNP-array	45

3.8	COBRA-FISH	46
3.9	Extracción de RNA	48
3.10	Microarrays de expresión	49
3.11	Análisis de los datos de los Microarrays de Expresión	50
3.12	SYBR Green Real Time RT-qPCR	53
3.13	Transducción lentiviral de líneas celulares de SE	57
3.14	Análisis de ciclo celular por citometría de flujo	59
3.15	Análisis de muerte celular por citometría de flujo	61
3.16	Ensayo de proliferación MTT	62
3.17	Ensayos in vivo de xenotransplantes en ratón	62
3.18	Extracción de proteínas y Western blot	63
3.19	Inmunohistoquímica	68
3.20	Clonación de las secuencias codificantes de WWOX y cdh13 en pMIG	69
3.21	Transducción retroviral de líneas celulares de SE	75
3.22	Análisis estadístico	77
4.	RESULTADOS	78
4.1	Resultados de los estudios de aCGH en líneas celulares y muestras tumorales de SE	79
4.2	Estudios del impacto de las ANC en los datos clínicos	86
4.3	Estudios transcriptómicos y de su relación con las ANC	91
4.4	Selección de un gen candidato en 1q, DTL, y validación de sus niveles mediante Real Time RT qPCR	97
4.5	Validación funcional del gen candidato DTL, in vitro e in vivo	98
4.6	Comprobación del efecto del silenciamiento de DTL en los sustratos del complejo CUL4/DDB1 ^{CDT2}	105
4.7	Estudios de microarray de expresión para comparar el perfil de expresión diferencial de los tumores 1qGSig y el debido al silenciamiento de DTL	108
4.8	Estudio inmunohistoquímico de la tasa proliferativa diferencial de los tumores 1qGSig	109
4.9	Validación de genes candidatos en 16q	110
5.	DISCUSIÓN	114
5.1	Consideración comparativa de la metodología usada para el abordaje del estudio de las ANC	115
5.2	Comparación de los resultados de aCGH obtenidos con la literatura previa	118
5.3	Las ANC contribuyen a la progresión del SE y son clínicamente relevantes.	119
5.4	Significado clínico y biológico de las ANC más recurrentes en SE, consideradas individualmente	121
5.5	Ganancia del brazo cromosómico 1q en cáncer	125
5.6	La ganancia de 1q se asocia con un perfil de desregulación del ciclo celular en SE, resultados bioinformáticos.	126
5.7	Selección del gen candidato en 1q	127
5.8	El gen candidato DTL, sobre-expresado en los tumores 1qGSig, ejerce un papel fundamental en el fenotipo de ciclo celular desregulado y mayor proliferación característico de este grupo de tumores.	128
5.9	Función del gen DTL, implicación en cáncer y potencialidad como diana terapéutica	131
5.9.1	Complejos E3 ligasa tipo RING finger	131

5.9.2	Expresión de DTL en tejidos humanos y de ratón_____	134
5.9.3	Regulación de la actividad de los complejos CRL y aplicación en terapia dirigida _____	135
5.10	Modelo molecular propuesto para la contribución de DTL a la oncogénesis y desarrollo del SE____	137
5.11	Otros genes candidatos en 1q validados en la literatura previa_____	140
6.	CONCLUSIONES _____	141
7.	BIBLIOGRAFÍA _____	144

ÍNDICE DE FIGURAS

Figura 1. Localización del SE y distribución por edad y sexo _____	4
Figura 2. Estudio morfológico e inmunohistoquímico de la biopsia de un paciente de SE _____	5
Figura 3. Ejemplo de SE localizado en el fémur y de tratamiento local quirúrgico _____	8
Figura 4. Esquema de la proteína de fusión EWS-FLI1 _____	11
Figura 5. Curva de supervivencia debida a la mutación del <i>locus</i> TP53 y a la delección homocigótica del <i>locus</i> CDKN2A _____	17
Figura 6. Ejemplo del patrón de bandeo en electroforesis en gel de agarosa de varias Aminolinking PCR _____	32
Figura 7. Ejemplo de una imagen adquirida tras el proceso de escaneado del BAC-array de la librería “Sanger 1 mb clone set” _____	37
Figura 8. Ampliación del bloque 1 de la figura anterior (bloque superior izquierdo) _____	38
Figura 9. Mapa del vector pMIG depositado en el repositorio Addgene (www.addgene.org) por su creador _____	71
Figura 10. Ejemplo de comprobación de 12 minipreps de construcciones pMIG-WWOX mediante restricción con la enzima ClaI _____	74
Figura 11. Ejemplo de comprobación de 8 minipreps de construcciones pMIG-cdh13 mediante restricción con la enzima EcoRI _____	75
Figura 12. Frecuencia de las ANC en las muestras tumorales y en las líneas celulares de SE _____	79
Figura 13. Representación de la posición y extensión de las ANC, en muestras tumorales (arriba) y en líneas celulares (abajo) de SE _____	80
Figura 14. Localización genómica y extensión de las regiones principales de ganancia en 1q _____	81
Figura 15. Comparación de los parámetros globales de ANC en muestras tumorales y líneas celulares de SE _____	82
Figura 16. Microdelección del <i>locus</i> CDKN2A detectada mediante aCGH BAC-array en la línea celular STAET2.1 _____	83
Figura 17. Microdelección en 16q23.2 que define la RMS de esta región, detectada en la línea celular SKES1 mediante aCGH BAC-array y delimitada de manera precisa mediante aCGH SNP-array _____	85
Figura 18. Cariotipo de la línea celular RM82, caracterizado mediante COBRA-FISH _____	86
Figura 19. Perfil de aCGH de la línea celular RM82 _____	86
Figura 20. Impacto de las ANC, consideradas de forma global, en la supervivencia total de los pacientes de SE _____	87
Figura 21. Definición de un grupo de muestras tumorales que comparten la firma de expresión local característica de la 1qG _____	92

Figura 22. Gráficas de enriquecimiento generadas por el análisis GSEA, correspondientes a las colecciones génicas _____	93
Figura 23. Resultado del análisis Ingenuity, mostrando las principales rutas canónicas enriquecidas en el perfil de expresión diferencial de los tumores 1qGSig _____	95
Figura 24. Resultado del análisis de clusterización jerárquica no supervisado, mostrando dentro del recuadro amarillo el cluster que agrupa a los tumores 1qGSig _____	96
Figura 25. Curvas de supervivencia de los pacientes con SE del grupo 1 y 2 _____	96
Figura 26. Validación por Real Time RT-qPCR de las diferencias en los niveles de mRNA de DTL entre los grupos de tumores 1qGSig y el resto de tumores _____	99
Figura 27. Comprobación mediante Real Time RT-qPCR de los niveles del mRNA de DTL en un panel de líneas celulares de SE y comparación con los valores de un microarray de expresión previo _____	100
Figura 28. Western blot que muestra la reducción de los niveles proteicos de DTL conseguidos mediante cada una de las construcciones shRNA ensayadas _____	101
Figura 29. Comprobación de la eficiencia de transducción en una población de TC32 transducida con la construcción pLKO.1-turboGFP _____	102
Figura 30. Análisis de ciclo celular y apoptosis mediante citometría de flujo tras el silenciamiento completo de DTL _____	103
Figura 31. Ensayo de silenciamiento gradual de DTL en las líneas celulares RM82 y TC32 _____	104
Figura 32. Resultados del ensayo <i>in vivo</i> de silenciamiento de DTL, que confirma el efecto inhibitor de la proliferación de las líneas celulares de SE _____	106
Figura 33. Western blot de 3 líneas celulares de SE con anticuerpos contra reguladores de la transición G1/S, tras silenciamiento completo de DTL _____	107
Figura 34. Niveles de señal de los “probesets” de TP53, CDKN1A y CDKN1B obtenidos en el estudio de microarrays de expresión del silenciamiento de DTL en líneas celulares de SE _____	108
Figura 35. Resultado del análisis GSEA comparando el perfil de expresión debido al silenciamiento de DTL con el perfil diferencial de los tumores 1qGSig _____	109
Figura 36. Resultados de la evaluación de la tasa proliferativa de los tumores 1qGSig mediante inmunohistoquímica del marcador Ki-67 _____	110
Figura 37. Valoración de los niveles de expresión de los supresores tumorales WWOX y cdh13 en el panel de líneas celulares de SE _____	111
Figura 38. Recuperación de los niveles de expresión de los supresores tumorales WWOX y cdh13 en las líneas células con delección de sus <i>loci</i> y niveles de expresión más bajos _____	112
Figura 39. Control GFP de la transducción retroviral con pMIG. Imagen de contraste de fases y de fluorescencia _____	113
Figura 40. Inmunohistoquímica con el anticuerpo anti-WWOX en un TMA de muestras tumorales de SE _____	113
Figura 41. Ejemplos de datos generados mediante aCGH SNP-array _____	116

Figura 42. Frecuencias de las ANC en SE según la literatura previa _____	119
Figura 43. Distribución genómica de las ANC en SE según la literatura previa _____	119
Figura 44. Frecuencia global de ANC del cromosoma 1 en muestras tumorales humanas _____	125
Figura 45. “Boxplot” de los niveles de expresión de DTL y p21 ^{WAF1/CIP1} en los tumores 1qGSig y el resto de tumores _____	128
Figura 46. Esquema de los diferentes complejos proteína ubiquitina ligasa y detalle del ensamblaje y regulación de los complejos de tipo CRL _____	133
Figura 47. Patrón de expresión del mRNA de DTL en tejidos humanos adultos y tumores humanos _____	135
Figura 48. Modelo molecular de la posible contribución de 1qG a la patogenia del SE _____	139

ÍNDICE DE TABLAS

Tabla 1. Frecuencia de las fusiones génicas más comunes en SE _____	9
Tabla 2. Evaluación de las RMS de frecuencia superior al 35% y tamaño inferior a 10 mb _____	84
Tabla 3. Resumen de los resultados del cribado bioinformático de las ANC con impacto clínico _____	90
Tabla 4. Resultados más representativos del análisis GSEA _____	94
Tabla 5. Información detallada de las construcciones shRNA usadas en el silenciamiento de DTL _____	100
Tabla 6. Combinaciones de elementos que conforman los principales complejos CRL _____	133

GLOSARIO DE ABREVIATURAS

Para las abreviaturas más comunes se ha optado por mantener los acrónimos según la nomenclatura en inglés, por motivos de claridad y congruencia.

I

1qG · Ganancia de 1q
1qGSig · Firma de expresión local (genes en 1q) asociada a la ganancia de 1q
1qN · 1q normal, no alterado

A

aCGH · Hibridación genómica comparativa sobre microarray
ANC · Alteraciones de número de copia

B

BAC · Cromosoma artificial bacteriano
BAC-array · Microarray de CGH basado en una librería genómica de BAC
BSA · Albúmina sérica bovina

C

CCD · Dispositivo de cargas eléctricas interconectadas
CDKI · Inhibidor de CDK (quinasa dependiente de ciclina)
cDNA · DNA complementario
CGH · Hibridación genómica comparativa
COBRA-FISH · Hibridación in situ fluorescente de tipo marcaje combinado binario y de ratio
CRL · E3 ligasa de tipo dedos RING con culina
cRNA · RNA complementario
CT · Ciclo umbral

D

dCTP · Nucleótido citosina trifosfato
DNA · Ácido desoxirribonucleico
dNTP · Desoxinucleótido trifosfato
DOP-PCR · PCR con cebadores de secuencia degenerada

E

ES · "Enrichment Score", valor de enriquecimiento, parámetro del análisis GSEA.
ESFT · Tumores de la familia del Sarcoma de Ewing
EWS-ETS · Gen quimérico que combina el gen EWS y a un miembro de la familia génica ETS, frecuentemente FLI1

F

FBS · Suero bovino fetal
FDR · Tasa de falsos positivos
FISH · Hibridación fluorescente in situ
Fw · Del inglés "Forward", hacia adelante, en sentido 5'-3'

G

GFP · Proteína verde fluorescente
GO · "Gene Ontology"
GSEA · Análisis de enriquecimiento de colecciones génicas

H

HIV-1 · Virus de la inmunodeficiencia humana adquirida
hMSC · Células madre mesenquimales humanas
HPV · Virus del papiloma humano

I

ICD-O-3 · Clasificación internacional de enfermedades oncológicas, versión O-3
IP · Ioduro de Propidio

K

kb · Kilobase
KDa · Kilodalton

L

log2ratio · Ratio logaritimizado en base 2

M

mb · Megabase
mCGH · Hibridación genómica comparativa sobre metafases
MEFs · Fibroblastos embrionarios murinos
Microarray · Microdisposición, micromatriz o microordenación de secuencias de DNA
milliQ · Agua filtrada y purificada por ósmosis inversa
moi · Multiplicidad de infección

MPC · Células progenitoras mesenquimales primarias
mRNA · RNA mensajero
MSC · Células madre mesenquimales
MSigDB · Bases de datos del análisis GSEA que contienen diferentes colecciones génicas
MTT · Metil tiazol tetrazolio

N

NCBI · "National center for biotechnology information", perteneciente al Instituto Nacional de la Salud (NIH) estadounidense
NES · "Normalized Enrichment Score", valor de enriquecimiento normalizado, parámetro del análisis GSEA.
NOD/SCID · Obesos no diabéticos con inmunodeficiencia severa combinada
NTA · Número total de alteraciones de número de copia
NTC · "Non-targeting control" o control no dirigido

O

oligo dT · Cebador compuesto de 18 nucleótidos de timidina

P

pb · Pares de bases
PBS · Tampón salino fosfato
PCR · Reacción en cadena de la polimerasa
PGA · Porcentaje de genoma alterado
PNET · Tumor primitivo neuroectodérmico
Probeset · Conjunto de sondas de un microarray de expresión dirigido contra los transcritos de un mismo locus
p-ter · Extremo p-terminal cromosómico

R

Real Time RT qPCR · Reacción de amplificación por PCR cuantitativa en tiempo real a partir de cDNA retrotranscrito

RMA · "Robust multi-array average", algoritmo de normalización por cuartiles de los datos de microarrays de expresión
RMS · Región mínima de solapamiento
RNA · Ácido ribonucleico
rpm · Revoluciones por minuto
rRNA · RNA ribosómico
Rv · Del inglés "Reverse", del revés, en sentido 3'-5'

S

SAM · "Significance analysis of microarrays", algoritmo de expresión diferencial
SDS · Dodecil sulfato sódico
SE · Sarcoma de Ewing
shRNA · RNA de interferencia ahorquillado
siRNA · RNA pequeño de interferencia
SNP · Polimorfismo de nucleótido único
SNP-array · Microarray de oligonucleótidos aplicable al estudio de ANC mediante aCGH
SSC · Citrato sódico salino
SV40 · Virus 40 vacuolado del simio

T

Taq polimerasa · Polimerasa de DNA termoestable recombinante de *Thermus aquaticus*
Tm · Temperatura de "melting"
TMA · Microarray o micromatriz de tejidos
tRNA · RNA transferente

U

UTR · Región no traducida del mRNA
UV · Ultravioleta

β

β ME · Beta mercaptoetanol

*“Allá al fondo está la muerte, pero no tenga miedo. Sujete el reloj con una mano, tome con dos dedos la llave de la cuerda, remóntela suavemente. **Ahora se abre otro plazo, los árboles despliegan sus hojas, las barcas corren regatas, el tiempo como un abanico se va llenando de sí mismo y de él brotan el aire, las brisas de la tierra, la sombra de una mujer, el perfume del pan.**”*

*¿Qué más quiere, qué más quiere? Átelo pronto a su muñeca, déjelo latir en libertad, imítelo anhelante. **El miedo herrumbra las áncoras, cada cosa que pudo alcanzarse y fue olvidada va corroyendo las venas del reloj, gangrenando la fría sangre de sus rubies.** Y allá en el fondo está la muerte si no corremos y llegamos antes y comprendemos que ya no importa.”*

Julio Cortázar. Instrucciones para dar cuerda al reloj.
Historias de cronopios y de famas

1. INTRODUCCIÓN

1.1 El Sarcoma de Ewing

El Sarcoma de Ewing (SE) es un tumor de células pequeñas y redondas que aparece generalmente en los huesos y, menos frecuentemente, en tejidos blandos y que afecta fundamentalmente a niños y adolescentes.

En rigor, el término SE representa a la entidad mayoritaria de un grupo más amplio de tumores con rasgos morfológicos e histológicos diversos que fueron definidos históricamente como neoplasias diferentes, pero que han sido aunados posteriormente a la luz de los descubrimientos en el campo de la biología molecular de estos tumores. El proceso de descubrimiento y clasificación ha sido el siguiente:

- En 1918 se describe el caso de un tumor compuesto de células indiferenciadas y redondas y dispuestas en rosetas localizado en el nervio ulnar de un paciente, al que se define como tumor primitivo neuroectodérmico o **PNET de tejido blando**¹.

- En 1921 un tumor primario de células pequeñas y redondas localizado en el hueso radio de una niña de 14 años es definido como “endotelioma difuso de hueso”. La designación de esta entidad quedó sin embargo vinculada al nombre de su descubridor, James Ewing, conociéndosele en adelante como **SE**².

- En 1975 se propone el primer caso de **SE extraesquelético**, derivado de un tejido blando³.

- En 1979 se define el **tumor de Askin**, con características histológicas similares al PNET, como un “tumor maligno de células pequeñas de la región toracopulmonar”⁴.

- En 1984, se describe el primer tumor neuroectodérmico (PNET) localizado en hueso, de nuevo compuesto por células pequeñas y redondas, al que se denomina **PNET de hueso**⁵.

- La identificación en ese mismo último año de un evento molecular común a la práctica totalidad de estas entidades, la translocación $t(11;22)(q24;q12)$ ⁶⁻⁸, y el posterior descubrimiento del gen quimérico EWS-FLI1 derivado de ella⁹ (unido a las evidencias del papel central que dicha alteración tiene en la tumorigénesis de estos sarcomas) sirvieron para englobarlas a todas en la categoría de **Tumores de la Familia del Sarcoma de Ewing o ESFT**¹⁰. Por motivos de claridad y para ajustarnos a la clasificación de la Organización

Mundial de la Salud (OMS) de tumores de hueso y partes blandas de 2002¹¹, aquí utilizaremos el término SE de forma extensiva para referirnos a todas estas entidades.

La media de edad de los pacientes afectados por SE es de 15 años, comprendiendo alrededor del 4% del total de tumores de niños y adolescentes. La incidencia es de 1 a 3 casos por millón de habitantes y año en los países del hemisferio Norte, siendo el segundo sarcoma de hueso más frecuente en pacientes menores de 20 años. Se localiza fundamentalmente en las extremidades distales y proximales, pelvis, tórax, cabeza y cuello. La incidencia en varones es ligeramente superior, en una proporción 1.3-1.5:1 respecto a mujeres¹². Por razas, la patología es más frecuente en caucásicos que en asiáticos, mientras que africanos y afro-americanos raramente la padecen¹³ (Figura 1).

1.2 Presentación clínica

Los síntomas más habituales son dolor e inflamación de la región afectada. También son comunes el incremento de la velocidad de sedimentación globular, la leucocitosis y la anemia. En los tumores localizados en la diáfisis de huesos largos es observable una progresiva reacción de la estructura ósea, con una característica reacción “en capas de cebolla” del periostio (Figura 3).

La localización en hueso es la más frecuente (85%) repartiéndose anatómicamente como sigue: extremidades distales (27%), proximales (25%), pelvis (20%), tórax (20%), columna vertebral y cráneo (9%).

Los SE extraesqueléticos o de partes blandas (15% de los casos) se distribuyen de la siguiente manera: tronco (32%), extremidades (26%), cabeza y cuello (18%), retroperitoneo (16%) y otras localizaciones (8%)¹⁴.

El SE tiene una gran tendencia a la metástasis, principalmente en hueso y pulmón. Más del 10% de pacientes presentan múltiples metástasis a hueso. La metástasis a hueso, pulmón y médula ósea es detectable al diagnóstico en un 25% de los pacientes, mientras que la metástasis a ganglios linfáticos es rara.

El diagnóstico anatomopatológico es complejo debido a la heterogeneidad histológica y la ausencia de marcadores inmunohistoquímicos patognomónicos. De este modo, sólo la detección molecular de la translocación específica $t(11;22)(q24;q12)$ y/o de la fusión génica asociada a ella (EWS-ETS) permiten diferenciar a esta entidad de otras como el neuroblastoma, el rhabdomyosarcoma, el linfoma no Hodgkin, el tumor desmoplásico de células pequeñas o el sarcoma sinovial¹⁵⁻¹⁶.

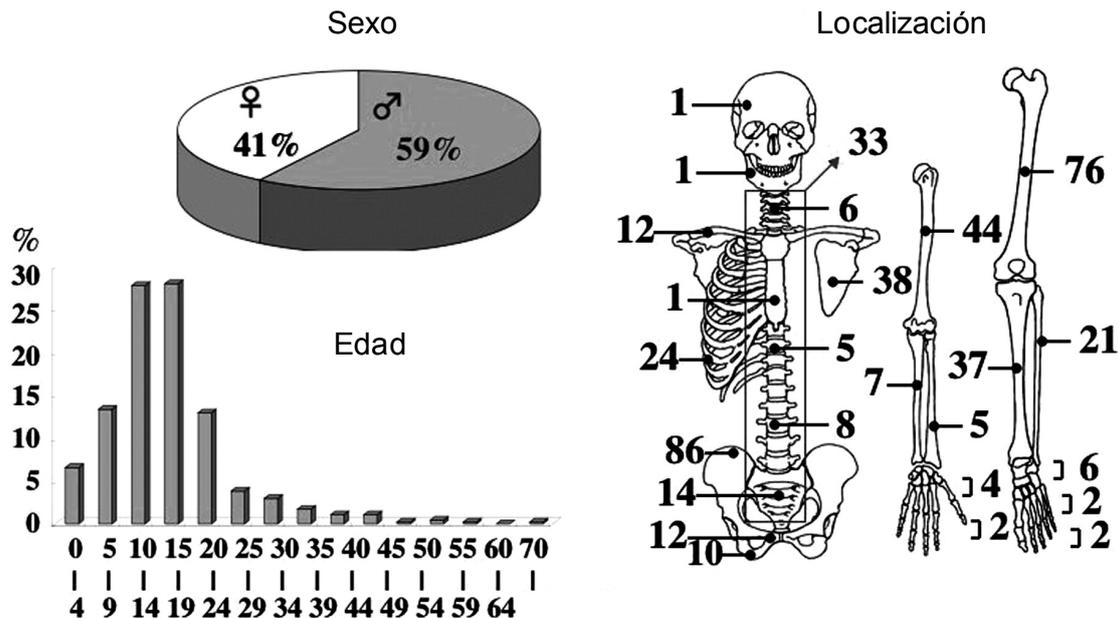


Figura 1. Localización del SE y distribución por edad y sexo. Adaptado de Iwamoto, Y. (2007). Diagnosis and treatment of Ewing's sarcoma. Jpn J Clin Oncol 37, 79-89)

1.3 Histología

El SE se caracteriza por presentar células pequeñas con núcleos prominentes, nucleolo ausente y cromatina dispersa (Figura 2), así como bajo índice mitótico. En ocasiones se aprecian células con núcleos hiper cromáticos, probablemente degenerativas, que conforman el patrón conocido como “células claras/células oscuras”. El citoplasma suele ser escaso, claro y rico en glucógeno, por lo que las células son positivas para la tinción del ácido periódico de Schiff (PAS). La disposición celular es densa, sin apenas matriz extracelular mientras que el patrón de crecimiento carece de una organización clara.

Los marcadores antigénicos típicos son CD99 (presente en un 90% de los tumores, Figura 2), la enolasa específica neural (NSE), S-100, CD57 y sinaptofisina. En un 20% de casos hay también presencia de citoqueratinas y neurofilamentos¹⁷.

En contraposición a este aspecto difuso típico del SE, las células del PNET se agrupan en nidos, trabéculas o rosetas, aunque en ocasiones también son apreciables grupos de células semejantes a la histología anteriormente descrita para el SE. Además, aunque los PNET presentan también inmunoreactividad para CD99, NSE y/o β -2 microglobulina, se distinguen por su positividad para marcadores neurales tales como Leu-7, sinaptofisina, neurofilamento, y S100. De hecho, la presencia de gránulos neurosecretorios es su principal elemento distintivo¹⁸⁻²⁰.

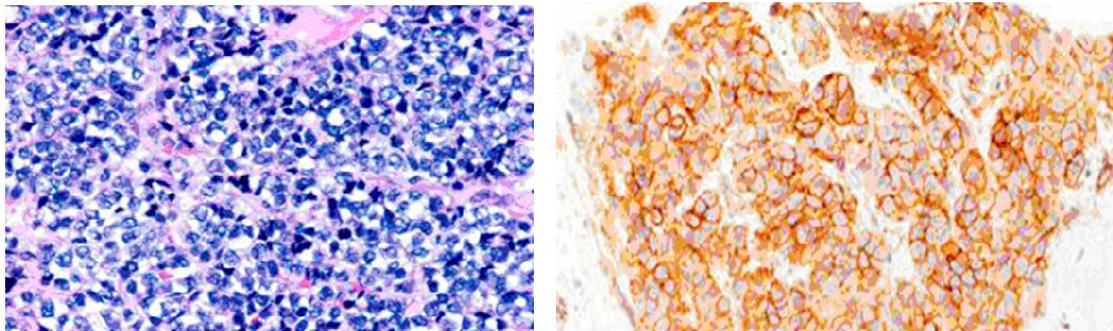


Figura 2. Estudio morfológico e inmunohistoquímico de la biopsia de un paciente de SE. Tinción de PAS con un patrón típico de células redondas azules con citoplasma escaso (izquierda). Inmunohistoquímica para CD99 con típico marcaje de membrana (derecha).

1.4 Factores con valor pronóstico

Los principales factores clínicos con valor pronóstico en SE son:

- Metástasis: la presencia de metástasis al diagnóstico (20-30% de los pacientes) es el factor con mayor impacto en la mala supervivencia de los pacientes, presentando una mejor supervivencia aquellos con metástasis a pulmón respecto a los que presentan metástasis en hueso. El resto de factores pronósticos no producen diferencias en supervivencia tan marcadas como las asociadas a este parámetro.
- Localización del tumor primario: las localizaciones en extremidades tienen mejor pronóstico que las centrales o pélvicas²¹⁻²².
- Tamaño del tumor primario: en la mayoría de los estudios el tamaño del tumor ha demostrado ser un factor pronóstico importante. Para definir los tumores más grandes se usa un límite de 100 ó 200 ml²³.

- Edad del paciente: los niños menores de 10 años tienen un mejor pronóstico que los adolescentes, los adultos jóvenes o los adultos²¹⁻²².
- Sexo del paciente: las mujeres afectadas de SE tienen un mejor pronóstico que los varones.
- Deshidrogenasa láctica: altos niveles de la enzima deshidrogenasa láctica (DHL) previos al tratamiento correlacionan con un mal pronóstico²². Los tumores primarios grandes y los metastáticos presentan una elevada concentración de DHL.

Los factores moleculares con valor pronóstico descrito son:

- Variantes del gen de fusión EWS-FLI1: aunque en estudios retrospectivos las distintas variantes de la fusión EWS-FLI1 (producidas por la existencia de diferentes sitios de ruptura en los intrones de ambos *loci* implicados, ver más adelante) parecían tener diferencias en el pronóstico clínico²⁴, un reciente trabajo prospectivo con muestras tumorales de más de 700 pacientes ha descartado este punto²⁵. Este resultado ha sido además refrendado por un trabajo independiente publicado al mismo tiempo²⁶.
- Detección del transcrito de la fusión EWS-ETS en muestras de médula ósea: la detección de pequeñas poblaciones de células tumorales en médula ósea mediante Real Time RT-qPCR se ha asociado a un mayor riesgo de recidiva²⁷.
- Delección de CDKN2A y mutaciones en TP53: consideradas por separado, cada una de estas alteraciones afecta a menos del 15% de los pacientes de SE. Sin embargo, aunque infrecuentes, se ha demostrado el peor pronóstico que confieren en varios estudios retrospectivos, el más completo de ellos el elaborado a partir de 60 muestras tumorales²⁸.

1.5 Tratamiento

Antes del uso de quimioterapia múltiple la supervivencia de los pacientes de SE era inferior al 10%. El uso de terapias multimodales ha conseguido mejorarla, incrementándola hasta alcanzar la tasa actual del 70%. Sin embargo, el pronóstico de supervivencia de aquellos pacientes con enfermedad primaria diseminada o recaída se mantiene aún por debajo del 20%²⁹.

El tratamiento del SE combina quimioterapia sistémica multimodal con tratamiento local (cirugía, radioterapia o ambas)³⁰⁻³³. Los agentes quimioterápicos que han demostrado mayor efectividad son doxorubicina (DXR), ciclofosfamida (CPA), vincristina (VCR), actinomicina-D (ACT), ifosfamida (IFM) y etopósido (VP16).

En 1968 se usó por primera vez una combinación de quimioterápicos y tratamiento local en SE. Dicha combinación incluía CPA, VCR y radioterapia. Era el inicio de la aplicación de la terapia multimodal en esta neoplasia³⁴. A partir de entonces se han ensayado diversas combinaciones que han incidido en un aumento paulatino de la supervivencia de los pacientes:

- VACD (VCR, ACT, CPA y DXR) y radioterapia: la inclusión de DXR en el esquema compuesto por sólo 3 agentes quimioterápicos demostró ser muy superior, incrementando la supervivencia libre de recaída del 24 al 60%³⁵. Además, se demostró una ulterior mejora en este factor al aumentar la dosis de DXR respecto a la usada inicialmente, alcanzando una tasa del 74% para los SE libres de metástasis³⁶. Estos ensayos se hicieron en el seno de los estudios estadounidense IESS-I e IESS-II y en el europeo CESS-81, y establecieron el esquema VACD como terapia estándar en SE³⁷.

- En Europa, el estudio CESS-86 ensayó la sustitución de DXR en el esquema VACD por IFM para el tratamiento de los pacientes con tumores con volúmenes superiores a 100ml. El resultado fue satisfactorio³².

- La implementación de IFM y VP16 (IE) en el esquema VACD (VACD-IE) en ensayos de fase II del estudio INT-0091 consiguió un notable incremento en la supervivencia libre de recaída (54% a 69%)³⁸, lo que supuso la adopción del esquema VACD-IE como nuevo estándar en el tratamiento del SE³⁹.

Respecto al manejo local, existe aún cierta controversia: aunque varios estudios muestran la eficacia de la resección del tumor reduciendo la tasa de recaída local (hasta niveles inferiores al 10%) e incrementando la supervivencia total, no existen ensayos aleatorizados comparando cirugía y radioterapia. En consecuencia, cuando es posible la resección del tumor con márgenes amplios normalmente prevalece la cirugía sin radioterapia, mientras que si no es claro que se alcanzarán los márgenes quirúrgicos necesarios se aplica radiación pre-operatoria y/o quimioterapia neoadyuvante. Además, si tras la cirugía los márgenes adecuados no han

sido alcanzados se puede hacer uso de radioterapia postoperatoria. Finalmente, cuando los márgenes son claramente inalcanzables por método alguno la amputación puede ser la única opción para el manejo de la enfermedad local (Figura 3). En contraposición, los tumores inoperables (debido normalmente a su localización central y/o tamaño masivo) son tratados localmente sólo con radioterapia⁴⁰⁻⁴¹.

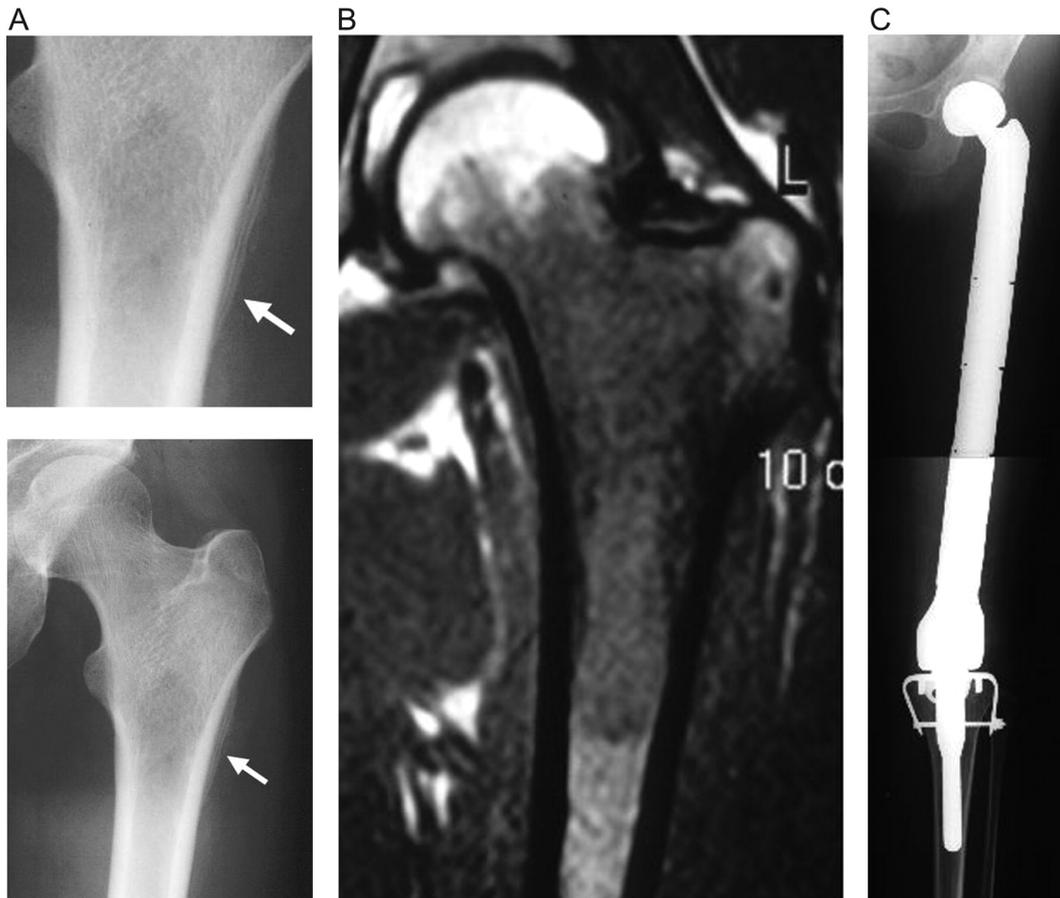


Figura 3. Ejemplo de SE localizado en el fémur y de tratamiento local quirúrgico (A) Radiografía mostrando una lesión invasiva metafiso-diafisaria femoral con una reacción perióstica en capas de cebolla. La imagen superior es una ampliación de la zona marcada en la inferior (*Flecha*). (B) Imagen de Resonancia Magnética mostrando un tumor intramedular expandiéndose desde el segmento largo de la diáfisis hasta el cuello femoral. (C) Reconstrucción de la extremidad afectada con una endoprótesis tras la resección. Reproducido de Iwamoto, Y. (2007). Diagnosis and treatment of Ewing's sarcoma. *Jpn J Clin Oncol* 37, 79-89.

A pesar de la evolución en el manejo clínico y su notable contribución a la supervivencia de los pacientes de SE, el tratamiento actual es insuficiente para un subgrupo de pacientes que o bien presenta enfermedad diseminada al diagnóstico o sufre recaída posterior. Estos pacientes muestran una buena respuesta inicial al tratamiento pero en la mayoría de los casos sólo se consigue controlar la enfermedad parcialmente y ésta acaba reapareciendo^{32,42-43}. Asimismo, no existe aún una estratificación clínica del SE debida a parámetros moleculares predictivos

ya que el conocimiento de estos es prácticamente nulo en esta enfermedad. En consecuencia es necesaria la búsqueda de nuevos factores pronósticos moleculares, así como de nuevas dianas terapéuticas que incidan en un mejor manejo clínico de este subtipo agresivo de SE.

1.6 La fusión quimérica EWS-FLI1

Para explicar las implicaciones moleculares de la alteración primaria del SE tomaremos como prototipo la fusión génica EWS-FLI1, producto de la translocación t(11;22)(q24;q12), ya que es la combinación génica EWS-ETS claramente mayoritaria (85% de los casos) y la más estudiada. Existen otras combinaciones génicas que involucran al gen EWS y a otros genes de la familia ETS, pero su frecuencia es notablemente menor (Tabla 1).

Translocación	Fusión Génica	Frecuencia
t(11;22)(q24;q12)	EWS-FLI1	85%
t(21;22)(q22;q12)	EWS-ERG	10%
t(7;22)(p22;q12)	EWS-ETV1	< 5%
t(17;22)(q12;q12)	EWS-E1AF	< 5%
t(2;22)(q33;q12)	EWS-FEV	< 5%

Tabla 1. Frecuencia de las fusiones génicas más comunes en SE. Adaptado de Burchill, S. A. (2003). Ewing's sarcoma: diagnostic, prognostic, and therapeutic implications of molecular abnormalities. *J Clin Pathol* 56, 96-102).

Empecemos por considerar brevemente las funciones nativas de los genes implicados:

1.6.1 EWS

La proteína correspondiente al *locus* EWSR1 pertenece a la familia de proteínas TET, caracterizada por la presencia de un dominio carboxiterminal de 87 aminoácidos que comprende un número variable de repeticiones RGG (arginina-glicina-glicina) a través del cual es capaz de unirse a complejos RNA-proteínas⁴⁴.

Aunque las funciones de EWS no son aún bien conocidas sus características estructurales y su capacidad de unión al RNA, junto con las conclusiones obtenidas de diversas experiencias realizadas, sugieren su participación en procesos de transcripción y procesamiento del RNA⁴⁴.

La expresión de EWS es ubicua y el alelo no translocado se expresa de manera normal⁴⁶. La localización habitual de la proteína es nuclear aunque su distribución subcelular puede verse alterada por diversos estímulos y por modificaciones post-traduccionales⁴⁷.

La región perteneciente a EWS involucrada en las fusiones EWS-ETS es indispensable para la acción oncogénica de esta ya que su sustitución por otros dominios proteicos elimina su potencial transformador⁴⁸.

1.6.2 FLI1

El *locus* de FLI1 fue inicialmente identificado como un lugar de integración frecuente de retrovirus inductores de leucemias murinas⁴⁹⁻⁵⁰. Su producto proteico pertenece a la familia de factores de transcripción ETS, caracterizada por la presencia de un dominio de unión al DNA (dominio ETS) altamente conservado, de 85 aminoácidos, cuya secuencia consenso de unión al DNA es GGA(A/T)⁵¹.

La expresión de FLI1 es elevada durante el desarrollo embrionario en células hematopoyéticas y endoteliales y en las células mesenquimales derivadas de la cresta neural⁵²⁻⁵³. En tejidos adultos su expresión está restringida al tejido hematopoyético, timo e hígado. En estos últimos la detección de su expresión se debe a la presencia de precursores de linfocitos, eritrocitos y megacariocitos⁵⁴. De hecho FLI1 activa la expresión de diversos genes específicos de megacariocitos y es capaz de inducir la diferenciación a este tipo celular⁵⁵⁻⁵⁷.

El papel de FLI1 como transformador oncogénico *per se*, especialmente en leucemias eritrocíticas, está mediado por el incremento de la supervivencia celular, inducción de factores anti-apoptóticos y la promoción de la progresión del ciclo celular⁵⁸⁻⁶¹.

1.6.3 EWS-FLI1

La proteína quimérica EWS-FLI1 (Figura 4) es el resultado de la unión del dominio N-terminal de EWS con el dominio C-terminal de FLI1 (que engloba al dominio ETS de unión al DNA), como consecuencia de la translocación t(11;22)(q24;q12). Se han descrito hasta 12 variantes diferentes de EWS-FLI1, según los puntos de ruptura y unión en los diversos

intrones de ambos genes en la translocación. La forma mayoritaria es la conocida como tipo 1, cuyo transcrito incluye la porción 5' del mRNA de EWS hasta el exón 7 y la porción 3' del de FLI1 desde el exón 6 hasta la 3' UTR. De manera análoga la fusión de tipo 2 comprende hasta el exón 7 de EWS y desde el exón 5 de FLI1, mientras que la fusión de tipo 3 llega hasta el exón 10 de EWS e incluye desde el exón 6 de FLI1.

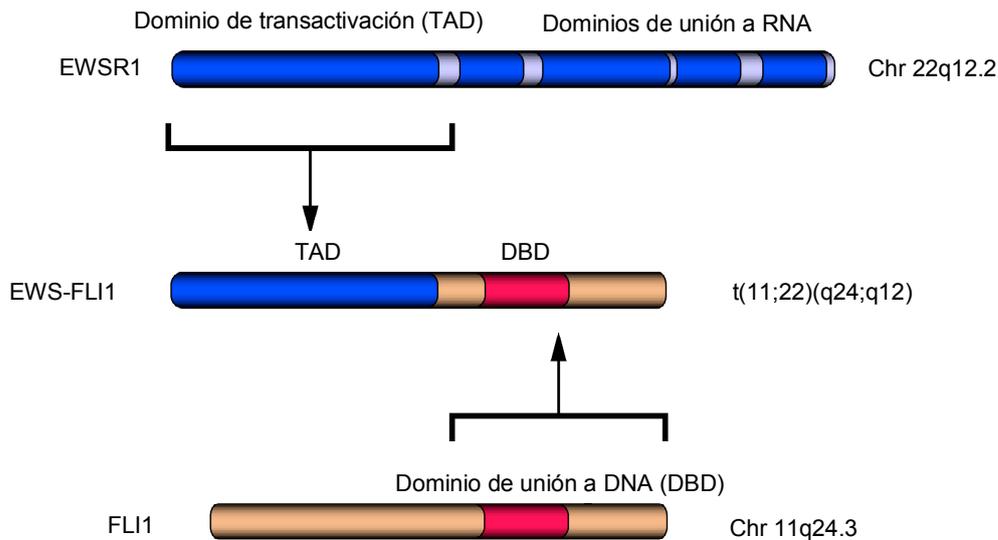


Figura 4. Esquema de la proteína de fusión EWS-FLI1

Existen evidencias que demuestran que el dominio aminoterminal aportado por EWS posee un potencial de transactivación superior al del dominio equivalente nativo de FLI1, que es reemplazado por aquel. De este modo, a pesar de que la fusión EWS-FLI1 comparte con la proteína FLI1 nativa la especificidad de unión al DNA, la proteína quimérica tiene funciones de activador/represor transcripcional más marcadas⁶²⁻⁶⁵. En consecuencia, la sobre-expresión exógena de FLI1 no es capaz de recapitular los efectos producidos por la fusión, tales como el crecimiento en agar blando o la formación de tumores en ratones inmunodeprimidos⁴⁸.

Los mutantes de EWS-FLI1 deficientes tanto en la región correspondiente a EWS como en la de FLI1 carecen de capacidad de transformación, lo cual demuestra la importancia de cada uno de ambos elementos para la acción oncogénica de la fusión quimérica⁶⁵.

Los genes desregulados por EWS-FLI1 mejor caracterizados son CDKN1A (p21), CCND1 (ciclina D1), MYC e Id2, que sufren un fuerte aumento de expresión y p57, el cual es

reprimido. La acción directa de la fusión sobre los promotores de estos genes ha sido demostrada⁶⁶⁻⁶⁸.

Además, EWS-FLI1 afecta a componentes de distintas vías de señalización, principalmente la de TGF β RII⁶⁹ y la de IGF1R. El efecto sobre esta última es especialmente relevante en la biología del SE⁷⁰⁻⁷¹. La acción sobre esta ruta es ejercida mediante la represión de IGFBP3, un importante regulador de esta vía⁷², y la activación de IGF1R⁷³.

Otros genes relevantes recientemente sugeridos como dianas de EWS-FLI1 son DAX1/NR0B1⁷⁴, PTPL1⁷⁵, PLD2⁷⁶, NKX2.2⁷⁷, RXR⁷⁸, VEGF⁷⁹ y Gli1⁸⁰.

Aparte de los hallazgos relativos a la acción de EWS-FLI1 sobre genes puntuales, un reciente trabajo⁸¹ ha profundizado en las consecuencias de la actividad desreguladora de esta sobre las principales funciones celulares, integrando los datos transcriptómicos de diversas líneas celulares de SE antes y después del silenciamiento de la fusión (mediante técnicas de siRNA o shRNA) con los de 59 tumores primarios de SE. Según los resultados de este estudio los genes afectados positivamente por la proteína quimérica desarrollan funciones de regulación del ciclo celular (siendo esta la función más afectada de todas), proliferación y respuesta a daño en el DNA, mientras que los genes reprimidos están principalmente involucrados en funciones de diferenciación y comunicación celular.

1.7 Origen celular y requerimiento de alteraciones secundarias

La heterogeneidad morfológica e histológica del SE, unida a la diversidad de localizaciones que presenta, ha dificultado la determinación de su origen y ha suscitado todo un abanico de teorías respecto a este punto, abarcando propuestas tan dispares como un inicio a partir de células madre de gran potencialidad o la existencia de diferentes células de origen.

Sin embargo, a pesar de que la diversidad de presentaciones parece apuntar a esto último (diversos tipos celulares originarios) se han aportado evidencias demostrando que las características neurales de los PNET, el subgrupo que más controversias plantea, se deben a la acción de la proteína quimérica EWS-ETS en lugar de ser una reminiscencia del “background” celular de origen⁸². Este hallazgo, junto con la permisividad diferencial que han mostrado las células madre mesenquimales (MSC) y sus progenitores a la expresión exógena

de EWS-FLI1 respecto a la refractariedad manifiesta de otros tipos celulares⁸³⁻⁸⁵, hacen de estas las principales candidatas para explicar el origen celular del SE. La facilidad de las MSC para migrar desde la médula ósea a localizaciones extraesqueléticas explicaría todo el abanico de entidades comprendidas dentro de la ESFT.

Por otra parte, los resultados y conclusiones de los trabajos que han pretendido desvelar el origen celular del SE han generado multitud de evidencias útiles para resolver otra cuestión importante: la del requerimiento o no de alteraciones moleculares secundarias. Y es que, aceptando la necesidad que el SE tiene de la fusión quimérica EWS-ETS para su oncogénesis⁸⁶, cabe plantearse ¿es suficiente este evento molecular para establecer y desarrollar la neoplasia?

Para intentar responder a esta pregunta repasemos los resultados obtenidos en algunos de los más importantes trabajos de la investigación relativa a la célula de origen del SE:

- La expresión exógena de EWS-FLI1 en fibroblastos embrionarios murinos (MEFs) induce parada de ciclo celular y apoptosis. En MEFs carentes de p16^{INK4a}, p19^{ARF} o p53 (procedentes de ratones knockout para estas funciones génicas) EWS-FLI1 es capaz de mantener su expresión pero no es capaz de lograr la transformación celular, la cual sólo se produce al co-introducir el antígeno T del virus SV-40⁸⁷.

- En fibroblastos humanos primarios, inmortalizados mediante expresión exógena de hTERT, la introducción de EWS-FLI1 provoca igualmente una pronta parada de ciclo mediada por p53, así como supone una sensible inducción de apoptosis. La introducción de EWS-FLI1 junto con la expresión exógena de la proteína E6 del HPV o de una forma de p53 dominante negativa consigue anular esta parada de ciclo inicial pero no libera a las células de sistemas de regulación ulteriores, puesto que el cultivo detiene de nuevo su crecimiento al cabo de unos días. De manera muy interesante, la expresión de EWS-FLI1 no consigue transformar estos fibroblastos ni siquiera ante la presencia conjunta de las proteínas E6 y E7 del HPV, del antígeno T del virus SV40 y del oncogén RAS^{V12}. De hecho, estos oncogenes son capaces de inducir *per se* la formación de colonias en agar blando, pero dicha capacidad es anulada al añadir al cóctel oncogénico la expresión de EWS-FLI1⁸³.

- La introducción de EWS-FLI1 en la línea celular C2C12 (una línea mioblástica murina que retiene la potencialidad para diferenciar no sólo a tejido muscular, si no también a óseo y

adiposo), induce una parada en el curso de la diferenciación miogénica. Cabe destacar que esta línea mioblástica sí resultó ser permisiva a la expresión exógena de la fusión quimérica⁸².

- Por otra parte la introducción de EWS-FLI1 en células progenitoras mesenquimales primarias (MPC) de ratón demostró la capacidad de estas células no sólo de mantener la expresión de EWS-FLI1 sino también de ser transformadas por su acción. Estas células conservaron intactos los *loci* tanto de p53 como de p16INK4A/p19ARF y mantuvieron una ruta funcional de p53 tras la introducción de la fusión quimérica⁸⁴.

- Un estudio paralelo con un planteamiento similar confirmó estos resultados. En este trabajo se demostró además que los tumores generados en ausencia de p53 eran más agresivos⁸⁸. De manera análoga, otro estudio en el que se introdujo EWS-FLI1 en MPC, esta vez carentes de p19^{ARF}, demostró un bloqueo de los programas de diferenciación adipogénico y osteogénico característicos de estas células, debido a los efectos moleculares de la proteína quimérica⁸⁹.

- Sin embargo, la introducción de la fusión EWS-FLI1 en células madre mesenquimales humanas (hMSC) ha producido resultados notablemente distintos a los obtenidos con las MPC de ratón: aunque permiten y mantienen la expresión de la proteína quimérica, ésta no consigue incrementar la baja tasa proliferativa y la viabilidad de las hMSC, ni tampoco es capaz de transformarlas (no producen tumores en xenotransplantes en ratón) o de retrasar su senescencia.

Tomadas en conjunto, todas estas evidencias sugieren la necesidad de otras alteraciones moleculares para el establecimiento o al menos para el desarrollo del SE ya que, incluso en modelos celulares humanos permisivos a la presencia de la fusión quimérica (como las hMSC), ésta no es capaz de inducir transformación por sí sola. Además, los resultados obtenidos en hMSC podrían simplemente estar reflejando la mayor permisividad de los entornos celulares pluripotentes debido a la ausencia en ellos de algunas de las restricciones moleculares que se oponen a la acción transformadora de EWS-ETS, como la ruta de p53 o la expresión de p16^{INK4a}, que de hecho anulan la reprogramación celular inducida⁹⁰⁻⁹³. Este punto no ha sido comprobado puesto que no se ha ensayado la permisividad de entornos tales como cultivos neuroectodérmicos primarios humanos o células madre embrionarias humanas.

1.8 Estudios previos de alteraciones moleculares secundarias en SE

La gran mayoría de la investigación desarrollada en el ámbito de la biología molecular del SE se ha centrado en las implicaciones moleculares de la fusión EWS-ETS, principalmente en sus acciones aberrantes, directas o indirectas, sobre la transcripción de los llamados “genes diana de la fusión” y en la relevancia de estos. Mientras tanto, la presencia y contribución de otras alteraciones, que llamaremos secundarias por considerarse EWS-ETS el evento primario, casi no han sido abordadas. Es por ello por lo que actualmente se tiene un escaso conocimiento relativo a qué eventos moleculares, independientes de la fusión génica, están involucrados en el desarrollo tumoral del SE.

La relevancia de dichas alteraciones secundarias estriba en su potencialidad en la estratificación clínica de la enfermedad ya que la fusión quimérica carece de valor pronóstico al estar presente en todos los tumores de esta entidad y debido a que, tal y como recientemente se ha demostrado²⁵⁻²⁶, no existen diferencias clínicas entre las diversas combinaciones génicas o a las variantes exónicas de cada una de ellas. En este escenario las alteraciones secundarias emergen como candidatos prometedores para explicar las evoluciones clínicas más adversas de la enfermedad y podrían asimismo suponer nuevas dianas para terapia dirigida.

Los pocos trabajos que han explorado esta parte de la biología molecular del SE se han centrado en la búsqueda de alteraciones de los genes tradicionalmente involucrados en cáncer y en estudios de alteraciones genómicas de número de copia. Repasemos los principales estudios de alteraciones secundarias realizados en esta área:

1.8.1 CDKN2A y TP53

Del estudio de *Huang et al.*²⁸ se desprende que las deleciones homocigóticas del *locus* CDKN2A afectan a alrededor de un 15% de pacientes de SE, mientras que un porcentaje similar se ve afectado por pérdidas de función de p53. Aunque claramente infrecuentes, ambas alteraciones confieren mal pronóstico en estudios tanto univariados como multivariados (Figura 5). Otros autores han publicado frecuencias diferentes para ambas alteraciones, pero las series muestrales usadas son sensiblemente más pequeñas, generalmente inferiores a 30 pacientes, y las técnicas de detección aplicadas no son las adecuadas⁹⁴⁻⁹⁷. Por

el contrario, en el estudio de *Huang et al.*²⁸ se incluyen muestras de 60 pacientes obteniendo unas frecuencias totales del 13% para cada una de las citadas alteraciones. El estudio además es el que utiliza la tecnología más indicada, teniendo en cuenta la diversidad de mecanismos de alteración que afectan a ambos *loci*:

- Las deleciones de CDKN2A son evaluadas mediante una sonda de FISH de tan sólo 100 kb, restringida a la posición del *locus* y a su tamaño e hibridada sobre improntas de las muestras tumorales. La microdelección más pequeña de CDKN2A publicada en SE⁹⁷, detectada en una línea celular y de 58 kb de extensión, abarca un 60% de dicha sonda por lo que el margen de falsos negativos de este estudio es menor que el de otros. De hecho, son más probables los falsos positivos si tenemos en cuenta que se consideraron como casos con delección de CDKN2A homocigótica (la única que suprimiría realmente las funciones proteicas de p16^{INK4a} y p14^{ARF}) a aquellos con al menos un 20% de células carentes de señal para la sonda (de más de 100 núcleos evaluados por muestra). Este umbral es bastante permisivo y considera positivas para la delección del *locus* a muestras que muy probablemente resultarían negativas mediante otras tecnologías (todas aquellas que parten de extractos moleculares de poblaciones celulares completas).

- Para la evaluación del *locus* TP53 se utilizó un abordaje combinando estudios de expresión en tejido (inmunohistoquímica), de RNA (microarrays de expresión) y de secuenciación exónica.

1.8.2 Alteraciones de número de copia (ANC)

Nos referiremos con este término (ANC) a las aberraciones moleculares que afectan al DNA genómico de manera que varían la dosis génica global (a nivel de población celular) de alguna de sus regiones o subregiones cromosómicas. Las técnicas capaces de obtener información acerca del estado de número de copias genómico a partir de DNA aislado de una población celular y marcado con fluorocromos se conocen como técnicas de hibridación genómica comparativa (CGH) y difieren en el soporte usado para la co-hibridación:

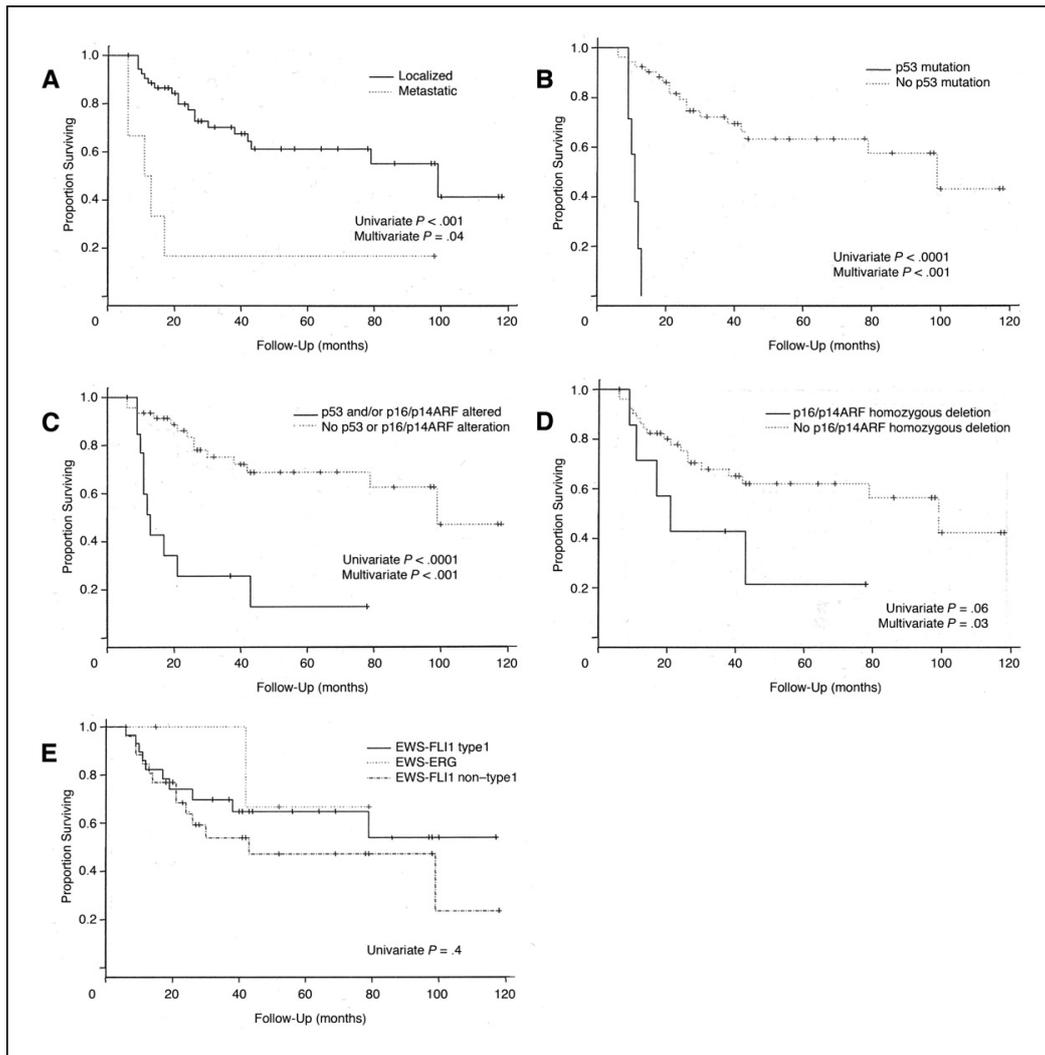


Figura 5. Curva de supervivencia debida a la mutación del locus TP53 (B) y a la delección homocigótica del locus CDKN2A (D) o de ambas consideradas conjuntamente (C) (Reproducido de Huang, H. *et al* (2005). Ewing sarcomas with p53 mutation or p16/p14ARF homozygous deletion: a highly lethal subset associated with poor chemoresponse. *J Clin Oncol* 23, 548-558.)

- **CGH sobre metafases (mCGH)⁹⁸**: en esta técnica el DNA problema (tumoral) es co-hibridado junto con un DNA de referencia (DNA sano o no alterado) con extensiones cromosómicas en metafase, sobre un porta de microscopia. Al estar el DNA tumoral y el de referencia marcados con distintos fluorocromos la co-hibridación de cada una de las sondas es en cada región cromosómica de las metafases proporcional a la presencia cuantitativa y correlativa de dicha región en los DNA tumoral y de referencia. La imagen capturada del microscopio de fluorescencia mediante un dispositivo CCD es analizada con programas informáticos específicos que detectan la presencia o no de desviación en cada región cromosómica respecto al balance de señales de fluorescencia esperable para una co-hibridación de muestras no alteradas. Dicha desviación es un reflejo de la ganancia o pérdida de material genómico en el DNA tumoral y, además de detectable, es cuantificable.

- **CGH array (aCGH)**: la tecnología es prácticamente idéntica salvo en el soporte sobre el cual se co-hibridan los DNA tumoral y referencia, puesto que en esta variante la extensión de cromosomas en metafase es sustituida por un microarray o micromatriz de secuencias genómicas. Tal y como se explica en el apartado de Material y Métodos correspondiente (3.4.1), una librería genómica humana contenida en vectores BAC es microdepositada sobre una superficie similar a la de un porta de microscopía (denominaremos a este subtipo de aCGH como BAC-array), o bien dicha superficie tiene adherida una colección de oligonucleótidos de síntesis química con secuencias pertenecientes al genoma humano (denominaremos a este otro subtipo SNP-array puesto que esta tecnología fue diseñada para el estudio de variaciones del tipo “single nucleotide polymorphisms” o polimorfismos de nucleótido único).

Estas técnicas informan en definitiva sobre variaciones de dosis genómica del tipo de delección, duplicación o amplificación genómica, pero son “ciegas” a los reordenamientos genómicos.

La diferencia fundamental entre ambos tipos de CGH es la resolución (el tamaño mínimo de las ANC detectables), de 5 mb para el mCGH⁹⁹, variable para el aCGH dependiendo de la distancia media genómica comprendida entre los elementos de la librería (BAC-array) o entre los oligonucleótidos (SNP-array) representados en el microarray. Además, el rango dinámico y la sensibilidad son notablemente superiores en las plataformas de aCGH¹⁰⁰.

Los estudios de ANC con estas tecnologías que se han llevado a cabo con muestras de SE han resultado muy fructíferos en la identificación de las ANC con mayor recurrencia y muy útiles para conocer sus respectivas frecuencias en la población tumoral. Sin embargo, son varias las carencias de estos trabajos previos:

- Uso de tecnologías hoy obsoletas caracterizadas por su escasa resolución y sensibilidad¹⁰¹⁻¹⁰⁶. Se trata de trabajos basados en cariotipado (que no aporta una información rigurosa a nivel poblacional) y/o mCGH.

- Ausencia de estudios de correlación de las ANC con los parámetros clínicos de los pacientes¹⁰¹ o ausencia de utilización de métodos de análisis bioinformático masivo y no sesgado: en los casos en que los autores evaluaron el impacto clínico de las ANC procedieron seleccionando las regiones testadas de una manera arbitraria o poco rigurosa en lugar de hacer

un cribado masivo, no sesgado, de la implicación clínica de todas y cada una de las regiones alteradas.

- Uso de series muestrales de menos de 30 pacientes^{101-103,107-108}.

- Ausencia de combinación de los datos de los estudios de ANC con datos de análisis transcriptómico y correlación de ambos con datos clínicos. Tan sólo dos trabajos¹⁰⁷⁻¹⁰⁸ integran estas 3 capas de información.

Todos los estudios publicados hasta la fecha adolecen de alguna de estas carencias, restando valor a las conclusiones obtenidas ya sea por su escaso rigor estadístico (series muestrales pequeñas), por la alta probabilidad de ANC no detectadas (tecnologías anticuadas) o por la falta de integración de los resultados de los estudios de ANC con otras capas de información (transcriptómica y/o clínica).

Revisaremos aquí los resultados de los trabajos más completos, ya sea por los tamaños muestrales utilizados, por la fiabilidad de las técnicas usadas y/o por la integración de varios niveles de información (molecular y clínica):

- Los resultados del estudio de mCGH realizado por Brisset *et al.*¹⁰⁴ con muestras de 43 tumores (21 primarios, 22 metastásicos) indican que las ganancias del brazo 1q, junto con las de los cromosomas 8 y 12 son las ANC más frecuentes en SE. El número medio de ANC parece ser mayor en las muestras de tumores metastásicos, al igual que la ganancia de 1q que parece estar sobrerrepresentada también en las metástasis. Sin embargo, ninguna de estas observaciones alcanza significación estadística.

- Ozaki *et al.*¹⁰⁵ llegan a conclusiones similares en otro estudio de mCGH, esta vez realizado a partir de 62 muestras (52 tumores primarios y 10 recaídas). El estudio apunta de nuevo a una mayor incidencia de ANC en los tumores de pacientes con recaída (5 frente a 2.5), esta vez con valor estadístico significativo. Las ANC más frecuentes son de nuevo las ganancias de los cromosomas 8 y 12, junto con la del 20 y la ganancia del brazo 1q. Las pérdidas más frecuentes afectan a los brazos cromosómicos 16q y 19q. No se encuentran asociaciones con la histología, el tamaño tumoral, el estadio, la localización o la respuesta histológica a la terapia. Se halla una diferencia estadística en la supervivencia total entre los pacientes con tumores con 5 o menos ANC y los de 5 o más (con menor supervivencia los pacientes de este último grupo). Las ganancias del 1q, 2q, 12 y 20 y la pérdida del 16q y 17p tienen asociadas

peores supervivencias en estudios estadísticos univariados. En multivariados la pérdida del 16q es la única ANC significativa.

- El estudio de Hattinger *et al.*¹⁰⁶ es el que utiliza la mayor serie muestral hasta la fecha (n=138 pacientes) haciendo uso de diversas técnicas citogenéticas (cariotipo, mCGH y FISH). Las alteraciones más recurrentes encontradas coinciden con las descritas anteriormente. Se halla la ganancia del 1q como la principal alteración con impacto en supervivencia (menor supervivencia asociada). También tienen menores supervivencias los pacientes con tumores con pérdida del 16q (asociada además a metástasis) y ganancia del cromosoma 12.

- Ferreira *et al.*¹⁰⁷ publican el primer estudio de SE hecho con aCGH (SNP-array), añadiendo además información de microarrays de expresión. Sin embargo el tamaño muestral es de tan sólo 25 pacientes. Se destacan como conclusiones la presencia de ANC en un 84% de las muestras tumorales, una media de 3 ANC por tumor, y una menor supervivencia de aquellos pacientes con tumores que albergan 3 o más ANC (grupo al que se considera “genómicamente inestable”). Los patrones transcriptómicos de estos dos grupos de tumores difieren en las funciones génicas relativas a la segregación cromosómica, la reparación de DNA y control del ciclo celular, sobre-expresadas en el grupo inestable.

- Hasta la fecha, el estudio de Savola *et al.*¹⁰⁸ es el más completo de todos puesto que, aunque con un planteamiento similar al anterior (n=31 muestras, estudio de aCGH SNP-array) y conclusiones muy parecidas (87% de tumores con ANC, se definen de nuevo como tumores inestables aquellos con más de 3 ANC y se confirma su peor supervivencia), se realizó una búsqueda de genes candidatos entre aquellos situados en las regiones genómicas más representativas y con mayor cambio a nivel de expresión debido a la alteración de dosis génica subyacente. El principal candidato, el gen HDGF, situado en 1q, tiene mayor expresión en las muestras tumorales respecto a muestras control (células CD34+ y tejido muscular), pero las diferencias de expresión no correlacionan con una peor supervivencia de los pacientes.

De estos trabajos (los más representativos hasta la fecha en el estudio de ANC) se pueden establecer una serie de conclusiones, comunes a todos ellos:

- Entre un 80 y un 90% de las muestras tumorales de SE presentan ANC.

- Las ANC más recurrentes en SE son las ganancias de los cromosomas 8, 12, 20 y 1q y las pérdidas del 16q.

- Los pacientes con tumores con 3 o más ANC tienen una peor supervivencia.

- Existen ANC que, consideradas individualmente, tienen un impacto en los parámetros clínicos, entendiendo por esto la existencia de asociación entre las ANC y ciertas circunstancias clínicas, así como con una peor supervivencia.

Sin embargo, muchas cuestiones quedan sin resolver:

- Los trabajos realizados con la técnica más fiable (aCGH) no cuentan con grupos muestrales suficientes para resolver las discrepancias acerca de qué ANC tienen impacto a nivel clínico, un aspecto sobre el cual existen notables diferencias.

- No se establece mediante qué mecanismos moleculares contribuyen las ANC más representativas, como la ganancia del 1q o la delección del 16q, al desarrollo tumoral y al peor resultado clínico.

- No se identifica qué genes situados en las regiones de ANC tienen un papel central en su efecto molecular y clínico.

Responder a estas preguntas es el principal objetivo de esta tesis.

2. HIPÓTESIS Y OBJETIVOS

2.1 Hipótesis

La hipótesis principal del presente estudio consiste en que las alteraciones moleculares secundarias tienen un papel biológico relevante en la oncogénesis y/o el desarrollo tumoral del SE y, dado que el evento molecular primario (EWS-ETS) carece de valor pronóstico, estas alteraciones moleculares son las únicas potencialmente aplicables a la estratificación clínica del SE.

Esta hipótesis general comprende la asunción de otras hipótesis menores:

1. Las ANC son alteraciones moleculares secundarias que tienen un probable papel biológico y clínico relevante en SE puesto que, en esta y en cualquier otra neoplasia, existe un perfil de ANC distintivo con ANC recurrentes específicas.
2. El impacto de las ANC en la clínica es un reflejo de su contribución a los procesos moleculares de oncogénesis y/o desarrollo tumoral.
3. El fenotipo molecular asociado a las ANC con mayor impacto clínico y los genes situados en ellas suponen una fuente potencial de marcadores pronósticos y dianas de tratamiento dirigido.

2.2 Objetivos

2.2.1 Objetivo general

La literatura científica actual del SE presenta muy pocos estudios de alteraciones moleculares secundarias y de su implicación en esta entidad tumoral. El trabajo presentado en esta tesis pretende identificar alteraciones de este tipo, así como caracterizar y evaluar su relevancia biológica y clínica, contribuyendo a definir subgrupos de pacientes de SE que se caractericen por la presencia de determinados marcadores pronósticos y dianas terapéuticas.

2.2.2 Objetivos específicos

Para lograr el objetivo general antes descrito se propusieron los siguientes objetivos específicos:

1. Realizar un estudio inicial de evaluación de ANC mediante técnicas de cribado masivo de alta resolución (basadas en la tecnología de microarrays) en una serie retrospectiva de muestras tumorales de al menos un tamaño doble al de las series estudiadas mediante aCGH en la literatura actual.
2. Evaluar mediante técnicas bioinformáticas y estadísticas el impacto clínico global de las ANC, como una confirmación inicial de la relevancia de este tipo de alteraciones en SE.
3. Aplicar una metodología de análisis bioinformático y bioestadístico que permita un cribado masivo de correlaciones de las ANC, consideradas individualmente, con los datos clínicos asociados a las muestras tumorales estudiadas, de modo que se identifiquen sin sesgo de elección las ANC más relevantes clínicamente en SE.
4. Estimar el efecto a nivel transcripcional de las ANC individuales con mayor impacto clínico, delimitando las principales funciones celulares afectadas por dichas ANC y definiendo así el fenotipo molecular asociado a ellas.
5. Seleccionar genes candidatos en las principales regiones de ANC con relevancia clínica y validarlos funcionalmente, demostrando *in vitro* e *in vivo* su implicación en el fenotipo molecular y funcional asociado a dichas ANC (hallado en los estudios *in silico*). Para ello se aplicarán técnicas que permitan recuperar los niveles del gen candidato (en el caso de deleciones de material genómico que conlleven la pérdida de su expresión), o bien que permitan reducir su sobre-expresión (en el caso de ganancias genómicas y aumento de los niveles de los genes asociados).
6. Valorar la aplicabilidad de los genes candidatos validados como biomarcadores y dianas de tratamiento dirigido.

3. MATERIAL Y MÉTODOS

3.1 Muestras y pacientes

Se utilizaron dos series de pacientes con SE, que denominaremos de ahora en adelante como grupo 1 y grupo 2.

El grupo de tumores utilizado para los estudios de aCGH, grupo 1 de tumores, comprende 66 tumores primarios y una metástasis pulmonar de pacientes con SE, recogidos por 2 instituciones europeas: la Universidad Católica de Lovaina (UCL), que aportó DNA de 40 tumores, y la Universidad Heinrich-Heine (HHU) de Düsseldorf, que cedió DNA de 27 tumores. Cuarenta y cuatro de estos pacientes eran varones y 23 mujeres, con edades comprendidas entre 1 y 84 años (mediana de edad = 15 años). La localización de los tumores primarios fue en hueso (n=47) y tejido blando (n=19). Diecisiete pacientes presentaron enfermedad diseminada al diagnóstico. Éste incluyó la evaluación histológica por parte de varios patólogos expertos en sarcomas siguiendo los criterios expuestos en la clasificación de tumores óseos y de partes blandas de la OMS, y la detección de la fusión EWS-ETS mediante FISH utilizando la sonda “break-apart” de EWS (Vysis), y/o RT-PCR¹⁰⁹. Se obtuvo la aprobación del los Comités Éticos de cada institución previamente al estudio.

El tiempo de seguimiento clínico fue de al menos 5 años o hasta la recaída.

Respecto al tratamiento recibido por los pacientes, aquellos cuyas muestras fueron recogidas en la HHU se adscribieron a los ensayos Euro-E.W.I.N.G. 99 (n=21)³⁰ o a los ensayos CESS'86³², EICESS'92³¹, o CWS'96³³ (el resto). Todos los pacientes recibieron tratamiento local consistente en resección y/o radioterapia.

Respecto a los pacientes cuyas muestras fueron colectadas por la UCL, 9 de ellos recibieron quimioterapia consistente en Vincristina, Dactinomicina/Actinomicina A y Doxorubicina/Adriamicina, suplementada con Ifosfamida o Ciclofosfamida. El resto de pacientes fueron incluidos en los protocolos EICESS'92 (n=8) o en el Euro-E.W.I.N.G. (n=23). Todos recibieron resección y/o radioterapia para el control local de la enfermedad, con la excepción de 2 de ellos, debido a la localización del tumor primario o por metástasis al diagnóstico.

Los esquemas de tratamiento de los pacientes de ambas instituciones son por lo tanto similares.

Se comprobó la correlación de los parámetros clínicos clásicos (localización del tumor primario, tamaño, presencia de metástasis o recaída local, etc.) con la supervivencia, para confirmar que eran los habituales de esta entidad tumoral.

La serie tumoral (n=38) que compone al grupo 2 de tumores, utilizada en los análisis transcriptómicos y en los TMA, ha sido ya publicada y validada por la institución Instituto Ortopédico Rizzoli (IOR), de Bolonia¹¹⁰, donde se realizaron la extracción de RNA, los microarrays de expresión y donde se compuso el TMA correspondiente a esta serie.

3.2 Líneas celulares

Las líneas celulares de SE utilizadas fueron A4573, A673, CADO-ES, CHP-100, EW3, RDES, RM82, SKES1, SKNMC, STAET1, STAET2.1, STAET10, TC32, TC71, TTC466, VH64 y WE68, obtenidas a través de la institución HHU de Düsseldorf, que las mantiene y caracteriza¹¹¹ como repositorio central de líneas celulares de la red de investigación europea EuroBoNet. Todas ellas fueron cultivadas en medio RPMI (Invitrogen), al 10-20% de FBS.

La línea celular 293T se empleó para la producción viral en los ensayos de transducción lenti y retroviral, crecida en DMEM (Invitrogen) adicionado con glutamina y 10% de FBS.

3.3 Extracción de DNA

El DNA genómico de las biopsias tumorales se extrajo en las instituciones HHU y UCL. Las muestras fueron seleccionadas por contener al menos un 70% de células tumorales (cuantificado mediante evaluación histológica) o, en aquellas en las que el porcentaje tumoral fue inferior el componente tumoral fue microdisecado. El DNA se extrajo haciendo uso de los kits “Qiagen DNA Extraction kit” (Qiagen) o “High Pure PCR Template Preparation kit” (Roche) siguiendo las recomendaciones de los fabricantes.

3.4 *Hibridación genómica comparativa, BAC-array*

3.4.1 Construcción del BAC-array

El microarray creado comprende una representación de la librería genómica humana desarrollada por el instituto Sanger (“The Wellcome Trust Sanger Institute”, Hinxton, GB), de libre distribución, conocida como “Sanger 1mb clone set”. La librería contiene elementos genómicos de secuencia y posición conocida espaciados 1 mb en el genoma humano, clonados en vectores de gran capacidad conocidos como BAC (cromosoma artificial bacteriano), que admiten como media insertos de unos 150 kb, estando compuesta en total por alrededor de 3500 clones BAC contenidos en un hospedador bacteriano (*E. coli*). El genoma humano no está por tanto representado de forma exhaustiva, sino que existe un inserto de 150 kb clonado en un vector BAC por cada mb de genoma no cubierto por la librería. Aunque esta resolución pueda parecer baja supone una notable mejora respecto a la resolución del mCGH y, además, la librería de 1mb está enriquecida con una sub-librería que comprende cientos de clones BAC específicamente seleccionados para cubrir los *loci* de los principales genes involucrados en cáncer. Este tamaño medio de la librería presenta también muchas otras ventajas como su versatilidad, precio y facilidad de uso.

La replicación de esta librería se realizó en el “Leids Universitair Medisch Centrum” (LUMC, Leiden), haciendo uso del biorrobot allí disponible y partiendo de una colección de minipreps (construcciones BAC ya extraídas del hospedador bacteriano *E. coli*). Esta replicación, mediante técnicas de PCR detalladas en el siguiente apartado, es necesaria debido a que los clones BAC son vectores de copia única cuyo cultivo bacteriano rinde una cantidad de DNA insuficiente para la creación del microarray.

3.4.2 DOP-PCR y Aminolinking PCR

El objetivo de estas técnicas de amplificación es producir una representación de la secuencia humana contenida en las construcciones BAC, obteniendo la suficiente cantidad de producto para construir el microarray. Para ello, las construcciones son amplificadas primero mediante DOP-PCR (“degenerated oligo-primed PCR”), que utiliza oligonucleótidos con secuencia

parcial al azar, y posteriormente se reamplifica este primer producto con una segunda PCR que hace uso de cebadores con modificación química amino en el extremo 5', con lo que se obtiene un producto final de amplificación amino-modificado. El grupo amino añadido se requiere para la unión del producto a la superficie donde se construye el microarray.

Las técnicas de PCR convencionales no son aplicables debido a los grandes tamaños de los insertos contenidos en los vectores BAC. Los cebadores usados en la DOP-PCR han sido además diseñados para diferenciar el inserto, de secuencia genómica humana, de la secuencia del vector BAC¹¹². La estructura y secuencia de los cebadores oligonucleotídicos es la que se explica en el siguiente esquema (secuencias en sentido 5'-3'):

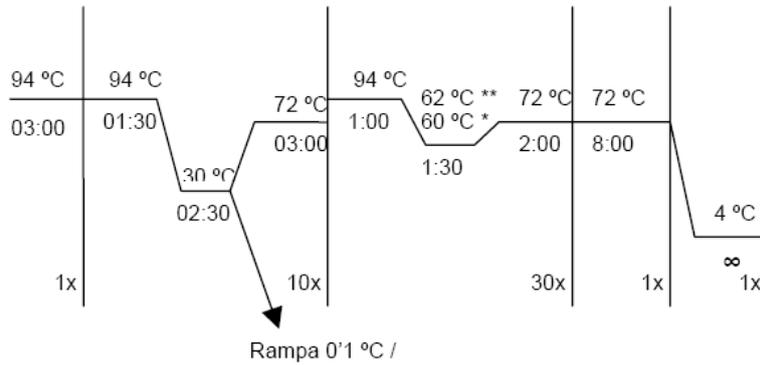


Para maximizar la proporción del inserto humano del clon BAC representado en la mezcla de secuencias amplificadas se realizan 3 DOP-PCRs independientes, cada una de las cuales usa un cebador de secuencia distinta en su porción discriminante del inserto genómico humano.

Se dispensa en una placa de 96 pocillos “Axygen” (Bonsai Technologies) 1 µl del DNA del BAC y se añade por cada pocillo la siguiente mezcla:

Solución TAPS-BSA-βME	5 µl
Cebador DOP (20 µM)	5 µl
dNTPs (10 mM cada uno)	1.25 µl
Solución W-1	2.5 µl
Taq polimerasa (Applied Biosystems)	0.5 µl
H ₂ O	34.75 µl
DNA del clon BAC	1 µl
Total	50 µl

El programa de PCR usado es el siguiente:



* Temperatura de anillamiento para los cebadores DOP1 y DOP3

** Temperatura de anillamiento para el cebador DOP2

Los 10 primeros ciclos de esta PCR están diseñados para amplificación de múltiples secuencias, para lo cual se programa una rampa creciente de temperatura de anillamiento que favorecerá en cada ciclo a un subconjunto distinto de los cebadores con secuencia al azar.

Preparación de soluciones:

Solución TAPS

Soluciones necesarias: MgCl₂ 1M, KOH 2M

Añadir a un vaso de precipitados lo siguiente:

6.08 g de TAPS (N-tris(hydroxymethyl)methyl-3-amino-proanesulphonic acid)

2.20 g de (NH₄)₂SO₄

2.5 ml de MgCl₂ 1M

60 ml de agua milliQ

Ajustar el pH a 9.3 con una solución de KOH 2M. Filtrar con jeringa y filtro de 0.22 µm y esterilizar con ultravioleta. Guardar a -20 °C. Una alícuota de 1 ml será suficiente para hacer un par de placas de 96 pocillos.

A la solución TAPS hay que adicionar en fresco, tras descongelarla, β Mercapto-etanol (βME) y una solución de BSA al 5% en agua (filtrado a través de 0.22 µm y esterilizado con Luz UV). Para 1 ml de TAPS completo:

- 960 µl de solución TAPS + 33 µl de BSA 5% + 7.5 µl de βME*

Solución W-1 (Polioxietileno ether W1, 1%)

0.25 g de W-1 ó Brij 58 (“Polyethylene glycol hexadecyl ether”, antes llamado por SIGMA W-1 y ahora Brij58) en 25 ml de agua milliQ.

Filtrar con jeringa y filtro de 0.22 μm y esterilizar con luz UV. Hacer alícuotas de 1 ml y conservar a $-20\text{ }^{\circ}\text{C}$

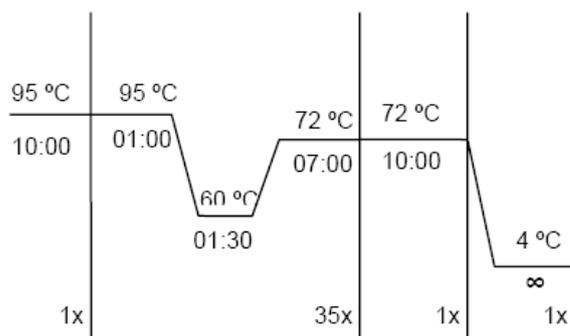
El resultado de la PCR se comprueba corriendo una pequeña porción (2 μl) en un gel de agarosa al 1%. Debe observarse una colección de bandas de entre 200 pb y 2 kb.

Tras esta la DOP-PCR, el producto es reamplificado mediante “**Aminolinking PCR**”, partiendo de 2 μl del producto de DOP-PCR en un volumen total de reacción de 50 μl .

Por cada pocillo de una placa de 96 la reacción comprende:

DNA DOP-PCR (mezcla de los productos DOP1, DOP2 y DOP3)	2 μl
Solución Aminolinking	5 μl
Cebador Aminolinking (20 μM)	4.16 μl
dNTPs (10mM cada uno)	1.25 μl
AmpliTaq polimerasa (Applied Biosystems)	0.5 μl
H ₂ O	37.09 μl
Total	50 μl

El programa de PCR es el siguiente:



De nuevo se comprueba el resultado de la PCR en un gel de agarosa al 1%. El patrón de bandas debe ser semejante al obtenido tras correr la DOP-PCR ya que la aminolinking PCR es tan solo una re-amplificación de los productos de la DOP-PCR (Figura 6).

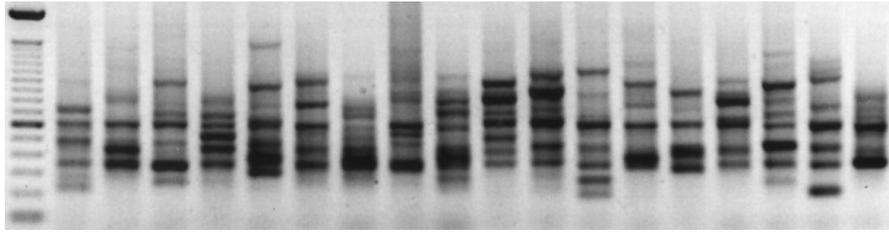


Figura 6. Ejemplo del patrón de bandeo en electroforesis en gel de agarosa de varias Aminolinking PCR (cada una de ellas producto de la reamplificación de los productos de DOP1-3 PCR de un clon BAC diferente), resuelta mediante electroforesis en gel de agarosa (la primera calle corresponde al marcador de pesos moleculares). Los tamaños de los fragmentos generados oscilan entre los 200 pb y las 2 kb.

Soluciones a preparar:

Solución de Aminolinking:

Soluciones requeridas: Tris 1M pH 8.5, MgCl₂ 1M, KCl 3M.

Añadir directamente a un falcon de 15 ml estéril (para 10 ml totales):

0.5 ml de Tris-HCl 1M pH 8.5 (50 mM en la solución final)

0.25 ml de MgCl₂ 1M (25 mM en la solución final)

1.67 ml KCl 3M (500 mM en la solución final)

7.58 ml de agua milliQ autoclavada

Filtrar con jeringa y filtro de 0.22 μ m y esterilizar con luz UV durante 30 min.

El cebador para la reacción de aminolinking PCR debe pedirse a las casas comerciales con modificación amino en el extremo 5', en el carbono C6.

Todos los cebadores una vez recibidos se resuspenden en agua milliQ a una concentración de 100 pmoles/ μ l (100 μ M). A partir de este vial stock se hacen alícuotas de 20 μ M para trabajar con ellas.

3.4.3 Microdeposición del producto de PCR

A 25 μ l de producto de Aminolinking-PCR se le añade 8.3 μ l de solución fosfato 1M pH 8.5 y se pasan los volúmenes de estas PCRs a placas de 384 pocillos, que se mantienen a -20 °C. La

placa de 384 descongelada para ser microdepositada debe precalentarse a 37 °C durante unos 5 minutos y ser agitada suavemente durante otros 5 minutos (en un agitador horizontal).

La microdeposición la realiza el biorrobot “MicroGrid II” (Genomic Solutions). Es necesario ajustar la humedad relativa de la cámara a 45 - 50%. La microdeposición se hace por triplicados de cada PCR (cada representación de un clon BAC) con una separación entre puntos de 0.225 μm , sobre portas “Codelink” (Surmodics). El biorrobot tiene un programa informático asociado (“TAS Application Suite”) que, tras introducirle los identificadores de cada pocillo del microarray (nombre del clon BAC, posición genómica, etc.) crea en un archivo informático una matriz o malla con la información referente a la identidad y posición de cada punto del microarray, que será utilizada posteriormente en el análisis.

La elección de la superficie de microdeposición es de gran importancia, condicionando dramáticamente la calidad del resultado final del experimento de aCGH que por lo demás casi no difiere de un protocolo a otro. Anteriormente se usaban distintas químicas de tratamiento de la superficie del porta, pero hoy el protocolo que aquí se describe es el más aceptado puesto que la calidad del resultado es claramente superior¹¹³. La superficie de los portas “Codelink” está recubierta por una micromatriz de poliacrilamida que contiene grupos químicos que reaccionan con el extremo amino del producto de la aminolinking PCR. En consecuencia y a diferencia del resto de superficies que se usaban inicialmente (recubiertas por grupos epoxi o poly-lisina):

- El DNA no queda directamente unido al vidrio sino que yace en una matriz tridimensional, lo cual mejora su accesibilidad para la hibridación e impide la formación de estructuras secundarias.
- El DNA depositado se une sólo por su extremo 5' y no por cualquier punto intermedio de su secuencia, lo cual también redundará en una mayor cantidad de secuencia disponible para la hibridación.

Tras la microdeposición los microarrays creados se mantienen en una atmósfera de humedad saturada al 75%, durante 20-24 horas (en un contenedor hermético con un centímetro de agua y sal en el fondo). De esta manera las “nanogotas” depositadas no se evaporan durante este tiempo, lo cual es de vital importancia para que tenga lugar la reacción química de unión del grupo amino del producto de PCR con los grupos químicos de la superficie del porta.

Una vez pasado este tiempo los microarrays se sumergen en una solución al 1% de Amoniaco en agitación suave (100 rpm) durante 5 minutos. De este modo se bloquean los grupos químicos que no han reaccionado con el DNA depositado. Tras ello se lavan 2-3 veces los portas con agua destilada y se sumergen en una solución de 0.1% SDS en agua, en agitación suave (100 rpm) durante 5 minutos.

De ahí los portas se pasan directamente a agua milliQ hirviendo, lo que causa que las hebras de DNA se separen, obteniéndose un microarray mayoritariamente compuesto de cadenas de DNA simples, con una mayor capacidad de hibridación. Tras un par de lavados en agua los portas se secan por centrifugación y se almacenan en un desecador (la ausencia de humedad es muy importante para que alcancen una durabilidad de hasta 1 año), listos para ser usados.

Solución a preparar:

Tampón fosfato 1M, 0.001% sarcosyl (100 ml)

Pesar 35.814 g de Sodio fosfato monobásico monohidrato ($\text{Na}_2\text{HPO}_4 \times \text{H}_2\text{O}$).

Añadir 90 ml de agua milliQ y agitar con calor suave (termoagitador) hasta que disuelva.

Preparar otra solución en 10 ml totales, en un falcon, para el ajuste de pH:

1.38 g di-Sodio hidrogenofosfato dodecahidrato ($\text{Na}_2\text{H}_2\text{P}_2\text{O}_7 \times 12 \text{H}_2\text{O}$)

Ajustar a pH 8.5 la primera solución usando la segunda. Enrasar a 100 ml totales con agua milliQ y añadir 3.3 μl de sarcosyl (“N-Lauroylsarcosine Sodium Salt”) al 30% (concentración final 0.001%). Filtrar con filtro de 0.22 μm para esterilizar.

3.4.4 Marcaje del DNA

Se parte de 500 ng de DNA, tanto de la muestra tumoral como referencia. El DNA debe presentar una ratio mínima 260/280 de 1.70 y una ratio 260/230 de al menos 1.6.

El DNA tumoral es marcado con el fluorocromo Cy5 (dCTP-Cy5, Amersham, GE) mientras que el DNA de referencia, consistente en un pool de 100 DNA sanos obtenidos del Banco Nacional de DNA tras aprobación de su Comités Ético y Científico Externos, se marca con Cy3 (dCTP-Cy3, Amersham, GE).

El marcaje consiste en una reacción tipo “Random Priming” para lo cual se utilizan oligonucleótidos de secuencia al azar (“Random Octamers”) resuspendidos en el tampón de reacción de la enzima DNA polimerasa de marcaje, “Exo-Minus Klenow” (Eppicentre), a una concentración final de 0.75 µg/µl. El tampón de reacción se diluye previamente 4 veces (es un tampón 10x) hasta una concentración final de 2.5x.

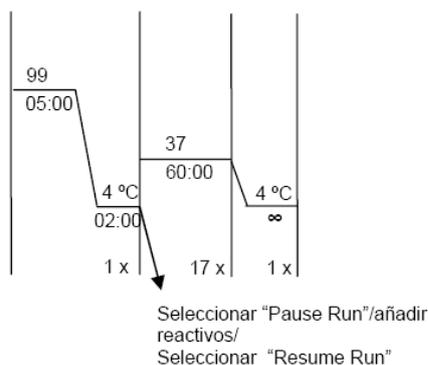
La reacción de marcaje se compone de:

Agua	21 – x µl
DNA (x = volumen correspondiente a 500 ng)	x µl
NaCl 1M	1 µl
“Random Octamers” en solución de reacción 2.5x	20 µl

Esta primera reacción se incuba en termociclador hasta el punto señalado en el esquema detallado más abajo (se somete a una desnaturalización inicial). Entonces se retiran los tubos del termociclador y, en hielo, se añade a los tubos lo siguiente.:

Solución 10x dCTPmix	5 µl
dCTP-Cy3 o dCTP-Cy5	2 µl
Enzima Exo-Minus Klenow	1 µl

Programa de marcaje para el termociclador (iCycler, BIO-RAD):



Tras esta incubación inicial, se continúa el programa de marcaje que incluye una incubación durante la noche.

La solución 10x dCTPmix, es una mezcla de nucleótidos con la mitad de concentración relativa de dCTP, lo cual favorece la incorporación del dCTP-fluorocromo. La concentración final de cada nucleótido trifosfato debe ser de 2 mM, salvo del dCTP (1 mM).

Tras el período de incubación se detiene la reacción añadiendo 5 μ l de tampón de parada (una solución de EDTA 0.5 M pH 8.0).

El DNA es luego purificado mediante el kit de columnas “G-50 Microspin Columns” (Amersham, GE), un sistema de cromatografía de exclusión que elimina de la solución los dCTP-Cy3/Cy5 no incorporados al DNA.

Las columnas son primero preparadas sometiéndose a vórtex y a una centrifugación a 735 g durante 1 minuto, para eliminar el tampón de equilibrado. Tras ello se aplica la muestra sobre la resina, con la columna situada sobre un microtubo opaco (para proteger al DNA finalmente marcado de la luz) y se centrifuga de nuevo durante 2 minutos a 735 g.

3.4.5 Hibridación

Inicialmente se precipita todo el DNA marcado y purificado (tumoral y referencia) junto con 100 μ l de Cot-1 humano (Invitrogen), una preparación de DNA genómico humano enriquecida en secuencias repetitivas que neutralizará el efecto de estas en la hibridación. En paralelo se precipita en otro microtubo una mezcla de 80 μ l de espermatozoides de arenque sonicado y de cadena única (SIGMA) junto con 80 μ l de Cot-1. Esta mezcla, la solución de prehibridación, se usará en el pretratamiento del microarray como bloqueo del ruido de fondo y de las secuencias repetitivas de éste. Se añade 1/10 del volumen de la mezcla de Acetato Na 3M pH 5.2 y 2.5 volúmenes de etanol 100% frío a cada tubo, se incuban a -80 °C durante 20 minutos y se centrifugan a 4 °C a máxima velocidad.

Tras eliminar el sobrenadante, el pellet de la mezcla de DNA se resuspende en 84 μ l de solución de hibridación más 6 μ l de tRNA de levadura (100 μ g/ μ l, Invitrogen, un “DNA carrier”). El tubo de pretratamiento se resuspende en 90 μ l de solución de hibridación. Se incuban ambos tubos a 70 °C durante 15 minutos, tras lo cual se resuspende el pellet por vórtex.

Mientras el tubo con la mezcla de DNA se incubaba a 37 °C durante 1 hora y 30 minutos, la solución de prehibridación se aplica sobre el microarray bajo un cubre de microscopía de 24x60 mm. El porta así preparado se introduce en la cámara de hibridación humidificada: una

caja de portas marca Kartell (para 25 portas) con un papel empapado en solución humidificadora. La cámara se incubaba a 37 °C en la estufa.

Transcurrida la incubación, se extrae el porta, se retira el cubre junto con la solución de prehibridación, y se aplica la solución de hibridación con los DNA marcados bajo un nuevo cubre de 24x60 mm. Se introduce en la cámara y esta en la estufa a 42 °C. La hibridación se incubaba durante 48 horas.

Tras ese tiempo el microarray es lavado para eliminar el exceso de sonda no unido. En una cubeta de vidrio se añaden 130 ml de solución de lavado 2 y se agita 10 minutos en un agitador horizontal a 250 rpm. Se pasa luego a una segunda cubeta, con 130 ml de solución de lavado 1 precalentada, que se agita esta vez durante 30 minutos a 48 °C, tras lo cual se repite el primer lavado en solución 2. Por último se secan los portas por centrifugación (1500 – 2000 rpm), listos para ser escaneados (Figuras 7 y 8).

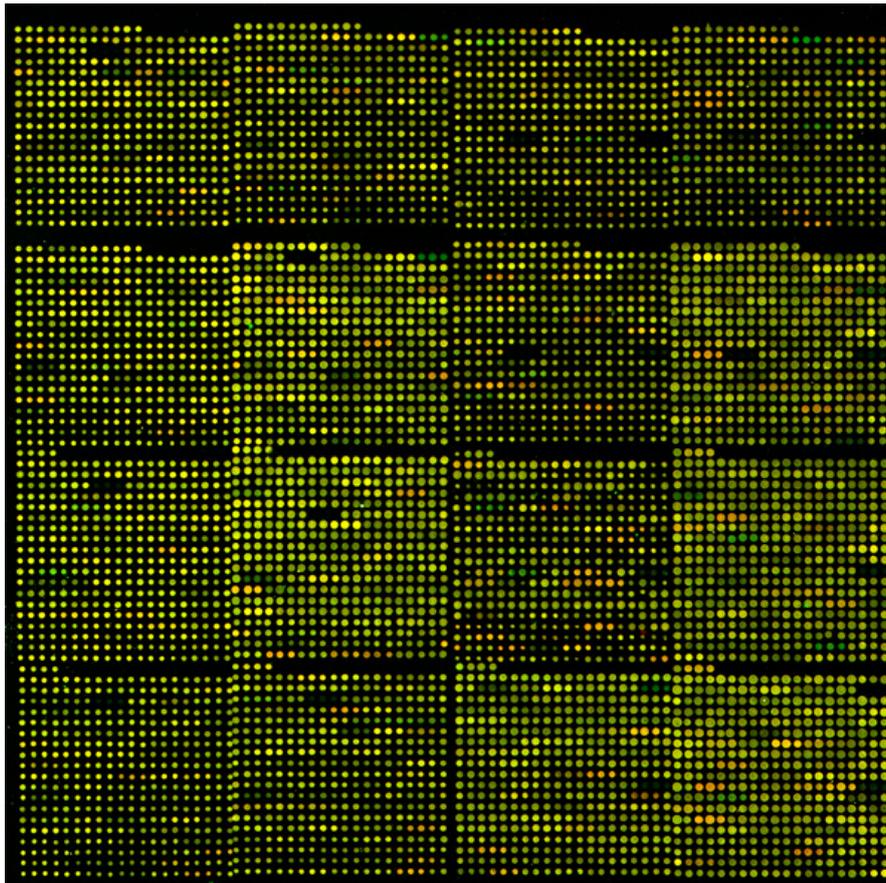


Figura 7. Ejemplo de una imagen adquirida tras el proceso de escaneado del BAC-array de la librería “Sanger 1 mb clone set”, hibridado con el DNA marcado de una línea celular de SE (CADO-ES). La imagen corresponde a la mitad del microarray completo (compuesto por otro cuadrante similar de 16 bloques de puntos).

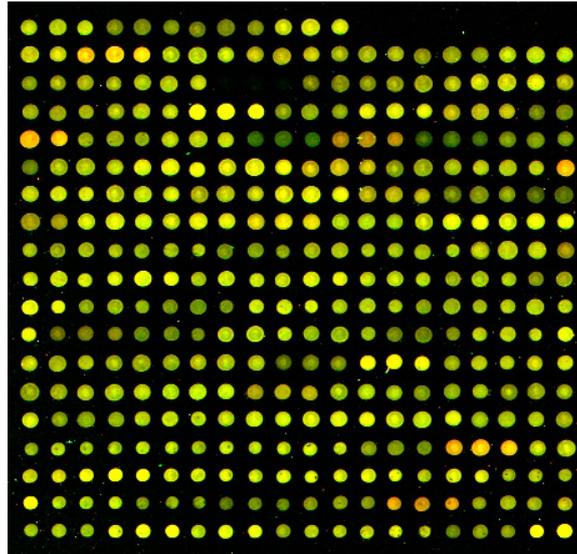


Figura 8. Ampliación del bloque 1 de la figura anterior (bloque superior izquierdo)

Soluciones a preparar:

Solución de hibridación (50 ml)

25 ml Formamida (50% Formamida)

5 g Dextrán Sulfato (10% Dextrán sulfato, SIGMA)

50 μ l Tween 20 (0.1% Tween 20)

5 ml SSC 20x (2x SSC)

500 μ l Tris 1M pH 7.4 (10 mM Tris pH 7.4)

Solución de lavado 1 (130 ml)

65 ml Formamida (50% Formamida)

13 ml SSC 20x (2x SSC)

52 ml H₂O milliQ limpia

Ajustar pH = 7.

Solución de lavado 2 (350 ml)

350 ml PBS

350 μ l Tween 20 (0.1% Tween 20)

Solución humidificadora (100 ml)

40 ml Formamida (20% Formamida)

10 ml 20xSSC (2xSSC)

50 ml H₂O**3.5 *Análisis de los datos de los aCGH BAC-array***

Los microarrays hibridados son escaneados en un aparato “Axon 3000B” (Axon laboratories), que obtiene las intensidades de señal de cada canal de fluorescencia y exporta multitud de datos numéricos, entre ellos el valor log₂ratio (ratio de las señales logaritmizada en base 2). El primer procesado de los datos crudos se hace con el programa informático “GenePix” (Molecular Devices): cada microarray es normalizado usando la media de la mediana de los ratios de todos los clones correspondientes a secuencias de cromosomas autosómicos y sin incluir los puntos marcados como de baja calidad, para lo cual se han establecido unos criterios de intensidad mínima de señal (“flagging scripts” implementados en GenePix). En aquellos casos en que este método de normalización es claramente incapaz de centrar correctamente la gráfica (lo cual pasa cuando se obtiene una gráfica especialmente alterada) se recurre a renormalizar los datos usando la mediana del valor log₂ratio de las regiones no alteradas, tomando como referencia el estado de los cromosomas X e Y.

Los archivos “.gpr” producidos por Genepix son posteriormente sometidos a análisis bioestadístico y bioinformático en R¹¹⁴, haciendo uso de los paquetes “snapCGH”¹¹⁵ y “CGHcall”¹¹⁶, obtenidos del repositorio bioinformático “Bioconductor”¹¹⁷. El paquete “CGHcall” permite realizar el siguiente procesamiento de los datos:

- Segmentación de los datos, reduciendo su complejidad de modo que se establecen fragmentos de genoma con un mismo estado de número de copia. Aunque en el paquete se implementan varios algoritmos de segmentación, en este trabajo se ha usado “DNAcopy”, basado en el algoritmo de segmentación binaria circular¹¹⁸, puesto que, a pesar de que los algoritmos basados en modelos de markov oculto (“hidden markov model”, HMM) están más extendidos, se ha demostrado que la segmentación binaria circular tiene un mejor rendimiento¹¹⁹.

- Proceso de “calling”, mediante el cual se atribuye a los segmentos genómicos establecidos por el algoritmo de segmentación un estado de número de copia respecto al estado 0 (sin alteración de número de copia). Para ello el programa realiza comparaciones en los datos intra e inter muestra. Permite además introducir un valor de porcentaje de celularidad tumoral para cada muestra.

Para un procesamiento posterior de los datos se utilizan los siguientes paquetes de Bioconductor:

- “CGHregions”¹²⁰, que realiza una reducción de la complejidad de los datos al eliminar los valores de los clones intermedios de cada segmento, manteniendo el valor final de estado de número de copias del segmento y su posición genómica (del primer y último clon BAC del segmento).

- “WECCA”¹²¹, paquete especialmente diseñado para estudios de clusterización jerárquica a partir de los datos de los estudios de ANC, que contempla las peculiaridades de este tipo de datos (los valores procesados por el algoritmo de “calling” se reducen a pérdida, ganancia, amplificación, lo cual no sucede con datos de expresión, que presentan valores continuos).

- “CGHtest”¹²², utilizado en las correlaciones con los datos clínicos. Incluye los “scripts” “CGHPermutations”, que permite encontrar las alteraciones genómicas de número de copia distintivas de un subgrupo clínico (comparaciones pareadas o múltiples, hace uso de pruebas no paramétricas), y “CGHlogrank”, que establece el impacto en supervivencia de cualquier alteración detectada, mediante un estudio Kaplan-Meier log rango. Aparte de los test estadísticos citados, ambos “scripts” establecen una tasa de falsos positivos (“false discovery rate”, FDR) mediante miles de permutaciones de los elementos respecto a los grupos definidos por el usuario.

- El “script” “CGH altered genome”, escrito por el doctorando en lenguaje R, permite valorar el tamaño global del genoma afectado por alteraciones genómicas de número de copia partiendo del archivo generado por el paquete “CGHcall”. El “script” comienza numerando ordinalmente cada segmento definido por el algoritmo “DNAcopy”, cromosoma por cromosoma, y calculando su tamaño (la posición genómica del último BAC del segmento menos la posición del primer BAC). Calcula después el tamaño genómico total que conforman todos los segmentos a los que el algoritmo de “calling” ha atribuido un estado de

número de copia alterado (pérdidas y ganancias), así como el número total de estos segmentos, parámetro al que llamaremos número total de alteraciones de número de copia (NTA). Finalmente, calcula el porcentaje que el tamaño total de genoma alterado representa, en cada muestra, respecto al tamaño total de genoma representado en el BAC-array, parámetro al que llamaremos porcentaje de genoma alterado (PGA).

Para valorar el impacto de estos dos parámetros en los valores de supervivencia de los pacientes se establecieron 6 intervalos, correspondientes a la división de los datos mediante 5 percentiles poblacionales, y se realizó un estudio de Kaplan-Meier log rango con los datos así divididos. Los grupos contiguos que no demostraron diferencias en supervivencia se unieron en un mismo grupo.

Los datos de microarrays producidos en esta investigación (tanto de los BAC-array como de los SNP-array y microarrays de expresión) han sido depositados en el repositorio público del NCBI, conocido como “Gene Expression Omnibus” (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) y son accesibles con el identificador GSE20368.

3.6 Hibridación de los SNP-array, modelo “GeneChip Human Mapping 500K Array Set” (Affymetrix)

Realizado por el Servicio de Genómica del CIC siguiendo el procedimiento a continuación resumido:

3.6.1 Preparación de los DNA genómicos

Se preparan 12 µl de cada muestra a una concentración de 50 ng/µl en TE-EDTA y se hacen alícuotas de 5 µl.

El DNA se somete a una digestión inicial. Cada muestra es digerida por separado con las enzimas de restricción NspI y Sty (NEB) y procesada por tanto por duplicado. Se añade a cada alícuota de 5 µl la siguiente mezcla:

Tampón de Restricción	2 µl
BSA 10x	2 µl
enzima restricción (NspI o Sty)	1 µl
agua	9.75 µl

Los tubos se incuban en el termociclador (“GeneAmp PCR System 9700”, Applied Biosystems) a 37 °C durante 2 horas, tras lo cual se inactivan las reacciones con una incubación final de 20 minutos a 65 °C.

Tras ello al DNA digerido se le acoplan unos “linkers” o adaptadores, mediante ligación, según la siguiente reacción:

Adaptador (específico para la NspI o Sty)	0.75 µl
Tampón 10 x de la T4 DNA ligasa	2.5 µl
T4 DNA ligasa (NEB)	2 µl

De esta mezcla se añaden 5.25 µl a los 20 µl totales de digestión del paso anterior y se incuba la reacción en el termociclador a 16 °C, 3 horas, seguido de un paso de inactivación enzimática a 70 °C durante 20 minutos. Tras la ligación se añaden 75 µl de agua para diluir el producto.

3.6.2 Amplificación del DNA mediante PCR y procesamiento del producto de PCR

Por cada dilución hecha en el apartado anterior se prepararán 3 reacciones de PCR además de un control negativo, para lo cual se reparte en 3 alícuotas de 10 µl (al tubo del control negativo se le añaden 10 µl de agua). Esta PCR se basa en el uso de cebadores con secuencia homóloga a la de los adaptadores acoplados en el paso anterior. A cada tubo se le añade la siguiente mezcla:

Agua	39.5 µl
Tampón 10x Taq PCR Titanium	10 µl
GC Melt 5M	20 µl
dNTPs (2.5 mM cada)	14 µl
PCR cebador 002 100µM	4.5 µl
Titanium Taq DNA polimerasa 50x	2 µl

Tras añadir los 90 μl de esta mezcla a cada muestra (y al control negativo) los tubos se llevan al termociclador y se aplica el siguiente programa de amplificación:

94°C	94°C	60°C	68°C	68°C	4°C
3min	30s	45s	15s	7min	∞
x30					

Al producto de PCR se le añaden 8 μl de EDTA 0.1 M tras lo cual las 3 PCR pertenecientes a una misma muestra y procesadas inicialmente con una misma enzima de restricción se mezclan. A continuación se procede a purificar las muestras con un sistema de vacío (“QIAvac multiwell unit”, Qiagen) adaptado al procesamiento de las muestras en placas de 96 pocillos. El volumen de las reacciones de PCR es eliminado por el vacío (el DNA procesado queda retenido) de 600 mbar durante 1.5-2 horas. Se realizan luego 2 lavados (con vacío de 15-20 minutos) tras añadir 50 μl de agua en cada lavado. En un tercer lavado similar el vacío se deja durante 45-70 minutos. Finalmente, con el vacío ya desconectado, se recoge el DNA retenido en los pocillos añadiendo 45 μl de solución RB y agitando las placas 10-15 minutos a 500 rpm. El volumen de cada pocillo se trasvasa a un microtubo.

Tras ello, el producto de PCR purificado se cuantifica utilizando un espectrofotómetro “ND1000” (Nanodrop) y se diluye a una concentración final de 2 $\mu\text{g}/\mu\text{l}$.

Finalmente, el producto de PCR es fragmentado, para lo cual se utilizan 90 μg totales de cada muestra y enzima de restricción (45 μl), para lo que se prepara la siguiente mezcla:

Agua	105 μl
Tampón de Fragmentación 10x	12 μl
Reactivo de Fragmentación (2U/ μl)	3 μl

A cada muestra se añaden 5 μl de Tampón de Fragmentación 10x y 5 μl de la mezcla anterior. Tras ello la reacción se incuba en el termociclador a 37 °C durante 35 minutos, con un paso final a 95 °C, 15 minutos.

3.6.3 Marcaje del producto de PCR

Por cada muestra y procesamiento con cada enzima de restricción, se añade por tubo la siguiente mezcla:

Tampón TdT 5x	14 μ l
“GeneChip DNA Labelling Reagent”	2 μ l
TdT 30 U/ μ l	3.5 μ l

La reacción de marcaje se incubaba en termociclador a 37 °C durante 4 horas, tras lo que se inactiva a 95 °C durante 15 minutos. Tras este paso, el producto de PCR ha quedado marcado con biotina.

3.6.4 Hibridación del array, lavados y escaneo

A los 70 μ l de muestra marcada del paso anterior se le añade la siguiente mezcla:

Solución MES 12X (Sigma)	12 μ l
Solución Denhardt 50x (Sigma)	13 μ l
EDTA 0.5M	3 μ l
DNA de Esperma de arenque (Promega)	3 μ l
OCR 100	2 μ l
DNA Cot-1 humano (1mg/ml) (Invitrogen)	3 μ l
Tween 20, 3%	1 μ l
DMSO 100%	13 μ l
TMACL 5M	140 μ l

Una vez añadida la mezcla se incubaba la muestra durante 10 minutos a 99 °C y 1 minuto a 49 °C, tras lo que se aplica al microarray (200 μ l por microarray, un porta específico para el DNA procesado con cada enzima de restricción) y se incuban 18 horas a 49 °C y 60 rpm en el horno de hibridación de Affymetrix (“Hybridization Oven 640”). Los modelos de microarray utilizados fueron “GeneChip Human Mapping 250K Nsp Arrays” y “GeneChip Human Mapping 250K Sty”, de Affymetrix.

Tras la hibridación se retira la solución de hibridación de los microarrays y se inyectan a la cámara de hibridación 300 μ l de solución “array holding” (4.15 ml MES 12X + 9.25 ml NaCl 5M + 50 μ l 10% Tween 20 + 36.55 ml agua). Se preparan las reacciones de tinción (por cada microarray hibridado) descritas a continuación.

- STAIN BUFFER:

Agua	800.04 μ l
SSPE 20x (Accugene)	360 μ l
Tween 20, 3%	3.96 μ l
Solución de Denhardt 50x	24 μ l

- SOLUCIÓN SAPE (tubo 1), preparar en un microtubo ámbar:

Solución “Stain buffer”	594 μ l
Estreptavidina-Ficoeritrina 1mg/ml (Invitrogen)	6 μ l

- SOLUCIÓN DEL ANTICUERPO (tubo 2):

Solución “Stain buffer”	594 μ l
anticuerpo biotinilado anti-estreptavidina (0.5mg/ml) (Vector Labs)	6 μ l

- SOLUCIÓN “ARRAY HOLDING BUFFER” (tubo 3):

Trasvasar 820 μ l de solución “array holding buffer” a un microtubo

Se colocan los microarrays y las tres soluciones en la estación “Fluidics Station 450” de Affymetrix y se aplica el programa de lavado indicado para el modelo del microarray (“GeneChip Human Mapping 500K Array Set”). Con los lavados la biotina con que se marcó el producto de PCR es reconocida por la avidina-ficoeritrina y además la señal es amplificada con un paso en el que se le une un anticuerpo anti-estreptavidina, también biotinilado.

Una vez finalizado el lavado se escanean los microarrays utilizando el “GeneChip Scanner 3000” de Affymetrix, que excita al fluorocromo ficoeritrina y recoge la señal que genera. En el proceso de escaneo se producen automáticamente los ficheros “.cel” y “.dat”. Además, el programa informático “GTYPE” de Affymetrix genera un informe final (fichero “.rpt”).

3.7 *Análisis de los datos de los aCGH SNP-array*

Para su análisis se hizo uso de programas informáticos desarrollados por Affymetrix: “GTYPE 4.1” y “CNAT 4.0”. Los archivos “.cel” generados se procesaron para el estudio de número de copias, con “smoothing” 0.1 y usando la colección de datos “Mapping 500K

Sample Data Set” generada por el proyecto HapMap¹²³ como referencias de DNA sano. Esta colección está compuesta de los archivos “.cel” producto de la hibridación con estos microarrays de 48 muestras sanas (96 hibridaciones teniendo en cuenta tanto el microarray NspI como el Sty). El programa informático utiliza un algoritmo de segmentación basado en modelos de markov oculto (HMM) y también realiza el “calling” de los datos.

3.8 COBRA-FISH

Esta técnica, acrónimo de “Combined Binary Ratio Labelling – Fluorescence In Situ hybridization”, permite distinguir cada uno de los cromosomas del genoma humano (incluso sus respectivos brazos cromosómicos) con un color de emisión de fluorescencia combinada distinto, haciendo uso de una colección de sondas marcadas con tan sólo 4 fluorocromos. Para conseguir un conjunto de sondas de este tipo se procede primero separando cada uno de los cromosomas humanos mediante citometría de flujo o microdissección. Tras ello cada cromosoma es amplificado por separado mediante la técnica de DOP-PCR de una manera idéntica a la explicada en el apartado de aCGH BAC-array, pero esta vez encadenando 4 DOP-PCRs consecutivas para generar grandes cantidades de producto de PCR. Dicho producto es finalmente marcado con fluorocromos mediante el método de “ULS” (“Universal Linkage System”, Kreatech), que utiliza los fluorocromos habitualmente usados en FISH pero acoplados a un derivado del cisplatino que permite su unión química directa al DNA, sin necesidad de reacciones enzimáticas. Para la obtención de tal cantidad de “colores” (24 en el método más común, pero se pueden conseguir incluso 48) a partir de tan pocos fluorocromos se aplica una combinación de “ratio labeling” (distintos niveles relativos de 3 fluorocromos que permiten una discriminación inicial de 12 cromosomas) y el “binary labeling” (la adición de un cuarto fluorocromo a un subgrupo de la mitad de los cromosomas marcados mediante “ratio labeling”, lo cual permite identificar 24 elementos en total, suficiente para cubrir el cariotipo humano completo).

Para una detallada descripción del protocolo de generación de las sondas marcadas, se puede consultar la siguiente referencia¹²⁴: Szuhai K *et al.* (2006). COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence *in situ* hybridization karyotyping. Nature Protocols 1, 264 – 275.

El procedimiento de hibridación con la muestra es muy similar al de cualquier otra técnica de pintado cromosómico:

- Al porta de microscopía con la extensión de metafases se le aplican 100 μ l de RNasa I (100 μ g/ml en 2 \times SSC) bajo un cubre de microscopía de 24 \times 60 mm y se incuba a 37 $^{\circ}$ C durante 10 minutos.

- Se lava 3 veces el porta en 2xSSC, lavados de 5 minutos. Tras el último lavado se añade una solución de pepsina (50 μ l de pepsina en 100 ml de 0.1 M HCl) precalentada y se incuba 5 minutos a 37 $^{\circ}$ C.

- Se lava el porta 2 veces en 1 \times PBS, lavados de 5 minutos a temperatura ambiente, tras lo cual se transfiere a una solución de 1% formaldehído/1 \times PBS (vol/vol) y se incuba 10 minutos a temperatura ambiente. Tras ello se hacen de nuevo 2 lavados en PBS de 5 minutos y se realiza un protocolo típico de deshidratación progresiva (70, 90 y 100% de etanol).

- Se aplican entonces 11 μ l de la preparación de la sonda marcada de COBRA-FISH (preparada como se detalla en la citada referencia en una solución de hibridación similar a la de aCGH BAC-array y en presencia también de Cot-1 humano, tRNA de levadura y DNA de esperma de arenque) y se coloca un cubre de microscopía de 20 \times 20 mm. Tras ello se desnaturaliza el porta y el DNA a 80 $^{\circ}$ C durante 45–90 segundos.

- El porta se pasa a una cámara de hibridación humedecida (similar a la descrita en el protocolo de aCGH BAC-array) y se incuba entre 40 y 72 horas (dependiendo del tipo de muestra) a 37 $^{\circ}$ C.

- Tras la hibridación, el porta se pasa a una solución 2 \times SSC y se incuba 5–10 minutos, tras lo cual se quita el cubre y se hacen 2 lavados de 5 minutos en 2 \times SSC/0.1% Tween 20 (vol/vol) a 37 $^{\circ}$ C, seguidos de otros 2 lavados de 5 minutos en 50% formamida, 2 \times SSC, pH 7.0 (vol/vol) a 44 $^{\circ}$ C. Se repite finalmente el lavado en 2 \times SSC/0.1% Tween 20 (vol/vol) a 37 $^{\circ}$ C y se finaliza con otros 2 lavados de 5 minutos, esta vez en 0.1 \times SSC a 60 $^{\circ}$ C.

- Se pasa el porta hibridado y lavado por una solución 1 \times TBS/0.05% Tween 20 y posteriormente se procede a una típica deshidratación progresiva en etanol (70, 90 y 100%, vol/vol).

- Se aplican 15 μ l de solución de montaje y “antifading” (que maximiza la duración de la actividad de los fluorocromos) “Citifluor AF1/DAPI” (400 ng/ml, casa comercial Citifluor) y se pone un cubre de 24 \times 60 mm.

- Para el proceso de adquisición de la imagen se requiere un microscopio de fluorescencia equipado con una cámara digital tipo CCD y los filtros fluorescentes para la obtención de imágenes de fluorescencia multicolor apropiados: los fluorocromos usados en el marcaje de la sonda son dietil aminometil coumarina (DEAC), cianina 3 (Cy3), cianina 5 (Cy5) y RG (Kreatech). La resolución óptica se realiza con un objetivo de al menos una apertura numérica = 1.30 y de 63 \times - 100 \times de aumento. Se deben adquirir las imágenes en los canales de fluorescencia correspondientes al infrarrojo, rojo lejano, rojo, amarillo, verde, azul y la contratinción de DAPI, por ese orden creciente de longitudes de onda (para evitar la pérdida de señal de los fluorocromos con mayor λ de excitación). Los tiempos de exposición deben ser ajustados para cada caso, oscilando entre 0.5 a 6 segundos. La imagen final es procesada por un programa informático especializado de la casa comercial Applied Imaging.

3.9 Extracción de RNA

La extracción de RNA total de las líneas celulares se realiza mediante el protocolo habitual de extracción con el reactivo “Trizol” (Invitrogen).

En todo momento se evita la confluencia de los cultivos celulares. Se retira el medio de cultivo y, sobre hielo, se hace un lavado con PBS frío. Tras ellos se depegan las células con un raspador en un pequeño volumen de PBS (menor a 1 ml), también sobre hielo. Las células despegadas se someten a centrifugación (1200 rpm, 3 minutos, 4 °C), se descarta el sobrenadante, se resuspenden en 1 ml de “Trizol” y se lisan por pipeteo.

Tras incubar 2 minutos a temperatura ambiente (para que se deshagan los complejos ribonucleoprotéicos) se añaden 200 μ l de cloroformo y se mezcla bien por inversión. Tras otros 2 minutos se centrifuga a 12000 rpm para propiciar la formación de las distintas fases. La fase superior, que contiene el RNA, se separa con cuidado de las demás (que retienen el resto de especies moleculares) y se deposita en un nuevo microtubo.

El RNA es entonces precipitado añadiendo 1 volumen de isopropanol (500 μ l), incubando en hielo 10 minutos y centrifugando posteriormente a máxima velocidad, 4 °C, durante 10 – 20 minutos. Tras este proceso el RNA queda formando un pellet en el fondo. Tras eliminar el sobrenadante y dejar secar lo suficiente (el pellet debe estar seco pero no debe volverse transparente, ya que entonces se hace muy difícil de resuspender) se resuspende en el volumen de agua estéril con DEPC (“DiethylenePyrocarbonate”, Ambion) acorde con el tamaño del pellet.

La concentración y pureza se valoran en el espectrofotómetro “ND1000” (Nanodrop). En caso de que fuera necesario ($260/280 < 1.8$; $260/230 < 1.6$), se purifican con el kit de columnas “RNeasy” (Qiagen), siguiendo las recomendaciones del fabricante. Además, los RNA se resuelven mediante electroforesis en gel de agarosa, comprobándose su integridad de acuerdo con los patrones de bandeo esperados de los rRNA 28S y 18S.

3.10 Microarrays de expresión

Realizados por el servicio de Genómica del CIC de acuerdo con el siguiente protocolo resumido:

Los RNA totales extraídos son primero evaluados en integridad y pureza (ausencia de contaminación de DNA) utilizando el chip “Agilent 2100 Bioanalyzer” (Agilent Technologies), un sistema miniaturizado de electroforesis que permite evaluar cantidades moleculares mínimas de RNA. La relación entre las áreas producidas por los rRNA 28S y 18S debe estar entre 1.8 y 2.1, siendo este parámetro el principal indicador de calidad. Un ratio distinto al mencionado refleja la degradación parcial de la muestra o la presencia de contaminantes.

A continuación se procede a retrotranscribir el RNA usando el kit “One Cycle cDNA Synthesis” (Affymetrix), que hace uso de un protocolo típico de retrotranscripción con oligo dT que incluye la secuencia del promotor de la polimerasa del fago T7. La reacción es purificada mediante fenol cloroformo y precipitación con etanol, incluyendo el reactivo “PLG -Phase Lock Gels” (Eppendorf), que forma una barrera inerte entre las fases acuosa y orgánica y facilita la extracción.

Tras ello el cDNA generado se transcribe *in vitro* en presencia de nucleótidos biotinilados y enzima T7 RNA polimerasa para generar cRNA marcado, para lo cual se utiliza el kit “BioArray RNA labeling” (Enzo), que además consigue una amplificación lineal de entre 40 y 80 veces. Al finalizar la reacción, ésta se somete a una nueva purificación.

El cRNA es entonces fragmentado obteniéndose tamaños de entre 35 y 200 pb, óptimos para la correcta hibridación con el microarray, mediante fragmentación alcalina por adición de 2 µl de tampón de fragmentación 5X (200 mM Tris-acetato pH 8,1, 500 mM KOAc, 150 mM MgOAc) por cada 8µl de cRNA e incubación a 94 °C durante 35 minutos. El resultado se comprueba mediante electroforesis en gel de agarosa.

Finalmente el cRNA fragmentado, junto con los controles internos de la transcripción, hibridación y marcaje, se añaden al microarray de Affymetrix “GeneChip® Human Genome U133 Plus 2.0 Array” y se incuban durante 16 horas a 45 °C con rotación constante a 0.4 g en el horno de hibridación “GeneChip Hybridization Oven 640” (Affymetrix). Tras ello se realiza la tinción con estreptavidina-ficoeritrina (SAPE) en la unidad de procesado “GeneChip Fluidics Station” (Affymetrix). Esta unidad se encarga igualmente de los lavados finales.

La unidad de escaneo “GeneArray” (Affymetrix) adquiere la señal de fluorescencia y digitaliza los resultados. El programa informático asociado “GCOS” (“Gene Chip Operating System”, Affymetrix) procesa inicialmente los datos, calculando los valores de calidad de la hibridación (entre otras la comparabilidad de las muestras hibridadas y la longitud de los mRNA hibridados).

3.11 Análisis de los datos de los Microarrays de Expresión

El análisis se realizó en R con los paquetes de Bioconductor “affy”¹²⁵ y “siggenes”¹²⁶.

Inicialmente, a los datos se les sustrajo el ruido de fondo y se normalizaron por cuartiles (a nivel de sonda) mediante el algoritmo “robust multi-array average” (RMA)^{125,127}. La coincidencia inter-chip de las distribuciones de las intensidades logaritmizadas se comprobó mediante la generación de histogramas, “boxplots” y “MVA-plots”.

Los estudios de expresión diferencial se llevaron a cabo aplicando el algoritmo “significance analysis of microarrays” (SAM)¹²⁶, que calcula un valor de diferencia de expresión (“d-value”) de tipo “signal-to-noise” (señal respecto a ruido) en el que se divide la diferencia directa de expresión (de los valores logaritmizados) entre grupos por la dispersión de los niveles intra-grupo, y que estima la FDR mediante la realización de cientos de permutaciones de los elementos respecto a los grupos definidos por el usuario.

Los estudios de clusterización jerárquica se realizaron aplicando el programa informático TIGR MeV4¹²⁸, usando parámetros típicos (“complete/average linkage”).

Para estudiar la coincidencia de los perfiles de expresión diferenciales generados con rutas génicas y funciones celulares se hizo uso del programa informático y análisis GSEA (“gene set enrichment analysis”)¹²⁹, versión 2.0. Este análisis evalúa estadísticamente la sobrerrepresentación o enriquecimiento de colecciones génicas o “gene sets” (generados por el usuario o, más frecuentemente, depositados en las bases de datos “MSigDB”, disponibles “on-line”) en los primeros puestos de la lista del perfil de expresión diferencial del usuario, ordenada de mayor a menor según un valor de expresión diferencial “Signal-to-Noise” (siendo los genes en los primeros puestos de la lista ordenada los más consistentemente sobre-expresados). El grado de enriquecimiento se refleja en el valor ES (“Enrichment Score” o valor de enriquecimiento), derivado de una computación global de los puestos de la lista en que se encontró a los elementos de la colección génica con que se comparó (mientras más bajos sean los puestos, más significativo es el enriquecimiento). En el caso que nos ocupa, la lista ordenada se corresponde con el perfil de expresión diferencial de los tumores 1qGSig (ver apartado 4.3).

El análisis calcula asimismo un valor de FDR (“q-value”) como resultado de la media de 1000 permutaciones de los elementos respecto a los grupos creados por el usuario (tumores 1qGSig frente a resto de tumores, en nuestro caso). Los autores del análisis GSEA aconsejan usar un umbral de FDR del 25% como límite de aceptación de colecciones génicas enriquecidas que podrían tener relevancia biológica. El valor de ES es normalizado con el valor de FDR para crear el parámetro NES (“normalized enrichment score”), el valor más representativo del análisis: aquellas colecciones génicas con elevado ES y bajo FDR tendrán los NES más significativos y serán los resultados más fiables de coincidencia con el perfil diferencial del usuario.

Por otra parte, los criterios seguidos para el establecimiento de las colecciones génicas difieren dependiendo de la base de datos utilizada, por lo que se pueden hacer distintos tipos de análisis según la “MSigDB” seleccionada:

- Base de datos MSigDB C1: los genes se agrupan en colecciones génicas según su posición genómica, por bandas citogenéticas.
- Base de datos MSigDB C2: colecciones génicas creadas a partir de los resultados de investigaciones publicadas. Por lo tanto, esta es una base de datos creada a partir de conocimiento empírico, experimental.
- Base de datos MSigDB C3: las colecciones génicas agrupan a genes que comparten un elemento de respuesta en su promotor o 3' UTR a un cierto factor de transcripción o a una familia de microRNA, respectivamente.
- Base de datos MSigDB C4: los genes se agrupan en las colecciones génicas según sus valores de “Gene Ontology” (términos GO).

Para la comparación del perfil diferencial de los tumores 1qGSig y el perfil obtenido tras el silenciamiento de DTL en las líneas celulares del SE, el procedimiento usado fue el siguiente:

1. Las tablas de valores normalizados por RMA del experimento de silenciamiento *in vitro* se sumaron de manera que de aquellos genes con varios “probesets” se mantuvo el elemento de valor máximo de señal.
2. Se hizo un análisis SAM pareado de expresión diferencial (líneas celulares control respecto a líneas celulares con DTL silenciado).
3. Se creó un “gene set” en formato “.gmt” incluyendo los 320 genes correspondiente al resultado del análisis SAM de delta 2.8 (FDR = 0.001) obtenido enfrentando a los tumores 1qGSig con el resto, al cual se llamó “SAM_38_CASES_MAX_FDR_0.001”.
4. La lista completa del análisis SAM de las líneas celulares silenciadas se ordenó descendientemente por “d-value”, ocupando las primeras posiciones por tanto los genes más consistentemente sobre-expresados en las líneas celulares control respecto a las silenciadas.

5. La lista ordenada se introdujo como “input” en el análisis GSEA, opción “pre-ranked Gene List” (lista pre-ordenada), fijando un nivel de 1000 permutaciones y seleccionando para confrontar tanto la colección génica creada localmente (“SAM_38_CASES_MAX_FDR_0.001”) como las 1892 colecciones génicas que componen la MsigDB C2. El cálculo de las FDR se hizo permutando las colecciones génicas.

Para el análisis de la integración de los genes diferencialmente expresados en los tumores 1qGSig en rutas de señalización conocidas se aplicó el programa informático “Analysis of Ingenuity Pathways” o IPA (Ingenuity® Systems, www.ingenuity.com), introduciendo como “input” la lista diferencial del análisis SAM de los tumores 1qGSig frente al resto, de FDR máximo 0.02.

Por último se recurrió al portal “BioGPS”¹³⁰ (antiguo “GeneAtlas”, <http://biogps.gnf.org>) para obtener los niveles de expresión de mRNA de DTL en tejidos humanos, a partir de las colecciones de datos de microarrays de expresión “HG_U133A/GNF1H y GNF1M Gene Atlas”¹³¹.

3.12 SYBR Green Real Time RT-qPCR

El RNA total extraído y comprobado en concentración, calidad e integridad se retrotranscribe usando cebadores oligo dT para enriquecer el DNA complementario (cDNA) final en secuencias de mRNA, descartando el rRNA, el cual no contiene información codificante y supone la mayor parte del RNA total de la célula.

Se parte de 2 µg de RNA, al que se le adiciona 1 µl de cebadores oligo dT y agua DEPC (Ambion) hasta un total de 11 µl. Todo ello en hielo.

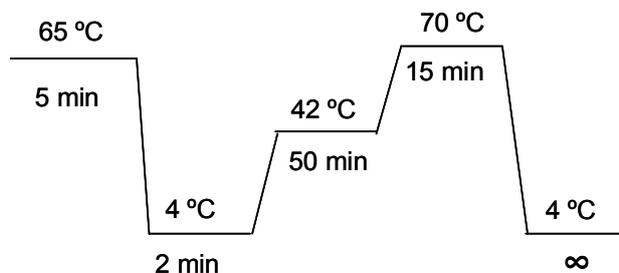
Con esta mezcla los tubos de 0.2 ml se introducen en el termociclador y se someten a un primer paso de desnaturalización de estructuras secundarias, incubándolos durante 5 minutos a 65%, tras lo que se programa un paso de enfriamiento a 4 °C. Llegado este punto se detiene el programa del termociclador y los tubos se sacan y se pasan a hielo.

Se añade entonces por tubo, en hielo, la mezcla de reacción consistente en:

Tampón 5x	4 μ l
0.1M DTT	2 μ l
dNTPs (10 mM cada uno)	1 μ l
Inhibidor de RNasa (Ambion)	1 μ l
Retrotranscriptasa SSII	1 μ l

Estos reactivos se incluyen en el kit de retrotranscripción “SuperScript II Reverse Transcriptase” (Invitrogen).

Al añadir los 9 μ l de este mix se consigue un volumen final de reacción de 20 μ l. Se introducen de nuevo los tubos en el termociclador y se prosigue el programa según el siguiente esquema:



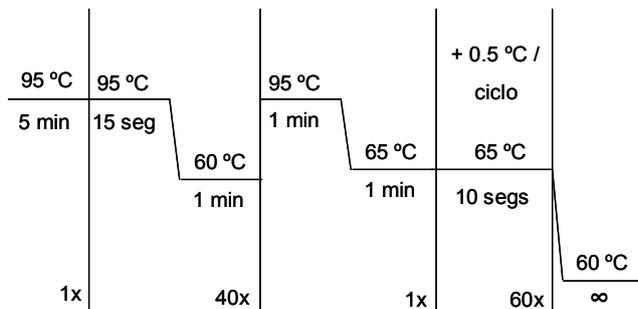
La reacción tiene lugar a 42 °C durante 50 minutos, seguido de un paso de inactivación enzimática a 70 °C. El volumen se lleva a un total de 100 μ l finales, quedando el cDNA a una concentración final de 20 ng/ μ l (se asume una eficiencia de retrotranscripción próxima al 100%).

El cDNA de una cadena se somete al ensayo de Real Time qPCR para lo cual se usan 20 ng (1 μ l) y se utiliza el kit “iQ SYBR Green Mix” (Bio-Rad). La amplificación se lleva a cabo en formato de placa de 96 pocillos en el termociclador óptico “iQ5” (Bio-Rad). Se incluyen en el experimento triplicados de cada condición y controles negativos (cDNA sustituido por agua) así como una curva patrón de calibrado de la eficiencia de los cebadores, con 3 puntos de una dilución seriada de 20 ng, 1 ng y 0.1 ng de cDNA (preferiblemente de una muestra con altos niveles de expresión del gen a testar).

La mezcla de reacción por gen y por cada pocillo de placa de 96 pocillos se describe en la siguiente tabla.

cDNA	1 μ l
“iQ SYBR Green Mix”	12.5 μ l
Cebador Fw (10 mM)	0.4 μ l
Cebador Rv (10 mM)	1 μ l
agua milliQ	10.7 μ l
Total	25 μ l

Programa del termociclador:



PCR: durante los 40 ciclos de PCR, la cámara captura el nivel de fluorescencia de cada pocillo en la fase de anillamiento-extensión.

Curva de disociación: durante los 60 ciclos de incremento gradual de temperatura (0.5 °C cada ciclo) la cámara captura la fluorescencia por pocillo. Conforme aumenta la temperatura el producto de PCR se desnaturaliza y la fluorescencia decrece (puesto que el reactivo “SYBR green” se intercala entre los pares de bases nitrogenadas del DNA de cadena doble). La representación de la derivada de la fluorescencia respecto al tiempo que se genera (una curva tipo campana) es una firma característica del amplicón, que depende no sólo de su composición de GC y su tamaño sino también de la propia secuencia, por lo que permite detectar amplificaciones inespecíficas (aquellas que no coincidan con la del control positivo).

El análisis de los datos se realiza con el programa informático “BioRad IQ5 2.0 Standard Edition Optical System Software” (Bio-Rad), basado en la comparación del ciclo de amplificación en que cada muestra alcanza un nivel de fluorescencia arbitrario conocido como “Threshold Cycle” (CT, ciclo umbral): las muestras que antes lleguen a ese umbral serán las que tengan una mayor expresión del transcrito. Sin embargo, hay que descartar los efectos debidos a errores de pipeteo, diferencias en la eficiencia de retrotranscripción o a niveles globales de transcripción mayores en unas muestras respecto a otras. Para ello se normaliza el CT de cada muestra para el gen problema con el CT de un gen de expresión constitutiva, un

“housekeeping gene”, cuyos niveles se asumen iguales en las distintas muestras comparadas. La fórmula matemática de esta normalización es compleja¹³².

Los cebadores oligonucleotídicos utilizados para medir los niveles de expresión de DTL, WWOX y cdh13 fueron (en sentido 5’-3’):

Nombre del Cebador	Secuencia (en sentido 5’-3’)
DTL Cebador Fw	TGTATCCCTACCCATCTCCTGGCA
DTL Cebador Rv	CTCATTACAGTTGACCTTTGGCACCC
WWOX Cebador Fw	GGCCTTTCACCAAGTCCATGCAAC
WWOX Cebador Rv	AAGTACATCCCTCCCAGACCCTCC
cdh13 Cebador Fw	TGCCCCGTTTCATTTACCCACAGT
cdh13 Cebador Rv	AGAAGGCTTACCAGGGCGTGTGT
GAPDH Cebador Fw	GCTCCTCCTGTTTCGACAGTCA
GAPDH Cebador Rv	AATCCGTTGACTCCGACCTTC

En el caso de DTL, los cebadores están diseñados sobre la secuencia del extremo 3’ del gen, produciendo un amplicón de 148 pb que coincide exactamente con la secuencia cubierta por el “probeset” que cubre el gen en el microarray de expresión.

Los cebadores han sido diseñados con el programa informático “Primer3”¹³³, optimizando los valores de autocomplementariedad, de complementariedad en 3’, valor de Tm (2-3 grados superior a la temperatura de anillamiento seleccionando la opción de tablas termodinámicas de SantaLucía 98 y considerando la concentración de 100mM de KCl del tampón de PCR), tamaño del amplicón (100 - 250 pb) y composición en GC (50 – 55%), entre otros. En consecuencia la eficiencia de la PCR con estos cebadores es cercana al 100%. Además, se comprobó mediante el programa “primer-blast”, del NCBI, la improbable amplificación inespecífica de otros mRNA con estos pares de cebadores.

Para el estudio de expresión de DTL se seleccionó como gen de referencia al gen TPT1, tras un estudio bioinformático que buscó el gen con menor cambio en los datos de expresión del grupo de 38 tumores de SE (grupo 2 de tumores) y posteriormente en las colecciones de expresión “HG_U133A/GNF1H y GNF1M Gene Atlas”¹³¹ con el siguiente procedimiento: se calculó como valor de dispersión de cada gen en el total de muestras la desviación estándar partida por la media. Se calculó asimismo la ratio entre el valor de expresión máximo y el

mínimo. Se tomó como valor discriminatorio el producto de ambos factores y se seleccionó el gen con menor valor para este parámetro. Se tuvo en cuenta también los altos niveles de expresión del gen, para que fuera fácilmente detectable. La función de TPT1, el gen con la expresión menos variable según este estudio, es típicamente la de un gen “housekeeping”, a pesar de su nombre (“tumor protein, translationally-controlled 1”), pues su función está relacionada con la unión de calcio y la estabilización de microtúbulos.

Los cebadores usados para TPT1 (diseñados de modo idéntico al descrito para los de DTL, sobre la secuencia cubierta por el “probeset” del microarray) fueron (secuencia en sentido 5’-3’):

TPT1 cebador Fw - TTGGACTACCGTGAGGATGGTGTG

TPT1 cebador Rv- TGTGGATGACAAGCAGAAGCCAGT

3.13 Transducción lentiviral de líneas celulares de SE

Se utilizó una colección de shRNA compuesta de 5 construcciones (MISSION, SIGMA) cada una de ellas dirigida contra una parte distinta del transcrito de DTL, clonadas en el vector pLKO.1. Junto con esta colección se utilizaron 2 vectores control: el pLKO.1-“non-targeting control” (NTC, que contiene una construcción shRNA que no está dirigida contra ningún transcrito celular en concreto) y el pLKO.1-“turboGFP”, que contiene la secuencia de la proteína verde fluorescente de copépodo (SIGMA). Asimismo, se utilizaron los vectores “pCMV-dr8.91” y “pMD.G” (tronolab), que aportan los genes del HIV-1 gag y pol (entre otros) y la proteína de la envuelta del virus de la estomatitis vesiculosa (vsv-g), respectivamente. Este es un sistema para la producción de virus defectivos en replicación de segunda generación: al estar los genes virales separados en 2 vectores que además no son empaquetados en las partículas virales, estas son incapaces de replicarse tras entrar en la célula diana y tan sólo actúan portando el transgén seleccionado.

Todos estos vectores se co-transfectaron en cultivos de células 293T al 80-90% de confluencia con un protocolo típico de lipofectamina. Estas células son utilizadas por tanto como células empaquetadoras o productoras virales. Al cultivo de 293T se le cambia el día anterior a la transfección el medio de cultivo, adicionando 12 ml de DMEM-glutamina, 10%

FBS, con una décima parte de la concentración habitual de penicilina-estreptomicina. La mezcla de transfección usada es la siguiente (por cada placa de 10 cm a transfectar):

pLKO.1-shRNA	9 µg
pCMV- dr8.91	10 µg
pMD.G	5 µl
Optimem (Invitrogen)	1.5 ml

En otro tubo se prepara la mezcla de 1.5 ml de “Optimem” (Invitrogen) con 60 µl de Lipofectamina 2000, 1 mg/ml (Invitrogen). La mezcla con la lipofectamina se adiciona suavemente, por goteo, sobre el volumen con la mezcla de las construcciones plasmídicas. Se mezcla suavemente por pipeteo y se incuba 20 minutos a temperatura ambiente. Tras ello se añaden los 3 ml suavemente al cultivo de 293T, por goteo, y se homogeniza inclinando la placa suavemente. Se incuba la transfección durante la noche, en el incubador de cultivos celulares.

Al día siguiente se retira el medio de transfección y se añaden 9 ml de medio con exceso de FBS (DMEM 40% FBS, mejora la producción viral) y 0.1x de penicilina/estreptomicina. Se considera este momento el punto de tiempo 0 de la producción viral.

Al día siguiente, 24 horas antes de la transducción, se siembran las células diana (las que van a ser transducidas) en placas de 6 pocillos, poniendo por pocillo 4×10^5 células, una placa por cada placa de 10 cm de células productoras transfectadas. Esto es para conseguir un silenciamiento completo de la proteína DTL.

Cuarenta y ocho horas tras las transfección se recoge el sobrenadante de las placas de células empaquetadoras con una jeringa y se filtra a través de un filtro de 0.22 µm de polisulfonato (de baja afinidad proteica, PALL). Se añaden 3 ml de medio fresco (DMEM 40% FBS). Además, se añade polibreno a una concentración final de 8 µg/ml (“hexadimethrine bromide”, SIGMA; este compuesto químico anula las cargas eléctricas, que producen una repulsión entre la membrana celular y la superficie de las partículas virales).

Los sobrenadantes virales procesados de este modo se añaden a las células diana (tras retirarles el medio), 2 ml por pocillo. Finalmente, las placas se centrifugan a 2000 rpm, 32 °C, durante una hora. Tras 8 horas de incubación en la estufa de cultivos (37 °C, 5% CO₂) se retira el medio de transducción y se añade el medio habitual del cultivo celular (frecuentemente RPMI al 10 – 20% de FBS para la mayoría de líneas celulares de SE).

En caso de los ensayos de silenciamiento progresivo, se toma el sobrenadante original, y se diluye más de lo habitual (se añaden 6 ml de medio fresco). Se toma una fracción de este sobrenadante y se diluye de manera progresiva, obteniendo distintos tubos falcon cada uno con un porcentaje inferior del sobrenadante original. No se centrifugan las placas.

Para el control de la eficiencia de transducción se transducen las células diana con la construcción pLKO.1-turboGFP, sin diluir o diluida de forma seriada con un factor de dilución 1/10 (desde 1 hasta 10^{-6} - 10^{-7}). Tras 48-72 horas se observa la población celular al microscopio de fluorescencia y se adquieren las células en PBS en el citómetro de flujo. Un control negativo de transducción (no transducido o transducido con el control NTC) se adquiere junto con ellas. Una vez delimitada la población negativa para GFP (la población control sin GFP), se calcula el porcentaje de población celular positiva para GFP para cada punto de la dilución seriada. Si se quiere titular la transducción, se selecciona el valor de la dilución seriada que consigue una transducción de un 20%-50% de células positivas (para evitar la subestimación debida a la presencia de células con multiplicidad de infección, o moi, superior a 1) y se calcula con este porcentaje el tamaño de la población transducida, teniendo en cuenta el número de células totales transducidas (para lo que es necesario realizar un conteo celular de un pocillo no transducido en el momento de la transducción), se multiplica por el factor de dilución y se divide el resultado por el volumen de sobrenadante aplicado por pocillo: de esa manera tenemos el número de partículas infectivas (diferente al número de partículas virales, ya que en la producción viral se generan multitud de partículas incompletas o sin el transgén) del sobrenadante viral, por ml. La moi se obtiene dividiendo este valor por el número de células que fueron transducidas por cada ml de sobrenadante viral.

3.14 Análisis de ciclo celular por citometría de flujo

Para este análisis las células son marcadas con yoduro de propidio (IP), que se une al DNA de manera proporcional a la cantidad de este, intercalándose entre los pares de bases nitrogenadas. La cantidad celular de DNA, y por tanto la señal debida al marcaje con IP, varía en las diferentes fases del ciclo celular, siendo el doble en la fase G2/M respecto a la fase G0/G1 y presentando todo el rango de valores intermedios en la fase S o de replicación del DNA.

Los cultivos celulares analizados estuvieron siempre en crecimiento activo, evitando la confluencia y subconfluencia. El número óptimo de células por pocillo de placa de 6, para las líneas celulares utilizadas en este estudio, es de $3 \times 10^5 - 4 \times 10^5$.

En primer lugar se aspira el medio y se tripsinizan las células con tripsina/EDTA (Invitrogen), durante el tiempo suficiente para que las células se despeguen y se separen bien entre ellas, pero acortándolo en lo posible para que no haya inducción de apoptosis. La reacción de la tripsina se detiene añadiendo un volumen de medio de cultivo-FBS igual al usado de tripsina (unos 200 – 300 μ l por pocillo). Se pasan las células despegadas a tubos de citometría y estos se centrifugan a 1200 rpm durante 3 minutos. Se aspira el sobrenadante y se lavan bien las células con PBS (con una nueva centrifugación y aspiración). Tras ello se añaden 700 μ l de etanol 70% frío (mantenido en congelador -20 °C) y se incuban los tubos cubiertos con “parafilm” a 4°C, entre 45 minutos a varias horas.

A continuación se hacen 2 lavados con PBS (centrifugando y aspirando el sobrenadante) y se añaden 200 -500 μ l de la siguiente mezcla:

- 1 ml de PBS, 5 μ l de RNasa y 5 μ l de IP.

Se incuban los tubos durante 1 – 2 horas, a temperatura ambiente y en agitación suave (tubos en horizontal). La RNasa eliminará rápidamente todo el RNA de las muestras y el IP se unirá estequiométricamente al DNA.

Finalmente, las muestras se adquieren en el citómetro de flujo (“FACScalibur”, Becton Dickinson) utilizando el programa “CellQuest” (BD). Normalmente se adquieren 2×10^5 células a una velocidad no superior a 1000 eventos por segundo. Las poblaciones se seleccionan con el programa informático “Paint-a-Gate”, descartando los dobletes y restos celulares (“debris”). El cálculo del reparto poblacional en cada fase del ciclo se realiza usando los algoritmos implementados en los programas informáticos “ModFit LT” (Verity Software House) y “FlowJo” (Treestar).

3.15 Análisis de muerte celular por citometría de flujo

El protocolo utilizado aprovecha el cambio en la distribución de la fosfatidilserina de la membrana celular durante el proceso de apoptosis, en el que este fosfolípido pasa de estar en la hemimembrana interna a la externa, y la gran afinidad y especificidad que tiene la proteína anexina V en la unión a este componente de la membrana. En células vivas la fosfatidilserina será inaccesible a la anexina V pero conforme transcurra la apoptosis pasará a estar de cara al exterior celular, produciéndose la unión con esta (en la etapa conocida como “apoptosis temprana”). El avance de la apoptosis provocará además una progresiva pérdida de la integridad de la membrana, con lo que el IP podrá penetrar en la célula e intercalarse en su DNA, por lo que en “apoptosis tardía” ambos compuestos marcarán las células. Finalmente, en la fase de muerte más avanzada, “necrosis”, la membrana estará desintegrada y sólo habrá unión del IP al DNA.

Para este ensayo se utilizó el Kit de AnexinaV-Dy-634 e IP de la empresa Immunostep.

Se retira el medio de las células (de un pocillo de placa de 6) y se añade a un tubo de citometría. Tras ello se tripsinizan las células y se detiene la reacción como se explica en el apartado de análisis de ciclo celular. Las células despegadas se pasan al tubo de citometría que contiene el sobrenadante del pocillo. Se precipitan las células por centrifugación.

Tras un lavado de las células con PBS se añade la siguiente mezcla (por pocillo):

tampón del kit (1x)	297 μ l
Anexina V – Dy-634	3 μ l

Tras mezclar bien e incubar 10 minutos en oscuridad se añaden 10 μ l del IP del kit, a lo que sigue una nueva incubación en oscuridad de 5 minutos.

Finalmente se adquieren 2×10^5 células en el citómetro de flujo, a una velocidad no superior a 2000 eventos por segundo. El análisis de datos se realiza utilizando el programa “Paint-a-Gate” (BD), delimitando el tamaño de las poblaciones celulares en cada fase apoptótica.

3.16 Ensayo de proliferación MTT

Este ensayo evalúa la proliferación celular haciendo uso del reactivo MTT “(3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide)”, Sigma) o sal de tetrazolio, que reacciona en las mitocondrias de las células metabólicamente activas y viables, siendo reducida por estas y transformándose en formazán. En este proceso el reactivo cambia de color (de amarillo pasa a azul) y se vuelve hidrofóbico, quedando capturado en la célula, por lo que se puede solubilizar con DMSO y cuantificar por colorimetría.

Para el ensayo de MTT se trabaja en formato de placa de 24 pocillos, cada condición ensayada por triplicado. Las células transducidas se siembran en diferentes cantidades para que no alcancen la confluencia en ningún momento del periodo de ensayo (5 días), sembrándose desde 1×10^5 a 5×10^3 células por pocillo (una placa por día de medida).

Se retira el medio y se añade por pocillo una mezcla de 270 μ l de medio completo y 30 μ l de la solución de MTT (5 μ g/ μ l en PBS). La placa se incuba a 37 °C una hora, tras lo que se retira el medio, se aplican 500 μ l de DMSO y se agita la placa suavemente durante 5 minutos, protegida de la luz.

Finalmente se mide la absorbancia a 570 nm en el lector de placas (“Ultra Evolution Microplate Reader”, TECAN). Para cada punto de tiempo se expresa el crecimiento relativo de las condiciones transducidas con el shRNA contra DTL respecto a la condición transducida con el NTC, como porcentaje de la absorbancia de este.

3.17 Ensayos in vivo de xenotransplantes en ratón

Para estos ensayos se transdujeron las líneas celulares RM82 y TC32 con las construcciones lentivirales de shRNA, tal y como se explica en el apartado correspondiente (3.13), con una dilución seriada de los sobrenadantes virales (100%, 75% y 50% del sobrenadante original).

Tras 48 horas las células se tripsinizaron, contaron y se prepararon suspensiones de 2×10^6 células en un volumen total de 0.2 ml compuesto por medio RPMI y “Matrigel” (BD) en una proporción relativa 1:1.

Estas suspensiones se inyectaron de forma subcutánea en el flanco izquierdo (células transducidas con el control NTC) o en el flanco derecho (células transducidas con las construcciones shRNA contra DTL) de ratones NOD/SCID (“Non-Obese-Diabetic/Severe-Compromised-ImmunoSystem”) de 4 semanas de edad (Charles River), concretamente ratones NOD/CrCrI- Prkdc^{scid}.

Se usaron 12 ratones por cada línea celular, haciendo cuadruplicados de cada condición de silenciamiento. Tras 19 días los animales fueron sacrificados mediante eutanasia consistente en sobredosis del anestésico Isofurano (Abbott Laboratories). Los tumores fueron pesados y la mitad de cada pieza tumoral se congeló en resina OCT y la otra mitad se fijó en paraformaldehído para estudios de análisis histopatológico.

El estudio fue previamente aprobado por el comité de bioética de la institución (Universidad de Salamanca) y se realizó de acuerdo con la normativa vigente Española y Europea (RD 1201/05 y 86/609/CEE, respectivamente).

3.18 Extracción de proteínas y Western blot

Para extraer proteínas de cultivos celulares se comienza retirando el medio de cultivo y se lavan las células con PBS frío, con las placas sobre hielo. Se añaden entonces 500 µl de PBS frío y se despegan las células con un raspador. El volumen se pasa a un tubo de 2 ml y se precipitan las células por centrifugación (1200 rpm, 4 °C, 3 minutos), tras lo cual se descarta el sobrenadante y se añade solución de lisis RIPA suplementada con inhibidores de proteasa y fosfatasa (el volumen dependerá de la cantidad de células, en general 100 µl por cada 3x10⁶ células).

A continuación se determina la concentración proteica mediante el uso de reactivo Bradford (Bio-Rad). Para ello se crea un patrón de concentración conocida, a partir de una solución de BSA de 2 mg/ml (Perkin Elmer) en volúmenes finales de 800 µl, desde 2 µg/ml a 10 µg/ml. De manera similar se preparan las muestras problema, añadiendo 1 µl del extracto proteico en el mismo volumen final. Se hacen duplicados de cada tubo. Se añaden 200 µl del reactivo Bradford, se mezcla bien por vórtex, se incuba 5 minutos y se mide la absorbancia en el

espectrofotómetro, a 595 nm. Las concentraciones de las muestras problema se calculan interpolando en la recta patrón (utilizando la fórmula del modelo de regresión creado).

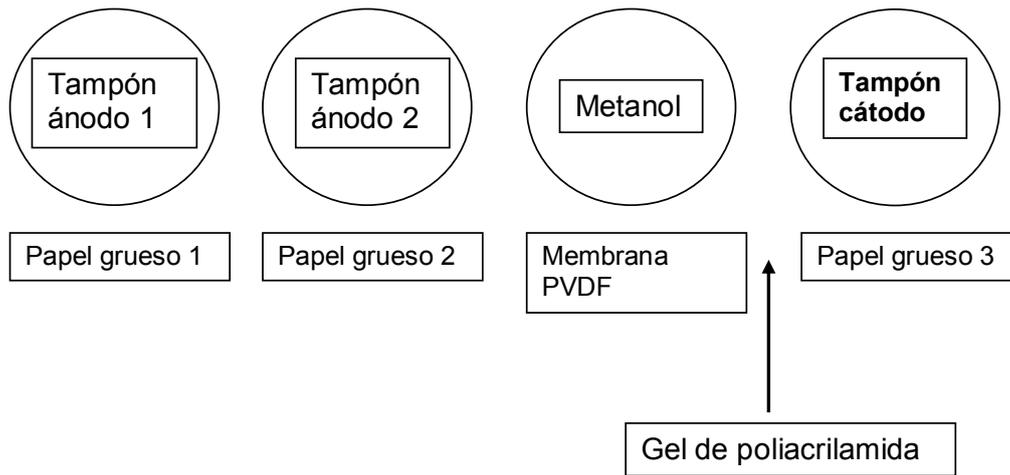
Los extractos proteicos se someten entonces a electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE). Para ello, 40 µg de proteína son primero desnaturalizados añadiéndoles tampón de carga 4x e hirviendo las muestras a 95 °C durante 5 minutos. En la preparación del gel se siguen las siguientes recetas, para 20 ml totales (los porcentajes usados dependen del tamaño de la proteína o proteínas que se pretenda visualizar, usándose mayores concentraciones cuanto menor sea el tamaño de la proteína de interés):

Porcentaje de poliacrilamida	H ₂ O	Acrilamida/Bis	Tris-HCl*	SDS 10%	TEMED	APS 10mg/100µl
4%	12.68	2	5	0.2	20µl	100 µl
6%	11.68	3	5	0.2	20µl	100 µl
8%	10.68	4	5	0.2	20µl	100 µl
10%	9.68	5	5	0.2	20µl	100 µl
12%	8.68	6	5	0.2	20µl	100 µl

* El Tris HCl varía en concentración y pH dependiendo de si el gel es el superior (“stacking gel”, 0.5M, pH 6.8, conocido como “Upper Tris”) o el inferior (“resolving gel”, 1.5 M, pH 8.8, conocido como “Lower Tris”)

Lo reactivos TEMED (“Tetramethylethylenediamine”, SIGMA) y APS (persulfato amónico, SIGMA) se añaden justo antes de la adición a los cristales montados en los soportes, para desencadenar la reacción de polimerización. Se añade primero el gel inferior (“resolving gel”), tras lo cual se añade isopropanol para facilitar su polimerización. Una vez esta ocurre, se retira el isopropanol y se añade el gel superior (“stacking gel”, al 4% de poliacrilamida) junto con el peine que creará los pocillos. Tras la polimerización se monta el sistema de electroforesis (“Miniprotean”, Bio-Rad), se rellena la cubeta con solución SDS-PAGE 1x, se depositan las muestras en los pocillos y se conecta la fuente de alimentación, iniciándose la electroforesis a 90 voltios hasta que las muestras alcanzan el gel inferior, momento en que se sube el voltaje a 120 – 150 voltios.

Una vez las muestras se resuelven hasta el punto deseado, se realiza una transferencia semiseca con el dispositivo “Trans-Blot Semi Dry” (Bio-Rad), durante 2 horas a 200 miliamperios, utilizando una membrana de PVDF (poli fluoruro de vinilideno, Bio-Rad). El montaje del sistema responde al esquema mostrado a continuación.



El esquema representa la posición de los distintos papeles empapados en los diversos tampones, así como la posición del gel y la membrana.

Tras la transferencia se confirma la electroforesis y transferencia mediante tinción de la membrana con solución de Rojo Ponceau, que marca las proteínas en la membrana y se elimina fácilmente con lavados de TBS-T (solución salina tamponada con Tris y Tween 20).

Finalmente la membrana se somete a la incubación con los distintos anticuerpos. En todos los casos la membrana se bloquea primero con 5% de leche en polvo en TBS-T, durante 1 hora a temperatura ambiente y se incuba el anticuerpo primario durante la noche en 1% BSA en TBS-T, a 4 °C. El anticuerpo secundario siempre se utiliza a una dilución 1:10000 en 1% BSA en TBS-T, incubado 1h a temperatura ambiente. También en todos los casos las incubaciones se hacen en agitación horizontal circular a una velocidad de 70 rpm. Los lavados entre soluciones se hacen con TBS-T.

Los anticuerpos primarios usados, las diluciones y el sistema de revelado para cada uno de ellos son los siguientes:

El anticuerpo contra DTL utilizado es un anticuerpo no comercial, amablemente cedido por los doctores Hung-Wei Pan y Hey-Chi Hsu (“Pathology Department”, “College of Medicine, National Taiwan University”, Taipei, Taiwán), ratón policlonal, 1:500, ECL plus (Amersham, GE).

Igualmente, el anticuerpo anti-WWOX ha sido desarrollado y suministrado por el laboratorio de los doctores Kay Huebner y Carlo Maria Croce (“Department of Molecular Virology, Immunology, and Medical Genetics, Comprehensive Cancer Center, and College of

Pharmacy”, “Ohio State University”, Columbia, EEUU), es un policlonal de ratón y fue utilizado en dilución 1:20.000 y revelado con el sistema ECL (Amersham, GE).

- Anticuerpos de Santa Cruz:

Anti Ciclina D1 (sc-718), conejo policlonal, 1:500, ECL

Anti Ciclina E (sc-247), ratón monoclonal, 1:500, ECL

Anti p21 (sc-397), conejo policlonal, 1:500, ECL plus

Anti p53 (mezcla de sc-126 y sc-98), ratón monoclonal ambos, 1:500 y 1:1000, ECL plus

- De otras casas comerciales:

Anti p27, Pharmingen (554069), ratón monoclonal, 1:1000, ECL plus

Anti CDT1, Abcam (ab70829), conejo policlonal, 1:500, ECL plus

Anti Ciclina A, SIGMA (c-4710), ratón monoclonal, 1:500, ECL

Anti β -tubulina, SIGMA (t-5293), ratón monoclonal, 1:10000, ECL

Anti actina, SIGMA (A2103), conejo policlonal, 1:5000, ECL

Anti cdh13, R&D Systems (AF3264), cabra policlonal, 1:1000, ECL

Anticuerpos secundarios:

Los que se incluyen en el kit ECL (Amersham, GE)

Para aplicar un anticuerpo nuevo sobre una membrana sobre la que ya se ha usado y revelado un anticuerpo primario anterior se procede incubando la membrana en la solución de “stripping” (precalentada a 50 °C) durante no más de 30 minutos a 50 °C, lavando la membrana posteriormente varias veces con TBS-T, lavados de 10 minutos. Tras ello se debe bloquear de nuevo la membrana con leche 5% en TBS-T.

Soluciones:

Solución de lisis RIPA (100 ml):

1 ml de Tris HCl 1M pH 7.5

2 ml de NaCl 5M

0.2 ml de EDTA 0.5M

1 ml de NP-40

0.5 g de deoxicolato sódico

1 ml de SDS 10%

Agua hasta 100 ml

A esta solución se le añade en fresco, cada vez que va a ser utilizado, para 1 ml total de solución de lisis:

900 μ l de solución RIPA + 40 μ l de inhibidor de proteasas + 50 μ l de NaF 1M + 10 μ l de NaOV 200 mM

Composición Tampón de carga 4x: Tris 200mM pH 6.8, 40% Glicerol, 4% β -ME, 0.04 azul de bromofenol.

Tampón del ánodo 1: 18.72 g de Tris (Sigma) en 500 ml de agua milliQ para una concentración final de 0.3 M (El pH de esta solución no necesita ajustarse).

Tampón del ánodo 2: 1.51 g de Tris (Sigma) en 500 ml de agua milliQ para una concentración final de 0.025 M (El pH de esta solución no necesita ajustarse).

Tampón del cátodo: 1.51 g de Tris (Sigma) en 500 ml de agua milliQ para una concentración final de 0.025 M. El pH debe ajustarse a 9.4 mediante la adición de 0.1 g de glicina (Merck).

Solución de rojo Ponceau:

3 g de rojo ponceau (SIGMA)

30 ml TCA (Ácido Tricloroacético al 100%, SIGMA)

Enrasar a 1 L con agua milliQ

Solución de “stripping”:

20 ml SDS 10%.

6.2 ml “Upper Tris” (Tris-HCl 0.5 M) pH 6.7

73.1 ml H₂O

Añadir 175 μ l β -ME a 25 ml de solución de “stripping” justo antes de usar.

3.19 Inmunohistoquímica

Se realizó sobre cortes de tejido fijados en formaldehído y embebidos en parafina, de 4 µm de grosor.

El protocolo se inicia con la desparafinación de los cortes calentándolos en una estufa (Memmert) a 60 °C durante toda la noche e introduciéndolos posteriormente en xilol (Panreac) durante 8 minutos. A continuación se sumergen los cortes en una solución de xilol/etanol 1:1, durante 5 minutos, tras lo que se procede a la hidratación, sumergiendo las preparaciones sucesivamente en alcoholes de gradación decreciente (100, 96 y 80%), 4 minutos en cada uno, y finalmente en agua destilada, durante 3 minutos.

Tras ello se realiza el desenmascaramiento antigénico, mediante calentamiento de los cortes en olla a presión con las muestras inmersas en tampón citrato pH 6 (Roche-Ventana Medical Systems), y el bloqueo inespecífico de las proteínas de las preparaciones con el Kit “Omnimap” (Roche-Ventana Medical Systems).

Se procede entonces a la inhibición de la peroxidasa endógena, incubando las preparaciones con H₂O₂ (Dako) durante 10 minutos y se aplica el anticuerpo primario dirigido contra Ki67 (MIB-1, Dako) o el anticuerpo primario contra WWOX (descrito en el apartado 3.18), diluido en la solución “Antibody Diluent” (Dako, Glostrup, Dinamarca). Se incuba el anticuerpo primario durante 1 hora a 37 °C, realizándose posteriormente tres lavados con PBS 1X, durante 5 minutos cada uno.

A continuación se incuban las preparaciones con el anticuerpo secundario (Roche-Ventana Medical Systems), diluyéndolos en la misma solución que el anticuerpo primario. Se incuba durante 30 - 40 minutos y se lava con PBS 1x, 3 veces, 5 minutos cada lavado.

Para el revelado se utiliza el método de visualización LSAB (Dako), tras lo cual se procede a deshidratar las muestras sumergiéndolas en una batería de soluciones de etanol de gradación creciente (80, 96 y 100%) y xilol (etanol 100% 1:1 y xilol 100%), 4 minutos en cada solución de etanol y 5 minutos en las de xilol, repitiendo el último paso dos veces más.

Finalmente se montan las preparaciones con medio “DEPEX” (Merck) y se guardan para posterior análisis al microscopio.

Para la evaluación de la inmunohistoquímica de Ki-67 las secciones fueron evaluadas por 2 patólogos de forma independiente. Sólo la tinción nuclear fue considerada positiva. La inmunoreactividad se puntuó de acuerdo con la siguiente escala: negativa (0) para menos de un 5% de células tumorales teñidas; débilmente positiva (1) para entre el 5% y el 10% de células tumorales teñidas; moderadamente positiva (2) para entre el 10% y el 50% de células tumorales teñidas; y fuertemente positiva (3) para más del 50% de células teñidas.

3.20 Clonación de las secuencias codificantes de *WWOX* y *cdh13* en *pMIG*

La secuencia codificante no mutada del transcrito 1 de *WWOX* (de 1245 pb, secuencia de referencia del NCBI: NM_016373.1) fue amplificada a partir del cDNA de la línea SKES, mientras que la de *cdh13* (de 2142 pb, secuencia referencia del NCBI: NM_001257.3) fue obtenida a partir de RNA de tejido sano de corazón y cerebro humanos (Ambion), los tejidos con mayor expresión de este gen.

Los cebadores para la amplificación fueron diseñados sobre las secuencias UTR y se les incorporaron colas con las dianas de restricción adecuadas para clonar el producto de PCR en el vector *pMIG*, tras realizar un estudio de restricción y seleccionar aquellas enzimas sin diana en las secuencias a clonar y con corte en el sitio de clonación múltiple del vector (utilizando el programa informático “on-line” “NEBcutter2”, <http://tools.neb.com/NEBcutter2/>). Los cebadores utilizados fueron (secuencias en sentido 5'-3'):

- Cebador para la clonación de *WWOX*, Fw:

CG GGATCC GCCAGGTGCCTCCACAGTCAGCCATGGCAGC
 apoyo BamHI secuencia *WWOX* 5'UTR inicio sec. codificante

- Cebador para la clonación de *WWOX*, Rv:

ACGC GTCGAC TGTGTGTGCCCATCCGCTCTG
 apoyo Sall secuencia homóloga a la 3' UTR de *WWOX*

- Cebador para la clonación de cdh13, Fw:

TA GGATCC GCCGCCACC ATGCAGCCGAGAACTCCGC
 apoyo BamHI secuencia Kozak inicio secuencia codificante cdh13

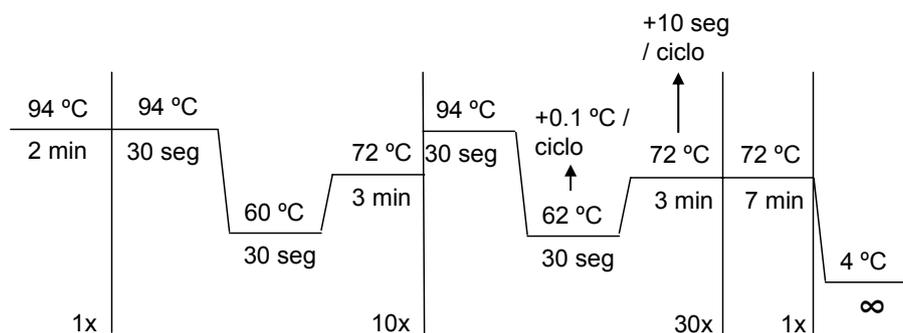
- Cebador para la clonación de cdh13, Rv:

CT GTTAAC TCACAGACAAGCTAAGCTGAAGAGGCTG
 apoyo HpaI (stop) secuencia homóloga al final de la secuencia codificante de cdh13

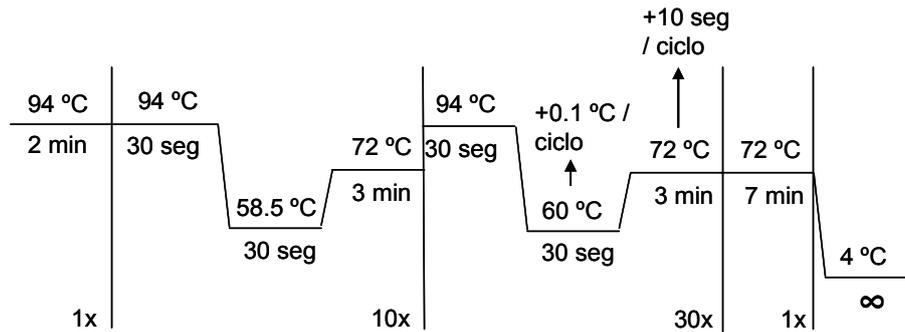
La amplificación se llevó a cabo utilizando una enzima DNA polimerasa termoestable de alta fidelidad (mezcla de una Taq polimerasa y de una enzima con actividad correctora de errores) y gran capacidad de extensión (hasta 5 kb), incluida en el kit “Expand High Fidelity PLUS PCR System” (Roche). Las condiciones de PCR utilizadas fueron (4 reacciones de restricción por cada clonación):

cDNA	50 ng
Tampón 5x del kit	5 µl
MgCl ₂ del kit	2 µl
dNTPs (100mM cada, Eppendorf)	0.5 µl
Cebador clonación Fw (20 mM)	0.4 µl
Cebador clonación Rv (20 mM)	0.4 µl
Enzima “High Expand Plus”	0.5 µl
Agua	hasta 25 µl

Condiciones de amplificación de la secuencia codificante de WWOX:



Condiciones de amplificación de la secuencia codificante de cdh13:



Una pequeña parte de la PCR (2 μ l) se resolvió en una electroforesis en gel de agarosa para comprobar la correcta amplificación (tamaño esperado de banda según el marcador de pesos moleculares “1kb DNA ladder” de Promega, cantidad de producto suficiente y ausencia de amplificación inespecífica) y el resto del volumen se guardó para la restricción.

El vector pMIG (6500 pb) se deriva del vector MSCV 2.2 tras incorporarle a este entre los sitios de clonación EcoRI-SalI la secuencia IRES-GFP perteneciente al vector pCITE IRES hGFP-3M. Es un vector retroviral que permite co-expresar en un mRNA bicistrónico la proteína de interés y la GFP, aunque carece de resistencia a antibiótico de selección. El mapa depositado en el repositorio “Addgene” (www.addgene.org) es el siguiente:

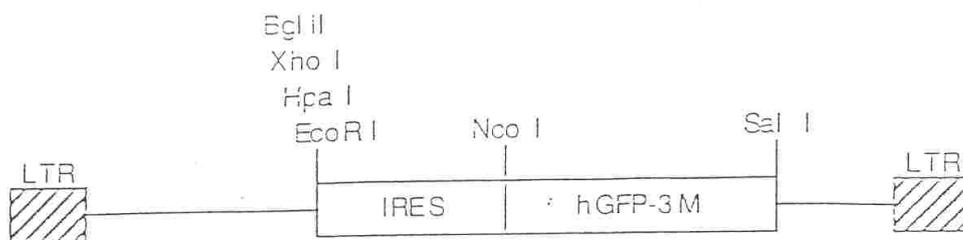


Figura 9. Mapa del vector pMIG depositado en el repositorio Addgene (www.addgene.org) por su creador

Tanto el vector como los productos de PCR purificados (usando el kit “QIAquick PCR Purification Kit”, Qiagen) se sometieron a una doble restricción para la clonación de la secuencia codificante de WWOX; en el caso de *cdh13*, el vector pMIG se sometió a restricciones secuenciales primero con la enzima BglII y luego con la enzima KspAI (isoesquizómero de HpaI) por incompatibilidad de los tampones de reacción, según las siguientes condiciones:

Restricción vector pMIG (para WWOX)		Restricción producto de PCR - WWOX	
Tampón Tango 10x (Fermentas)	4 μ l	Tampón BamHI (Fermentas)	4 μ l
DNA	2 μ g	DNA (elución kit purificación)	30 μ l
BglII (Fermentas)	1 μ l	BamHI (Fermentas)	1.4 μ l
XhoI (Fermentas)	1 μ l	SalI (Fermentas)	2.6 μ l
agua	hasta 20 μ l	agua	hasta 40 μ l
Restricción 1 vector pMIG (para cdh13)		Restricción producto de PCR – cdh13	
Tampón O 10x (Fermentas)	4 μ l	Tampón BamHI (Fermentas)	4 μ l
DNA	2 μ g	DNA (elución kit purificación)	30 μ l
BglII (Fermentas)	2 μ l	BamHI (Fermentas)	1.4 μ l
		KspAI (Fermentas)	2.6 μ l
agua	hasta 20 μ l	agua	hasta 40 μ l
Restricción 2 vector pMIG (para cdh13)			
Tampón B 10x (Fermentas)	4 μ l		
Eluido Restricción 1 purificada	30 μ l		
KspAI (Fermentas)	4 μ l		
agua	hasta 40 μ l		

Las restricciones (incubadas durante toda la noche a 37 °C) se resolvieron mediante electroforesis en un gel al 1% de agarosa (de bajo punto de fusión, Conda) tras lo que se extrajeron las bandas correspondientes al vector linearizado y al producto de PCR con extremos cohesivos. Se purificaron las bandas con el kit de columnas “Illustra GFX PCR DNA and Gel Band Purification Kit” (Amersham, GE). El eluido del vector, de 30 μ l, se defosforiló según la siguiente reacción:

Vector abierto y purificado	30 μ l
Tampón NEB 3	6 μ l
CIP (fosfatasa alcalina de intestino de ternera, NEB)	3 μ l
agua	11 μ l

La reacción de fosforilación se incubó durante 2 horas a 37 °C, tras lo cual se volvió a purificar, de nuevo con el citado kit de Qiagen. El eluido final, con el vector ya abierto y defosforilado, fue cuantificado en el espectrofotómetro “ND1000” (Nanodrop), al igual que el producto de PCR con extremos cohesivos. Tras ello ambas secuencias (vector e inserto) se

ligaron usando la ligasa de DNA del fago T4 (Fermentas). Se ensayaron diversas condiciones, especialmente en el caso de la construcción pMIG-cdh13, puesto que esta ligación fue más complicada de obtener por ser uno de los extremos ligados romo (el generado por el corte con la enzima KspAI). Para esta construcción fueron necesarios ratios moleculares iguales o superiores a 12:1 (inserto:vector) y la adición de polietilenglicol (PEG) a la reacción, mientras que la construcción pMIG-WWOX se generó con facilidad sin necesidad de añadir PEG, tanto con ratios moleculares de 3:1 como de 6:1. Se incluyeron asimismo ligaciones control carentes de inserto. Las reacciones de ligación óptimas (según el resultado de la posterior transformación en *E. coli* DH5 α y cribado de colonias) fueron:

Ligación pMIG (BglII/XhoI) – WWOX (BamHI/SalI)		Ligación pMIG (BglII/HpaI) – cdh13 (BamHI/HpaI)	
Tampón T4 DNA Ligasa (Fermentas)	1 μ l	Tampón T4 DNA Ligasa (Fermentas)	4 μ l
pMIG	90 ng	pMIG	20 ng
WWOX	30 ng	cdh13	80 ng
T4 DNA ligasa (Fermentas)	1 μ l	T4 DNA ligasa (Fermentas)	2.5 μ l
agua	hasta 10 μ l	PEG (Fermentas)	1 μ l
		agua	hasta 10 μ l

Para la transformación en el hospedador bacteriano la mitad de la reacción de ligación (5 μ l) se añadió a 100 μ l de *E. coli* DH5 α competentes (protocolo de competencia realizado por la Unidad de Lavado y Esterilización del CIC) descongeladas en hielo y se incubó la mezcla en hielo durante 30 minutos. Tras ello se sometió a las bacterias a un choque térmico de 40 segundos a 37 °C seguido de 1 minuto en hielo. Finalmente se les añadió 900 μ l de medio LB (Luria-Bertani), se incubaron en el agitador a 37 °C y 250 rpm durante una hora y se sembraron en placas de agar-LB-ampicilina (100 μ g/ml de ampicilina, Unidad de Lavado y Esterilización del CIC) en cantidades crecientes (para evitar el posible césped bacteriano).

Tras incubar boca abajo las placas durante 15 – 20 horas en la estufa a 37 °C, evitando la formación de colonias satélite, se hizo una primera evaluación del resultado comparando el número de colonias en las placas vector más inserto con las colonias de las placas control (ligaciones con sólo vector), confirmando su mayor número en las placas con la ligación completa. Se inocularon cultivos líquidos (7 ml de LB con ampicilina a 100 μ g/ml) a partir de 6 - 12 colonias de cada construcción y condición, que fueron sometidos a miniprep (“Nucleobond plasmid mini”, Macherey-Nagel) tras ser incubadas durante 16 horas a 37 °C, agitación 250 rpm, y a posterior restricción para la selección de las construcciones positivas.

En el caso de WWOX, dado que ambos sitios de restricción se perdieron durante la clonación (ligaciones BamHI-BglII y XhoI-SalI), se cortó con la enzima de restricción ClaI, que tiene un corte dentro del inserto (a 800 pb del sitio de clonación BamHI) y otro en el vector, debiendo liberarse una banda de 1.9 kb en las construcciones correctamente formadas. De igual manera, los productos de miniprep de la construcción pMIG-cdh13 se comprobaron por restricción con EcoRI (ya que el sitio de restricción en 5' se perdió en la clonación, BamHI-BglII), que corta dos veces en la secuencia codificante de cdh13 (a su inicio y final) liberando también una banda de 1.9 kb. Las restricciones se resolvieron en gel de agarosa al 1% y se seleccionaron las construcciones positivas (las que liberaron un inserto del tamaño correcto, 1.9 kb, según el marcador de pesos moleculares "1kb DNA Ladder" de Promega, resuelto en paralelo, Figuras 10 y 11) para ser posteriormente comprobadas por secuenciación (Servicio de Genómica del CIC), utilizando los siguientes cebadores oligonucleotídicos. Las secuenciaciones con estos cebadores permiten cubrir toda la secuencia clonada con una cobertura de 2x (la mayoría de la secuencia es cubierta por al menos 2 lecturas).

Nombre Cebador	Secuencia (en sentido 5'-3')
pMIG 5'	CTTTATCCAGCCCTCAC
pMIG 3'	GCCTTATTCCAAGCGG
WWOX Real Time Fw	(ver apartado Real Time)
WWOX Real Time Rv	(ver apartado Real Time)
cdh13 sec 1 Fw	GGCCGGTGCCTCTGGAA
cdh13 sec 1 Rv	GCAGGAGGGCATTATCGGTG
cdh13 sec 2 Fw	AGCCACCGTCCACATCACT

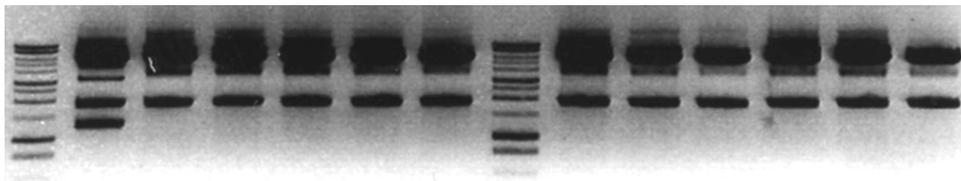


Figura 10. Ejemplo de comprobación de 12 minipreps de construcciones pMIG-WWOX mediante restricción con la enzima ClaI. Las calles 1 y 8 corresponden al marcador de pesos molecular "1kb DNA ladder" de Promega. Tan sólo la construcción resuelta en la calle 2 es negativa. Todas las demás liberan una banda del tamaño esperado (1.9 kb).

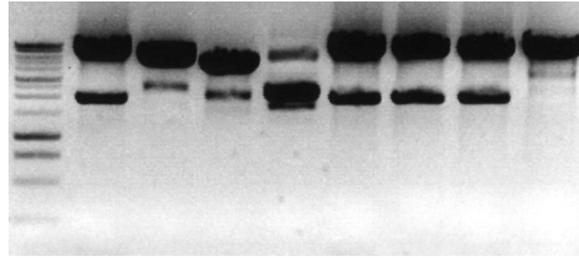


Figura 11. Ejemplo de comprobación de 8 minipreps de construcciones pMIG-cdh13 mediante restricción con la enzima EcoRI. La calle 1 corresponde al marcador de pesos molecular “1kb DNA ladder” de Promega. Las construcciones resueltas en las calles 3, 4, 5 y 9 son negativas (seguramente eventos de recombinación en *E. coli*) mientras que las resueltas en las calles 2, 6, 7 y 8 son positivas pues liberan una banda del tamaño esperado (1.9 kb).

La secuenciación descartó la presencia de mutaciones aleatorias originadas durante la amplificación por PCR en varias de las construcciones, las cuales fueron re-transformadas en *E. coli* DH5 α , crecidas en cultivos de 100 ml de LB-ampicilina y sometidas a midiprep (“QIAGEN Plasmid Midi Kit”, Qiagen). Los DNA finales se volvieron a comprobar por restricción, para descartar la eliminación o alteración de las construcciones durante el cultivo bacteriano (frecuente en el caso de los vectores retro y lentivirales crecidos en cepas de *E. coli* positivas para el gen recA, como es la DH5 α , debido a las secuencias LTR que contienen) y se utilizaron en la transfección de cultivos celulares sólo aquellos con unos parámetros de concentración y pureza mínimos (concentración de al menos 1 $\mu\text{g}/\mu\text{l}$, ratio 260/280 de al menos 1.8 y ratio 260/230 de al menos 2.0, según la media de 2 mediciones con el espectrofotómetro “ND1000”, Nanodrop).

3.21 Transducción retroviral de líneas celulares de SE

El sistema utilizado se basa en la producción de retrovirus defectivos en replicación de segunda generación, similar al descrito anteriormente para la generación de lentivirus (apartado 3.13). Sin embargo, a diferencia del protocolo usado para la transducción lentiviral, los vectores de producción retroviral (en este caso los genes gag-pol transfectados pertenecen al virus de la leucemia murina de Moloney, Mo MuLV) se co-transfectaron en cultivos de células 293T a baja confluencia mediante un protocolo típico de transfección con fosfato cálcico (“CalPhos Mammalian Transfection Kit”, Clontech).

El día anterior a la transfección se siembran 2×10^6 células 293T por placa de 10 cm, en un volumen total de 8 ml de DMEM, 10% FBS, suplementado con glutamina y

penicilina/estreptomicina. El día de la transfección, por la tarde, se añade 2 ml adicionales de medio y se realizan en dos tubos falcon de 15 ml las siguientes mezclas (por cada placa de 10 cm de 293T a transfectar):

- Tubo 1:

pMIG-WWOX/cdh13 o vector vacío	8 µg
pUMVC (genes gag-pol del Mo MuLV)	10 µg
pMD.G (gen de la envuelta vsv-g)	5 µl
CaCl ₂ 2M	124 µl
agua estéril	hasta 1 ml

- Tubo 2:

1 ml de solución HBS 2x

Tras mezclar debidamente el contenido del tubo 1, este se añade sobre el volumen del tubo 2 lentamente por goteo mientras se mantiene este agitándose por vórtex a una velocidad de 1000 rpm. Una vez hecho esto se incuba la mezcla durante 5 minutos y después, tras resuspender el precipitado lentamente por pipeteo, se adicionan los 2 ml lentamente por goteo a la placa de 293T y se homogeniza bien agitando suavemente la placa con movimientos en cruz. Se debe comprobar al microscopio la presencia de un fino precipitado correspondiente a las partículas de fosfato cálcico que contienen los DNA transfectados.

Al día siguiente se retira el medio de transfección, se lava la placa con PBS para eliminar los restos del precipitado y se añaden 9 ml de medio fresco (de nuevo DMEM, 10% FBS suplementado). Este es el tiempo 0 de producción viral.

Un día después, 24 horas antes de la transducción, se siembran las células diana (las que van a ser transducidas) en placas de 6 pocillos, poniendo por pocillo 4×10^5 células y sembrando una placa de 6 pocillos por cada placa de 10 cm de células productoras transfectadas.

Cuarenta y ocho horas tras las transfección se procede de forma idéntica a la descrita para la transducción lentiviral: el sobrenadante es filtrado, se añaden 3 ml de medio fresco (DMEM 40% FBS) y polibreno a una concentración final de 8 µg/ml y se reparten 2 ml por pocillo de

las células a transducir, tras lo que las placas se centrifugan a 2000 rpm, 32 °C, durante una hora. La transducción se incuba durante 8 horas.

3.22 Análisis estadístico

Para la estadística se aplicó el programa informático SPSS versión 17. Todos los test estadísticos realizados y sus resultados se exponen con detalle en los apartados correspondientes de la sección Resultados.

4. RESULTADOS

4.1 Resultados de los estudios de aCGH en líneas celulares y muestras tumorales de SE

El resultado de los análisis de aCGH BAC-array de 67 muestras de tumores (grupo 1 de tumores, ver Material y Métodos apartado 3.1) y de 16 líneas celulares de SE se detalla en las Figuras 12 y 13.

El perfil global de las ANC más recurrentes coincide con el ya descrito por los trabajos previos: las alteraciones más frecuentes corresponden a la ganancia del cromosoma 8 (en especial el brazo 8q) y de los brazos 1q, 3p, 9p, 16q y 17p. Sin embargo, las frecuencias totales detectadas en el presente estudio son superiores en general a las previamente descritas (ver apartado 5.2 para una posible explicación de este hecho).

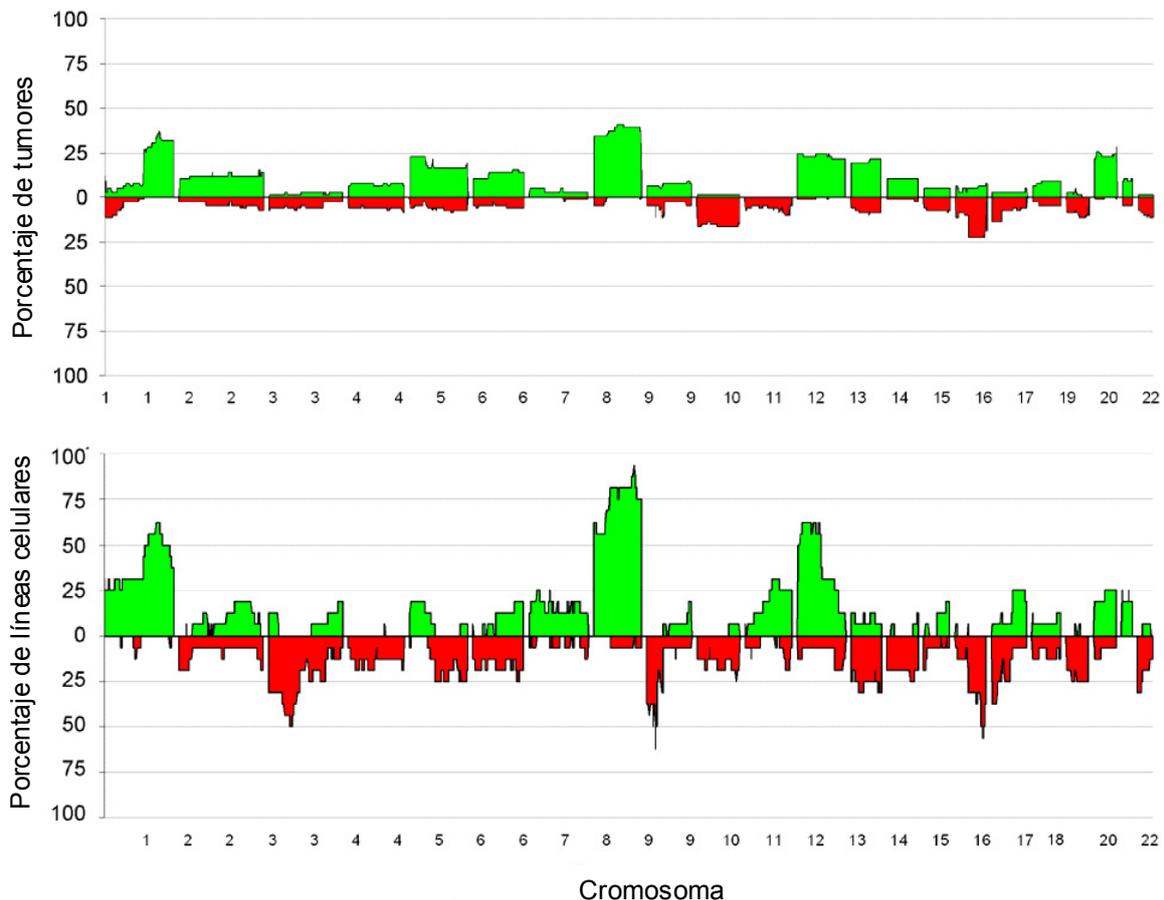


Figura 12. Frecuencia de las ANC en las muestras tumorales y en las líneas celulares de SE. En el panel superior: muestras tumorales de SE; en el panel inferior: líneas celulares de SE. El eje Y representa la frecuencia expresada como porcentaje muestral y el eje X el número de cromosoma. El color verde representa las ganancias de material genómico mientras que el rojo corresponde a las pérdidas.

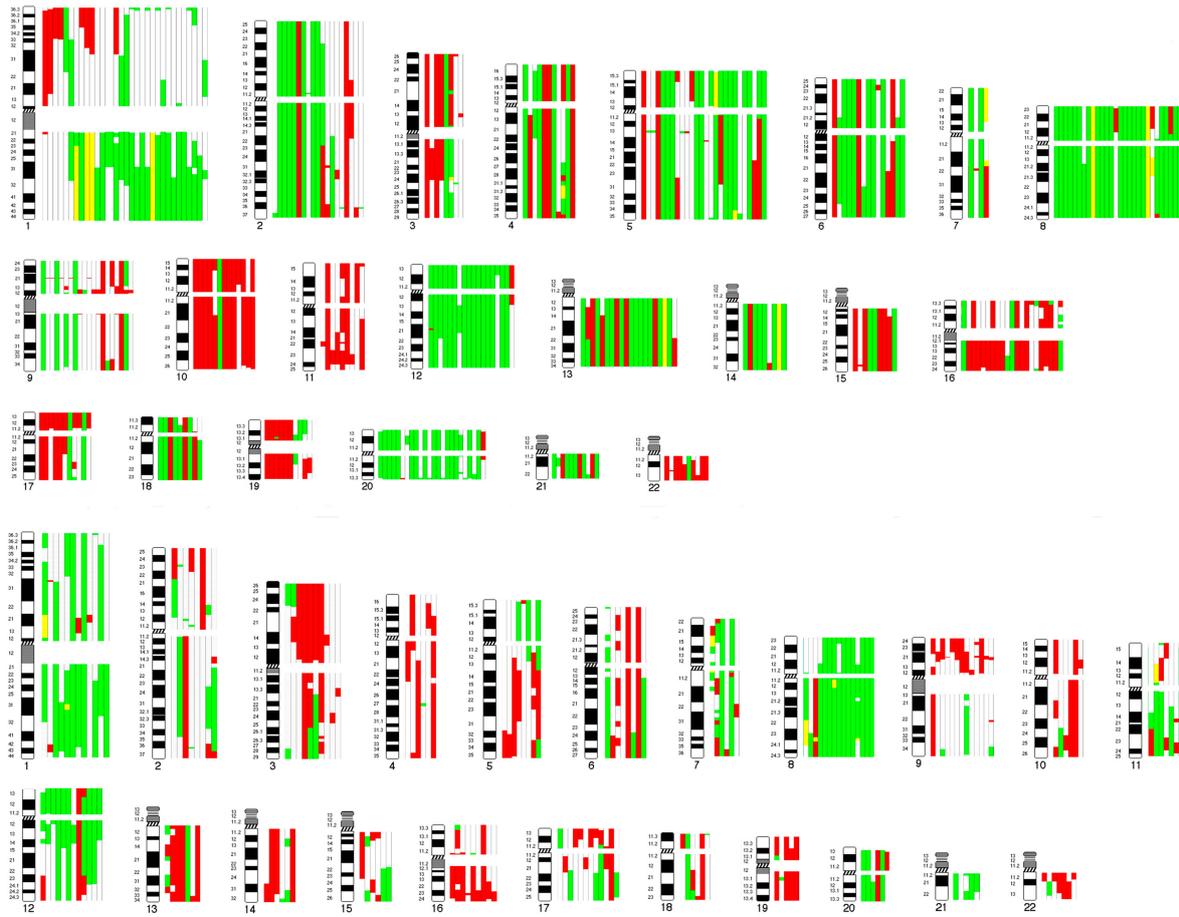


Figura 13. Representación de la posición y extensión de las ANC, en muestras tumorales (arriba) y en líneas celulares (abajo) de SE. El color verde representa ganancias mientras que el color rojo representa pérdidas de material genómico. Cada barra vertical a la derecha del ideograma cromosómico representa una muestra.

Dado que, tal y como se explica más adelante, la ganancia de 1q (1qG) fue la ANC con mayor relevancia clínica de todas las detectadas, cabe describir aquí con más detalle su frecuencia y extensión genómica en las muestras estudiadas. En total se identificaron 25 muestras tumorales (37.3%) con ANC en esta región. De ellas 15 (22.4%) presentaban ganancia del brazo 1q completo mientras que el resto, de acuerdo con el resultado del paquete informático “CGHregions”, se puede dividir en 2 grupos principales: el de aquellas que afectaban al 1q proximal (1q22-1q25.2; 154.6 – 180.3 mb desde el p-ter), 4 muestras adicionales respecto a las 15 con todo el 1q ganado (28% en total); y las que afectaban al 1q distal (1q31.3-1q44; 196.24 - 245 mb desde el p-ter), presentes en 6 muestras adicionales (31% en total). Una tercera región queda definida en el espacio genómico comprendido entre ambas regiones, pero no fue considerada de relevancia puesto que en los estudios de correlación con los datos clínicos (ver siguiente apartado) alcanzó una significación estadística menor, indicando que

esta región es tan sólo el producto del solapamiento de las regiones proximal y distal y que carece por tanto de relevancia biológica y clínica *per se* (Figura 14).

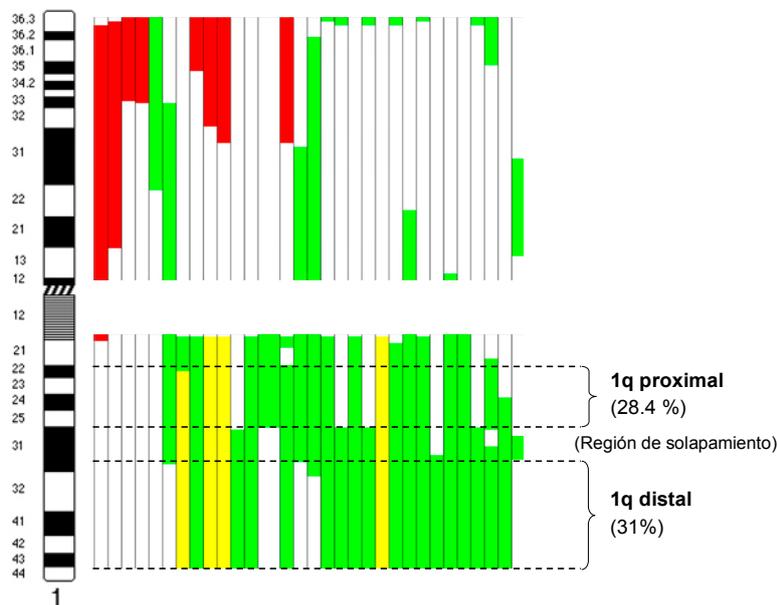


Figura 14. Localización genómica y extensión de las regiones principales de ganancia en 1q. El color rojo representa pérdidas de material genómico, el color verde ganancias y el color amarillo amplificación (dos copias extra o más).

Los parámetros Porcentaje de Genoma Alterado (PGA) y Número Total de ANC (NTA) fueron calculados para estimar la presencia global de ANC en cada muestra, tal y como se explica en el apartado de Material y Métodos 3.5. Mientras el primero (nunca usado previamente en la literatura de SE) es un indicador fiable del total del genoma afectado por ANC en cada muestra, el segundo (usado por varios autores previamente) no tiene en consideración la notable variación en tamaño de las ANC y es por tanto un estimador menos confiable. De hecho, ambos parámetros tienden a discrepar como se puede observar en la Figura 15B, en la que se representan sus respectivos valores en cada una de las líneas celulares analizadas por aCGH BAC-array.

De los valores de estos parámetros se deriva que:

- Alrededor de un 16% de las muestras tumorales analizadas no presentan ninguna ANC y cerca de un 22% tiene menos de un 0.1% del genoma alterado. Estos valores coinciden con los de trabajos previos, que presentaron una cifra de un 80-90% de tumores afectados por ANC.

- Las muestras tumorales y líneas celulares analizadas presentan tamaños medios de ANC similares (mediana 61.8% frente a 47.3%, respectivamente), pero difieren en PGA (mediana 10% frente a 28%) y NTA (mediana 5 frente a 15.5), siendo ambos parámetros inferiores en las muestras tumorales (Prueba U de Mann-Whitney, p-valor < 0.01 en ambos casos, Figura 15A). Es decir, aunque las ANC de ambos tipos de muestras tienen tamaños parecidos, las líneas celulares están más afectadas globalmente por las ANC.

- Existe una gran variabilidad en la presencia global de ANC en las muestras tumorales, abarcando desde muestras sin ANC hasta otras con casi un 70% del genoma alterado.

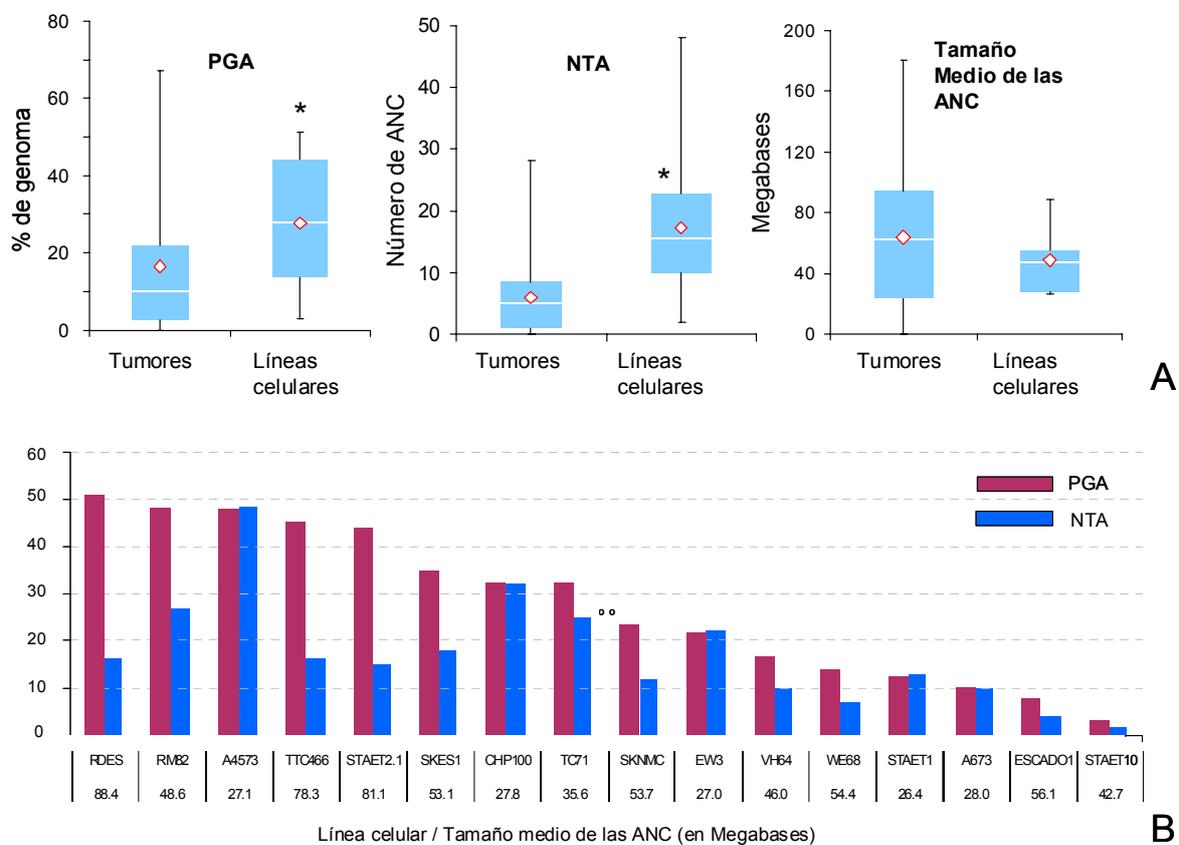


Figura 15. Comparación de los parámetros globales de ANC en muestras tumorales y líneas celulares de SE. (A) “Boxplots” de los parámetros PGA (izquierda), NTA (centro) y tamaño medio de las ANC (derecha), que revelan la mayor presencia de ANC en las líneas celulares respecto a los tumores de SE. (B) La comparación de los parámetros PGA y NTA en el panel de líneas celulares de SE demuestra que estos dos parámetros no siempre coinciden. Como ejemplo, la línea celular RDES tiene un NTA bajo-medio pero es la que mayor PGA tiene de todo el panel. El eje Y representa en una misma escala tanto el PGA como el NTA. Además, los rótulos del eje X muestran el tamaño medio de las ANC en cada línea celular.

Considerando las ANC de manera individual (ANC que afectan a regiones genómicas concretas), cabe destacar especialmente que la delección del *locus* de CDKN2A, aunque una de

las ANC más frecuentes en líneas celulares (62.5% de ellas la presentan), sólo fue detectada en 8 de las muestras tumorales (12%), en 5 de ellas en forma de microdelección de un único BAC. Respecto a este punto, la Figura 16 muestra la capacidad del BAC-array utilizado de detectar las microdelecciones de este *locus* gracias a la presencia de la sub-librería genómica dirigida contra los *loci* de los genes tradicionalmente involucrados en cáncer (ver apartado 3.4.1). De manera similar, las deleciones que englobaron al *locus* de TP53, aunque de nuevo frecuentes en líneas celulares (37.5%), fueron infrecuentes en las muestras tumorales (13.4%). Es de resaltar además que la mayoría de estas deleciones afectaron a una sola de las copias alélicas (deleciones hemicigóticas) y por tanto no pudieron suponer *per se* la pérdida de función de estos supresores tumorales, constituyendo a lo sumo un primer “hit” en esa dirección.

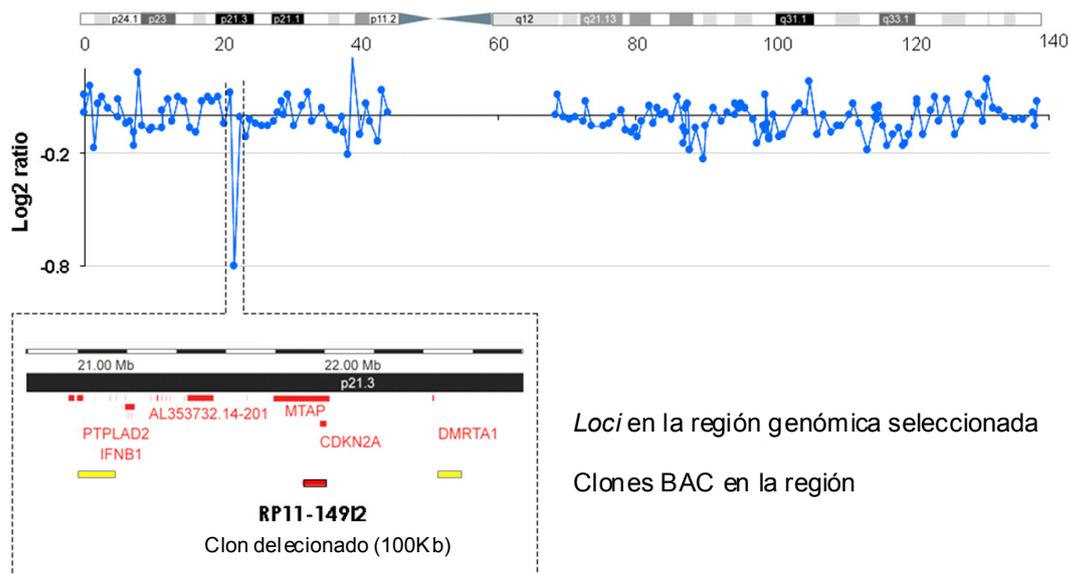


Figura 16. Microdelección del *locus* CDKN2A detectada mediante aCGH BAC-array en la línea celular STAET2.1. La imagen muestra el ideograma del cromosoma 9 y la correspondiente gráfica de aCGH BAC-array, en la que el eje Y representa el valor \log_2 ratio y el eje X la posición genómica en mb desde el p-ter. El recuadro muestra una ampliación de la región 9p21.3, los *loci* de la región (incluido CDKN2A) y los clones BAC del microarray que la cubren, representados a tamaño escalado y en su respectiva posición genómica.

Se definieron las regiones mínimas de solapamiento (RMS) con tamaño inferior a 10 mb y frecuencia superior a 35% (considerando conjuntamente las muestras de tumores y líneas celulares). Las RMS son las zonas genómicas mínimas en tamaño de una cierta región recurrentemente alterada en número de copias y son muy útiles en la definición de breves listas de genes candidatos cuya ganancia o pérdida de función contribuye de manera más

relevante a la oncogénesis y/o desarrollo tumoral. Las RMS así definidas se resumen en la Tabla 2.

<i>Tipo de ANC</i>	<i>Región Genómica</i>	<i>Posición (mb del p-ter)</i>	<i>Tam. Máx. (mb)</i>	<i>Frec. en líneas cel.</i>	<i>Frec. en tumores</i>	<i>Locus principal</i>
Pérdida	3p21.1 – 3p14.2	51.8 – 62.2	10.4	50%	6%	FHIT
Pérdida	9p21.3	21.15 – 22.5	1.35	62.5%	12%	CDKN2A
Pérdida	16q23.2	79.21 – 80.25	1.04	56.25%	22.4%	(Figura 17)
Pérdida	17p13.3 – 17p13.1	1.5 – 7.5	6	37.5%	13.4%	TP53
Ganancia	8q24.13 – q24.21	123.5 – 130.34	6.84	93.75%	39%	MYC

Tabla 2. Evaluación de las RMS de frecuencia superior al 35% y tamaño inferior a 10 mb (Tam. Máx: tamaño máximo de la RMS calculado considerando el espaciado entre el primer/último BAC alterado y los inmediatamente contiguos no alterados; Frec: frecuencia).

Cabe resaltar que la RMS del cromosoma 8, situada en 8q24.13 – q24.21 y de 6.8 mb de tamaño, se encontró amplificada en 3 líneas celulares y ganada en todas ellas salvo en una (93.75% de líneas celulares), y además constituye la ANC más frecuente en tumores (39% de tumores). Esta RMS comprende, entre otros, el *locus* del gen MYC (o *c-myc*), una de las dianas de la fusión EWS-ETS mejor caracterizada.

Para delimitar con mayor precisión la RMS situada en 16q23.2, una delección de sólo 3 BAC detectada en la línea celular SKES1, se realizó un estudio de aCGH SNP-array que permitió delimitar una RMS de tan sólo 1.04 mb, con márgenes de error muy reducidos (4982 y 12603 pb). La región es de un gran interés puesto que la pérdida de material genómico en 16q constituye la delección más frecuente en SE y la RMS delimitada permite definir una lista de genes candidatos muy restringida (Figura 17).

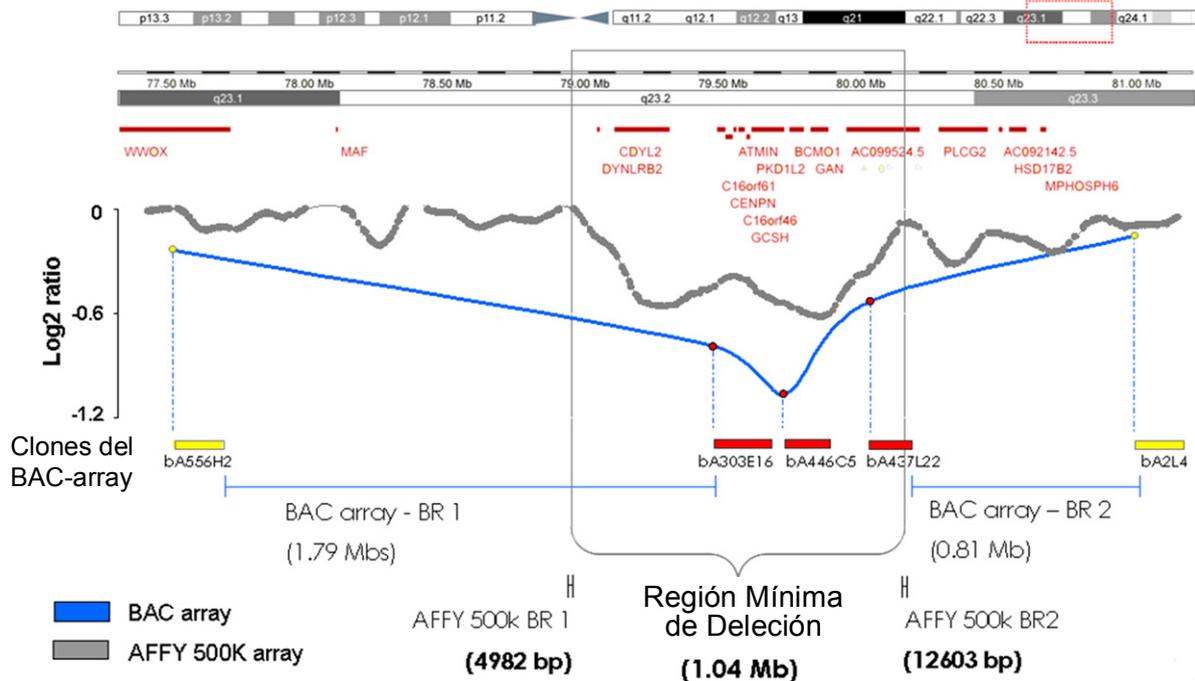


Figura 17. Microdeleción en 16q23.2 que define la RMS de esta región, detectada en la línea celular SKES1 mediante aCGH BAC-array y delimitada de manera precisa mediante aCGH SNP-array, restringiendo su tamaño a 1.04 mb. Los genes candidatos comprendidos en la RMS pueden observarse en la región encuadrada. La región genómica representada corresponde a la zona marcada en rojo sobre el ideograma del cromosoma 16. Las gráficas de BAC-array y SNP-array están representadas a escala y expresadas en valor log2ratio. Se incluyen asimismo los clones BAC del BAC-array que abarcan la región, también representados a escala. BR: “breakpoint region” o región de ruptura.

Dado que la línea celular RM82 es la que presenta unos mayores niveles proteicos de DTL (ver más adelante, apartado 4.5 y Figura 27) y ya que además está, a diferencia de las otras 2 líneas celulares usadas en los ensayos funcionales, afectada por una clara y marcada inestabilidad cromosómica (Figura 19), se caracterizaron con mayor detalle sus aberraciones genómicas aplicando la técnica COBRA-FISH¹²⁴. El resultado confirmó su complejo cariotipo, con numerosos reordenamientos cromosómicos (no detectables por aCGH) y una ploidía cercana a 3n (58 a 63 cromosomas en 18 metafases evaluadas). En la Figura 18 se muestra este resultado, mientras que la Figura 19 incluye el perfil de aCGH (completamente procesado), siendo ambos altamente coincidentes. Se observa como la 1qG es el resultado de una translocación derivativa al cromosoma 16, t(1;16)der(1)t(1;1). La línea celular RM82 tiene por tanto 4 copias genómicas de DTL. Además, la translocación t(21;22) confirma la presencia de la variante de la fusión génica EWS-ERG.

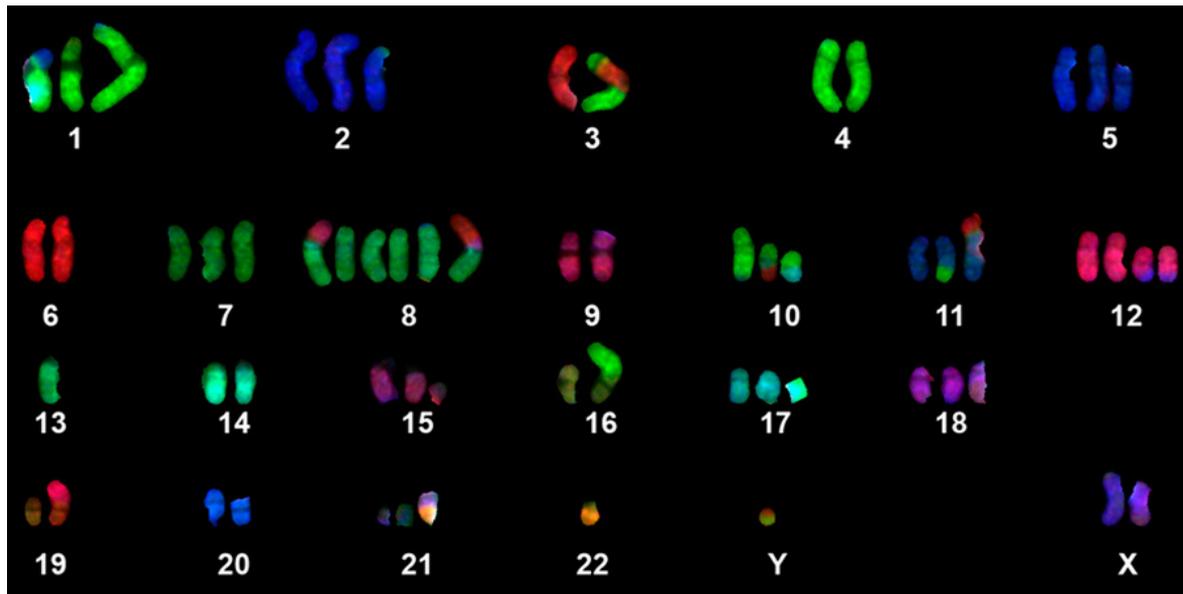


Figura 18. Cariotipo de la línea celular RM82, caracterizado mediante COBRA-FISH. Resumen del cariotipo (basado en la evaluación de 18 metafases): 58-63 <3n-> X.del(X),der(Y)t(Y;3),der(1)t(1;5),der(1)t(1;16)der(1)t(1;1),der(2)t(2;8)del(2)(q3?),der(2)t(2;2),der(3)t(Y;4)t(Y;3)t(1;3),-3,-4,-5,-6,+der(8;17)(q10;p10)t(9;17),+der(8)t(2;8),+der(8;18)(q10;q10)t(9;18),-9,del(9)(p),der(10;17)(p10;q10),der(10)t(3;10),der(11)t(4;11),der(11)t(Y;11)t(Y;3),der(12)t(12;21),+der(12)t(12;21),-13,-14,del(15)(q1?)-16,der(16)t(1;16)(p2?:q2?),der(18)t(15;18),-19,der(19)t(6;19),add(21)(q2?),der(21)t(21;?:22),-22,-22[cp18]

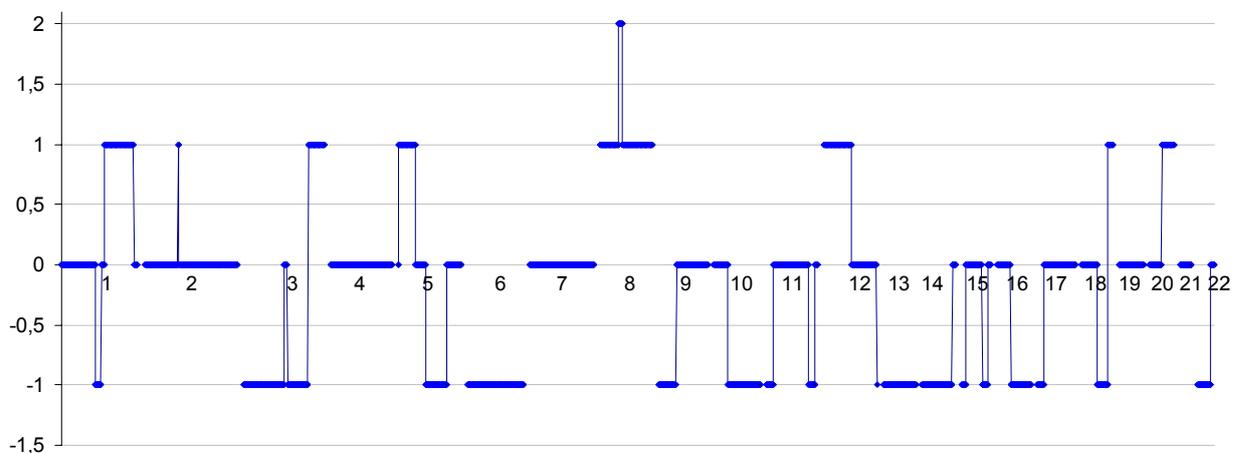


Figura 19. Perfil de aCGH de la línea celular RM82. El eje Y representa los valores del algoritmo de "calling" (-1 = pérdida, 0 = normal, 1 = ganancia, 2 = amplificación) mientras el eje X muestra el número de cromosoma.

4.2 Estudios del impacto de las ANC en los datos clínicos

Inicialmente se indagó el impacto de las ANC, consideradas de forma global, en los parámetros clínicos de los pacientes. A diferencia de lo realizado por otros autores¹⁰⁷⁻¹⁰⁸ se decidió no fijar un umbral de NTA arbitrario con el que dividir las muestras en dos grupos y

comprobar la supervivencia diferencial. En su lugar, se dividió la distribución poblacional del parámetro PGA en 5 intervalos, de acuerdo con percentiles poblacionales equidistantes, y se realizó un estudio estadístico de Kaplan-Meier log rango. Los intervalos que no mostraron diferencias entre sí fueron unidos. Esta prueba estadística mostró como cada grupo de pacientes comprendido en un intervalo de PGA tenía una peor supervivencia total asociada respecto a la de los grupos con un PGA menor (Figura 20, izquierda), indicando la existencia de una relación inversa entre este parámetro y la supervivencia total. Del mismo modo, las muestras se dividieron por percentiles para el parámetro NTA y se realizó un estudio análogo. Aunque la asociación de este parámetro con una peor supervivencia también fue buena, los intervalos no se ordenaron de una manera correlativa, ya que el segundo y tercer intervalo quedaron en posiciones invertidas entre sí (Figura 20, derecha).

Además, fue muy interesante hallar que el grupo de pacientes cuyas muestras tumorales carecían de ANC (PGA < 0.1%, 22% de los pacientes) presentó una supervivencia extraordinariamente favorable (cerca al 85% a 5 años).

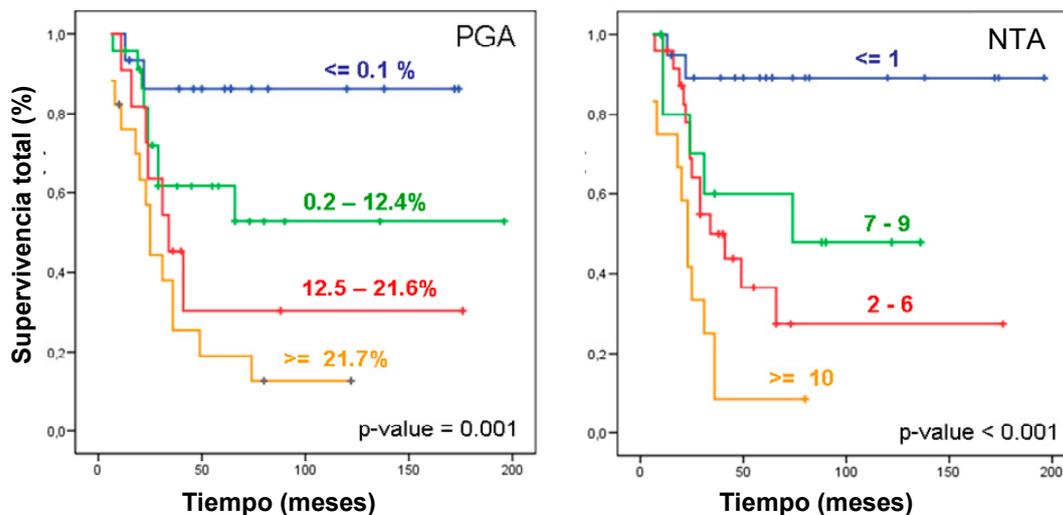


Figura 20. Impacto de las ANC, consideradas de forma global, en la supervivencia total de los pacientes de SE. Resultados del test Kaplan-Meier log rango de los parámetros PGA (izquierda) y NTA (derecha), organizados en intervalos (denotados sobre la correspondiente curva de supervivencia). El parámetro PGA (izquierda) muestra una mejor correlación con la supervivencia total que el parámetro NTA (derecha, nótese que los intervalos 7-9 y 2-6 tienen supervivencias invertidas respecto a lo esperado).

Una vez demostrada la relevancia clínica de las ANC consideradas de forma global, se continuó el estudio con la búsqueda de ANC individuales con impacto en los parámetros clínicos. Para ello, se hizo uso de un abordaje bioinformático de cribado exhaustivo usando el paquete de Bioconductor “CGHtest”, que permite 2 tipos de análisis:

- Aplicando el “script” “CGHpermutations”, se compararon las muestras tumorales divididas en subgrupos clínicos. El programa realiza un test no paramétrico de rangos y calcula una FDR a partir de 10000 permutaciones para identificar cualquier ANC diferencialmente sobrerrepresentada en alguno de los subgrupos clínicos enfrentados. Como ejemplo, uno de los test comprendió la confrontación de los datos de aCGH BAC-array de los tumores primarios sin progresión de la enfermedad (sin recaída), con los de aquellos que sí tuvieron recaída (ya fuera por metástasis o por reaparición del tumor en la localización primaria), detectando la presencia diferencial de la 1qG en los tumores primarios del segundo grupo (pacientes con recaída).

- Utilizando el “script” “CGHlogrank”, se encontraron las ANC con impacto en la supervivencia total o libre de enfermedad de los pacientes. Este programa realiza un test Kaplan-Meier log rango para cada una de las regiones con ANC e implementa el cálculo de la FDR a partir de 10000 permutaciones.

Utilizando estas aproximaciones informáticas se hallaron diversas ANC con impacto clínico, según los resultados descritos en la Tabla 3. De todas ellas la 1qG, presente en un 28-31% de los tumores (dependiendo de la región considerada), fue seleccionada para posteriores estudios ya que:

- Fue la segunda ANC con mayor incidencia en el grupo muestral (28-31% de tumores la presentaron).

- Fue la única ANC hallada diferencialmente presente en los tumores con progresión de la enfermedad (presente en los tumores de 22 de los 40 pacientes con recaída y tan sólo en 3 de los 26 sin ella; p -valor < 0.01 , FDR < 0.05).

- Fue la ANC que demostró un mayor impacto en la supervivencia de los pacientes (Figura 25).

Más en concreto, teniendo en cuenta las diversas regiones en las que el algoritmo “CGHregions” dividió al brazo 1q (Figura 14), tanto las regiones proximal como distal demostraron los valores más bajos de los estadísticos mientras que la región intermedia, aunque también significativa, presentó valores claramente mayores y por tanto menos significativos.

Además, el impacto detectado de la 1qG en los datos clínicos es independiente del parámetro PGA. Esta posibilidad fue descartada mediante una prueba U de Mann Whitney, en la que se compararon los valores de PGA de los tumores con 1qG (considerando aquellas muestras tumorales con $PGA > 0\%$) frente al resto (tumores con 1q normal, 1qN) y que no alcanzó significación estadística. Sin embargo, el script “CGHPermutations” detectó varias ANC diferencialmente asociadas con 1qG (Tabla 3, condición testada “tumores con 1qG vs. tumores con 1qN”): la ganancia del cromosoma 12 y del cromosoma 20 y la pérdida de brazo 16q. Esta última asociación confirma las observaciones previas que describen translocaciones no balanceadas entre ambas cromosomas, también detectada en el estudio de COBRA-FISH de la línea RM82 (Figura 18)¹³⁴. En cualquier caso, de acuerdo con los análisis Kaplan Meier log rango realizados, no existen diferencias en la supervivencia total de los pacientes con tumores que albergan tan sólo la 1qG respecto a los que muestran la 1qG junto con alguna de estas ANC acompañantes (dato no mostrado).

Tal y como era esperable debido a la asociación detectada entre la 1qG y la recaída o progresión de la enfermedad (el principal factor pronóstico en SE), los análisis multivariados de Cox no identificaron a esta ANC como un factor pronóstico independiente.

Respecto al impacto de otras ANC en los parámetros clínicos es especialmente destacable la asociación de la ganancia del cromosoma 8 con la presencia de metástasis en pulmón (Tabla 3, condición testada “Metástasis a sitios comprendiendo el pulmón vs. metástasis a sitios no pulmón”). Es decir, se observa una presencia diferencial de la ganancia del cromosoma 8 en los tumores primarios de los pacientes que, tanto si se trata de metástasis única como de metástasis múltiple, tienen metástasis en pulmón: la ganancia del cromosoma 8 está presente en los tumores primarios de 17 de 24 pacientes (71%) con metástasis en pulmón. De estos, entre los pacientes con metástasis única en pulmón, 13 de 15 (86%) tienen esta ANC en el tumor primario, mientras que de entre los que desarrollaron metástasis múltiple con al menos una de ellas en pulmón, 4 de 9 (44%) presentan esta ANC. Por contra, solo 1 paciente de 12 (8.3%) mostró ganancia del cromosoma 8 en el tumor primario sin haber desarrollado metástasis pulmonar (en el resto de casos, metástasis a otros órganos).

Tipo de test: "CGHPermutations"		
Condición clínica testada	ANC significativas	p-valor / FDR
Tumor primario con localización ósea vs. no ósea	NS	
Progresión de la enfermedad vs. no progresión	Ganancia de 1q	0.01 / 0.046
Metástasis vs. no metástasis	NS	
Metástasis al diagnóstico vs. nueva metástasis	NS	
Metástasis única vs. metástasis múltiple	NS	
Metástasis a sitios comprendiendo el pulmón vs. metástasis a sitios no pulmón	Ganancia del 8p (todo el brazo)	0.017 / 0.03
	Ganancia del 8q (todo el brazo)	0.002 / 0.005
Pacientes menores de 18 años vs. mayores de 18	Pérdida del 16p (p-ter – 2 mb)	0.03 / 0.057
Pacientes menores de 24 años vs. mayores de 24	Pérdida del 16p (p-ter – 2 mb)	0.026 / 0.066
Líneas celulares vs. muestras tumorales con PGA > 1%	Ganancia del 8q (124 mb – 129 mb desde p-ter)	0.003 / 0.06
	Pérdida del 3p (40 mb – 75 mb desde p-ter)	0.004 / 0.02
	Pérdida del 9p (p-ter – 24 mb desde p-ter)	0.005 / 0.014
Tumores con 1qG vs. tumores con 1qN	Ganancia del cr. 12 (completo)	0.005 / 0.005
	Ganancia del cr. 20 (completo)	0.002 / 0.002
	Pérdida del 16q (45 – 90 mb desde p-ter)	0.02 / 0.07

Tipo de test: "CGHlogrank"		
ANC	Regiones cromosómicas significativas (mb desde p-ter)	p-valor / FDR
<i>Supervivencia Total</i>		
1qG	141.4 – 152.8	0.004 / 0.02
	154.6 – 180.3	< 10 ⁻³ / < 10 ⁻³
	183.4 – 195.0	0.001 / 0.016
	196.2 – 245.0	0.002 / 0.021
Ganancia del cr. 20	0.2 - 53.4	0.01 / 0.035
	54.5 – 61	0.005 / 0.023
	61.5 - 62.8	0.003 / 0.02
Pérdida del cr. 22	15.6 – 50	0.014 / 0.078
<i>Supervivencia Libre de Enfermedad</i>		
1qG	141.4 – 152.8	0.002 / 0.21
	154.6 – 180.3	< 10 ⁻³ / < 10 ⁻³
	183.4 – 195.0	0.011 / 0.07
	196.2 – 245.0	0.004 / 0.03
Pérdida del cr. 10	1 – 135.5	0.005 / 0.06

Tabla 3. Resumen de los resultados del cribado bioinformático de las ANC con impacto clínico (mb: megabases; cr.: cromosoma; NS: no significativo, ninguna ANC enriquecida en las condiciones testadas).

También cabe resaltar que las diferencias de supervivencia debidas a la delección del 16q no alcanzaron significación estadística, en contraste con lo publicado en un trabajo anterior¹⁰⁶ y a pesar de ser la delección más frecuente en el grupo muestral estudiado.

Para finalizar esta sección, el estudio de clusterización de los tumores según sus patrones de ANC, utilizando el "script" "WECCA", no produjo coincidencias evidentes entre la clasificación debida a la similitud de los perfiles de ANC y las diversas clasificaciones debidas a los parámetros clínicos habituales (dato no mostrado), descartando la presencia de perfiles de ANC con significado clínico.

Como breve conclusión de este bloque de análisis, que luego extenderemos en la sección Discusión, las ANC son alteraciones secundarias clínicamente relevantes en SE y la 1qG tiene un profundo impacto en la supervivencia total de los pacientes, en gran medida debido a su asociación con el parámetro clínico recaída.

4.3 Estudios transcriptómicos y de su relación con las ANC

Para establecer el efecto a nivel transcripcional de la ANC con mayor relevancia clínica de las detectadas en las muestras del grupo 1 de tumores, la 1qG, se hizo un re-análisis de datos de microarrays de expresión pertenecientes a una serie independiente de 38 muestras tumorales de SE, a la que llamaremos grupo 2 de tumores (ver Material y Métodos, apartado 3.1). Dichos datos de expresión fueron tomados del repositorio GEO (“Gene Expression Omnibus”, NCBI) y son accesibles con el identificador GSE12102. Un subgrupo de 17 de estas muestras han sido además previamente estudiadas mediante aCGH SNP-array¹⁰⁸, comprendiendo 5 muestras con 1qG (en todos los casos ganancia del 1q completo). Los datos de aCGH fueron depositados por sus autores en el repositorio público CanGEM (www.cangem.org) y son accesibles con el identificador CG-SER-18.

Para incluir en el estudio al grupo completo de 38 tumores se creó una firma de expresión local (formada sólo por genes situados en 1q) característica de la 1qG. Dicha firma fue generada mediante un análisis SAM de expresión diferencial del subgrupo de 17 tumores divididos según su estado conocido de número de copias de 1q (1qG frente a 1qN). Se seleccionaron los 68 genes situados en 1q con mayores valores de “d-value” del análisis, comprendiendo los genes con “d-value” > 4. Esta firma (1qGSig) se usó para dividir las 38 muestras tumorales mediante clusterización no supervisada. Tal y como se muestra en la Figura 21, un cluster claro de 9 muestras tumorales compartía esta firma de expresión (5 de ellas las que se conocía que tenían 1qG gracias a los datos de aCGH SNP-array).

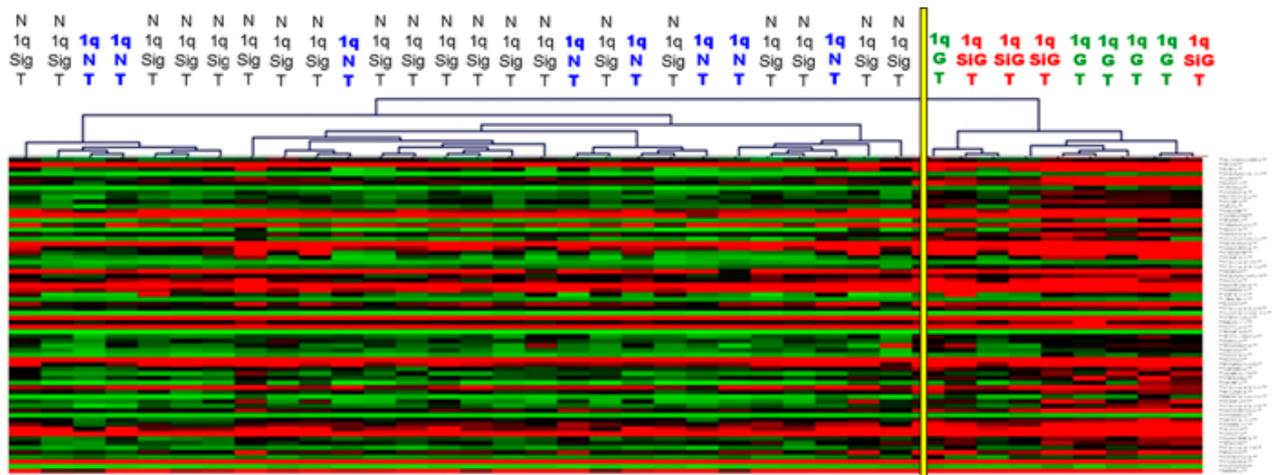


Figura 21. Definición de un grupo de muestras tumorales que comparten la firma de expresión local característica de la 1qG (cluster de la derecha separado por la línea vertical amarilla) mediante clusterización jerárquica. 1qGT: tumores con estado de número de copia conocido y 1qG; 1qNT: tumores con estado de número de copia conocido y 1qN; 1qGSig T: tumores que comparten la firma relacionada con 1q, 1qGSig; N 1qGSig T: tumores que no comparten la firma 1qGSig.

A continuación se aplicó el análisis de enriquecimiento de colecciones génicas (“Gene set enrichment análisis”, GSEA¹²⁹) a los datos completos de expresión de las 38 muestras tumorales divididas en 2 grupos de acuerdo con el perfil 1qGSig. Una explicación de este análisis se encuentra en la sección de Material y Métodos 3.11. Los resultados obtenidos al comparar el patrón diferencial de expresión de los tumores 1qGSig (calculado por el propio análisis GSEA) con las diversas colecciones génicas o “gene sets” de las bases de datos del análisis GSEA (MSigDB) presentaron una gran coincidencia entre sí:

- 1.- La comparación con las colecciones génicas de la MSigDB C2 dio como resultado el enriquecimiento en numerosas colecciones génicas derivadas de estudios experimentales relacionados con el ciclo celular. Como ejemplo, la colección más enriquecida en la lista diferencial de los tumores 1qGSig fue la llamada “G1 to S cell cycle reactome”, un grupo de genes implicados en la regulación de la transición de la fase G1 a la fase S del ciclo celular.
- 2.- La comparación con las colecciones génicas de la MSigDB C5, compuestas por genes agrupados según un término GO común, encontró enriquecimientos muy significativos en genes implicados en la regulación de la progresión del ciclo celular y del metabolismo del DNA.

3.- La comparación con las colecciones génicas de la MSigDB C3 (genes que comparten secuencias de respuesta a factores de transcripción o microRNA) descubrió el enriquecimiento en numerosos genes que son dianas directas del factor de transcripción E2F1, regulador central del ciclo celular, en especial de la transición de G1/S.

Todas estas coincidencias con las colecciones génicas de las bases de datos del análisis GSEA presentaron valores de NES y FDR muy significativos. En la Figura 22 se muestran ejemplos de las gráficas de enriquecimiento que se generaron en el análisis y la Tabla 4 resume las listas génicas enriquecidas más representativas.

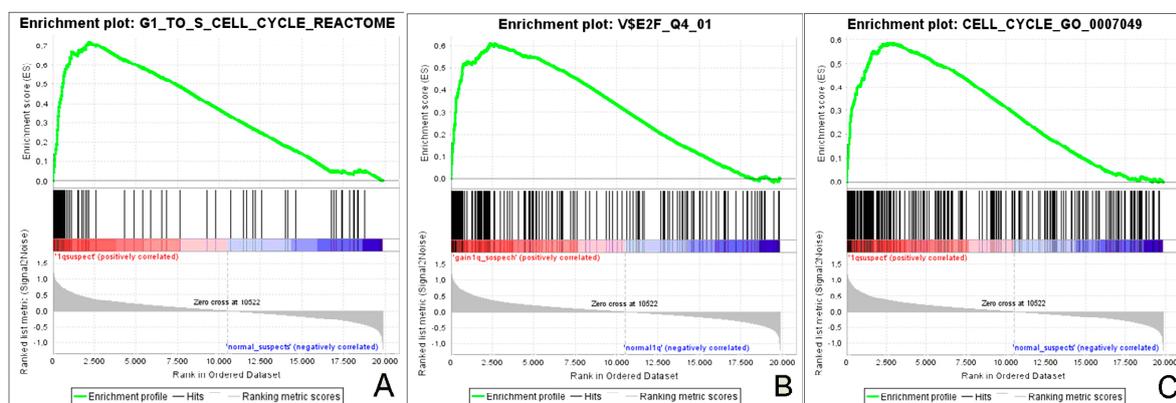


Figura 22. Gráficas de enriquecimiento generadas por el análisis GSEA, correspondientes a las colecciones génicas (A) “G1 to S cell cycle reactome”, de la MSigDB C2, (B) “E2F_Q4_01”, perteneciente a la MSigDB C3 y (C) “Cell_Cycle_GO_0007049”, de la MSigDB C5. El Eje Y representa el ES

La composición génica de estas colecciones, además de otros detalles, puede consultarse en la página web <http://www.broadinstitute.org/gsea/msigdb/genesets.jsp>, accediendo por el nombre representado en la primera columna de la Tabla 4.

Se descartó además que existiera un enriquecimiento génico debido a las ANC acompañantes de 1qG (ganancia de los cromosomas 12 y 20 y pérdida del 16q) mediante el análisis de GSEA utilizando la colección MSigDB C1, cuyas colecciones génicas agrupan a los genes por su posición genómica, y que sólo halló enriquecida la expresión de los genes situados en 1q.

Colecciones génicas de la MSigDB C2	Tamaño	NES	NOM p-val	FDR q-val
“G1_to_s_cell_cycle_reactome”	68	2.193	0.000	0.001
“Cell_cycle_kegg”	83	2.154	0.000	0.001
“Brentani_cell_cycle”	81	2.136	0.002	0.001
“Hsa04110_cell_cycle”	109	2.130	0.000	0.001
“Cell_cycle”	76	2.107	0.000	0.001
“Pyrimidine_metabolism”	57	1.989	0.000	0.005
“P21_any_dn”	34	1.967	0.000	0.006
“Ren_e2f1_targets”	38	1.945	0.000	0.006
“Schumacher_myc_up”	49	1.938	0.000	0.006
Colecciones génicas de la MSigDB C3	Tamaño	NES	NOM p-val	FDR q-val
“V\$E2F_Q4_01”	185	2.128	0.000	0.002
“V\$E2F_Q3_01”	181	2.109	0.000	0.001
“V\$E2F_Q6_01”	182	2.098	0.000	0.001
“V\$E2F1_Q4_01”	177	2.085	0.000	0.001
“V\$E2F1DP2_01”	185	2.009	0.002	0.003
“V\$E2F1DP1_01”	185	2.009	0.002	0.002
“V\$E2F4DP2_01”	185	2.009	0.002	0.002
“V\$E2F1_Q3”	191	2.001	0.000	0.002
“V\$E2F1_Q6_01”	192	1.998	0.002	0.002
“V\$MYC_Q2”	152	1.799	0.000	0.010
“V\$MYC_MAX_01”	215	1.763	0.000	0.015
Colecciones génicas de la MSigDB C5	Tamaño	NES	NOM p-val	FDR q-val
“Cell_cycle_go_0007049”	294	2.114	0.000	0.015
“Cell_cycle_checkpoint_go_0000075”	45	2.111	0.000	0.008
“Regulation_of_cell_cycle”	170	2.108	0.000	0.006
“Cell_cycle_phase”	157	2.105	0.000	0.004
“Mitotic_cell_cycle”	140	2.098	0.000	0.004
“Dna_replication”	93	2.095	0.000	0.004
“Cell_cycle_process”	180	2.077	0.000	0.004
“Regulation_of_dna_metabolic_process”	41	2.041	0.002	0.005
“G1_s_transition_of_mitotic_cell_cycle”	27	1.971	0.002	0.006
“Regulation_of_cyclin_dependent_protein_kinase_activity”	42	1.942	0.000	0.008

Tabla 4. Resultados más representativos del análisis GSEA, enfrentando el patrón diferencial de expresión de los tumores 1qGSig con las colecciones génicas de las MSigDB. Tamaño: número de genes que compone la colección génica; NES: valor de enriquecimiento normalizado; NOM p-val: p-valor nominal; FDR q-val: q-valor de la FDR.

Los resultados del análisis GSEA fueron validados mediante un programa similar pero de funcionamiento diferente, el análisis de las rutas Ingenuity. Este análisis encuentra también enriquecimientos de colecciones génicas en el dato de expresión diferencial pero esta vez las colecciones génicas de comparación han sido establecidas según un criterio bibliográfico. Se realizó un análisis SAM (dividiendo de nuevo los tumores según la firma 1qGSig) que se

utilizó como “input” del análisis Ingenuity (en concreto una lista de expresión diferencial de umbral $FDR < 0.02$). El resultado confirmó el hallado por el análisis GSEA: las funciones más significativamente enriquecidas obtenidas fueron ciclo celular y cáncer, mientras que las vías canónicas principalmente halladas enriquecidas fueron, entre otras, metabolismo de DNA (de purinas y pirimidinas), regulación del “checkpoint” G1/S y ubiquitinación de proteínas (Figura 23). Ya veremos más adelante lo significativo de este último resultado, no encontrado por el análisis GSEA.

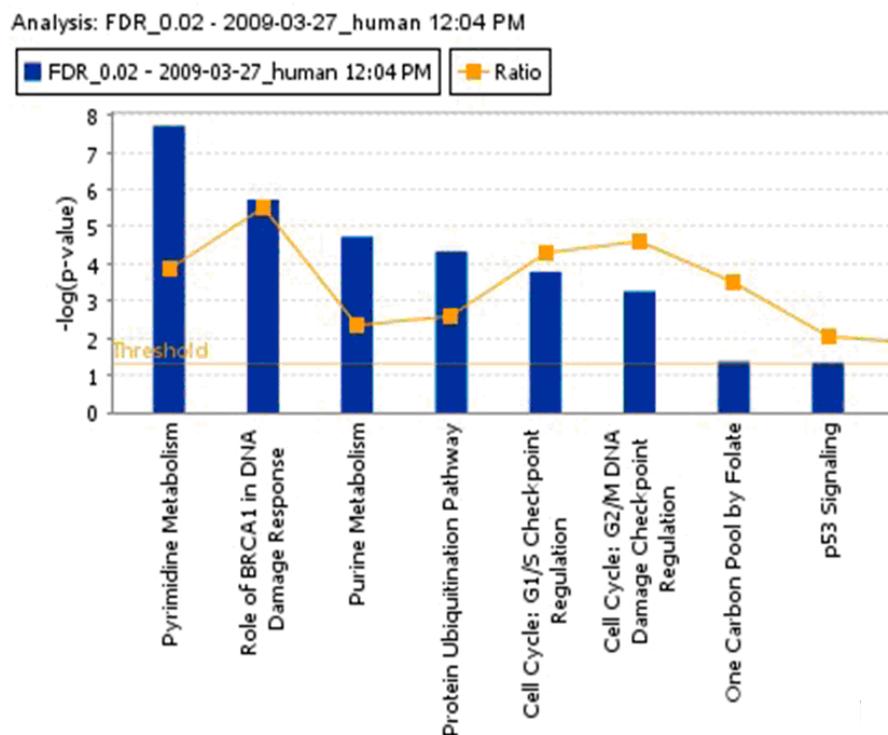


Figura 23. Resultado del análisis Ingenuity, mostrando las principales rutas canónicas enriquecidas en el perfil de expresión diferencial de los tumores 1qGSig. El eje Y muestra el valor de enriquecimiento, y el umbral considerado significativo (línea amarilla horizontal).

Además de estos análisis supervisados se llevó a cabo un estudio de clusterización jerárquica no supervisado, usando para ello los perfiles completos de expresión del grupo 2 de tumores. El agrupamiento obtenido coincidió de manera satisfactoria con la clasificación según el perfil 1qGSig, indicando este resultado el fuerte impacto de la ANC 1qG en el transcriptoma (Figura 24). Es especialmente interesante el hecho de que los tumores “mal clasificados” por la clusterización jerárquica resultaron ser tumores no 1qGSig incluidos dentro del cluster de los 1qGSig, lo cual podría estar indicando que estas muestras comparten el perfil de desregulación de ciclo celular pero por mecanismos distintos a la 1qG.

4.4 Selección de un gen candidato en 1q, DTL, y validación de sus niveles mediante Real Time RT qPCR

Hasta este punto los hallazgos respecto al fenotipo molecular asociado a la 1qG se restringían a predicciones bioinformáticas. Por ello, para demostrar empíricamente la relación entre la 1qG y la desregulación del ciclo celular asociada a esta ANC se establecieron los siguientes requisitos de selección de genes candidatos:

1. Localización en 1q
2. Sobre-expresión significativa en tumores 1qGSig según el parámetro “d-value”. Este parámetro combina la diferencia de expresión media entre los grupos confrontados con una medida de la desviación intra-grupo. En consecuencia su valor es indicativo no sólo de la diferencia de expresión entre los grupos comparados sino también de una expresión similar dentro de cada grupo y es por tanto el parámetro de expresión diferencial más significativo y confiable.
3. Incremento de expresión neto (“fold-change”) superior a lo esperado por el cambio de dosis génica (debida a la 1qG subyacente): “R-fold” (parámetro de “fold-change” del análisis SAM) superior a 3 (el esperado sería 1.5)
4. Función proteica claramente capaz de ejercer una regulación de la progresión del ciclo celular.
5. Función proteica involucrada además en cuantas más rutas canónicas del resultado del análisis Ingenuity como sea posible (al menos implicada en 3 de dichas rutas).

Sólo un gen cumplió convenientemente con todos estos requisitos: DTL, también conocido como CDT2, L2DTL2 o RAMP. Este gen cumple los requisitos propuestos de la siguiente manera:

- 1.- Se localiza en 1q, concretamente en 1q32.3, a 212.2 mb desde el p-ter.
- 2.- Ocupa el primer puesto de la lista del análisis de expresión diferencial SAM ordenada descendientemente por “d-value”, con el mayor valor de este parámetro (11.88) y por tanto

con la expresión diferencial más significativa y consistente de todos los genes diferencialmente expresados en las muestras 1qGSig.

3.- El incremento de expresión neto de DTL (“fold-change”) entre los tumores 1qGSig y los tumores no 1qGSig es de 4.78, ocupando el puesto 13° de la lista completa del análisis SAM ordenada por este factor y siendo el gen con el mayor “fold-change” de todos los situados en 1q.

4 y 5.- Su función está directamente involucrada en el control del ciclo celular: la proteína DTL forma parte del complejo proteína-ubiquitina ligasa CUL4/DDB1^{CDT2}, dentro del cual cumple una función de identificador de los sustratos que serán modificados con una cola de poliubiquitina por el complejo para ser posteriormente degradados por el proteasoma 26S¹³⁵⁻¹³⁶ (ver el apartado 5.9 para mayor información sobre DTL). Los complejos de este estilo utilizan distintas proteínas receptoras de sustrato cada una de las cuales reconoce a un subconjunto de sustratos diferentes. Cuando el complejo CUL4/DDB1 utiliza a DTL como receptor los sustratos reconocidos son, entre otros, p53, p21^{WAF1/CIP1} y p27^{KIP1} (reguladores de la transición G1/S)¹³⁷⁻¹⁴² y CDT1 (regulador de la replicación del DNA en la fase S)¹⁰⁶. Por lo tanto la función de DTL está involucrada en 4 de las rutas del análisis Ingenuity (transición G1/S, metabolismo de purinas y pirimidinas y ubiquitinación de proteínas).

La validación de la diferencia de expresión de DTL en las muestras tumorales se realizó mediante Real Time RT-qPCR a partir del RNA de un subgrupo de 14 de los 38 tumores (comprendiendo 7 tumores 1qGSig) del grupo 2 de tumores. Este ensayo confirmó un incremento de expresión (“fold-change”) de 3.9, considerando las medianas de cada grupo, y de 4.79 (exactamente el valor predicho por el microarray de expresión) considerando las medias (Figura 26).

4.5 Validación funcional del gen candidato DTL, in vitro e in vivo

Para comprobar la capacidad como regulador del ciclo celular del gen candidato seleccionado en líneas celulares de SE se llevó a cabo una aproximación metodológica mediante silenciamiento de su transcrito en líneas celulares con 1qG y sobre-expresión de DTL.

Antes de ello se realizó un estudio de los niveles del mRNA de DTL mediante Real Time RT-qPCR en un panel de líneas celulares de SE del cual existían datos previos de microarrays de expresión, realizados por un grupo colaborador¹⁴³ y accesibles en el repositorio “Array Express Archive” (www.ebi.ac.uk/microarray-as/ae/) con el identificador E-MEXP-1142. Los niveles obtenidos mediante ambas técnicas coincidieron de manera casi exacta, indicando la estabilidad de la expresión de este gen en las líneas celulares ensayadas, a pesar de que ambos estudios se hicieron con varios años de diferencia y partiendo de cultivos con número de subcultivaciones muy distintos (Figura 27).

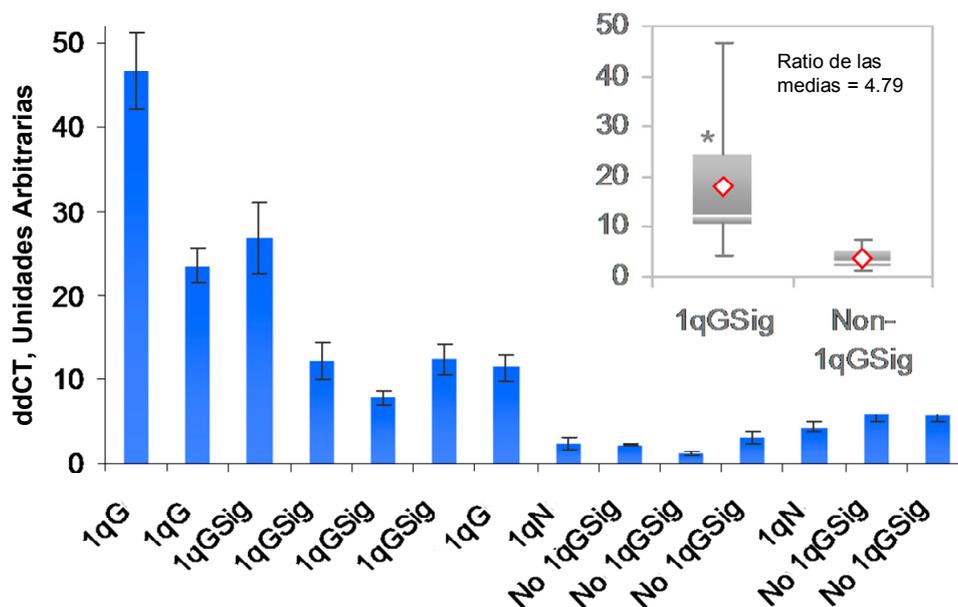


Figura 26. Validación por Real Time RT-qPCR de las diferencias en los niveles de mRNA de DTL entre los grupos de tumores 1qGSig y el resto de tumores. El eje Y muestra los valores de ddCT de DTL, normalizados con el gen de referencia TPT1 (media de duplicados). El cuadro interno muestra en un “boxplot” el resultado en los 2 grupos estudiados y el “fold-change” usando la media de cada grupo. 1qG: tumores con estado de número de copias conocido y 1qG; 1qN: tumor con estado de número de copias conocido y 1qN; 1qGSig: tumores que comparten la firma de expresión relacionada con la ganancia de 1q; No 1qGSig: tumores que no comparten la firma de expresión relacionada con la ganancia de 1q.

La línea celular RM82 demostró tener los niveles más elevados del mRNA de DTL. Otras 2 líneas celulares con 1qG, no incluidas en el estudio de microarrays de expresión, también demostraron niveles especialmente altos: TC32 y STAET10 (dato no mostrado).

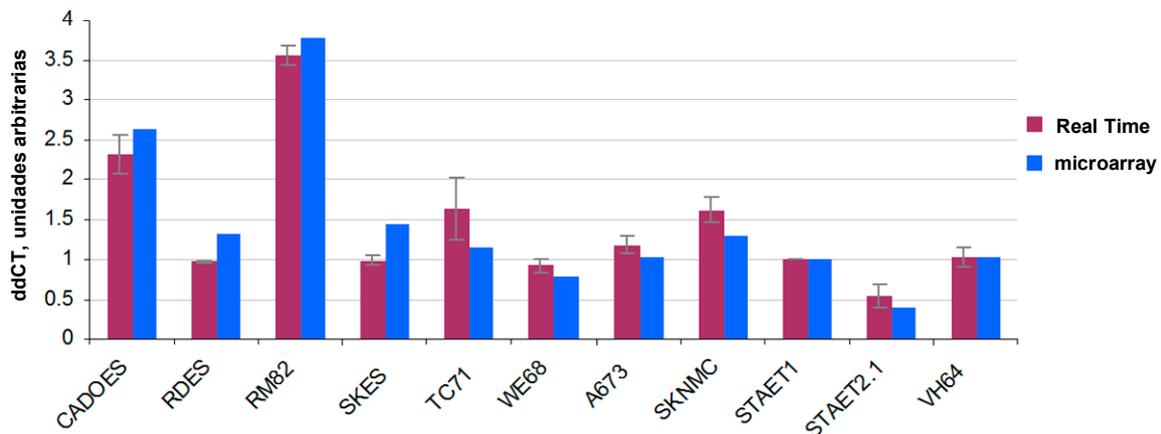


Figura 27. Comprobación mediante Real Time RT-qPCR de los niveles del mRNA de DTL en un panel de líneas celulares de SE y comparación con los valores de un microarray de expresión previo. El eje Y muestra los valores de ddCT normalizados respecto a GAPDH (media de duplicados).

Para el silenciamiento del gen candidato se seleccionó una colección de construcciones de shRNA, cada una de ellas dirigida contra una región diferente del transcrito de DTL. En la Tabla 5 se ofrecen los detalles acerca de estas construcciones, que están clonadas en el vector lentiviral pLKO.1. Asimismo, en todos los ensayos de silenciamiento realizados las células fueron transducidas en paralelo con una construcción pLKO.1-“non-targeting control” (NTC) o control no dirigido, que contiene una secuencia shRNA capaz de disparar la actividad de la maquinaria de silenciamiento celular sin afectar a los niveles de ningún transcrito en concreto, lo cual permite valorar los efectos inespecíficos de esta técnica.

Construcción shRNA	Secuencia de la construcción shRNA (en sentido 5'-3')	Exones de DTL cubiertos
sh280	GCTCCAATATGGAACATGTA /TACATGTTCCATATTGGGAGC	Exones 2 – 3
sh424	CCTGGTGAACCTAAACTTGTT /AACAAAGTTTAAAGTTCACCAGG	Exones 4 – 5
sh425	CTGGTGAACCTAAACTTGTTA /TAACAAGTTTAAAGTTCACCAG	Exones 4 – 5
sh1385	GCCTAGTAACAGTAACGAGTA /TACTCGTACTGTTACTAGGC	Exón 14
sh2495	CCGAGGATGAATGCTGTGTTT /AAACACAGCATTTCATCCTCGG	3' UTR
NTC	CAACAAGATGAAGAGCACCAA /TTGGTGCTCTTCATCTTGTTG	(NA)

Tabla 5. Información detallada de las construcciones shRNA usadas en el silenciamiento de DTL. NTC: control “non-targeting”, NA: no aplicable.

Inicialmente se comprobó la capacidad de las 5 construcciones para reducir la expresión de DTL en la línea celular RM82 mediante Western Blot. Como se observa en la Figura 28, las 5 redujeron en mayor o menor medida los niveles proteicos de DTL, así como todas indujeron apoptosis.

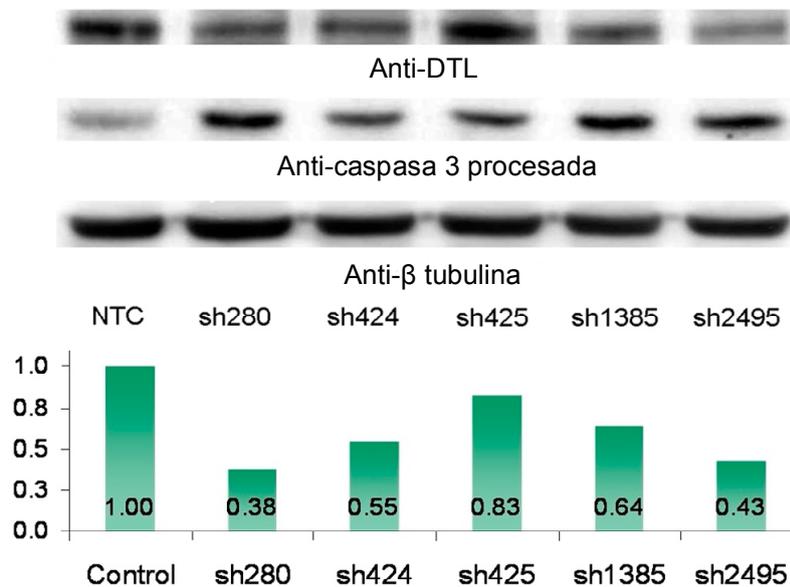


Figura 28. Western blot que muestra la reducción de los niveles proteicos de DTL conseguidos mediante cada una de las construcciones shRNA ensayadas. Cuantificación en el panel inferior: niveles de DTL normalizados respecto a los de β-tubulina, cuantificados con el programa “Image J” (<http://rsbweb.nih.gov/ij/>). Se observa asimismo una considerable inducción de apoptosis, representada por el incremento de caspasa 3 procesada. NTC: control “non-targeting”.

La eficiencia de la transducción lentiviral fue evaluada mediante el uso de un vector pLKO.1-turboGFP, con cuyo sobrenadante viral se transdujo la línea TC32 en dilución seriada (diluyendo el sobrenadante viral con un factor 1/10, desde 1 a 10^{-6} - 10^{-7}). Setenta y dos horas tras la transducción se observaron las células por microscopía de fluorescencia y se adquirieron mediante citometría de flujo. El porcentaje de población celular transducida fue en los distintos ensayos de transducción realizados del 100% de células para la condición transducida sin diluir el sobrenadante (Figura 29) y del 50-80% para la transducida con el factor de dilución 10^{-2} , con lo que todos los ensayos descritos (incluidos los de silenciamiento progresivo o gradual, ver más adelante) han sido llevados a cabo con poblaciones celulares totalmente transducidas.

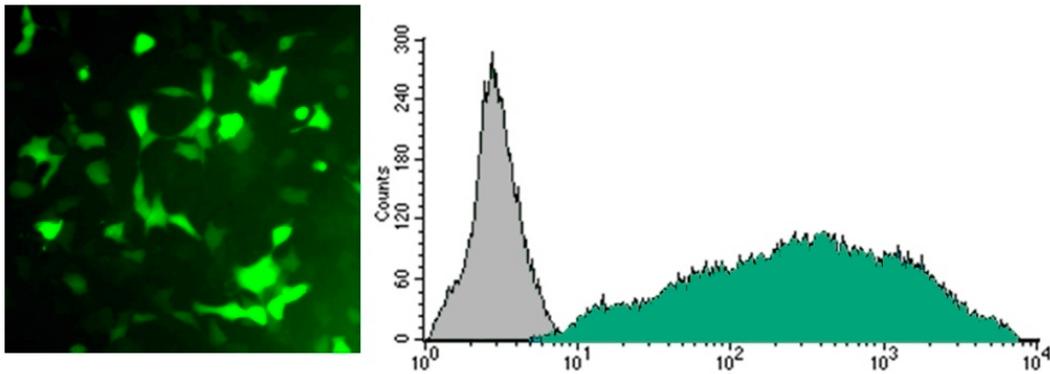


Figura 29. Comprobación de la eficiencia de transducción en una población de TC32 transducida con la construcción pLKO.1-turboGFP, mediante microscopía de fluorescencia (izquierda, magnificación 200x) y citometría de flujo (derecha). Las imágenes corresponden a la población transducida con el sobrenadante no diluido. La imagen de citometría representa en un histograma de fluorescencia una población no transducida (en gris) frente a la población transducida con el vector de GFP (en verde), donde el eje X representa la intensidad de fluorescencia en escala logarítmica.

Para el resto de ensayos se seleccionaron las dos construcciones que indujeron un mayor silenciamiento y que cubren un mayor rango de variantes transcripcionales, según la predicción bioinformática de la base de datos “Ensembl” (www.ensembl.org) de isoformas generadas a partir del *locus* de DTL mediante corte y empalme alternativo (o “splicing” alternativo). Las dos construcciones elegidas fueron la sh280 y la sh2495.

Los ensayos funcionales se realizaron en las líneas celulares RM82 y TC32, ambas con 1qG y altos niveles del transcrito de DTL. Se optó por una estrategia de silenciamiento progresivo, parecida a lo que se consigue con un sistema de expresión inducible: el sobrenadante viral de las células empaquetadores fue diluido con medio fresco (dentro del rango de eficiencia de transducción poblacional completa estimado con los ensayos de GFP) para conseguir sobrenadantes con diferentes cargas virales que al aplicarlos a la población celular consiguen varios niveles de silenciamiento, produciendo un efecto de silenciamiento gradual o progresivo. Cuarenta y ocho horas tras aplicar este método de transducción las células se tripsinizaron y sembraron y se sometieron a análisis de ciclo celular, apoptosis y proliferación. Esta aproximación metodológica de silenciamiento gradual presenta las siguientes ventajas:

- El estudio planteado tiene como objetivo comprobar el efecto de la eliminación de la sobre-expresión de DTL, no su supresión, puesto que muy probablemente cualquier célula en proceso de proliferación necesita unos niveles mínimos de la proteína, lo cual se consigue con este abordaje.

- El silenciamiento completo de la proteína tiene un efecto de parada de ciclo de alrededor del 90% de la población celular (detenida en G1) e inducción de apoptosis masiva, lo cual impide la realización de estudios de media-larga duración (Figura 30).

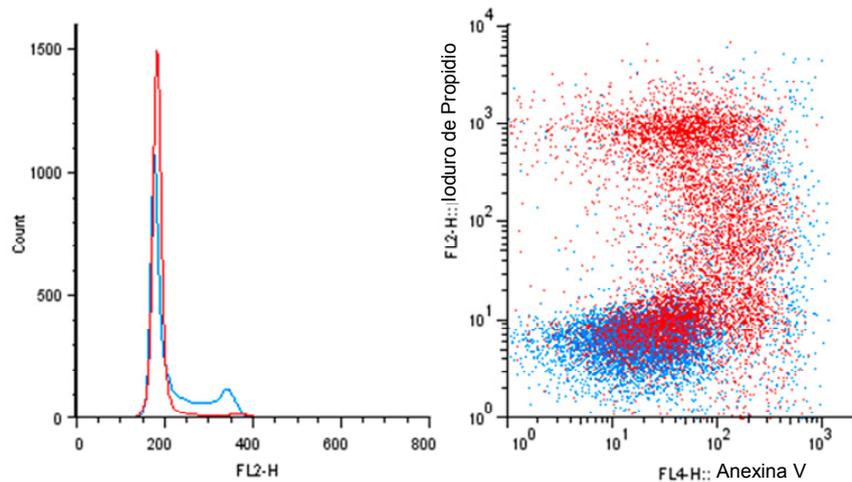


Figura 30. Análisis de ciclo celular y apoptosis mediante citometría de flujo tras el silenciamiento completo de DTL. La línea celular RM82, transducida con un sobrenadante viral muy concentrado de las construcciones sh280 y sh2495, fue sometida al estudio del ciclo celular (izquierda, histograma de señal del canal FL2, IP) y de apoptosis (derecha, “dotplot” de anexina V frente a IP), como se detalla en Material y Métodos, 48 horas tras la transducción viral. La población representada en azul corresponde a la transducción con el shRNA control (NTC) mientras la representada en rojo fue transducida con las construcciones shRNA contra DTL (sh280 y sh2495). Las poblaciones celulares comparadas tienen tamaños poblacionales iguales.

El silenciamiento gradual *in vitro* de DTL (expresado en la Figura 31 como el porcentaje de sobrenadante viral usado respecto a la condición sin diluir) resultó en:

- Una disminución gradual de la población celular en fase S y acumulación celular en fase G0/G1. La correlación entre el grado de silenciamiento y la disminución de la fase S es muy ajustada (valores de regresión lineal $R^2 = 0.97$).
- Un aumento progresivo de la apoptosis. La línea celular TC32 presentó una mayor inducción de apoptosis respecto a RM82. Además, en esta última línea celular el incremento de la apoptosis con el silenciamiento gradual parece cesar a partir de un cierto punto (a partir de la dilución del 50%), donde se satura.
- Una disminución progresiva de la proliferación celular, estudiada durante 5 días, en respuesta al silenciamiento creciente.

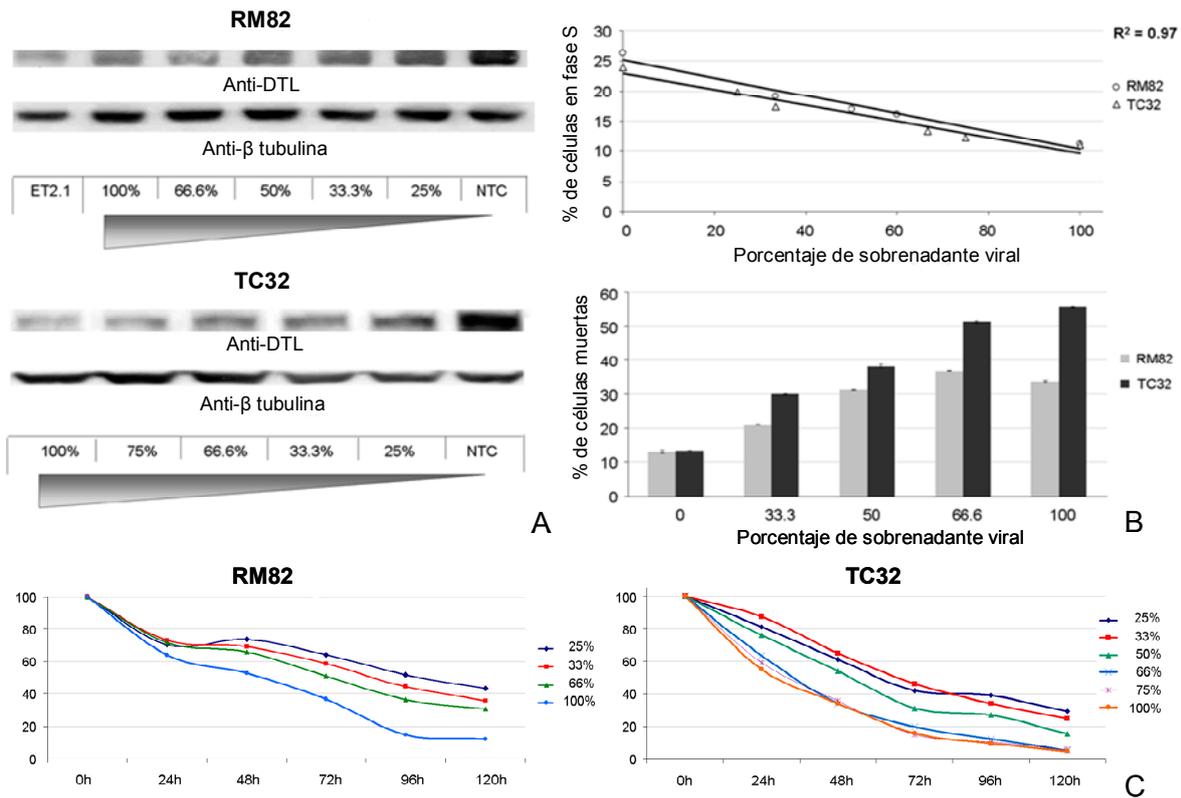


Figura 31. Ensayo de silenciamiento gradual de DTL en las líneas celulares RM82 y TC32. (A) Western blot mostrando la reducción gradual de los niveles de DTL; el triángulo gris marca la reducción del silenciamiento de izquierda a derecha. Se incluye en el panel superior el extracto proteico de la línea celular STAET2.1 como referencia, ya que esta línea celular no tiene IqG ni sobre-expresión de DTL (B) Descenso progresivo del % de población celular en fase S al incrementar el silenciamiento (arriba) y aumento paralelo de la apoptosis (abajo), media de duplicados en ambos casos, resultados de un experimento representativo. Las barras de error no se incluyen en el panel superior por motivos de claridad. (C) Descenso gradual de la proliferación celular al aumentar el silenciamiento (barras de error no incluidas por motivos de claridad). El eje X muestra el punto temporal ensayado mientras el eje Y representa el porcentaje de absorbancia del ensayo MTT de cada nivel de silenciamiento respecto a la absorbancia del control NTC correspondiente.

De manera similar, para los ensayos *in vivo* las líneas RM82 y TC32 fueron transducidas con un sobrenadante viral correspondiente a las construcciones shRNA sh280 y sh2495, diluido de forma gradual. Tras 48 horas las poblaciones celulares transducidas fueron inyectadas subcutáneamente en ratones NOD/SCID (el tumor control, transducido con el NTC, en el flaco izquierdo y el silenciado en el derecho). El rango de silenciamiento fue restringido a un mínimo del 50% del sobrenadante no diluido para evitar el sobrecrecimiento de poblaciones celulares no transducidas o transducidas con moi bajas, ya que la duración de este ensayo es notablemente superior a la del *in vitro*.

Los resultados del ensayo *in vivo* (Figura 32) confirmaron los obtenidos en el *in vitro* aunque las diferencias encontradas fueron algo más atenuadas. Cabe destacar sin embargo que la

condición de mayor silenciamiento produjo una casi completa supresión del crecimiento tumoral en varios de los animales. Para el estudio estadístico los datos fueron primeramente pareados, dividiendo el peso final del tumor control por el del tumor silenciado (tumores de un mismo animal) y los ratios obtenidos fueron logaritmizados ($\log_2\text{ratio}$) para mantener la normalidad de los datos, lo cual se comprobó con la prueba Kolmogorov-Smirnov ($p\text{-valor} < 0.05$). El test estadístico ANOVA descartó una diferencia en el ratio de pesos debida al factor “línea celular” por lo que los datos de ambas líneas celulares fueron considerados conjuntamente, aumentando así los tamaños muestrales y por tanto la significación estadística final. El estudio de ANOVA encontró diferencias estadísticamente significativas en los datos debidas al efecto del silenciamiento ($p\text{-valor} = 0.05$). Las pruebas *post-hoc* (Scheffe, Sidak y Bonferroni) confirmaron la diferencia de $\log_2\text{ratio}$ entre el nivel de silenciamiento mayor y menor ($p\text{-valor} < 0.05$), aunque no las del resto de comparaciones múltiples, que mostraron valores sólo marginalmente significativos (comparación del nivel intermedio de silenciamiento con el resto). Finalmente, un estudio t de Student de dos colas confirmó el efecto del shRNA en cada nivel de silenciamiento respecto al ratio esperable (ratio = 1) debido a la ausencia de diferencias ($H_0 = 0$, ya que el $\log_2\text{ratio}$ de 1 = 0) en todas las condiciones ($p\text{-valor} < 0.05$) salvo en la de menor silenciamiento.

4.6 Comprobación del efecto del silenciamiento de DTL en los sustratos del complejo CUL4/DDB1^{CDT2}

Una vez comprobado el importante papel ejercido por DTL en el control de la transición G1/S del ciclo celular en las líneas celulares de SE con 1qG y sobre-expresión de este gen, se realizó un nuevo silenciamiento para confirmar el efecto sobre los niveles proteicos de los sustratos del complejo CUL4/DDB1^{CDT2}. Para ello se transdujeron con un sobrenadante viral no diluido 3 líneas celulares de SE: las anteriormente usadas RM82 y TC32, y STAET10, una línea celular cuya única ANC es 1qG y que presenta también niveles medios-altos de DTL.

Los extractos proteicos de las 3 líneas celulares silenciadas (utilizando en la transducción las construcciones sh280 y sh2495 conjuntamente) y de sus respectivos controles NTC se sometieron a Western blot, usando anticuerpos contra reguladores de la transición G1/S, incluyendo varios de los sustratos conocidos del complejo CUL4/DDB1^{CDT2} (Figura 33).

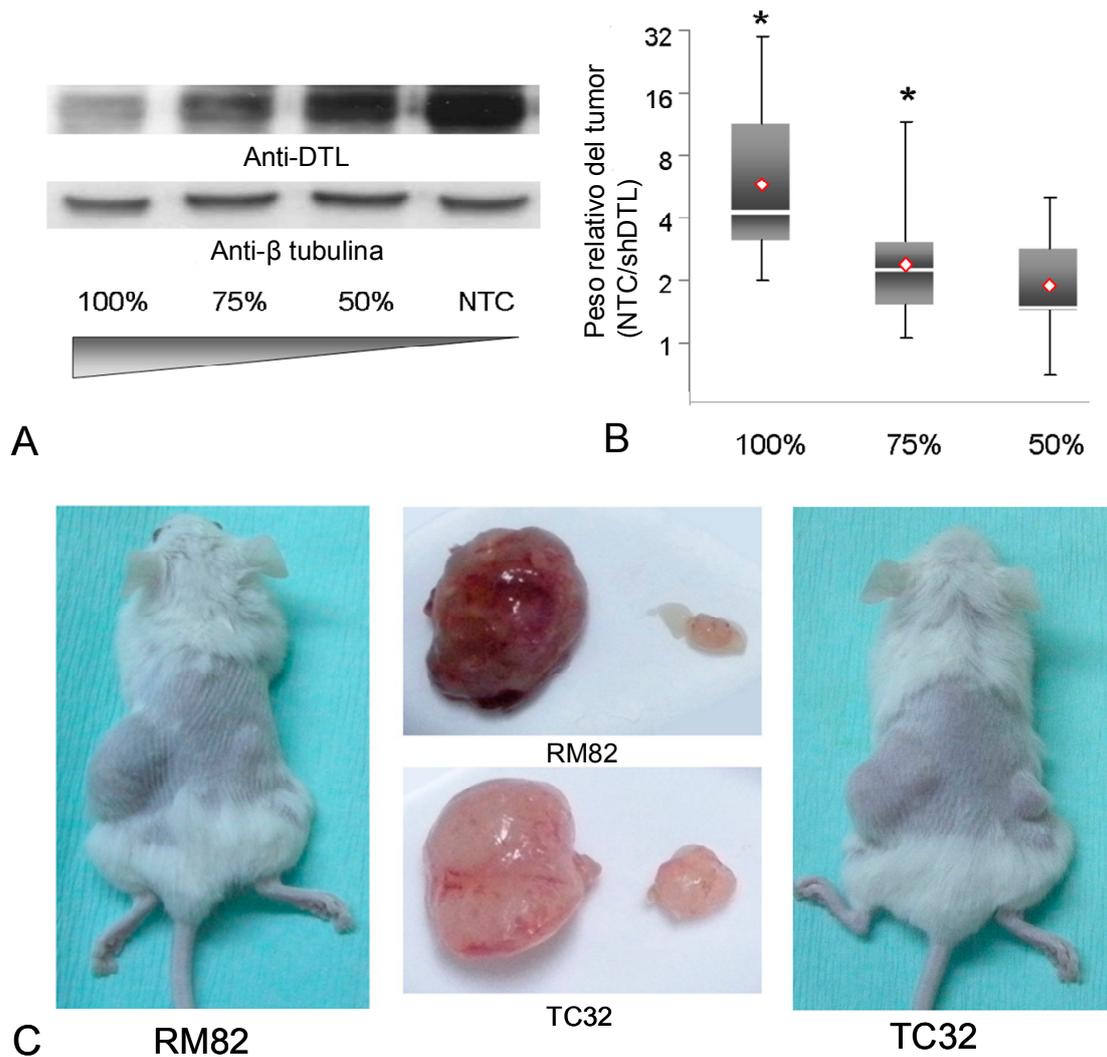


Figura 32. Resultados del ensayo *in vivo* de silenciamiento de DTL, que confirma el efecto inhibitor de la proliferación de las líneas celulares de SE. (A) Western blot del silenciamiento progresivo de DTL en RM82. (B) “Boxplot” representando el tamaño tumoral relativo de los tumores. Los valores representados están logaritmizados (\log_2 ratio) pero la escala del eje Y está expresada en valores naturales. Los asteriscos reflejan el resultado del t-test para una muestra ($H_0 = 0$) de cada condición de silenciamiento. (C) Imágenes correspondientes a los tumores de dos animales de la condición de mayor silenciamiento (100% de sobrenadante viral). En cada caso el tumor de la izquierda o flanco izquierdo corresponde al control NTC y el de la derecha o flanco derecho al del tumor silenciado (shRNA: sh280 + sh2495).

Los resultados obtenidos revelan un incremento dramático de los niveles proteicos de $p21^{WAF1/CIP1}$ y $p27^{KIP1}$, así como un aumento no tan marcado de $p53$, confirmando los trabajos previos que identificaron a estas proteínas como sustratos del complejo. Los cambios no obstante no son uniformes en las 3 líneas, reflejando los distintos entornos moleculares que representan (Figura 33):

- La línea celular STAET10, la menos afectada por ANC, presentó incrementos de los 3 sustratos del complejo.

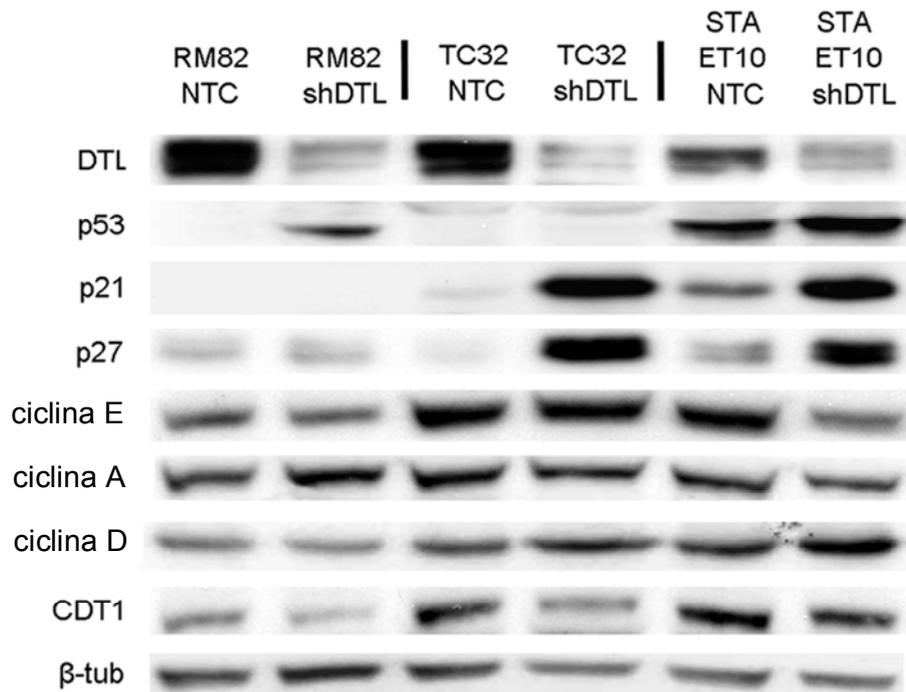


Figura 33. Western blot de 3 líneas celulares de SE con anticuerpos contra reguladores de la transición G1/S, tras silenciamiento completo de DTL. shDTL: condiciones con DTL silenciado usando las construcciones sh280 y sh2495; NTC: control “non-targeting”; β -tub: beta tubulina, usada como control de carga.

- La línea celular TC32, con tan sólo las ANC 1qG y ganancia del cromosoma 8¹⁴⁴, demostró grandes incrementos de p21^{WAF1/CIP1} (posiblemente independientes de p53 puesto que los niveles de esta proteína no son detectables mediante esta técnica) y de p27^{KIP1}.

- La línea celular RM82, la segunda más afectada por ANC de las 16 estudiadas por aCGH BAC-array, parece carecer de expresión de p21^{WAF1/CIP1} y sólo es detectable en sus extractos proteicos un incremento moderado de p53. Sin embargo, esta línea celular presenta una mutación dominante negativa R273H de p53¹¹¹ por lo que dicho incremento probablemente no tiene efecto sobre las dianas transcripcionales de esta proteína. De este resultado y de la comprobación del potente efecto que el silenciamiento de DTL tuvo en RM82 (tanto *in vitro* como *in vivo*, ver apartado anterior) se deduce que otros inhibidores de la transición G1/S, aún no descritos, deben de estar bajo la acción del complejo CUL4/DDB1^{CDT2}.

Los microarrays de expresión hibridados con RNA total de las líneas celulares silenciadas y sus respectivos controles NTC (ver siguiente apartado) no mostraron cambios significativos en los niveles detectados por los “probeset” correspondientes a los genes CDKN1A, CDKN1B y TP53 (*loci* respectivos de p21^{WAF1/CIP1}, p27^{KIP1} y p53), (Figura 34), por lo que el efecto del silenciamiento de DTL sobre los niveles de estas especies proteicas tiene lugar a

nivel postranscripcional, resultado consistente con la función de ubiquitinación descrita para el complejo CUL4/DDB1^{CDT2}.

Sólo en el caso de TC32 se observa un incremento de los niveles del mRNA del *locus* CDKN1A (que codifica el CDKI p21^{WAF1/CIP1}), de alrededor de 3 veces, claramente insuficiente para explicar el cambio detectado a nivel proteico.

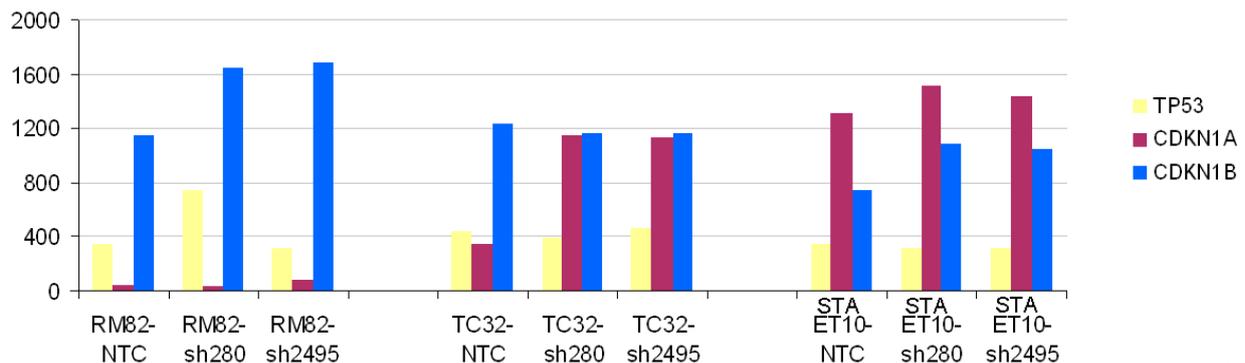


Figura 34. Niveles de señal de los “probesets” de TP53, CDKN1A y CDKN1B obtenidos en el estudio de microarrays de expresión del silenciamiento de DTL en líneas celulares de SE. El eje Y muestra los niveles de señal tras normalizar los datos del microarray de expresión mediante RMA.

4.7 Estudios de microarray de expresión para comparar el perfil de expresión diferencial de los tumores 1qGSig y el debido al silenciamiento de DTL

Los RNA totales extraídos de las líneas celulares RM82, TC32 y STAET10 tras silenciar DTL con las construcciones sh280 y sh2495 (esta vez transducidas por separado) y de sus respectivos controles NTC se hibridaron con microarrays de expresión. La lista génica del perfil de expresión diferencial (análisis SAM pareado) obtenida al enfrentar los datos de expresión de los controles NTC con los de las condiciones silenciadas fue comparada con el perfil de expresión diferencial de los tumores 1qGSig, mediante análisis GSEA (tal y como se explica en Material y Métodos, apartado 3.11). El resultado obtenido muestra un fuerte enriquecimiento (NES = 3.16, Figura 35) de la colección génica de 320 genes que compone el perfil diferencial de los tumores 1qGSig (generada como se explica en el apartado 3.11) en el perfil diferencial debido al silenciamiento de DTL, de lo que se deriva que ambos patrones de expresión son muy similares y, por tanto, que la sobre-expresión de DTL contribuye de manera notable al perfil de expresión diferencial de los tumores 1qGSig y por extensión de los 1qG.

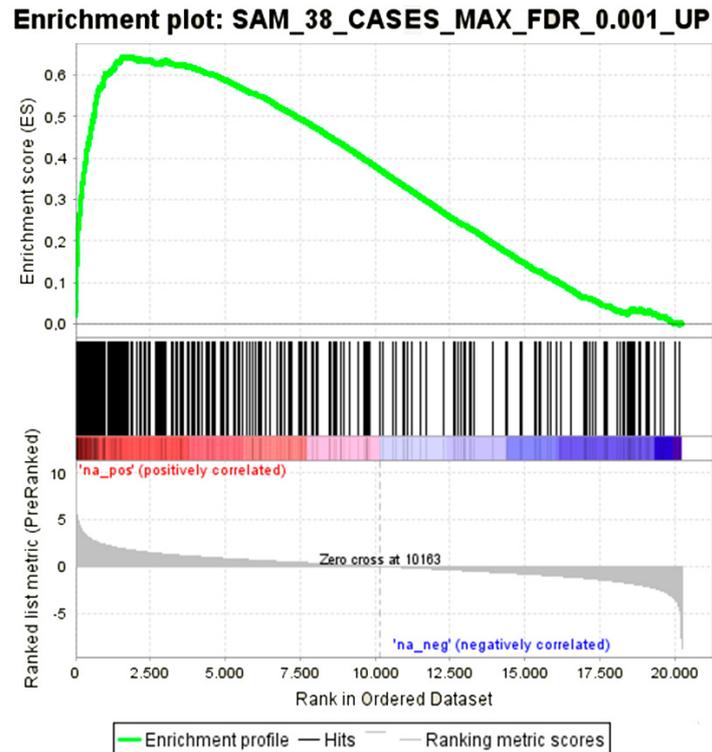


Figura 35. Resultado del análisis GSEA comparando el perfil de expresión debido al silenciamiento de DTL con el perfil diferencial de los tumores 1qGSig. El eje Y muestra el valor del parámetro ES: valor de enriquecimiento.

4.8 Estudio inmunohistoquímico de la tasa proliferativa diferencial de los tumores 1qGSig

De acuerdo con los resultados tanto de los análisis bioinformáticos como de los ensayos *in vitro* e *in vivo* se concluyó que era esperable encontrar una mayor tasa proliferativa en las muestras tumorales 1qGSig respecto al resto de muestras tumorales de SE.

Para validar esta asunción se recurrió al estudio inmunohistoquímico del marcador Ki-67, un marcador contrastado en la evaluación de la proliferación celular, sobre cortes de tejidos de 33 de los 38 tumores del grupo 2 de tumores dispuestos en un microarray de tejidos (TMA). El resultado fue evaluado de acuerdo con 4 puntuaciones de 0 a 3 según el porcentaje de células con núcleos positivos para el marcador (ver apartado 3.19). El análisis estadístico de los datos confirmó la mayor inmunotinción de las muestras tumorales 1qGSig (Kendal-Tau, p-valor = 0.008; U de Mann Whitney, p-valor = 0.02, Figura 36).

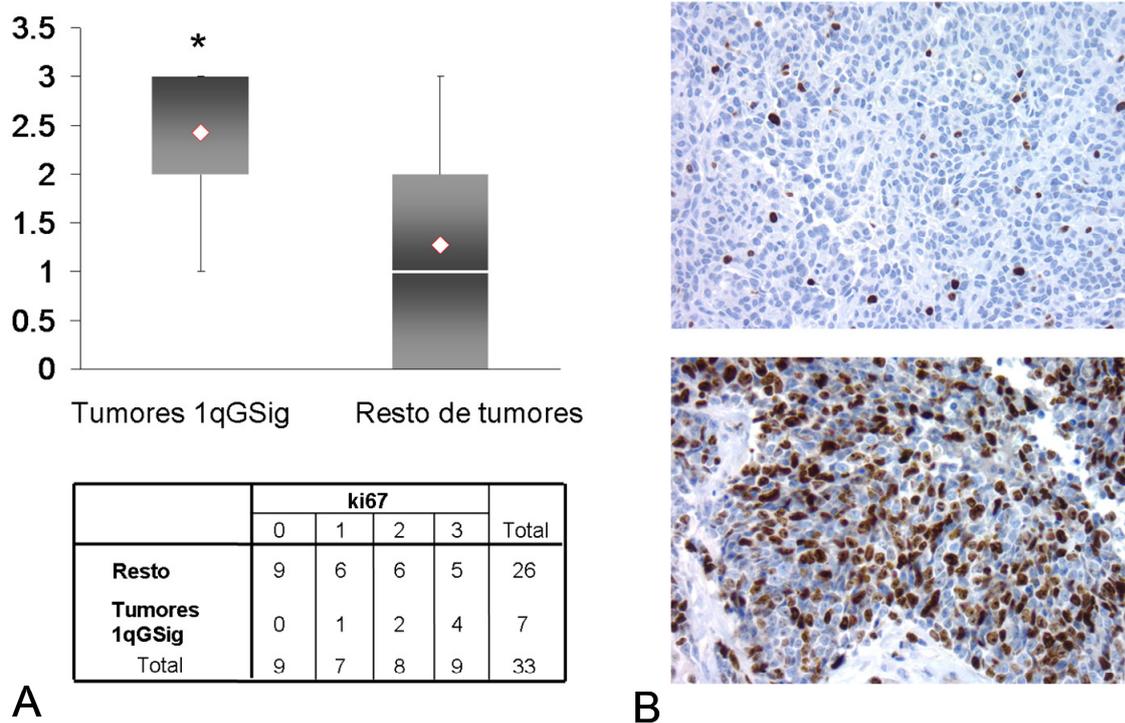


Figura 36. Resultados de la evaluación de la tasa proliferativa de los tumores 1qGSig mediante inmunohistoquímica del marcador Ki-67. (A) “Boxplot” representando los valores de tinción de cada uno de los grupos (arriba) y la correspondiente tabla de contingencia (abajo), el asterisco denota la significación estadística. (B) Imágenes mostrando un corte negativo para la inmunotinción (arriba, puntuación 0) y otro positivo (abajo, puntuación 3). Magnificación 200x.

Al igual que sucedió con los estudios de clusterización no supervisada, el resultado de este ensayo sugiere la existencia de tumores de SE que comparten la desregulación del ciclo celular pero no presentan la 1qG: mientras que los tumores 1qGSig son todos positivos para el marcador de proliferación (la mayoría, 57%, recibiendo el máximo valor de puntuación), el resultado en el resto de tumores es heterogéneo y aunque la mayoría de casos carecen de marcaje o tienen un marcaje muy débil (58%, valores 0 ó 1 de puntuación), un grupo nada despreciable (19%) recibe un valor máximo de puntuación.

4.9 Validación de genes candidatos en 16q

Aunque la delección de 16q no alcanzó significación estadística en el estudio del impacto de las ANC en los datos clínicos, se trata de la delección más frecuente en SE y cabría la posibilidad de que los genes de la región con funciones de supresor tumoral estuvieran afectados en las muestras tumorales por otros mecanismos de pérdida de función tales como

la mutación de su secuencia o el silenciamiento de su expresión mediante metilación de las islas CpG de sus promotores (en aquellos que las tengan), pudiendo de esta manera alcanzar la relevancia clínica que el estudio de ANC no consigue detectar.

Con el objeto de testar esta hipótesis se ensayó la recuperación de la expresión de los principales genes candidatos en la segunda RMS más restringida en extensión de la región 16q, dado que la RMS mostrada en la Figura 16 no contiene genes con actividades supresoras tumorales descritas. Esta segunda RMS está situada en 16q23.1-q24.1 (76.37 – 84.75 mb desde el p-ter) y es de 8.38 mb de extensión. Los principales supresores tumorales de la región son WWOX (FOR o FRA16D)¹⁴⁵⁻¹⁵² y cdh13 (caderina13, H-caderina o T-caderina)¹⁵³⁻¹⁵⁸, cuya pérdida de función ha sido ampliamente involucrada en procesos tumorales.

En primer lugar se comprobaron sus niveles de expresión en un panel de 10 líneas celulares de SE, mediante Real Time RT qPCR y Western Blot. Como se observa en la Figura 37, los niveles de WWOX son heterogéneos y correlacionan bien con el estado de número de copias genómico subyacente de 16q. En el caso de cdh13 los niveles proteicos son bajos en todas las líneas independientemente del estado genómico y de los niveles de mRNA, lo cual sugiere la presencia de una regulación postraduccional supresora de su proteína.

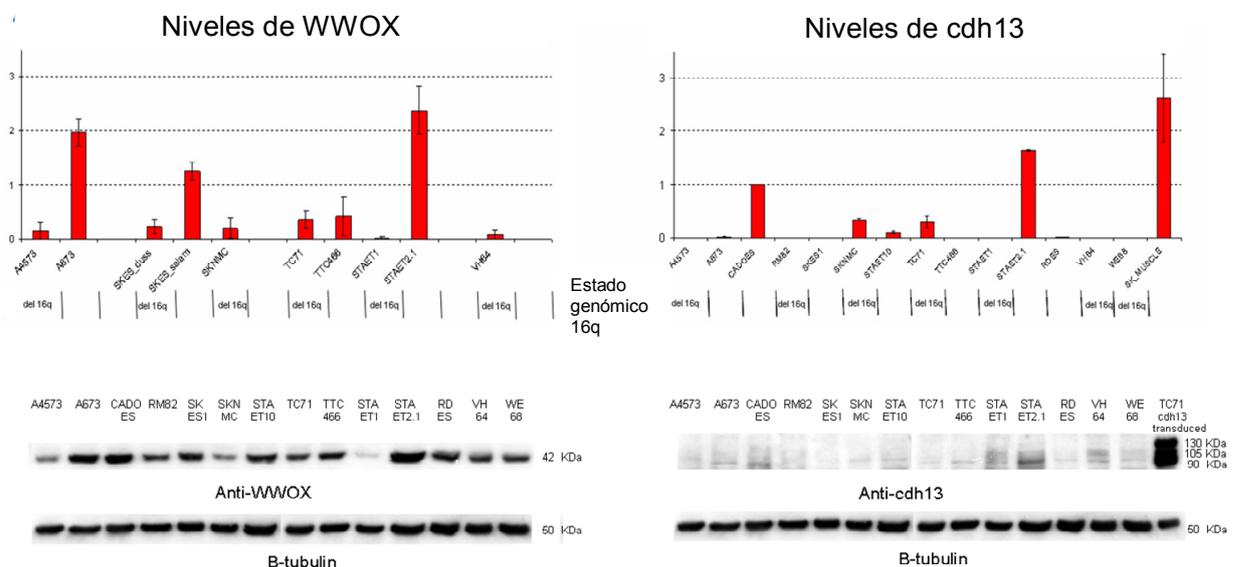


Figura 37. Valoración de los niveles de expresión de los supresores tumorales WWOX y cdh13 en el panel de líneas celulares de SE, mediante Real Time RT qPCR (paneles superiores) y Western blot (paneles inferiores). En el caso de cdh13 se incluyó una muestra de cDNA de músculo esquelético (SK_MUSCLE) en los ensayos de Real Time, como referencia de niveles fisiológicos de expresión y en el Western Blot se incluyó un control positivo consistente en la línea TC71 transducida con la construcción pMIG-cdh13. Los ensayos de Real Time RT qPCR están representados en unidades arbitrarias, normalizados ddCT respecto a los niveles de GAPDH y son la media de duplicados.

A continuación, se obtuvo el cDNA de ambos genes a partir de muestras sin mutación de la secuencia y elevada expresión (la línea celular SKES1 en el caso de WWOX y RNA de tejido sano de corazón y cerebro, en el caso de *cdh13*). La secuencia codificante completa de ambos cDNA fue amplificada y clonada en el vector retroviral pMIG. Este vector incorpora una secuencia IRES-GFP tras el sitio de clonación, generando a partir del promotor lentiviral 5' LTR, en un mismo transcrito bicistrónico, los mRNA de ambas proteínas pero con una traducción independiente (por lo que el gen de interés y la GFP se expresan en distintos péptidos). La Figura 38 muestra la efectividad de esta aproximación recuperando los niveles proteicos de ambos supresores tumorales.

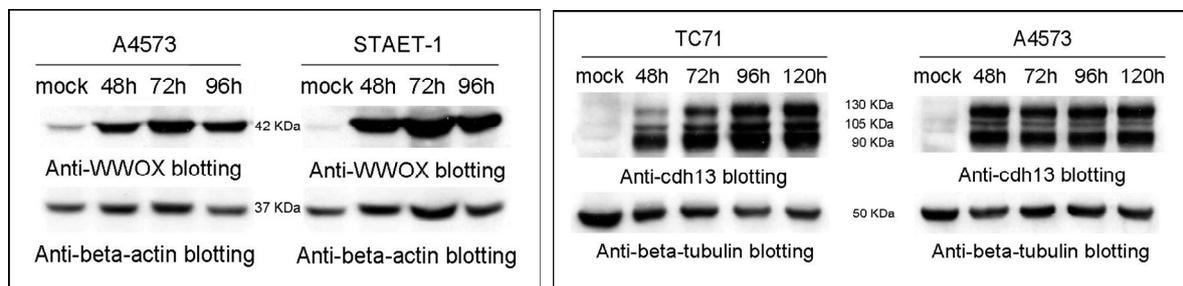


Figura 38. Recuperación de los niveles de expresión de los supresores tumorales WWOX y *cdh13* en las líneas celulares con delección de sus *loci* y niveles de expresión más bajos. A la izquierda, niveles proteicos de WWOX varios días tras la transducción de la construcción pMIG-WWOX en las líneas celulares A4573 y STAET1. A la derecha, niveles proteicos de *cdh13* varios días tras la transducción de la construcción pMIG-*cdh13* en las líneas celulares TC71 y A4573. Las distintas bandas de *cdh13* son producto de la modificación postraduccional por glicosilación y se observan también en los Western blot de la proteína endógena. Mock: control transducido con el vector vacío.

Como era de esperar, la integración viral en el genoma de las líneas celulares ensayadas permitió eficiencias de transducción poblacional del 100% (Figura 39) y estabilidad de la expresión ectópica durante una ventana de tiempo de al menos una semana, sin necesidad de selección con antibiótico, que pudo ser fácilmente monitorizada por la co-expresión de la proteína verde fluorescente (Figura 39).

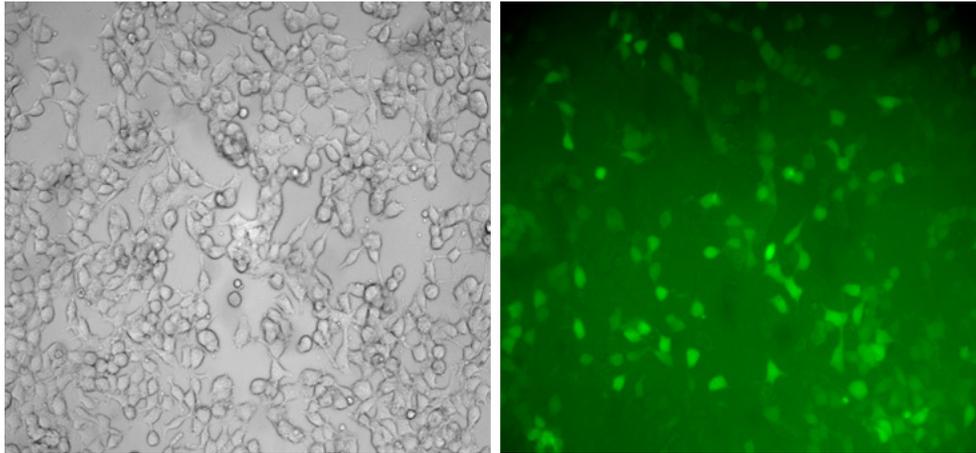


Figura 39. Control GFP de la transducción retroviral con pMIG. Imagen de contraste de fases (izquierda) y de fluorescencia (derecha) de una población celular de STAET1 96 horas tras la transducción con la construcción pMIG-WWOX (magnificación 100x)

A pesar de la fuerte actividad supresora tumoral de ambos genes descrita en la literatura, las pruebas a las que se sometieron las células transducidas no mostraron ningún efecto de la recuperación de su expresión en movilidad, proliferación, patrón de ciclo celular o apoptosis (datos no mostrados).

En el caso de WWOX también se realizó un estudio de inmunohistoquímica (Figura 40) sobre cortes de TMA que, aunque de nuevo mostró una alta heterogeneidad en los niveles de expresión de las distintas muestras tumorales ensayadas (coincidente con el patrón de expresión observado en el panel de líneas celulares), no reflejó ninguna correlación con los datos clínicos.

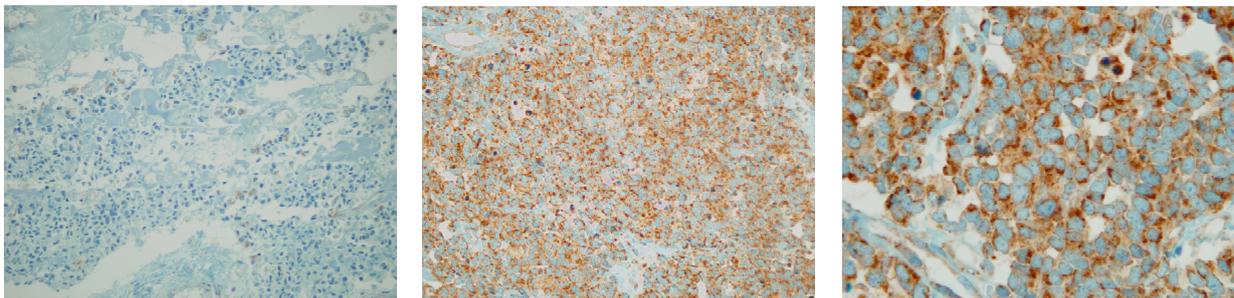


Figura 40. Inmunohistoquímica con el anticuerpo anti-WWOX en un TMA de muestras tumorales de SE. A la izquierda un ejemplo de corte de tejido negativo para WWOX (magnificación 100x), en el centro un ejemplo de un corte de tejido positivo que muestra un claro marcaje citoplasmático en la práctica totalidad de las células (magnificación 100x) y a la derecha este mismo corte con mayor detalle (magnificación 400x).

5. DISCUSIÓN

5.1 Consideración comparativa de la metodología usada para el abordaje del estudio de las ANC

Consideremos las diferencias de la tecnología utilizada en el presente estudio respecto a la utilizada en los trabajos más recientes de aCGH en SE¹⁰⁷⁻¹⁰⁸, que hicieron uso de plataformas comerciales de SNP-array. A pesar de la evidente mayor resolución de los SNP-array, varios son los puntos a favor de la plataforma de BAC-array:

- Como puede observarse en la Figura 17, aparte de la incapacidad de los BAC-array para delimitar con precisión los puntos de ruptura de las ANC detectadas, se aprecia una marcada diferencia en la sensibilidad de ambas plataformas, alcanzando el BAC-array para la misma ANC (el DNA hibridado es el mismo) un valor de log₂ratio de prácticamente el doble del obtenido, en términos absolutos, por el SNP-array de Affymetrix. Esta situación ha sido la tónica general en los resultados obtenidos con ambas plataformas.

- El ruido basal de los BAC-array (entendido como la oscilación en el valor log₂ratio no debida a una situación biológica sino intrínseca a la plataforma) es notablemente inferior al de los SNP-array. Este valor es para la práctica totalidad de los resultados de los BAC-array generados en este trabajo de ± 0.2 de log₂ratio, como se observa en la Figura 17, mientras que en los SNP-array oscila entre ± 0.5 y ± 1 (Figura 41). Esta diferencia es fácilmente explicable considerando que la longitud de los oligonucleótidos del SNP-array es de 22-70 pb (según la casa comercial, más largos los de Agilent) mientras que los fragmentos de aminolinking PCR (los elementos que componen las sondas del BAC-array) tienen tamaños de 200 pb a 2Kb, además de por las peculiaridades de la superficie de los portas “Codewalk” (ver apartado 3.4.3).

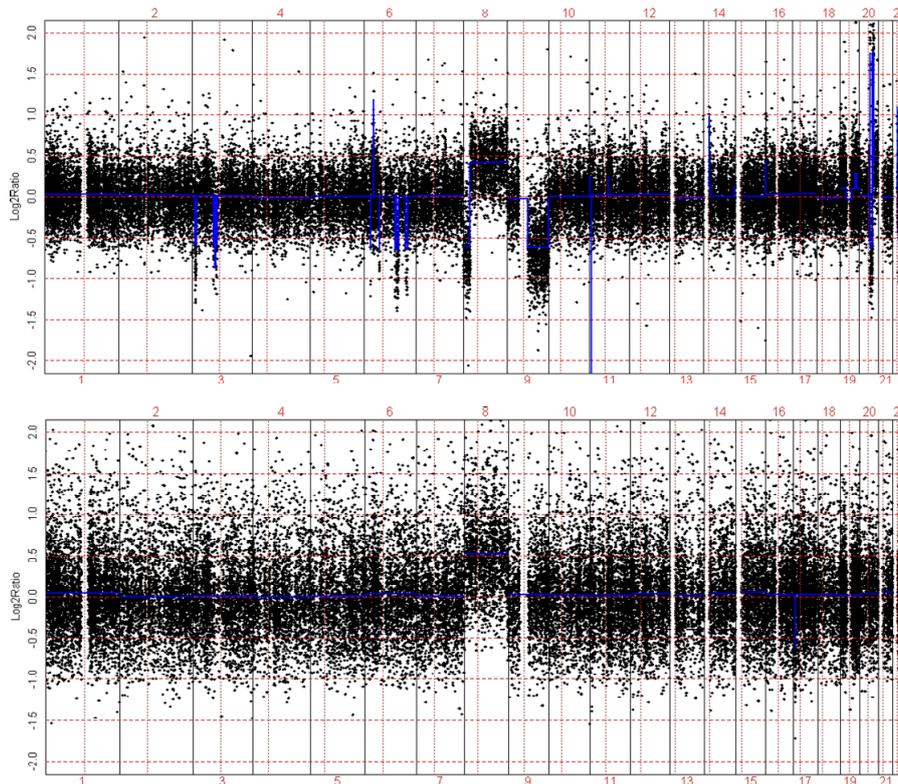


Figura 41. Ejemplos de datos generados mediante aCGH SNP-array. Hibridaciones correspondientes a dos muestras del subgrupo de 17 muestras tumorales (pertenecientes al grupo 2 de tumores) caracterizados con el SNP-array de Agilent “Human Genome CGH Microarray Kit 44K”. Los datos en crudo han sido obtenidos del repositorio público CanGEM y procesados con el paquete “snapCGH” del mismo modo que los datos presentados en la Figura 17.

La relevancia de estas diferencias técnicas estriba en que limitan la sensibilidad final alcanzable por cada una de las plataformas, condicionando su capacidad de detección. Cuando se trabaja con una población 100% homogénea para las ANC, sin presencia de células no tumorales y de ploidía normal ($2n$), como es una línea celular diploide, las diferencias en sensibilidad son prácticamente irrelevantes puesto que cualquier plataforma será capaz de detectar las ANC de la muestra. Sin embargo, las muestras tumorales se alejan marcadamente de esta situación ideal de la siguiente manera:

- Presencia de células no tumorales: salvo en el casos de muestras microdisseccionadas las poblaciones celulares de un tumor presentan con frecuencia células sanas, ya sean células de soporte, células infiltrantes del sistema inmune o células no transformadas del tejido adyacente. El DNA de estas subpoblaciones (carente de ANC) está lógicamente presente en el extracto utilizado para las hibridaciones, reduciendo las ratios finales esperables para una

ANC típica (1.5 para las ganancias, 0.5 para las pérdidas) y haciéndolas más difíciles de detectar.

- Ploidías aberrantes: los tumores presentan con frecuencia valores de ploidía cercanas a $3n$ (como muestra la Figura 18 del estudio de COBRA-FISH de la línea celular RM82), $4n$ e incluso superiores. Las ratios esperables se atenúan para estas muestras siendo por ejemplo la ratio esperable para una ganancia en una población celular triploide de 1.33 (frente al ratio esperable de 1.5 correspondiente a una ganancia de una copia en un “background” diploide).

- Heterogeneidad genómica de la población tumoral: las ANC son generalmente alteraciones secundarias y en un modelo de evolución tumoral típico aparecen en una sola célula como consecuencia de un suceso de replicación del DNA o de segregación cromosómica aberrante. A partir de ese momento la supuesta ventaja conferida para la progresión tumoral hace que la ANC sea mantenida en la población derivada de la célula en que surgió la alteración. Sin embargo, toda la masa tumoral previa al suceso no la presentará, haciendo que la localización de la pieza tumoral usada para la extracción de DNA influya en la detección de las ANC y que una misma pieza pueda comprender una población mixta de células positivas y negativas para ella. De nuevo una situación así disminuirá la ratio esperable.

La mayor coincidencia del ratio observado con el esperado (en la Figura 17, \log_2 ratio esperado -1) y el menor umbral de ruido técnico hacen que los BAC-array sean una plataforma ideal para un primer cribado de las muestras, enfocado a la maximización de la sensibilidad, mientras que los SNP-array pueden ser posteriormente utilizados para delimitar las ANC halladas más interesantes, por ejemplo en las muestras decisivas para la delimitación de las RMS. Un esquema de trabajo así presenta también la ventaja de maximizar el tamaño muestral analizable, si incluimos en el balance global de las plataformas su respectivos precios (60 euros por hibridación el BAC-array, más de 500 euros el SNP-array) y la limitación de los recursos económicos, el motivo principal de que los estudios de aCGH SNP-array hechos en SE comprendan tan sólo 25 y 31 muestras¹⁰⁷⁻¹⁰⁸. Este esquema de trabajo es el que aquí ha sido utilizado, permitiendo el estudio total de 83 muestras (entre líneas celulares y muestras tumorales), la mayor serie muestral de SE estudiada mediante aCGH hasta la fecha.

Respecto a la evidente diferencia de resolución de ambas plataformas cabe argumentar que, tomando como ejemplo la microdelección más frecuente y significativa que se conoce en SE,

la que afecta al *locus* de CDKN2A (*locus* de 27.55 kb de tamaño), ésta es detectada por nuestro BAC-array (Figura 16) gracias a la sub-librería dirigida contra los *loci* de los principales genes involucrados en cáncer que incluye la librería de BAC utilizada (“Sanger 1 mb clone set”). De hecho, suponiendo la presencia de microdeleciones tan reducidas como la más pequeña detectada en SE a día de hoy⁹⁷, de 58 kb de extensión y que eliminó por completo el *locus* de CDKN2A, ésta también sería detectada por el BAC-array utilizado en nuestro estudio puesto que una alteración así abarcaría un 60% del BAC de la librería localizado en la región (de 100 kb de tamaño, Figura 16).

Las diferencias en frecuencias globales de las ANC detectadas en SE por otros autores y las obtenidas con nuestra metodología son por tanto, teniendo en cuenta todas estas consideraciones, fácilmente explicables.

5.2 Comparación de los resultados de aCGH obtenidos con la literatura previa

Como ya se ha comentado en el apartado anterior la utilización de un método de cribado de aCGH con gran sensibilidad como el usado en la presente investigación es de gran importancia. Es probablemente por ello que, aunque las ANC recurrentes detectadas en este estudio son las mismas y el perfil global es muy parecido al de estudios previos, las frecuencias de alteración por cromosoma parecen ser en general superiores: en las Figuras 42 y 43 se muestran las frecuencias totales de ANC en SE obtenidas de la base de datos Progenetix¹⁵⁹ (www.progenetix.net), correspondientes a los grupos de enfermedad (códigos de la “International Classification of Diseases for Oncology”, ICD-O-3) 9260/3 – “Ewing Sarcoma” y 9473/3 – “Primitive Neuroectodermal Tumors”. Las frecuencias obtenidas provienen de los datos recopilados de 196 muestras tumorales (las líneas celulares fueron descartadas del análisis) analizadas tanto por mCGH como por aCGH y pertenecen a 14 publicaciones. Comparadas con las frecuencias de ANC detectadas en el estudio aquí presentado son en general más bajas.

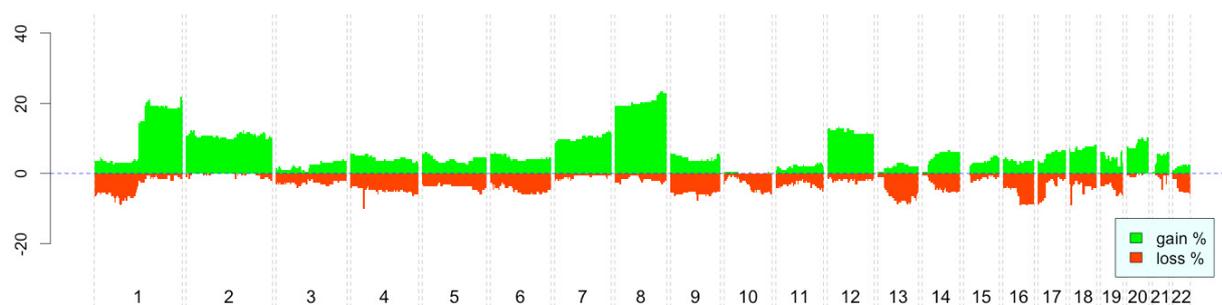


Figura 42. Frecuencias de las ANC en SE según la literatura previa. El eje X representa el número de cromosoma mientras que el eje Y muestra las frecuencias expresadas en porcentaje muestral. Los datos han sido obtenidos de la base de datos Progenetix a partir de 14 publicaciones y comprenden 196 muestras tumorales. El color rojo representa pérdidas de material genómico mientras que el color verde representa ganancias.

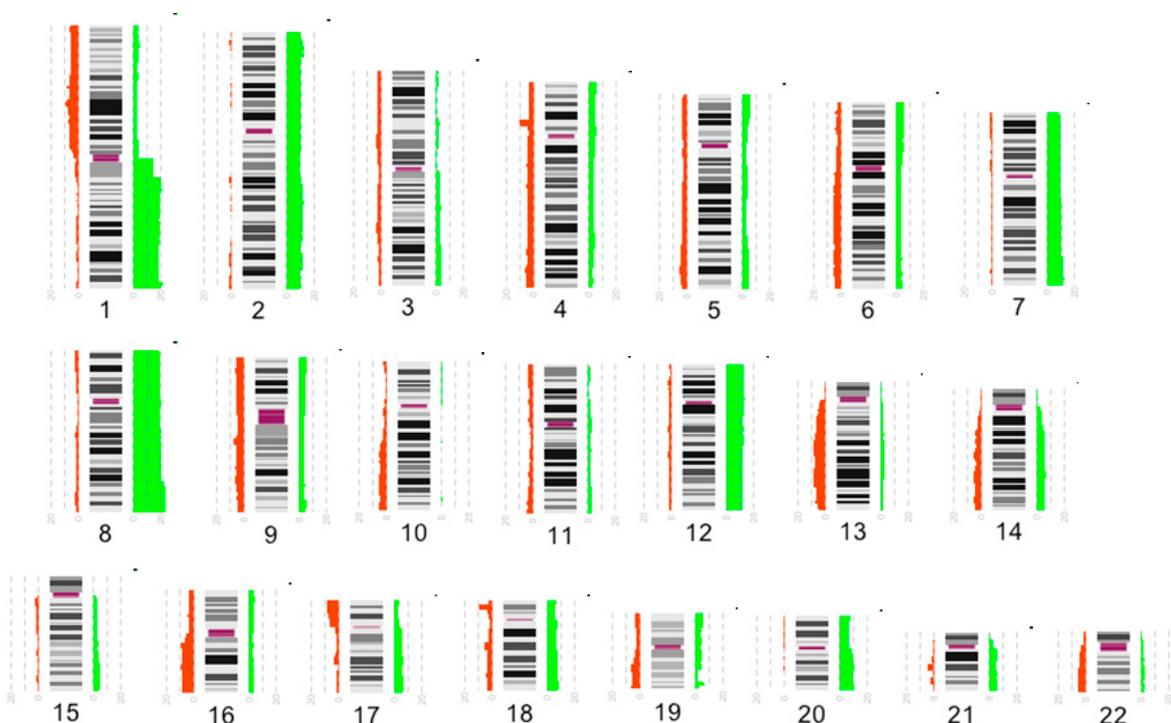


Figura 43. Distribución genómica de las ANC en SE según la literatura previa. La imagen muestra, junto a los ideogramas cromosómicos, la frecuencia global de ANC expresada en porcentaje muestral. El color rojo representa pérdidas de material genómico mientras que el color verde representa ganancias.

5.3 Las ANC contribuyen a la progresión del SE y son clínicamente relevantes.

Tal y como muestra el resultado de los estudios de diferencias de supervivencia de los pacientes debidas a los parámetros PGA y NTA, las ANC no sólo tienen impacto en la clínica del SE sino que además este efecto es muy marcado, correlacionando el valor creciente de estos estimadores globales de ANC con una supervivencia gradualmente peor.

Los autores que previamente han evaluado el efecto global de las ANC en la supervivencia del SE lo han hecho mediante la definición de umbrales de número de copia arbitrarios, asumiendo sin mayor base científica que los tumores que rebasan este umbral están afectados por procesos de inestabilidad cromosómica, siendo el umbral más frecuentemente usado el de $NTA = 3$ ^{105,107-108}.

En el presente trabajo se hicieron inicialmente las dos siguientes asunciones:

- Los valores globales de estimación de las ANC tienen probablemente un efecto progresivo sobre la supervivencia, siendo esta peor mientras mayores sean sus valores.
- El parámetro PGA probablemente tiene una mejor correlación con los datos clínicos de supervivencia que el parámetro NTA, ya que éste último no tiene en cuenta las notables variaciones en el tamaño medio de las ANC que existen de unas muestras a otras (como muestra la Figura 15). Por el contrario el valor de PGA refleja de una manera menos ambigua la alteración genómica global de número de copia.

Nuestros resultados confirman ambas asunciones: las ANC consideradas en conjunto tienen un efecto marcado en supervivencia, que es peor cuanto mayor es el valor del estimador PGA, y además este parámetro correlaciona mejor con la supervivencia que el NTA (Figura 20).

Enfocando los resultados de manera opuesta, casi un 20% de tumores carecen de ANC y la supervivencia de este conjunto de pacientes es especialmente favorable. Este hallazgo es de especial interés pues es en potencia de gran utilidad como factor predictor de pronóstico favorable.

En contraste con el umbral de $NTA = 3$ fijado por otros autores, el hecho de que la mayoría de ANC detectadas no parezcan mostrar valor clínico consideradas individualmente, unido a la evidencia de la peor supervivencia diferencial de los tumores con ANC (cualquiera que sea el estimador utilizado), sugiere la existencia de dos grupos de tumores de SE:

- Tumores que no desarrollan ANC: la evolución biológica de estos tumores probablemente no propicia la aparición de inestabilidad cromosómica, y el transcurso clínico resultante es muy favorable.

- Tumores con ANC: en algún momento suceden eventos que propician la aparición de inestabilidad cromosómica, de manera que las ANC surgen y se incrementan con el transcurso del desarrollo tumoral. El hecho de que sean escasas las ANC que parecen tener impacto en supervivencia consideradas individualmente (teniendo en cuenta que la mayoría de estas se asocian a la ANC 1qG y su efecto en supervivencia es seguramente un reflejo de aquella) apunta a que, salvo la 1qG, el resto de alteraciones no tienen un efecto de “driver” genético individualmente y son sólo un reflejo de la acumulación progresiva de ANC en este grupo de inestabilidad cromosómica. Según este modelo, algunas combinaciones de ANC tendrían un mayor efecto en la progresión tumoral, realimentándola y propiciando la aparición de subsecuentes ANC e incrementando con ello la probabilidad de un pronóstico adverso. El parámetro PGA sería entonces un interesante indicador del estado de progresión tumoral en estos pacientes, siendo el verdadero factor pronóstico la presencia de ANC, por pequeño que fuera su número.

5.4 Significado clínico y biológico de las ANC más recurrentes en SE, consideradas individualmente

La comparación de las frecuencias de las ANC de tumores y líneas celulares permite obtener interesantes conclusiones acerca de su relevancia y contribución real al SE. La evidente diferencia en frecuencias globales entre ambos tipos de muestras está principalmente causada por la presencia de casi un 20% de muestras tumorales carentes por completo de ANC. Sin embargo, si se elimina de la comparación este grupo de tumores, a pesar de la alta coincidencia de sus perfiles la confrontación estadística usando el programa “CGHPermutations” detecta la sobrerrepresentación en los perfiles de aCGH de las líneas celulares de las ANC ganancia de 8q (de 124mb a 129 mb), pérdida de 3p (de 40mb a 75mb desde p-ter) y pérdida de 9p (desde el p-ter a 24 mb), regiones que comprenden en su interior 3 de las 5 RMS definidas por el estudio de aCGH y que incluyen los *loci* de los genes MYC, FHIT y CDKN2A, respectivamente (Tabla 2, y Tabla 3 condición testada “Líneas celulares vs. muestras tumorales con PGA > 1%”).

La interpretación de estas diferencias no es clara, pudiendo estar causadas por el sesgo debido al establecimiento en cultivo de las líneas celulares de SE a partir de los tumores más agresivos, lo que indicaría que las ANC sobrerrepresentadas son alteraciones características

de dicho grupo de tumores y destacaría su relevante contribución al desarrollo tumoral. Sin embargo también podrían reflejar, sin descartar que jueguen un papel en el tumor *in vivo*, una contribución principal de estas ANC en la adaptación de los tumores del SE al cultivo celular *in vitro*, restando importancia a su relevancia biológica real. El hecho de que, como es el caso de la delección del 3p que engloba al *locus* de FHIT, su incidencia en las muestras tumorales sea tan infrecuente (6% para la delección de 3p) apunta más bien a lo segundo.

Aunque carente de rigor estadístico, hemos observado una acumulación especialmente alta de copias del cromosoma 8 en la línea celular RM82 (6 copias), una línea con evidente inestabilidad cromosómica, como demuestran las Figura 19 y 20, reforzando la idea de que la ganancia de esta región es también especialmente ventajosa en el cultivo *in vitro*.

Por otra parte, pocas son las alteraciones que alcanzan significación estadística en la evaluación de la correlación con los datos clínicos de las ANC consideradas individualmente. Claramente la 1qG debe ejercer efectos biológicos relevantes para la evolución tumoral puesto que no sólo ella, sino también las ANC que aparecen asociadas a ella, impactan de manera muy marcada en la supervivencia total de los pacientes. El efecto de las ANC acompañantes de la 1qG en supervivencia es atribuible a 1qG puesto que no suponen una diferencia de supervivencia respecto a la causada por 1qG *per se*. El resto de ANC con impacto en supervivencia (pérdida del cromosoma 22 en supervivencia total y pérdida del cromosoma 10 en supervivencia libre de recaída) podrían tener también un papel relevante, pero su efecto biológico no ha sido investigado en el presente trabajo debido a sus frecuencias muestrales más bajas (12% para la primera, 15% para la segunda).

Es interesante además que, mientras que existe un escaso número de tumores que en lugar de pérdida presentan ganancia de estas regiones (1.5% tanto para el cromosoma 22 como para el 10), esto no sucede en absolutamente ningún tumor en el caso de la región 1q (0% de tumores con pérdida del 1q). Es este un criterio muy útil para discriminar las ANC que probablemente se acumulan de manera aleatoria por efecto de una inestabilidad cromosómica subyacente y no suponen una contribución biológica o clínica relevante. Como ejemplo, el cromosoma 4 está ganado y perdido en esta serie de muestras con la misma frecuencia (6%). Algo parecido, aunque sin ser un empate tan marcado entre las alteraciones de ambos signos, ocurre con los cromosomas 2, 3, 4, 5, 6, 9 (salvo 9p) y 13.

Especialmente relevante es la ausencia de asociación entre la 1qG y las ANC que afectan a los *loci* de las otras alteraciones moleculares secundarias con relevancia clínica en SE (TP53 y CDKN2A), ya que esta ausencia de asociación hace improbable el solapamiento de las muestras tumorales afectadas por estas alteraciones y la 1qG, sugiriendo que se trata de eventos independientes con repercusión clínica propia.

Respecto a las ANC que no alcanzan significación estadística en los estudios de correlación con los datos clínicos cabe destacar, a pesar de su baja frecuencia, la pérdida del 1p36, probablemente la ANC que más ha sido implicada en carcinogénesis¹⁶⁰.

En general es destacable la ausencia de impacto en supervivencia de las pérdidas de material genómico, más proclives a producir efectos clínicos dado que la pérdida de dosis génica de genes con actividad supresora tumoral es un evento que predispone a una pérdida total de función por un segundo evento (según el modelo de “double hit” de Knudson) o que incluso puede tener un efecto directo si el alelo que resta no es suficiente para mantener los niveles proteicos y funcionales requeridos (modelo de haploinsuficiencia). En contraste, al menos intuitivamente, la ganancia de una sola copia génica no debería suponer un fuerte efecto de sobre-expresión génica de los *loci* albergados en la región ganada (sí las amplificaciones que ocurren en otros tumores, como la de MYCN en neuroblastoma¹⁶¹) y por tanto sería esperable una menor contribución oncogénica de las ANC de este signo. Posiblemente esta inesperada mayor frecuencia y relevancia clínica de las ganancias genómicas refleja las peculiaridades de la biología de un tumor como el SE donde existe un evento primario que probablemente “aprovecha” una circunstancia hormonal transitoria (reflejada por la mediana de edad de incidencia y la dependencia del tumor de la ruta de IGF1-R), para establecer un proceso canceroso con necesidad de sólo otros pocos eventos moleculares, una situación que se aleja del paradigma de oncogénesis y desarrollo tumoral habitual de los carcinomas¹⁶². Las ganancias de número de copia podrían ser en este escenario más efectivas propiciando la progresión tumoral. De hecho, el modelo que se propone más adelante para explicar la contribución de la 1qG y la sobre-expresión de DTL al desarrollo del SE (apartado 5.10) se basa en esta idea.

Respecto a la correlación de la ganancia del cromosoma 8 con la localización de las metástasis en pulmón es este un resultado difícil de interpretar ya que ninguno de los genes en la RMS definida permite establecer una relación evidente entre esta ANC y la selección del

órgano de metástasis. De hecho la región incluye un gen relacionado con la supresión de las metástasis, MTSS1¹⁶³. La RMS contiene sin embargo el *locus* de MYC. Este conocido oncogén está siendo recientemente implicado con el proceso de la metástasis: Ma *et al*¹⁶⁴ han observado que tanto MYC como MYCN actúan aumentando la expresión del microRNA mir 9, que propicia la metástasis suprimiendo la expresión de la E-caderina, además de la consiguiente activación de la expresión de VEGF a través de la liberación de β -catenina. Estos resultados coinciden con los obtenidos por Lujambio *et al.*¹⁶⁵ quienes, investigando los patrones de metilación diferencial característicos de tumores metastáticos encontraron este mismo microRNA, miR-9, junto con otros 2 (miR-148a, miR-34b/c), diferencialmente metilados en tumores metastáticos. La lista de genes silenciados por estos microRNAs incluye a MYC, además de a otros estrechamente relacionados con la progresión del control celular (E2F3, CDK6 y TGIF2).

Sin embargo, la ganancia del cromosoma 8 en SE no parece correlacionar con la presencia de metástasis, sino con la localización diferencial de estas en pulmón, lo cual tiene interés clínico ya que las metástasis de los SE en esta localización tienen un mejor pronóstico asociado que los que metastatizan a la otra localización mayoritaria (hueso). Además, MYC es una de las dianas de la fusión EWS-ETS mejor conocidas así que la ganancia de 8q no actuaría causando su sobre-expresión sino que a lo sumo tendría un efecto amplificador mediante el incremento de dosis génica (de una sola copia extra en la mayoría de los casos). Si los posibles niveles diferenciales de MYC entre los tumores sin ganancia del cromosoma 8 y los de aquellos que presentan esta ANC podrían determinar la localización metastática es un punto que deberá ser explorado en futuros estudios.

Finalmente la ANC pérdida de 16q, una de las alteraciones más atractivas *a priori* por ser la deleción más frecuente en SE y por haber sido previamente correlacionada con un peor pronóstico de los pacientes¹⁰⁶, no alcanza la significación estadística necesaria en la serie de datos aquí estudiada. Aún así, se ensayó la recuperación de genes supresores tumorales (WWOX y cdh13) situados en una región algo más amplia que la RMS, cuyos resultados han sido mostrados en el apartado 4.9 y que sugieren una escasa relevancia biológica y clínica de esta ANC y de los genes situados en ella en SE. Además, este resultado negativo resalta la dificultad de encontrar funciones génicas no relacionadas con el evento molecular primario (EWS-ETS) con importancia en la oncogénesis y/o desarrollo de esta patología.

5.5 Ganancia del brazo cromosómico 1q en cáncer

Gracias a la base de datos Progenetix¹⁵⁹ (www.progenetix.net) podemos evaluar la prevalencia global de la 1qG en cáncer a partir de los datos de 118 publicaciones comprendiendo 4492 muestras tumorales (descartando líneas celulares) analizadas mediante aCGH y pertenecientes a la práctica totalidad de neoplasias existentes (345 subtipos de enfermedad según los códigos ICD-O-3). Como se puede apreciar en la Figura 44 la 1qG alcanza una frecuencia del 15% de las muestras y presenta valores similares para todo el 1q sin que aparentemente existan regiones de alteración circunscritas a porciones concretas del brazo cromosómico. Cabe destacar la frecuencia doble de esta ANC en SE (28-31% hallada en este estudio) respecto a esta estimación global, lo cual resalta su contribución oncogénica especialmente significativa en esta entidad tumoral.

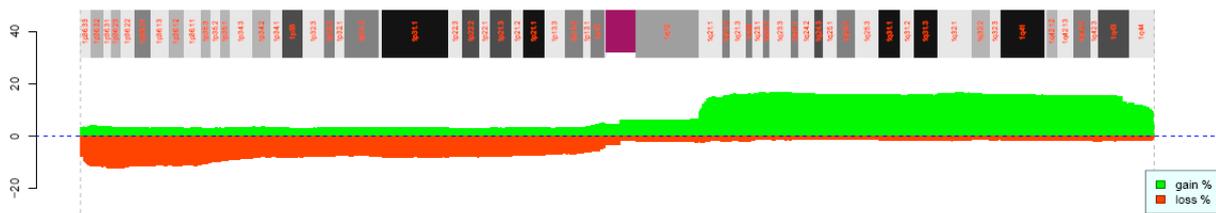


Figura 44. Frecuencia global de ANC del cromosoma 1 en muestras tumorales humanas. Obtenido de la base de datos Progenetix. El color rojo representa pérdidas de material genómico mientras que el color verde equivale a ganancias.

Respecto al impacto de la 1qG en la clínica de otras entidades tumorales, se ha descrito una asociación de esta ANC con una mala supervivencia de los pacientes en neuroblastoma¹⁶⁶, meduloblastoma¹⁶⁷, mieloma múltiple¹⁶⁸, tumor de Wilms¹⁶⁹ y carcinoma papilar de tiroides¹⁷⁰, lo cual plantea la posibilidad de que el fenotipo molecular o mecanismo hallado asociado a ella en el presente estudio pueda ser compartido por estas otras patologías oncológicas.

5.6 La ganancia de 1q se asocia con un perfil de desregulación del ciclo celular en SE, resultados bioinformáticos.

El estudio de expresión de la serie independiente de 38 tumores, el grupo 2 de tumores, arroja un consistente resultado que sugiere la presencia de un ciclo celular desregulado en los tumores con 1qG. Aunque los resultados de las diferentes bases de datos MSigDB son de naturaleza muy diversa, todos coinciden en apuntar a esta situación. De especial significación es el análisis usando la MSigDB C3, puesto que esta base de datos carece de sesgo en la creación de las colecciones génicas que contiene al agruparse los genes según la predicción bioinformática de su respuesta común a factores de transcripción y microRNA, partiendo de las secuencias de sus promotores y extremos 3' UTR¹⁷¹. Como se ha explicado en el correspondiente apartado de resultados, el análisis GSEA usando esta base de datos detecta un fuerte enriquecimiento de la lista diferencial de expresión de los tumores 1qGSig en genes regulados por E2F1, activador central de la progresión del ciclo celular. No sucede lo mismo con las bases de datos C2 y C5, que, aunque muy completas y significativas, no están libres de sesgo, la primera porque está enriquecida en colecciones génicas pertenecientes a las principales áreas de interés de la investigación actual, y la segunda (términos GO) porque se basa en el conocimiento previo de los genes anotados. Además, y muy significativamente, el análisis Ingenuity refuerza este resultado, lo cual confiere una consistencia global aún mayor al análisis bioinformático.

Por otra parte, el resultado de la clusterización no supervisada sugiere un importante efecto de la 1qG en el transcriptoma, identificando un cluster que coincide de una manera muy aceptable con la clasificación según la firma 1qGSig. Es interesante el hallazgo de muestras tumorales que no comparten la ANC y que se agrupan también en este cluster, lo cual sugiere que comparten el patrón de expresión de los 1qGSig y posiblemente presentan también una desregulación del ciclo celular, pero adquirida mediante eventos moleculares distintos a la 1qG. De manera consistente con esta interpretación, el estudio de inmunohistoquímica con el marcador Ki-67 detectó un subgrupo de muestras tumorales del grupo no 1qGSig con una alta tasa proliferativa.

5.7 Selección del gen candidato en 1q

Ya que los resultados hasta este punto consistían en datos de microarrays y de estudios bioinformáticos vimos la necesidad de probar empíricamente la capacidad de genes específicos localizados en 1q y cuya expresión respondiera a la presencia de la ANC subyacente (1qG) de llevar a cabo efectos regulatorios sobre el ciclo celular y cuyos niveles de expresión fueran necesarios para la progresión del ciclo celular de las líneas celulares de SE con esta alteración.

Las validaciones se centraron en un solo gen candidato por la sorprendente consistencia de los datos numéricos y funcionales que lo hacían encajar como piedra de clave en todo el almacén de datos bioinformáticos creado. Especialmente sorprendente fue encontrar a una de las “probeset” del gen situada en la primera posición del estudio de expresión diferencial ranqueado por “d-value”, siendo además el de mayor “fold-change” de los situados en 1q, así como comprobar que su función enlaza las predicciones del análisis Ingenuity de enriquecimiento génico de las rutas de metabolismo de DNA, control de la transición G1/S y ubiquitinación de proteínas. Sin embargo, son varias las cuestiones que merecen una discusión más en profundidad:

- Respecto al cambio en expresión (“fold-change”) de este gen a la ANC subyacente, varias veces superior al esperado (4.78 frente a 1.5), proponemos la existencia de una retroalimentación negativa de los niveles de expresión de este gen y los de p21^{WAF1/CIP1}. Tal retroalimentación fue sugerida por el análisis GSEA utilizando la MSigDB C2, que identificó el enriquecimiento significativo de la colección génica “P21_any_dn” en el perfil diferencial de los tumores 1qGSig (Tabla 4). Este “gene set” se generó en el estudio de Wu *et al.*¹⁷² en el que los autores sobre-expresaron p21^{CIP1/WAF1} en líneas celulares de carcinoma de ovario y comprobaron la respuesta transcriptómica generada. El mRNA de DTL fue uno de los que se hallaron más significativamente reprimidos, en una lista de tan sólo 38 genes. El efecto podría no ser directo pues DTL se encuentra entre los genes que respondieron a la expresión de p21 a medio plazo (estudio de “time-course” realizado entre 4 y 24 horas). De hecho, en los datos de los microarrays de expresión del grupo 2 de tumores existe una diferencia estadística (prueba U de Mann Whitney = 0.008) en los niveles de expresión de p21^{CIP1/WAF1} en las muestras tumorales 1qGSig respecto al resto (menor en las muestras 1qGSig), lo cual es mostrado en la Figura 45.

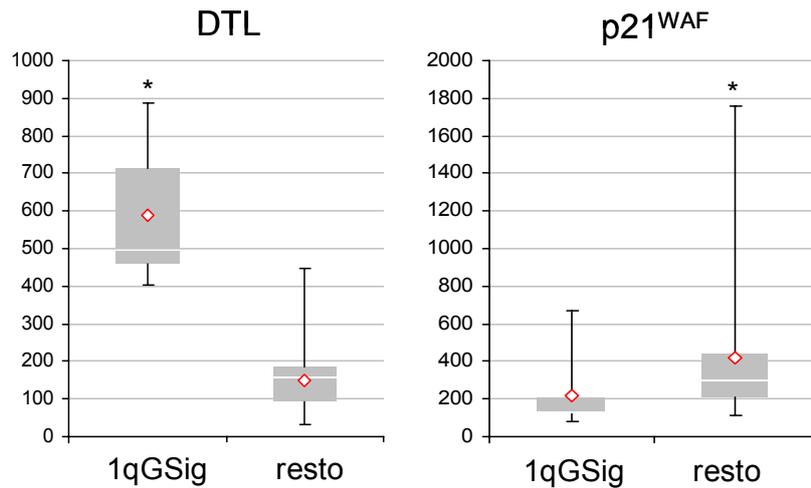


Figura 45. “Boxplot” de los niveles de expresión de DTL y p21^{WAF1/CIP1} en los tumores 1qGSig y el resto de tumores, según los niveles de sus “probeset” en los datos de los microarray de expresión del grupo 2 de tumores.

Según este modelo, el incremento inicial de DTL debido a la 1qG podría redundar en una disminución de los niveles de p21, lo que a su vez retroalimentaría la subida de los niveles de DTL, creando un “feedback” negativo que terminaría provocando el nivel de expresión más alto de lo esperado de DTL detectado en los tumores 1qGSig.

5.8 El gen candidato DTL, sobre-expresado en los tumores 1qGSig, ejerce un papel fundamental en el fenotipo de ciclo celular desregulado y mayor proliferación característico de este grupo de tumores.

Las validaciones funcionales muestran la dependencia de las líneas celulares de SE con 1qG y sobre-expresión de DTL de la función de este gen para la progresión del ciclo celular, tanto *in vitro* como *in vivo*. De este modo, los ensayos de validación del gen candidato confirman las predicciones informáticas relativas a la acción desreguladora de ciclo de la 1qG. Este es un hallazgo muy novedoso en el campo del SE dado que, tal y como se ha apuntado en el apartado de Introducción, son pocos los factores conocidos independientes del evento primario EWS-ETS con una contribución relevante a su biología y clínica.

En el diseño experimental se optó por un silenciamiento progresivo de DTL para evitar la completa supresión de la expresión génica, la cual es seguramente requerida en unos niveles mínimos para la correcta división celular de cualquier célula en proliferación. De este modo,

el abordaje presentado demuestra que es la disminución de la sobre-expresión de DTL en los cultivos celulares, y no sólo la supresión total de sus niveles, la que se traduce en una ralentización de la progresión del ciclo celular, proporcional al grado de silenciamiento.

Cabe destacar que las transducciones en paralelo con un vector idéntico al de las construcciones shRNA (pLKO.1) conteniendo el cDNA de la GFP, que se hicieron con el sobrenadante viral diluido de forma seriada con un factor 1/10 (desde 1 hasta 10^{-7}), garantizaron la completa transducción poblacional en todo el rango de dilución usado en los ensayos de silenciamiento, no sólo porque la condición transducida con el sobrenadante sin diluir mostró por citometría de flujo un 100% de población GFP positiva, sino también porque la población transducida con el sobrenadante correspondiente al factor de dilución 10^{-2} mostró un 50% - 80% de señal de GFP, en experimentos independientes.

Además, la especificidad del ensayo de silenciamiento realizado ha sido demostrada de diversas formas:

- Se han utilizado 5 construcciones shRNA distintas, cada una con una secuencia y dirigidas contra una porción diferente del mRNA de DTL. Todas ellas demostraron capacidad para disminuir los niveles proteicos de DTL y de inducir apoptosis (blot de caspasa-3). Además, se comprobó la capacidad de 3 de ellas de inducir una parada de ciclo en G0/G1 (dato no mostrado).

- El análisis informático mediante el algoritmo BLAST (con la opción “somewhat similar sequences”, “blastn”, que favorece el hallazgo de homologías poco restrictivas) de la secuencia de la construcción sh280 muestra un solo “hit” en el transcriptoma humano, coincidente con la entrada “NM_016448.2”, la secuencia de referencia de DTL, mientras que el mismo análisis llevado a cabo con la secuencia de la construcción sh2495 sólo da un hit no correspondiente a DTL, con homología total de sólo 14 de los 21 nucleótidos con la secuencia de referencia “NM_001136004.1”, correspondiente al gen MICAL3, “microtubule associated monooxygenase, calponin and LIM domain containing 3, transcript variant 2, mRNA”. El silenciamiento de este gen, por otra parte improbable debido a la escasa homología, difícilmente podría explicar el resultado obtenido en los ensayos *in vitro* e *in vivo*.

- El estudio mediante Western Blot de los efectos del silenciamiento de DTL sobre diversos reguladores de la transición G1/S sólo encontró variación en aquellos que han sido

extensamente validados como principales sustratos del complejo proteína ubiquitina ligasa CUL4/DDB1^{CDT2}; p21^{CIP1/WAF1}, p27^{KIP1} y p53. Además, los cambios detectados a nivel proteico no fueron acompañados de variaciones de los niveles de los mRNA, resultado consistente con el efecto de regulación postraducciona propia de la actividad de este complejo.

Por otra parte, aunque pueda parecer reduccionista validar uno sólo de los cientos de genes localizados en 1q (712 *loci* con esta localización incluidos en el microarray 133 Plus 2.0 de Affymetrix), cabe argumentar que el propósito de los ensayos de validación era demostrar la implicación de genes candidatos de este brazo cromosómico en el fenotipo molecular detectado (desregulación del ciclo celular) y que en ningún momento se propone a DTL como el único responsable de dicho fenotipo sino como uno de los principales, lo cual es demostrado por la alta coincidencia del patrón de expresión diferencial de los tumores 1qGSig y el patrón diferencial debido al silenciamiento de DTL en 3 líneas celulares con 1qG. Además, DTL es el único gen que cumple los requisitos de selección: es el único que combina una diferencia de expresión fiable y consistente (el mayor “d-value” de todo el análisis SAM) con un “fold-change” neto claramente superior al esperado y, sobre todo, es el único de entre los pocos genes que sobrepasan estos límites numéricos capaz de encajar funcionalmente en las predicciones de los análisis GSEA e Ingenuity.

Además, y aunque ello no formaba parte de los criterios de selección, el complejo proteína ubiquitina ligasa del que forma parte DTL, compuesto de numerosas actividades enzimáticas, supone una excelente diana terapéutica que podría constituir desde un punto de vista farmacológico un “talón de Aquiles” de los SE con 1qG (ver siguiente apartado). A este respecto es especialmente representativo el efecto *in vivo* conseguido por el silenciamiento de DTL, que fue capaz de frenar casi por completo la proliferación tumoral de las líneas celulares xenotransplantadas en los ratones NOD/SCID. Un fármaco convenientemente dirigido contra las actividades enzimáticas (como el ya desarrollado por la farmacéutica Takeda, ver apartado 5.9.3) de este complejo podría por tanto tener una actividad dramática sobre el desarrollo tumoral.

Finalmente, la demostración directa de la sobre-expresión de DTL sobre los tejidos de tumores de SE en los que se ha demostrado la mayor tasa proliferativa de los tumores 1qGSig mediante inmunohistoquímica no ha sido posible dado que no existe actualmente

ningún anticuerpo contra esta proteína que funcione en este tipo de ensayo. Sin embargo, la validación realizada mediante Real-Time RT qPCR se llevó a cabo con RNA perteneciente a los mismos tumores que comprende el TMA sobre el que se estudió Ki-67, y la correlación de los niveles de mRNA con elevados niveles proteicos ha sido demostrada en las líneas celulares de SE.

5.9 Función del gen DTL, implicación en cáncer y potencialidad como diana terapéutica

5.9.1 Complejos E3 ligasa tipo RING finger

La proteína DTL forma parte de un complejo proteína ubiquitina ligasa, conocido como CUL4/DDB1^{CDT2}, con evidentes funciones reguladoras de la progresión del ciclo celular. Este tipo de complejos multiproteicos se encargan de marcar con una cola de poliubiquitina a diversas especies proteicas que son subsiguientemente degradadas por el proteasoma 26S. El proceso conlleva la unión covalente de la cola de poliubiquitina a uno o varios residuos de lisina de la proteína marcada a través de 3 reacciones enzimáticas en cadena que comprenden la acción de una enzima E1-activadora que activa los monómeros de ubiquitina en una reacción dependiente de ATP, una enzima E2-conjugadora que transfiere el monómero activado mediante una reacción de transtiolación, y una enzima E3-ligasa que es la que finalmente conjuga la ubiquitina con el sustrato¹⁷³. Es precisamente este último paso el que confiere la especificidad por el sustrato y en el que DTL juega su papel.

Las E3 ligasas se dividen en 3 tipos principales según el tipo de dominio proteico que compone su núcleo enzimático (Figura 46, izquierda):

- Ligasas con dominios HECT (“homologous to E6-AP carboxy terminus”)¹⁷⁴, que conjugan directamente la ubiquitina activada con el sustrato formando un intermediario tioéster con ella.

- Ligasas con dominios de dedos RING (de “really interesting new gene”), dominios ricos en cisteína e histidina que coordinan dos átomos de Zn en una estructura de abrazaderas cruzadas. Este subtipo de E3 ligasas no transfieren directamente la ubiquitina al sustrato sino

que lo hacen de manera indirecta, activando de forma alostérica a la E2, de modo que, junto con otras proteínas como la culina, conforman un armazón sobre el cual la ubiquitina activada se transfiere de la E2 directamente al sustrato. Dentro de este grupo existen dos subclases: las que contienen en un mismo polipéptido los dominios de dedos RING y los dominios de unión al sustrato (siendo la proteína MDM2 el ejemplo prototípico) y las de tipo ligasas de dedos RING con culina (CRL, “cullin-containing RING-finger E3 ligases”), en las que el sustrato se une a una proteína intermedia, llamada culina, a su vez por medio de diversos adaptadores¹⁷⁵ (Figura 46, derecha).

- Ligasas con dominios U-box¹⁷⁶, que componen una familia menor de complejos de estructura similar a las E3 ligasas de dedos RING pero que carecen de los residuos quelantes de metal típicos de los dominios de dedos RING.

El subtipo de complejo CRL se compone de un núcleo del que forman parte un miembro de la familia culina (CUL1, 2, 3, 4A, 4B, 5 o 7 en humanos, cada uno codificado por un *locus* distinto) estrechamente unido a la proteína E3 ligasa de tipo dedos RING Rbx1 (también llamada ROC1) o Rnf7 (también llamada Rbx2 o ROC2). Esta base estructural conforma un armazón alargado cuya finalidad es hacer de ensamblaje de las actividades enzimáticas y el sustrato. El extremo C-terminal de la proteína de dedos RING (Rbx1 o Rnf7) se une a la E2 conjugada con ubiquitina activada, mientras que su extremo N-terminal permanece unido a un complejo multiproteico al que, a través de la culina y de diversos adaptadores, se une el sustrato¹⁷⁷ (Figura 46, derecha).

Los complejos CRL comprenden una amplia variedad de posibilidades de ensamblaje, que se diferencian principalmente en la culina que forma la base del armazón y, especialmente, en el receptor de sustrato, que es el elemento que determina la “elección” de la proteína que será marcada con ubiquitina para ser degradada y que es precisamente la función que desempeña DTL dentro de su correspondiente complejo CRL.

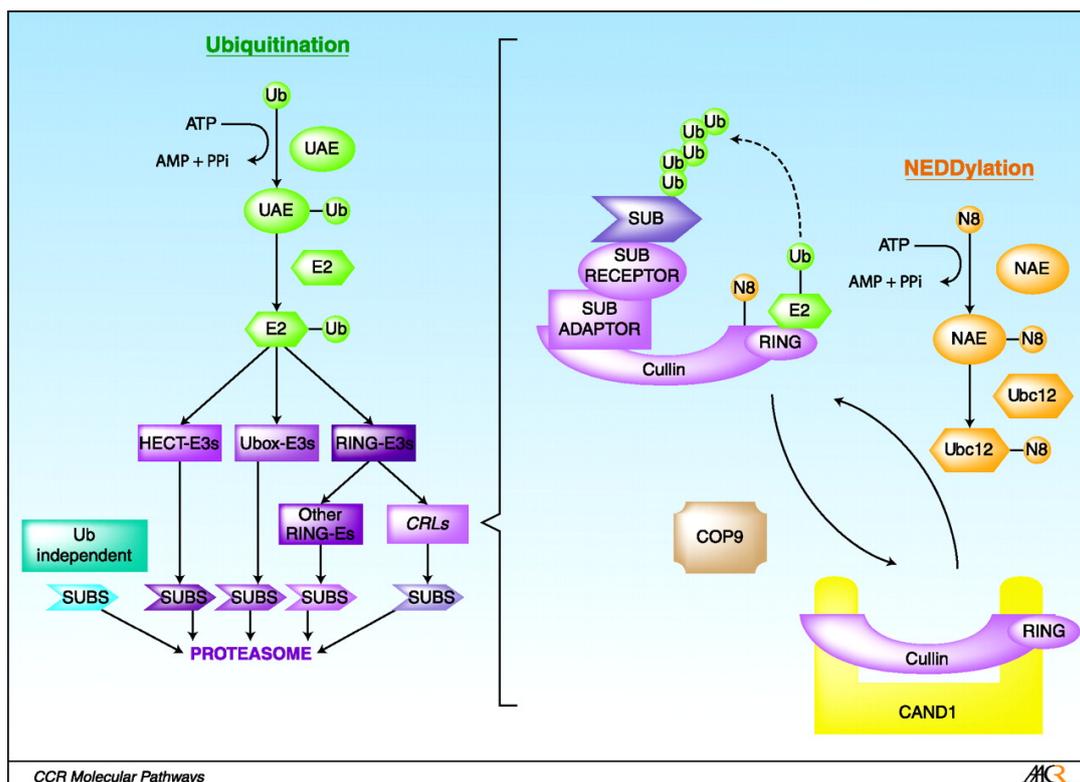


Figura 46. Esquema de los diferentes complejos proteína ubiquitina ligasa y detalle del ensamblaje y regulación de los complejos de tipo CRL. Reproducido de Soucy TA *et al.*¹⁷⁸ Targeting NEDD8-activated cullin-RING ligases for the treatment of cancer. Clin Cancer Res. 2009, 15(12):3912-6.

La tabla 6 resume los principales complejos CRL que se ensamblan a partir de las diferentes culinas, detallando los elementos variables asociados (adaptador y receptor de sustrato) y los sustratos reconocidos por cada variedad de complejo.

Culina	Adaptador / Tipo de Receptor de sustrato	Receptores de sustrato principales	Sustratos
CUL1	SKP1/F-box	SKP2, β-TrCP, FBW7	p27, p-IκBα, β-catenina, ciclina E, c-MYC, c- JUN, mTOR
CUL2,5	Elongina-BC/ SOCS-box	VHL	HIF1α
CUL3	Dominio BTB	KEAP1	NRF2
CUL4A	DDB1 / DCAF	DDB2 / DTL (CDT2)	Histonas H2A, H3 y H4, XPC / CDT1, p27, p21, p53,

Tabla 6. Combinaciones de elementos que conforman los principales complejos CRL

DTL forma parte del complejo CUL4/DDB1, o CRL4, en el que la culina 4A ó 4B (prácticamente idénticas, aunque con ciertas variaciones funcionales) es la base estructural a la que está unida la E3 ligasa ROC1 (Rbx1) o ROC2 (Rnf7) y que tiene como adaptador a la

proteína DDB1, una proteína de gran tamaño (127 KDa) compuesta de 7 láminas β , muy conocida por sus funciones en la reparación del DNA¹⁷⁹. El adaptador DDB1 une con alta afinidad a una familia de proteínas con dominios WD40, conocida como “DCAFs” (“Ddb1 and Cul4 associated factors”), que son las que hacen de receptor de sustrato en este complejo. Las DCAF se caracterizan por poseer varias repeticiones del dominio WD40, un dominio proteico de unos 40 residuos de extensión que contiene un dipéptido Gly-His separado por 20-30 residuos de otro dipéptido Trp-Asp (WD) y que conforman estructuras tipo lámina beta. DTL y DDB2, ambas proteínas DCAF, son los principales receptores de sustrato del complejo CRL4¹⁷⁹.

Cuando el complejo CRL4 funciona “guiado” por DTL se le denomina CUL4/DDB1^{CDT2} y los sustratos reconocidos y marcados con ubiquitina son CDT1^{136,138,180-181}, p21^{WAF1/CIP1 140-142}, p27^{KIP1 139}, p53¹³⁷ y E2F1¹⁸². Además, al menos para la ubiquitinación de CDT1 y p21^{WAF1/CIP1} parece ser necesario el reclutamiento del complejo CRL4 a la cromatina puesto que para que se produzca la interacción con estos sustratos se requiere la presencia del complejo PCNA-DNA^{138,183-184}.

5.9.2 Expresión de DTL en tejidos humanos y de ratón

La expresión de DTL está restringida en ratón a los estadios tempranos del desarrollo embrionario y el modelo “knockout” muere pronto tras las primeras divisiones del cigoto¹⁸⁵.

En humanos no existe expresión de DTL en tejidos adultos, salvo en aquellos con proliferación activa, como los hematopoyéticos, según se deduce de los patrones de expresión observables en la colección de microarrays de expresión del “GeneAtlas” (Figura 47).

Por tanto, los patrones de expresión de DTL en ratón y humano redundan en su función activadora del ciclo celular, puesto que su expresión está restringida al desarrollo embrionario temprano, momento de la vida del individuo donde la división celular es más activa siendo además su función indispensable para tal proceso. Del mismo modo, estas características funcionales y de expresión confieren a DTL una evidente potencialidad oncogénica. A pesar de ello pocos trabajos han involucrado aún a este gen en procesos de oncogénesis. Hasta el

momento sólo se le ha relacionado con carcinoma hepático¹⁸⁶, cáncer de mama¹⁸⁷ y ha sido sugerido como potencial gen candidato en 1q en rhabdomyosarcomas con 1qG¹⁸⁸.

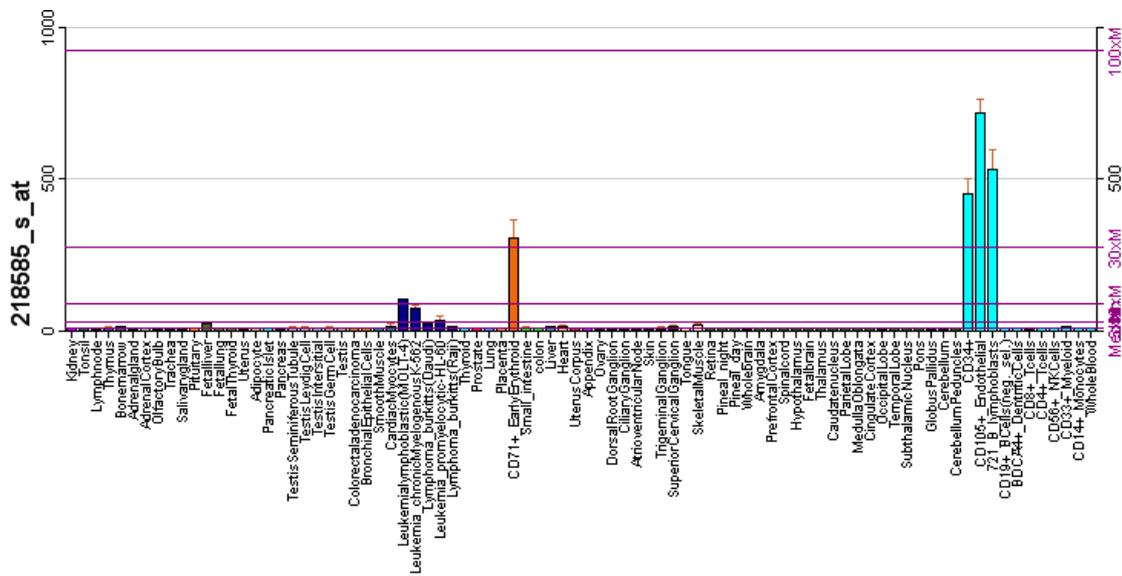


Figura 47. Patrón de expresión del mRNA de DTL en tejidos humanos adultos y tumores humanos. Extraído de BioGPS (<http://biogps.gnf.org>).

5.9.3 Regulación de la actividad de los complejos CRL y aplicación en terapia dirigida

Finalmente, respecto a la potencialidad de este gen como diana terapéutica, cabe destacar que los complejos CRL están sometidas a procesos de regulación de su actividad tanto activadores como inhibidores muy específicos (Figura 46, derecha):

- Los complejos CRL están regulados positivamente por complejos de nedilación. Dichos complejos son a su vez maquinarias enzimáticas similares a las de ubiquitinación, salvo que la ubiquitina es sustituida por una proteína muy similar (“ubiquitin-like”) llamada NEDD8 y que su sustrato es la propia culina de los complejos CRL. Comprenden actividades enzimáticas análogas a las de lo complejos de ubiquitinación¹⁸⁹: la “Nedd8-Activating Enzyme” o NAE, heterodímero de NAE1 y UBA3 y análogo a la E1, usa ATP para producir primero un intermediario activo NEDD8-adenilato a partir del cual une luego la NEDD8 a un residuo cisteína de su estructura, formando un complejo NAE-NEDD8 tioéster activo. La enzima UBC12 (o UB2M), análoga a la E2, recibe en un residuo cisteína activo la NEDD8 activada, de nuevo mediante enlace tioéster. Finalmente una E3 ligasa conjuga la NEDD8 al sustrato,

un residuo lisina próximo al extremo C-terminal de la culina del complejo CRL, que queda así “nedilado”. La homología de esta maquinaria con el propio complejo CRL al que activa es tal que recientes trabajos han identificado a la proteína SCCRO (DCN1) como la actividad E3 ligasa¹⁹⁰ del proceso de nedilación, poseedora también de funciones de armazón similares a los de las culinas.

El complejo CRL debe tener una copia de NEDD8 conjugada para comportarse como una holoenzima con actividad proteína ubiquitina ligasa¹⁹¹⁻¹⁹². Es decir, sólo es funcional tras ser modificado de esta forma, ya que sólo tras la unión de la NEDD8 el complejo es capaz de reclutar a la E2-ubiquitina¹⁹³. Además, se ha demostrado que tras la nedilación se producen grandes cambios conformacionales en el complejo que aproximan entre sí a la E2-ubiquitina y al sustrato¹⁹⁴.

- De manera complementaria, los complejos CRL son inactivados mediante denedilación, llevada a cabo por la actividad isopeptidasa de la subunidad metaloproteasa del complejo COP9 signalosoma¹⁹⁵, que elimina la NEDD8 conjugada con la culina del complejo CRL. La denedilación promueve la disociación de los componentes del complejo y favorece la captura del conjunto culina-E3 ligasa por parte de la proteína CAND1 (“cullin associated NEDD8 dissociated-1”), que inhibe estéricamente el ensamblaje de sus elementos¹⁹⁶⁻¹⁹⁷. Este proceso favorece además el reciclaje del complejo, de manera que el conjunto culina-E3 ligasa queda disponible para ser ensamblado de nuevo con diferentes adaptadores y receptores de sustrato.

Estos sistemas de regulación son de especial interés puesto que muy recientemente la casa farmacéutica Takeda (ahora propietaria de Millenium) ha desarrollado un fármaco, conocido como MLN4924, dirigido específicamente a la inhibición de la enzima NAE, primera enzima del proceso de nedilación¹⁹⁸. En este estudio el fármaco demostró una potente actividad citotóxica *in vitro* e *in vivo*, además de una alta especificidad, ya que las actividades enzimáticas análogas “ubiquitin-activating enzyme” (UAE o UBA1, actividad E1) y “SUMO-activating enzyme” (SAE, heterodímero de SAE1 y UBA2) presentaron una IC50 de inhibición de entre 375 y 2500 veces superiores (la IC50 de inhibición de la actividad NAE fue de tan sólo 0.004 µM). La inducción de apoptosis obtenida en las líneas celulares testadas en este trabajo pareció derivar de una síntesis descontrolada de DNA (re-replicación) y acumulación de las células en la fase S del ciclo celular y del consecuente daño de DNA acumulado.

La ventaja de este nuevo fármaco respecto al ya extensamente validado bortezomib, que actúa directamente inhibiendo la acción del proteasoma 26S, es que el fármaco MLN4924 sólo bloquea un 9% del tráfico de reciclado de proteínas de la vía de ubiquitinación (en contraste con la inhibición del 50% que produce bortezomib), ya que inhibiendo el proceso de nedilación sólo se ven afectadas las actividades E3 ligasa de los complejos CRL. Además, en la línea celular de cáncer colorrectal HCT-116 el fármaco MLN4924 indujo la acumulación de p27^{KIP1} y CDT1, entre otros conocidos sustratos de los complejos CRL, demostrando su especificidad. La inducción de re-replicación de DNA y acumulación de células en fase S apunta a que es precisamente la acumulación de CDT1 (uno de los principales sustratos del complejo CUL4/DDB1^{CDT2}) la que actuaría como principal mediadora del efecto del fármaco puesto que esta proteína se asocia a los complejos pre-replicativos del DNA y su degradación es fundamental para evitar que se vuelvan a disparar los orígenes de replicación que ya han sido abiertos¹⁹⁹. De hecho la sobre-expresión de CDT1 o el silenciamiento de su inhibidor, la proteína Geminina, producen un efecto prácticamente idéntico al debido a la acción del fármaco MLN4924²⁰⁰⁻²⁰¹. Además, la detección de fosforilación en la Ser317 de la proteína CHK1 tras la aplicación de la droga es consistente con esta idea, puesto que el proceso de re-replicación del DNA tiene efectos conocidos de daño en el DNA, que son señalizados a través de la ruta ATM/ATR iniciada con dicha fosforilación.

De acuerdo con el portal “web” del NCI (“National Cancer Institutes”), en Estados Unidos existen actualmente 4 ensayos clínicos activos en fase I con este fármaco, que están siendo aplicados a pacientes de melanoma, leucemia mielógena aguda y síndrome mielodisplásico de alto grado, linfoma y mieloma múltiple, y de tumores no hematológicos, respectivamente.

5.10 Modelo molecular propuesto para la contribución de DTL a la oncogénesis y desarrollo del SE

Pocos son los trabajos previos que han cribado muestras de SE en busca de alteraciones secundarias. En estos escasos estudios se ha demostrado que estas alteraciones, aunque infrecuentes en el caso de p53 y p16^{INK4}/p14^{ARF}, tienen un importante impacto clínico en SE. Es destacable que la sobre-expresión del gen validado en este trabajo, DTL, tiene un efecto en la reducción de los niveles de importantes supresores tumorales involucrados en funciones similares (inhibidores de la progresión de ciclo celular tales como el propio p53, p21^{WAF1/CIP1}

y p27^{KIP1}), lo cual podría explicar además la ausencia de su expresión proteica detectada en buena parte de las muestras tumorales de SE²⁰².

Como ya se ha comentado en el apartado 1.7, aunque la opinión tradicional es que la fusión EWS-ETS es no sólo necesaria sino también suficiente para el establecimiento y desarrollo del SE, diversas evidencias apuntan al requerimiento de otras alteraciones para la consecución de estos procesos. Especialmente concluyentes son las evidencias obtenidas de la expresión exógena de la fusión quimérica en diversos contextos celulares, en ninguno de los cuales se ha conseguido la transformación celular, incluidos aquellos que han demostrado permisividad a su presencia^{83,85}.

Por otra parte, tanto las alteraciones secundarias actualmente descritas, incluida la 1qG/sobre-expresión de DTL, como la propia alteración primaria (EWS-ETS) inciden en la desregulación de un mismo proceso, el control de la progresión del ciclo celular. En este sentido, un reciente trabajo que ha analizado el impacto molecular de la fusión quimérica sobre los diversos procesos y funciones celulares haciendo uso de una gran colección de perfiles de expresión tumorales (pertenecientes a 59 tumores primarios y al silenciamiento de EWS-ETS en 5 líneas celulares) ha señalado precisamente a esta (la regulación del ciclo celular) como la principal función alterada por la fusión EWS-FLI1⁸¹. De hecho, tal y como se explica en el apartado 1.6.3, MYC, ciclina D y p57, todos ellos importantes reguladores de esta función celular, están entre las dianas génicas mejor conocidas de la acción transactivadora aberrante de EWS-ETS.

Teniendo todo ello en cuenta se propone el modelo molecular representado en la Figura 48 para explicar el marcado impacto de la 1qG y de su perfil de expresión local (1qGSig) en la supervivencia de los pacientes, especialmente considerando la sobre-expresión de DTL asociada a ella. De acuerdo con este modelo, EWS-ETS tendría un potente efecto promoviendo la progresión del ciclo celular especialmente a través de la sobre-expresión de ciclina D1, lo cual podría conllevar la independencia de las señales mitogénicas externas pero no la liberación del resto de controles del ciclo. La alteración 1qG, especialmente considerando la subsiguiente sobre-expresión de DTL, incidiría precisamente a este nivel eliminando buena parte de los inhibidores del ciclo a través de su potente efecto pleiotrópico sobre ellos. De esta manera los tumores que por un proceso de inestabilidad cromosómica incipiente ganaran el brazo 1q en una fase temprana del desarrollo tumoral podrían verse

favorecidos por una pronta ventaja de crecimiento, que además propiciaría la acumulación del resto de alteraciones necesarias para la evolución tumoral, tales como las requeridas para la producción de metástasis. Este modelo es consistente con la mayor tasa proliferativa encontrada en los tumores 1qGSig.

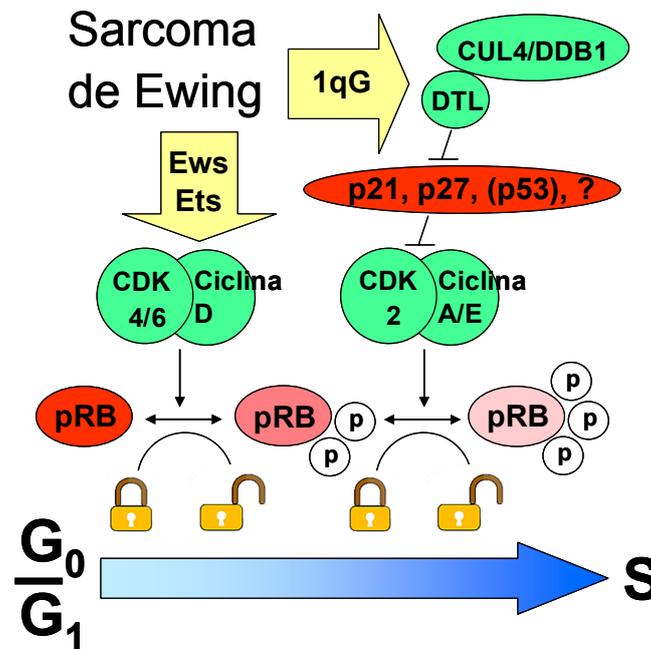


Figura 48. Modelo molecular de la posible contribución de 1qG a la patogénesis del SE. Los inhibidores del ciclo celular se representan en color rojo, mientras los activadores aparecen en verde. El esquema no pretende ser exhaustivo y por ello sólo se han representado las actividades directamente involucradas en dicho modelo.

Como se observa en la Tabla 6, los complejos CRL (especialmente los CRL1 y CRL4) tienen actividades directas muy diversificadas sobre todo tipo de reguladores del ciclo. De hecho, tal y como se aprecia en la Figura 33, es de esperar el descubrimiento de nuevas funciones inhibitoras de la proliferación celular bajo el control del complejo CUL4/DDB1^{CDT2} puesto que la línea celular RM82, la cual carece de todos sus sustratos conocidos con dicha función (pues presenta una forma p53 dominante negativa R273H¹¹¹), se ve igualmente afectada de manera dramática por el silenciamiento de DTL, tanto *in vitro* como *in vivo*. Esto, unido al requerimiento indispensable de la función de DTL en el desarrollo temprano embrionario (lo cual no sucede en el caso de su equivalente en el complejo CRL1, SKP2²⁰³) sugiere una función de favorecimiento de la proliferación celular de este complejo CRL4, que se ha empezado a estudiar tan recientemente como en el año 2003, más extensa de la actualmente conocida.

5.11 Otros genes candidatos en 1q validados en la literatura previa

Pocos trabajos han validado genes candidatos en el brazo cromosómico 1q. En un estudio reciente²⁰⁴ los autores definieron una RMS de 3.06 mb situada en 1q32.1 como resultado de un estudio de pérdidas genómicas mediante qPCR multiplexada de 14 STS (“sequence tagged sites”) localizadas en 1q25.3-1q41, en líneas celulares y muestras tumorales de diversos tipos tumorales, especialmente retinoblastoma (n=55). El estudio de expresión diferencial de las muestras tumorales respecto a los tejidos sanos correspondientes identificó al gen KIF14 como el principal gen candidato de la región, sobre-expresado en la mayoría de muestras tumorales y correlacionado con una peor supervivencia de los pacientes.

En el presente estudio este gen también presenta sobre-expresión en los tumores 1qGsig, aunque con menor significación que DTL (“d-value” = 6.6, “fold-change” = 3.37). Además, los niveles de KIF14 se reducen a la mitad tras el silenciamiento de DTL, según los datos de los microarrays de expresión. Este resultado es fácil de explicar si tenemos en cuenta que DTL se comporta como un fuerte regulador del ciclo celular en las líneas celulares en que ha sido estudiado y que KIF14 es una proteína de la familia de las kinesinas, involucradas en el ensamblaje de los microtúbulos durante la mitosis, por lo que no es de extrañar que los niveles de este gen caigan tras la parada de ciclo inducida por el silenciamiento de DTL. Por lo tanto la sobre-expresión de KIF14 podría estar simplemente reflejando el mayor estado proliferativo que caracteriza a las células tumorales y que las diferencia de sus respectivos tejidos sanos. De una forma similar, en nuestros datos reflejaría la mayor proliferación detectada en los tumores con 1qG. Lo que parece claro es que este gen, KIF14, carece de funciones regulatorias del ciclo celular y su sobre-expresión es probablemente un reflejo del fenotipo asociado a la 1qG, careciendo de actividad como “driver” genético *per se*.

6. CONCLUSIONES

1. Las ANC consideradas globalmente tienen relevancia clínica ya que su acumulación correlaciona con una peor supervivencia de los pacientes con SE, siendo el parámetro PGA un mejor estimador global de las ANC que el previamente usado por otros investigadores, NTA.
2. A nivel individual, la 1qG es la ANC que demuestra un mayor impacto en la clínica de los pacientes de SE, mostrando una fuerte asociación con la recaída o progresión de la enfermedad y con una notable peor supervivencia de los pacientes, tanto total como libre de recaída.
3. La 1qG es independiente del valor global de PGA, por lo que no es un reflejo de la cantidad global de ANC en las muestras tumorales y la presencia de ANC individuales asociadas a ella no supone una diferencia clínica significativa.
4. La 1qG no se asocia a las ANC que afectan a los *loci* CDKN2A y TP53, las otras 2 únicas alteraciones moleculares secundarias con relevancia clínica halladas hasta ahora en SE, por lo que una ulterior asociación entre la 1qG y la pérdida de función de la proteínas codificadas por estos *loci* es improbable.
5. La 1qG tiene una incidencia poblacional de alrededor del doble de la de estas alteraciones secundarias previamente descritas y por su naturaleza (ganancia cromosómica y sobre-expresión génica asociada) no requiere de un segundo “hit”, a diferencia de aquellas.
6. La 1qG se asocia con un perfil transcripcional de ciclo celular desregulado.

7. El gen DTL, localizado en 1q, es el principal candidato de la región puesto que es el que presenta una sobre-expresión más significativa y consistente en las muestras tumorales con 1qG, a la vez de ser el gen en 1q con la mayor diferencia neta de expresión, y puesto que sus función esta estrechamente relacionada con el fenotipo molecular predicho por los estudios *in silico*.

8. La función del gen DTL ha demostrado ser indispensable para la progresión del ciclo celular y la supervivencia de las líneas celulares de SE con 1qG que lo sobre-expresan, tanto *in vitro* como *in vivo*.

9. La acción reguladora del ciclo que ejerce DTL en las líneas celulares de SE es ejercida a través de, al menos, p21^{WAF1/CIP1}, p27^{KIP1} y p53, aunque en la línea celular RM82 lo hace probablemente a través de otros efectores.

10. La sobre-expresión de DTL es uno de los principales elementos que contribuyen al perfil de expresión diferencial de los tumores de SE con 1qG o con el perfil asociado de expresión 1qGSig.

11. Los tumores de SE con 1qG o con el perfil asociado de expresión 1qGSig, en los que además ha sido demostrada la sobre-expresión de DTL, tienen una mayor tasa de proliferación que el resto de tumores.

12. Las características funcionales y de regulación enzimática del complejo proteína ubiquitina ligasa del que forma parte DTL, CUL4/DDB1^{CDT2}, lo convierten en un prometedor objetivo de terapia dirigida en el subtipo agresivo de SE con 1qG.

7. BIBLIOGRAFÍA

1. Stout, A.P. A tumor of ulnar nerve. *Proc NY Pathol Soc* **21**, 2-12 (1918).
2. Ewing, J. Diffuse endotelioma of bone. *Proc Nat Pathol Soc* **21**, 17-24 (1921).
3. Angervall, L. & Enzinger, F.M. Extraskeletal neoplasm resembling Ewing's sarcoma. *Cancer* **36**, 240-251 (1975).
4. Askin, F.B., Rosai, J., Sibley, R.K., Dehner, L.P. & McAlister, W.H. Malignant small cell tumor of the thoracopulmonary region in childhood: a distinctive clinicopathologic entity of uncertain histogenesis. *Cancer* **43**, 2438-2451 (1979).
5. Jaffe, R., *et al.* The neuroectodermal tumor of bone. *Am J Surg Pathol* **8**, 885-898 (1984).
6. Aurias, A., Rimbaut, C., Buffe, D., Zucker, J.M. & Mazabraud, A. Translocation involving chromosome 22 in Ewing's sarcoma. A cytogenetic study of four fresh tumors. *Cancer Genet Cytogenet* **12**, 21-25 (1984).
7. Turc-Carel, C., Philip, I., Berger, M.P., Philip, T. & Lenoir, G.M. Chromosome study of Ewing's sarcoma (ES) cell lines. Consistency of a reciprocal translocation t(11;22)(q24;q12). *Cancer Genet Cytogenet* **12**, 1-19 (1984).
8. Whang-Peng, J., *et al.* Chromosome translocation in peripheral neuroepithelioma. *N Engl J Med* **311**, 584-585 (1984).
9. Delattre, O., *et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162-165 (1992).
10. Delattre, O., *et al.* The Ewing family of tumors--a subgroup of small-round-cell tumors defined by specific chimeric transcripts. *N Engl J Med* **331**, 294-299 (1994).
11. Fletcher CDM, U.K., Mertens F. *World Health Organization: Classification of Tumours. Pathology and Genetics of Tumours of Soft Tissue and Bone.*, (IARC Press, Lyon, 2002).

-
12. Ginsberg, J., Woo, SY, Hicks, MJ, Horowitz, ME. Ewing's sarcoma family of tumors: Ewing's sarcoma of bone and soft tissue and the peripheral primitive neuroectodermal tumors. *In: Principles and Practice of Pediatric Oncology 4th ed, Pizzo, PA, Poplack, DG (Eds), Lippincott, Williams and Wilkins, Philadelphia (2002).*
 13. Dorfman, H.D. & Czerniak, B. Bone cancers. *Cancer* **75**, 203-210 (1995).
 14. Raney, R.B., *et al.* Ewing's sarcoma of soft tissues in childhood: a report from the Intergroup Rhabdomyosarcoma Study, 1972 to 1991. *J Clin Oncol* **15**, 574-582 (1997).
 15. Meier, V.S., Kuhne, T., Jundt, G. & Gudat, F. Molecular diagnosis of Ewing tumors: improved detection of EWS-FLI-1 and EWS-ERG chimeric transcripts and rapid determination of exon combinations. *Diagn Mol Pathol* **7**, 29-35 (1998).
 16. Dagher, R., *et al.* Molecular confirmation of Ewing sarcoma. *J Pediatr Hematol Oncol* **23**, 221-224 (2001).
 17. Dejana, E. Endothelial cell-cell junctions: happy together. *Nat Rev Mol Cell Biol* **5**, 261-270 (2004).
 18. Schmidt, D., Herrmann, C., Jurgens, H. & Harms, D. Malignant peripheral neuroectodermal tumor and its necessary distinction from Ewing's sarcoma. A report from the Kiel Pediatric Tumor Registry. *Cancer* **68**, 2251-2259 (1991).
 19. Banerjee, S.S., Agbamu, D.A., Eyden, B.P. & Harris, M. Clinicopathological characteristics of peripheral primitive neuroectodermal tumour of skin and subcutaneous tissue. *Histopathology* **31**, 355-366 (1997).
 20. Ehrig, T., Billings, S.D. & Fanburg-Smith, J.C. Superficial primitive neuroectodermal tumor/Ewing sarcoma (PN/ES): same tumor as deep PN/ES or new entity? *Ann Diagn Pathol* **11**, 153-159 (2007).
 21. Cotterill, S.J., *et al.* Prognostic factors in Ewing's tumor of bone: analysis of 975 patients from the European Intergroup Cooperative Ewing's Sarcoma Study Group. *J Clin Oncol* **18**, 3108-3114 (2000).
-

-
22. Bacci, G., *et al.* Prognostic factors in non-metastatic Ewing's sarcoma tumor of bone: an analysis of 579 patients treated at a single institution with adjuvant or neoadjuvant chemotherapy between 1972 and 1998. *Acta Oncol* **45**, 469-475 (2006).
 23. Ahrens, S., *et al.* Evaluation of prognostic factors in a tumor volume-adapted treatment strategy for localized Ewing sarcoma of bone: the CESS 86 experience. Cooperative Ewing Sarcoma Study. *Med Pediatr Oncol* **32**, 186-195 (1999).
 24. de Alava, E., *et al.* EWS-FLI1 fusion transcript structure is an independent determinant of prognosis in Ewing's sarcoma. *J Clin Oncol* **16**, 1248-1255 (1998).
 25. Le Deley, M.C., *et al.* Impact of EWS-ETS Fusion Type on Disease Progression in Ewing's Sarcoma/Peripheral Primitive Neuroectodermal Tumor: Prospective Results From the Cooperative Euro-E.W.I.N.G. 99 Trial. *J Clin Oncol* (2010).
 26. van Doorninck, J.A., *et al.* Current treatment protocols have eliminated the prognostic advantage of type 1 fusions in Ewing sarcoma: a report from the Children's Oncology Group. *J Clin Oncol* **28**, 1989-1994 (2010).
 27. Schleiermacher, G., *et al.* Increased risk of systemic relapses associated with bone marrow micrometastasis and circulating tumor cells in localized ewing tumor. *J Clin Oncol* **21**, 85-91 (2003).
 28. Huang, H.Y., *et al.* Ewing sarcomas with p53 mutation or p16/p14ARF homozygous deletion: a highly lethal subset associated with poor chemoresponse. *J Clin Oncol* **23**, 548-558 (2005).
 29. Haeusler, J., *et al.* The value of local treatment in patients with primary, disseminated, multifocal Ewing sarcoma (PDMES). *Cancer* **116**, 443-450 (2010).
 30. Juergens, C., *et al.* Safety assessment of intensive induction with vincristine, ifosfamide, doxorubicin, and etoposide (VIDE) in the treatment of Ewing tumors in the EURO-E.W.I.N.G. 99 clinical trial. *Pediatr Blood Cancer* **47**, 22-29 (2006).
 31. Paulussen, M., *et al.* Results of the EICESS-92 Study: two randomized trials of Ewing's sarcoma treatment--cyclophosphamide compared with ifosfamide in standard-
-

-
- risk patients and assessment of benefit of etoposide added to standard treatment in high-risk patients. *J Clin Oncol* **26**, 4385-4393 (2008).
32. Paulussen, M., *et al.* Localized Ewing tumor of bone: final results of the cooperative Ewing's Sarcoma Study CESS 86. *J Clin Oncol* **19**, 1818-1829 (2001).
33. Klingebiel, T., *et al.* Treatment of children with metastatic soft tissue sarcoma with oral maintenance compared to high dose chemotherapy: report of the HD CWS-96 trial. *Pediatr Blood Cancer* **50**, 739-745 (2008).
34. Hustu, H.O., Holton, C., James, D., Jr. & Pinkel, D. Treatment of Ewing's sarcoma with concurrent radiotherapy and chemotherapy. *J Pediatr* **73**, 249-251 (1968).
35. Nesbit, M.E., Jr., *et al.* Multimodal therapy for the management of primary, nonmetastatic Ewing's sarcoma of bone: a long-term follow-up of the First Intergroup study. *J Clin Oncol* **8**, 1664-1674 (1990).
36. Burgert, E.O., Jr., *et al.* Multimodal therapy for the management of nonpelvic, localized Ewing's sarcoma of bone: intergroup study IESS-II. *J Clin Oncol* **8**, 1514-1524 (1990).
37. Jurgens, H., *et al.* Multidisciplinary treatment of primary Ewing's sarcoma of bone. A 6-year experience of a European Cooperative Trial. *Cancer* **61**, 23-32 (1988).
38. Kung, F.H., *et al.* Ifosfamide/etoposide combination in the treatment of recurrent malignant solid tumors of childhood. A Pediatric Oncology Group Phase II study. *Cancer* **71**, 1898-1903 (1993).
39. Grier, H.E., *et al.* Addition of ifosfamide and etoposide to standard chemotherapy for Ewing's sarcoma and primitive neuroectodermal tumor of bone. *N Engl J Med* **348**, 694-701 (2003).
40. Schuck, A., *et al.* Local therapy in localized Ewing tumors: results of 1058 patients treated in the CESS 81, CESS 86, and EICESS 92 trials. *Int J Radiat Oncol Biol Phys* **55**, 168-177 (2003).
-

-
41. Sailer, S.L., Harmon, D.C., Mankin, H.J., Truman, J.T. & Suit, H.D. Ewing's sarcoma: surgical resection as a prognostic factor. *Int J Radiat Oncol Biol Phys* **15**, 43-52 (1988).
 42. Pinkerton, C.R., *et al.* Treatment strategies for metastatic Ewing's sarcoma. *Eur J Cancer* **37**, 1338-1344 (2001).
 43. Bernstein, M.L., *et al.* Intensive therapy with growth factor support for patients with Ewing tumor metastatic at diagnosis: Pediatric Oncology Group/Children's Cancer Group Phase II Study 9457--a report from the Children's Oncology Group. *J Clin Oncol* **24**, 152-159 (2006).
 44. Bertolotti, A., Lutz, Y., Heard, D.J., Chambon, P. & Tora, L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J* **15**, 5022-5031 (1996).
 45. Ohno, T., *et al.* The EWS gene, involved in Ewing family of tumors, malignant melanoma of soft parts and desmoplastic small round cell tumors, codes for an RNA binding protein with novel regulatory domains. *Oncogene* **9**, 3087-3097 (1994).
 46. Aman, P., *et al.* Expression patterns of the human sarcoma-associated genes FUS and EWS and the genomic structure of FUS. *Genomics* **37**, 1-8 (1996).
 47. Belyanskaya, L.L., Delattre, O. & Gehring, H. Expression and subcellular localization of Ewing sarcoma (EWS) protein is affected by the methylation process. *Exp Cell Res* **288**, 374-381 (2003).
 48. Chan, D., *et al.* Transformation induced by Ewing's sarcoma associated EWS/FLI-1 is suppressed by KRAB/FLI-1. *Br J Cancer* **88**, 137-145 (2003).
 49. Bergeron, D., Poliquin, L., Kozak, C.A. & Rassart, E. Identification of a common viral integration region in Cas-Br-E murine leukemia virus-induced non-T-, non-B-cell lymphomas. *J Virol* **65**, 7-15 (1991).

-
50. Ben-David, Y., Giddens, E.B., Letwin, K. & Bernstein, A. Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes Dev* **5**, 908-918 (1991).
 51. Rao, V.N., Ohno, T., Prasad, D.D., Bhattacharya, G. & Reddy, E.S. Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene* **8**, 2167-2173 (1993).
 52. Truong, A.H. & Ben-David, Y. The role of Fli-1 in normal cell function and malignant transformation. *Oncogene* **19**, 6482-6489 (2000).
 53. Melet, F., Motro, B., Rossi, D.J., Zhang, L. & Bernstein, A. Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Mol Cell Biol* **16**, 2708-2718 (1996).
 54. Maroulakou, I.G. & Bowe, D.B. Expression and function of Ets transcription factors in mammalian development: a regulatory network. *Oncogene* **19**, 6432-6442 (2000).
 55. Deveaux, S., *et al.* Analysis of the thrombopoietin receptor (MPL) promoter implicates GATA and Ets proteins in the coregulation of megakaryocyte-specific genes. *Blood* **87**, 4678-4685 (1996).
 56. Bastian, L.S., Kwiatkowski, B.A., Breininger, J., Danner, S. & Roth, G. Regulation of the megakaryocytic glycoprotein IX promoter by the oncogenic Ets transcription factor Fli-1. *Blood* **93**, 2637-2644 (1999).
 57. Athanasiou, M., *et al.* Increased expression of the ETS-related transcription factor FLI-1/ERGB correlates with and can induce the megakaryocytic phenotype. *Cell Growth Differ* **7**, 1525-1534 (1996).
 58. Howard, J.C., Yousefi, S., Cheong, G., Bernstein, A. & Ben-David, Y. Temporal order and functional analysis of mutations within the Fli-1 and p53 genes during the erythroleukemias induced by F-MuLV. *Oncogene* **8**, 2721-2729 (1993).
 59. Pereira, R., *et al.* FLI-1 inhibits differentiation and induces proliferation of primary erythroblasts. *Oncogene* **18**, 1597-1608 (1999).
-

-
60. Tamir, A., *et al.* Fli-1, an Ets-related transcription factor, regulates erythropoietin-induced erythroid proliferation and differentiation: evidence for direct transcriptional repression of the Rb gene during differentiation. *Mol Cell Biol* **19**, 4452-4464 (1999).
 61. Lesault, I., Quang, C.T., Frampton, J. & Ghysdael, J. Direct regulation of BCL-2 by FLI-1 is involved in the survival of FLI-1-transformed erythroblasts. *EMBO J* **21**, 694-703 (2002).
 62. Mao, X., Miesfeldt, S., Yang, H., Leiden, J.M. & Thompson, C.B. The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. *J Biol Chem* **269**, 18216-18222 (1994).
 63. Bailly, R.A., *et al.* DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. *Mol Cell Biol* **14**, 3230-3241 (1994).
 64. Ohno, T., Rao, V.N. & Reddy, E.S. EWS/Fli-1 chimeric protein is a transcriptional activator. *Cancer Res* **53**, 5859-5863 (1993).
 65. May, W.A., *et al.* The Ewing's sarcoma EWS/FLI-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than FLI-1. *Mol Cell Biol* **13**, 7393-7398 (1993).
 66. Nakatani, F., *et al.* Identification of p21WAF1/CIP1 as a direct target of EWS-Fli1 oncogenic fusion protein. *J Biol Chem* **278**, 15105-15115 (2003).
 67. Dauphinot, L., *et al.* Analysis of the expression of cell cycle regulators in Ewing cell lines: EWS-FLI-1 modulates p57KIP2 and c-Myc expression. *Oncogene* **20**, 3258-3265 (2001).
 68. Fukuma, M., Okita, H., Hata, J. & Umezawa, A. Upregulation of Id2, an oncogenic helix-loop-helix protein, is mediated by the chimeric EWS/ets protein in Ewing sarcoma. *Oncogene* **22**, 1-9 (2003).
 69. Hahm, K.B. Repression of the gene encoding the TGF-beta type II receptor is a major target of the EWS-FLI1 oncoprotein. *Nat Genet* **23**, 481 (1999).
-

-
70. Scotlandi, K., *et al.* Insulin-like growth factor I receptor-mediated circuit in Ewing's sarcoma/peripheral neuroectodermal tumor: a possible therapeutic target. *Cancer Res* **56**, 4570-4574 (1996).
 71. Strammiello, R., *et al.* Impact of IGF-I/IGF-IR circuit on the angiogenetic properties of Ewing's sarcoma cells. *Horm Metab Res* **35**, 675-684 (2003).
 72. Prieur, A., Tirode, F., Cohen, P. & Delattre, O. EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3. *Mol Cell Biol* **24**, 7275-7283 (2004).
 73. Toretsky, J.A., Kalebic, T., Blakesley, V., LeRoith, D. & Helman, L.J. The insulin-like growth factor-I receptor is required for EWS/FLI-1 transformation of fibroblasts. *J Biol Chem* **272**, 30822-30827 (1997).
 74. Mendiola, M., *et al.* The orphan nuclear receptor DAX1 is up-regulated by the EWS/FLI1 oncoprotein and is highly expressed in Ewing tumors. *Int J Cancer* **118**, 1381-1389 (2006).
 75. Abaan, O.D., *et al.* PTPL1 is a direct transcriptional target of EWS-FLI1 and modulates Ewing's Sarcoma tumorigenesis. *Oncogene* **24**, 2715-2722 (2005).
 76. Kikuchi, R., *et al.* Ewing's sarcoma fusion protein, EWS/Fli-1 and Fli-1 protein induce PLD2 but not PLD1 gene expression by binding to an ETS domain of 5' promoter. *Oncogene* **26**, 1802-1810 (2007).
 77. Smith, R., *et al.* Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell* **9**, 405-416 (2006).
 78. Ramakrishnan, R., *et al.* Role of protein-protein interactions in the antiapoptotic function of EWS-Fli-1. *Oncogene* **23**, 7087-7094 (2004).
 79. Fuchs, B., Inwards, C.Y. & Janknecht, R. Vascular endothelial growth factor expression is up-regulated by EWS-ETS oncoproteins and Sp1 and may represent an

-
- independent predictor of survival in Ewing's sarcoma. *Clin Cancer Res* **10**, 1344-1353 (2004).
80. Zwerner, J.P., *et al.* The EWS/FLI1 oncogenic transcription factor deregulates GLI1. *Oncogene* **27**, 3282-3291 (2008).
81. Kauer, M., *et al.* A molecular function map of Ewing's sarcoma. *PLoS ONE* **4**, e5415 (2009).
82. Eliazer, S., Spencer, J., Ye, D., Olson, E. & Ilaria, R.L., Jr. Alteration of mesodermal cell differentiation by EWS/FLI-1, the oncogene implicated in Ewing's sarcoma. *Mol Cell Biol* **23**, 482-492 (2003).
83. Lessnick, S.L., Dacwag, C.S. & Golub, T.R. The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer Cell* **1**, 393-401 (2002).
84. Riggi, N., *et al.* Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res* **65**, 11459-11468 (2005).
85. Riggi, N., *et al.* EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res* **68**, 2176-2185 (2008).
86. Herrero-Martin, D., *et al.* Stable interference of EWS-FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target. *Br J Cancer* **101**, 80-90 (2009).
87. Deneen, B. & Denny, C.T. Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation. *Oncogene* **20**, 6731-6741 (2001).
88. Castellero-Trejo, Y., Eliazer, S., Xiang, L., Richardson, J.A. & Ilaria, R.L., Jr. Expression of the EWS/FLI-1 oncogene in murine primary bone-derived cells Results in EWS/FLI-1-dependent, ewing sarcoma-like tumors. *Cancer Res* **65**, 8698-8705 (2005).
89. Torchia, E.C., Jaishankar, S. & Baker, S.J. Ewing tumor fusion proteins block the differentiation of pluripotent marrow stromal cells. *Cancer Res* **63**, 3464-3468 (2003).
-

-
90. Hong, H., *et al.* Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* (2009).
 91. Kawamura, T., *et al.* Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* (2009).
 92. Li, H., *et al.* The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* (2009).
 93. Marion, R.M., *et al.* A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* (2009).
 94. Tsuchiya, T., *et al.* Analysis of the p16INK4, p14ARF, p15, TP53, and MDM2 genes and their prognostic implications in osteosarcoma and Ewing sarcoma. *Cancer Genet Cytogenet* **120**, 91-98 (2000).
 95. Wei, G., *et al.* Prognostic impact of INK4A deletion in Ewing sarcoma. *Cancer* **89**, 793-799 (2000).
 96. Lopez-Guerrero, J.A., Pellin, A., Noguera, R., Carda, C. & Llombart-Bosch, A. Molecular analysis of the 9p21 locus and p53 genes in Ewing family tumors. *Lab Invest* **81**, 803-814 (2001).
 97. Savola, S., Nardi, F., Scotlandi, K., Picci, P. & Knuutila, S. Microdeletions in 9p21.3 induce false negative results in CDKN2A FISH analysis of Ewing sarcoma. *Cytogenet Genome Res* **119**, 21-26 (2007).
 98. Kallioniemi, A., *et al.* Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* **258**, 818-821 (1992).
 99. Lichter, P., Joos, S., Bentz, M. & Lampel, S. Comparative genomic hybridization: uses and limitations. *Semin Hematol* **37**, 348-357 (2000).
 100. Pinkel, D., *et al.* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**, 207-211 (1998).
-

-
101. Knuutila, S., *et al.* Comparative genomic hybridization study on pooled DNAs from tumors of one clinical-pathological entity. *Cancer Genet Cytogenet* **100**, 25-30 (1998).
 102. Kullendorff, C.M., *et al.* Cytogenetic aberrations in Ewing sarcoma: are secondary changes associated with clinical outcome? *Med Pediatr Oncol* **32**, 79-83 (1999).
 103. Tarkkanen, M., *et al.* Clinical correlations of genetic changes by comparative genomic hybridization in Ewing sarcoma and related tumors. *Cancer Genet Cytogenet* **114**, 35-41 (1999).
 104. Brisset, S., *et al.* CGH analysis of secondary genetic changes in Ewing tumors: correlation with metastatic disease in a series of 43 cases. *Cancer Genet Cytogenet* **130**, 57-61 (2001).
 105. Ozaki, T., *et al.* Genetic imbalances revealed by comparative genomic hybridization in Ewing tumors. *Genes Chromosomes Cancer* **32**, 164-171 (2001).
 106. Hattinger, C.M., *et al.* Prognostic impact of chromosomal aberrations in Ewing tumours. *Br J Cancer* **86**, 1763-1769 (2002).
 107. Ferreira, B.I., *et al.* Array CGH and gene-expression profiling reveals distinct genomic instability patterns associated with DNA repair and cell-cycle checkpoint pathways in Ewing's sarcoma. *Oncogene* **27**, 2084-2090 (2008).
 108. Savola, S., *et al.* Combined use of expression and CGH arrays pinpoints novel candidate genes in Ewing sarcoma family of tumors. *BMC Cancer* **9**, 17 (2009).
 109. Friedrichs, N., *et al.* Pitfalls in the detection of t(11;22) translocation by fluorescence in situ hybridization and RT-PCR: a single-blinded study. *Diagn Mol Pathol* **15**, 83-89 (2006).
 110. Scotlandi, K., *et al.* Overcoming resistance to conventional drugs in Ewing sarcoma and identification of molecular predictors of outcome. *J Clin Oncol* **27**, 2209-2216 (2009).

-
111. Ottaviano, L., *et al.* Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort. *Genes Chromosomes Cancer* **49**, 40-51.
 112. Larsen, J., Ottesen, A.M., Lundsteen, C., Leffers, H. & Larsen, J.K. Optimization of DOP-PCR amplification of DNA for high-resolution comparative genomic hybridization analysis. *Cytometry* **44**, 317-325 (2001).
 113. Carter, N.P., Fiegler, H. & Piper, J. Comparative analysis of comparative genomic hybridization microarray technologies: report of a workshop sponsored by the Wellcome Trust. *Cytometry* **49**, 43-48 (2002).
 114. Dessau, R.B. & Pipper, C.B. ["R"--project for statistical computing]. *Ugeskr Laeger* **170**, 328-330 (2008).
 115. Smith, M.L., Marioni, J.C., Hardcastle, T.J., Thorne, N.P. snapCGH: Segmentation, Normalization and Processing of aCGH Data Users' Guide. *Bioconductor* (2006).
 116. van de Wiel, M.A., *et al.* CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics* **23**, 892-894 (2007).
 117. Gentleman, R.C., *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).
 118. Olshen, A.B., Venkatraman, E.S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5**, 557-572 (2004).
 119. Willenbrock, H. & Fridlyand, J. A comparison study: applying segmentation to array CGH data for downstream analyses. *Bioinformatics* **21**, 4084-4091 (2005).
 120. Mark A. van de Wiel, a.W.N., van Wieringen. CGHregions: Dimension Reduction for Array CGH Data with Minimal Information Loss. *Cancer informatics* **3**, 9 (2007).
 121. Van Wieringen, W.N., Van De Wiel, M.A. & Ylstra, B. Weighted clustering of called array CGH data. *Biostatistics* **9**, 484-500 (2008).
-

-
122. van de Wiel, M.A., Smeets, S.J., Brakenhoff, R.H. & Ylstra, B. CGHMultiArray: exact P-values for multi-array comparative genomic hybridization data. *Bioinformatics* **21**, 3193-3194 (2005).
 123. Thorisson, G.A., Smith, A.V., Krishnan, L. & Stein, L.D. The International HapMap Project Web site. *Genome Res* **15**, 1592-1593 (2005).
 124. Szuhai, K. & Tanke, H.J. COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nat Protoc* **1**, 264-275 (2006).
 125. Irizarry, R.A., *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* **31**, e15 (2003).
 126. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121 (2001).
 127. Irizarry, R.A., *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264 (2003).
 128. Saeed, A.I., *et al.* TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374-378 (2003).
 129. Subramanian, A., *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
 130. Wu, C., *et al.* BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* **10**, R130 (2009).
 131. Su, A.I., *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* **101**, 6062-6067 (2004).
 132. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45 (2001).

-
133. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-386 (2000).
 134. Hattinger, C.M., *et al.* Demonstration of the translocation der(16)t(1;16)(q12;q11.2) in interphase nuclei of Ewing tumors. *Genes Chromosomes Cancer* **17**, 141-150 (1996).
 135. Higa, L.A., *et al.* CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat Cell Biol* **8**, 1277-1283 (2006).
 136. Jin, J., Arias, E.E., Chen, J., Harper, J.W. & Walter, J.C. A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol Cell* **23**, 709-721 (2006).
 137. Banks, D., *et al.* L2DTL/CDT2 and PCNA interact with p53 and regulate p53 polyubiquitination and protein stability through MDM2 and CUL4A/DDB1 complexes. *Cell Cycle* **5**, 1719-1729 (2006).
 138. Higa, L.A., *et al.* L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. *Cell Cycle* **5**, 1675-1680 (2006).
 139. Higa, L.A., *et al.* Involvement of CUL4 ubiquitin E3 ligases in regulating CDK inhibitors Dacapo/p27Kip1 and cyclin E degradation. *Cell Cycle* **5**, 71-77 (2006).
 140. Abbas, T., *et al.* PCNA-dependent regulation of p21 ubiquitylation and degradation via the CRL4Cdt2 ubiquitin ligase complex. *Genes Dev* **22**, 2496-2506 (2008).
 141. Kim, Y., Starostina, N.G. & Kipreos, E.T. The CRL4Cdt2 ubiquitin ligase targets the degradation of p21Cip1 to control replication licensing. *Genes Dev* **22**, 2507-2519 (2008).
 142. Nishitani, H., *et al.* CDK inhibitor p21 is degraded by a proliferating cell nuclear antigen-coupled Cul4-DDB1Cdt2 pathway during S phase and after UV irradiation. *J Biol Chem* **283**, 29045-29052 (2008).

-
143. Schaefer, K.L., *et al.* Microarray analysis of Ewing's sarcoma family of tumours reveals characteristic gene expression signatures associated with metastasis and resistance to chemotherapy. *Eur J Cancer* **44**, 699-709 (2008).
 144. Szuhai, K., Ijszenga, M., Tanke, H.J., Rosenberg, C. & Hogendoorn, P.C. Molecular cytogenetic characterization of four previously established and two newly established Ewing sarcoma cell lines. *Cancer Genet Cytogenet* **166**, 173-179 (2006).
 145. Uren, A.G., *et al.* Large-scale mutagenesis in p19(ARF)- and p53-deficient mice identifies cancer genes and their collaborative networks. *Cell* **133**, 727-741 (2008).
 146. Chang, N.S., Hsu, L.J., Lin, Y.S., Lai, F.J. & Sheu, H.M. WW domain-containing oxidoreductase: a candidate tumor suppressor. *Trends Mol Med* **13**, 12-22 (2007).
 147. Gaudio, E., *et al.* Physical association with WWOX suppresses c-Jun transcriptional activity. *Cancer Res* **66**, 11585-11589 (2006).
 148. Aqeilan, R.I., *et al.* Targeted deletion of Wwox reveals a tumor suppressor function. *Proc Natl Acad Sci U S A* **104**, 3949-3954 (2007).
 149. Qin, H.R., *et al.* A role for the WWOX gene in prostate cancer. *Cancer Res* **66**, 6477-6481 (2006).
 150. Fabbri, M., *et al.* WWOX gene restoration prevents lung cancer growth in vitro and in vivo. *Proc Natl Acad Sci U S A* **102**, 15611-15616 (2005).
 151. Aqeilan, R.I., *et al.* Inactivation of the Wwox gene accelerates forestomach tumor progression in vivo. *Cancer Res* **67**, 5606-5610 (2007).
 152. Gourley, C., *et al.* WWOX gene expression abolishes ovarian cancer tumorigenicity in vivo and decreases attachment to fibronectin via integrin alpha3. *Cancer Res* **69**, 4835-4842 (2009).
 153. Brock, M.V., *et al.* DNA methylation markers and early recurrence in stage I lung cancer. *N Engl J Med* **358**, 1118-1128 (2008).
-

-
154. Jin, Z., *et al.* Promoter hypermethylation of CDH13 is a common, early event in human esophageal adenocarcinogenesis and correlates with clinical risk factors. *Int J Cancer* **123**, 2331-2336 (2008).
 155. Chan, D.W., Lee, J.M., Chan, P.C. & Ng, I.O. Genetic and epigenetic inactivation of T-cadherin in human hepatocellular carcinoma cells. *Int J Cancer* **123**, 1043-1052 (2008).
 156. Ghosh, S., *et al.* Use of multicellular tumor spheroids to dissect endothelial cell-tumor cell interactions: a role for T-cadherin in tumor angiogenesis. *FEBS Lett* **581**, 4523-4528 (2007).
 157. Riou, P., *et al.* Expression of T-cadherin in tumor cells influences invasive potential of human hepatocellular carcinoma. *FASEB J* **20**, 2291-2301 (2006).
 158. Philippova, M., *et al.* RhoA and Rac mediate endothelial cell polarization and detachment induced by T-cadherin. *FASEB J* **19**, 588-590 (2005).
 159. Baudis, M. & Cleary, M.L. Progenetix.net: an online repository for molecular cytogenetic aberration data. *Bioinformatics* **17**, 1228-1229 (2001).
 160. Bagchi, A. & Mills, A.A. The quest for the 1p36 tumor suppressor. *Cancer Res* **68**, 2551-2556 (2008).
 161. Schwab, M., *et al.* Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. *Nature* **305**, 245-248 (1983).
 162. Fearon, E.R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-767 (1990).
 163. Parr, C. & Jiang, W.G. Metastasis suppressor 1 (MTSS1) demonstrates prognostic value and anti-metastatic properties in breast cancer. *Eur J Cancer* **45**, 1673-1683 (2009).
 164. Ma, L., *et al.* miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol* **12**, 247-256 (2010).
-

-
165. Lujambio, A., *et al.* A microRNA DNA methylation signature for human cancer metastasis. *Proc Natl Acad Sci U S A* **105**, 13556-13561 (2008).
 166. Pezzolo, A., *et al.* Presence of 1q gain and absence of 7p gain are new predictors of local or metastatic relapse in localized resectable neuroblastoma. *Neuro Oncol* (2008).
 167. Lo, K.C., *et al.* Gain of 1q is a potential univariate negative prognostic marker for survival in medulloblastoma. *Clin Cancer Res* **13**, 7022-7028 (2007).
 168. Balcarikova, J., *et al.* Gain of chromosome arm 1q in patients in relapse and progression of multiple myeloma. *Cancer Genet Cytogenet* **192**, 68-72 (2009).
 169. Hing, S., *et al.* Gain of 1q is associated with adverse outcome in favorable histology Wilms' tumors. *Am J Pathol* **158**, 393-398 (2001).
 170. Kjellman, P., *et al.* Gain of 1q and loss of 9q21.3-q32 are associated with a less favorable prognosis in papillary thyroid carcinoma. *Genes Chromosomes Cancer* **32**, 43-49 (2001).
 171. Xie, X., *et al.* Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338-345 (2005).
 172. Wu, Q., *et al.* Transcriptional regulation during p21WAF1/CIP1-induced apoptosis in human ovarian cancer cells. *J Biol Chem* **277**, 36329-36337 (2002).
 173. Haas, A.L. & Rose, I.A. The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis. *J Biol Chem* **257**, 10329-10337 (1982).
 174. Kee, Y. & Huibregtse, J.M. Regulation of catalytic activities of HECT ubiquitin ligases. *Biochem Biophys Res Commun* **354**, 329-333 (2007).
 175. Sun, Y. Targeting E3 ubiquitin ligases for cancer therapy. *Cancer Biol Ther* **2**, 623-629 (2003).
 176. Hatakeyama, S. & Nakayama, K.I. U-box proteins as a new family of ubiquitin ligases. *Biochem Biophys Res Commun* **302**, 635-645 (2003).
-

-
177. Zheng, N., *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703-709 (2002).
178. Soucy, T.A., Smith, P.G. & Rolfe, M. Targeting NEDD8-activated cullin-RING ligases for the treatment of cancer. *Clin Cancer Res* **15**, 3912-3916 (2009).
179. Jackson, S. & Xiong, Y. CRL4s: the CUL4-RING E3 ubiquitin ligases. *Trends Biochem Sci* **34**, 562-570 (2009).
180. Ralph, E., Boye, E. & Kearsey, S.E. DNA damage induces Cdt1 proteolysis in fission yeast through a pathway dependent on Cdt2 and Ddb1. *EMBO Rep* **7**, 1134-1139 (2006).
181. Sansam, C.L., *et al.* DTL/CDT2 is essential for both CDT1 regulation and the early G2/M checkpoint. *Genes Dev* **20**, 3117-3129 (2006).
182. Shibutani, S.T., *et al.* Intrinsic negative cell cycle regulation provided by PIP box- and Cul4Cdt2-mediated destruction of E2f1 during S phase. *Dev Cell* **15**, 890-900 (2008).
183. Senga, T., *et al.* PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. *J Biol Chem* **281**, 6246-6252 (2006).
184. Havens, C.G. & Walter, J.C. Docking of a specialized PIP Box onto chromatin-bound PCNA creates a degron for the ubiquitin ligase CRL4Cdt2. *Mol Cell* **35**, 93-104 (2009).
185. Liu, C.L., Yu, I.S., Pan, H.W., Lin, S.W. & Hsu, H.C. L2dtl is essential for cell survival and nuclear division in early mouse embryonic development. *J Biol Chem* **282**, 1109-1118 (2007).
186. Pan, H.W., *et al.* Role of L2DTL, cell cycle-regulated nuclear and centrosome protein, in aggressive hepatocellular carcinoma. *Cell Cycle* **5**, 2676-2687 (2006).
187. Ueki, T., *et al.* Involvement of elevated expression of multiple cell-cycle regulator, DTL/RAMP (denticleless/RA-regulated nuclear matrix associated protein), in the growth of breast cancer cells. *Oncogene* **27**, 5672-5683 (2008).
-

-
188. Missiaglia, E., *et al.* Genomic imbalances in rhabdomyosarcoma cell lines affect expression of genes frequently altered in primary tumors: an approach to identify candidate genes involved in tumor development. *Genes Chromosomes Cancer* **48**, 455-467 (2009).
 189. Gong, L. & Yeh, E.T. Identification of the activating and conjugating enzymes of the NEDD8 conjugation pathway. *J Biol Chem* **274**, 12036-12042 (1999).
 190. Kim, A.Y., *et al.* SCCRO (DCUN1D1) is an essential component of the E3 complex for neddylation. *J Biol Chem* **283**, 33211-33220 (2008).
 191. Podust, V.N., *et al.* A Nedd8 conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. *Proc Natl Acad Sci U S A* **97**, 4579-4584 (2000).
 192. Read, M.A., *et al.* Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. *Mol Cell Biol* **20**, 2326-2333 (2000).
 193. Kawakami, T., *et al.* NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J* **20**, 4003-4012 (2001).
 194. Duda, D.M., *et al.* Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134**, 995-1006 (2008).
 195. Wei, N., Serino, G. & Deng, X.W. The COP9 signalosome: more than a protease. *Trends Biochem Sci* **33**, 592-600 (2008).
 196. Zheng, J., *et al.* CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. *Mol Cell* **10**, 1519-1526 (2002).
 197. Lo, S.C. & Hannink, M. CAND1-mediated substrate adaptor recycling is required for efficient repression of Nrf2 by Keap1. *Mol Cell Biol* **26**, 1235-1244 (2006).
 198. Soucy, T.A., *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732-736 (2009).
 199. Kim, Y. & Kipreos, E.T. Cdt1 degradation to prevent DNA re-replication: conserved and non-conserved pathways. *Cell Div* **2**, 18 (2007).
-

200. Melixetian, M., *et al.* Loss of Geminin induces rereplication in the presence of functional p53. *J Cell Biol* **165**, 473-482 (2004).
201. Vaziri, C., *et al.* A p53-dependent checkpoint pathway prevents rereplication. *Mol Cell* **11**, 997-1008 (2003).
202. Maitra, A., Roberts, H., Weinberg, A.G. & Geradts, J. Aberrant expression of tumor suppressor proteins in the Ewing family of tumors. *Arch Pathol Lab Med* **125**, 1207-1212 (2001).
203. Wang, H., *et al.* Skp2 is required for survival of aberrantly proliferating Rb1-deficient cells and for tumorigenesis in Rb1^{+/-} mice. *Nat Genet* **42**, 83-88 (2010).
204. Corson, T.W., Huang, A., Tsao, M.S. & Gallie, B.L. KIF14 is a candidate oncogene in the 1q minimal region of genomic gain in multiple cancers. *Oncogene* **24**, 4741-4753 (2005).

C
A
R
L
O
S

M
A
C

K

I

N

T
O
S
H

T
E
S
I
S

IDENTIFICACIÓN Y CARACTERIZACIÓN DE SUBTIPOS DE ENFERMEDAD EN SARCOMA DE EWING

MEDIANTE TÉCNICAS DE CRIBADO MASIVO ENFOCADAS
AL DESCUBRIMIENTO DE BIOMARCADORES Y DIANAS TERAPÉUTICAS

D
O
C
T
O
R
A
L

SALAMANCA
2010



CSIC



VNIVERSIDAD
D SALAMANCA