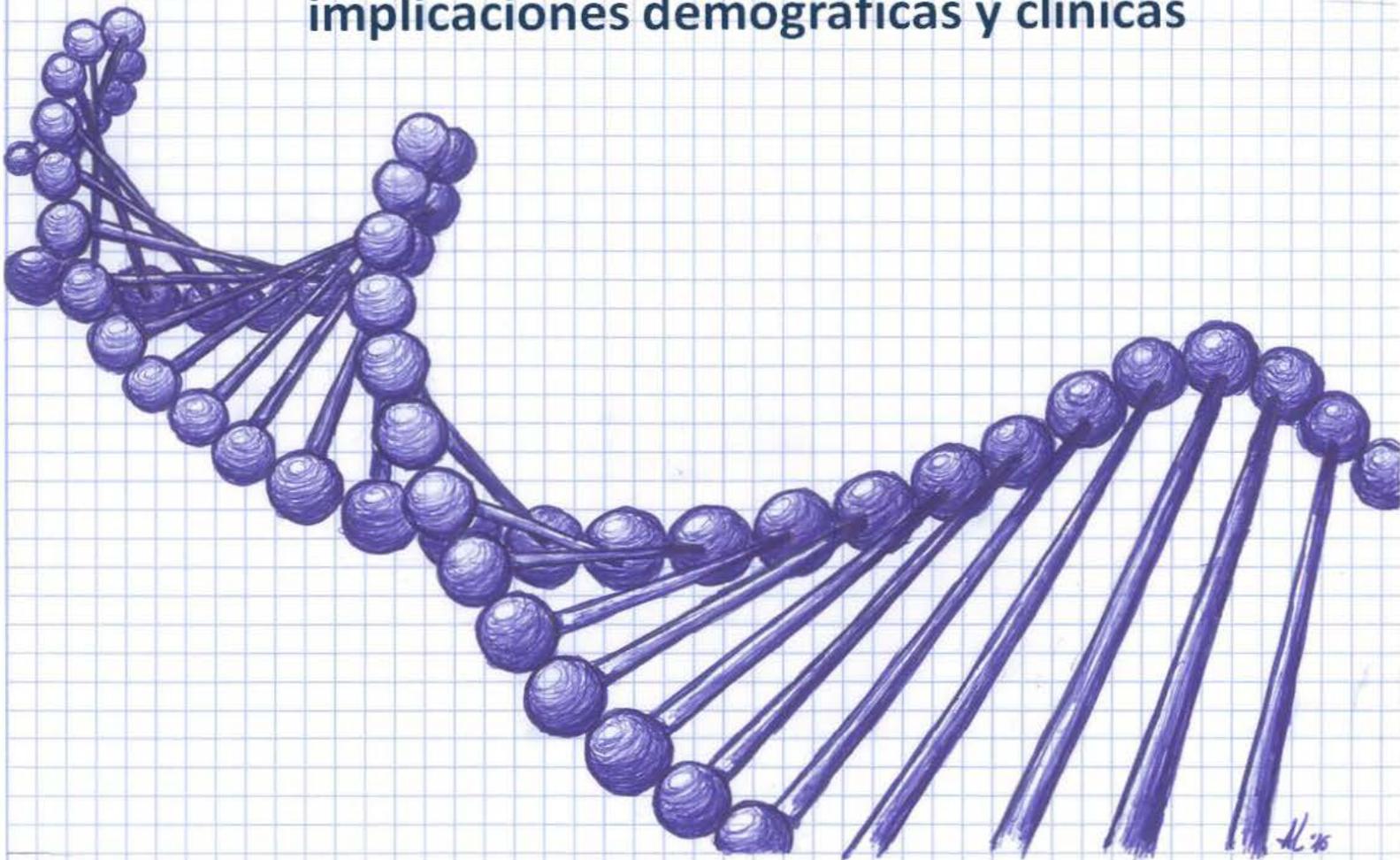




UNIVERSIDAD DE LAS PALMAS
DE GRAN CANARIA

Escuela de Doctorado de la ULPGC
Departamento de Ciencias Clínicas

Estudio del papel de polimorfismos en genes de reparación del ADN y del metabolismo de andrógenos en pacientes con cáncer de próstata: implicaciones demográficas y clínicas



TESIS DOCTORAL

Programa de Doctorado de Investigación Aplicada a las Ciencias Sanitarias

Almudena Valenciano Folgado

Las Palmas de Gran Canaria
2017

A mi madre, mi estrella polar, y a mis hermanos, Mufasa y Scar

A Jesús, no conozco amor más desinteresado

A todas las personas que luchan contra este gigante

Agradecimientos

Agradecimientos

Este trabajo es el resultado del esfuerzo de muchas personas y de la influencia de otras tantas. Hay muchas cosas que no soy, pero sí soy agradecida, lo cual agradezco enormemente, porque eso me hace ser feliz cada día.

Es por ello que debo agradecer en primer lugar la gran oportunidad que me brindó mi director de tesis y jefe del Servicio de Oncología Radioterápica del Hospital Universitario de Gran Canaria Dr. Negrín, el Dr. Pedro Carlos Lara Jiménez, al ofrecerme un lugar en su equipo de trabajo. Gracias por tus conocimientos, por tu dedicación y por confiar en mi capacidad.

Como parte imprescindible de este proyecto, quiero agradecer a mi compañero y director de tesis, el Dr. Luis Alberto Henríquez Hernández. Admiro tu capacidad de trabajo, tu perseverancia, tu organización, aún es un misterio para mí cómo eres capaz de hacer tantas cosas a la vez y todas tan bien. Gracias por ser mi guía durante los años de trabajo, por tu motivación constante, por tu confianza, por mostrarme la luz en los momentos de oscuridad. Sin tu esfuerzo, nada de esto sería posible.

Al igual que hubiese sido imposible sin la colaboración desinteresada de todos los pacientes que forman parte de este estudio. Gracias a todos los pacientes que durante años y de forma altruista han contribuido a que otros muchos diagnosticados cada día no oigan la palabra “muerte” cuando su médico les dice “cáncer”.

Gracias a todo el personal del Servicio de Oncología Radioterápica del Hospital Universitario de Gran Canaria Dr. Negrín, porque son ellos los que luchan por los pacientes. A todo el personal médico, Bernardino, Marta, Bea, Gustavo, Raquel, Ruth, Nacho, Suli, Irene, Nieves, Íñigo y Mario. A los residentes Alba, Jesús y Laura, y a todos los residentes en rotación que han compartido algún mes conmigo en el laboratorio y me han aguantado “el rollo”. Agradezco especialmente la paciencia del personal de enfermería. Gracias Gerardo, Eugenio, Minga, Ana Clara, Bea, Lola, Laura, Lali, Pilar y Loli por no darme una patada cada vez que aparecía por la puerta con mis tubitos y papeles. Gracias a Vanessa y a todos los técnicos, auxiliares, administrativos y físicos.

Quiero agradecer especialmente, por haber compartido conmigo las horas de trabajo, las risas, los queques y los piscolabis, a todas las niñas (y algún niño también) del Servicio de Inmunología del Hospital Universitario de Gran Canaria Dr. Negrín. Gracias al Dr. Carlos Rodríguez Gallego por sus consejos y por hacerme un huequito en el laboratorio. Gracias a todo el personal, Mari Carmen, Rita, Maica, Belén, Isa, Ana, Carmen, Juan Carlos, Jota, Itahisa, Maite, Florentino, Toni y Vero; vosotros me habéis adoptado como canariona y me habéis hecho sentir como en casa. No me olvido de vosotras, Fany, Marta, Nereida y Yanira; muchísimas gracias por vuestro soporte técnico e intelectual, sois unas profesionales excepcionales.

En mi vida personal, son muchas las personas que han influido sobre mí y las que de algún modo me han hecho ser quien soy y llegar tan lejos. Sin duda la más importante es y será siempre mi madre, Rosa. A ti te debo esta vida que me has regalado, en todos los sentidos y no solo en el literal. Eres la persona más increíble que conozco y gracias a tu fortaleza cambiaste el destino de tus hijos. Gracias mamá, por apoyarme siempre y respetarme en todas las decisiones que he tomado. Gracias a mis hermanos, Manuel y Jesús, por ser tan locos como yo y por perseguir la felicidad incansablemente. Gracias a Jesús, mi padre “en funciones”, por quererme y por admirarme tanto. A mis abuelos, que cuidaron de mí de forma tan cariñosa. A mi padre, Manuel, por mostrarme el camino correcto. Quiero agradecer a toda mi familia, tíos, primos, abuelos, etc., porque en mis mejores recuerdos siempre aparecéis vosotros.

A mi querido amigo Gerardo le debo especial gratitud por apoyarme durante los años de este trabajo. Por ser mi pañuelo de lágrimas y desahogo. Por hacerme la vida más fácil. Y a toda su familia por convertirse en mi familia en esta maravillosa isla. A todas las personas que han aparecido por mi vida y que, aun estando lejos hoy, siguen formando parte de ella, quiero agradecerles que me hayan regalado parte de su tiempo y de su esencia, que me acompañan cada día. A mis *biolocos*, Arancha, Quique, Isa, Rocío, Gloria y Virginia, por tantas historias y horas de fotosíntesis. A mis niñas de Mariehem, Fátima y Raquel, con vosotras todo es maravillosamente surrealista. A Carla, mi vecina de isla. A *El clan del pececito*, mi familia canaria (casi todos peninsulares), Igor, Natalia, Laia, Sabrina, Juanma, Mireia y Fede, que me acompañaron en los primeros años de mi estancia en Gran Canaria, y a todos los demás que se incorporaron después, gracias por todos los momentos, incontables. ¡Viva el zumo de papaya! Por último, quiero dar las gracias a las personas que han convivido conmigo durante este último año

y que ahora forman parte del nuevo proyecto de mi vida. Gracias a la sección 12 y a su tutor. Y gracias a Iván, mi codirector extraoficial, por sus críticas “tan constructivas” y por fijarse en cada mínimo detalle; has contribuido a esto más de lo que crees.

Para terminar, sé que he nombrado a mucha gente y con ello he corrido el riesgo de olvidar a alguien. Si es así no me lo tengas en cuenta, pues ha sido solo un despiste y en tu caso sabrás perfectamente que deberías estar incluido.

Índice

ÍNDICE

1.- Introducción	3
1.1.- Epidemiología	3
1.2.- Diagnóstico clínico del cáncer de próstata y su tratamiento	5
1.3.- Biología molecular del cáncer de próstata	8
1.3.1.- Maquinaria de reparación de ADN	8
1.3.1.1.- Mecanismos de reparación de daño al ADN	9
1.3.1.2.- Genes implicados en la reparación del ADN	16
1.3.1.3.- SNPs en genes de reparación del ADN y CaP	19
1.3.2.- Metabolismo de andrógenos	22
1.3.2.1.- SNPs en genes del metabolismo de andrógenos y CaP	23
1.3.3.- Polimorfismos de nucleótido simple (SNPs)	25
1.3.3.1.- SNPs y CaP	28
1.3.4.- Biomarcadores clínicos en cáncer de próstata.	29
1.3.4.1.- Antígeno Prostático Específico (PSA) y sus derivados	31
1.3.4.2.- Otros biomarcadores proteicos	33
1.3.4.3.- Plataformas de diagnóstico	37
2.- Pacientes y métodos	43
2.1.- Pacientes	43
2.2.- Aislamiento y cuantificación de ADN	43
2.3.- Criterios de selección de los SNPs	44
2.4.- Genes y SNPs	44
2.5.- Genotipado	49
2.6.- Análisis estadísticos	50
3.- Hipótesis	55
4.- Objetivos	59
5.- Capítulo I: La heterogeneidad en la distribución de polimorfismos entre poblaciones del mismo origen étnico como factor de confusión en estudios de asociación genética	63
5.1.- Introducción	63
5.2.- Resultados	65
5.3.- Discusión	70
5.4.- Conclusiones	74
6.- Capítulo II: Variaciones intra-étnicas en el gen 5-alfa reductasa como factor implicado en la variabilidad de respuesta clínica a los inhibidores de esta enzima	77

6.1.- Introducción.....	77
6.2.- Resultados	79
6.3.- Discusión.....	84
6.4.- Conclusión	87
7.- Capítulo III: Polimorfismos de nucleótido simple en genes de reparación del ADN como factores de riesgo asociados a la progresión del cáncer de próstata.	91
7.1.- Introducción.....	91
7.2.- Resultados	93
7.3.- Discusión.....	98
7.4.- Conclusión	101
8.- Capítulo IV: Asociación entre polimorfismos de nucleótido simple en genes de reparación de rotura doble de cadena del ADN y agresividad del cáncer de próstata en la población española.	105
8.1.- Introducción.....	105
8.2.- Resultados	108
8.2.1.- Variación intra-étnica de SNPs.....	108
8.2.2.- Variables clínicas asociadas a los SNPs individualmente	108
8.2.3.- Análisis combinados de múltiples SNPs.....	113
8.3.- Discusión.....	114
8.4.- Conclusiones.....	119
9.- Capítulo V: La presencia de determinadas variaciones genéticas en genes implicados en el metabolismo de la testosterona están asociadas con la progresión del cáncer de próstata: un estudio multicéntrico español.	123
9.1.- Introducción.....	123
9.2.- Resultados	124
9.3.- Discusión.....	128
9.4.- Conclusión	130
10.- Conclusiones.....	133
11.- Referencias bibliográficas.....	137
12.- Anexo: artículos generados a partir de la presente Tesis Doctoral.	153

Introducción

1.- Introducción

1.1.- Epidemiología

El cáncer de próstata (CaP) es el cuarto tipo tumoral más frecuente en el mundo. En varones, es el segundo cáncer con mayor incidencia y la quinta causa de muerte por cáncer. Según los datos obtenidos por la serie de registros GLOBOCAN, en 2012 se diagnosticaron 1 112 000 nuevos casos de cáncer de próstata en todo el mundo (un 15% del total de tumores diagnosticados en varones), mientras que el número de muertes producidas por este tipo tumoral fue de 307 000 (el 6.6% de las muertes producidas por cáncer en varones) (Ferlay et al. 2015). Sin embargo, esta distribución global presenta diferencias notables entre las regiones consideradas desarrolladas y las regiones en vías de desarrollo. El cáncer en general es una enfermedad asociada a las regiones desarrolladas, donde la tasa de incidencia es dos veces superior a la que encontramos en las regiones en vías de desarrollo (Jemal et al. 2011). Esta relación es especialmente destacable en el CaP, ya que el 70% de los nuevos casos diagnosticados se localizan en países desarrollados (759 000 casos de los 1 112 000 diagnosticados en todo el mundo), pasando a ser el tumor más frecuente en varones en estas regiones (Ferlay et al. 2015). Una de las principales causas de esta diferencia es el empleo del método de diagnóstico basado en la determinación de los niveles de antígeno prostático específico (PSA) en suero, que identifica no solo los tumores clínicamente relevantes, sino también aquellos de crecimiento lento que de otro modo escaparían al diagnóstico. Es por ello que las regiones que presentan una tasa de incidencia más elevada, como Australia, Nueva Zelanda, Norteamérica y el norte de Europa, son aquellas en las que esta técnica de cribado y de diagnóstico está más extendida, coincidiendo además el repunte de la incidencia con la incorporación en los años 90 de esta prueba (Jemal et al. 2011). El envejecimiento de la población y el incremento de la esperanza de vida en las regiones occidentales justifican igualmente estas altas tasas de incidencia, ya que el CaP está estrechamente asociado a la edad.

En Europa, el CaP es el tercero más frecuente, con 416 700 nuevos casos diagnosticados en el año 2012 (lo que supone un 12.1% del total de casos) y una tasa de incidencia ajustada por edad de 96 casos por 100 000 habitantes (Ferlay et al. 2013). En España, los datos son similares a los de los países de su entorno, con una tasa de

incidencia ajustada por edad en el año 2012 de 96.8 casos por 100 000 habitantes, lo que en cifras equivale a 27 850 nuevos casos diagnosticados durante ese año (Ferlay et al. 2013). Según el Centro Nacional de Epidemiología, en 2012 se produjeron en España 6038 muertes por CaP. Las figuras 1a y 1b muestran los mapas de distribución geográfica por Comunidades Autónomas de las muertes y la tasa de mortalidad, respectivamente. En las regiones en las que se realizó este estudio (Andalucía, Canarias, Cataluña y País Vasco) se observan diferencias relevantes. Andalucía es la región con mayor número de muertes por CaP (con 912 muertes, también es la Comunidad Autónoma con más muertes en el conjunto de España), seguida de Cataluña (777), País Vasco (363) y Canarias (237) (Figura 2a). Si por el contrario tenemos en cuenta la tasa de mortalidad bruta (casos por 100 000 habitantes), se observa que el País Vasco presenta una tasa de mortalidad muy superior al resto de las Comunidades Autónomas (35.53, frente a las tasas del 22.23, 22.45 y 21.82 de Andalucía, Canarias y Cataluña, respectivamente) (Figura 2b), así como a la tasa de mortalidad bruta media de España (26.64). Esta diferencia entre las regiones del norte y del sur se observa en todo el territorio español, tal y como muestra la figura 1b.

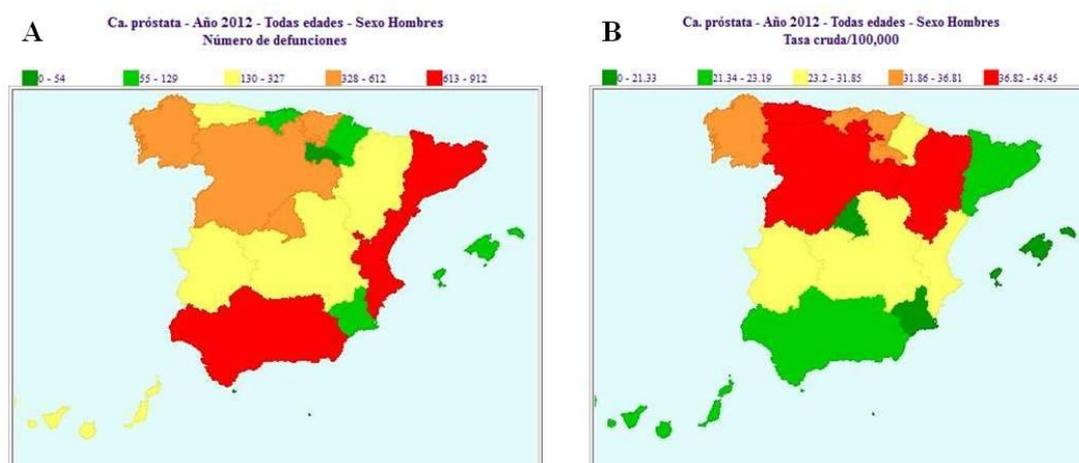
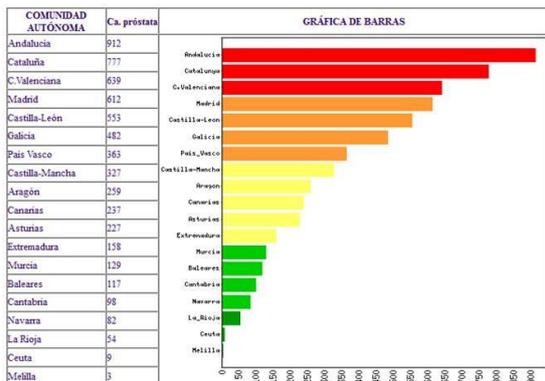


Figura 1, mapas de distribución geográfica por Comunidades Autónomas del número de defunciones (A) y de la tasa de mortalidad (B) como consecuencia del CaP.

A Ca. próstata – 2012 – número de defunciones.



B Ca. próstata – 2012 – Tasa cruda/100,000

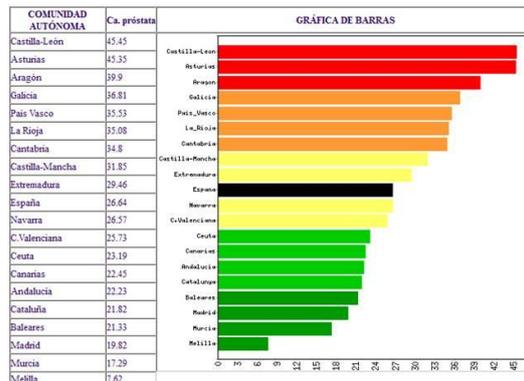


Figura 2, representación gráfica por Comunidades Autónomas del número de defunciones (A) y tasa de mortalidad (B) como consecuencia del CaP.

1.2.- Diagnóstico clínico del cáncer de próstata y su tratamiento

Las altas tasas de incidencia del CaP registradas en las regiones desarrolladas se deben en gran medida al empleo del test de diagnóstico consistente en la determinación de PSA en suero, el cual se realiza habitualmente para la detección precoz de dichos tumores en aquellos pacientes que se encuentran en edad de riesgo. Es el principal método de diagnóstico para el CaP, junto con el tacto rectal (TR). El objetivo de este test, sin embargo, no es incrementar el número de casos detectados, sino reducir la mortalidad de los pacientes diagnosticados de CaP. La detección precoz ha contribuido al incremento del diagnóstico de tumores en estadios iniciales, localizados, lo que probablemente ha favorecido la reducción de la mortalidad por CaP recogida en algunos países occidentales como Estados Unidos, Canadá o Suecia, donde la determinación de PSA en suero está ampliamente implantada (Helgesen et al. 1996; Farkas et al. 1998; Mettlin and Murphy 1998; Hankey et al. 1999; Meyer et al. 1999). Sin embargo, no existen evidencias concluyentes sobre esta posible relación directa (Ilic et al. 2007). No obstante, la detección tumoral en estadios localizados permite abordar el tratamiento de estos pacientes con estrategias dirigidas a la curación de la enfermedad. Además de la conducta de vigilancia activa, el tratamiento del CaP se completa con técnicas de radioterapia (radioterapia externa y/o braquiterapia), cirugía (prostatectomía radical) y tratamientos sistémicos (quimioterapia, deprivación androgénica de primera y segunda línea y nuevos agentes terapéuticos). El esquema de tratamiento es específico para cada

caso en particular y depende de varios factores tales como la edad del paciente, sus condiciones de morbilidad previas al tratamiento, la esperanza de vida, las toxicidades generadas por los tratamientos, así como de las características clínicas y el estadio clínico tumorales.

En el CaP, las características clínico-patológicas vienen determinadas por el estadio clínico tumoral y el *Gleason score* (GS). El estadio clínico representa la extensión tumoral; la forma en que se expresa más habitualmente es mediante el sistema de clasificación TNM (*tumor-node-metastasis*), propuesto por el *American Joint Committee on Cancer* (AJCC) y la *International Union against Cancer* (UICC) con el objetivo de desarrollar un sistema de estadiaje universal y aplicable a cualquier localización tumoral (Sobin et al. 2009; Edge et al. 2010). Este sistema viene expresado en valores alfanuméricos que representan la extensión anatómica de la enfermedad, tanto en el tumor primario (T), como en los ganglios linfáticos regionales (N) y la metástasis a distancia (M). El GS, por su parte, hace referencia al grado histológico del CaP y su procedimiento de determinación atiende a la especial distribución anatómica que presentan los tumores en la próstata (pequeños nódulos cancerígenos distribuidos por toda la glándula prostática) (Gleason and Mellinger 1974). Estas características clínico-patológicas son relevantes a la hora de determinar el riesgo de cada paciente, lo que sirve de guía para la selección del tratamiento más conveniente (Tabla 1). Los sistemas clásicos de clasificación del riesgo emplean los niveles de PSA en suero, el estadio T clínico y los valores de GS para dividir a los pacientes de CaP en tres grupos de riesgo -bajo, intermedio y alto- según la probabilidad de recaída que presentan tras recibir diversos tratamientos locales (D'Amico et al. 1998; Blasko et al. 2000; Zelefsky et al. 2000; Chism et al. 2004). Estos sistemas de clasificación han sido la base para establecer otros sistemas más actuales, como el propuesto por la *National Comprehensive Cancer Network* (NCCN 2013) o el sistema de la Universidad de San Francisco-California, conocido como *CAPRA score* (Cooperberg et al. 2005). La combinación de estos factores clínicos también ha sido utilizada por otros autores para desarrollar modelos matemáticos predictivos (Roach 1993; Roach et al. 1994), tablas y nomogramas (Partin et al. 1993; Partin et al. 1997; Partin et al. 2001; Makarov et al. 2007; Zelefsky et al. 2007) a partir de extensas series de pacientes, con varios fines: i) seleccionar aquellas personas sobre las que se debe realizar la prueba diagnóstica, ii) interpretar de forma más exacta los nivel de PSA medidos en los pacientes, iii)

determinar la necesidad de realizar una biopsia y iv) seleccionar de forma específica el tratamiento más adecuado para cada paciente (Helfand et al. 2015).

Tabla 1. Grupos de pronóstico según el sistema de clasificación propuesto por AJCC y UICC

Grupo	TNM			PSA	GS
Grupo I	T1a-c	N0	M0	< 10	≤ 6
	T2a	N0	M0	< 10	≤ 6
Grupo IIA	T1a-c	N0	M0	< 20	7
	T1a-c	N0	M0	≥ 10 y < 20	≤ 6
Grupo IIB	T2a, b	N0	M0	< 20	≤ 7
	T2c	N0	M0	Cualquiera	Cualquiera
	T1-2	N0	M0	≥ 20	Cualquiera
Grupo III	T1-2	N0	M0	Cualquiera	≥ 8
	T3a, b	N0	M0	Cualquiera	Cualquiera
Grupo IV	T4	N0	M0	Cualquiera	Cualquiera
	Cualquier T	N1	M0	Cualquiera	Cualquiera
	Cualquier T	Cualquier N	M1	Cualquiera	Cualquiera

Nota: cuando no existan datos bien de PSA, o bien de GS, el grupo deberá determinarse mediante la categoría T y cualquiera de los datos de PSA o GS que estén disponibles. Cuando no existan datos de ninguno de los dos valores, no es posible establecer el grupo; en ese caso utilizar la división del estadio.

Las técnicas empleadas para el tratamiento de CaP engloban los siguientes métodos: vigilancia activa, prostatectomía radical (PR), radioterapia externa (RTE), braquiterapia (BT), crioablación (CA) y terapias de deprivación androgénica (TDA) (Mottet et al. 2015). La PR es el procedimiento más extendido para el tratamiento de pacientes con CaP localizado. Sin embargo, en las últimas décadas las técnicas basadas en la radiación, tanto la RTE como la BT, han cobrado importancia como opción de primera línea de tratamiento debido a sus buenos resultados, tanto de forma exclusiva como en combinación con TDA (Bracarda et al. 2005; Cooperberg et al. 2010). Con relación a los resultados clínicos obtenidos tras el empleo de técnicas quirúrgicas o radioterápicas, hay que subrayar que no existen estudios *randomizados* concluyentes hasta la fecha. No obstante, los estudios observacionales sugieren que los resultados de supervivencia libre de enfermedad (SLE) y supervivencia global (SG) para pacientes de CaP estratificados en función del riesgo son similares tras el empleo de cualquiera de estas dos técnicas (en el caso del tratamiento radioterápico, ajustando los esquemas dosimétricos y empleando técnicas de conformación modernas) (Kupelian et al. 2004; DiBiase and Roach 2015). Como alternativa al tratamiento curativo, la tendencia en los últimos años es adoptar lo que se conoce como vigilancia activa en aquellos casos en los que es improbable el desarrollo de CaP clínica o biológicamente significativo. Esto es debido a que la determinación de PSA en suero no discrimina con eficiencia el CaP

agresivo de aquel que no lo es, de modo que muchos pacientes reciben un tratamiento que no necesitan (PR o TDA), con los efectos secundarios y la pérdida en calidad de vida que ello conlleva (Lund et al. 2014).

1.3.- Biología molecular del cáncer de próstata

1.3.1.- Maquinaria de reparación de ADN

La integridad del ADN es un requisito esencial para garantizar la correcta proliferación y funcionalidad celulares. Los daños producidos en dicha molécula juegan un papel primordial en los procesos de mutagénesis y carcinogénesis. El ADN contiene toda la información necesaria para sintetizar las proteínas que participan en los procesos celulares y las rutas metabólicas, imprescindibles para el correcto funcionamiento del organismo. La secuencia de ADN codifica dicha información y pequeñas modificaciones pueden suponer la síntesis de una proteína con deficiencias funcionales o incluso la total ausencia de producción de la misma. Son varios los mecanismos que de forma endógena o exógena pueden alterar la secuencia genética. Uno de los mecanismos endógenos que generan mutaciones es la propia replicación del ADN. Cada vez que una célula se divide, debe replicar su material genético para transmitirlo a las células descendientes. Este mecanismo de replicación es muy preciso y tiene una tasa de mutación muy pequeña, pero teniendo en cuenta la enorme longitud (medida en pares de bases) de la molécula de ADN, la probabilidad de que el mecanismo de replicación cometa errores no es nula. Por otro lado, de forma también endógena, algunas moléculas y productos metabólicos inducen lesiones en el ADN, como es el caso de las especies reactivas de oxígeno y otras sustancias reactivas. Los factores endógenos son los principales generadores de daño en el ADN; en total, se ha estimado que las lesiones producidas por causas endógenas son superiores a 20 000 por cada célula en el transcurso de un día (De Bont and van Larebeke 2004). Además de los factores endógenos, existen multitud de agentes exógenos con capacidad mutagénica, principalmente sustancias químicas, radiación ionizante y no ionizante, agentes biológicos, etc. En estudios de epidemiología se ha determinado que la alta incidencia del cáncer no puede ser atribuida únicamente a procesos endógenos y que la participación de factores exógenos es una condición necesaria en el 75 – 80% de los casos de cáncer (Doll and Peto 1981). No obstante, del mismo modo que los propios procesos celulares generan riesgo de producir lesiones en el ADN, también las células

tiene capacidad para reparar dichas lesiones mediante los diversos mecanismos de reparación del ADN, encargados de mantener la integridad del material genético.

Las lesiones de ADN pueden ser de diversa naturaleza según la causa que las provoque (Helleday et al. 2014). De estas causas, son frecuentes las pérdidas espontáneas de una base nitrogenada, ya que el enlace N-glicosídico entre la base nitrogenada y la pentosa es lábil. Se estima que este fenómeno ocurre en una proporción de 10^4 por célula y día (Lindahl 1993). También de forma endógena son habituales los procesos de desaminación (pérdida de un grupo amino de una base nitrogenada), lo que da lugar a una base nitrogenada diferente y como consecuencia, a un cambio en la secuencia de nucleótidos. Los agentes alquilantes, generalmente de origen exógeno, producen una metilación del átomo 6 de oxígeno de la guanina (O⁶-metilguanina), lo que hace que tenga mayor afinidad por la timina y no por la citosina como sería lo correcto (Eadie et al. 1984), favoreciendo así la sustitución de una base por otra. En condiciones normales, las especies reactivas de oxígeno participan en el metabolismo celular y son necesarias para algunos procesos celulares, pero producidas en exceso generan un estrés oxidativo que puede originar lesiones en el ADN. Entre otras modificaciones, las especies reactivas de oxígeno producen dobles enlaces en las bases nitrogenadas, así como la sustracción de un átomo de hidrógeno (H) de la timina y de la pentosa (Cooke et al. 2003). El resultado de estas modificaciones sigue siendo la alteración de la secuencia de nucleótidos de la molécula original de ADN, pero en este caso también se puede generar una modificación física de la molécula mediante la rotura de la cadena de ADN. Las radiaciones ionizantes son agentes muy lesivos que producen la rotura completa de la molécula, bien sea por la rotura de una de las cadenas o de forma más extrema, mediante la rotura de la doble hélice (*single strand break*, SSB y *double strand break*, DSB, respectivamente) (Santivasi and Xia 2014). Por el contrario, las radiaciones no ionizantes tienden a generar modificaciones bioquímicas, entre ellas los dímeros de pirimidinas.

1.3.1.1.- Mecanismos de reparación de daño al ADN

Para cada tipo de daño al ADN existe un mecanismo de reparación específico que en situaciones de normalidad identifica y repara dicha lesión, manteniendo la integridad del material genético (Wood 1996). Como hemos visto, son muchos los factores que modifican el ADN a lo largo de la vida de una persona en cada una de las

células de su cuerpo, de tal forma que si no existiera esta capacidad intrínseca de reparación tan eficiente, la propia esencia del ADN como transmisor de la información hereditaria se vería comprometida.

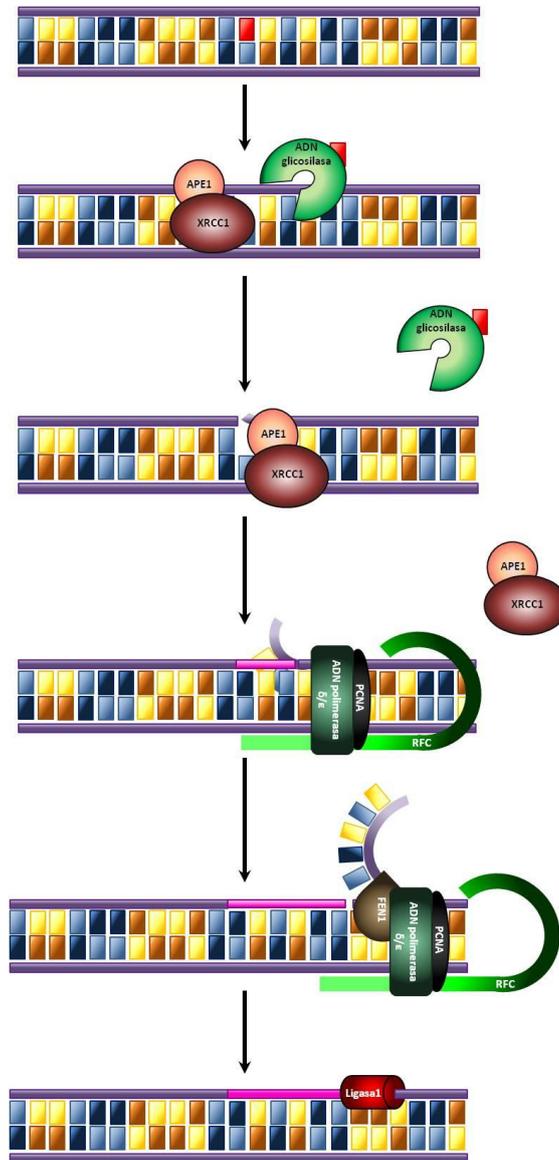


Figura 3, resumen esquemático del mecanismo de reparación del ADN por escisión de una base (BER).

Abreviaturas: APE1, *apurinic/apryrimidinic endonuclease type 1*; XRCC1, *X-ray repair cross-complementing protein 1*; RFC, *replication factor C*; PCNA, *proliferating cell nuclear antigen*; FEN1, *flap endonuclease 1*.

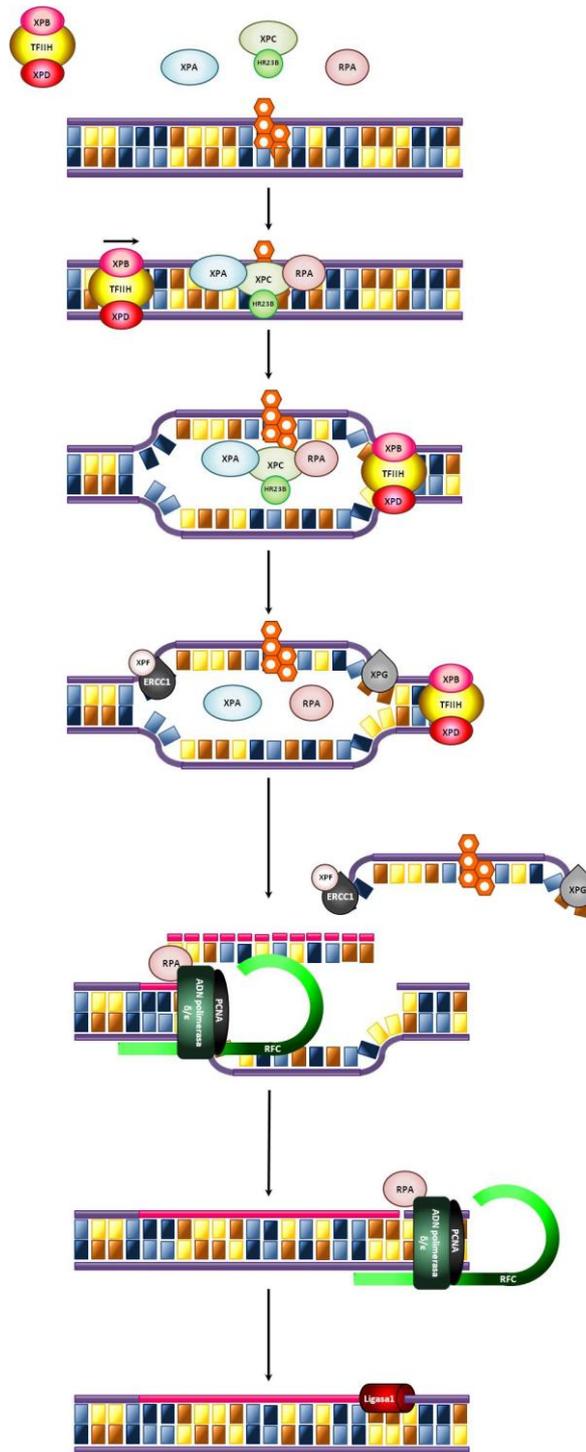


Figura 4, resumen esquemático del mecanismo de reparación del ADN por escisión de un nucleótido (NER). Abreviaturas: XPA, *xeroderma pigmentosum complementation group A*; XPC, *xeroderma pigmentosum complementation group C*; HR23B, *RAD23 homolog B*; RPA, *replication protein A*; XPB, *xeroderma pigmentosum complementation group B*; XPD, *xeroderma pigmentosum complementation group D*; TFIIH, *human transcription factor II*; XPG, *xeroderma pigmentosum complementation group G*; XPF, *xeroderma pigmentosum complementation group F*; ERCC1, *excision repair cross-complementing rodent repair deficiency, complementation group 1*; RFC, *replication factor C*; PCNA, *proliferating cell nuclear antigen*; FEN1, *flap endonuclease 1*.

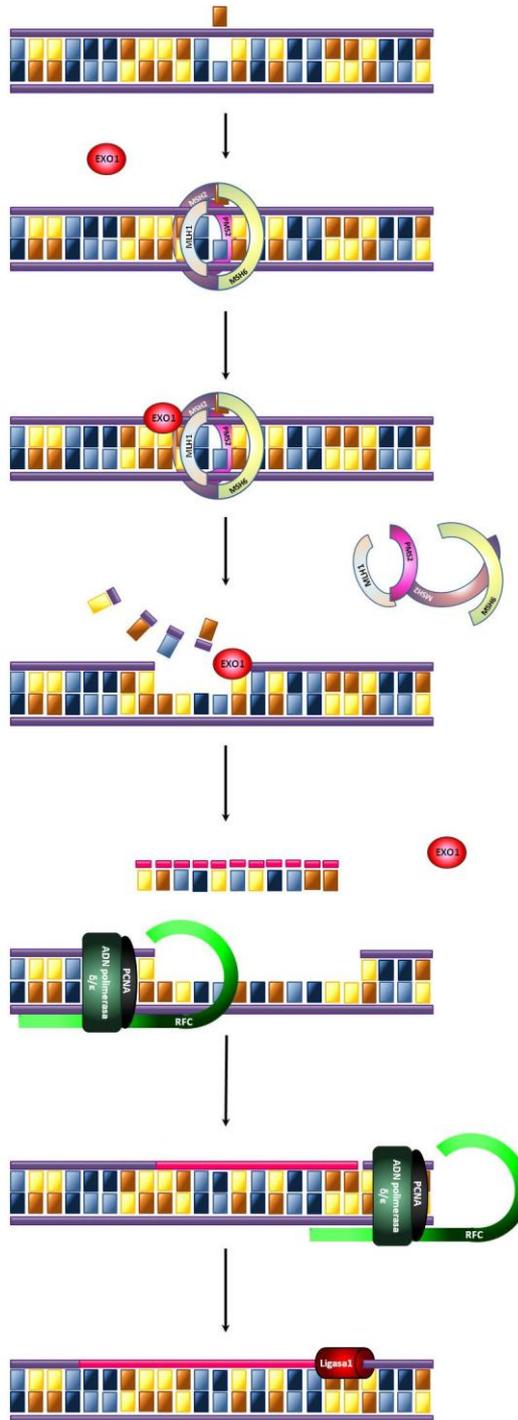


Figura 5, resumen esquemático del mecanismo de reparación del ADN por desapareamiento (*mismatch repair*). Abreviaturas: MSH2, *MutS protein homolog 2*; MSH6, *MutS protein homolog 6*; MLH1, *MutL protein homolog 1*; PMS2, *PMS1 protein homolog 1*; EXO1, *exonuclease 1*; RFC, *replication factor C*; PCNA, *proliferating cell nuclear antigen*; FEN1, *flap endonuclease 1*.

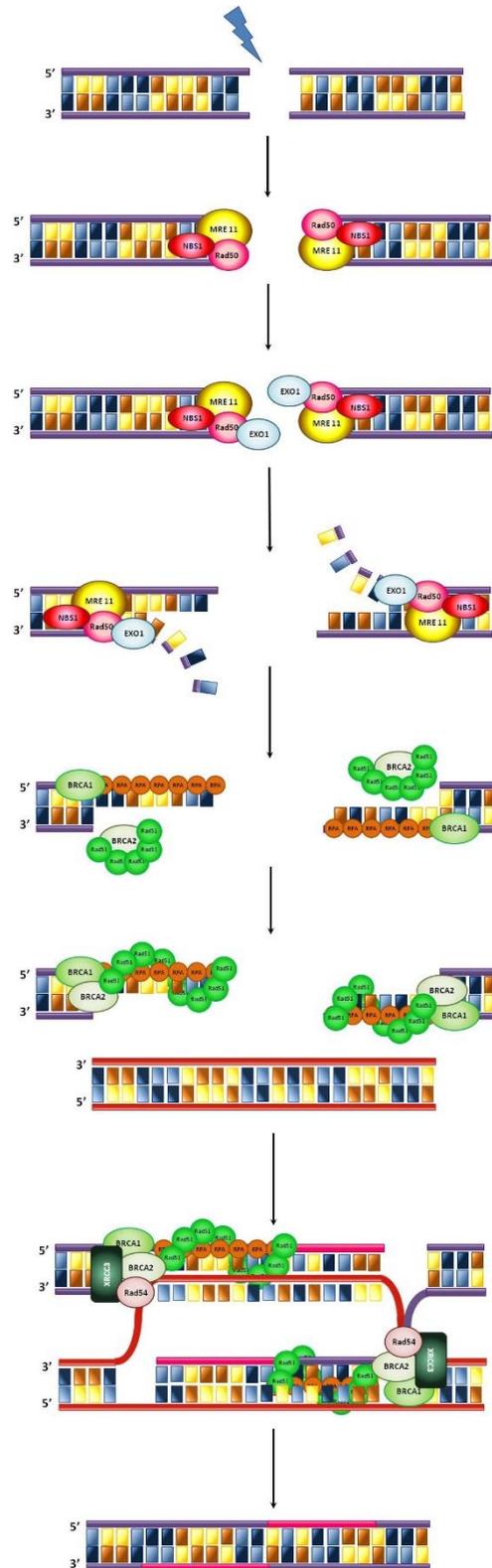


Figura 6, resumen esquemático del mecanismo de reparación por recombinación homóloga de las roturas dobles de cadena de ADN. Abreviaturas: MRE11, *meiotic recombination protein 11*; NBS1, *nijmegen breakage syndrome 1 protein*; EXO1, *exonuclease 1*; RPA, *replication protein A*; BRCA1, *breast cancer 1 protein*; BRCA2, *breast cancer 2 protein*; XRCC3, *X-ray repair cross-complementing protein 3*.

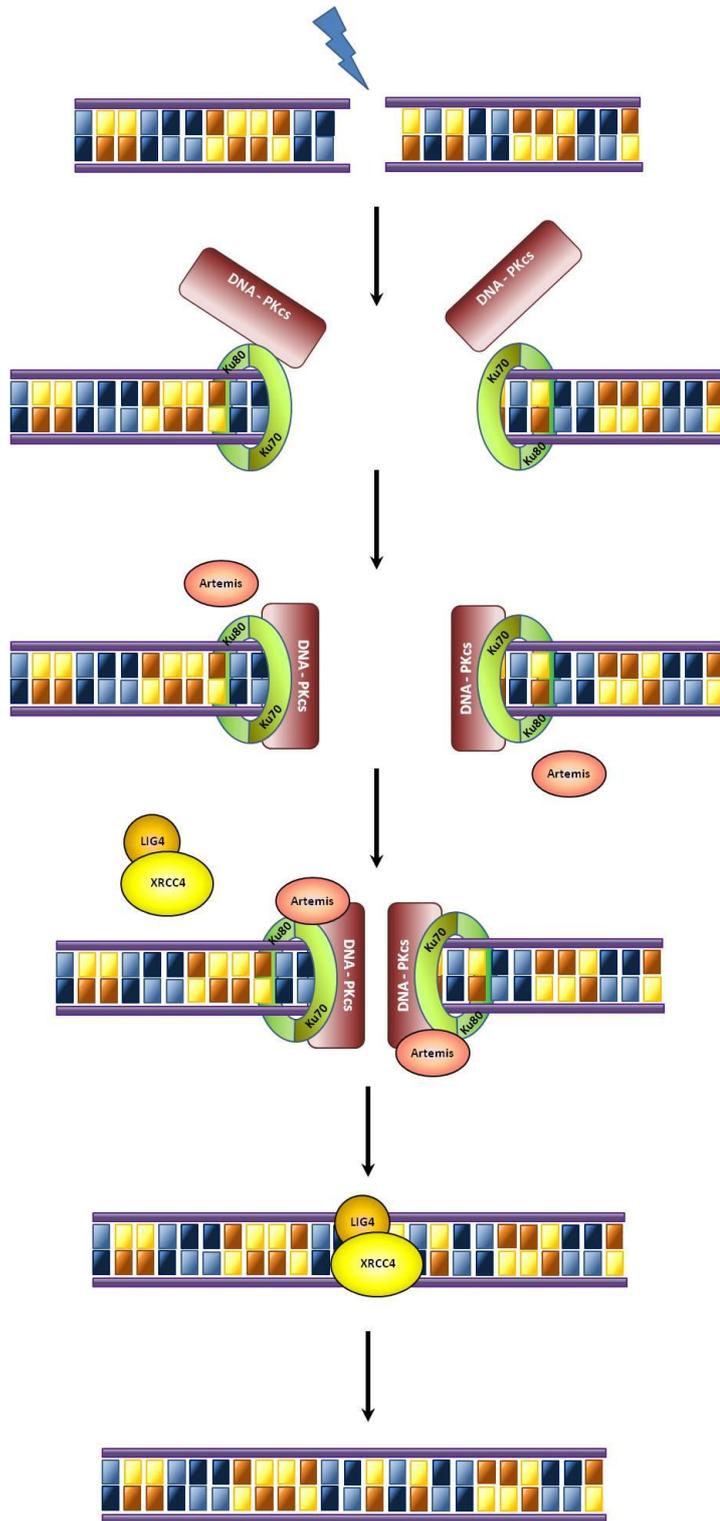


Figura 7, resumen esquemático del mecanismo de reparación no homólogo de unión de extremos de las roturas dobles de cadena de ADN. Abreviaturas: DNA-PKcs, *DNA-dependent protein kinase, catalytic subunit*; Ku70, *X-ray repair cross-complementing protein 6*; Ku80, *X-ray repair cross-complementing protein 5*; Artemis, *DNA cross-link repair 1C*; LIG4, *ligase 4*; XRCC4, *X-ray repair cross-complementing protein 4*.

Para revertir la metilación de la guanina es necesaria la participación de la enzima O⁶-metilguanina-DNA metiltransferasa, que actúa de forma directa sobre la lesión de ADN, uniendo el grupo metilo de la guanina a uno de sus residuos de cisteína, perdiendo su funcionalidad al finalizar (Pegg 1990). Sin embargo, la mayoría de procesos para reparar las lesiones del ADN cuentan con la participación coordinada de un conjunto de enzimas y proteínas. El mecanismo de reparación por escisión de una base (*base excision repair*, BER) (Figura 3) se encarga de reparar aquellas lesiones que consisten en la sustitución de una base nitrogenada por otra que no es la correcta. El procedimiento consiste en varios pasos mediante los que se produce la hidrólisis de los enlaces N-glicosídicos de la base incorrecta para su escisión y la incorporación de la base correspondiente en el lugar que queda libre (Nemec et al. 2010). En ese mecanismo de reparación también participan las enzimas ADN-glicosilasas, que son las encargadas de corregir los errores por desaminación de las bases nitrogenadas (Krokan and Bjoras 2013). Cuando el error se origina por la pérdida de una base, quedando un lociapurínico o apirimidínico, son las AP endonucleasas las que reconocen la lesión. Los mecanismos de reparación por escisión de un nucleótido (*nucleotide excision repair*, NER) (Figura 4) son similares a los anteriores, pero actúan sobre el nucleótido completo. Estos mecanismos se encargan principalmente de eliminar los aductos más visibles, como los dímeros de pirimidinas, aductos cisplatino-purina y aductos formados por carcinógenos policíclicos como el acetilaminofloreno (Wood 1996). La reparación del ADN por desapareamiento (*mismatch repair*, MMR) (Figura 5) es el mecanismo que media el reconocimiento y la reparación de los errores producidos durante la replicación del ADN en los que la enzima ADN-polimerasa introduce una base incorrecta en un loci determinado. Del mismo modo que en el caso anterior, este mecanismo requiere de la actuación conjunta de una serie de enzimas que participan en equilibrio y de forma coordinada (Li 2008). Si hablamos de la lesión más drástica que puede sufrir la molécula de ADN, las roturas de cadena son sin duda una de ellas, en especial cuando son las dos cadenas las que se ven dañadas, quedando la molécula dividida en fragmentos. La reparación de las roturas simples de cadena (*single strand breaks*, SSB) requieren de una enzima ligasa que sea capaz de unir los extremos libres de la cadena dañada. Por su parte, la reparación de las roturas dobles de cadena (*double strand breaks*, DSB) requieren de un mecanismo mucho más complicado, que incluye una serie de complejos proteicos y otras enzimas. Los fallos producidos en la reparación de este tipo de lesiones pueden suponer la pérdida de fragmentos enteros de cromosomas o la

unión de fragmentos no consecutivos que hagan perder el sentido de la codificación genética, lo que supone la muerte o senescencia celulares, así como el origen de células cancerosas. Las roturas dobles de cadena requieren de un complicado proceso para su reparación. Este puede producirse mediante mecanismos de recombinación homóloga (*homologous recombination*, HR) (Figura 6), en los que previamente se produce una resección de cada una de las cadenas dañadas, utilizando de modelo el cromosoma homólogo para la síntesis de las cadenas lesionadas (Le Guen et al. 2015). O bien, la reparación de DSB puede llevarse a cabo mediante un mecanismo no homólogo de unión de extremos (*non-homologous end joining*, NHEJ) (Figura 7), en el cual se procede a la unión de los extremos libres producidos en la lesión sin emplear el cromosoma homólogo como molde. Debido a que este segundo mecanismo no es selectivo a la hora de la reparación, ya que puede unir cualquier extremo de ADN que se encuentre libre, clásicamente se ha tendido a pensar que la NHEJ favorece los procesos de mutación, mientras que la HR es un mecanismo conservador de la secuencia de ADN (Lieber 2010).

Los mecanismos de reparación del ADN son imprescindibles para mantener la integridad celular, ya que las lesiones en la molécula de ADN pueden generar mutaciones que impidan el normal funcionamiento de la célula, provocando la muerte de la misma o en algunos casos contribuyendo a su *malignización*, lo que puede derivar en una célula cancerígena. Sin embargo, pequeñas modificaciones genéticas son imprescindibles para generar la variabilidad necesaria para permitir a los organismos adaptarse a los cambios del medio. Los mecanismos de reparación deben mantener el equilibrio entre la estabilidad genómica y la diversidad genética (Le Guen et al. 2015).

1.3.1.2.- Genes implicados en la reparación del ADN

Los mecanismos de reparación del ADN son procesos celulares muy complejos en los que participan una gran cantidad de moléculas. Para cada mecanismo de reparación existen un conjunto de enzimas con diversas funciones que van desde el reconocimiento de las lesiones o errores, la eliminación de las bases o nucleótidos erróneos, la generación de extremos apropiados para la incorporación de nuevos nucleótidos, la propia incorporación de nucleótidos y por último, la unión de los extremos libres (Wood et al. 2001). Además, los mecanismos de reparación están en estrecha relación con otros procesos celulares, como la proliferación celular, la

replicación del ADN o la apoptosis, lo que facilita el mantenimiento de un equilibrio entre todos ellos, evitando que células inviables continúen dividiéndose. Detrás de esta regulación se encuentran una serie de moléculas que gestionan todos estos procesos celulares y que tradicionalmente se han dado en llamar “guardianes del genoma” (Tong et al. 2001; Meulmeester and Jochemsen 2008). En este apartado vamos a desarrollar aquellos genes relacionados con algunos mecanismos de reparación del ADN que han sido incluidos en el presente estudio (Tabla 2).

Tabla 2. Genes de reparación del ADN. Tabla modificada de Wood et al. 2001

Nombre del gen (sinónimos)	Actividad	Localización cromosómica
<i>XRCC1</i>	Función de ligamiento de enzimas PARP	19q13.2
<i>ERCC2 (XPD)</i>	Función helicasa de extremo 5' a 3' del ADN	19q13.2-q13.3
<i>ERCC1</i>	Subunidad de incisión en 5'	19q13.2-q13.3
<i>XRCC6 (Ku70)</i>	Unión a los extremos de ADN	22q13.2-q13.31
<i>XRCC5 (Ku80)</i>	Unión a los extremos de ADN	2q35
<i>LIG4</i>	Función ligasa entre los extremos libres de ADN	13q33-q34
<i>PARP1</i>	Unión a los extremos de cadena simple de ADN	
<i>ATM</i>	Organización de procesos celulares	11q22-q23
<i>MVP</i>	Organización de procesos celulares	16p13.1-p11.2
<i>p53</i>	Organización de procesos celulares	17p13

Abreviaturas: *XRCC1*, *x-ray repair cross-complementing group 1*; *ERCC2*, *excision repair cross-complementing rodent repair deficiency, complementation group 2*; *ERCC1*, *excision repair cross-complementing rodent repair deficiency, complementation group 1*; *XPD*, xeroderma pigmentosa grupo D; *XRCC6*, *x-ray repair complementing defective repair in Chinese hamster cell 6*; *XRCC5*, *x-ray repair complementing defective repair in Chinese hamster cell 5*; *LIG4*, ligasa 4; *PARP1*, *poly (ADP-ribose) polymerase*; *ATM*, ataxia telangiectasia mutada; *MVP*, *major vault protein*.

La mayoría de los genes que se han seleccionado codifican proteínas con actividad enzimática que de una forma directa participan en los procesos descritos anteriormente. El gen que codifica para la proteína *x-ray repair cross-complementing group 1* (*XRCC1*), se localiza en el cromosoma 19q13.2. Esta proteína se ha descrito como uno de los agentes imprescindibles para la reparación de tipo SSB, mediante su interacción con la mayoría de enzimas operantes en este mecanismo (Thompson and West 2000). Se ha reportado su interacción física con la ADN ligasa III α (Caldecott et al. 1996), la ADN polimerasa β (Caldecott et al. 1996; Kubota et al. 1996) y las poli (ADP-ribosa) polimerasas 1 y 2 (Masson et al. 1998; Schreiber et al. 2002). En el trabajo de Brem et al. se demuestra claramente que *XRCC1* es necesaria para la reparación eficiente de SSBs y el mantenimiento de la estabilidad genómica (Brem and Hall 2005). Muchas de estas enzimas participan a su vez en los mecanismos BER, por lo que también es imprescindible la coordinación de *XRCC1* en estos mecanismos (Marsin et al. 2003). También en el cromosoma 19q13.2-q13.3 se localiza los genes que

codifican para las proteínas *excision repair cross-complementing rodent repair deficiency, complementation group 1* (ERCC1) y *excision repair cross-complementing rodent repair deficiency, complementation group 2* (ERCC2). Ambas forman parte del complejo proteico que actúa en los mecanismos NER con actividades diferentes: mientras ERCC2 tiene actividad helicasa, ERCC1 se encarga de formar el extremo 5' (Evans et al. 1997). En los pacientes que sufren xeroderma pigmentosa (una alteración de genes relacionados con la reparación del ADN), la proteína ERCC2 es deficiente. Estos pacientes se caracterizan por ser muy sensibles a la radiación UV y son 2000 veces más propensos a padecer cáncer de piel, así como a padecer otras afecciones de la piel a causa de su exposición al sol (Lehmann 2001). La proteína Ligasa 4 (LIG4) está codificada en el cromosoma 13q33-q34 y su actividad enzimática consiste en unir los extremos libres de ADN. Es la enzima ligasa más flexible conocida, es capaz de ligar extremos de ADN incompatibles (Gu et al. 2007) y participa en la reparación de DSB mediante el mecanismo no homólogo (NHEJ) (Grawunder et al. 1998). Otros genes que tienen un papel imprescindible en el mecanismo de reparación NHEJ son los que codifican para las proteínas *X-ray repair complementing defective repair in Chinese hamster cell 6* (XRCC6, también conocida como ku70) y *X-ray repair complementing defective repair in Chinese hamster cell 5* (XRCC5, también conocida como ku80) que se localizan en los cromosomas 22q13.2-13.31 y 2q35 respectivamente. Estas proteínas reconocen los extremos de la doble cadena de ADN y se une a estos favoreciendo el reclutamiento del resto de enzimas participantes en su reparación (Lieber 2008). La familia de proteínas polimerasas (*poly (ADP-ribose) polymerases*, PARP) está compuesta por hasta 17 proteínas diferentes. La proteína PARP1 es una enzima que tiene afinidad por los extremos de cadena simple de ADN. Una vez se une a la cromatina, induce el reclutamiento del resto de maquinaria de reparación de SSB (Berger 1985).

Otros de los genes que han sido seleccionados para este estudio codifican para proteínas que hasta el momento no se han descrito como enzimas con actividad directa en los mecanismos de reparación. Sin embargo, son proteínas que tienen papeles muy relevantes en el control de muchos procesos celulares, principalmente el control del ciclo celular y de la senescencia y muerte celulares en respuesta a las lesiones genéticas; estas proteínas son iniciadoras de los mecanismos de reparación del ADN. La proteína kinasa ataxia-telangiectasia mutada (ATM) viene codificada en el cromosoma 11q22-

q23. No participa directamente como parte de ningún mecanismo de reparación de ADN, si no que regula la respuesta del ciclo celular cuando se produce una lesión (Shiloh and Rotman 1996). En particular, regula la ruta que conduce a la muerte celular programada (apoptosis), así como varios puntos de control del ciclo de proliferación celular. Las mutaciones que provocan defectos en la proteína ATM son las responsables de causar el síndrome de ataxia-telangiectasia, que se caracteriza por una neurodegeneración progresiva (ataxia), así como una dilatación de los vasos sanguíneos (telangiectasia). Además, los pacientes que sufren este síndrome padecen infecciones recurrentes por deficiencias en la respuesta inmune y presentan mayor predisposición a padecer algunos tipos de cáncer, siendo también particularmente sensible a los efectos de la radiación ionizante (Canman and Lim 1998). La proteína p53 aparece mutada en el 60% de las neoplasias (Levine et al. 1991) y el gen que la codifica (situado en el cromosoma 17p13) es considerado uno de los primeros genes supresores de tumores descritos como tal. Es una proteína clave en varios procesos celulares, regulando el ciclo celular y la apoptosis en función de las posibles lesiones producidas en el ADN (Levine and Oren 2009) (interacciona con la proteína ATM). El gen que codifica para *major vault protein* (MVP) se localiza en el cromosoma 16p13.1-p11.2. Esta proteína forma parte de un complejo proteico que de forma genérica se denomina como *Vault* por su similitud estructural a la de un barril, el cual está formado a su vez por otras dos partículas proteicas más pequeñas (vPARP y una telomerasa, TEP1), así como por un conjunto de pequeños fragmentos de ARN no codificante (vRNA) (van Zon et al. 2003). En un inicio, MVP se relacionó con la resistencia a fármacos. Con posterioridad se ha asociado con otros procesos celulares, entre ellos con una posible participación en el mecanismo de reparación NHEJ por su interacción con las proteínas Ku (Lloret et al. 2009).

1.3.1.3.- SNPs en genes de reparación del ADN y CaP

La capacidad de reparación de las lesiones de ADN es una característica celular esencial para el mantenimiento de la integridad genética. Las alteraciones que afectan a esta capacidad están asociadas con el desarrollo de ciertas enfermedades, entre ellas el cáncer. Previamente se han nombrado algunas deficiencias en las que, además de otros efectos, se observa una mayor predisposición a desarrollar determinados tipos de tumores (ataxia-telangiectasia y la xeroderma pigmentosa). Una célula cancerígena no es más que una célula que ha adquirido capacidades que le permiten sobrevivir y

proliferar sin control y en condiciones extremas. Las nuevas capacidades adquiridas son el resultado de la acumulación de mutaciones genéticas y para ello, las deficiencias en los mecanismos de reparación del ADN tienen un papel primordial. Estas variaciones genéticas hacen que se produzca un fallo generalizado de los sistemas de control celular de modo que, aun siendo células inestables genéticamente, no entran en apoptosis, sino que siguen proliferando de forma ilimitada. Las células afectadas, cuanto más sobreviven sin control, presentan más probabilidad de seguir acumulando errores que le proporcionan la facultad de escapar al control inmunológico o de migrar a nuevos tejidos. Es por ello que las células tumorales son tan difíciles de eliminar, haciendo del cáncer una enfermedad compleja en su tratamiento. En la última década, probablemente cientos de polimorfismos de nucleótido simple (SNPs) en genes de reparación del ADN han sido estudiados con el fin de establecer una relación con el riesgo de padecer distintos tipos de cáncer. La revisión de Köberle et al. analiza de una forma muy extensa y detallada las publicaciones que existen hasta la fecha sobre SNPs localizados en genes de reparación del ADN y el riesgo de desarrollar cáncer, así como con la respuesta al tratamiento para algunos de ellos (Koberle et al. 2016).

Además de su relación con la aparición de errores genéticos en las células tumorales, los mecanismos de reparación del ADN también han sido asociados con la respuesta al tratamiento con fármacos (Zamble and Lippard 1995) y a radioterapia (Parliament and Murray 2010). Tanto la quimioterapia como la radiación ionizante son factores que inducen daño en el ADN, siendo este su principal mecanismo de acción. Sin embargo, algunos tejidos normales se ven afectados por sus efectos; que estos presenten mayor o menor grado de toxicidad depende de su capacidad para reparar las lesiones genéticas. Esta variabilidad entre pacientes, tanto en la respuesta al tratamiento radioterápico, como en el desarrollo de toxicidad dependiente de la radiación en el tejido normal, es uno de los retos a los que se enfrentan los médicos en el manejo clínico de los pacientes con cáncer. En la búsqueda de estrategias que faciliten la programación de un tratamiento individualizado, muchos de los análisis de variabilidad genética realizados en los últimos años van dirigidos a estudiar la radiotoxicidad en tejido sano; en este campo, los genes involucrados en la reparación del ADN son los más estudiados. Del mismo modo que con el riesgo de cáncer, son muchas las publicaciones que existen en relación a los tratamientos, por lo que ya hay varias revisiones que tratan de recoger toda la información relacionada. La revisión de

Borchiellini et al. contiene los SNPs de genes de reparación del ADN que se han estudiado en relación a la respuesta a la radioterapia (Borchiellini et al. 2012) y lo mismo pero para quimioterapia viene recogido en la revisión de Bouwman et al. (Bouwman and Jonkers 2012).

En CaP existen varios estudios en los que se analiza el valor pronóstico de determinados SNPs localizados en genes involucrados en la reparación del ADN, en relación con el riesgo de padecer este tipo tumoral, así como las posibilidades que estos resultados ofrecen para el desarrollo de terapias personalizadas basadas en biomarcadores específicos, tal y como se está aplicando en otros tipos tumorales (Flaherty et al. 2012). Recientemente, la red que investiga el atlas genómico del cáncer (*The Cancer Genome Atlas Research Network*, TCGA), ha publicado la secuencia exónica completa de 333 pacientes con CaP localizado y en 62 de los 333 pacientes (19%) los tumores presentaban aberraciones somáticas o de línea germinal en genes clave de los procesos de reparación del ADN (*BRCA1*, *BRCA2*, *CDK12*, *ATM*, *FANCD2* y *RAD51C*) (Cancer Genome Atlas Research 2015). En cuanto al riesgo de padecer CaP, las mutaciones germinales en *BRCA1* aumentan este riesgo (Leongamornlert et al. 2012). En *XRCC1* se han analizado diversos polimorfismos en estudios de casos y controles en los que se ha determinado que el SNP R399Q (codón Arg399Gly) influye en el riesgo de desarrollar CaP (Xu et al. 2007). Sin embargo, este resultado no se ha validado con un número mayor de pacientes en la población caucásica (Rybicki et al. 2004). En el gen *PARP*, el polimorfismo rs1136410 está relacionado con el riesgo de CaP y con disminución en la función enzimática en respuesta a daños oxidativos (Lockett et al. 2004). En *XPD* se han estudiado varios polimorfismos, pero de la mayoría de los estudios no se sustraen resultados significativos (Rybicki et al. 2004; Ritchey et al. 2005; Bau et al. 2007).

Aunque existen algunas asociaciones, en general los estudios de variaciones genéticas no muestran resultados concluyentes. Muchos de ellos se realizan en poblaciones muy pequeñas y los resultados que se obtienen no son validados cuando se incluye un mayor número de pacientes. Sin embargo, el convencimiento de que la variabilidad genética es un elemento clave para la individualización del tratamiento de los pacientes hace que cada año aumente el número de publicaciones en relación a este tema de estudio.

1.3.2.- Metabolismo de andrógenos

El CaP es un cáncer hormono-dependiente, esto es, que requiere de la presencia de andrógenos para su iniciación, promoción y progresión (Huggins 1967). Es por ello que la ablación androgénica ha sido clásicamente la principal intervención terapéutica para el tratamiento de CaP. Sin embargo, en un momento determinado del desarrollo tumoral, un gran número de los tumores de próstata progresan a un estado de insensibilidad androgénica, de modo que estos tratamientos dejan de tener efecto (Eisenberger et al. 1985). Las terapias anti-androgénicas se han dirigido en primera instancia a la supresión de los andrógenos de origen testicular, bien mediante castración quirúrgica o mediante el empleo de hormonas sintéticas. En estos casos, la producción de andrógenos continúa en las glándulas suprarrenales, lo que dificulta la supresión completa de la síntesis de estas hormonas (Eisenberger et al. 1998). El desarrollo de tumores resistentes al tratamiento se manifiesta mediante un aumento de los niveles de PSA en suero, progresión de la enfermedad y síntomas de empeoramiento. No obstante, no todos los tumores resistentes a hormonas manifiestan la misma respuesta a tratamientos de deprivación androgénica (TDA); se ha desarrollado una clasificación de estos tumores resistentes a hormonas basada en dicha respuesta al tratamiento y se han propuesto distintas terapias hormonales para cada uno de ellos (Rini and Small 2002). La resistencia a los TDA supone en la actualidad un reto para el manejo de los pacientes con CaP.

Los andrógenos son las hormonas sexuales masculinas derivadas del ciclopentanoperhidrofenantreno, siendo su principal y más abundante representante en suero la testosterona (T). La T se sintetiza principalmente en los testículos y en menor proporción (10 %) en las glándulas suprarrenales. Los andrógenos son los precursores de las hormonas estrogénicas, por lo que pequeñas cantidades de T también se producen en los ovarios y en las glándulas suprarrenales de las mujeres. La T es la responsable de los caracteres sexuales masculinos y su principal función en adultos es el correcto desarrollo espermático. La figura 8 muestra un esquema de los procesos metabólicos mediante los que se sintetiza la T. La síntesis de T está mediada por la enzima 17-hidroxisilasa/17, 20-liasa, codificada por el gen del citocromo P450 17A1 (*CYP17A1*). La dihidrotestosterona (DHT) se conoce como el derivado prostático de la T, aunque se produce no solo en la próstata, sino también en los testículos, las glándulas suprarrenales y los folículos capilares. El proceso metabólico por el cual la T se

convierte en DHT está catalizado por la enzima 5-alfa-esteroide reductasa (SRD5A) (McConnell 1995). Esta hormona tiene más afinidad por los receptores androgénicos que la propia T. El papel de la DHT en el desarrollo del CaP fue conocido mediante el descubrimiento de la deficiencia del polipéptido alfa 2 de la enzima 5-alfa-esteroide reductasa (SRD5A2), pues los individuos que padecían esta alteración no presentaban ni hiperplasia benigna de próstata (HBP) ni CaP. Desde que esto se hizo patente, ha aumentado el interés por desarrollar inhibidores de 5-alfa-esteroide reductasa (5 α -RI). De entre los diferentes compuestos desarrollados, solo dos fármacos han sido aprobados por la *U.S. Food and Drug Administration* (FDA) para el uso clínico: finasterida, que inhibe SRD5A2, aprobado para el tratamiento de HBP y alopecia masculina; y dutasterida, el cual inhibe SRD5A1 y 2, aprobado para el tratamiento de HBP (Edwards and Moore 2002; Andriole and Kirby 2003). Teniendo en cuenta que la DHT es el principal andrógeno prostático, el empleo de 5 α -RIs parece un método razonable para la prevención del CaP (Thompson et al. 2003; Andriole et al. 2010). Sin embargo, los estudios dirigidos a analizar estos resultados no determinan si el efecto de los fármacos se debe a su capacidad de prevención o a la reducción de tumores ya existentes (Walsh 2010).

1.3.2.1.- SNPs en genes del metabolismo de andrógenos y CaP

Desde que los andrógenos han sido considerados como un factor de riesgo para el CaP, ha aumentado el interés por estudiar los polimorfismos relacionados con su síntesis y metabolismo. La revisión de Schleutker recoge algunos de los genes que participan en la síntesis y metabolismo de la T, así como sus SNPs más relevantes en relación al CaP (Schleutker 2012).

El gen *CYP17A1* se localiza en el cromosoma 10q24.3 y codifica para un miembro perteneciente a la familia de enzimas del citocromo P450. Tiene actividad 17 α -hidroxilasa y 17,20-liasa en el metabolismo de la T. La existencia de relación entre la presencia de ciertos polimorfismos en el gen *CYP17A1* y el CaP es un hecho cuestionado (Lindstrom et al. 2007; Setiawan et al. 2007; Mononen and Schleutker 2009). Sin embargo, se han sugerido que algunos polimorfismos de *CYP17A1* podrían estar asociados con el riesgo de CaP (Wang et al. 2011). A pesar de los resultados contradictorios, *CYP17A1* sigue considerándose un factor importante en la progresión del CaP y en la actualidad es diana terapéutica para el fármaco acetato de abiraterona, el

cual ha demostrado ser muy activo en el tratamiento del CaP mediante el bloqueo específico de esta enzima (Yang 2011). No obstante, solo un estudio analiza la relación entre SNPs localizados en el gen *CYP17A1* y la respuesta al tratamiento con acetato de abiraterona, realizado en pacientes metastásicos de CaP resistente a castración. En este estudio se relaciona el polimorfismo rs2486758 con menor tiempo de progresión bioquímica (Binder et al. 2016).

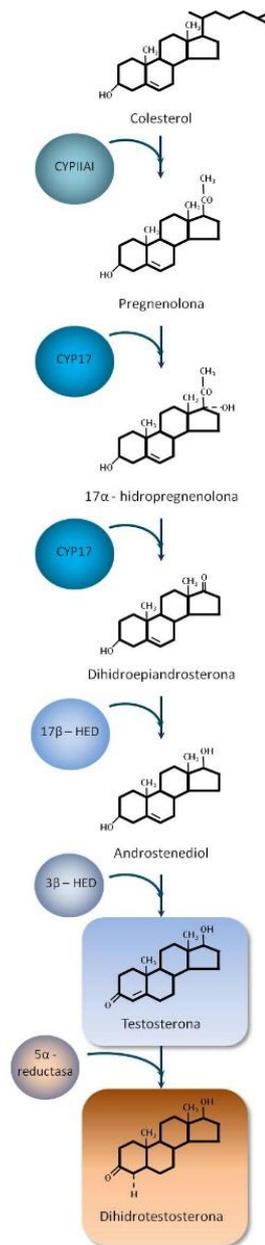


Figura 8, esquema de la ruta metabólica de la testosterona y su posterior derivación a dihidrotestosterona, así como de las enzimas que participan en ella. *Abreviaturas: CYP11A1, cytochrome P450sc; CYP17, steroid 17α-monooxygenase; 17β-HED, 17-beta-hidroxiesteroide deshidrogenasa; 3β-HED, 3-beta-hidroxiesteroide deshidrogenasa.*

Los genes que codifican para las enzimas SRD5A1 y SRD5A2 se localizan en los cromosomas 5p15.31 y 2p23, respectivamente. Ambas enzimas convierten la T en DHT mediante un proceso irreversible. En estos genes se han identificado varios polimorfismos relacionados con el riesgo de padecer CaP (Paz-y-Mino et al. 2009; Li et al. 2010; Setlur et al. 2010; Li et al. 2013). Sin embargo, para la mayoría de SNPs, los resultados son bastante contradictorios. Hasta la fecha, no se han publicado estudios que relacionen SNPs localizados en estos genes con la respuesta al tratamiento con finasterida o dutasterida, a pesar de que es sabido que la presencia de determinados polimorfismos en un paciente puede condicionar la efectividad de determinados fármacos (Ma and Lu 2011).

Los genes implicados en la síntesis y metabolismo de andrógenos son claros candidatos a presentar relación con el riesgo de padecer CaP, así como con la respuesta a los tratamientos que inhiben de forma específica la actividad de las enzimas que codifican. No obstante, son muy pocos los estudios que han tratado de establecer esta relación y los que existen no muestran resultados concluyentes. Esto no ha sido impedimento para que en la actualidad los fármacos que actúan como inhibidores de estas enzimas tengan una aplicación relevante en la práctica clínica de los pacientes con CaP.

1.3.3.- Polimorfismos de nucleótido simple (SNPs)

El material genético de las células humanas se organiza en un total de cuarenta y seis cromosomas, veintitrés de ellos procedentes de un progenitor y los otros veintitrés procedentes del otro. Cada molécula de ADN (consistente en todo el material genético contenido en los cuarenta y seis cromosomas), está formada por una fracción codificante, es decir, que contiene la información necesaria para la síntesis de proteínas (los exones dentro de los genes) y por una fracción no codificante, que contiene elementos del ADN como los intrones, las secuencias estructurales y algunas otras regiones de las cuales aún no se conoce su función. Solo aproximadamente el 1.1% del ADN está formado por estructuras codificantes (Venter et al. 2001). No toda la secuencia de nucleótidos de la molécula de ADN es igual en todos los individuos. Son precisamente las variaciones en dicha secuencia de ADN las que proporcionan las diferencias de fenotipo entre los distintos sujetos. Incluso dentro de la misma célula, el material genético procedente de cada progenitor contiene una secuencia distinta. Los

alelos son las formas alternativas que puede presentar un mismo gen y que se pueden manifestar con diferencias en la función de dicho gen, produciendo variaciones en las características hereditarias. Dentro de una población existe un número determinado de alelos para cada gen, que viene representado por las secuencias de ADN más frecuentes en esa población para ese gen en concreto. Cuando en la estructura del ADN se produce una alteración en la secuencia que genera una variación diferente a la de los alelos más frecuentes, entonces hablamos de mutación genética. Estas alteraciones suelen generarse a lo largo de la vida de un individuo de forma natural, debido a fallos en el proceso de replicación del ADN o mediante sustancias que el propio organismo genera, y de forma no natural mediante la acción de diferentes agentes externos. La mayoría de las mutaciones se corrigen. Generalmente no son transmitidas a la descendencia; solo en el caso de que se produzcan en las células germinales se trataría de mutaciones hereditarias.

Los polimorfismos son variaciones genéticas que se generan mediante mutaciones hereditarias y que son transmitidas a varias generaciones de descendientes, de modo que su presencia alcanza cierta relevancia en la población (por encima del 1% de los integrantes). Estas variaciones pueden localizarse en zonas no codificantes del genoma, considerándose variaciones silenciosas, pues no tienen repercusión en el fenotipo del individuo; o bien pueden conferir al portador alguna característica diferente al resto de miembros de la población, pudiendo aportarle alguna ventaja adaptativa, o por el contrario, favorecer el desarrollo de determinadas enfermedades (Guttmacher and Collins 2002). Dentro de las posibles modalidades de variaciones genéticas, los polimorfismos de nucleótido simple (*single nucleotide polymorphisms*, SNP) son los más frecuentes, constituyendo hasta el 90% de las variaciones genéticas humanas. Estos polimorfismos consisten en la sustitución de un único nucleótido de la secuencia de ADN y aparecen de media a lo largo del genoma humano cada 1300 bases. Las variaciones en una única base pueden modificar la lectura de la secuencia de ADN y dar origen a proteínas alteradas en su estructura lo que, en general, supone una pérdida de la funcionalidad proteica. Este es el principal mecanismo mediante el cual una variación genética puede dar origen a una enfermedad. Sin embargo, existen otros mecanismos. Puede por ejemplo darse el caso contrario, en el que una proteína aumenta su capacidad funcional originando una situación de toxicidad o pérdida de control de procesos (Guttmacher and Collins 2002). En cualquier caso, desde que se publicó la secuencia

completa del genoma humano, el interés por relacionar las variaciones genéticas con el desarrollo de enfermedades ha crecido exponencialmente, viéndose reflejado en la multitud de publicaciones científicas que se han reportado desde entonces.

Los avances en la tecnología de genotipado, que permite la obtención de resultados más rápidos y baratos, han contribuido a que cada vez sean más los estudios observacionales que tratan de relacionar variaciones en genes candidatos con el riesgo de padecer diversas enfermedades, entre las que destaca la patología tumoral (Lin et al. 2006). La metodología llamada “del gen candidato” consiste en realizar estudios individuales en los que es el investigador el que selecciona los polimorfismos que va a someter a análisis, habitualmente localizados en genes involucrados en procesos celulares previamente relacionados con el desarrollo de la enfermedad que esté estudiando (cáncer en este caso). Una de las limitaciones que presentan este tipo de estudios de asociación genética es el número de casos incluidos. Así, para que los resultados estadísticos sean válidos y representativos de la población, el tamaño de la muestra ha de ser lo suficientemente grande (Altmuller et al. 2001). En las primeras publicaciones de genes candidatos se cuestionó la validez de los resultados estadísticamente significativos debido a que el número de pacientes era bastante reducido. Para tratar de validar estos resultados, se procedió al desarrollo de meta-análisis, aumentando de este modo el tamaño muestral. No obstante, algunos autores discutieron la validez de esta solución al presentar una heterogeneidad estadísticamente significativa entre los estudios que componían cada uno de los meta-análisis (Ioannidis et al. 2001). Además de la limitación del tamaño muestral existen otros factores limitantes tales como el coste excesivo de este tipo de diseños experimentales “gen a gen”, factores de confusión debidos a la estructura poblacional y geográfica o sesgos en la selección. Con todo, este tipo de diseño ha sido mejorado y en la actualidad ha sido sustituido en parte por estudios de tipo *genome-wide association studies* (GWAS), o lo que es lo mismo, el estudio de las variaciones genéticas de todo el genoma o de regiones muy amplias de este, en relación a la presencia de la enfermedad (Wang et al. 2005). Los estudios GWAS representan un diseño más eficiente a la hora de diseccionar genéticamente las enfermedades complejas si lo comparamos con los estudios de asociación tradicionales. Una de las principales ventajas que ofrecen estos estudios es que no parten de una hipótesis (como ocurre en los estudios de gen candidato, en los que se seleccionan los genes a estudiar), sino que permiten poner a prueba al genoma

completo (Seng and Seng 2008). Ambos tipos de estudios han servido para evaluar la asociación entre variaciones genéticas y el riesgo de padecer cáncer, y de forma concreta con el riesgo de padecer cáncer de próstata.

1.3.3.1.- SNPs y CaP.

La revisión de Dong et al. recoge los resultados de todos los estudios de meta-análisis observacionales publicados hasta 2008 en los que se evaluaba la asociación entre polimorfismos genéticos y el riesgo de padecer cáncer. Para ello, se seleccionaron aquellos estudios que cumplían ciertos requisitos: i) que incluyese al menos 500 casos entre todos los estudios que comprendían el meta-análisis, ii) que evaluaran el riesgo de padecer cáncer, iii) que no trataran sobre marcadores genéticos de los antígenos leucocitarios humanos (*human leukocyte antigen*, HLA) y iv) que estuvieran publicados en inglés (Dong et al. 2008). De esta recopilación se sustrae que en esa fecha se habían realizado estudios en 18 localizaciones tumorales diferentes (siendo las más abundantes cáncer de mama, cáncer de próstata y cáncer de pulmón), para polimorfismos en un total de 99 genes diferentes (principalmente genes relacionados con la reparación de ADN) (Dong et al. 2008). De igual modo, Ioannidis et al. publicaron más tarde una revisión recopilando los estudios más relevantes que mediante GWAS evaluaban la relación de las variaciones genéticas con el desarrollo del cáncer (Ioannidis et al. 2010).

Los estudios epidemiológicos hasta la fecha sugieren que el desarrollo del CaP está fuertemente influenciado por factores genéticos (Lichtenstein et al. 2000; Schaid 2004). Es por ello que existe un gran interés en determinar el papel que presentan los SNPs, no solo en el riesgo de padecer cáncer, sino también su papel como marcador pronóstico y predictivo. Muestra de ello son las decenas de publicaciones que se han generado a partir de estas premisas, especialmente basadas en estudios GWA y que Van den Broeck et al. recogen excepcionalmente en una reciente revisión (Van den Broeck et al. 2014). Actualmente, la determinación de polimorfismos no tiene una aplicación real en la práctica clínica como factor de riesgo del CaP. Sin embargo, cuando se combina con la determinación de los niveles de PSA en suero, el papel predictivo de esta última se ve incrementado (Nam et al. 2009; Aly et al. 2011). Más aún, resultados recientes sugieren que los niveles de PSA en suero también están sujetos a variabilidad genética, lo que podría explicar entre el 40 y el 45% de las variaciones de los niveles de PSA observados en la población (Gudmundsson et al. 2010). Hasta la fecha, alrededor

de 100 SNPs han sido relacionados con el riesgo de desarrollar CaP, los cuales explicarían el 35% del componente heredable de este tipo tumoral (Eeles et al. 2013). Con el desarrollo de la tecnología y el incremento en el conocimiento sobre polimorfismos, los estudios de variaciones genéticas incluyen cada vez más SNPs. Los resultados sugieren que incrementando el número de SNPs en estudio, aumenta su valor en la predicción del riesgo de desarrollar CaP. En el trabajo de Lindstrom et al. (Lindstrom et al. 2012) se evaluaron 25 SNPs relacionados con el riesgo de CaP en 7509 pacientes y 7652 controles, encontrando que los portadores de mayor número de estos SNPs presentaban cinco veces más riesgo de ser diagnosticados de CaP. Sin embargo, los resultados no son tan claros cuando se trata de relacionar los polimorfismos con la agresividad tumoral; además hay que tener en cuenta que para muchas de las asociaciones encontradas aún no se conoce su significado biológico con claridad (Xu et al. 2010).

1.3.4.- Biomarcadores clínicos en cáncer de próstata.

Para la mayoría de tumores sólidos, la extensión anatómica de la enfermedad es considerada el principal factor predictivo de supervivencia. Desde su desarrollo en 1958, el sistema de clasificación del estadio tumoral TNM es la principal herramienta de determinación de la extensión anatómica y es la base para la elección del tipo y la intensidad del tratamiento inicial (Greene and Sobin 2008). Con posterioridad, otros factores como el tipo y el grado histológicos, la edad del paciente o su comorbilidad, han contribuido a completar la información de supervivencia y la respuesta terapéutica, permitiendo establecer un tratamiento más específico a las condiciones del paciente. Sin embargo, en el camino hacia el tratamiento individualizado, son los marcadores biológicos o biomarcadores los que están adquiriendo especial relevancia. En las últimas décadas, la incorporación de los marcadores biológicos en el estudio clínico del paciente oncológico está permitiendo dividir las clases tumorales tradicionales en subgrupos que presentan diferentes comportamientos y de este modo ajustar la predicción de la supervivencia y la aplicación del tratamiento de una forma más específica a cada paciente concreto. Además de su función clínica, los marcadores biológicos también se emplean en la actualidad como dianas terapéuticas para el desarrollo de fármacos que actúan solo en aquellos tipos tumorales en los que una determinada molécula esté alterada. Uno de los casos más significativos es la determinación de los receptores de estrógeno y progesterona y del marcador Her2/neu

en el cáncer de mama, los cuales determinan el subtipo tumoral, condicionan el pronóstico y establecen qué pacientes se beneficiarían de un posible tratamiento hormonal y cuáles de inmunoterapia anti-Her2 (Buzdar and Howell 2001). La relevancia de la incorporación de marcadores biológicos en la determinación de la supervivencia o del tratamiento a aplicar es controvertida y no es igual para todos los casos. Sin embargo, con los resultados de que se disponen en la actualidad, es razonable pensar que los biomarcadores contribuyen a establecer una terapia más individualizada.

Solo una pequeña parte de los diferentes tipos de cáncer tiene un origen hereditario (alrededor del 10%) (Balmain 2001). No obstante, la genética y la aparición de células cancerígenas están íntimamente relacionadas ya que son las alteraciones genéticas las que habilitan a dichas células de su capacidad de proliferación y metástasis. Y aún las alteraciones genéticas en sí no reflejan la complejidad de los procesos oncológicos. Determinados tipos de marcadores biológicos establecen el estado de algunos fenómenos relacionados con el tumor, como son la angiogénesis, la apoptosis, la proliferación, el metabolismo, la hipoxia o la resistencia a determinados fármacos. Cuando hablamos de marcadores biológicos debemos diferenciar entre marcadores moleculares (receptores de membrana, proteínas con actividad kinasa, antígenos tumorales, péptidos liberados por el tumor...) y marcadores genéticos. En este último se distinguen a su vez marcadores asociados al ADN (polimorfismos de nucleótido simple, aberraciones cromosómicas, alteración del número de copias de fragmentos de ADN, inestabilidad de microsatélites, alteraciones en la metilación de regiones promotoras...) y marcadores asociados al ARN (sobre-expresión o infra-expresión de transcriptomas y modificación de la expresión de ARN regulador - microARN-) (Ludwig and Weinstein 2005). Para evaluar y validar el uso de un marcador biológico en la práctica clínica se han establecido cinco fases: i) exploración preclínica (fase I), ii) análisis clínico y validación (fase II), iii) estudio longitudinal retrospectivo (fase III), iv) cribado prospectivo (fase IV) y v) control del cáncer (fase V) (Pepe et al. 2001). Realizar estas fases de forma que los resultados tengan relevancia científica requiere de una planificación complicada, que en muchos casos es difícil de conseguir (Barker 2003; Mandrekar and Sargent 2009). De la gran cantidad de marcadores biológicos que se han determinado como relevantes en relación al diagnóstico y pronóstico del cáncer, solo una pequeña parte se ha incorporado a la práctica clínica diaria (Ludwig and Weinstein 2005).

1.3.4.1.- Antígeno Prostático Específico (PSA) y sus derivados

La determinación de PSA en suero como test de diagnóstico del CaP es uno de los ejemplos que muestran la gran importancia que representan algunos de estos marcadores biológicos en la práctica clínica actual, y no solo en su papel como factor diagnóstico, sino también como uno de los factores necesarios en la determinación del riesgo de la enfermedad. La incorporación de la determinación de PSA como herramienta de diagnóstico para el CaP ha supuesto un incremento en el número de pacientes diagnosticados de esta enfermedad y sin embargo, desde ese momento, la mortalidad por CaP se ha visto reducida (datos disponibles en: <http://seer.cancer.gov/statfacts/html/prost.html>). Más aún, los niveles de PSA elevados están relacionados con un mayor riesgo de padecer CaP, así como con mayor grado y estadio tumorales y con un mayor riesgo de presencia de metástasis (Catalona et al. 1993). A pesar de ello, son varias las razones que hacen que la determinación de PSA en suero no sea un biomarcador ideal. La principal de ellas es que la determinación del nivel de PSA en suero no permite distinguir entre tumores agresivos de aquellos que no lo son. En la práctica clínica, cuando la medición de PSA es elevada (tradicionalmente por encima de 4 ng/ml de sangre), subsecuentemente se realizan un examen de tacto rectal (TR) y una biopsia de tejido que permitan confirmar el diagnóstico. Sin embargo, de las biopsias que se realizan a pacientes con niveles por encima de esos 4 ng/ml, un bajo porcentaje (del 25% al 30%) revelarían la presencia de cáncer (Parekh et al. 2007), aunque estos resultados no están del todo claros. Del mismo modo, el diagnóstico mediante PSA es suficientemente sensible para detectar muchos de los tumores de bajo riesgo, lo que supone una sobredimensión de los casos diagnosticados y consecuentemente un exceso de tratamiento de los casos de CaP (Howrey et al. 2013), estimando que el porcentaje de pacientes sobrediagnosticados mediante PSA está entre el 25% y el 50% (Auvinen et al. 2016). Pese a todas sus limitaciones, la determinación de PSA en suero es una muestra de la importancia que tienen los marcadores biológicos en el diagnóstico, pronóstico y tratamiento del CaP y en extensión, de la importancia de los biomarcadores en la práctica clínica del cáncer.

La principal debilidad que presenta la estimación de PSA en suero es su baja especificidad. Con el fin de aumentar la especificidad de dicha determinación, se han propuesto otros test derivados del PSA en suero, como son la velocidad de PSA (tasa de cambio de PSA en el tiempo), densidad de PSA (relación entre PSA y volumen

prostático), niveles de PSA específicos por edad y tiempo de duplicación de PSA (Benson et al. 1992; Carter et al. 1992; Oesterling et al. 1995; Schmid 1995). Sin embargo, estas variantes no han mostrado un incremento del valor pronóstico del test clásico de PSA en suero. En los años noventa comenzaron a descubrirse isoformas de PSA que han permitido el desarrollo de índices de diagnóstico basados en dichas isoformas (Tabla 3). La mayor parte del PSA que se determina en suero (PSA total o tPSA) aparece ligada a una proteína plasmática denominada α -1-antiquimotripsina. Sin embargo, una parte del PSA se encuentra libre, no asociándose a ninguna proteína (PSA libre o fPSA). Debido a esto, es posible la inmunodetección selectiva de la proteína α -1-antiquimotripsina, pudiendo calcular de este modo el porcentaje de fPSA en suero mediante la relación $tPSA = fPSA + PSA \text{ ligado a } \alpha\text{-1-antiquimotripsina}$ (de La Taille et al. 1997). La determinación del porcentaje de fPSA ha sido aprobada por la FDA (*Food and Drug Administration*) para la detección precoz de CaP en aquellos hombres con niveles de PSA (tPSA) entre 4 y 10 ng/ml, en base a los resultados obtenidos en el ensayo multicéntrico de Partin et al. (Partin et al. 1998). Por otro lado, el fPSA está compuesto a su vez por tres isoformas: i) proPSA, ii) fPSA intacto y iii) PSA benigno. En el estudio de estas isoformas no se ha detectado un incremento del valor predictivo con respecto al % fPSA, salvo por la isoforma [-2]proPSA, que sí parece tener un papel relevante en la detección precoz del CaP, así como en la identificación de las formas agresivas del tumor (Lazzeri et al. 2013; Heidegger et al. 2014). Tomando como referencia la determinación de PSA y sus derivados, recientemente se han desarrollado varios test que tienen por finalidad mejorar la especificidad del diagnóstico de CaP mediante los niveles de PSA en suero. Estos índices, junto con otros test empleados en la práctica clínica actual del CaP, vienen recogidos en la revisión de Gaudreau et al. (Gaudreau et al. 2016). El *prostate health index* (PHI), desarrollado por Beckman Coulter Inc., consiste en una fórmula matemática que incorpora tres isoformas de PSA, $([-2]proPSA/fPSA) \times PSA^{1/2}$. Esta fórmula se emplea para distinguir características de malignidad del CaP en pacientes mayores de 50 años, con niveles de PSA entre 4 y 10 ng/ml y con tacto rectal normal (TR). En el estudio de Lazzeri et al., el PHI aparece como un factor capaz de discriminar entre CaP de inflamación prostática crónica e hiperplasia benigna de próstata en pacientes con PSA en niveles de entre 4 y 10 ng/ml y TR normal (Lazzeri et al. 2013). Por otro lado, el conocido como *4KScore*, desarrollado por OPKO Lab Inc., representa la combinación de cuatro proteínas calicreínas: i) tPSA, ii) fPSA, iii) PSA intacto y iv) péptido relacionado con calicreína 2 (hK2). Este índice

también incorpora información clínica, como la edad y el historial de biopsias negativo precedente, y sirve para determinar la probabilidad de presencia de CaP en una biopsia determinada. En el estudio retrospectivo de Vickers et al. se mostró que la medida de *4KScore* aumenta la exactitud predictiva para el diagnóstico clínico del CaP comparado con el nivel de tPSA y la edad (Vickers et al. 2011).

Tabla 3. Biomarcadores derivados de la determinación de PSA en suero. Modificada de Gaudreau et al. 2016

Biomarcador	Descripción	Aplicaciones	Muestra
Antígeno prostático específico (PSA)	Antígeno prostático específico	Diagnóstico del CaP	Suero
PSA libre (fPSA)	PSA no enlazado a proteínas del plasma	Seguimiento del CaP	Suero
[-2]proPSA	Cadena sencilla de PSA no procesada completamente	Diagnóstico del CaP	Suero
<i>Prostate Health Index</i> (PHI)	Fórmula matemática de tres biomarcadores: ([-2]proPSA/fPSA) x PSA ^{1/2}	Distinguir condiciones prostáticas de malignidad en pacientes ≥ 50 años con PSA entre 4 – 10 ng/ml y TR normal	Suero
<i>4K Score</i>	Inmunoensayo de cuatro calicreínas: tPSA, fPSA, PSA intacto y péptido relacionado con calicreínas 2 (hK2)	Estimación de la probabilidad de presentar CaP agresivo en la primera biopsia y sucesivas	Plasma sanguíneo

Abreviaturas: CaP, cáncer de próstata; TR, tacto rectal.

1.3.4.2.- Otros biomarcadores proteicos

El CaP, al igual que la mayoría de tipos tumorales, presenta una gran heterogeneidad celular que dificulta en muchos casos la elección del tratamiento más adecuado para cada caso en particular. En el contexto de un tumor de próstata, encontramos subpoblaciones de células resistentes a cualquier intervención terapéutica, posibilitando la aparición de células potencialmente metastásicas (Rycaj et al. 2016). De hecho, se ha determinado que existen notables diferencias en el genotipo y fenotipo de células metastásicas de CaP, no solo entre diferentes pacientes, sino incluso entre células del mismo paciente (Shah et al. 2004). Por lo tanto, es imprescindible ampliar el conocimiento sobre la biología del CaP, de modo que el tratamiento pueda adecuarse de forma más específica a cada paciente. En este sentido, los biomarcadores han sido objeto de estudio durante las últimas décadas, con el fin de encontrar herramientas que

puedan ayudar a los médicos a prescribir un tratamiento cada vez más personalizado (Moschini et al. 2016).

Los biomarcadores clásicamente estudiados han sido las proteínas, en especial aquellas que participan en los mecanismos celulares con un claro papel en los procesos de tumorigénesis, como son la proliferación celular, la apoptosis, la supervivencia celular, la motilidad y la reparación de daños en el ADN. Los estudios dirigidos a evaluar el valor pronóstico de estas proteínas se basan en el análisis de la expresión proteica en el propio tejido celular. Las proteínas que intervienen en el control y regulación de la proliferación celular han sido las primeras en asociarse con los procesos de carcinogénesis. La pérdida de control de la proliferación celular es el primer paso de una compleja serie de cambios celulares que desembocan en la aparición del cáncer. La proteína Ki-67, la cual se expresa en aquellas células en proceso de expansión, es marcador no solo de la presencia tumoral, sino que también lo es de tumores metastásicos en CaP (el índice de Ki-67 es mayor en tumores que en hiperplasia benigna de próstata y a su vez es mayor en aquellos tumores con metástasis) (Verma et al. 2015). Esta proteína es reconocida como un importante marcador de metástasis a distancia y de muerte relacionada con CaP (Khor et al. 2009; Fisher et al. 2013). En relación a la proliferación tumoral, también cabe destacar el factor de crecimiento dependiente de insulina (*insuline-like growth factor*, IGF) en sus isoformas I y II, así como su receptor de membrana (IGF-1R). Algunos de los estudios realizados hasta la fecha sugieren que niveles elevados de estas proteínas están asociados con mayor riesgo de padecer CaP (Pollak et al. 1998; Biernacka et al. 2012). Es de destacar, que la relevancia de IGF en el desarrollo de CaP radica en su capacidad de activar la traslocación nuclear del receptor de andrógeno en ausencia de hormona androgénica (Culig et al. 1995). El factor de crecimiento transformante- β (*transforming growth factor- β* , TGF- β), interviene en varios procesos celulares tales como la formación de órganos o la inmunomodulación, así como en la progresión del cáncer. En condiciones normales, TGF- β participa en el mantenimiento de la homeostasis celular; sin embargo, cuando presenta desregulación contribuye a la formación tumoral (Barrack 1997). En CaP, altos niveles de TGF- β está relacionado con peor *Gleason score*, metástasis ósea, angiogénesis y pronóstico clínico desfavorable (Wikstrom et al. 1998; Elliott and Blobel 2005). El receptor de membrana del factor de crecimiento epidérmico humano (*human epidermal growth factor receptor type 2*, HER-2), es otra de las proteínas que

intervienen en la proliferación celular. Esta proteína ha sido muy estudiada en algunos tipos tumorales y hoy representa un importante marcador biológico para la clasificación del cáncer de mama, así como una significativa diana terapéutica para el tratamiento de diversos tipos tumorales (Yarden 2001; Sacco and Worden 2016). Los resultados sobre el análisis de los niveles de HER-2 en CaP no son tan esclarecedores. Sin embargo, existen estudios que relacionan este marcador con la progresión tumoral hacia una enfermedad más agresiva (Di Lorenzo et al. 2004). El factor de crecimiento del endotelio vascular (*vascular endothelial growth factor*, VEGF) es una proteína implicada en los procesos de angiogénesis, tanto normales como patológicos y por lo tanto, tiene un papel relevante en el desarrollo tumoral. En CaP, así como en otros tumores sólidos, VEGF presenta una expresión superior a la de células normales (Wegiel et al. 2005).

En el desarrollo de células cancerígenas, el ADN sufre múltiples alteraciones y mutaciones que confieren a las células una capacidad descontrolada de proliferación. En condiciones normales, las células disponen de mecanismos que les permiten reparar estas lesiones del ADN. Algunas de las proteínas que participan en estos procesos de reparación han sido propuestas como posibles factores biológicos para la predicción del resultado clínico en CaP. La proteína p53 es una proteína reguladora, clave en varios procesos celulares, la cual tiene por objetivo mantener la integridad celular, de modo que si se produce un daño en el ADN, p53 interviene en la activación de los mecanismos de reparación del ADN. Si la lesión perdura, p53 favorece la apoptosis celular, evitando la duplicación de células aberrantes. La proteína p53 fue una de las primeras cuya mutación se asoció al desarrollo del cáncer, siendo así uno de los primeros marcadores biológicos conocidos. No en vano, Levine et al. determinaron que p53 aparece mutada en el 60% de los tumores humanos (Levine et al. 1991). En CaP no es un marcador muy estudiado, sin embargo algunos resultados sugieren que p53 se expresa por encima de lo normal en los tumores de próstata comparado con su expresión en las enfermedades benignas, en las que dicha expresión no es significativamente diferente a situaciones de normalidad (Verma et al. 2015).

Las proteínas Ku70 y Ku80 forman un complejo que interviene en el mecanismo de reparación no homólogo de las roturas de la doble cadena de ADN. Este complejo reconoce los extremos libres que resultan de la rotura y se une a ellos para estimular el inicio de la reparación mediante la unión de otras enzimas en la zona lesionada. En CaP, la expresión de Ku70 mejora el valor predictivo del PSA al combinarla con el GS, de

manera que en los pacientes con $GS \leq 7$ o niveles bajos de Ku70 tratados con radioterapia y privación androgénica no se observó recaída bioquímica (aumento de los niveles de PSA) (Hasegawa et al. 2016).

La proteína *Ataxia Telangiectasia Mutated* (ATM) es otra proteína reguladora que mantiene la integridad celular actuando de “llave” de distintos procesos celulares tales como arresto del ciclo celular, apoptosis y reparación del ADN. Es una proteína ampliamente estudiada en su papel de marcador biológico de distintos tipos de carcinomas; sin embargo, en CaP no existen muchos estudios en los que se evalúe dicho papel. En el estudio de Angele et al. se determinó que los niveles de ATM en las muestras de CaP eran significativamente más elevados que los de tejido normal (Angele et al. 2004).

Otra proteína encargada de la regulación de la integridad celular es la proteína Ligasa 4 (LIG4), la cual se encarga de regular la replicación, la recombinación y la reparación del ADN. En CaP, los niveles de expresión de esta proteína son significativamente mayores con respecto al epitelio de próstata normal (Grupp et al. 2015); también en este estudio se observa una relación significativa entre la expresión de LIG4 y recaída bioquímica temprana.

Un proceso celular clave en la malignización celular es la pérdida de control sobre la apoptosis. Este mecanismo de muerte celular programada es un mecanismo de regulación sobre aquellas células que bien por senescencia, bien por aberraciones en su material genético, o por cualquier otro motivo, se convierten en células inviables incapaces de cumplir su función. En las células cancerígenas, el equilibrio que regula la apoptosis se rompe, de modo que células con multitud de fallos en su ADN son capaces de sobrevivir y proliferar sin ningún control. En el proceso de regulación de la apoptosis hay dos proteínas clave, la proteína *B-cell lymphoma 2* (Bcl-2) y sus proteínas relacionadas (una familia de proteínas de hasta 25 miembros) y la proteína *Bcl-2 associated X protein* (BAX). Estas dos proteínas tienen funciones antagonistas en la regulación de la apoptosis y mantienen el equilibrio mediante sus niveles de expresión. Bcl-2 es una proteína antiapoptótica (Korsmeyer 1992) que se expresa en aquellas células que deben sobrevivir, como células precursoras o células del sistema nervioso, descendiendo sus niveles de expresión en las células maduras. Por otro parte, BAX es una proteína proapoptótica, de modo que cuando los niveles de BAX aumentan en el

citoplasma, se induce la muerte celular (Oltvai et al. 1993). Tal y como explica Oltvai et al., estas proteínas tienen la capacidad de formar homodímeros entre sí y al mismo tiempo unirse entre ellas formando heterodímeros Bcl-2/BAX. Cuando los niveles de Bcl-2 son elevados, las moléculas de BAX son “secuestradas” formando heterodímeros, hasta que por el exceso de Bcl-2 ésta comienza a unirse en homodímeros Bcl-2/Bcl-2, los cuales inhiben la apoptosis. En determinadas circunstancias, los niveles de BAX se ven incrementados (mediante la acción de p53 entre otras), de manera que el exceso de moléculas de BAX comienza a formar homodímeros BAX/BAX, los cuales inducen la apoptosis. La expresión de estas proteínas se ha relacionado con la aparición de neoplasia (Bagg and Cossman 1992; Hickman 2002). En CaP, la expresión de estas proteínas se ha asociado significativamente con un incremento de la recaída en pacientes tratados con radioterapia externa como tratamiento exclusivo (Pollack et al. 2003). El índice obtenido entre las expresiones de las dos proteínas, Bcl-2/BAX, se relaciona de forma significativa con los tumores que no responden al tratamiento radioterápico (Mackey et al. 1998).

Todos los marcadores anteriores y algunos más han sido estudiados para establecer su valor como factores predictivos, o bien de pronóstico, del riesgo de padecer CaP, así como de los resultados clínicos o de la respuesta a un determinado tratamiento de los pacientes. En el meta-análisis de Zhao et al. se llegan a recoger 92 proteínas diferentes evaluadas como posibles biomarcadores de pronóstico en diversos estudios (Zhao et al. 2014).

1.3.4.3.- Plataformas de diagnóstico

En los últimos años, el avance en las técnicas de proteómica y genómica ha acelerado la comprensión de la biología del CaP, además de hacer más accesible el uso de estas tecnologías para el diagnóstico de rutina en la práctica clínica. En este contexto, varias compañías han desarrollado plataformas o paneles de diagnóstico y pronóstico, las cuales incluyen el análisis de una serie de proteínas o bien, de mutaciones o variaciones genéticas que previamente han sido relacionadas con los procesos de génesis tumoral del CaP (Cuzick et al. 2011). En la tabla 4, publicada por Gaudreau et al. (Gaudreau et al. 2016), vienen recogidas las plataformas de diagnóstico más relevantes.

Tabla 4. Plataformas de diagnóstico del CaP. Modificada de Gaudreau et al. 2016

Biomarcador	Descripción	Aplicaciones	Muestra
<i>Progensa</i> , medición de <i>prostate cancer antigen 3</i> (PCA3)	Test de amplificación de ácidos nucleicos de la concentración de PCA3 y ARN de PSA en orina	Decisión de repetir una biopsia en hombres ≥ 50 años con ≥ 1 biopsias negativas	Orina
<i>Prostarix</i>	Algoritmo de regresión logística de combinación de cuatro metabolitos de la orina: sarcosina, alanina, glicina y glutamato	Decisión de primera biopsia o repetidas en pacientes con TR negativo y niveles de PSA moderados	Orina
<i>Promark</i>	Test proteómico de 8 biomarcadores	Diferenciar tumores agresivos de los que no lo son en biopsia	Tejido
<i>ConfirmMDx</i>	Test de metilación múltiple de ADN	Decisión de repetir una biopsia	Tejido
<i>Oncotype DX</i>	Medición de expresión de 17 genes relacionados con 4 rutas moleculares diferentes	Individualización del tratamiento del CaP en función de su agresividad	Tejido
<i>Prolaris</i>	Expresión de 46 genes para dar un índice de progresión del ciclo celular	Individualización del tratamiento del CaP en función de su agresividad	Tejido
<i>Decipher</i>	Test genómico de medición de 22 biomarcadores de ARN de múltiples rutas biológicas	Clasificación de riesgo de metástasis en pacientes con riesgo intermedio/alto post-cirugía	Tejido

Abreviaturas: CaP, cáncer de próstata; TR, tacto rectal.

El test *Progensa* (Hologic Inc.) consiste en un test de amplificación de ácidos nucleicos que mide la concentración de ARN del antígeno prostático 3 (*prostate cancer antigen*, PCA3) y de PSA en orina después de la realización de un tacto rectal. El fin de este test es determinar la necesidad de realizar una segunda biopsia cuando la primera ha sido negativa. Según los resultados, la medición de PCA3 mediante *Progensa* (PCA3-mRNA/PSA-mRNA) tiene una capacidad de diagnóstico de CaP considerablemente superior a la de los niveles de PSA en suero (Pepe and Aragona 2011).

El panel de diagnóstico *Prolaris*, desarrollado por Myriad Genetics Inc., consiste en un test de medición de la expresión de ARN de 31 genes que participan en la progresión del ciclo celular y 15 genes de referencia en muestras de tumor primario. Este test clasifica a los pacientes con CaP localizado en función de la agresividad de su enfermedad, junto con otros parámetros clínicos como el GS o niveles de PSA (Crawford et al. 2014). Este algoritmo fue primeramente elaborado en 2011 (Cuzick et al. 2011), estableciéndose como el factor predictivo de recurrencia más significativo junto con los niveles de PSA. Desde esta fecha se han realizado varios estudios de

validación (Cooperberg et al. 2013; Freedland et al. 2013; Bishoff et al. 2014; Cuzick et al. 2015).

El test *OncotypeDX Genomic Prostate Score* (GPS) (Genomic Health Inc.) incluye 12 genes relacionados con la progresión del CaP (metabolismo de andrógenos, proliferación, organización celular y respuesta del estroma) y 5 genes de referencia (Knezevic et al. 2013). En el estudio de validación, este test permitía predecir el riesgo de alto grado y/o confinamiento tumoral en el órgano en biopsias de tumores extraídos mediante prostatectomía radical (Klein et al. 2014). Recientemente, el test GPS se ha relacionado con patología adversa en pacientes de riesgo muy bajo, bajo o intermedio después de ser tratados con prostatectomía radical (Cullen et al. 2015).

El test *Decipher PC*, diseñado por GenomeDX Biosciences, mide 22 biomarcadores de ARN no codificante con el fin de clasificar en categorías según el riesgo de metástasis a los pacientes con riesgo intermedio/alto tras recibir prostatectomía radical, (Badani et al. 2013). También el test *Decipher* se ha mostrado como mejor factor predictivo de metástasis clínica temprana comparado con otros marcadores genéticos, plataformas multi-gen y el resto de variables clínico-patológicas (Erho et al. 2013).

La compañía MDxHealth ha desarrollado la plataforma *ConfirmMDx*, la cual determina ciertos marcadores epigenéticos de metilación de algunos genes relacionados con la proliferación celular. El objetivo de este test es diferenciar entre los verdaderos resultados negativos en biopsia de próstata de aquellos que pueden ocultar la presencia de células cancerígenas (Trock et al. 2012). Con respecto a este test, se han llevado a cabo algunos estudios de validación; en el estudio DOCUMENT, este test se asoció de forma independiente con la detección del CaP en una segunda biopsia tomada a los 13 meses (de media) de haber obtenido un resultado negativo inicial, mostrando un valor predictivo de resultado negativo del 88% (Partin et al. 2014).

El test *ProMark* (Metamark Genetics Inc.) es una determinación proteómica compuesta de 8 biomarcadores medidos mediante un mecanismo cuantitativo automatizado de inmunofluorescencia que se realiza en los casos de biopsia negativa (Shipitsin et al. 2014). En los estudios de validación, el test *ProMark* proporcionó información independiente acerca del sistema de estratificación del riesgo (Blume-Jensen et al. 2015).

Por último, Metabolon Inc. ha desarrollado el índice *Prostarix Risk Score* con la finalidad de facilitar la decisión de realizar una biopsia (inicial o secundaria) en pacientes con tacto rectal negativo y niveles medios de PSA. Este test permite distinguir entre enfermedades benignas de próstata, CaP clínicamente localizado y metástasis (Jain et al. 2012; McDunn et al. 2013), para ello mide cuatro metabolitos en la orina (sarcosina, alanina, glicerina y glutamato), los cuales son combinados para obtener un algoritmo de regresión logístico.

Estas plataformas de diagnóstico están aprobadas en su mayoría por la Asociación estadounidense FDA, así como incluidas en múltiples guías de práctica clínica de CaP. Su desarrollo se fundamenta en el hecho de que es bastante improbable que un único biomarcador sea capaz de guiar las decisiones clínicas futuras. El uso de plataformas de diagnóstico requiere de un soporte estadístico complejo que obliga a un diseño previo en ensayos clínicos o meta-análisis (Gaudreau et al. 2016).

Pacientes y Métodos

2.- Pacientes y métodos

2.1.- Pacientes

En este estudio, se incluyeron un total de 601 pacientes diagnosticados con cáncer de próstata localizado no metastásico, de origen español, con etnicidad establecida hasta la segunda generación parental. Las muestras de los pacientes presentaron la siguiente distribución geográfica (Tabla 5): 91 de Andalucía (15.14%), 51 del País Vasco (8.48%), 238 de Canarias (39.61%) y 221 de Cataluña (36.77%). La relación de Instituciones participantes en el reclutamiento de pacientes y recogida de muestras de sangre aparece descrita en la tabla 5. El estudio fue aprobado por el Comité de Investigación y Ética de cada uno de los hospitales participantes.

Tabla 5. Origen regional de los participantes en el estudio

Origen regional	n	(%)	No. de hospitales
Andalucía	91	(15.14)	2
País Vasco	51	(8.48)	1
Canarias	238	(36.60)	1
Cataluña	221	(36.77)	4
Total	601	(100)	8

Para todos los pacientes incluidos en el estudio se recogieron los siguientes datos clínicos: tamaño tumoral clínico (cT), niveles de antígeno prostático específico (PSA) en suero al diagnóstico y el *Gleason score* (D'Amico et al. 2000). El tamaño tumoral clínico se evaluó mediante tacto rectal (TR), seguido de ecografía transrectal (ET) e imagen por resonancia magnética (RM). Los niveles de PSA en suero se determinaron mediante quimioluminiscencia en un analizador Architect i2000 (Laboratorios Abbott, IL, EEUU). El *Gleason score* fue determinado en biopsia por un patólogo. Los individuos se clasificaron en tres grupos de riesgo de acuerdo con la clasificación D'Amico (D'Amico et al. 1998): riesgo bajo, intermedio y alto.

2.2.- Aislamiento y cuantificación de ADN

Las Instituciones participantes fueron las encargadas de realizar las extracciones de sangre de los pacientes con CaP. Todas las muestras se recogieron en tubos de hemograma de 3 ml con EDTA-3K (Venosafe, Terumo Medical Corporation, Japón) y

fueron enviadas al Hospital Universitario de Gran Canaria Dr. Negrín para su análisis genético. El ADN se aisló a partir de 380 µl de sangre mediante el sistema de purificación iPrep (Applied Biosystems, Foster City, CA, EEUU) con los kits específicos de extracción de ADN genómico a partir de sangre iPrep™ PureLink™ (Invitrogen, por Life Technologies, Carlsbad, CA, EEUU). Una vez aislado, tanto la concentración de ADN de cada muestra como su calidad, se determinaron usando un NanoDrop ND-1000 (Thermo Scientific, Wilmington, DE, EEUU).

2.3.- Criterios de selección de los SNPs

Los SNPs fueron seleccionados a partir del proyecto HapMap de datos procedentes de individuos con ascendencia europea (CEU) (disponible en www.hapmap.org). La determinación del desequilibrio de ligamiento por pares (*pairwise linkage disequilibrium*, LD) se realizó en el software Haploview versión 4.2 (<http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>) (Barrett et al. 2005). Esta estrategia se basa en los bloques de haplotipos formados por los agrupamientos de SNPs (Gabriel et al. 2002). El proyecto HapMap permite la disponibilidad de una colección de SNPs genotipados, así como el software de análisis para facilitar la construcción de nuevos haplotipos (International HapMap 2003). De esta forma, seleccionando un modesto número de SNPs en uno de los bloques es posible recuperar más del 95% de la heterocigosis presente en un locus. La herramienta Haploview permite seleccionar los SNPs etiquetados basándose en los bloques de haplotipos identificados mediante los datos de genotipos procedentes del proyecto HapMap. El valor de R^2 proporcionado por Haploview fue mayor al 95% para todos los genes seleccionados en este estudio. Por tanto, aunque el presente estudio es de tipo “gen candidato”, los SNPs incluidos en el estudio fueron propuestos por Hapmap y Haploview, contribuyendo de esta forma a disminuir el sesgo propio de una selección manual y dirigida de los polimorfismos en base a lo establecido en la bibliografía. Además, seleccionando los SNPs de esta forma, se proceden a estudiar SNPs menos conocidos y por tanto, se contribuye en mayor medida al conocimiento científico.

2.4.- Genes y SNPs

La radioterapia basa su eficacia y efectividad en la producción de lesiones en la estructura del ADN. El genotipo que presenta cada individuo condiciona la capacidad

de reparación del ADN y por lo tanto, la respuesta al daño que sobre éste produce la radiación. Como consecuencia, la variabilidad genética entre individuos determina la diferente susceptibilidad que presentan los pacientes a la radioterapia, así como la probabilidad de desarrollar efectos adversos tras el tratamiento. Por todo ello, SNPs localizados en genes de reparación del ADN (ya sea reparación homóloga o no homóloga) se han asociado tradicionalmente con la toxicidad inducida por radiación en varios tipos tumorales. Por esta razón, se estudiaron un total de 10 SNPs en 6 genes diferentes que codifican para proteínas involucradas en la reparación homóloga de ADN: *x-ray repair cross-complementing protein 1 (XRCC1)* (Langsenlehner et al. 2011; Mangoni et al. 2011), *excision repair cross-complementing rodent repair deficiency, complementation group 2 (ERCC2)* (Zhang et al. 2010), *excision repair cross-complementing rodent repair deficiency, complementation group 1 (ERCC1)* (Biggerstaff et al. 1993), ligasa IV (*LIG4*) (Damaraju et al. 2006), ataxia telangiectasia mutada (*ATM*) (Tanteles et al. 2012), y proteína tumoral p53 (*TP53*) (Chang-Claude et al. 2009). Además, se estudiaron un total de 22 SNPs localizados en 4 genes relacionados con la reparación no homóloga de ADN: *X-ray repair complementing defective repair in Chinese hamster cell 6 (XRCC6)* (Suga et al. 2008), *X-ray repair complementing defective repair in Chinese hamster cell 5 (XRCC5)*, *major vault protein (MVP)* y *poly (ADP-ribose) sintetase 1 (PARP1)*. La descripción de todos estos polimorfismos se recoge en las tablas 6a y 6b, respectivamente.

Dado que el CaP es un tumor hormono-dependiente, donde los andrógenos ejercen un papel primordial, es lógico pensar que polimorfismos en genes relacionados con el metabolismo de los andrógenos pueden tener algún papel en la respuesta a los tratamientos o en el comportamiento biológico de este tipo tumoral. En este sentido, se estudiaron un total de 32 SNPs localizados en tres genes diferentes: citocromo P450 17A1 o esteroide 17-alfa monooxigenasa (*CYP17A1*) (Hamada et al. 2007), 5-alfa-esteroide reductasa, polipéptido alfa 1 (*SRD5A1*) y polipéptido alfa 2 (*SRD5A2*) (Aumuller et al. 1996). Las enzimas codificadas por estos genes forman parte de la ruta metabólica de producción de testosterona (T) (en el caso de *CYP17*) y de su conversión a dihidrotestosterona (DHT). Las variaciones genéticas localizadas en estos genes condicionan la respuesta prostática a dichos andrógenos (Lindstrom et al. 2006). La descripción de todos estos polimorfismos se recoge en la tabla 6c.

Tomadas en conjunto, se estudiaron un total de 64 variaciones genéticas.

Tabla 6a. Descripción de los SNPs en genes de reparación homóloga del ADN incluidos en el estudio y analizados con OpenArray

Gen	Símbolo	ID Assay	ID SNP	Alelos	Chr	Posición
<i>X-ray repair complementing defective repair in Chinese hamster cells 1</i>						
	<i>XRCC1</i>	C___622564_10	rs25487	C/T	19q13	44055726
		C___622570_10	rs25489	C/T	19q13	44056412
		C__11463404_10	rs1799782	A/G	19q13	44057574
<i>Excision repair cross-complementing rodent repair deficiency, complementation group 2</i>						
	<i>ERCC2/XPD</i>	C___3145033_10	rs13181	G/T	19q13	45854919
<i>Excision repair cross-complementing rodent repair deficiency, complementation group 1</i>						
	<i>ERCC1</i>	C___2532959_1_	rs11615	A/G	19q13	45923653
Ligasa IV						
	<i>LIG4</i>	C__11427969_20	rs1805388	A/G	13q33	108863591
		C__11427968_10	rs1805386	A/G	13q33	108861913
Ataxia telangiectasia mutada						
	<i>ATM</i>	C__33307908_10	rs17503908	G/T	11q22	108215397
		C__45273750_10	rs1800057	C/G	11q22	108143456
Proteína tumoral P53						
	<i>P53</i>	C___2403545_10	rs1042522	C/G	17p13	7579472

Abreviaturas: Chr, cromosoma; C, citosina; T, timina; A, adenina; G, guanina. Todos las sondas están comercialmente disponibles en Applied Biosystems (ver referencia en ID Assay).

Tabla 6b. Descripción de los SNPs en genes de reparación no homóloga del ADN incluidos en el estudio y analizados con OpenArray

Gen	Símbolo	ID Assay	ID SNP	Alelos	Chr	Posición
<i>X-ray repair complementing defective repair in Chinese hamster cells 6</i>						
	<i>XRCC6</i>	C__30443011_20	rs5751131	A/G	22	41629053
		C__15872242_20	rs2267437	C/G	22	41620695
		C__128361_10	rs11912946	C/T	22	41658157
	AHQJH31		rs7291732	A/C	22	41642253
	AHRSF99		rs881092	A/C	22	41653206
<i>X-ray repair complementing defective repair in Chinese hamster cells 5</i>						
	<i>XRCC5</i>	C__32572528_10	rs16855458	A/C	2	216146098
		C__3231053_10	rs9288516	A/T	2	216188541
		C__8838367_1_	rs1051677	C/T	2	216205525
		C__8838368_1_	rs1051685	A/G	2	216205653
<i>Poly (ADP-ribose) polymerase 1</i>						
	<i>PARP1</i>	C__9632806_10	rs8679	A/G	1	226360853
		C__32303980_10	rs3219123	A/G	1	226367647
		C__11639250_10	rs1805410	C/T	1	226380964
		C__32303938_10	rs3219062	G/T	1	226383047
		C__11639249_1_	rs1805414	G/A	1	226385663
		C__11639199_10	rs1805404	A/G	1	226402257
		C__27478754_10	rs3219027	A/G	1	226402712
<i>Major vault protein</i>						
	<i>MVP</i>	C__2851907_10	rs9923649	G/A	16	29820817
		C__30952394_10	rs12149514	A/G	16	29820296
		C__2851919_20	rs4788186	G/A	16	29829904
		C__2851933_20	rs3815824	T/C	16	29841539
		C__25599707_10	rs35916172	G/A	16	29844761
		C__27480544_10	rs3764944	G/A	16	29844810

Abreviaturas: Chr, cromosoma; C, citosina; T, timina; A, adenina; G, guanina. Todas las sondas están comercialmente disponibles en Applied Biosystems (ver referencia en ID Assay), excepto para los SNPs rs7291732 y rs881092, que fueron específicamente diseñados para este estudio.

Tabla 6c. Descripción de los SNPs en genes relacionados con el metabolismo de andrógenos incluidos en el estudio y analizados con OpenArray

Gen	Símbolo	ID Assay	ID SNP	Alelos	Chr	Posición
<i>Steroid-5-alpha-reductase, alpha polypeptide 1</i>						
	<i>SRD5A1</i>	C__2935721_10	rs166050	A/G	5p13	6643847
		C__7569675_10	rs501999	C/T	5p13	6635601
		C__9766953_10	rs518673	A/G	5p13	6629930
		C__2935715_10	rs3822430	A/G	5p13	6651970
		C__820231_20	rs500182	G/T	5p13	6635799
		C__1238034_20	rs8192120	A/C	5p13	6632320
		C__27891058_10	rs4702378	C/T	5p13	6643874
		AHS1EGH*	rs1691053	A/G	5p13	6677165
		C__7569805_10	rs39848	C/T	5p13	6669762
		C__29424231_10	rs3797179	A/G	5p13	6666922
<i>Steroid-5-alpha-reductase, alpha polypeptide 2</i>						
	<i>SRD5A2</i>	C__2791128_10	rs2208532	A/G	2p13	31788989
		C__345976_10	rs12470143	C/T	2p13	31763558
		C__15876107_20	rs2281546	G/T	2p13	31757024
		C__2791143_10	rs3754838	T/C	2p13	31808300
		C__30116919_10	rs4952222	A/C	2p13	31799863
		C__2791134_10	rs7562326	C/T	2p13	31798765
		C__15755333_10	rs2300702	C/G	2p13	31788018
		AHUACMP*	rs4952197	A/G	2p13	31767131
		C__2362605_10	rs676033	C/T	2p13	31808970
		C__2362601_10	rs523349	C/G	2p13	31805706
		C__2486459_10	rs9332975	C/T	2p13	31750417
		C__11160708_10	rs7594951	C/T	2p13	31791793
<i>Cytochrome P450 17A1</i>						
	<i>CYP17A1</i>	C__11201601_10	rs3781287	G/T	10p10	104595420
		C__11201597_10	rs1004467	A/G	10p10	104594507
		C__2852784_30	rs743572	A/G	10p10	104597152
		C__3284585_10	rs10883782	A/G	10p10	104583932
		C__3284587_10	rs619824	A/C	10p10	104581288
		C__15807798_10	rs2486758	C/T	10p10	104597480
		C__25597854_10	rs17115100	G/T	10p10	104591393
		C__11201596_10	rs4919686	A/C	10p10	104592249
		C__3284577_10	rs10786712	C/T	10p10	104596396
		C__12119916_1_	rs6163	A/C	10p10	104596924

Abreviaturas: Chr, cromosoma; C, citosina; T, timina; A, adenina; G, guanina. Todos las sondas están comercialmente disponibles en Applied Biosystems (ver referencia en ID Assay), excepto para los SNPs rs1691053 y rs4952197, que fueron específicamente diseñados para este estudio.

2.5.- Genotipado

El genotipado de SNPs se realizó con un termociclador Biotrove OpenArray[®] (Applied Biosystems, Foster City, CA, EEUU). Las muestras de ADN utilizadas para el genotipado con OpenArray (OA) se encontraban dentro del rango de calidad de 1.5 – 2.2 para los ratios de absorbancia A260/280 y A260/230. Previamente al proceso de genotipado, todas las muestras se diluyeron a una concentración de ADN de 50 ng/μl tal y como especifican las recomendaciones del fabricante. Así mismo, en cada array se incorporó un control libre de ADN (con la denominación de *non-template control*, NTC) consistente en agua doblemente destilada y libre de ADNasas. Una vez mezclado el ADN con la *master mix*, se transfirieron las muestras a los arrays mediante un cargador automático. Posteriormente y de acuerdo con las recomendaciones del fabricante, estos se introdujeron en el casete con líquido de inmersión y fueron sellados con pegamento para evitar la evaporación de la muestra. Los ensayos múltiples de TaqMan se llevaron a cabo en un sistema Dual Flat Block (384-well) GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, EEUU) con el siguiente ciclo de PCR: paso inicial a 93°C durante 10 minutos, seguido de 50 ciclos de 45 segundos a 95°C, 13 segundos a 94°C y 2 minutos y 14 segundos a 53°C; para terminar, un paso final durante 2 minutos a 25°C y el proceso de *holding* a 4°C.

Los resultados de fluorescencia se midieron utilizando el software OpenArray[®] SNP Genotyping Analysis versión 1.0.5 (Applied Biosystems, Foster City, CA, EEUU). El análisis del genotipado se realizó con el software TaqMan Genotyper versión 1.0.1 (disponible en: <http://www.lifetechnologies.com/es/en/home/global/forms/taqman-genotyper-software-download-reg.html>). Para este análisis se empleó el método *autocalling* como método de determinación. Como valor de calidad de los datos de genotipado se seleccionó un umbral de 0.95 de determinación. Los análisis de genotipado se realizaron de forma separada para cada población y de forma automática sin ninguna asignación manual de genotipo (Figura 9).

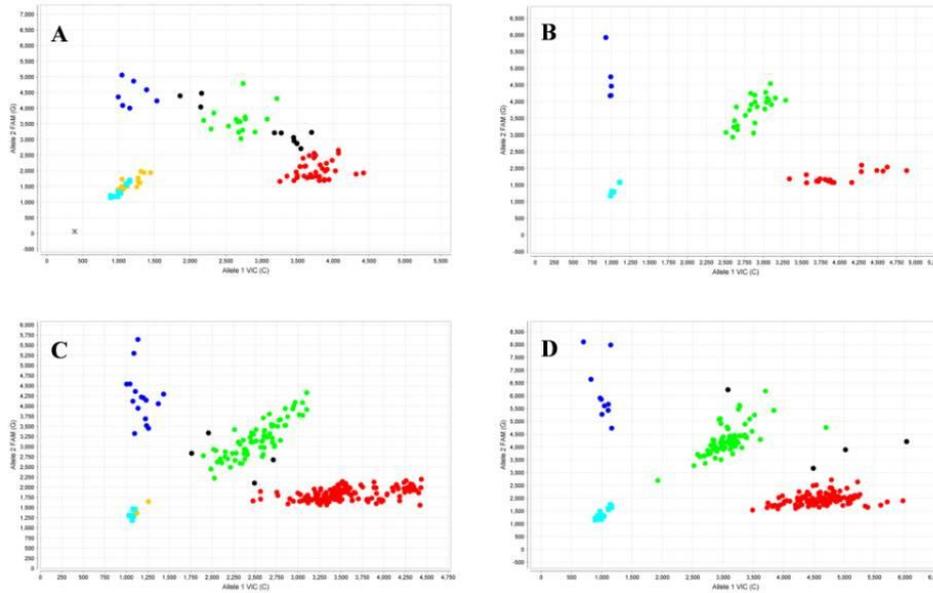


Figura 9, diagramas de dispersión del genotipado del SNP rs1042522 en las poblaciones de Andalucía (A), País Vasco (B), Canarias (C) y Cataluña (D), mediante el ciclador Biotrove OpenArray® NT. Cada gráfica representa un diagrama de dispersión de uno de los alelos (sonda FAM) y el otro alelo (sonda VIC). Las muestras homocigotas aparecen en color azul o rojo; los heterocigotos emiten fluorescencia de ambas sondas y aparecen representados en verde. Las muestras NTC aparecen en forma de cuadrados azul claro y representan la fluorescencia base de aquellas muestras que no poseen ADN para amplificar. Las muestras “no determinadas” aparecen como puntos negros y las “no amplificadas” como puntos naranjas. Estos diagramas de dispersión se obtuvieron mediante el software TaqMan Genotyper software versión 1.0.1.

2.6.- Análisis estadísticos

Las frecuencias genotípicas y alélicas se determinaron utilizando el entorno de web SNPator (*SNP Analysis to Results*, del Centro Nacional de Genotipado Español y del Instituto Nacional de Bioinformática) (Morcillo-Suarez et al. 2008). El exceso relativo de heterocigosis fue determinado con el fin de comprobar las compatibilidades de las frecuencias genotípicas con el equilibrio de Hardy-Weinberg (*Hardy-Weinberg equilibrium*, HWE). De este modo, los valores de p para el test *standard exact HWE lack of fit* se calcularon utilizando SNPator. Las comparaciones de las frecuencias genotípicas y alélicas entre las poblaciones, así como los análisis de haplotipos, también se realizaron con ese entorno de web.

El análisis de componentes principales (*principal component analysis*, PCA) se realizó con las aplicaciones Clases y Métodos de SnpMatrix y XSnMatrix (Clayton 2012) implementados como paquete de R y disponible en Bioconductor (versión 2.11;

<http://bioconductor.org>). Este análisis consiste en transformar un grupo de variables originales en otro grupo de variables -los componentes principales- obtenidos como una combinación lineal de los primeros. Las nuevas variables contienen toda la información, pero la mayoría de los componentes principales presentan una variabilidad tan pequeña que pueden ser ignoradas. De esta manera, pocos componentes (generalmente 3 o menos) pueden representar y explicar razonadamente bien el grupo de objetos de la muestra sin perder información significativa. Los PCA reducen la complejidad de los datos y permiten la representación gráfica de las variables.

Se llevaron a cabo la construcción de agrupamientos jerárquicos no supervisados (Eisen et al. 1998) de los SNPs de cada población, para lo cual se usó el programa MultiExperiment Viewer versión 4.9 (disponible en www.tigr.org). Los clústeres o agrupamientos se realizaron utilizando la distancia de correlación Euclidiana. Para obtener los clústeres de forma satisfactoria, los homocigotos salvajes se codificaron como -1, los heterocigotos como 0 y los homocigotos mutados como 1.

Los análisis estadísticos adicionales se llevaron a cabo con el programa PAWS Statistics 15 (IBM Corporation, Armonk, NY, EEUU).

Hipótesis

3.- Hipótesis

1. La homogeneidad genotípica de una población es un factor clave a la hora de hacer estudios de asociación genética y de respuesta a tratamientos farmacológicos. Comprobar este hecho daría garantías de fiabilidad en los subsecuentes estudios realizados.

2. La presencia de determinadas variaciones genéticas relacionadas con los procesos de reparación del daño al ADN (tanto en los mecanismos de reparación homóloga como no homóloga) podrían condicionar el comportamiento biológico del cáncer de próstata en términos de agresividad y progresión.

3. La presencia de determinadas variaciones genéticas relacionadas con genes involucrados en el metabolismo de los andrógenos podrían condicionar el comportamiento biológico del cáncer de próstata en términos de agresividad y progresión.

Objetivos

4.- Objetivos

1. Comprobar la homogeneidad de la distribución alélica y genotípica de genes relacionados con la reparación del daño al ADN y genes involucrados en el metabolismo de andrógenos en una serie de pacientes españoles con cáncer de próstata, para conocer así el comportamiento genotípico de los polimorfismos incluidos en el estudio.

2. Valorar si la presencia de polimorfismos de nucleótido simple en genes relacionados con la reparación homóloga del daño al ADN condicionan el comportamiento clínico del cáncer de próstata.

3. Valorar si la presencia de polimorfismos de nucleótido simple en genes relacionados con la reparación no homóloga del daño al ADN condicionan el comportamiento clínico del cáncer de próstata.

4. Valorar si la presencia de polimorfismos de nucleótido simple en genes relacionados con el metabolismo de andrógenos condicionan el comportamiento clínico del cáncer de próstata.

Capítulo I

5.- Capítulo I: La heterogeneidad en la distribución de polimorfismos entre poblaciones del mismo origen étnico como factor de confusión en estudios de asociación genética

5.1.- Introducción

Los polimorfismos genéticos son variantes del genoma que se caracterizan por aparecer como mutaciones en algunos individuos, tener la facultad de ser transmitidos a la descendencia (son heredables) y, tras varias generaciones, adquirir cierta frecuencia dentro de la población (al menos el 1%). Los polimorfismos son la base de la evolución; algunos polimorfismos pueden estar silenciados, pero los que no lo están pueden conferir beneficios al individuo que los porta o por el contrario, pueden estar involucrados en el desarrollo de enfermedades (Guttmacher and Collins 2002). Los polimorfismos más frecuentes son los polimorfismos de nucleótido simple (*single nucleotide polymorphisms*, SNPs). La distribución genotípica de una población viene determinada por su origen étnico y es diferente a la de poblaciones con un origen étnico distinto. Sin embargo, las diferencias de distribución genotípica observadas dentro de las poblaciones de una misma etnia sugieren que este no es un factor suficiente para asegurar la homogeneidad de una muestra. En este sentido, se han descrito varios ejes de estratificación significativos en la distribución genotípica de la población europea, el principal sigue una tendencia norte-sur-este, pero también se ha determinado una tendencia a lo largo del eje este-oeste (Bauchet et al. 2007). En el caso de España, aunque la población que habita en la península ibérica muestra una sustancial homogeneidad genética (Laayouni et al. 2010), existen resultados que sugieren que la influencia de la población del noroeste de África en algunas poblaciones españolas produce ciertas diferencias que pueden incrementar el riesgo de falsos positivos en los estudios de epidemiología genética (Pino-Yanes et al. 2011).

La radioterapia (RT) es un tratamiento eficiente que se ofrece a pacientes con cáncer de próstata localizado como una alternativa viable a la cirugía (Bracarda et al.

2005). Aunque ambos tratamientos muestran resultados comparables en términos de supervivencia (Kupelian et al. 2004), las principales diferencias entre ambas terapias están relacionadas con los efectos adversos que producen cada una de ellas. El control tumoral mediante RT requiere el uso de dosis máximas que son liberadas manteniendo un riesgo tolerable de toxicidad en el tejido sano no tumoral circundante, siendo la toxicidad clínica el principal factor limitante de la eficacia del tratamiento radioterápico (Johansson et al. 2000). El papel de la genética en la respuesta del tejido sano a la RT es un hecho aceptado por la comunidad científica y podría contribuir a explicar la gran variación de toxicidad del tejido sano que experimentan los pacientes tratados con esta técnica, incluso cuando se administran dosis y esquemas de tratamiento similares (Bentzen and Overgaard 1994). La radiación causa la pérdida de estructura y funcionalidad de la mayoría de moléculas biológicas, incluido el ADN. La capacidad de reparación del ADN engloba una variedad de mecanismos tal y como se ha descrito anteriormente (reparación por escisión de nucleótido o base, recombinación homóloga, unión no homóloga de extremos y reparación por desapareamiento), y que condicionan en cada individuo de forma particular la respuesta del tejido tumoral y del tejido sano a la radiación (Henriquez-Hernandez et al. 2012). Debido a ello, los estudios de tipo “gen candidato” se han centrado en aquellos genes involucrados principalmente en el reconocimiento y la reparación del daño en el ADN (como los genes *ATM*, *XRCC*, *XPD*, *ERCC*, *LIG4* y *TP53*, entre otros), así como genes relacionados con la expulsión y eliminación de radicales libres (*SOD2*) o la respuesta anti-inflamatoria (*TGFBI*).

La asociación entre los SNPs y la toxicidad inducida por radiación se ha estudiado extensamente (West and Barnett 2011) y en este contexto se han formado numerosos consorcios con el fin de identificar las variaciones genéticas que se relacionan comúnmente con el desarrollo de toxicidad inducida por radiación (West et al. 2010). Aunque prometedores, los resultados hasta ahora obtenidos fallan en la fase de validación (Barnett et al. 2012) y hoy por hoy, el desarrollo de una firma de SNPs que determine la predicción de la toxicidad es todavía un objetivo lejano. Esta falta de resultados validados puede ser explicada por diferentes razones: presencia de factores de confusión, un tamaño muestral insuficiente, falta de consenso en la metodología empleada para determinar el genotipo, diferencias en la estadística o diferencias en los métodos de valoración de la toxicidad inducida por RT (West et al. 2012); de estas

razones, la heterogeneidad de la población estudiada es un factor cuyo efecto ha sido habitualmente infravalorado.

Con todas estas presunciones en mente, hemos diseñado un estudio con el objetivo de evaluar la distribución genotípica de 10 SNPs localizados en 6 genes diferentes involucrados en la reparación de ADN y que se asocian clásicamente con la toxicidad inducida por radiación. El estudio se ha realizado en una amplia serie de pacientes de cáncer de próstata españoles, tal y como se ha descrito en la sección de “Pacientes y Métodos” de esta Tesis Doctoral, para determinar la homogeneidad de la población y revelar potenciales factores de confusión que hayan podido ser subestimados en los estudios de asociación entre los SNPs y la toxicidad inducida por radiación.

5.2.- Resultados

Todas las muestras genotipadas en este estudio cumplían con los criterios de calidad expuestos en el apartado “Pacientes y Métodos” y todas ellas fueron genotipadas con el mismo lote de material y en el mismo momento. Se genotiparon un total de 601 muestras para 10 SNPs. De las 6010 posibles determinaciones, el 94.36% fueron genotipadas exitosamente. Los niveles de determinación para los SNPs considerados fueron (mediana (rango)): 79.5% (68.1 – 91.2%) para la población de Andalucía, 100% (80.4 – 100%) para la población del País Vasco, 97.7% (94.5 – 99.2%) para la población de Canarias y 97.9% (83.3 – 99.1%) para la población de Cataluña.

Las frecuencias genotípicas y alélicas obtenidas tras el análisis aparecen recogidas en la tabla 7. Se encontró que, en total, 4 SNPs de dos poblaciones diferentes presentaban desviación del HWE, indicando un exceso relativo de heterocigosis para esos SNPs: rs25487 (*XRCCI*) en sujetos de Cataluña y rs13181 (*ERCC2*), rs11615 (*ERCC1*) y rs180057 (*ATM*) en sujetos de Andalucía (Tabla 7). La distribución genotípica fue diferente entre las poblaciones de estudio para 4 de los 10 SNPs: rs25487, rs13181, rs11615 y rs1805386 (*LIG4*) (test de χ^2 , Tabla 7), lo que muestra una distribución diferencial de genotipos entre las distintas poblaciones para esos SNPs concretos.

Se realizó un análisis jerárquico no supervisado con el fin de visualizar dichas diferencias de distribución genotípica entre las cuatro poblaciones en estudio. Tal y como se muestra en la figura 10, los polimorfismos quedan distribuidos en dos clústeres

principales, cada uno de ellos con un número y una identidad de SNPs diferente, sugiriendo cierto grado de heterogeneidad entre las poblaciones. Además de lo anterior, se utilizó la herramienta de entorno de web SNPator para comparar las poblaciones de forma individualizada, una frente a otra. Las diferencias en las distribuciones genotípicas se encontraron principalmente entre la población de Andalucía y el resto de poblaciones (test de χ^2 , Tabla 8). Más aún, las mayores diferencias se encontraron entre las poblaciones de Andalucía y Cataluña, que presentaron 3 SNPs (rs25487, rs13181 y rs11615) distribuidos diferencialmente entre los pacientes de CaP de ambas poblaciones.

Tabla 7. Frecuencias alélicas y genotípicas de los polimorfismos incluidos en el estudio. Se incluyen las diferencias significativas entre las poblaciones, así como la desviación del equilibrio de Hardy-Weinberg (HWE).

	% Determinación	Genotipos			HWE	Alelos	
<i>XRCC1</i>							
rs25487		CC	CT	TT		C	T
Andalucía	0.79	0.49	0.34	0.18	ns	0.65	0.35
País Vasco	0.80	0.44	0.51	0.05	ns	0.70	0.30
Canarias	0.95	0.48	0.41	0.11	ns	0.68	0.32
Cataluña	0.83	0.36	0.55	0.09	*	0.63	0.37
<i>p value</i>				0.012			
rs25489		CC	CT	TT		C	T
Andalucía	0.91	0.81	0.19	0.00	ns	0.90	0.10
País Vasco	1.00	0.86	0.14	0.00	ns	0.93	0.07
Canarias	0.97	0.87	0.13	0.00	ns	0.93	0.07
Cataluña	0.98	0.90	0.09	0.01	ns	0.95	0.05
<i>p value</i>				0.178			
rs1799782		AA	AG	GG		A	G
Andalucía	0.85	0.00	0.09	0.91	ns	0.05	0.95
País Vasco	1.00	0.00	0.08	0.92	ns	0.04	0.96
Canarias	0.99	0.01	0.12	0.87	ns	0.07	0.93
Cataluña	0.98	0.01	0.11	0.88	ns	0.06	0.94
<i>p value</i>				0.936			
<i>ERCC2</i>							
rs13181		GG	GT	TT		G	T
Andalucía	0.74	0.19	0.15	0.66	*	0.27	0.73
País Vasco	1.00	0.06	0.37	0.57	ns	0.25	0.75
Canarias	0.98	0.11	0.45	0.44	ns	0.33	0.67
Cataluña	0.97	0.09	0.53	0.38	ns	0.35	0.65
<i>p value</i>				0.0001			
<i>ERCC1</i>							
rs11615		AA	AG	GG		A	G
Andalucía	0.70	0.58	0.20	0.22	*	0.68	0.32
País Vasco	1.00	0.43	0.43	0.14	ns	0.65	0.35
Canarias	0.98	0.43	0.41	0.16	ns	0.63	0.37
Cataluña	0.99	0.32	0.52	0.16	ns	0.58	0.42
<i>p value</i>				0.001			
<i>LIG4</i>							
rs1805388		AA	AG	GG		A	G
Andalucía	0.74	0.06	0.12	0.82	ns	0.12	0.88

País Vasco	0.98	0.04	0.38	0.58	ns	0.23	0.77
Canarias	0.99	0.03	0.25	0.72	ns	0.15	0.85
Cataluña	0.99	0.05	0.22	0.73	ns	0.16	0.84
<i>p value</i>				0.051			
rs1805386		AA	AG	GG		A	G
Andalucía	0.85	0.78	0.16	0.06	ns	0.85	0.15
País Vasco	0.98	0.84	0.16	0.00	ns	0.92	0.08
Canarias	0.96	0.73	0.25	0.02	ns	0.85	0.15
Cataluña	0.98	0.66	0.28	0.06	ns	0.80	0.20
<i>p value</i>				0.029			
ATM							
rs17503908		GG	GT	TT		G	T
Andalucía	0.81	0.03	0.08	0.89	ns	0.07	0.93
País Vasco	0.98	0.00	0.20	0.80	ns	0.10	0.90
Canarias	0.99	0.01	0.20	0.79	ns	0.10	0.90
Cataluña	0.98	0.01	0.17	0.82	ns	0.09	0.91
<i>p value</i>				0.088			
rs1800057		CC	CG	GG		C	G
Andalucía	0.80	1.00	0.00	0.00	*	1.00	0.00
País Vasco	1.00	0.94	0.06	0.00	ns	0.97	0.03
Canarias	0.97	0.95	0.04	0.01	ns	0.97	0.03
Cataluña	0.99	0.92	0.08	0.00	ns	0.96	0.04
<i>p value</i>				0.186			
TP53							
rs1042522		CC	CG	GG		C	G
Andalucía	0.68	0.63	0.26	0.11	ns	0.76	0.24
País Vasco	1.00	0.41	0.49	0.10	ns	0.66	0.34
Canarias	0.97	0.61	0.32	0.07	ns	0.77	0.23
Cataluña	0.98	0.60	0.35	0.05	ns	0.78	0.22
<i>p value</i>				0.059			

Abreviaturas: ns, no significativo. Las diferencias de distribución se llevaron a cabo mediante el test estadístico de χ^2 . Las poblaciones y SNPs desviadas del HWE están marcadas con un asterisco (*) ($p < 0.01$).

Tabla 8. Comparación de las frecuencias alélicas y genotípicas entre las poblaciones en estudio. Se incluyen solo los SNPs diferencialmente distribuidos

Poblaciones	Frecuencia alélica	Frecuencia genotípica
Can vs. And	-	rs13181
Can vs. Vasc	-	-
Can vs. Cat	-	-
And vs. Vasc	-	rs1805388
And vs. Cat	-	rs13181
		rs25487
		rs13181
		rs11615
Vasc vs. Cat	rs1805386	-

Abreviaturas: Can, Canarias; And, Andalucía; Vasc, País Vasco; Cat, Cataluña. Los análisis se realizaron incluyendo las tres variables genotípicas. Las diferencias fueron significativas cuando $p < 0.01$.

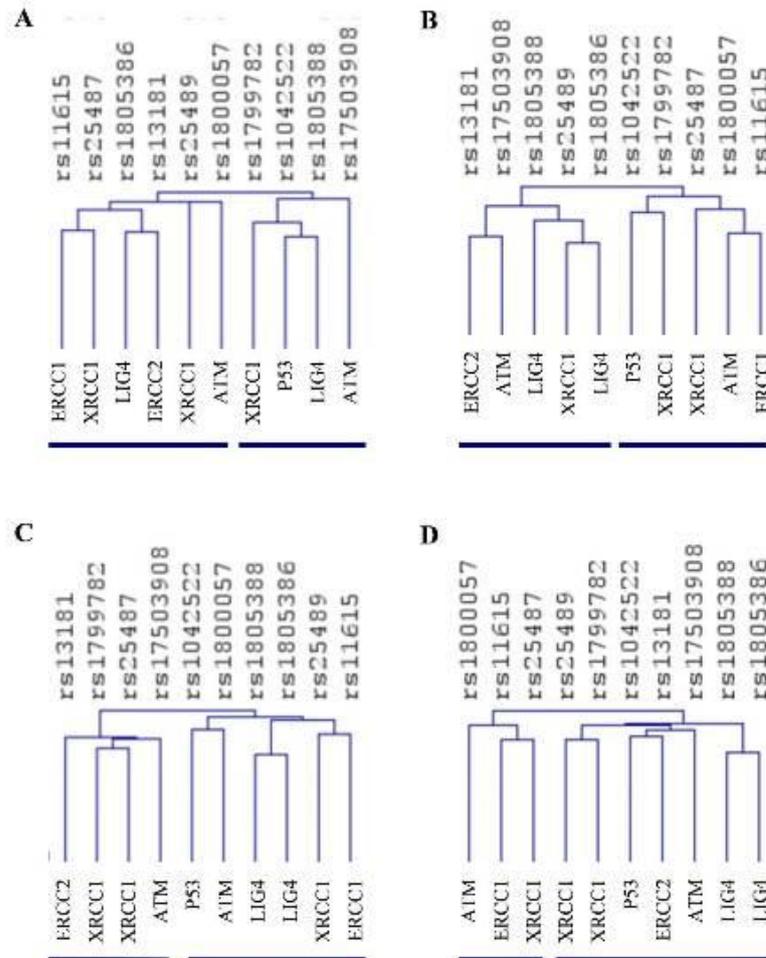


Figura 10, *non-supervised hierarchical clustering* de los distintos SNPs en pacientes de CaP de las poblaciones de Andalucía (A), País Vasco (B), Canarias (C) y Cataluña (D). Los clústeres fueron realizados utilizando la correlación de distancia Euclidiana y el ligamiento medio y fueron procesados con el programa MultiExperiment Viewer (<http://www.tigr.org>). Los dendogramas muestran los clústeres de los SNPs. A ellos se ha unido el símbolo del gen para identificar cada SNP. Las líneas azules debajo de cada panel muestran los dos clústeres principales generados.

Para identificar las diferencias globales entre las poblaciones, se llevó a cabo un análisis de componentes principales (PCA). Los componentes 1 y 2 fueron los responsables del 15.3% y del 14.3% de la diferencia entre poblaciones, respectivamente. En la representación gráfica conjunta de dichos componentes no se apreciaba discriminación de las poblaciones entre sí (Figura 11A). Sin embargo, cuando los componentes se analizaron por separado, el primero fue capaz de distinguir entre las poblaciones de Andalucía y Cataluña (Figura 11B), corroborando los resultados observados en la tabla 8 y mostrando diferencias claras en las distribuciones de genotipos entre las poblaciones analizadas.

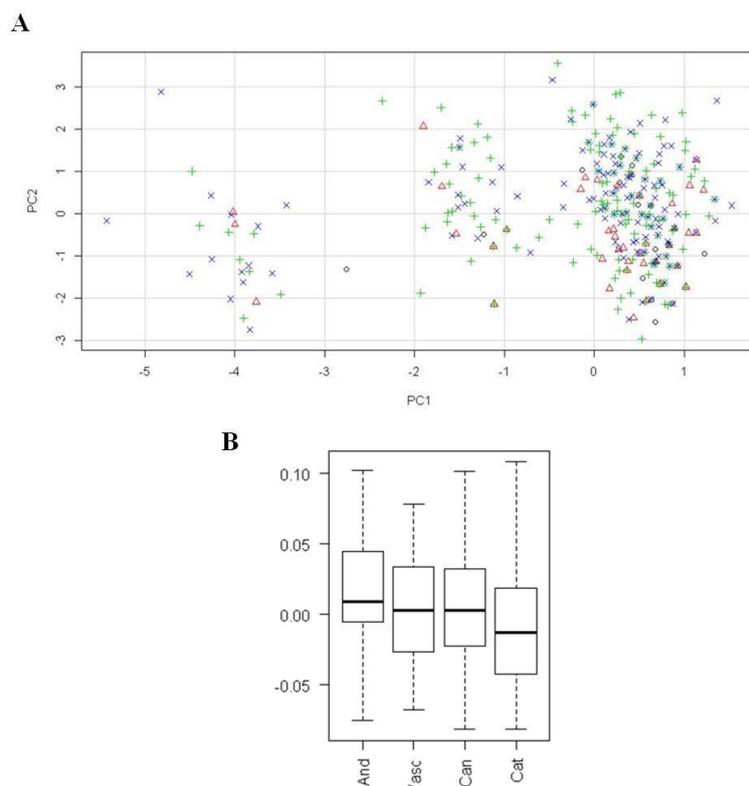


Figura 11, diagrama de representación de los dos componentes principales obtenidos del análisis de las poblaciones (A) y diagrama de cajas del componente principal 1 entre las diferentes poblaciones (B). Símbolos en el diagrama A: ° (negro), Andalucía; Δ (rojo), País Vasco; x (azul), Canarias y + (verde), Cataluña. Abreviaturas en el diagrama B: And, Andalucía; Vasc, País Vasco; Can, Canarias y Cat, Cataluña.

Por último, el análisis de haplotipos se llevó a cabo con la herramienta de entorno de web SNPator. Tal como se muestra en la tabla 9, se encontró que los haplotipos más frecuente en cada población eran diferentes para cada una de ellas. Para los SNPs del cromosoma 11 (donde está codificado el gen *ATM*), el haplotipo GG se encontraba ausente en la población de Andalucía. Para los SNPs del cromosoma 13 (donde está codificado el gen *LIG4*), los haplotipos GG y AA mostraron una distribución diferente en las distintas poblaciones. En el caso de los SNPs del cromosoma 19 (donde están codificados los genes *XRCC1*, *ERCC2* y *ERCC1*), el haplotipo CCGGG estaba presente solo en los pacientes de CaP de las poblaciones de Cataluña y Canarias, mientras que el haplotipo CCGTG solo apareció en los pacientes de Andalucía y País Vasco. El hecho de que los haplotipos más frecuentes sean los mismos en todas las poblaciones sugiere un elevado grado de similitud entre los individuos de la misma etnia, algo por otra parte esperable.

Tabla 9. Análisis de los tres haplotipos más frecuentes (%) en los cromosomas 11 (gen *ATM*), 13 (gen *LIG4*) y 19 (genes *XRCCI*, *ERCC2* y *ERCCI*) entre las poblaciones estudiadas

Haplotipo/chr	Andalucía	fr	País Vasco	fr	Canarias	fr	Cataluña	fr
Hap 1	CT	93.18	CT	89.82	CT	89.15	CT	91.27
Hap 2	CG	6.80	CG	7.24	CG	8.23	CG	4.87
Hap 3	GT	0.02	GG	2.74	GG	2.39	GG	3.86
Haplotipo/chr13								
Hap 1	AG	76.08	AG	69.81	AG	70.17	AG	64.14
Hap 2	GG	11.89	AA	22.27	AA	15.22	GG	19.60
Hap 3	AA	10.53	GG	7.11	GG	14.49	AA	16.23
Haplotipo/chr19								
Hap 1	CCGTA	32.52	CCGTA	31.19	CCGTA	27.37	CCGTA	26.11
Hap 2	TCGTA	14.86	CCGTG	14.81	TCGTA	17.03	TCGTA	14.31
Hap 3	CCGTG	11.37	TCGTA	14.29	CCGGG	12.32	CCGGG	12.37

Abreviaturas: chr, cromosoma; fr, frecuencia; Hap, haplotipo. Los haplotipos en el cromosoma 11 están formados por el locus rs1800057 y rs17503908, respectivamente; los haplotipos en el cromosoma 13 están formados por el locus rs1805386 y rs1805388, respectivamente; los haplotipos en el cromosoma 19 están formados por el locus rs25487, rs25489, rs1799782, rs13181 y rs11615, respectivamente.

5.3.- Discusión

La radiogenómica es el estudio de la variabilidad genética, principalmente de los polimorfismos de nucleótido simple (SNPs), en relación con el desarrollo de toxicidad inducida por los tratamientos radioterápicos. Estos estudios van dirigidos principalmente a la búsqueda de un test capaz de predecir qué pacientes de cáncer tienen mayor probabilidad de desarrollar efectos adversos tras la administración de radioterapia (RT) (Henriquez-Hernandez et al. 2012). La predicción de la toxicidad en tejido normal permitiría ajustar la dosis de radiación para cada paciente de forma individual, con el fin de evitar estos efectos adversos, poniendo especial interés en aquellos casos de pacientes con CaP en los que altos niveles de radiación están asociados con un mejor resultado del control bioquímico y una reducción de la metástasis a distancia (Zelevsky et al. 2011). El papel que desempeña la variabilidad genética en el desarrollo de la toxicidad inducida por radiación ya ha sido probado (Andreassen and Alsner 2009). En este aspecto, la gran variabilidad interindividual observada entre los distintos casos parece estar explicada en parte por la genética, ya que esta variabilidad llega a producirse entre pacientes con características clínicas muy similares y tratados con los mismos esquemas de tratamiento (Buchholz 1999). Existe bastante bibliografía que hace referencia al papel predictivo de algunos SNPs concretos con respecto a la toxicidad del tejido normal, sin embargo, los estudios de validación no

han corroborado dichos resultados, poniendo en cuestión la utilidad de los SNPs como herramienta predictiva de la toxicidad inducida por radiación (Barnett et al. 2012).

En una población, la asociación estadística que pudiera existir entre el genotipo de un locus determinado y la aparición de un rasgo binario concreto (como sería el caso de ausencia/presencia de toxicidad inducida por radiación) puede mostrarse de tres maneras (Clayton 2001): i) el locus puede estar relacionado con el rasgo de forma causal (la presencia de diferentes alelos conlleva diferente riesgo de presentar toxicidad), ii) el locus puede no ser la causa por sí mismo, pero podría estar suficientemente cerca de un locus que sí sea causal, estando en desequilibrio de ligamiento con él, o iii) la relación observada puede ser el efecto de una estratificación o una mezcla de la población. Esta última relación puede actuar como factor de confusión, creando asociaciones aún en ausencia de una unión causal real o bien encubrir dicha relación de causalidad. Es por ello que, con el fin de excluir falsas asociaciones, es imprescindible llevar a cabo un diseño y/o análisis experimental adecuado, teniendo en cuenta para ello que los sesgos que resultan de errores sistemáticos (como sesgos en la selección de los pacientes o en los resultados de las mediciones) persisten conforme aumenta el tamaño de la muestra. Algunas de las causas que podrían suponer la inclusión de este tipo de errores sistemáticos serían: i) la presencia de varios grupos étnicos en la población estudiada, ii) variaciones entre los grupos respecto a las diferencias alélicas del locus en estudio y iii) diferencias en la frecuencia de la enfermedad entre dichos grupos por razones que no estén relacionadas con el locus de interés. En relación con lo anterior, se sabe que la etnicidad es uno de los factores que influye en la aplicabilidad de la farmacogenética (Sai and Saito 2011).

Todas las poblaciones incluidas en este estudio son consideradas como caucásicas. Sin embargo, la historia natural de algunas de las poblaciones, como la población de Canarias o la del País Vasco, son consideradas genéticamente diferentes. Mientras que la población canaria está influenciada por la migración de África Noroccidental y la colonización europea (Rando et al. 1999), la población vasca tiene un origen muy distinto (Calafell and Bertranpetit 1994). No obstante, en una publicación reciente se genotiparon 120 SNPs de 30 individuos procedentes de 10 poblaciones diferentes de España (la población canaria no estaba incluida en este estudio), concluyendo que las poblaciones estudiadas eran genotípicamente similares (Laayouni et al. 2010). Sin embargo, ninguno de los SNPs considerados en el presente

estudio estaba incluido en el artículo reseñado. Nuestros resultados indican que la distribución genotípica de 4 SNPs fue diferente entre las poblaciones de Andalucía, País Vasco, Canarias y Cataluña. Estos resultados se compararon con la cohorte de pacientes de cáncer de próstata más grande analizada en España (Fachal et al. 2012), con un total de 698 pacientes de CaP procedentes de Galicia analizados para 14 SNPs localizados en los genes ataxia telangiectasia mutada (*ATM*), *excision repair cross-complementing rodent repair deficiency, complementation group 2* (*ERCC2*), ligasa IV (*LIG4*), *MutL protein homolog 2* (*MLH2*) y *x-ray repair cross-complementing protein 3* (*XRCC3*). Tres de los SNPs analizados también se incluyeron en este estudio: rs1805388 (*LIG4*), rs1805386 (*LIG4*) y rs1800057 (*ATM*). La distribución genotípica de los SNPs rs1805388 y rs1805386 fue significativamente diferente entre la población gallega y las poblaciones incluidas en nuestro estudio (test de χ^2 , $p = 0.001$ y $p = 0.007$, respectivamente). Esto pone de manifiesto la variabilidad genética para determinados SNPs que pueden presentar poblaciones de la misma etnia procedentes del mismo país. De acuerdo con nuestros resultados, la población de Andalucía es la que presenta mayores diferencias en la distribución genotípica de algunos SNPs, mostrando las mayores disparidades con la población de Cataluña (resultados observados con el test de χ^2 y el PCA). Las diferencias entre poblaciones también fueron evidentes en el análisis de haplotipos. Estos resultados sugieren que debe tenerse en cuenta cada SNP de forma independiente a la hora de buscar variables de confusión que puedan influir de forma importante en la interpretación de los resultados en estudios de asociación de variabilidad genética. Así, aquellos estudios diseñados para determinar la asociación entre SNPs y la toxicidad inducida por radiación, siguen habitualmente un diseño de casos y controles. En ellos se asume que las dos series de sujetos pueden emplearse para proporcionar estimaciones no sesgadas de las correspondientes distribuciones genotípicas de miembros afectados y no afectados de cualquier población subyacente (Clayton 2001). Esta suposición fundamental puede no darse en la realidad, dando lugar a hallazgos sesgados que pueden ser clasificados en dos categorías: i) sesgo de selección, causado por un muestreo inapropiado de casos y controles y ii) sesgo de información, causado por errores de medidas diferenciales entre casos y controles. Cuando una variable de confusión es detectada en un estudio, el método clásico empleado en los estudios de epidemiología consiste en la estratificación de los análisis en función de dicha variable potencial, probando las posibles asociaciones entre los factores de interés (como el genotipo) y la enfermedad en cada uno de los estratos

(como los diferentes grados de toxicidad inducida por la radiación). En lo que concierne a la presencia de sesgos derivados de la estratificación de la población en los estudios genéticos de casos y controles, estos deben ser atenuados mediante un diseño y análisis de casos y controles apropiados, la evaluación de la probabilidad de que los sesgos más habituales puedan presentarse en un estudio determinado (Wacholder et al. 2000) y, si es necesario, el empleo de métodos de corrección (Reich and Goldstein 2001).

Este estudio también presenta algunas limitaciones que deben ser tenidas en cuenta. Primero, todas las muestras procedían de pacientes con CaP, los cuales pueden presentar una frecuencia genotípica diferente a la que presentan los sujetos de la población sana. Sin embargo, en estudios diseñados para evaluar posibles asociaciones entre SNPs y toxicidad inducida por radiación, los sujetos que actúan de control no son individuos sanos, sino que son pacientes con nulo o escaso grado de toxicidad, mientras que los casos son aquellos pacientes con alto grado de toxicidad, de modo que no se comparan sujetos de ambas poblaciones, sino que todos los sujetos son enfermos de cáncer. De este modo, estos estudios reproducen el diseño estándar de los estudios de casos y controles. En segundo lugar, el número de sujetos de las diferentes poblaciones varía extensamente. Sin embargo, el hecho de que las mayores diferencias no se encuentren en la población con el menor número de pacientes (población del País Vasco, con 51 muestras) sugiere que esta limitación no es decisiva en la interpretación de los resultados. Además, si la heterogeneidad entre poblaciones es considerada un sesgo sistemático, este es independiente del tamaño poblacional. En último lugar, para hacer un análisis ciego, los datos clínicos de los pacientes (TNM, estadio clínico, grado tumoral, fallo bioquímico o *Gleason score*) no estuvieron disponibles durante este análisis. En este sentido, es posible que alguno de los polimorfismos pueda influir en las características clínicas del tumor, pudiendo plantearse además como factores de riesgo para otras características de la enfermedad (Henriquez-Hernandez et al. 2010; Henriquez-Hernandez et al. 2010).

Por otro lado, deben destacarse algunas de las ventajas que presenta este estudio: i) el número de sujetos incluidos en el estudio es suficiente como para obtener datos fidedignos sobre la distribución genotípica de estos 10 SNPs en las poblaciones de pacientes de CaP estudiadas (especialmente para Canarias y Cataluña), ii) todos los sujetos fueron hombres, evitando un posible sesgo asociado al sexo y iii) todas las determinaciones (6010 determinaciones en total) fueron realizadas con el mismo

método (OpenArray[®], Applied Biosystems), con el mismo lote de material y por el mismo investigador, minimizando los errores de origen técnico.

5.4.- Conclusiones

Las diferencias en la distribución de genotipos entre diferentes poblaciones de la misma etnia puede ser un factor de confusión importante y podría influir en la falta de validación de los resultados obtenidos en los estudios de SNPs asociados con la toxicidad inducida por radiación. Esto es de especial relevancia cuando se llevan a cabo extensos meta-análisis con sujetos procedentes de distintos países (Barnett et al. 2012). Los resultados aquí expuestos sugieren que la homogeneidad entre las personas (especialmente entre aquellos sujetos considerados controles del estudio) debería ser probada antes de proceder con cualquier otro análisis.

Capítulo II

6.- Capítulo II: Variaciones intra-étnicas en el gen 5-alfa reductasa como factor implicado en la variabilidad de respuesta clínica a los inhibidores de esta enzima

6.1.- Introducción

El cáncer de próstata (CaP) es un tipo de cáncer hormono-dependiente que necesita de la presencia de andrógenos para su iniciación y la posterior progresión de la enfermedad (Huggins 1967). Es por ello que la ablación androgénica ha sido la principal intervención terapéutica para el tratamiento de aquellos CaP hormono-sensibles (Eisenberger et al. 1998). Aunque la testosterona (T) es el andrógeno más abundante en suero, la dihidrotestosterona (DHT) es el principal andrógeno prostático. El proceso metabólico por el cual la T pasa a ser DHT está catalizado por la enzima 5-alfa-esteroide reductasa (SRD5A). Este proceso de conversión se lleva a cabo en la próstata, los testículos, los folículos capilares y las glándulas adrenales (McConnell 1995). El papel de la DHT fue conocido con posterioridad al descubrimiento de la deficiencia del polipéptido alfa 2 de dicha enzima (SRD5A2), de tal manera que los individuos que padecían esta alteración no presentaban ni hiperplasia benigna de próstata (HBP) ni CaP.

Se han identificado tres isoenzimas de SRD5A: polipéptido alfa 1 (SRD5A1), polipéptido alfa 2 (SRD5A2) y polipéptido alfa 3 (SRD5A3) (Aumuller et al. 1996). Las isoenzimas SRD5A1 y SRD5A2 son las mejor estudiadas y están implicadas en la reducción de esteroides 5-alfa (*steroid 5-alfa-reduction*). Desde que se descubrieron las implicaciones de la deficiencia de SRD5A2 se ha incrementado el interés en el desarrollo de inhibidores para 5-alfa-esteroide reductasa (5 α -RI). De entre los diferentes compuestos desarrollados, dos fármacos han sido aprobados por la *U.S. Food and Drug Administration* (FDA) para el uso clínico: finasterida, que inhibe SRD5A2, aprobado para el tratamiento de HBP y alopecia masculina; y dutasterida, el cual inhibe SRD5A1

y 2, aprobado para el tratamiento de HBP (Edwards and Moore 2002; Andriole and Kirby 2003).

En las décadas más recientes, la incidencia del CaP ha aumentado en la mayoría de países desarrollados, incluyendo España. Por el contrario, la tasa de mortalidad debida al CaP se ha reducido simultáneamente (Quaglia et al. 2003). Esto, junto con el hecho de que la incidencia del CaP tiene una elevada dependencia de la edad, hace de esta enfermedad un blanco ideal para la quimioprevención. Dado que el CaP es altamente dependiente de los andrógenos y que la DHT es el principal andrógeno prostático, el empleo de 5 α -RIs parece un método razonable para la prevención del CaP (Thompson et al. 2003; Andriole et al. 2010). Algunos autores han mostrado una reducción de la incidencia de CaP en biopsias aleatorias, sin embargo, este hallazgo podría representar una disminución o una inhibición del crecimiento de tumores ya existentes y no un indicador de prevención tumoral (Andriole et al. 2010). Otros autores, por el contrario, no han apreciado un decrecimiento significativo de la incidencia de CaP (Wilt et al. 2008; Murtola et al. 2009). En la actualidad, se ha llegado a la conclusión de que dutasterida y finasterida no previenen el CaP, sino que simplemente reducen tumores con poco potencial de letalidad (Walsh 2010). En el CaP, se ha descrito una disminución de la expresión del gen *SRD5A2* y consecuentemente, una reducción de la actividad enzimática (Luo et al. 2003), un hecho que posiblemente condiciona la respuesta a los inhibidores de 5-alfa-esteroide reductasa.

Los polimorfismos se definen como variaciones genéticas que aparecen como mutaciones en algunos individuos y son transmitidas a la descendencia hasta alcanzar una frecuencia en la población de al menos el 1%. Estas variaciones son consideradas la base de la evolución, de tal forma que aquellas variaciones que se consolidan en la población pueden aparecer silenciadas, aportar algún tipo de beneficio a los individuos que las portan, o bien pueden estar involucradas en el desarrollo de ciertas enfermedades (Guttmacher and Collins 2002). Los polimorfismos más frecuentes son los polimorfismos de nucleótido simple (SNPs). En los genes que codifican para las enzimas *SRD5A1* y 2 se han identificado varios polimorfismos relacionados con el riesgo de padecer CaP (Paz-y-Mino et al. 2009; Setlur et al. 2010). También la actividad enzimática de *SRD5A* está modulada por ciertos SNPs localizados en estos genes (Giwercman et al. 2005; Cussenot et al. 2007). Del mismo modo, tanto los niveles de hormonas esteroides (sexuales) circulantes y de globulina transportadora de hormonas

esteroides (SHBG), como el resultado clínico del CaP (en términos de fallo bioquímico), están influenciados por algunos SNPs localizados en los genes que codifican para estas enzimas (Ahn et al. 2009; Audet-Walsh et al. 2011). Estas variaciones genéticas en genes clave de las rutas de andrógenos parecen ser importantes en la respuesta prostática a dichos andrógenos (Lindstrom et al. 2006). Como se ha argumentado en el capítulo anterior, las diferencias en la distribución genotípica entre distintas poblaciones de una misma etnia pueden ser un importante factor de confusión en estudios de asociación, especialmente cuando se llevan a cabo extensos meta-análisis con sujetos procedentes de diferentes países (Henriquez-Hernandez et al. 2013).

Con todas estas premisas hemos diseñado este estudio que persigue evaluar la distribución genotípica de 22 SNPs localizados en los genes *SRD5A1* y *SRD5A2*, en un amplio conjunto de pacientes de CaP españoles, con el fin de determinar la homogeneidad de la población y revelar potenciales sesgos asociados con el fallo de 5 α -RIs cuando estos tratamientos son prescritos.

6.2.- Resultados

Todas las muestras genotipadas en este estudio cumplían los criterios de calidad indicados en el apartado “Pacientes y Métodos” y todas ellas fueron genotipadas con el mismo lote de material y al mismo tiempo. Un total de 601 pacientes españoles de CaP fueron genotipados para 22 SNPs: 10 SNPs localizados en el gen *SRD5A1* y 12 SNPs localizados en el gen *SRD5A2*. De las 13 222 posibles determinaciones genotípicas, el 94.31% fueron genotipadas con éxito.

Los resultados de frecuencias genotípicas y alélicas se muestran en la tabla 10. En 6 de los 10 SNPs localizados en el gen *SRD5A1* se observó un exceso relativo de heterocigosis, lo que es indicativo de una desviación del HWE. Lo mismo ocurrió en 9 de los 12 SNPs localizados en el gen *SRD5A2*.

Para comparar las poblaciones, una frente a otra de forma individual, se utilizó la herramienta SNPator. Los resultados muestran diferencias de distribución genotípica entre las poblaciones. Se hallaron diferencias en la distribución de genotipos entre las diferentes poblaciones en estudio en 12 de los 22 SNPs: rs166050, rs501999, rs518673, rs3822430, rs8192120, rs39848 (*SRD5A1*); y rs2208532, rs12470143, rs2281546, rs3754838, rs523349, rs9332975 (*SRD5A2*) (test de χ^2 , Tabla 10). De acuerdo con estos resultados, las poblaciones de Cataluña y Andalucía mostraron las mayores diferencias,

presentando una distribución genotípica diferente en 11 de los 22 SNPs (6 localizados en el gen *SRD5A1* y 5 en el gen *SRD5A2*).

Tabla 10. Frecuencias alélicas y genotípicas de SNPs localizados en los genes *SRD5A1* y *SRD5A2* entre las diferentes poblaciones

	% determinación	Genotipos			Alelos		Consecuencia funcional	MAF
<i>SRD5A1</i>								
rs166050		AA	AG	GG	A	G	<i>Intron variation</i>	0.12
Andalucía	0.75	0.65	0.21	0.14	0.75	0.25		
País Vasco	0.98	0.52	0.42	0.06	0.73	0.27		
Canarias	0.99	0.62	0.33	0.05	0.79	0.21		
Cataluña	0.94	0.57	0.37	0.06	0.76	0.24		
<i>p value</i>				0.025				
rs501999		CC	CT	TT	C	T	<i>Intron variation</i>	0.49
Andalucía	0.70	0.48	0.20	0.32	0.59	0.41		
País Vasco	1.00	0.29	0.49	0.22	0.54	0.46		
Canarias	0.99	0.29	0.49	0.22	0.53	0.47		
Cataluña	0.98	0.30	0.43	0.27	0.51	0.49		
<i>p value</i>				0.003				
rs518673		AA	AG	GG	A	G	<i>Intron variation</i>	0.32
Andalucía	0.75	0.19	0.19	0.62	0.29	0.71		
País Vasco	1.00	0.16	0.37	0.47	0.34	0.66		
Canarias	0.96	0.07	0.42	0.51	0.28	0.72		
Cataluña	0.96	0.10	0.40	0.50	0.30	0.70		
<i>p value</i>				0.006				
rs3822430		AA	AG	GG	A	G	<i>Synonymous codon</i>	0.30
Andalucía	0.77	0.37	0.24	0.39	0.49	0.51		
País Vasco	1.00	0.27	0.61	0.12	0.58	0.42		
Canarias	0.97	0.33	0.51	0.16	0.58	0.42		
Cataluña	0.97	0.37	0.47	0.16	0.60	0.40		
<i>p value</i>				<0.001				
rs500182		GG	GT	TT	G	T	<i>Intron variation</i>	0.20
Andalucía	0.81	0.01	0.14	0.85	0.08	0.92		
País Vasco	1.00	0.02	0.18	0.80	0.11	0.89		
Canarias	0.98	0.01	0.17	0.82	0.09	0.91		
Cataluña	0.99	0.01	0.18	0.81	0.10	0.90		
<i>p value</i>				0.961				
rs8192120		AA	AC	CC	A	C	<i>Intron variation</i>	0.43
Andalucía	0.80	0.14	0.22	0.64	0.25	0.75		
País Vasco	0.99	0.08	0.42	0.50	0.29	0.71		
Canarias	0.97	0.10	0.42	0.48	0.31	0.69		
Cataluña	0.98	0.10	0.48	0.42	0.34	0.66		
<i>p value</i>				0.015				
rs4702378		CC	CT	TT	C	T	<i>Intron variation</i>	0.24
Andalucía	0.82	0.07	0.23	0.70	0.18	0.82		
País Vasco	1.00	0.06	0.20	0.74	0.16	0.84		
Canarias	0.99	0.04	0.33	0.63	0.21	0.79		
Cataluña	0.99	0.06	0.33	0.61	0.23	0.77		
<i>p value</i>				0.304				
rs1691053		AA	AG	GG	A	G	ND	0.15
Andalucía	0.87	0.90	0.09	0.01	0.95	0.05		
País Vasco	1.00	0.90	0.08	0.02	0.94	0.06		
Canarias	0.98	0.86	0.13	0.01	0.93	0.07		
Cataluña	0.99	0.85	0.15	0.00	0.93	0.07		
<i>p value</i>				0.341				
rs39848		CC	CT	TT	C	T	<i>Downstream variation</i>	0.43
Andalucía	0.70	0.32	0.22	0.46	0.43	0.57		
País Vasco	0.96	0.22	0.43	0.35	0.44	0.56		
Canarias	0.97	0.11	0.50	0.39	0.36	0.64		
Cataluña	0.94	0.16	0.45	0.39	0.39	0.61		
<i>p value</i>				<0.001				
rs3797179		AA	AG	GG	A	G	<i>Intron variation</i>	0.12
Andalucía	0.70	0.00	0.17	0.83	0.09	0.91		

País Vasco	0.96	0.00	0.29	0.71	0.14	0.86		
Canarias	0.99	0.05	0.25	0.70	0.17	0.83		
Cataluña	0.96	0.03	0.26	0.71	0.17	0.83		
<i>p value</i>				0.213				
SRD5A2								
rs2208532		AA	AG	GG	A	G	<i>Intron variation</i>	0.44
Andalucía	0.70	0.45	0.28	0.27	0.59	0.41		
País Vasco	0.99	0.20	0.42	0.38	0.41	0.59		
Canarias	0.99	0.34	0.46	0.20	0.57	0.43		
Cataluña	0.92	0.28	0.54	0.18	0.55	0.45		
<i>p value</i>				0.001				
rs12470143		CC	CT	TT	C	T	<i>Intron variation</i>	0.37
Andalucía	0.81	0.45	0.24	0.31	0.57	0.43		
País Vasco	1.00	0.26	0.53	0.21	0.52	0.48		
Canarias	0.99	0.32	0.51	0.17	0.57	0.43		
Cataluña	0.98	0.29	0.54	0.17	0.56	0.44		
<i>p value</i>				0.001				
rs2281546		GG	GT	TT	G	T	<i>Intron variation</i>	0.17
Andalucía	0.83	0.09	0.07	0.85	0.12	0.88		
País Vasco	1.00	0.00	0.27	0.73	0.14	0.86		
Canarias	0.99	0.03	0.22	0.75	0.14	0.86		
Cataluña	0.99	0.02	0.25	0.73	0.14	0.86		
<i>p value</i>				0.001				
rs3754838		CC	CT	TT	C	T	ND	0.12
Andalucía	0.78	0.10	0.11	0.79	0.16	0.84		
País Vasco	1.00	0.00	0.23	0.77	0.12	0.88		
Canarias	0.96	0.03	0.20	0.77	0.12	0.88		
Cataluña	0.98	0.02	0.20	0.78	0.12	0.88		
<i>p value</i>				0.010				
rs4952222		AA	AC	CC	A	C	<i>Intron variation</i>	0.04
Andalucía	0.85	0.00	0.01	0.99	0.01	0.99		
País Vasco	1.00	0.00	0.00	1.00	0.00	1.00		
Canarias	0.99	0.00	0.00	1.00	0.00	1.00		
Cataluña	0.97	0.00	0.00	1.00	0.00	1.00		
<i>p value</i>				0.086				
rs7562326		CC	CT	TT	C	T	<i>Intron variation</i>	0.14
Andalucía	0.80	0.05	0.12	0.83	0.12	0.88		
País Vasco	1.00	0.00	0.23	0.77	0.12	0.88		
Canarias	0.98	0.03	0.20	0.77	0.13	0.87		
Cataluña	0.97	0.02	0.21	0.77	0.12	0.88		
<i>p value</i>				0.321				
rs2300702		CC	CG	GG	C	G	<i>Intron variation</i>	0.45
Andalucía	0.81	0.28	0.34	0.38	0.45	0.55		
País Vasco	0.96	0.12	0.49	0.39	0.37	0.63		
Canarias	0.99	0.21	0.47	0.32	0.44	0.56		
Cataluña	0.94	0.16	0.50	0.34	0.41	0.59		
<i>p value</i>				0.127				
rs4952197		AA	AG	GG	A	G	<i>Intron variation</i>	0.25
Andalucía	0.80	0.12	0.19	0.69	0.22	0.78		
País Vasco	0.99	0.04	0.28	0.68	0.18	0.82		
Canarias	0.99	0.06	0.31	0.63	0.22	0.78		
Cataluña	0.97	0.06	0.36	0.58	0.24	0.76		
<i>p value</i>				0.117				
rs676033		CC	CT	TT	C	T	ND	0.35
Andalucía	0.81	0.57	0.31	0.12	0.72	0.28		
País Vasco	0.99	0.60	0.32	0.08	0.76	0.24		
Canarias	0.99	0.50	0.39	0.11	0.70	0.30		
Cataluña	0.98	0.50	0.43	0.07	0.71	0.29		
<i>p value</i>				0.402				
rs523349		CC	CG	GG	C	G	<i>Missense</i>	0.35
Andalucía	0.78	0.55	0.27	0.18	0.68	0.32		
País Vasco	0.92	0.64	0.32	0.04	0.80	0.20		
Canarias	0.71	0.18	0.52	0.30	0.45	0.55		
Cataluña	0.89	0.52	0.40	0.08	0.72	0.28		
<i>p value</i>				<0.001				
rs9332975		CC	CT	TT	C	T	<i>utr variation</i>	0.14
Andalucía	0.77	0.10	0.07	0.83	0.14	0.86		

País Vasco	1.00	0.00	0.26	0.74	0.13	0.87		
Canarias	0.99	0.03	0.19	0.78	0.13	0.87		
Cataluña	0.98	0.02	0.20	0.78	0.12	0.88		
<i>p value</i>				0.002				
rs7594951		CC	CT	TT	C	T	<i>Intron variation</i>	0.11
Andalucía	0.75	0.84	0.10	0.06	0.89	0.11		
País Vasco	1.00	0.76	0.24	0.00	0.88	0.12		
Canarias	0.99	0.77	0.20	0.03	0.88	0.12		
Cataluña	0.98	0.78	0.20	0.02	0.88	0.12		
<i>p value</i>				0.193				

Abreviaturas: MAF, frecuencia del alelo menor; ND, no disponible.
Las consecuencias funcionales y las MAF se encuentran disponibles en: <http://www.ncbi.nlm.nih.gov/projects/SNP/>
Las diferencias estadísticas entre las distribuciones genotípicas fueron calculadas mediante el test χ^2 .

Para visualizar las distribuciones genotípicas entre las cuatro poblaciones se realizó un análisis de agrupamiento jerárquico no supervisado. Tal y como se muestra en la figura 12, los polimorfismos localizados en *SRD5A1* se distribuyen en tres clústeres principales, cada uno con diferente número e identidades de SNPs, sugiriendo heterogeneidad entre las poblaciones. Para *SRD5A2*, los SNPs se distribuyeron en dos clústeres principales, muy similares para las poblaciones del País Vasco, Canarias y Cataluña, pero claramente diferente a la población de Andalucía (Figura 13).

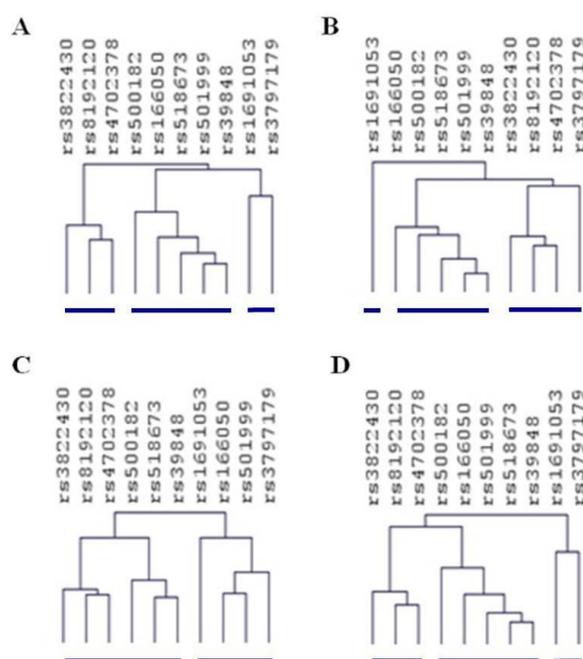


Figura 12, agrupamiento jerárquico no supervisado de los distintos SNPs del gen *SRD5A1* de las poblaciones de Andalucía (A), País Vasco (B), Canarias (C) y Cataluña (D). Los clústeres fueron realizados utilizando la correlación de distancia Euclidiana y el ligamiento medio y fueron procesados con el programa MultiExperiment Viewer (<http://www.tigr.org>). Los dendogramas muestran los clústeres de los SNPs. Las líneas azules debajo de cada panel muestran los clústeres principales generados.

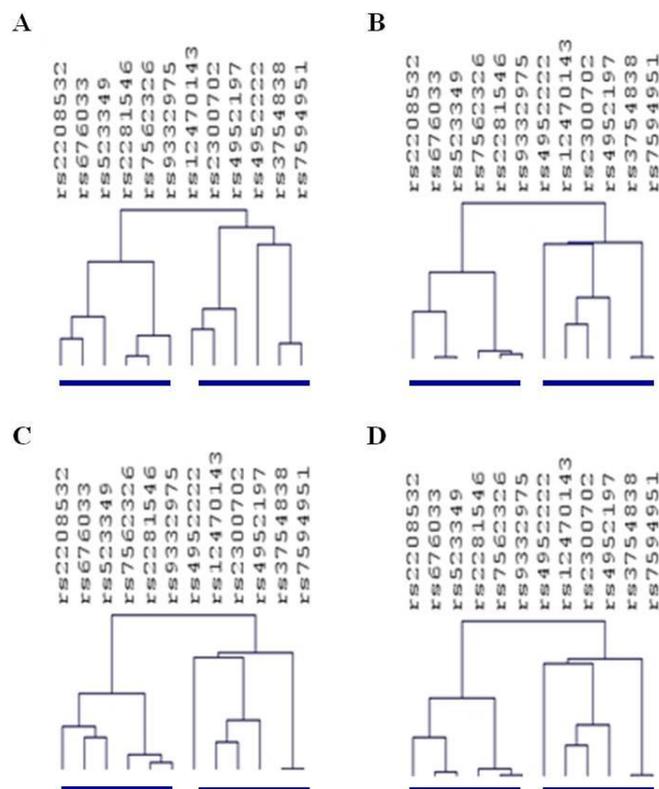


Figura 13, agrupamiento jerárquico no supervisado de los distintos SNPs del gen *SRD5A2* de las poblaciones de Andalucía (A), País Vasco (B), Canarias (C) y Cataluña (D). Los clústeres fueron realizados utilizando la correlación de distancia Euclidiana y el ligamiento medio y fueron procesados con el programa MultiExperiment Viewer (<http://www.tigr.org>). Los dendogramas muestran los clústeres de los SNPs. Las líneas azules debajo de cada panel muestran los dos clústeres principales generados.

El análisis de los componentes principales (PCA) se llevó a cabo para identificar diferencias globales entre las poblaciones. Los componentes 1, 2 y 3 fueron responsables del 43.7%, 22.6% y 15.3% de la varianza para *SRD5A1*, respectivamente (porcentaje acumulado de 81.6%); y del 48.8%, 33.2% y 9.1% de la varianza para *SRD5A2* (porcentaje acumulado de 91.1%). El primer componente se distinguió frente a los otros componentes para ambos genes (Figuras 14A y 14B), mostrando las diferencias en la distribución de genotipos entre las poblaciones analizadas, especialmente para el gen *SRD5A1* (lo que concuerda con lo observado en los análisis de agrupamiento no jerarquizado).

Se realizó un análisis de haplotipos con la herramienta SNPator. Para *SRD5A1* (localizado en el cromosoma 5), el haplotipo GCCTATGGCA sólo estaba presente entre los sujetos de la población de Andalucía. Para *SRD5A2* (localizado en el cromosoma 2),

se observó una mayor homogeneidad en la distribución de haplotipos. Estos resultados sugieren mayor variabilidad en el gen *SRD5A1* que en el gen *SRD5A2*, tal y como se sustrae de los resultados del resto de los análisis realizados. No obstante, el hecho de que los haplotipos más frecuentes sean iguales en todas las poblaciones, alude a una similitud entre los individuos de la misma etnia.

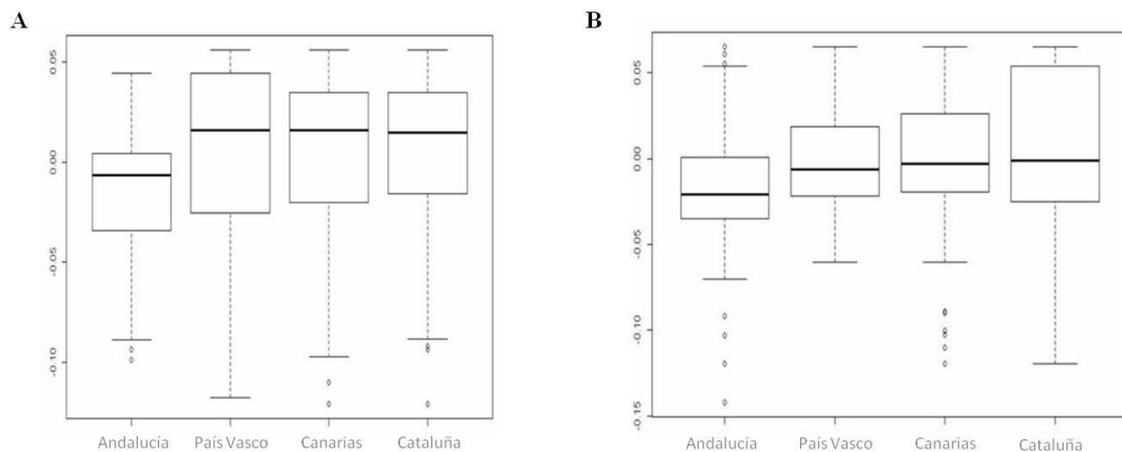


Figura 14, diagrama de cajas del componente 1 en las diferentes poblaciones después del análisis del componente principal (PCA) para *SRD5A1* (A) y *SRD5A2* (B).

6.3.- Discusión

La presencia de determinados polimorfismos en un paciente puede condicionar profundamente la efectividad de determinados tratamientos farmacológicos (Ma and Lu 2011). De forma específica, se ha descrito que ciertos SNPs en *SRD5A2* (como el polimorfismo rs523349) modifican la actividad enzimática y condicionan la conversión de T a DHT (Cussenot et al. 2007). Se sabe que las frecuencias genotípicas y alélicas varían mucho según la etnia del individuo. Al mismo tiempo, se está incrementando el reconocimiento de las diferencias intra-étnicas como importante factor condicionante de la variabilidad de la respuesta a fármacos observada entre diferentes individuos. Como consecuencia, se ha creado una nueva rama de conocimiento denominada fármaco-etnia (Yasuda et al. 2008), la cual podría describirse como el estudio de influencia de la diversidad étnica en la respuesta a fármacos y la toxicidad generada por estos. La apreciación de diferencias asociadas a la etnia queda a menudo oculta por la necesidad de grandes poblaciones para la realización de ensayos clínicos, que debido precisamente al gran tamaño, presentan inevitablemente una gran diversidad de individuos. Así, tal y

como indica la hoja de datos de finasterida, “el efecto de la raza en la farmacocinética de finasterida no ha sido estudiado”. En el capítulo anterior de esta Tesis Doctoral se han reportado las diferencias en la distribución de genotipos entre poblaciones de la misma etnia (Henriquez-Hernandez et al. 2013), resaltando la importancia de las diferencias asociadas a la etnia cuando se hacen estudios de asociación genética. Hemos querido explorar este campo no solo en lo referido a genes relacionados con la reparación del ADN, sino también en genes relacionados con el metabolismo de los andrógenos, moléculas importantes en el desarrollo de la patología tumoral prostática. Por ello, se diseñó este estudio con el objetivo de evaluar la distribución genotípica de 22 SNPs localizados en los genes *SRD5A1* y *SRD5A2* en un amplio conjunto de pacientes españoles de CaP, con el fin de determinar la homogeneidad de la población y revelar potenciales sesgos asociados, en este caso, a la prescripción de 5 α -RIs.

Nuestros resultados muestran que la distribución genotípica de 12 de los 22 SNPs estudiados fue estadísticamente diferente entre las poblaciones incluidas en el ensayo. La población española es considerada caucásica. Sin embargo, la historia natural de cada subpoblación presenta significativas variaciones, ya que cada una de ellas ha recibido de diferente manera el impacto de las distintas fuentes poblacionales originadas por la gran complejidad en los procesos migratorios del Mediterráneo, generando así la composición genética característica de la población española (Ambrosio et al. 2010). Mientras que la población de la península ibérica parece haber sido originada por migraciones de la península arábiga, *fertile crescent*, la región balcánica y el norte de África (Ambrosio et al. 2010), la población canaria ha sido influenciada por la migración del noroeste de África y la posterior colonización europea (Rando et al. 1999). Las diferencias entre poblaciones fueron también evidentes en el análisis de haplotipos, lo cual sugiere que cada SNP debe ser considerado individualmente cuando el fin sea encontrar posibles variables de confusión que puedan ser cruciales para la interpretación de los resultados. El hecho de que la variabilidad encontrada en *SRD5A1* sea mayor que la obtenida para *SRD5A2* apoya esta observación. A pesar de estos resultados, como norma general, ni la etnia ni la dotación genética de los pacientes es tomada en cuenta a la hora de diseñar un ensayo clínico. Los inhibidores 5 α -RI, bloqueadores específicos de la enzima responsable de sintetizar DHT, despiden como tratamientos prometedores para la prevención del CaP. Sin embargo, los más estudiados hasta la fecha, dutasterida y finasterida, parecen no ser útiles para esta

prevención (Walsh 2010). En el presente trabajo sugerimos que los SNPs localizados en *SRD5A1* y *SRD5A2* podrían ser un factor relevante a tener en cuenta cuando el objetivo sea determinar la respuesta a estos fármacos. El hecho de que rs523349 – un SNP claramente involucrado en el éxito de la quimioprevención del CaP mediante 5 α -RI (Cussenot et al. 2007) – aparezca diferencialmente distribuido en nuestro estudio, refuerza estos hallazgos. Esta hipótesis se basa en la premisa de que ciertos polimorfismos localizados en estos genes parecen estar asociados no solo con los resultados clínicos (Audet-Walsh et al. 2011), sino también con el riesgo de padecer CaP (Li et al. 2010; Li et al. 2013) y con las diferencias en los niveles de globulinas circulantes portadoras de hormonas sexuales y hormonas sexuales esteroideas (Ahn et al. 2009). Además del papel conductor que desempeñan los SNPs en la farmacogenética, existen otros factores que interactúan de forma combinada para determinar el resultado final (como diferencias ambientales, diferencias en las prácticas locales o diferencias en la interacción fármaco-fármaco) (O'Donnell and Dolan 2009). Teniendo en cuenta todo lo anterior, es necesario que los estudios y ensayos clínicos encaminados a determinar la utilidad de 5 α -RI en la quimioprevención del CaP se sustenten sobre un buen diseño de la población de estudio.

El presente trabajo tiene algunas limitaciones que deben observarse. Primero, para cegar el estudio, los datos clínicos de los pacientes no estuvieron disponibles, esto es, no se utilizaron datos sobre estadio TNM, grado tumoral, fallo bioquímico o *Gleason score*. La presencia de ciertos SNPs del gen *SRD5A2* en hombres jóvenes con CaP está asociada con una enfermedad más agresiva en el momento del diagnóstico (Scariano et al. 2008). Del mismo modo, algunas características patológicas del CaP parecen estar influenciadas por SNPs localizados en el gen *SRD5A2* (Jaffe et al. 2000). Estas observaciones enfatizan la necesidad de seleccionar los pacientes cuidadosamente, no solo en cuanto a sus características clínicas y demográficas, sino también en relación a su etnia. Segundo, el estudio está enfocado en los genes *SRD5A1* y *SRD5A2*, ignorando posibles asociaciones con otros genes y polimorfismos. Se ha establecido una relación entre SNPs localizados en el gen 3-beta-hidroxiesteroide deshidrogenasa tipo II (*HSD3B2*), los cuales se asocian con el incremento del riesgo y la agresividad al CaP, y SNPs de *SRD5A2*, que modifican estas cualidades (Neslund-Dudas et al. 2007). En el mismo sentido, se ha sugerido una relación entre SNPs del gen *CYP17A1* (un gen que codifica una enzima clave en la síntesis de andrógenos) y SNPs del gen *SRD5A2* (Onen

et al. 2007). Tercero, el número de sujetos de cada una de las poblaciones varía ampliamente. Sin embargo, el hecho de que las mayores diferencias no hayan sido encontradas en las poblaciones con menor número de pacientes (País Vasco, con 51 muestras) sugiere que esta limitación puede no ser decisiva en la interpretación de los resultados. Finalmente, hemos observado que diferentes SNPs se desvían del HWE, especialmente en la población de Andalucía. El HWE se define como una ecuación predictiva para las frecuencias genotípicas en grandes poblaciones en términos de frecuencias alélicas de un locus. Además de lo limitado del tamaño muestral ($n = 91$ para pacientes de Andalucía), es posible que la distribución genotípica en esta población esté fuertemente influenciada por características tumorales no tenidas en consideración. Además, dado que las condiciones bajo las cuales el HWE es estrictamente válido son muy severas y que las poblaciones reales no satisfacen la mayoría de estas condiciones, la aplicación del HWE para la predicción de las frecuencias genotípicas es cuestionable (Lewontin and Hartl 1991; Chakraborty 2002). En este sentido, un estudio de casos y controles, especialmente en la población de Andalucía, se presenta necesario para confirmar y encontrar explicación a estos hallazgos.

Por otro lado, deben resaltarse varias ventajas: i) el estudio incluye un número de sujetos suficiente para obtener datos fidedignos sobre la distribución de estos 22 SNPs en las poblaciones de CaP estudiadas (especialmente para Canarias y Cataluña); ii) todos los sujetos son varones (los 5α -RI están contraindicados en mujeres), eliminando así el posible sesgo derivado del género y iii) todas las determinaciones (13 222 en total) fueron realizadas con la misma metodología (OpenArray, Applied Biosystems), con el mismo lote de material y por el mismo investigador, minimizando así los sesgos de origen técnico. Todos los análisis han sido automáticos y no ha habido determinación manual de genotipos en ningún caso.

6.4.- Conclusión

La diferencia en la distribución de genotipos de *SRD5A1* y *SRD5A2* entre las diferentes poblaciones de la misma etnia puede ser un factor responsable del fallo de los tratamientos basados en 5α -RI. Nuestros resultados sugieren que la dotación genética debe ser tomada en cuenta (especialmente en ensayos clínicos que incluyan gente de diferentes etnias) en la determinación de la efectividad de este tipo de fármacos.

Capítulo III

7.- Capítulo III: Polimorfismos de nucleótido simple en genes de reparación del ADN como factores de riesgo asociados a la progresión del cáncer de próstata.

7.1.- Introducción

El cáncer de próstata (CaP) es una enfermedad compleja, cuyo comportamiento tumoral parece estar influenciado por factores hormonales y genéticos. El estadio tumoral, el grado tumoral en términos de *Gleason score* y los niveles de antígeno prostático específico (PSA) en suero son factores utilizados clínicamente para la clasificación de los pacientes en los diferentes grupos de riesgo, condicionando las decisiones que los facultativos toman con respecto al tratamiento de este tipo de pacientes. Sin embargo, se estima que para impedir una muerte relacionada con CaP, 293 hombres deben ser cribados y 12 de ellos tratados, lo que sugiere que los factores predictivos usados diariamente en la clínica necesitan ser mejorados. La investigación en el área de la genómica apunta a la identificación de nuevos marcadores para el CaP, sin embargo, los resultados no son concluyentes (Kote-Jarai et al. 2008) y es patente la necesidad de determinar nuevos biomarcadores que pronostiquen el comportamiento tumoral y ayuden así a tomar decisiones críticas respecto al tratamiento.

El ADN de las células está expuesto a numerosos elementos que lo dañan continuamente, como los radicales libres de oxígeno endógenos y algunos químicos de origen exógeno. Para invertir las lesiones producidas en el ADN, las células disponen de diferentes rutas de reparación (Wood et al. 2001). Los defectos en estas rutas de reparación del ADN pueden incrementar el número de mutaciones persistentes que son transmitidas a las generaciones de células hijas, provocar inestabilidad genómica y en última instancia, condicionar la agresividad de la enfermedad tumoral (Park et al. 2009). Existe una ruta de reparación diferente para cada tipo de daño que puede sufrir el ADN, y todas ellas son esenciales para el mantenimiento de la integridad del genoma (Yu et al. 1999). Así, encontramos la reparación por escisión de una base (*base excision repair*,

BER), reparación por escisión de un nucleótido (*nucleotide excision repair*, NER), reparación por despareamiento (*mismatch repair*, MMR) o reparación de roturas de doble cadena (*double strand break repair*, DSBR). La variabilidad genética de los genes involucrados en dicha reparación podría conferir susceptibilidad a la aparición de determinados tumores y además, estar asociada con la agresividad de la enfermedad (Zhu et al. 2004), pudiendo incluso inducir la transformación del tejido normal en tumoral mediante la adquisición de propiedades oncológicas. Los polimorfismos de nucleótido simple (SNPs) se definen como variaciones del genoma que están asentadas en la población, presentando al menos una proporción del 1%. Teniendo en cuenta que hay millones de SNPs en el genoma humano, el hecho de que se seleccionen aquellos SNPs que puedan afectar a las funciones fenotípicas y que en último término contribuyan al desarrollo de enfermedades es un hecho poco probable. Los estudios de gen candidato se centran en la selección de genes y polimorfismos que han sido previamente relacionados con una enfermedad y de los cuales ya se conoce su función. Entre los genes involucrados directamente en la reparación de ADN en humanos (Ronen and Glickman 2001), seis de ellos ya han sido estudiados en amplias series de pacientes españoles con CaP, seleccionados por su relevancia en los mecanismos que dirigen la enfermedad (Henriquez-Hernandez et al. 2013): *x-ray repair cross-complementing protein 1 (XRCC1)*, *excision repair cross-complementing rodent repair deficiency, complementation group 2 (ERCC2)*, *excision repair cross-complementing rodent repair deficiency, complementation group 1 (ERCC1)*, ligasa 4 (*LIG4*), ataxia telangiectasia mutada (*ATM*) y proteína tumoral p53 (*TP53*).

Como hemos visto en los Capítulos I y II de esta Tesis Doctoral, el origen étnico de la población en estudio es un factor imprescindible a tener en cuenta en los estudios de asociación de genes. De hecho, en la literatura encontramos multitud de publicaciones que estudian variaciones genéticas que son factores de riesgo para ciertas enfermedades entre sujetos de un determinado origen étnico, pero cuyos resultados no son válidos en sujetos de otras etnias (Correa-Cerro et al. 1999; Giwercman et al. 2008). Además, se ha publicado que las diferencias en la distribución genotípica entre diferentes poblaciones de la misma etnia pueden ser un importante factor de confusión en los estudios de asociación genética (Henriquez-Hernandez et al. 2013). En relación a lo anterior, mientras que las cohortes estudiadas son a menudo multi-étnicas, la guía STROGAR (*STrengthening the Reporting Of Genetic Association studies in*

Radiogenomics) recomienda indicar si la etnia está identificada y codificada en la determinación de las asociaciones genotipo-fenotipo y alienta al uso de cohortes de poblaciones étnicamente uniformes (Kerns et al. 2014).

La hipótesis de este trabajo sugiere que las variaciones genéticas en genes involucrados en la reparación del ADN confieren diferente comportamiento a las células del CaP y pueden suponer fenotipos clínicos diferentes. Es por ello que el objetivo se fundamenta en determinar la relación existente entre 10 SNPs localizados en 6 genes que participan en la reparación del ADN y que han sido clásicamente asociados con el riesgo de padecer CaP (Park et al. 2009), con la agresividad tumoral, en una amplia muestra de pacientes españoles con CaP.

7.2.- Resultados

Todas las muestras genotipadas cumplían con los criterios de calidad establecidos en el apartado “Pacientes y Métodos” y todas ellas fueron genotipadas con el mismo lote de material y al mismo tiempo. Vistos los resultados obtenidos en el Capítulo I de esta Tesis Doctoral, excluimos del presente estudio a los pacientes procedentes de Andalucía, para garantizar una homogeneidad de la muestra en cuanto a la distribución de los polimorfismos en estudio, minimizando así el sesgo debido a variación intra-étnica. Así, el número total de pacientes incluidos en este Capítulo III fue de 494. Se genotiparon 10 SNPs en un total de 494 pacientes de CaP. Del total de las 4940 posibles determinaciones, el 97.17% fueron genotipadas satisfactoriamente. Las frecuencias genotípicas y alélicas se muestran en la tabla 11. Las frecuencias del alelo menor (*minor allele frequencies*, MAF) obtenidas fueron similares a las reportadas en la literatura. Todos los SNPs se encontraban en HWE.

Tabla 11. Frecuencias alélicas y genotípicas de los pacientes incluidos en el estudio (n = 494)

Gen/SNP	Función	chr	n	Frecuencias genotípicas					Frecuencias alélicas				MAF#	Consecuencia funcional#	
<i>XRCC1</i>	BER	19q13	436	CC	0.42	CT	0.48	TT	0.10	C	0.66	T	0.34	0.26	<i>missense</i>
rs25487*			483	CC	0.88	CT	0.12	TT	0.00	C	0.94	T	0.06	0.06	<i>missense</i>
rs25489										A	0.06	G	0.94	0.13	<i>missense</i>
rs1799782*	NER	19q13	487	AA	0.00	AG	0.12	GG	0.88						
<i>ERCC2</i>			482	GG	0.10	GT	0.47	TT	0.43	G	0.34	T	0.66	0.24	<i>missense</i>
rs13181	NER	19q13								A	0.61	G	0.39	0.36	<i>synonymous codon</i>
<i>ERCC1</i>			488	AA	0.39	AG	0.46	GG	0.16						
rs11615	DSBR	13q23								A	0.17	G	0.83	0.15	<i>missense</i>
<i>LIG4</i>			488	AA	0.04	AG	0.25	GG	0.71						
rs1805388†			480	AA	0.70	AG	0.26	GG	0.04	A	0.83	G	0.17	0.10	<i>synonymous codon</i>
rs1805386†	DSBR	11q22								G	0.09	T	0.91	0.06	<i>intron variant</i>
<i>ATM</i>			486	GG	0.00	GT	0.18	TT	0.81						
rs17503908			486	CC	0.94	CG	0.06	GG	0.00	C	0.97	G	0.03	0.02	<i>missense</i>
rs1800057	DSBR	17p13													
<i>TP53</i>			486	CC	0.59	CG	0.35	GG	0.06	C	0.76	G	0.24	0.39	<i>missense</i>
rs1042522			484	CC	0.59	CG	0.35	GG	0.06						

Abreviaturas: BER, reparación por escisión de base; NER, reparación por escisión de nucleótido; DSBR, reparación de rotura doble de cadena; chr, cromosoma; MAF, frecuencia del alelo menor.

#Información disponible en: <http://www.ncbi.nlm.nih.gov/projects/SNP/>

*SNPs en desequilibrio de ligamiento.

† SNPs en desequilibrio de ligamiento.

La distribución de las variables clínicas viene detallada en la tabla 12. La mayoría de los pacientes de CaP eran cT1a – cT2a (54.7%), PSA < 10 ng/ml (61.9%) y *Gleason score* < 7 (45.7%). No se observaron diferencias en las variables clínicas entre las poblaciones de las diferentes regiones de España (dato no mostrado).

Tabla 12. Descripción de las variables clínicas		
Variables clínicas	n	(%)
Tamaño tumoral clínico (cT)		
cT1a – cT2a	270	(54.7)
cT2b – cT2c	141	(25.8)
cT3 – cT4	66	(13.4)
ND	17	(3.4)
PSA al diagnóstico (ng/mL)		
< 10	306	(61.9)
10 – 19.99	103	(20.9)
> 20	79	(16.0)
NA	6	(1.2)
<i>Gleason score</i>		
< 7	226	(45.7)
7	195	(39.5)
> 7	71	(14.4)
NA	2	(0.4)
D'Amico grupo de riesgo		
Bajo	120	(24.3)
Intermedio	184	(37.2)
Alto	173	(35.0)
ND	17	(3.4)
<i>Abreviaturas:</i> PSA, antígeno prostático específico; ND, no disponible.		

De los 10 SNPs analizados, rs11615 (MAF = 0.39) y rs17503908 (MAF = 0.09), localizados en los genes *ERCC1* y *ATM* respectivamente, mostraron una distribución significativamente diferente entre los pacientes con CaP en relación a las variables clínicas (Tabla 13). De este modo, rs11615 se asoció significativamente al tamaño clínico tumoral (test de χ^2 , $p = 0.002$), mientras que rs17503908 lo fue con los valores de *Gleason score* (test de χ^2 , $p = 0.005$). Con respecto a rs11615, se observó que de entre los 259 pacientes diagnosticados con un tamaño tumoral cT1a – cT2a, 175 fueron portadores del alelo G (67.57%). Por otro lado, de entre los 66 pacientes diagnosticados como cT3 – cT4, 31 portaban el alelo G (46.97%), siendo estadísticamente significativa esta diferencia de distribución (Figura 15A). En lo referente a rs17503908, 169 de los 224 pacientes (75.45%) clasificados con un *Gleason score* < 7 fueron genotipados como

TT, mientras que 59 de los 70 pacientes (84.29%) clasificados con un *Gleason score* > 7 fueron genotipados como TT, siendo estadísticamente significativa esta diferencia de distribución (Figura 15B).

Tabla 13. Asociaciones significativas entre variables clínicas y SNPs

SNP	Genotipos	Tamaño tumoral clínico (cT)			p#
		cT1a – cT2a	cT2b – cT2c	cT3 – cT4	
rs11615	AA	84	63	35	0.002
	AG	140	51	23	
	GG	35	23	8	
		Gleason score			
		<7	7	>7	
rs17503908	GG	0	2	0	0.005
	GT	55	22	11	
	TT	169	166	59	

Test χ^2 .

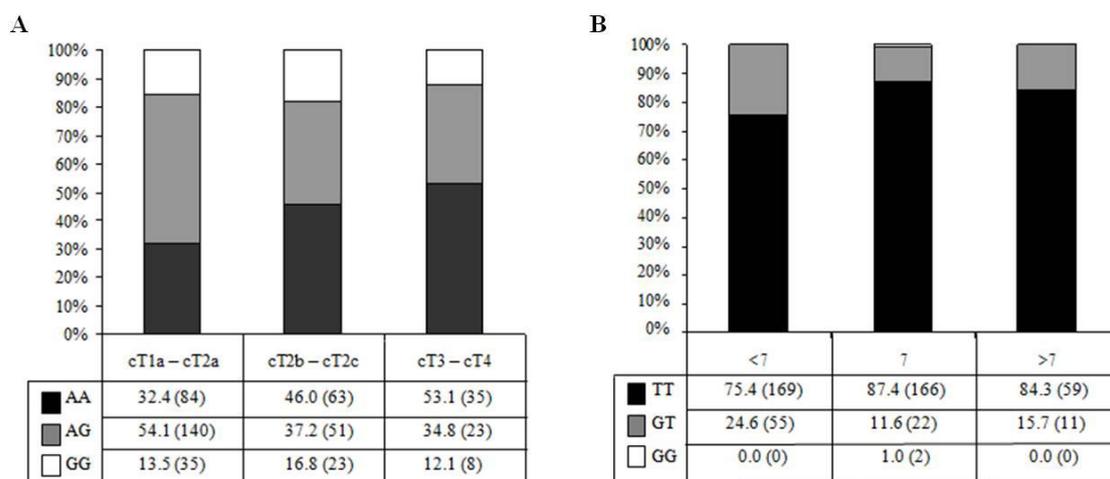


Figura 15, distribución genotípica del SNP rs11615 entre los distintos tamaños tumorales (A) y la distribución genotípica del SNP rs17503908 entre los distintos valores de *Gleason Score* (B).

En posteriores análisis se exploró el papel específico de los SNPs rs11615 y rs17503908 en relación con las variables clínicas a las que se asociaron. Para ello se desarrollaron los análisis de acuerdo con varios modelos genéticos: recesivo, dominante, homocigoto y heterocigoto (Tabla 14). Se observó que los pacientes portadores de los genotipos GG + AG para el SNP rs11615 presentaron menos riesgo de desarrollar tumores cT2b – cT4 que aquellos pacientes de CaP con genotipos AA (OR = 0.504 (Intervalo de Confianza (CI) del 95% = 0.346 – 0.734), $p < 0.0001$). Los mismos

resultados se obtuvieron en el modelo heterocigoto (Tabla 14). La tendencia fue similar para el SNP rs17503908. Los pacientes de CaP portadores de los genotipos GG + GT presentaron menos riesgo de desarrollar tumores con mayor *Gleason score* que aquellos pacientes de CaP portadores del genotipo TT (OR = 0.478 (CI 95% = 0.299 – 0.764), $p = 0.002$), y similares resultados se obtuvieron para el modelo heterocigoto (Tabla 14). Tanto A como T son los alelos ancestrales para los SNPs rs11615 y rs17503908, respectivamente. De acuerdo con nuestros resultados, los genotipos homocigotos salvajes están asociados con factores de peor pronóstico y aquellos pacientes de CaP AA – rs11615 y TT – rs17503908 presentan mayor riesgo de desarrollar tumores cT2b – cT4 (OR = 2.207 (CI 95% = 1.472 – 3.310), $p < 0.001$) y *Gleason score* ≥ 7 (OR = 2.219 (CI 95% = 1.379 – 3.570), $p < 0.001$), respectivamente.

Tabla 14. Análisis univariante de los polimorfismos rs11615 (*ERCC1*) y rs17503908 (*ATM*) en relación a las variables clínicas tamaño clínico tumoral (cT) y *Gleason score*

Variable clínica	n	SNP	Modelo recesivo		Modelo dominante		Modelo homocigoto		Modelo heterocigoto	
			OR (95%CI)	p	OR (95%CI)	p	OR (95%CI)	p	OR (95%CI)	p
cT, 1a-2a/2b-4	270/207	rs11615	1.02 (0.62 – 1.68)	0.928	0.50 (0.35 – 0.73)	<0.0001	0.67 (0.39 – 1.16)	0.152	0.45 (0.30 – 0.70)	<0.0001
			GG vs. AG + AA		GG + AG vs. AA		GG vs. AA		AG vs. AA	
			GG vs. GT + TT		GG + GT vs. TT		GG vs. TT		GT vs. TT	
<i>Gleason</i> , <7/≥7	226/266	rs17503908	NA	–	0.48 (0.30 – 0.76)	0.002	NA	–	0.45 (0.28 – 0.72)	0.001

Abreviaturas: OR, odds ratio; CI, intervalo de confianza; NA, no aplicable (debido al escaso número de muestras; solo 2 genotipados como GG-rs17503908). Prueba estadística: regresión logística binaria (categoría de referencia para SNP rs11615: AA. Categoría de referencia para SNP rs17503908: TT).

Del mismo modo, se estudió el papel de los SNPs rs11615 y rs17503908 en el contexto de los grupos de riesgo según la clasificación D’Amico, la cual estima la agresividad biológica de los tumores de próstata en función de los parámetros clínicos y los agrupa en diferentes categorías de riesgo que reflejan la capacidad de crecimiento y expansión. Los resultados mostraron que los pacientes de CaP con genotipo TT para el SNP rs17503908 presentaban mayor riesgo de desarrollar tumores de alto riesgo según la clasificación D’Amico (OR = 1.694 (CI 95% = 1.015 – 2.827), $p = 0.044$). Por el contrario, no se observó esta tendencia para el SNP rs11615 (Figura 16). No obstante, aquellos pacientes de CaP portadores de los genotipos AA – TT para rs11615 y rs17503908 conjuntamente, presentaron el mayor riesgo de desarrollar tumores de alto riesgo según la clasificación D’Amico (OR = 2.566 (CI 95% = 1.277 – 5.156), $p = 0.008$) (Figura 16). Los resultados fueron similares cuando las series se dividieron en riesgo bajo vs. riesgo intermedio – alto según la clasificación D’Amico (OR = 1.978 (CI 95% = 1.022 – 3.830), datos no mostrados). Estos resultados están en consonancia con los resultados mostrados anteriormente y sugieren que estos genotipos específicos se asocian con factores de mal pronóstico (Figura 17).

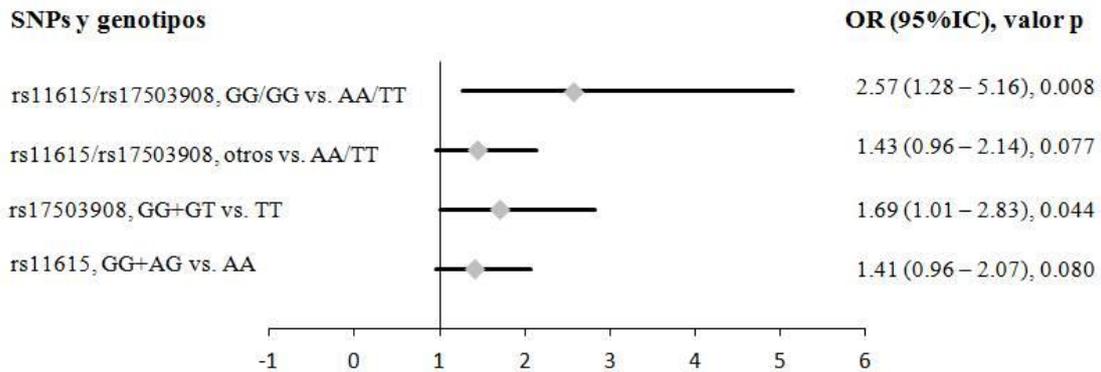


Figura 16, diagrama Forrest de los *odds ratios* (OR) con un 95% de intervalo de confianza (IC) para los polimorfismos rs11615 (*ERCCI*) y rs17503908 (*ATM*) y los grupos de riesgo de recaída según D'Amico. Cada rombo representa el OR y la línea horizontal indica el 95% de IC. Para la regresión binaria logística, los pacientes fueron dicotomizados en dos grupos según el siguiente criterio: riesgo bajo – intermedio vs. riesgo alto según la clasificación D'Amico

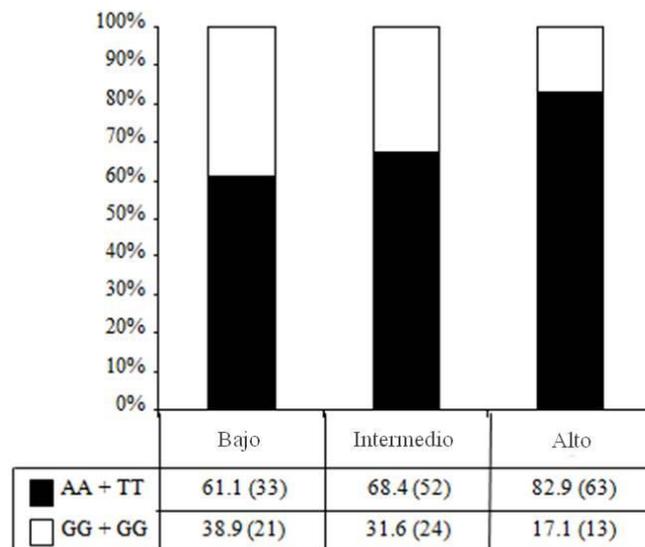


Figura 17, distribución genotípica combinada de los SNPS rs11615 y rs17503908 de acuerdo con los grupos de riesgo según la clasificación D'Amico. La distribución genotípica es estadísticamente diferente (test χ^2 , $p = 0.017$).

Ningún otro polimorfismo mostró asociaciones significativas con las variables clínicas.

7.3.- Discusión

Las lesiones en el ADN son muy frecuentes y conducen a variaciones genéticas tales como deleciones, amplificaciones, reordenamientos y traslocaciones cuyo resultado es la alteración de la homeostasis celular y la aparición de comportamientos tumorales (Wood et al. 2001). En este estudio se investiga la posible asociación entre 10 SNPs localizados en genes que codifican para proteínas involucradas en la reparación del ADN y la agresividad tumoral de un amplio grupo de pacientes españoles con CaP, siguiendo un sistema de gen candidato basado en los estudios publicados previamente (Henriquez-Hernandez et al. 2013). Se observó una fuerte asociación de los genotipos AA – rs11615 y TT – rs17503908 con variables clínicas de mal pronóstico.

Durante la última década se ha producido un incremento en el interés que suscita el estudio del papel de los SNPs en el desarrollo y progresión del CaP. En esta línea, se han explorado en profundidad algunos SNPs localizados en genes que codifican proteínas de reparación del ADN, especialmente en relación con la predicción de la toxicidad inducida por radiación (Van den Broeck et al. 2014). *ERCC1* es un gen localizado en el cromosoma 19q13 que codifica para una proteína relacionada con la reparación por escisión de nucleótido, formando junto con XPF una endonucleasa libre específica nuclear (Ronen and Glickman 2001). En el contexto del CaP, se ha reportado que polimorfismos localizados en *ERCC1* podrían predisponer a las células del epitelio prostático a una transformación hacia células cancerígenas (Matoka et al. 2012). Sin embargo, existe una falta de información sobre el papel de este gen en la agresividad de la enfermedad. Las variables genéticas del cromosoma 19q13 han sido evaluadas en 7370 casos de CaP sin que se haya encontrado relación con la agresividad tumoral (Kote-Jarai et al. 2008). A pesar de lo amplio de la serie, los pacientes incluidos proceden de países diferentes y en él no se tiene en cuenta la etnicidad como factor de confusión, más aún si apreciamos que incluso las diferencias observadas entre poblaciones dentro de la misma etnia sugieren que la raza no es factor suficiente para asegurar la homogeneidad de una muestra (Bauchet et al. 2007). En nuestro estudio se observa que los pacientes de CaP portadores del genotipo AA – rs11615 presentaron mayor riesgo de desarrollar tumores de mayor tamaño. Por el contrario, el alelo G en combinación con rs3212986 (también en el gen *ERCC1*) se ha relacionado con una baja expresión de *ERCC1*, lo que supone una reducción de la capacidad de reparación del ADN y en consecuencia con una mejor respuesta a quimioterapia/radioterapia

(Woelfelschneider et al. 2008). Por extensión, la presencia del alelo A mantendría los niveles normales de ERCC1, condicionando así la malignidad del tumor y la respuesta al tratamiento. En este sentido, el alelo G, que no es el alelo ancestral, podría conferir ventajas clínicas en términos de tamaño tumoral.

El gen *ATM* está codificado en el cromosoma 11q22. En respuesta a las roturas de doble cadena del ADN (*double-strand breaks*, DSBs), ATM fosforila una variedad de proteínas que están involucradas en la reparación de este tipo de lesiones genéticas y en el control del ciclo celular (Ronen and Glickman 2001). La inactivación de *ATM/ATR* es un paso crucial en la promoción de la inestabilidad genómica inducida por andrógenos y en la carcinogénesis de próstata (Chiu et al. 2012); además, algunas variantes del gen *ATM* podrían conferir un incremento moderado en el riesgo de padecer cáncer de próstata. Las variables genéticas de *ATM* se han asociado con la toxicidad inducida por radiación (Cesaretti et al. 2007; Barnett et al. 2012), aunque ninguna de estas asociaciones se ha confirmado en estudios de validación (Barnett et al. 2012). Se ha observado una tendencia similar en el contexto de la agresividad tumoral (Browning et al. 2006). En nuestro trabajo, los pacientes con CaP que portaban el genotipo TT – rs17503908 presentaron mayor riesgo de desarrollar tumores de alto grado. De este modo, la presencia del alelo G, el cual no es el alelo ancestral, podría conferir ventajas clínicas en términos de *Gleason score*. Aunque este es un resultado novedoso en CaP, existe una gran variedad de mutaciones tipo *missense* y variantes de *ATM* entre los pacientes de cáncer de mama, pudiendo contribuir algunas de ellas en la etiología y progresión de la malignidad de la enfermedad (Dork et al. 2001).

Dado que los genotipos salvajes parecen representar un factor de riesgo asociado con la agresividad tumoral, evaluamos el papel combinado de ambos genotipos en relación con los grupos de riesgo establecidos por D'Amico. Observamos que los pacientes que portaban los genotipos homocigotos de ambos polimorfismos (AA + TT) presentaron mayor riesgo de desarrollar tumores de alto riesgo de acuerdo con esta clasificación. Los genes *ERCCI* y *ATM* están codificados en cromosomas diferentes y por tanto no hay una combinación conservada de los polimorfismos. Sin embargo, las frecuencias genotípicas para AA – rs11615 y TT – rs17503908 fueron de 0.39 y 0.81 respectivamente, con un total de 148 pacientes de CaP portadores de ambos genotipos (29.9% del total de la serie). En consecuencia, un análisis combinado podría dar una idea sobre el papel de ambos polimorfismos como factores de predicción de la

agresividad tumoral. No obstante, debe tenerse en consideración que las consecuencias funcionales de rs11615, así como rs17503908, no son de tipo *missense*, esto es, no se produce una sustitución de aminoácidos en la traducción proteica, por lo que la relevancia biológica de este tipo de alteraciones genéticas no es clara. A pesar de ello, las evidencias sugieren que las mutaciones sinónimas también son importantes y ya existe una lista de mutaciones sinónimas que crece con rapidez y que se asocian al desarrollo de enfermedades humanas (Cartegni et al. 2002; Chamary et al. 2006). Aunque de manera más sutil que en el caso de las mutaciones tipo *missense*, los SNPs sinónimos pueden afectar a la función de la proteína alterando las estructuras secundarias del ARN y en consecuencia, reduciendo la expresión proteica (Nackley et al. 2006) y posiblemente afectando el plegamiento y la función de dicha proteína (Kimchi-Sarfaty et al. 2007). Por ello, es posible que ambos polimorfismos puedan ser importantes en la determinación de las características tumorales del CaP.

Los estudios de asociación de genes candidatos son a menudo criticados por su falta de validación. Para ello debe considerarse en un futuro próximo la realización de estudios de validación, utilizando una cohorte de pacientes de CaP independiente y *randomizada* (Patnala et al. 2013). Del presente estudio deben subrayarse algunos factores limitantes: i) aunque 494 pacientes con CaP pueden parecer suficientes para obtener datos estadísticos de confianza, es posible que algunos resultados tengan naturaleza estocástica, especialmente para aquellos SNPs con baja MAF; ii) no se han tenido en cuenta otros factores asociados con el CaP, tales como la edad, la historia familiar, hábitos tóxicos o algunos tipos de dieta y iii) estas observaciones deben ser confirmadas en una cohorte independiente. A pesar de lo anterior, este estudio también presenta una serie de ventajas que contribuyen a su credibilidad: i) es un estudio multicéntrico que proporciona pacientes de diferentes regiones del país, minimizando los sesgos que se producen en los estudios llevados a cabo en un único hospital; ii) todos los sujetos fueron de origen español y en la selección de individuos se ha reducido la posible influencia de las variaciones intra-étnicas (Henriquez-Hernandez et al. 2013) y iii) todas las determinaciones (4940 en total) fueron llevadas a cabo con la misma metodología (OpenArray, Applied Biosystems), con el mismo lote de material y por el mismo investigador, minimizando así los sesgos de origen técnico.

7.4.- Conclusión

Los pacientes con cáncer de próstata portadores de los genotipos homocigotos salvajes para rs11615 o rs17503908 presentaron variables clínicas de peor pronóstico, un hecho acentuado en aquellos pacientes homocigotos para ambos polimorfismos.

Capítulo IV

8.- Capítulo IV: Asociación entre polimorfismos de nucleótido simple en genes de reparación de rotura doble de cadena del ADN y agresividad del cáncer de próstata en la población española.

8.1.- Introducción

El cáncer de próstata (CaP) es una enfermedad compleja, condicionada por factores hormonales y genéticos. Tanto el estadio tumoral como el grado tumoral (medido en términos de *Gleason score*) y los niveles de PSA al diagnóstico son utilizados en la rutina de la práctica clínica diaria con el fin de clasificar a los pacientes en distintos grupos de riesgo, lo que condiciona la decisión sobre el tratamiento de la enfermedad. Sin embargo, estos factores de pronóstico clínico solo son capaces de explicar una parte de la variabilidad en los resultados clínicos que se observan entre los pacientes (Fraser et al. 2015), haciéndose necesaria la introducción de nuevos factores que mejoren la predicción del pronóstico y de la respuesta al tratamiento de manera individualizada.

La rotura doble de cadena de ADN (*DNA double strand break*, DSB) es la forma más perjudicial de lesión en esta molécula, conduciendo a las células a la rotura de sus cromosomas y condicionando el posterior reordenamiento cromosómico (Khanna and Jackson 2001). Las DSBs pueden tener un origen exógeno (mediante la exposición a radiación ionizante) o endógeno (exposición a especies reactivas de oxígeno); también pueden originarse durante la recombinación somática, la replicación del ADN y cuando se producen roturas de cadena simple (*DNA single strand breaks*) (Khanna and Jackson 2001). Las DSBs no reparadas o reparadas erróneamente pueden provocar muerte celular, inestabilidad genómica y/o transformación oncogénica (Sharpless et al. 2001). En células eucariotas, las DSBs se reparan mediante dos mecanismos diferentes: la recombinación homóloga (*homologous recombination*, HR) y la unión no homóloga de

extremos de ADN (*non-homologous end joining*, NHEJ) (Shrivastav et al. 2008). Algunos polimorfismos de nucleótido simple (SNP) – definidos como mutaciones hereditarias que están presentes en al menos el 1% de una población – en genes que codifican para proteínas implicadas en la HR se han propuesto como nuevos factores asociados con la agresividad del CaP (Henriquez-Hernandez et al. 2014). Debido a la complejidad del genoma de los mamíferos, la HR no es un mecanismo demasiado eficiente ya que encontrar una secuencia homóloga a un fragmento determinado de ADN es un proceso muy complejo. Es por ello que la NHEJ juega un papel predominante en la reparación de las DSBs (West 2003). Esta afirmación tiene sentido si se tiene en cuenta que la mayoría del genoma humano está formado por secuencias no codificantes. Por todo lo anterior, algunas de las variaciones genéticas más comunes en aquellos genes que codifican para proteínas involucradas en la reparación de DSBs de ADN son buenas candidatas como biomarcadores de baja penetrabilidad del cáncer (Tseng et al. 2009).

El mecanismo de reparación no homóloga comienza con el reclutamiento de las proteínas Ku70 y Ku80 que se unen a los extremos del ADN. La proteína Ku70 está codificada en el gen *XRCC6* (*X-ray repair complementing defective repair in Chinese hamster cell 6*), mientras que la proteína Ku80 está codificada en el gen *XRCC5* (*X-ray repair complementing defective repair in Chinese hamster cell 5*). Existen otras proteínas que participan en la unión de los extremos del ADN (tales como la proteína quinasa de ADN (DNA-PK), *XRCC4* y ligasa 4 (*LIG4*)). No obstante, el buen funcionamiento del complejo Ku70/Ku80 es crítico al inicio del mecanismo de reparación no homóloga (Bassing and Alt 2004). Sin embargo, no se han realizado estudios en CaP sobre las variantes polimórficas de los genes relacionados con la recombinación no homóloga y su potencial asociación con variables clínicas de agresividad tumoral y pronóstico de la enfermedad.

Las partículas conocidas como *Vaults*, son partículas ribonucleoproteicas con una estructura en forma de barril (Kedersha et al. 1991) formadas por tres proteínas: *major vault protein* (MVP), *vault poly (adenosine diphosphate-ribose) polymerase* (VPARP) y *telomerase-associated protein 1* (TEP1), junto con pequeños fragmentos de ARN no traducidos (vRNA). Clásicamente, MVP se ha asociado con el desarrollo de resistencia a fármacos (Scheffer et al. 1995). Sin embargo, en los últimos años ha crecido el interés en esta partícula, especialmente en cómo los niveles de transcripción

del gen *MVP* y de la proteína en sí, se ven incrementados en respuesta a agentes lesivos del ADN, incluyendo la radiación ionizante (Shimamoto et al. 2006). Además de su implicación en la regulación de varios procesos celulares, incluyendo mecanismos de transporte, transducción de señales y respuesta inmunitaria (Lara et al. 2011), su interacción con otras proteínas como PTEN o Bcl-2 sugiere que MVP interviene de alguna forma en el destino celular (Chung et al. 2005; Ryu et al. 2008). Es más, MVP se ha relacionado recientemente con la maquinaria de reparación de DSBs a través de su asociación con la expresión de Ku70/Ku80 (Lloret et al. 2009). Es por ello que SNPs localizados en el gen *MVP* puedan influir en el comportamiento del CaP, bien por su papel en la resistencia al tratamiento, la regulación de la apoptosis o con los mecanismos de reparación de ADN (Lara et al. 2011).

La *poly (ADP-ribose) sintetase 1* es una proteína codificada por el gen *PARP1*, que participa en la diferenciación, proliferación y transformación tumoral (Thomas and Tulin 2013). La proteína PARP1 interviene en la reparación de DSBs de ADN mediante su interacción con BRCA1/2, así como con otras proteínas (Schultz et al. 2003). PARP1 y Ku compiten por la reparación de DSBs del ADN mediante diferentes mecanismos de reparación no homóloga (Wang et al. 2006). En los últimos años, el empleo de inhibidores de PARP1 en pacientes con cáncer de ovario y mama con mutación en los genes *BRCA1/2* ha mostrado unos resultados prometedores (Fong et al. 2009), así como en pacientes con CaP (Zhang 2014). El hecho de que PARP1 participe como nexo entre los mecanismos de reparación homóloga y no homóloga del ADN, así como su creciente interés como diana terapéutica, hacen de este gen un buen candidato para ser estudiado como condicionante de la agresividad tumoral. En la actualidad, no se ha estudiado el papel de SNPs localizados en este gen en el contexto del CaP.

El objetivo de este estudio ha sido explorar la hipótesis de que las variaciones de genes involucrados directa (*XRCC6* y *XRCC5*) o indirectamente (*MVP* y *PARP1*) con la reparación de DSBs del ADN confieren un incremento en el riesgo de presentar un CaP más agresivo. Según nuestro conocimiento, esta es la primera vez que se estudian SNPs localizados en estos genes en relación con el CaP.

8.2.- Resultados

8.2.1.- Variación intra-étnica de SNPs

Todas las muestras genotipadas cumplieron con los requisitos de calidad expuestos en el apartado “Pacientes y Métodos” de esta Tesis Doctoral. Se genotiparon un total de 22 SNPs en 601 pacientes (Tabla 6). Del total de las 13 222 posibles determinaciones, 12 608 fueron genotipadas satisfactoriamente (95%).

Como se ha expuesto en los Capítulos I y II, las diferencias en la distribución de genotipos dentro de las diferentes poblaciones de una misma etnia son un importante factor de confusión en los estudios de epidemiología genética (Henriquez-Hernandez et al. 2013; Henriquez-Hernandez et al. 2015). Por ello, en un primer lugar analizamos la distribución genotípica entre sujetos de diferentes regiones de España, resaltando que todos los pacientes de CaP de este estudio son de origen caucásico. Las frecuencias genotípicas y alélicas se muestran en la tabla 15. Se observa que 9 de los 22 SNPs analizados presentan una distribución diferente de acuerdo a la región geográfica a la que pertenecen: rs5751131, rs2267437, rs7291732, rs881092 (*XRCC6*); rs1051677 (*XRCC5*); rs3219123, rs1805410, rs1805414 (*PARP1*) y rs4788186 (*MVP*). Tal y como se ha observado para otros SNPs (ver Capítulos I y II), la población de Andalucía mostró una distribución genotípica y alélica significativamente diferente de las observadas en las poblaciones del País Vasco, Canarias y Cataluña. Es por ello que los pacientes de Andalucía se excluyeron del resto de los análisis con el objetivo de homogeneizar la serie y disminuir los sesgos. Hemos seguido esta estrategia previamente en el Capítulo III.

8.2.2.- Variables clínicas asociadas a los SNPs individualmente

Por las razones anteriormente comentadas, en los subsiguientes análisis se incluyeron un total de 494 pacientes de CaP. La distribución de las variables clínicas aparece detallada en la tabla 12. De manera sucinta, observamos que la mayoría de los pacientes de CaP presentaban un tamaño tumoral clínico cT1a – cT2a (54.7%), un nivel de PSA al diagnóstico < 10 ng ml⁻¹ (61.9%) y un *Gleason score* < 7 (45.7%). Así, de los 494 pacientes incluidos, 304 (61.5%) se clasificaron como tumores de riesgo bajo-intermedio de acuerdo con la clasificación de D’Amico.

Tabla 15. Frecuencias alélicas y genotípicas de los genes considerados en las distintas poblaciones en estudio. Se muestran las diferencias estadísticas entre los diferentes genotipos

	% determinación	Genotipos			Alelos	
XRCC6						
rs5751131		AA	AG	GG	A	G
Andalucía	0.97	0.78	0.22	0.00	0.89	0.11
País Vasco	1.00	1.00	0.00	0.00	1.00	0.00
Canarias	0.99	0.99	0.01	0.00	0.99	0.01
Cataluña	0.99	1.00	0.00	0.00	1.00	0.00
<i>p value</i>				0.0001		
rs2267437		CC	CG	GG	C	G
Andalucía	0.70	0.53	0.25	0.22	0.66	0.34
País Vasco	0.98	0.40	0.34	0.26	0.57	0.43
Canarias	0.86	0.42	0.34	0.24	0.59	0.41
Cataluña	0.94	0.40	0.47	0.13	0.63	0.37
<i>p value</i>				0.008		
rs11912946		CC	CT	TT	C	T
Andalucía	0.93	0.89	0.11	0.00	0.95	0.05
País Vasco	0.92	0.96	0.04	0.00	0.98	0.02
Canarias	0.96	0.96	0.04	0.00	0.98	0.02
Cataluña	0.98	0.94	0.06	0.00	0.97	0.03
<i>p value</i>				n.s.		
rs7291732		AA	AC	CC	A	C
Andalucía	0.92	0.00	0.12	0.88	0.06	0.94
País Vasco	1.00	0.00	0.00	1.00	0.00	1.00
Canarias	0.99	0.00	0.00	1.00	0.00	1.00
Cataluña	0.99	0.00	0.00	1.00	0.00	1.00
<i>p value</i>				0.0001		
rs881092		AA	AC	CC	A	C
Andalucía	0.91	0.00	0.12	0.88	0.06	0.94
País Vasco	1.00	0.00	0.02	0.98	0.01	0.99
Canarias	0.99	0.00	0.01	0.99	0.01	0.99
Cataluña	0.99	0.00	0.03	0.97	0.01	0.99
<i>p value</i>				0.0001		
XRCC5						
rs16855458		AA	AC	CC	A	C
Andalucía	0.73	0.09	0.20	0.71	0.19	0.81
País Vasco	0.96	0.02	0.45	0.53	0.24	0.76
Canarias	0.98	0.03	0.36	0.61	0.21	0.79
Cataluña	0.96	0.05	0.32	0.63	0.21	0.79
<i>p value</i>				n.s.		
rs9288516		AA	AT	TT	A	T
Andalucía	0.84	0.00	0.01	0.99	0.01	0.99
País Vasco	1.00	0.00	0.00	1.00	0.00	1.00
Canarias	0.99	0.01	0.05	0.94	0.03	0.97
Cataluña	0.98	0.00	0.05	0.95	0.02	0.98
<i>p value</i>				n.s.		
rs1051677		CC	CT	TT	C	T
Andalucía	0.81	0.07	0.15	0.78	0.14	0.86
País Vasco	0.86	0.02	0.12	0.86	0.08	0.92
Canarias	1.00	0.00	0.22	0.78	0.11	0.89
Cataluña	0.98	0.01	0.21	0.78	0.11	0.89
<i>p value</i>				0.001		

rs1051685		AA	AG	GG	A	G
Andalucía	0.87	0.81	0.14	0.05	0.88	0.12
País Vasco	1.00	0.84	0.14	0.02	0.91	0.09
Canarias	0.99	0.82	0.15	0.03	0.90	0.10
Cataluña	0.99	0.80	0.19	0.01	0.90	0.10
<i>p value</i>				n.s.		
PARP1						
rs8679		AA	AG	GG	A	G
Andalucía	0.83	0.75	0.18	0.07	0.84	0.16
País Vasco	0.96	0.55	0.39	0.06	0.74	0.26
Canarias	0.96	0.59	0.33	0.08	0.75	0.25
Cataluña	0.95	0.60	0.34	0.06	0.77	0.23
<i>p value</i>				n.s.		
rs3219123		AA	AG	GG	A	G
Andalucía	0.82	0.03	0.01	0.96	0.03	0.97
País Vasco	1.00	0.02	0.16	0.82	0.10	0.90
Canarias	1.00	0.01	0.06	0.93	0.04	0.96
Cataluña	0.98	0.00	0.16	0.86	0.07	0.93
<i>p value</i>				0.001		
rs1805410		CC	CT	TT	C	T
Andalucía	0.77	0.10	0.23	0.67	0.21	0.79
País Vasco	0.98	0.00	0.30	0.70	0.15	0.85
Canarias	0.98	0.03	0.30	0.67	0.17	0.83
Cataluña	0.97	0.03	0.33	0.64	0.20	0.80
<i>p value</i>				0.047		
rs3219062		GG	GT	TT	G	T
Andalucía	0.82	1.00	0.00	0.00	1.00	0.00
País Vasco	1.00	0.98	0.02	0.00	0.99	0.01
Canarias	0.99	0.99	0.01	0.00	0.99	0.01
Cataluña	0.99	0.99	0.01	0.00	0.99	0.01
<i>p value</i>				n.s.		
rs1805414		GG	GA	AA	A	G
Andalucía	0.75	0.30	0.23	0.47	0.59	0.41
País Vasco	0.94	0.17	0.42	0.41	0.63	0.37
Canarias	0.98	0.09	0.38	0.53	0.72	0.28
Cataluña	0.92	0.10	0.42	0.48	0.69	0.31
<i>p value</i>				0.0001		
rs1805404		AA	AG	GG	A	G
Andalucía	0.77	0.04	0.16	0.80	0.12	0.88
País Vasco	1.00	0.02	0.29	0.69	0.17	0.83
Canarias	0.99	0.01	0.20	0.79	0.11	0.89
Cataluña	0.99	0.02	0.19	0.79	0.11	0.89
<i>p value</i>				n.s.		
rs3219027		AA	AG	GG	A	G
Andalucía	0.82	0.07	0.27	0.69	0.19	0.81
País Vasco	1.00	0.06	0.33	0.61	0.23	0.77
Canarias	1.00	0.04	0.26	0.70	0.17	0.83
Cataluña	0.99	0.02	0.30	0.68	0.17	0.83
<i>p value</i>				n.s.		
MVP						
rs9923649		GG	GA	AA	G	A
Andalucía	0.74	0.81	0.16	0.03	0.89	0.11
País Vasco	0.98	0.76	0.22	0.02	0.87	0.13
Canarias	0.98	0.73	0.25	0.02	0.85	0.15
Cataluña	0.95	0.74	0.24	0.02	0.85	0.15

<i>p value</i>					n.s.	
rs12149514		AA	AG	GG	A	G
Andalucía	0.76	1.00	0.00	0.00	1.00	0.00
País Vasco	0.96	0.92	0.08	0.00	0.96	0.04
Canarias	0.99	0.96	0.04	0.00	0.98	0.02
Cataluña	0.94	0.96	0.04	0.00	0.98	0.02
<i>p value</i>					n.s.	
rs4788186		GG	GA	AA	G	A
Andalucía	0.76	0.48	0.20	0.32	0.58	0.42
País Vasco	1.00	0.35	0.45	0.20	0.58	0.42
Canarias	0.99	0.35	0.49	0.16	0.59	0.41
Cataluña	0.97	0.44	0.42	0.14	0.65	0.35
<i>p value</i>					0.001	
rs3815824		TT	TC	CC	T	C
Andalucía	0.78	0.04	0.17	0.79	0.13	0.87
País Vasco	0.98	0.02	0.20	0.78	0.12	0.88
Canarias	0.98	0.02	0.24	0.74	0.14	0.86
Cataluña	0.98	0.02	0.23	0.75	0.14	0.86
<i>p value</i>					n.s.	
rs35916172		GG	GA	AA	G	A
Andalucía	0.85	1.00	0.00	0.00	1.00	0.00
País Vasco	1.00	1.00	0.00	0.00	1.00	0.00
Canarias	1.00	1.00	0.00	0.00	1.00	0.00
Cataluña	0.99	1.00	0.00	0.00	1.00	0.00
<i>p value</i>					n.s.	
rs3764944		GG	GA	AA	G	A
Andalucía	0.81	1.00	0.00	0.00	1.00	0.00
País Vasco	1.00	1.00	0.00	0.00	1.00	0.00
Canarias	0.99	0.99	0.01	0.00		
Cataluña	0.99	1.00	0.00	0.00	1.00	0.00
<i>p value</i>					n.s.	
<i>Abreviaturas:</i> n.s., no significativo.						
Nota: diferencias en la distribución genotípica determinadas por el test χ^2						

De los 22 SNPs analizados, 4 de ellos mostraron una distribución significativamente diferente entre los pacientes de CaP de acuerdo con las variables clínicas consideradas (Tabla 16): rs2267437 (*XRCC6*), rs1051685 (*XRCC5*), rs3219123 (*PARP1*) y rs3815824 (*MVP*). La distribución genotípica de dichos SNPs estaba en equilibrio de Hardy-Weinberg.

Tabla 16. Asociaciones significativas entre las variables clínicas y los SNPs. La distribución de los pacientes está expresada en números absolutos y en porcentajes (entre paréntesis)

SNP (gen)	Genotipos	Tamaño clínico tumoral (cT)			p#
		cT1a – cT2a	cT2b – cT2c	cT3 – cT4	
rs2267437 (<i>XRCC6</i>)	CC	105 (24.3)	41 (9.5)	29 (6.7)	0.004
	CG	107 (24.8)	45 (10.4)	20 (4.6)	
	GG	37 (8.6)	38 (8.8)	10 (2.3)	
		PSA al diagnóstico (ng/mL)			
		<10	10 – 19.99	≥20	
rs1051685 (<i>XRCC5</i>)	AA	253 (52.4)	73 (15.1)	71 (14.7)	0.007
	AG	46 (9.5)	25 (5.2)	6 (1.2)	
	GG	3 (0.6)	4 (0.8)	2 (0.4)	
rs3219123 (<i>PARP1</i>)	AA	1 (0.2)	0 (0.0)	1 (0.2)	0.041
	AG	29 (6.0)	18 (3.7)	4 (0.8)	
	GG	274 (56.6)	84 (17.4)	73 (15.1)	
rs3815824 (<i>MVP</i>)	TT	4 (0.8)	2 (0.4)	5 (1.0)	0.025
	TC	61 (12.7)	29 (6.0)	18 (3.8)	
	CC	238 (49.6)	69 (14.4)	54 (11.3)	
		Gleason score			
		<7	7	>7	
rs2267437 (<i>XRCC6</i>)	CC	91 (20.4)	58 (13.0)	31 (7.0)	0.037
	CG	82 (18.4)	78 (17.5)	19 (4.3)	
	GG	31 (7.0)	41 (9.2)	14 (3.1)	
		Grupo de riesgo D'Amico			
		Bajo	Intermedio	Alto	
rs2267437 (<i>XRCC6</i>)	CC	52 (12.0)	58 (13.4)	64 (14.8)	0.015
	CG	46 (10.6)	76 (17.6)	52 (12.0)	
	GG	12 (2.8)	34 (7.9)	39 (9.2)	

Test de χ^2 .

De los SNPs estudiados, se encontró que rs2267437 estaba significativamente asociado con la variable cT (test χ^2 , $p = 0.004$). Así, 37 de los 249 pacientes de CaP con tumores cT1a – cT2a eran portadores del genotipo GG (14.8%), mientras que de los 183 que presentaban tumores cT2b – cT4, 48 portaban el genotipo GG (26.2%). Los mismos resultados se observaron en relación con las variables *Gleason score* y los grupo de riesgo según D'Amico (test χ^2 , $p = 0.037$ y $p = 0.015$, respectivamente) (Tabla 16). En ambos casos, los portadores del genotipo GG presentaron unas características tumorales de mayor agresividad. En esta línea, el genotipo GG se detectó en 12 de los 110 pacientes con tumores de riesgo bajo, mientras que apareció en 73 de los 323 pacientes con tumores de riesgo intermedio-alto (10.9% vs. 22.6%, respectivamente). En los análisis univariantes, el genotipo GG del polimorfismo rs2267437 (*XRCC6*) aparece como un factor de riesgo en relación con el desarrollo de tumores de próstata más agresivos (Tabla 17). El hecho de que los modelos recesivos y homocigotos muestren resultados significativos, refuerza el papel del genotipo GG como factor pronóstico de

la presencia de características tumorales más desfavorables. En este sentido, aquellos pacientes de CaP portadores del genotipo GG presentan mayor riesgo de desarrollar tumores de mayor tamaño (OR = 2.04 (CI 95% = 1.26 – 3.29), $p = 0.004$), presentan mayores niveles de PSA al diagnóstico (OR = 2.12 (CI 95% = 1.19 – 3.78), $p = 0.044$) y presentan tumores de mayor riesgo según D'Amico (OR = 2.38 (CI 95% = 1.24 – 4.58), $p = 0.0009$) que aquellos pacientes que porten los genotipos CC/CG (Tabla 17).

Tabla 17. Análisis univariante del SNP rs2267437 (*XRCC6*) en relación a las variables clínicas

Variables clínicas	n	Modelo recesivo		Modelo dominante		Modelo homocigoto		Modelo heterocigoto	
		GG vs. CG + CC		CG + GG vs. CC		CC vs. GG		CG vs. CC	
		OR (95%CI)	p	OR (95%CI)	p	OR (95%CI)	p	OR (95%CI)	p
cT, 1a-2a/2b-4	270/207	2.04 (1.26 – 3.29)	0.004	1.18 (0.80 – 1.74)	0.413	1.95 (1.15 – 3.29)	0.013	0.91 (0.59 – 1.40)	0.673
PSA, <20/≥20	409/79	2.12 (1.17 – 3.78)	0.011	1.30 (0.75 – 2.21)	0.355	2.07 (1.08 – 3.98)	0.029	0.96 (0.52 – 1.77)	0.892
Gleason, <7/≥7	226/266	1.65 (1.01 – 2.68)	0.044	1.37 (0.94 – 2.01)	0.100	1.81 (1.07 – 3.08)	0.027	1.21 (0.80 – 1.83)	0.368
Grupo riesgo, Bajo/Int-Alto	120/357	2.38 (1.24 – 4.58)	0.009	1.48 (0.95 – 2.29)	0.080	2.59 (1.30 – 5.18)	0.007	1.19 (0.74 – 1.89)	0.475

Abreviaturas: OR, odds ratio; CI, intervalo de confianza; Int, intermedio.
Test estadístico: regresión logística binaria (Categoría de referencia: CC).

Tal y como se observa en la tabla 16, SNPs localizados en diferentes genes estaban relacionados con los niveles de PSA al diagnóstico. Entre los 404 pacientes que mostraron niveles de PSA al diagnóstico < 20 ng ml⁻¹, 78 de ellos eran portadores del alelo G en el polimorfismo rs1051685 (*XRCC5*) (19.3%), mientras que 8 de los 79 pacientes con niveles iniciales de PSA ≥ 20 ng ml⁻¹ portaban el alelo G (10.1%) (test χ^2 , $p = 0.007$). Estos resultados sugieren que el alelo G actuaría como factor protector, reforzado este hecho por los resultados obtenidos en el análisis univariante (datos no mostrados en tabla). Así, aquellos pacientes de CaP portadores del genotipo AG mostraron menor riesgo de presentar niveles altos de PSA al diagnóstico (OR = 0.388 (CI 95% = 0.16 – 0.93), $p = 0.033$) que aquellos pacientes portadores del genotipo AA. Por el contrario, los pacientes portadores del genotipo TT para el polimorfismo rs3815824 (*MVP*) mostraron mayor riesgo de presentar niveles altos de PSA al diagnóstico (OR = 4.74 (CI 95% = 1.40 – 16.07), $p = 0.013$) que aquellos pacientes portadores del genotipo CC. Finalmente, aunque el polimorfismo rs3219123 (*PARP1*) estaba significativamente asociado con los niveles de PSA al diagnóstico (test χ^2 , $p = 0.041$, Tabla 16), este SNP no parece ser un factor de riesgo para esta variable en el análisis univariante.

8.2.3.- Análisis combinados de múltiples SNPs

Los resultados obtenidos sugieren que los genes considerados en el presente estudio pueden contribuir al comportamiento tumoral, debido tanto a interacción directa como indirecta con el mecanismo de reparación no homólogo. Se estimó el efecto

combinado de los genes asociados significativamente con las diferentes variables clínicas (Figura 18). Para ello, consideramos el polimorfismo rs2267437 (*XRCC6*) como la base de este análisis, debido a su alto impacto en las diferentes características tumorales (tal y como se muestra en las tablas 16 y 17). Aunque la asociación con *XRCC5* no fue significativa (OR = 2.11 (CI 95% = 0.91 – 4.94)), sí lo fue la asociación con los polimorfismos rs3219123 (*PARP1*) o rs3815824 (*MVP*), siendo esta un factor de riesgo asociado a los grupos de riesgo según D’Amico. De este modo, los pacientes de CaP portadores de genotipos diferentes a los considerados como referencia (CC + CC y CC + GG para rs2267437 + rs3815824 y rs2267437 + rs3219123, respectivamente) presentaron mayor riesgo de desarrollar tumores de riesgo intermedio-alto según D’Amico (Figura 18). Sorprendentemente, la combinación de estos tres SNPs no mostró un incremento significativo en el riesgo, sugiriendo que rs2267437 (*XRCC6*) y rs3815824 (*MVP*) tienen un papel predominante en el comportamiento del CaP (OR = 3.33 (CI 95% de 1.56 – 7.17), $p = 0.002$).

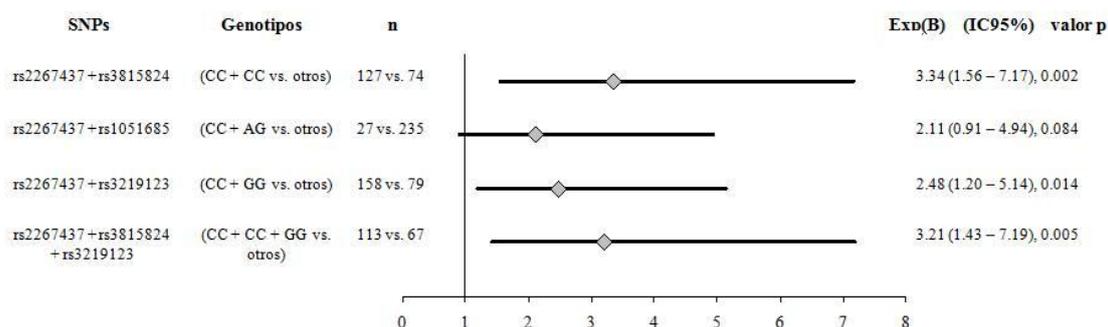


Figura 18, diagrama Forrest de los *odds ratios* (OR) con un 95% de intervalo de confianza (IC) para los polimorfismos rs2267437 (*XRCC6*), rs3815824 (*MVP*), rs1051685 (*XRCC5*) y rs3219123 (*PARP1*) analizados de forma combinada y los grupos de riesgo de recaída según D’Amico. Cada rombo representa el OR y la línea horizontal indica el 95% de IC. Para la regresión binaria logística, los pacientes fueron dicotomizados en dos grupos según el siguiente criterio: riesgo bajo vs. riesgo intermedio-alto según la clasificación D’Amico.

8.3.- Discusión

Aunque algunos alelos polimórficos de los genes relacionados con la reparación no homóloga podrían predisponer a sus portadores a un mayor riesgo de desarrollar tumores más agresivos, hasta la fecha no se han obtenido evidencias clínicas directas de que esto sea así en pacientes con CaP. En capítulos anteriores de esta Tesis Doctoral, se

ha reportado que las variaciones genéticas más comunes en genes involucrados en la reparación homóloga (como *XRCC1*, *ERCC1* o *ERCC2*, entre otros), están asociadas con variables clínicas de mal pronóstico para el CaP (Henriquez-Hernandez et al. 2014). En este capítulo, nos centramos en el estudio de los genes clásicos involucrados en la reparación no homóloga del daño al ADN (*XRCC6* y *XRCC5*), así como de otros genes asociados con la reparación de DSBs (*PARP1* y *MVP*). Hasta donde nosotros sabemos, este es el primer estudio que analiza la influencia de estos cuatro genes en el comportamiento tumoral del CaP. Aunque es un estudio de tipo “gen candidato”, en este caso se utilizó un enfoque de selección de SNPs haciendo uso de la plataforma informática HapMap, lo que reduce los sesgos inherentes a la selección manual de los SNPs ampliamente estudiados y permite la inclusión de nuevas variantes génicas con posibles papeles biológicos.

En este estudio se incluyeron un total de 601 pacientes de cáncer de próstata de cuatro regiones diferentes de España. Las variaciones intra-étnicas en las frecuencias genómicas y alélicas son un importante factor de confusión que es a menudo obviado. La confusión puede aparecer si la población contiene varios grupos étnicos, si las frecuencias alélicas en el *locus* de interés difieren entre grupos o si la frecuencia de la enfermedad también difiere entre los grupos por causas no relacionadas con el *locus* de interés (Clayton 2001). Este fenómeno ha sido descrito en nuestra serie de pacientes (Henriquez-Hernandez et al. 2013; Henriquez-Hernandez et al. 2015), un efecto explicado por las diferencias en el devenir de la historia natural de las distintas poblaciones (Calafell and Bertranpetit 1994; Rando et al. 1999). Tal y como era de esperar, a la vista de los resultados reportados en los Capítulos I y II de esta Tesis Doctoral, la población de Andalucía mostró diferencias significativas en la distribución genotípica y alélica, y por tanto, esa *subserie* de pacientes fue excluida de los análisis de asociación con el fin de eliminar los sesgos de tipo intra-étnicos.

Ku70/Ku80 es un complejo proteico codificado por los genes *XRCC6* y *XRCC5*, el cual está involucrado en la iniciación del proceso de reparación no homóloga. Específicamente, Ku70/Ku80 es responsable del reconocimiento de las terminaciones de ADN en aquellos lugares donde se ha producido una lesión de tipo DSB (Bassing and Alt 2004; Rivera-Calzada et al. 2007; Gomes et al. 2010). Debido a que Ku70/Ku80 es un complejo formado por dos monómeros, la funcionalidad del complejo puede verse comprometida si se producen cambios estructurales en las proteínas, y es bien conocido

que los polimorfismos genéticos pueden comprometer no solo la unión de un complejo proteico sino también la fuerza de los enlaces entre proteínas y ADN (Gomes et al. 2010). El polimorfismo rs2267437 de *XRCC6* aparece como un factor de riesgo asociado a variables clínicas de mal pronóstico para el CaP en términos de tamaño tumoral, nivel de PSA al diagnóstico, *Gleason score* e incluso con los grupos de riesgo según la clasificación de D'Amico. Se ha investigado el polimorfismo rs2267437 para establecer su papel como factor de riesgo de desarrollar cáncer con resultados inconclusos. Sin embargo, una revisión sistemática reciente y un meta-análisis han concluido que dicho SNP está asociado con un aumento significativo del riesgo de sufrir cáncer de mama, carcinoma de células renales y carcinoma hepatocelular (Jia et al. 2015). No obstante, este estudio no incluyó ningún paciente con CaP, ni tampoco se analizó ninguna asociación con las variables clínicas. También parece haberse establecido una relación con el riesgo de desarrollo tumoral para el polimorfismo rs2267437 – específicamente el genotipo GG – en otro meta-análisis (Xu et al. 2013). Nuestros resultados refuerzan estos hallazgos anteriores y postulan al SNP rs2267437 como un factor genético novedoso en relación al comportamiento biológico del CaP.

El polimorfismo rs1051685 del gen *XRCC5* ha sido previamente evaluado en pacientes con mieloma. Aquellos pacientes portadores del genotipo GG presentaron menor riesgo de desarrollar mieloma (Hayden et al. 2007). Estos resultados se han probado también en pacientes españoles con mieloma múltiple (Cibeira et al. 2011) y coinciden con los resultados presentados en este estudio de pacientes con CaP. Aunque sería esperable encontrar un vínculo entre SNPs localizados en el gen *XRCC5* y *XRCC6*, cuando los polimorfismos rs2267437 y rs1051685 son analizados de forma conjunta, no se observa dicha asociación, posiblemente debido al pequeño tamaño de la muestra de pacientes que son portadores de lo que podemos llamar este “genotipo protector” ($n = 27$). La frecuencia del alelo menor para rs1051685 fue de aproximadamente el 10%, lo cual coincide con lo determinado específicamente para la población europea (datos disponibles en <http://www.ncbi.nih.gov/projects/SNP/>). Con los datos disponibles actualmente, el papel potencial del polimorfismo rs1051685 necesita ser explorado más profundamente.

En la búsqueda de bibliografía sobre las variables genéticas de *PARP1*, no se han encontrado publicaciones previas que versen sobre el estudio de la asociación entre polimorfismos localizados en este gen y el cáncer de próstata. Sin embargo, algunos

SNPs localizados en *PARP1* parecen estar relacionados con glioma (Adel Fahmideh et al. 2014) y cáncer gastrointestinal (Hu et al. 2014), pero no hay resultados en relación al polimorfismo rs3219123 o en el contexto del CaP. En nuestra serie, se observa una asociación entre este SNP y los niveles de PSA al diagnóstico, pero los análisis univariantes no arrojaron resultados significativos. Incluso cuando se analizó este SNP en combinación con rs2267437, los resultados continuaron siendo no significativos, sugiriendo un papel poco relevante del polimorfismo rs3219123 (*PARP1*) en la biología del CaP.

Existe una única publicación que haga referencia al papel de las variables genéticas del gen *MVP*. Se trata de un artículo que tiene por objetivo explorar el papel de los SNPs en genes de resistencia a múltiples drogas en pacientes con epilepsia en el lóbulo temporal medial con esclerosis del hipocampo (Balan et al. 2013). Los autores de este estudio no observaron ningún papel de los diferentes SNPs de *MVP* – incluido rs3815824 – en este tipo de pacientes. En este aspecto, *MVP* es un gen desconocido cuyo interés se ha visto incrementado desde que fuera asociado a Ku70/Ku80 (Lloret et al. 2009) y a su papel en la respuesta a la radioterapia (Lloret et al. 2008; Henriquez-Hernandez et al. 2011; Henriquez-Hernandez et al. 2012). En este estudio se reporta por primera vez que los pacientes de CaP portadores del genotipo TT para este gen tienen mayor riesgo de desarrollar tumores de próstata de peor pronóstico en términos de nivel de PSA al diagnóstico (un factor pronóstico clásico de este tipo de cáncer). Además, en combinación con el polimorfismo rs2267437 (*XRCC6*), se incrementa el riesgo de presentar tumores de mayor riesgo al diagnóstico según la clasificación de D'Amico. Los mecanismos que están detrás de esta asociación no son aún conocidos, pero apoyan la hipótesis de la unión potencial entre *MVP* y los mecanismos de reparación no homóloga (Lloret et al. 2009).

Las consecuencias funcionales de los SNPs incluidos en este estudio son difíciles de determinar, ya que tanto rs2267437 como rs3815824 son variaciones localizadas en intrones, las cuales no se traducen como proteínas. Sin embargo, mientras que los intrones son eliminados, los exones se unen entre sí mediante una compleja modificación post-transcripcional de ARN denominada *splicing*. Este proceso de *splicing* está conducido por los elementos reguladores cis. Por ello, aunque las consecuencias que pueden tener los SNPs localizados en las regiones intrónicas de estos elementos reguladores son desconocidas, podrían influir en el *spliciosoma* (Jensen et al.

2009). Además, se han asociado algunos SNPs localizados en regiones no codificantes con algunas enfermedades humanas y, en cualquier caso, se necesita un mayor conocimiento sobre su impacto en la función génica y la salud del individuo (Shastri 2009). Para finalizar, los SNPs localizados en regiones intrónicas son considerados herramientas *in silico*, tales como “F-SNP: una colección de SNPs funcionales, priorizados específicamente para los estudios de asociación a enfermedades” (disponible en: <http://compbio.cs.queensu.c/F-SNP/>). Esto sugiere que este tipo de variables genéticas tienen relevancia biológica en las enfermedades humanas (Lee and Shatkay 2008; Lee and Shatkay 2008). En esta línea, se necesitan más estudios funcionales para conocer en profundidad las consecuencias de los polimorfismos rs2267437 y rs3815824 en relación al CaP.

Este estudio presenta una serie de ventajas que le confieren credibilidad. Primero, la gran confianza que ofrece tanto el sistema de genotipado, como los resultados de las variables clínicas consideradas (rutinarias en la práctica clínica diaria) minimizan la posibilidad de resultados inferidos debido a la subjetividad del observador. Segundo, este estudio se ha llevado a cabo en una serie de pacientes españoles de CaP que ha sido previamente homogenizada en relación a la variabilidad intra-étnica, un factor que se ha determinado como relevante en los estudios de asociación con polimorfismos. Y tercero, todas las determinaciones (13 222 en total) fueron realizadas mediante la misma metodología (OpenArray, Applied Biosystems), con el mismo lote de chips y por el mismo investigador, minimizando así los sesgos de origen tecnológico. Por otro lado, este estudio también presenta algunas deficiencias que deben ser resaltadas: i) aunque 601 pacientes de CaP parecen suficientes para obtener resultados estadísticamente creíbles, es posible que algunos resultados sean de naturaleza estocástica (error de tipo I), especialmente para esos SNPs con una frecuencia del alelo menor más baja y teniendo en cuenta que los estudios de asociación se llevaron a cabo en una serie final de 494 pacientes de CaP; ii) otros factores asociados con el CaP (como son la edad, la historia familiar, hábitos tóxicos, volumen prostático, nivel de testosterona, otros genes involucrados en la reparación de ADN o algunos tipos de dieta) no han sido tenidos en cuenta en este estudio; iii) en este estudio no hemos llevado a cabo el testado múltiple (correcciones de Bonferroni), limitando así la fortaleza de algunos de los análisis estadísticos y iv) es necesario realizar estudios

funcionales dirigidos a determinar las consecuencias biológicas de determinados SNPs para entender el papel real de estas variaciones genéticas en el CaP.

Nuestros resultados, de forma conjunta, sugieren que las variaciones genéticas en genes involucrados en procesos de reparación de DSBs y otros genes íntimamente ligados a la reparación no homóloga de ADN pueden ser relevantes para el CaP. La inhabilitación celular para responder de forma correcta o reparar los daños en el ADN conduce a una inestabilidad genética que, al final, puede incrementar la tasa de desarrollo tumoral o condicionar la agresividad del tumor. De cualquier modo, con el fin de comprender mejor las consecuencias biológicas de estas variaciones genéticas, deben tenerse en cuenta otros genes involucrados en la reparación de ADN (como *p53*, *ARM*, *BRCA1* o *BRCA2*), especialmente a medida que se acumulan evidencias que conectan las deficiencias de la respuesta celular a las DSBs de ADN y la tumorigénesis (Khanna and Jackson 2001).

8.4.- Conclusiones

En conclusión, este estudio muestra por primera vez que los polimorfismos rs2267437 (*XRCC6*) y rs3815824 (*MVP*) están asociados con las variables clínicas que identifican los tumores de CaP de mayor agresividad. Estos resultados deben ser considerados como generadores de hipótesis y por ello deben realizarse estudios adicionales con cohortes de hombres más extensas con el objetivo de validar los mismos, así como su efecto combinado. Más aún, deben realizarse estudios funcionales para comprender el papel de ambos SNPs en el CaP.

Capítulo V

9.- Capítulo V: La presencia de determinadas variaciones genéticas en genes implicados en el metabolismo de la testosterona están asociadas con la progresión del cáncer de próstata: un estudio multicéntrico español.

9.1.- Introducción

El cáncer de próstata (CaP) es una enfermedad hormono-dependiente que requiere de la presencia de andrógenos para su desarrollo (Huggins 1967). La síntesis de testosterona (T) está mediada por la enzima 17-hidroxilasa/17, 20-liasa, codificada por el gen del citocromo P450 17A1 (*CYP17A1*). Sin embargo, la hormona efectora es la dihidrotestosterona (DHT), considerada por tanto como el metabolito activo de la testosterona. La T se convierte en DHT a través de la ruta 5 α -reductasa, siendo la funcionalidad de esta ruta la que condiciona la disponibilidad tumoral de DHT y el subsecuente desarrollo y progresión tumorales (Hamada et al. 2007).

Los polimorfismos de nucleótido simple (SNPs) son variaciones genéticas en la secuencia de ADN que pueden condicionar la eficacia funcional de la enzima codificada en un gen determinado (Roses 2000). Es de esperar que las alteraciones genéticas que afecten a enzimas involucradas en la ruta metabólica de la testosterona puedan alterar la homeostasis hormonal y de este modo influir en el desarrollo y posterior progresión del CaP.

La existencia de una relación entre la presencia de ciertos polimorfismos en el gen *CYP17A1* y el cáncer de próstata es un hecho cuestionado (Lindstrom et al. 2007; Setiawan et al. 2007; Mononen and Schleutker 2009). Aunque se ha sugerido que algunos polimorfismos de *CYP17A1* están asociados con el riesgo de CaP (Wang et al. 2011) y su supervivencia (Wright et al. 2010), no está claro el mecanismo por el que SNPs localizados en este gen podrían afectar a variables clínicas del CaP tales como el tamaño tumoral, los niveles de antígeno prostático específico (PSA) o el *Gleason score*.

En cualquier caso, *CYP17A1* es considerado un factor importante para la progresión del CaP, hasta el punto de ser en la actualidad una diana terapéutica para el tratamiento de dicha enfermedad. El acetato de abiraterona es un nuevo fármaco usado en el tratamiento del CaP que actúa mediante el bloqueo específico de *CYP17A1* (Yang 2011). Sin embargo, no se ha estudiado el papel de los SNPs localizados en dicho gen en la respuesta al tratamiento con acetato de abiraterona.

Las enzimas 5 α -reductasa están codificadas por los genes esteroide-5 α -reductas, α -polipéptidos 1 (*SRD5A1*) y 2 (*SRD5A2*). Del mismo modo, el papel de los polimorfismos en *SRD5A1* y *SRD5A2* en relación con el CaP ha sido muy poco explorado, aunque existen estudios que muestran asociaciones positivas de algunas variaciones genéticas en *SRD5A1* y *SRD5A2* en la predicción del resultado del CaP, en términos de recaída bioquímica (Audet-Walsh et al. 2011). A pesar de ello, *SRD5A1/A2* también son dianas terapéuticas para el tratamiento del CaP y el bloqueo específico de estas enzimas con dutasterida y finasterida ha mostrado que la disminución de la producción de DHT mediante inhibidores de 5 α -reductasa (5 α -RI) disminuye la incidencia del CaP clínicamente localizado (Thompson et al. 2003; Andriole et al. 2004).

Dado que la presencia de variaciones genéticas en aquellos genes involucrados en la biosíntesis y metabolismo de andrógenos puede alterar la disponibilidad hormonal, en el presente estudio partimos de la hipótesis de que ciertos polimorfismos localizados en estos genes podrían ser importantes para definir las características biológicas de la enfermedad y posiblemente influir en los factores pronósticos clásicos del CaP. Por primera vez, se ha llevado a cabo un estudio para explorar la asociación entre variaciones hereditarias en *SRD5A1/A2* y *CYP17A1* y la progresión del CaP en una cohorte de pacientes caucásicos españoles con CaP.

9.2.- Resultados

Todas las muestras genotipadas cumplían con los criterios de calidad establecidos en el apartado “Pacientes y Métodos” de esta Tesis Doctoral, y todas ellas fueron genotipadas con el mismo lote de material y al mismo tiempo. Un total de 494 pacientes de CaP fueron genotipados para 32 SNPs (Tabla 6). De las 15 808 posibles determinaciones, el 97.33% fueron genotipadas con éxito. Las frecuencias genotípicas y alélicas se muestran en la tabla 18. Todos los SNPs se encontraban en equilibrio de

Hardy-Wienberg (HWE). La distribución de las variables clínicas está detallada en la tabla 12. Tal y como se ha expuesto con anterioridad, la mayoría de los pacientes eran cT1a – cT2a (54.7%), con un PSA al diagnóstico < 10ng/ml (61.9%) y un *Gleason score* < 7 (45.7%). No se observaron diferencias de proporción en la agresividad del CaP entre las diferentes regiones de España.

Tabla 18. Frecuencias alélicas y genotípicas de los pacientes incluidos en el estudio

Gen/SNP	n	Frecuencias genotípicas						Frecuencias alélicas			
SRD5A1											
rs166050	479	AA	0.59	AG	0.36	GG	0.05	A	0.77	G	0.23
rs501999	488	CC	0.28	CT	0.47	TT	0.24	C	0.52	T	0.48
rs518673	476	AA	0.09	AG	0.41	GG	0.50	A	0.29	G	0.71
rs3822430	481	AA	0.33	AG	0.51	GG	0.16	A	0.59	G	0.41
rs500182	489	GG	0.01	GT	0.18	TT	0.82	G	0.09	T	0.91
rs8192120	483	AA	0.10	AC	0.45	CC	0.46	A	0.32	C	0.68
rs4702378	490	CC	0.06	CT	0.31	TT	0.63	C	0.21	T	0.79
rs1691053	488	AA	0.87	AG	0.13	GG	0.00	A	0.93	G	0.07
rs39848	472	CC	0.14	CT	0.47	TT	0.39	C	0.38	T	0.62
rs3797179	482	AA	0.04	AG	0.26	GG	0.70	A	0.17	G	0.83
SRD5A2											
rs2208532	475	AA	0.30	AG	0.49	GG	0.21	A	0.54	G	0.46
rs12470143	488	CC	0.30	CT	0.52	TT	0.17	C	0.57	T	0.43
rs2281546	490	GG	0.02	GT	0.24	TT	0.74	G	0.14	T	0.86
rs3754838	480	CC	0.02	CT	0.20	TT	0.78	C	0.12	T	0.88
rs4952222	490	AA	0.00	AC	0.00	CC	1.00	A	0.00	C	1.00
rs7562326	483	CC	0.02	CT	0.21	TT	0.77	C	0.13	T	0.87
rs2300702	477	CC	0.18	CG	0.49	GG	0.33	C	0.42	G	0.58
rs4952197	485	AA	0.06	AG	0.33	GG	0.61	A	0.22	G	0.78
rs676033	487	CC	0.51	CT	0.40	TT	0.09	C	0.71	T	0.29
rs523349	398	CC	0.40	CG	0.43	GG	0.17	C	0.62	G	0.38
rs9332975	487	CC	0.02	CT	0.20	TT	0.78	C	0.12	T	0.88
rs7594951	489	CC	0.78	CT	0.20	TT	0.02	C	0.88	T	0.12
CYP17A1											
rs3781287	485	GG	0.21	GT	0.47	TT	0.32	G	0.45	T	0.55
rs1004467	486	AA	0.79	AG	0.20	GG	0.02	A	0.89	G	0.11
rs743572	465	AA	0.32	AG	0.48	GG	0.20	A	0.56	G	0.44
rs10883782	477	AA	0.65	AG	0.32	GG	0.04	A	0.81	G	0.20
rs619824	487	AA	0.22	AC	0.50	CC	0.28	A	0.47	C	0.53
rs2486758	482	CC	0.04	CT	0.29	TT	0.67	C	0.18	T	0.82
rs17115100	490	GG	0.80	GT	0.19	TT	0.01	G	0.89	T	0.11
rs4919686	492	AA	0.48	AC	0.41	CC	0.10	A	0.69	C	0.31
rs10786712	491	CC	0.33	CT	0.47	TT	0.20	C	0.57	T	0.43
rs6163	474	AA	0.19	AC	0.48	CC	0.32	A	0.43	C	0.57

De entre los 32 SNPs estudiados, solo dos de ellos presentaron una distribución significativamente diferente entre los pacientes de CaP en relación con los niveles de PSA (Tabla 19): rs3822430 (frecuencia del alelo menor (*minor allele frequency*, MAF) = 0.41) y rs1691053 (MAF = 0.07), ambos localizados en el gen *SRD5A1*. Así, de entre

los 157 pacientes con CaP que presentaban niveles de PSA al diagnóstico < 10ng/ml, 92 de ellos (58.6%) portaban el genotipo AA para el polimorfismo rs3822430; mientras que de los 180 pacientes con CaP con niveles de PSA al diagnóstico ≥ 10 ng/ml, 115 de ellos (63.9%) portaban el alelo G para el mismo SNP (test de χ^2 , $p = 0.015$). La misma tendencia se observó para el SNP rs1691053 (test de χ^2 , $p = 0.008$). Al estudiar el papel combinado de ambos SNPs en relación con el nivel de PSA al diagnóstico, se observó que los pacientes con CaP portadores del alelo G en ambos SNPs mostraron niveles de PSA al diagnóstico mayores que aquellos portadores AA – AA (73% vs. 27%, test de χ^2 , $p = 0.002$) (Tabla 19). Además, los portadores del alelo G para ambos SNPs rs3822430 y rs1691053 presentaban mayor riesgo de tener niveles de PSA al diagnóstico > 20 ng/ml (test de regresión logística binaria, Exp(B) = 2.812 (CI del 95% = 1.397 – 5.657), $p = 0.004$) que aquellos portadores del genotipo AA – AA (datos no mostrados). Ningún otro polimorfismo mostró asociaciones significativas con las variables clínicas.

Tabla 19. Asociaciones significativas entre el nivel de PSA al diagnóstico y SNPs

SNP	Genotipos	PSA al diagnóstico (ng/mL)			p#
		< 10	10-19.99	> 20	
rs3822430	AA	92	44	21	0.015
	AG	44	44	14	
	GG	21	50	7	
rs1691053	AA	272	89	59	0.008
	AG	30	13	18	
	GG	0	1	0	
rs3822430 - rs1691053	AA + AA	80	38	10	0.002
	G portadores	213	64	66	

Test de χ^2 .

Se estudió el papel de los tres haplotipos más frecuentes detectados en nuestra serie, localizados en los cromosomas (chr) 2 (donde se localiza el gen *SRD5A2*), chr 5 (localización del gen *SRD5A1*) y chr 10 (localización del gen *CYP17A1*) (Tabla 20). Los análisis de haplotipo se realizaron con el programa SNPator. Para el chr 2, el 70.9% de los pacientes con CaP presentaba uno de los tres haplotipos mostrados en la tabla; para el chr 5, el 54.9% presentaba uno de los tres haplotipos más frecuentes mostrados; y para el chr 10, el 62% de los pacientes presentaba uno de los tres haplotipos mostrados en la tabla. Se observó que entre los 23 pacientes homocigotos para el haplotipo GCTTGTAGTA (localizado en el chr 5), 19 de ellos (82.6%) eran cT1a – cT2a,

comparado con el 55.4% (251 pacientes de 453) observado entre los pacientes con CaP no homocigotos para este haplotipo (test de χ^2 , $p = 0.018$, datos no mostrados). Se observó una tendencia similar para la variable clínica *Gleason score* < 7 (69.6% vs. 44.9%, test de χ^2 , $p = 0.032$). Además, tal y como se muestra en la tabla 21, observamos que los pacientes con CaP no homocigotos para el haplotipo GCTTG TAGTA presentaron un mayor riesgo de desarrollar tumores con un tamaño clínico mayor (test de regresión logística binaria, $\text{Exp(B)} = 3.823$ (CI 95% = 1.280 – 11.416), $p = 0.016$). En el mismo sentido, los pacientes con CaP no homocigotos para el haplotipo GCTTG TAGTA presentaron un mayor riesgo de tener un *Gleason score* mayor (test de regresión logística binaria, $\text{Exp(B)} = 2.808$ (CI 95% = 1.134 – 6.953), $p = 0.026$). Ningún otro haplotipo en el chr 5 o en los chrs 2 y 10 mostraron asociación significativa con las variables clínicas consideradas en este estudio.

Tabla 20. Frecuencia de detección (%) de los tres haplotipos más frecuentes localizados en los cromosomas 2 (*SRD5A2*), 5 (*SRD5A1*) y 10 (*CYP17A1*)

	Cromosoma 2	frec	Cromosoma 5	frec	Cromosoma 10	frec
Hap1	TTGGACTCCTC	32.8	GCTTG TAGTA	21.6	AAGCAGTAGT	28.1
Hap2	TTCACGCTCGTT	21.0	ACCTATGGCA	18.2	CGGAATCCAT	17.9
Hap3	TTCGGGCTCCTC	11.1	GATTACAGTA	15.1	CAGAATCCAC	16.0

Abreviaturas: hap, haplotipo; frec, frecuencia.

Haplotipos en cromosoma 2 están formados por los SNPs: rs9332975, rs2281546, rs12470143, rs4952197, rs2300702, rs2208532, rs7594951, rs7562326, rs4952222, rs523349, rs3754838, rs676033.

Haplotipos en cromosoma 5 están formados por los SNPs: rs518673, rs8192120, rs501999, rs500182, rs166050, rs4702378, rs3822430, rs3797179, rs39848, rs1691053.

Haplotipos en cromosoma 10 están formados por los SNPs: rs619824, rs10883782, rs17115100, rs4919686, rs1004467, rs3781287, rs10786712, rs6163, rs743572, rs2486758.

Tabla 21. Haplotipos significativamente asociados a las variables clínicas

Haplotipo	chr	Exp(B)	C.I. 95%	p
cT1-T2a vs. cT2b-T4				
Homocigoto para GCTTG TAGTA#	5	1		
Otros haplotipos		3.823	(1.280-11.416)	0.016
<i>Gleason score</i> (<7 vs. ≥7)				
Homocigoto para GCTTG TAGTA#	5	1		
Otros haplotipos		2.808	(1.134-6.953)	0.026

Abreviaturas: chr, cromosoma; C.I., intervalo de confianza.

Categoría de referencia (Test estadístico: regresión logística binaria)

9.3.- Discusión

En el presente trabajo se ha estudiado el papel de 32 SNPs localizados en tres genes importantes para el metabolismo de la testosterona (*SRD5A1*, *SRD5A2* y *CYP17A1*), en una cohorte española de pacientes con cáncer de próstata, asumiendo que los diferentes niveles de testosterona y de dihidrotestosterona observados entre los individuos de una población podría estar influenciado por la presencia de polimorfismos hereditarios localizados en estos genes, condicionando este hecho las características clínicas del tumor, máxime teniendo en cuenta el importante papel que los andrógenos tienen en el desarrollo y comportamiento biológico del cáncer de próstata.

En esta cohorte, la presencia de ciertos SNPs en *SRD5A1* parece estar asociada con algunas variables clínicas del CaP, tales como el nivel inicial de PSA. De este modo, los portadores del alelo G para los SNPs rs3822430 y rs1691053 mostraron mayor riesgo de presentar niveles de PSA al diagnóstico > 20ng/ml. Este resultado podría tener una importante influencia en la práctica clínica rutinaria. En pacientes con CaP tratados con 5 α -RI descienden los niveles de PSA (Ross et al. 2012). Más aún, se ha sugerido que el nivel sérico de DHT podría servir como un potencial marcador de diagnóstico de la actividad de la enzima 5 α -reductasa intra-prostática durante el tratamiento de los pacientes con 5 α -RI (Stanczyk et al. 2013). Por otro lado, el efecto de los 5 α -RI en términos del nivel de PSA y DHT podría estar altamente influenciado por la presencia de determinados polimorfismos en el gen *SRD5A1*. Curiosamente, no hay estudios que hayan investigado el papel de los SNPs localizados en este gen en relación con la respuesta a 5 α -RI. Sí se conoce que SNPs específicos localizados en *SRD5A1* están asociados con baja actividad enzimática (Signorelli et al. 2008), pero en cualquier caso, el conocimiento sobre los polimorfismos de *SRD5A1* en relación con el CaP es muy limitado y contradictorio. En un estudio reciente, los varones con genotipo AG o GG en el polimorfismo rs1691053, en contraposición a los que portaban el genotipo AA en dicho polimorfismo, presentaban mayor riesgo de desarrollar CaP (Setlur et al. 2010), lo que apoya los resultados obtenidos en este estudio y sugiere que este SNP podría estar relacionado con el CaP. Con respecto al polimorfismo rs3822430, se ha estudiado en el contexto del fallo bioquímico sin haberse obtenido una asociación significativa (Audet-Walsh et al. 2011). La relevancia de los SNPs del gen *SRD5A1* en relación al CaP parece reforzada por el hecho de que un haplotipo específico de este gen (GCTTG TAGTA) esté fuertemente asociado con las variables clínicas de tamaño

tumoral y *Gleason score*. La influencia que este haplotipo específico pueda tener en la función de la enzima aún no ha sido examinada. Un análisis sistemático tanto de variables constitucionales, como de variables somáticas de la enzima esteroide-5 α -reductasa tipo II (en CaP), indicaron variaciones farmacogenéticas significativas en la respuesta a finasterida y dutasterida y permitieron mapear las áreas de la enzima salvaje responsables de la inhibición tiempo-específica para cada uno (o ambos) de los inhibidores (Makridakis and Reichardt 2005). Los SNPs son la base de la farmacogenética y es posible que tanto SNPs individuales como haplotipos específicos puedan explicar la heterogeneidad de respuesta a fármacos observada en la práctica clínica. Sin embargo, en este trabajo ningún SNP o haplotipo en el gen *SRD5A2* parece asociarse con cualquiera de las variables clínicas de la serie. No obstante, otros autores sí han publicado la relación de los SNPs rs4952197 o rs523349 (ambos incluidos en nuestro análisis) con la recurrencia bioquímica del CaP (Audet-Walsh et al. 2011). De hecho, rs523349 es uno de los polimorfismos mejor caracterizados del gen *SRD5A2* y se ha asociado con mal pronóstico en pacientes de CaP en términos de fallo bioquímico (incremento de los niveles de PSA después del tratamiento) (Shibata et al. 2002) y de invasión extracapsular (Jaffe et al. 2000). Shibata et al. y Jaffe et al. publicaron sus resultados en cohortes más pequeñas que las nuestras y Audet-Walsh et al. incluyó en su serie sujetos de distintos orígenes étnicos. Como hemos expuesto anteriormente, la homogeneidad étnica de la muestra tiene una fuerte influencia en los análisis de variaciones genéticas (Henriquez-Hernandez et al. 2013; Henriquez-Hernandez et al. 2015). Las diferencias en las técnicas de determinación genotípica o en las estrategias para el análisis de los SNPs podrían ser razones que expliquen las diferencias observadas entre estos resultados y los publicados por otros autores.

En esta serie, ninguno de los polimorfismos en *CYP17A1* presentó asociación estadísticamente significativa con las variables clínicas consideradas. Existen publicaciones que muestran una asociación entre los SNPs rs743572, rs619824 y rs2486758 y el riesgo de desarrollar CaP (Mononen and Schleutker 2009; Wang et al. 2011). Sin embargo, los resultados son contradictorios y están condicionados por el grupo étnico (Hamada et al. 2007). Los resultados de otro trabajo mostraron una relación entre el polimorfismo rs10883783 y la supervivencia del CaP tras una media de seguimiento de 13.2 años (Wright et al. 2010). La función de este gen en referencia al CaP parece estar limitado a su desarrollo en sí mismo, sin alterar las características

biológicas una vez desarrollado. Esta hipótesis se ve reforzada por el hecho de que el tratamiento con acetato de abiraterona presenta los mejores resultados cuando la producción testicular de testosterona es suprimida mediante un tratamiento hormonal; de otro modo, la inhibición específica de *CYP17A1* no tiene éxito (Shibata et al. 2002).

Los polimorfismos que afectan a la funcionalidad del gen *SRD5A1* podrían alterar la tasa de producción de DHT y subsecuentemente la exposición local a andrógenos de las células prostáticas sensibles a estas sustancias, influenciando así las características biológicas del tumor: tamaño tumoral, nivel de PSA en suero y *Gleason score*. El modo en que estos SNPs y haplotipos pueden influir en la función de las enzimas y en los niveles de testosterona es una cuestión que aún debe ser respondida. La alta fiabilidad de los sistemas de genotipado y la objetividad de los valores clínicos considerados (dado que se trata de técnicas rutinarias en la práctica clínica) minimizan las posibilidades de resultados inferidos por causa de la subjetividad de las observaciones. El presente estudio se ha realizado en una serie de pacientes de cáncer de próstata españoles, con un buen tamaño muestral (casi 500 sujetos) y homogeneidad en cuanto a la etnia de los mismos, un factor que como hemos visto anteriormente, se ha demostrado importante en estudios de asociación con polimorfismos.

9.4.- Conclusión

Este estudio revela la importancia de los polimorfismos de nucleótido simple en relación con las características biológicas al diagnóstico del cáncer de próstata, definiendo nuevas acciones del gen *SRD5A1* en una cohorte de pacientes españoles con CaP. Experimentos futuros deberán explorar el papel de los genotipos en el seguimiento de esta enfermedad.

Conclusiones

10.- Conclusiones

1. Las diferencias en la distribución de genotipos entre diferentes poblaciones de la misma etnia puede ser un factor de confusión importante y podría influir en la falta de validación de los resultados obtenidos en los estudios de SNPs asociados con la toxicidad inducida por radiación. Esto es de especial relevancia cuando se llevan a cabo extensos meta-análisis con sujetos procedentes de distintos países. Los resultados aquí expuestos sugieren que la homogeneidad entre las personas (especialmente entre aquellos sujetos considerados controles del estudio) debería ser probada antes de proceder con cualquier otro análisis.

2. La diferencia en la distribución de genotipos de *SRD5A1* y *SRD5A2* entre las diferentes poblaciones de la misma etnia puede ser un factor responsable del fallo de los tratamientos basados en 5 α -RI. Nuestros resultados sugieren que la dotación genética debe ser tomada en cuenta (especialmente en ensayos clínicos que incluyan gente de diferentes etnias) en la determinación de la efectividad de este tipo de fármacos.

3. Los pacientes con cáncer de próstata portadores de los genotipos homocigotos salvajes para los SNPs rs11615 o rs17503908, localizados en los genes *ERC1* y *ATM*, presentaron variables clínicas de peor pronóstico, un hecho acentuado en aquellos pacientes homocigotos para ambos polimorfismos.

4. Los polimorfismos rs2267437 (*XRCC6*) y rs3815824 (*MVP*) están asociados con variables clínicas que identifican los tumores de CaP de mayor agresividad. En concreto, la presencia de variables genotípicas en estos dos genes, por separado o en conjunto, se asocian a mayores niveles de PSA al diagnóstico así como a tumores de mayor riesgo según la clasificación de riesgo de D'Amico.

5. Determinados polimorfismos de nucleótido simple en el gen *SRD5A1* parecen ser relevantes en relación con las características biológicas al diagnóstico del cáncer de próstata, definiendo nuevos haplotipos en el cromosoma 5 que podrían condicionar el comportamiento biológico de la enfermedad.

*Referencias
bibliográficas*

11.- Referencias bibliográficas

- Adel Fahmideh, M., J. Schwartzbaum, et al. (2014). "Association between DNA repair gene polymorphisms and risk of glioma: a systematic review and meta-analysis." Neuro Oncol **16**(6): 807-814.
- Ahn, J., F. R. Schumacher, et al. (2009). "Quantitative trait loci predicting circulating sex steroid hormones in men from the NCI-Breast and Prostate Cancer Cohort Consortium (BPC3)." Hum Mol Genet **18**(19): 3749-3757.
- Altmuller, J., L. J. Palmer, et al. (2001). "Genomewide scans of complex human diseases: true linkage is hard to find." Am J Hum Genet **69**(5): 936-950.
- Aly, M., F. Wiklund, et al. (2011). "Polygenic risk score improves prostate cancer risk prediction: results from the Stockholm-1 cohort study." Eur Urol **60**(1): 21-28.
- Ambrosio, B., C. Hernandez, et al. (2010). "Searching the peopling of the Iberian Peninsula from the perspective of two andalusian subpopulations: a study based on Y-chromosome haplogroups J and E." Coll Antropol **34**(4): 1215-1228.
- Andreassen, C. N. and J. Alsner (2009). "Genetic variants and normal tissue toxicity after radiotherapy: a systematic review." Radiother Oncol **92**(3): 299-309.
- Andriole, G., D. Bostwick, et al. (2004). "Chemoprevention of prostate cancer in men at high risk: rationale and design of the reduction by dutasteride of prostate cancer events (REDUCE) trial." J Urol **172**(4 Pt 1): 1314-1317.
- Andriole, G. L., D. G. Bostwick, et al. (2010). "Effect of dutasteride on the risk of prostate cancer." N Engl J Med **362**(13): 1192-1202.
- Andriole, G. L. and R. Kirby (2003). "Safety and tolerability of the dual 5alpha-reductase inhibitor dutasteride in the treatment of benign prostatic hyperplasia." Eur Urol **44**(1): 82-88.
- Angele, S., A. Falconer, et al. (2004). "ATM protein overexpression in prostate tumors: possible role in telomere maintenance." Am J Clin Pathol **121**(2): 231-236.
- Audet-Walsh, E., J. Bellemare, et al. (2011). "SRD5A polymorphisms and biochemical failure after radical prostatectomy." Eur Urol **60**(6): 1226-1234.
- Aumuller, G., W. Eicheler, et al. (1996). "Immunocytochemical evidence for differential subcellular localization of 5 alpha-reductase isoenzymes in human tissues." Acta Anat (Basel) **156**(4): 241-252.
- Auvinen, A., S. M. Moss, et al. (2016). "Absolute Effect of Prostate Cancer Screening: Balance of Benefits and Harms by Center within the European Randomized Study of Prostate Cancer Screening." Clin Cancer Res **22**(1): 243-249.
- Badani, K., D. J. Thompson, et al. (2013). "Impact of a genomic classifier of metastatic risk on postoperative treatment recommendations for prostate cancer patients: a report from the DECIDE study group." Oncotarget **4**(4): 600-609.
- Bagg, A. and J. Cossman (1992). "BCL-2: physiology and role in neoplasia." Cancer Treat Res **63**: 141-166.
- Balan, S., S. K. Radhab, et al. (2013). "Major vault protein (MVP) gene polymorphisms and drug resistance in mesial temporal lobe epilepsy with hippocampal sclerosis." Gene **526**(2): 449-453.
- Balmain, A. (2001). "Cancer genetics: from Boveri and Mendel to microarrays." Nat Rev Cancer **1**(1): 77-82.
- Barker, P. E. (2003). "Cancer biomarker validation: standards and process: roles for the National Institute of Standards and Technology (NIST)." Ann N Y Acad Sci **983**: 142-150.

- Barnett, G. C., C. E. Coles, et al. (2012). "Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study." Lancet Oncol **13**(1): 65-77.
- Barnett, G. C., R. M. Elliott, et al. (2012). "Individual patient data meta-analysis shows no association between the SNP rs1800469 in TGFB and late radiotherapy toxicity." Radiother Oncol **105**(3): 289-295.
- Barrack, E. R. (1997). "TGF beta in prostate cancer: a growth inhibitor that can enhance tumorigenicity." Prostate **31**(1): 61-70.
- Barrett, J. C., B. Fry, et al. (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-265.
- Bassing, C. H. and F. W. Alt (2004). "The cellular response to general and programmed DNA double strand breaks." DNA Repair (Amst) **3**(8-9): 781-796.
- Bau, D. T., H. C. Wu, et al. (2007). "Association of XPD polymorphisms with prostate cancer in Taiwanese patients." Anticancer Res **27**(4C): 2893-2896.
- Bauchet, M., B. McEvoy, et al. (2007). "Measuring European population stratification with microarray genotype data." Am J Hum Genet **80**(5): 948-956.
- Benson, M. C., I. S. Whang, et al. (1992). "The use of prostate specific antigen density to enhance the predictive value of intermediate levels of serum prostate specific antigen." J Urol **147**(3 Pt 2): 817-821.
- Bentzen, S. M. and J. Overgaard (1994). "Patient-to-Patient Variability in the Expression of Radiation-Induced Normal Tissue Injury." Semin Radiat Oncol **4**(2): 68-80.
- Berger, N. A. (1985). "Poly(ADP-ribose) in the cellular response to DNA damage." Radiat Res **101**(1): 4-15.
- Biernacka, K. M., C. M. Perks, et al. (2012). "Role of the IGF axis in prostate cancer." Minerva Endocrinol **37**(2): 173-185.
- Biggerstaff, M., D. E. Szymkowski, et al. (1993). "Co-correction of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro." EMBO J **12**(9): 3685-3692.
- Binder, M., B. Y. Zhang, et al. (2016). "Common Genetic Variation in CYP17A1 and Response to Abiraterone Acetate in Patients with Metastatic Castration-Resistant Prostate Cancer." Int J Mol Sci **17**(7).
- Bishoff, J. T., S. J. Freedland, et al. (2014). "Prognostic utility of the cell cycle progression score generated from biopsy in men treated with prostatectomy." J Urol **192**(2): 409-414.
- Blasko, J. C., P. D. Grimm, et al. (2000). "Palladium-103 brachytherapy for prostate carcinoma." Int J Radiat Oncol Biol Phys **46**(4): 839-850.
- Blume-Jensen, P., D. M. Berman, et al. (2015). "Development and clinical validation of an in situ biopsy-based multimarker assay for risk stratification in prostate cancer." Clin Cancer Res **21**(11): 2591-2600.
- Borchiellini, D., M. C. Etienne-Grimaldi, et al. (2012). "The impact of pharmacogenetics on radiation therapy outcome in cancer patients. A focus on DNA damage response genes." Cancer Treat Rev **38**(6): 737-759.
- Bouwman, P. and J. Jonkers (2012). "The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance." Nat Rev Cancer **12**(9): 587-598.
- Bracarda, S., O. de Cobelli, et al. (2005). "Cancer of the prostate." Crit Rev Oncol Hematol **56**(3): 379-396.
- Brem, R. and J. Hall (2005). "XRCC1 is required for DNA single-strand break repair in human cells." Nucleic Acids Res **33**(8): 2512-2520.
- Browning, R. E. t., H. Li, et al. (2006). "ATM polymorphism IVS62+60G>A is not associated with disease aggressiveness in prostate cancer." Urology **67**(6): 1320-1323.
- Buchholz, T. A. (1999). "Finding our sensitive patients." Int J Radiat Oncol Biol Phys **45**(3): 547-548.

- Buzdar, A. and A. Howell (2001). "Advances in aromatase inhibition: clinical efficacy and tolerability in the treatment of breast cancer." *Clin Cancer Res* **7**(9): 2620-2635.
- Calafell, F. and J. Bertranpetit (1994). "Principal component analysis of gene frequencies and the origin of Basques." *Am J Phys Anthropol* **93**(2): 201-215.
- Caldecott, K. W., S. Aoufouchi, et al. (1996). "XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro." *Nucleic Acids Res* **24**(22): 4387-4394.
- Cancer Genome Atlas Research, N. (2015). "The Molecular Taxonomy of Primary Prostate Cancer." *Cell* **163**(4): 1011-1025.
- Canman, C. E. and D. S. Lim (1998). "The role of ATM in DNA damage responses and cancer." *Oncogene* **17**(25): 3301-3308.
- Cartegni, L., S. L. Chew, et al. (2002). "Listening to silence and understanding nonsense: exonic mutations that affect splicing." *Nat Rev Genet* **3**(4): 285-298.
- Carter, H. B., J. D. Pearson, et al. (1992). "Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease." *JAMA* **267**(16): 2215-2220.
- Catalona, W. J., D. S. Smith, et al. (1993). "Detection of organ-confined prostate cancer is increased through prostate-specific antigen-based screening." *JAMA* **270**(8): 948-954.
- Cesaretti, J. A., R. G. Stock, et al. (2007). "A genetically determined dose-volume histogram predicts for rectal bleeding among patients treated with prostate brachytherapy." *Int J Radiat Oncol Biol Phys* **68**(5): 1410-1416.
- Cibeira, M. T., C. F. de Larrea, et al. (2011). "Impact on response and survival of DNA repair single nucleotide polymorphisms in relapsed or refractory multiple myeloma patients treated with thalidomide." *Leuk Res* **35**(9): 1178-1183.
- Clayton, D. (2001). "Population association. Handbook of statistical genetics." 519-540.
- Clayton, D. (2012). "snpStats: SnpMatrix and XSNPMatrix classes and methods. R package version 1.8.1."
- Cooke, M. S., M. D. Evans, et al. (2003). "Oxidative DNA damage: mechanisms, mutation, and disease." *FASEB J* **17**(10): 1195-1214.
- Cooperberg, M. R., J. M. Broering, et al. (2010). "Time trends and local variation in primary treatment of localized prostate cancer." *J Clin Oncol* **28**(7): 1117-1123.
- Cooperberg, M. R., D. J. Pasta, et al. (2005). "The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy." *J Urol* **173**(6): 1938-1942.
- Cooperberg, M. R., J. P. Simko, et al. (2013). "Validation of a cell-cycle progression gene panel to improve risk stratification in a contemporary prostatectomy cohort." *J Clin Oncol* **31**(11): 1428-1434.
- Correa-Cerro, L., G. Wöhr, et al. (1999). "(CAG)_nCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population." *Eur J Hum Genet* **7**(3): 357-362.
- Crawford, E. D., M. C. Scholz, et al. (2014). "Cell cycle progression score and treatment decisions in prostate cancer: results from an ongoing registry." *Curr Med Res Opin* **30**(6): 1025-1031.
- Culig, Z., A. Hobisch, et al. (1995). "Activation of the androgen receptor by polypeptide growth factors and cellular regulators." *World J Urol* **13**(5): 285-289.
- Cullen, J., I. L. Rosner, et al. (2015). "A Biopsy-based 17-gene Genomic Prostate Score Predicts Recurrence After Radical Prostatectomy and Adverse Surgical Pathology in a Racially Diverse Population of Men with Clinically Low- and Intermediate-risk Prostate Cancer." *Eur Urol* **68**(1): 123-131.
- Cussenot, O., A. R. Azzouzi, et al. (2007). "Low-activity V89L variant in SRD5A2 is associated with aggressive prostate cancer risk: an explanation for the adverse effects observed in chemoprevention trials using 5-alpha-reductase inhibitors." *Eur Urol* **52**(4): 1082-1087.

- Cuzick, J., S. Stone, et al. (2015). "Validation of an RNA cell cycle progression score for predicting death from prostate cancer in a conservatively managed needle biopsy cohort." *Br J Cancer* **113**(3): 382-389.
- Cuzick, J., G. P. Swanson, et al. (2011). "Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study." *Lancet Oncol* **12**(3): 245-255.
- Chakraborty, R. (2002). Hardy-Weinberg Equilibrium. *Biostatistical Genetics and Genetic Epidemiology*. R. Elston, J. Olson and L. Palmer. West Sussex, John Wiley & Sons Ltd.: 369-371.
- Chamary, J. V., J. L. Parmley, et al. (2006). "Hearing silence: non-neutral evolution at synonymous sites in mammals." *Nat Rev Genet* **7**(2): 98-108.
- Chang-Claude, J., C. B. Ambrosone, et al. (2009). "Genetic polymorphisms in DNA repair and damage response genes and late normal tissue complications of radiotherapy for breast cancer." *Br J Cancer* **100**(10): 1680-1686.
- Chism, D. B., A. L. Hanlon, et al. (2004). "A comparison of the single and double factor high-risk models for risk assignment of prostate cancer treated with 3D conformal radiotherapy." *Int J Radiat Oncol Biol Phys* **59**(2): 380-385.
- Chiu, Y. T., J. Liu, et al. (2012). "Inactivation of ATM/ATR DNA damage checkpoint promotes androgen induced chromosomal instability in prostate epithelial cells." *PLoS One* **7**(12): e51108.
- Chung, J. H., M. E. Ginn-Pease, et al. (2005). "Phosphatase and tensin homologue deleted on chromosome 10 (PTEN) has nuclear localization signal-like sequences for nuclear import mediated by major vault protein." *Cancer Res* **65**(10): 4108-4116.
- D'Amico, A. V., R. Whittington, et al. (1998). "Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer." *JAMA* **280**(11): 969-974.
- D'Amico, A. V., R. Whittington, et al. (2000). "Utilizing predictions of early prostate-specific antigen failure to optimize patient selection for adjuvant systemic therapy trials." *J Clin Oncol* **18**(18): 3240-3246.
- Damaraju, S., D. Murray, et al. (2006). "Association of DNA repair and steroid metabolism gene polymorphisms with clinical late toxicity in patients treated with conformal radiotherapy for prostate cancer." *Clin Cancer Res* **12**(8): 2545-2554.
- De Bont, R. and N. van Larebeke (2004). "Endogenous DNA damage in humans: a review of quantitative data." *Mutagenesis* **19**(3): 169-185.
- de La Taille, A., A. Houlgatte, et al. (1997). "[The incidence of the variability of the free PSA/total PSA ratio on the early diagnosis of prostate cancer]." *Prog Urol* **7**(3): 455-463.
- Di Lorenzo, G., R. Autorino, et al. (2004). "HER-2/neu receptor in prostate cancer development and progression to androgen independence." *Tumori* **90**(2): 163-170.
- DiBiase, S. J. and M. Roach (2015). External beam radiation therapy for localized prostate cancer. *UpToDate*. M. E. Ross. Waltham, Massachusetts.
- Doll, R. and R. Peto (1981). "The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today." *J Natl Cancer Inst* **66**(6): 1191-1308.
- Dong, L. M., J. D. Potter, et al. (2008). "Genetic susceptibility to cancer: the role of polymorphisms in candidate genes." *JAMA* **299**(20): 2423-2436.
- Dork, T., R. Bendix, et al. (2001). "Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients." *Cancer Res* **61**(20): 7608-7615.
- Eadie, J. S., M. Conrad, et al. (1984). "Mechanism of mutagenesis by O6-methylguanine." *Nature* **308**(5955): 201-203.
- Edge, S., D. R. Byrd, et al. (2010). *AJCC Cancer staging manual*, Springer.
- Edwards, J. E. and R. A. Moore (2002). "Finasteride in the treatment of clinical benign prostatic hyperplasia: a systematic review of randomised trials." *BMC Urol* **2**: 14.

- Eeles, R. A., A. A. Olama, et al. (2013). "Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array." *Nat Genet* **45**(4): 385-391, 391e381-382.
- Eisen, M. B., P. T. Spellman, et al. (1998). "Cluster analysis and display of genome-wide expression patterns." *Proc Natl Acad Sci U S A* **95**(25): 14863-14868.
- Eisenberger, M. A., B. A. Blumenstein, et al. (1998). "Bilateral orchiectomy with or without flutamide for metastatic prostate cancer." *N Engl J Med* **339**(15): 1036-1042.
- Eisenberger, M. A., R. Simon, et al. (1985). "A reevaluation of nonhormonal cytotoxic chemotherapy in the treatment of prostatic carcinoma." *J Clin Oncol* **3**(6): 827-841.
- Elliott, R. L. and G. C. Blobe (2005). "Role of transforming growth factor Beta in human cancer." *J Clin Oncol* **23**(9): 2078-2093.
- Erho, N., A. Crisan, et al. (2013). "Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy." *PLoS One* **8**(6): e66855.
- Evans, E., J. G. Moggs, et al. (1997). "Mechanism of open complex and dual incision formation by human nucleotide excision repair factors." *EMBO J* **16**(21): 6559-6573.
- Fachal, L., A. Gomez-Caamano, et al. (2012). "Association of a XRCC3 polymorphism and rectum mean dose with the risk of acute radio-induced gastrointestinal toxicity in prostate cancer patients." *Radiother Oncol* **105**(3): 321-328.
- Farkas, A., D. Schneider, et al. (1998). "National trends in the epidemiology of prostate cancer, 1973 to 1994: evidence for the effectiveness of prostate-specific antigen screening." *Urology* **52**(3): 444-448; discussion 448-449.
- Ferlay, J., I. Soerjomataram, et al. (2015). "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012." *Int J Cancer* **136**(5): E359-386.
- Ferlay, J., E. Steliarova-Foucher, et al. (2013). "Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012." *Eur J Cancer* **49**(6): 1374-1403.
- Fisher, G., Z. H. Yang, et al. (2013). "Prognostic value of Ki-67 for prostate cancer death in a conservatively managed cohort." *Br J Cancer* **108**(2): 271-277.
- Flaherty, K. T., C. Robert, et al. (2012). "Improved survival with MEK inhibition in BRAF-mutated melanoma." *N Engl J Med* **367**(2): 107-114.
- Fong, P. C., D. S. Boss, et al. (2009). "Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers." *N Engl J Med* **361**(2): 123-134.
- Fraser, M., A. Berlin, et al. (2015). "Genomic, pathological, and clinical heterogeneity as drivers of personalized medicine in prostate cancer." *Urol Oncol* **33**(2): 85-94.
- Freedland, S. J., L. Gerber, et al. (2013). "Prognostic utility of cell cycle progression score in men with prostate cancer after primary external beam radiation therapy." *Int J Radiat Oncol Biol Phys* **86**(5): 848-853.
- Gabriel, S. B., S. F. Schaffner, et al. (2002). "The structure of haplotype blocks in the human genome." *Science* **296**(5576): 2225-2229.
- Gaudreau, P. O., J. Stagg, et al. (2016). "The Present and Future of Biomarkers in Prostate Cancer: Proteomics, Genomics, and Immunology Advancements." *Biomark Cancer* **8**(Suppl 2): 15-33.
- Giwerzman, C., A. Giwerzman, et al. (2008). "Polymorphisms in genes regulating androgen activity among prostate cancer low-risk Inuit men and high-risk Scandinavians." *Int J Androl* **31**(1): 25-30.
- Giwerzman, Y. L., P. A. Abrahamsson, et al. (2005). "The 5alpha-reductase type II A49T and V89L high-activity allelic variants are more common in men with prostate cancer compared with the general population." *Eur Urol* **48**(4): 679-685.
- Gleason, D. F. and G. T. Mellinger (1974). "Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging." *J Urol* **111**(1): 58-64.

- Gomes, B. C., S. N. Silva, et al. (2010). "The role of common variants of non-homologous end-joining repair genes XRCC4, LIG4 and Ku80 in thyroid cancer risk." Oncol Rep **24**(4): 1079-1085.
- Grawunder, U., D. Zimmer, et al. (1998). "DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes." Mol Cell **2**(4): 477-484.
- Greene, F. L. and L. H. Sobin (2008). "The staging of cancer: a retrospective and prospective appraisal." CA Cancer J Clin **58**(3): 180-190.
- Grupp, K., L. Roettger, et al. (2015). "Expression of DNA ligase IV is linked to poor prognosis and characterizes a subset of prostate cancers harboring TMPRSS2:ERG fusion and PTEN deletion." Oncol Rep **34**(3): 1211-1220.
- Gu, J., H. Lu, et al. (2007). "XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps." EMBO J **26**(4): 1010-1023.
- Gudmundsson, J., S. Besenbacher, et al. (2010). "Genetic correction of PSA values using sequence variants associated with PSA levels." Sci Transl Med **2**(62): 62ra92.
- Guttmacher, A. E. and F. S. Collins (2002). "Genomic medicine--a primer." N Engl J Med **347**(19): 1512-1520.
- Hamada, A., R. Danesi, et al. (2007). "Association of a CYP17 polymorphism with overall survival in Caucasian patients with androgen-independent prostate cancer." Urology **70**(2): 217-220.
- Hankey, B. F., E. J. Feuer, et al. (1999). "Cancer surveillance series: interpreting trends in prostate cancer--part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates." J Natl Cancer Inst **91**(12): 1017-1024.
- Hasegawa, T., M. Someya, et al. (2016). "Expression of Ku70 predicts results of radiotherapy in prostate cancer." Strahlenther Onkol.
- Hayden, P. J., P. Tewari, et al. (2007). "Variation in DNA repair genes XRCC3, XRCC4, XRCC5 and susceptibility to myeloma." Hum Mol Genet **16**(24): 3117-3127.
- Heidegger, I., H. Klocker, et al. (2014). "[-2]proPSA is an early marker for prostate cancer aggressiveness." Prostate Cancer Prostatic Dis **17**(1): 70-74.
- Helfand, B. T., W. J. Catalona, et al. (2015). "A genetic-based approach to personalized prostate cancer screening and treatment." Curr Opin Urol **25**(1): 53-58.
- Helgesen, F., L. Holmberg, et al. (1996). "Trends in prostate cancer survival in Sweden, 1960 through 1988: evidence of increasing diagnosis of nonlethal tumors." J Natl Cancer Inst **88**(17): 1216-1221.
- Helleday, T., S. Eshtad, et al. (2014). "Mechanisms underlying mutational signatures in human cancers." Nat Rev Genet **15**(9): 585-598.
- Henriquez-Hernandez, L. A., E. Bordon, et al. (2012). "Prediction of normal tissue toxicity as part of the individualized treatment with radiotherapy in oncology patients." Surg Oncol **21**(3): 201-206.
- Henriquez-Hernandez, L. A., M. Lloret, et al. (2011). "BCL-2, in combination with MVP and IGF-1R expression, improves prediction of clinical outcome in complete response cervical carcinoma patients treated by radiochemotherapy." Gynecol Oncol **122**(3): 585-589.
- Henriquez-Hernandez, L. A., M. Moreno, et al. (2012). "MVP expression in the prediction of clinical outcome of locally advanced oral squamous cell carcinoma patients treated with radiotherapy." Radiat Oncol **7**: 147.
- Henriquez-Hernandez, L. A., A. Murias-Rosales, et al. (2010). "Distribution of TYMS, MTHFR, p53 and MDR1 gene polymorphisms in patients with breast cancer treated with neoadjuvant chemotherapy." Cancer Epidemiol **34**(5): 634-638.
- Henriquez-Hernandez, L. A., L. F. Perez, et al. (2010). "TYMS, MTHFR, p53 and MDR1 gene polymorphisms in breast cancer patients treated with adjuvant therapy." Cancer Epidemiol **34**(4): 490-493.

- Henriquez-Hernandez, L. A., A. Valenciano, et al. (2013). "Polymorphisms in DNA-repair genes in a cohort of prostate cancer patients from different areas in Spain: heterogeneity between populations as a confounding factor in association studies." PLoS One **8**(7): e69735.
- Henriquez-Hernandez, L. A., A. Valenciano, et al. (2015). "Intraethnic variation in steroid-5-alpha-reductase polymorphisms in prostate cancer patients: a potential factor implicated in 5-alpha-reductase inhibitor treatment." J Genet **94**(2): 335-341.
- Henriquez-Hernandez, L. A., A. Valenciano, et al. (2014). "Single nucleotide polymorphisms in DNA repair genes as risk factors associated to prostate cancer progression." BMC Med Genet **15**: 143.
- Hickman, J. A. (2002). "Apoptosis and tumorigenesis." Curr Opin Genet Dev **12**(1): 67-72.
- Howrey, B. T., Y. F. Kuo, et al. (2013). "The impact of PSA screening on prostate cancer mortality and overdiagnosis of prostate cancer in the United States." J Gerontol A Biol Sci Med Sci **68**(1): 56-61.
- Hu, Y., M. Zhou, et al. (2014). "Two DNA repair gene polymorphisms on the risk of gastrointestinal cancers: a meta-analysis." Tumour Biol **35**(3): 1715-1725.
- Huggins, C. (1967). "Endocrine-induced regression of cancers." Cancer Res **27**(11): 1925-1930.
- Ilic, D., D. O'Connor, et al. (2007). "Screening for prostate cancer: a Cochrane systematic review." Cancer Causes Control **18**(3): 279-285.
- International HapMap, C. (2003). "The International HapMap Project." Nature **426**(6968): 789-796.
- Ioannidis, J. P., P. Castaldi, et al. (2010). "A compendium of genome-wide associations for cancer: critical synopsis and reappraisal." J Natl Cancer Inst **102**(12): 846-858.
- Ioannidis, J. P., E. E. Ntzani, et al. (2001). "Replication validity of genetic association studies." Nat Genet **29**(3): 306-309.
- Jaffe, J. M., S. B. Malkowicz, et al. (2000). "Association of SRD5A2 genotype and pathological characteristics of prostate tumors." Cancer Res **60**(6): 1626-1630.
- Jain, M., R. Nilsson, et al. (2012). "Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation." Science **336**(6084): 1040-1044.
- Jemal, A., F. Bray, et al. (2011). "Global cancer statistics." CA Cancer J Clin **61**(2): 69-90.
- Jensen, C. J., B. J. Oldfield, et al. (2009). "Splicing, cis genetic variation and disease." Biochem Soc Trans **37**(Pt 6): 1311-1315.
- Jia, J., J. Ren, et al. (2015). "Association between the XRCC6 polymorphisms and cancer risks: a systematic review and meta-analysis." Medicine (Baltimore) **94**(1): e283.
- Johansson, S., H. Svensson, et al. (2000). "Timescale of evolution of late radiation injury after postoperative radiotherapy of breast cancer patients." Int J Radiat Oncol Biol Phys **48**(3): 745-750.
- Kedersha, N. L., J. E. Heuser, et al. (1991). "Vaults. III. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry." J Cell Biol **112**(2): 225-235.
- Kerns, S. L., D. de Ruyscher, et al. (2014). "STROGAR - STrengthening the Reporting Of Genetic Association studies in Radiogenomics." Radiother Oncol **110**(1): 182-188.
- Khanna, K. K. and S. P. Jackson (2001). "DNA double-strand breaks: signaling, repair and the cancer connection." Nat Genet **27**(3): 247-254.
- Khor, L. Y., K. Bae, et al. (2009). "MDM2 and Ki-67 predict for distant metastasis and mortality in men treated with radiotherapy and androgen deprivation for prostate cancer: RTOG 92-02." J Clin Oncol **27**(19): 3177-3184.
- Kimchi-Sarfaty, C., J. M. Oh, et al. (2007). "A "silent" polymorphism in the MDR1 gene changes substrate specificity." Science **315**(5811): 525-528.
- Klein, E. A., M. R. Cooperberg, et al. (2014). "A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling." Eur Urol **66**(3): 550-560.

- Knezevic, D., A. D. Goddard, et al. (2013). "Analytical validation of the Oncotype DX prostate cancer assay - a clinical RT-PCR assay optimized for prostate needle biopsies." BMC Genomics **14**: 690.
- Koberle, B., B. Koch, et al. (2016). "Single nucleotide polymorphisms in DNA repair genes and putative cancer risk." Arch Toxicol.
- Korsmeyer, S. J. (1992). "Bcl-2 initiates a new category of oncogenes: regulators of cell death." Blood **80**(4): 879-886.
- Kote-Jarai, Z., D. F. Easton, et al. (2008). "Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL Consortium." Cancer Epidemiol Biomarkers Prev **17**(8): 2052-2061.
- Krokan, H. E. and M. Bjoras (2013). "Base excision repair." Cold Spring Harb Perspect Biol **5**(4): a012583.
- Kubota, Y., R. A. Nash, et al. (1996). "Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein." EMBO J **15**(23): 6662-6670.
- Kupelian, P. A., L. Potters, et al. (2004). "Radical prostatectomy, external beam radiotherapy <72 Gy, external beam radiotherapy > or =72 Gy, permanent seed implantation, or combined seeds/external beam radiotherapy for stage T1-T2 prostate cancer." Int J Radiat Oncol Biol Phys **58**(1): 25-33.
- Laayouni, H., F. Calafell, et al. (2010). "A genome-wide survey does not show the genetic distinctiveness of Basques." Hum Genet **127**(4): 455-458.
- Langsenlehner, T., W. Renner, et al. (2011). "Association between single nucleotide polymorphisms in the gene for XRCC1 and radiation-induced late toxicity in prostate cancer patients." Radiother Oncol **98**(3): 387-393.
- Lara, P. C., M. Pruschy, et al. (2011). "MVP and vaults: a role in the radiation response." Radiat Oncol **6**: 148.
- Lazzeri, M., A. Haese, et al. (2013). "Clinical performance of serum prostate-specific antigen isoform [-2]proPSA (p2PSA) and its derivatives, %p2PSA and the prostate health index (PHI), in men with a family history of prostate cancer: results from a multicentre European study, the PROMetheuS project." BJU Int **112**(3): 313-321.
- Le Guen, T., S. Ragu, et al. (2015). "Role of the double-strand break repair pathway in the maintenance of genomic stability." Mol Cell Oncol **2**(1): e968020.
- Lee, P. H. and H. Shatkay (2008). "F-SNP: computationally predicted functional SNPs for disease association studies." Nucleic Acids Res **36**(Database issue): D820-824.
- Lee, P. H. and H. Shatkay (2008). "Ranking single nucleotide polymorphisms by potential deleterious effects." AMIA Annu Symp Proc: 667-671.
- Lehmann, A. R. (2001). "The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases." Genes Dev **15**(1): 15-23.
- Leongamornlert, D., N. Mahmud, et al. (2012). "Germline BRCA1 mutations increase prostate cancer risk." Br J Cancer **106**(10): 1697-1701.
- Levine, A. J., J. Momand, et al. (1991). "The p53 tumour suppressor gene." Nature **351**(6326): 453-456.
- Levine, A. J. and M. Oren (2009). "The first 30 years of p53: growing ever more complex." Nat Rev Cancer **9**(10): 749-758.
- Lewontin, R. C. and D. L. Hartl (1991). "Population genetics in forensic DNA typing." Science **254**(5039): 1745-1750.
- Li, G. M. (2008). "Mechanisms and functions of DNA mismatch repair." Cell Res **18**(1): 85-98.
- Li, J., R. J. Coates, et al. (2010). "Steroid 5- α -reductase Type 2 (SRD5a2) gene polymorphisms and risk of prostate cancer: a HuGE review." Am J Epidemiol **171**(1): 1-13.

- Li, Q., Y. Zhu, et al. (2013). "Steroid 5-alpha-reductase type 2 (SRD5A2) V89L and A49T polymorphisms and sporadic prostate cancer risk: a meta-analysis." Mol Biol Rep **40**(5): 3597-3608.
- Lichtenstein, P., N. V. Holm, et al. (2000). "Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland." N Engl J Med **343**(2): 78-85.
- Lieber, M. R. (2008). "The mechanism of human nonhomologous DNA end joining." J Biol Chem **283**(1): 1-5.
- Lieber, M. R. (2010). "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway." Annu Rev Biochem **79**: 181-211.
- Lin, B. K., M. Clyne, et al. (2006). "Tracking the epidemiology of human genes in the literature: the HuGE Published Literature database." Am J Epidemiol **164**(1): 1-4.
- Lindahl, T. (1993). "Instability and decay of the primary structure of DNA." Nature **362**(6422): 709-715.
- Lindstrom, S., H. O. Adami, et al. (2007). "Inherited variation in hormone-regulating genes and prostate cancer survival." Clin Cancer Res **13**(17): 5156-5161.
- Lindstrom, S., F. R. Schumacher, et al. (2012). "Common genetic variants in prostate cancer risk prediction--results from the NCI Breast and Prostate Cancer Cohort Consortium (BPC3)." Cancer Epidemiol Biomarkers Prev **21**(3): 437-444.
- Lindstrom, S., F. Wiklund, et al. (2006). "Germ-line genetic variation in the key androgen-regulating genes androgen receptor, cytochrome P450, and steroid-5-alpha-reductase type 2 is important for prostate cancer development." Cancer Res **66**(22): 11077-11083.
- Lockett, K. L., M. C. Hall, et al. (2004). "The ADPRT V762A genetic variant contributes to prostate cancer susceptibility and deficient enzyme function." Cancer Res **64**(17): 6344-6348.
- Ludwig, J. A. and J. N. Weinstein (2005). "Biomarkers in cancer staging, prognosis and treatment selection." Nat Rev Cancer **5**(11): 845-856.
- Lund, L., N. Svolgaard, et al. (2014). "Prostate cancer: a review of active surveillance." Res Rep Urol **6**: 107-112.
- Luo, J., T. A. Dunn, et al. (2003). "Decreased gene expression of steroid 5 alpha-reductase 2 in human prostate cancer: implications for finasteride therapy of prostate carcinoma." Prostate **57**(2): 134-139.
- Lloret, M., P. C. Lara, et al. (2009). "Major vault protein may affect nonhomologous end-joining repair and apoptosis through Ku70/80 and bax downregulation in cervical carcinoma tumors." Int J Radiat Oncol Biol Phys **73**(4): 976-979.
- Lloret, M., P. C. Lara, et al. (2008). "MVP expression is related to IGF1-R in cervical carcinoma patients treated by radiochemotherapy." Gynecol Oncol **110**(3): 304-307.
- Ma, Q. and A. Y. Lu (2011). "Pharmacogenetics, pharmacogenomics, and individualized medicine." Pharmacol Rev **63**(2): 437-459.
- Mackey, T. J., A. Borkowski, et al. (1998). "bcl-2/bax ratio as a predictive marker for therapeutic response to radiotherapy in patients with prostate cancer." Urology **52**(6): 1085-1090.
- Makarov, D. V., B. J. Trock, et al. (2007). "Updated nomogram to predict pathologic stage of prostate cancer given prostate-specific antigen level, clinical stage, and biopsy Gleason score (Partin tables) based on cases from 2000 to 2005." Urology **69**(6): 1095-1101.
- Makridakis, N. and J. K. Reichardt (2005). "Pharmacogenetic analysis of human steroid 5 alpha reductase type II: comparison of finasteride and dutasteride." J Mol Endocrinol **34**(3): 617-623.
- Mandrekar, S. J. and D. J. Sargent (2009). "Clinical trial designs for predictive biomarker validation: theoretical considerations and practical challenges." J Clin Oncol **27**(24): 4027-4034.

- Mangoni, M., S. Bisanzi, et al. (2011). "Association between genetic polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT genes and radiosensitivity in breast cancer patients." *Int J Radiat Oncol Biol Phys* **81**(1): 52-58.
- Marsin, S., A. E. Vidal, et al. (2003). "Role of XRCC1 in the coordination and stimulation of oxidative DNA damage repair initiated by the DNA glycosylase hOGG1." *J Biol Chem* **278**(45): 44068-44074.
- Masson, M., C. Niedergang, et al. (1998). "XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage." *Mol Cell Biol* **18**(6): 3563-3571.
- Matoka, D. J., V. Yao, et al. (2012). "Deficiency of DNA repair nuclease ERCC1-XPF promotes prostate cancer progression in a tissue recombination model." *Prostate* **72**(11): 1214-1222.
- McConnell, J. D. (1995). "Prostatic growth: new insights into hormonal regulation." *Br J Urol* **76 Suppl 1**: 5-10.
- McDunn, J. E., Z. Li, et al. (2013). "Metabolomic signatures of aggressive prostate cancer." *Prostate* **73**(14): 1547-1560.
- Mettlin, C. J. and G. P. Murphy (1998). "Why is the prostate cancer death rate declining in the United States?" *Cancer* **82**(2): 249-251.
- Meulmeester, E. and A. G. Jochemsen (2008). "p53: a guide to apoptosis." *Curr Cancer Drug Targets* **8**(2): 87-97.
- Meyer, F., L. Moore, et al. (1999). "Downward trend in prostate cancer mortality in Quebec and Canada." *J Urol* **161**(4): 1189-1191.
- Mononen, N. and J. Schleutker (2009). "Polymorphisms in genes involved in androgen pathways as risk factors for prostate cancer." *J Urol* **181**(4): 1541-1549.
- Morcillo-Suarez, C., J. Alegre, et al. (2008). "SNP analysis to results (SNPator): a web-based environment oriented to statistical genomics analyses upon SNP data." *Bioinformatics* **24**(14): 1643-1644.
- Moschini, M., M. Spahn, et al. (2016). "Incorporation of tissue-based genomic biomarkers into localized prostate cancer clinics." *BMC Med* **14**: 67.
- Mottet, N., J. Bellmunt, et al. (2015). The European Association of Urology (EAU) Guidelines on Prostate Cancer.
- Murtola, T. J., T. L. Tammela, et al. (2009). "Prostate cancer incidence among finasteride and alpha-blocker users in the Finnish Prostate Cancer Screening Trial." *Br J Cancer* **101**(5): 843-848.
- Nackley, A. G., S. A. Shabalina, et al. (2006). "Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure." *Science* **314**(5807): 1930-1933.
- Nam, R. K., W. W. Zhang, et al. (2009). "Utility of incorporating genetic variants for the early detection of prostate cancer." *Clin Cancer Res* **15**(5): 1787-1793.
- NCCN (2013). National Comprehensive Cancer Network. Prostate Cancer.
- Nemec, A. A., S. S. Wallace, et al. (2010). "Variant base excision repair proteins: contributors to genomic instability." *Semin Cancer Biol* **20**(5): 320-328.
- Neslund-Dudas, C., C. H. Bock, et al. (2007). "SRD5A2 and HSD3B2 polymorphisms are associated with prostate cancer risk and aggressiveness." *Prostate* **67**(15): 1654-1663.
- O'Donnell, P. H. and M. E. Dolan (2009). "Cancer pharmacoethnicity: ethnic differences in susceptibility to the effects of chemotherapy." *Clin Cancer Res* **15**(15): 4806-4814.
- Oesterling, J. E., S. J. Jacobsen, et al. (1995). "The use of age-specific reference ranges for serum prostate specific antigen in men 60 years old or older." *J Urol* **153**(4): 1160-1163.
- Oltvai, Z. N., C. L. Milliman, et al. (1993). "Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death." *Cell* **74**(4): 609-619.

- Onen, I. H., A. Ekmekci, et al. (2007). "The association of 5alpha-reductase II (SRD5A2) and 17 hydroxylase (CYP17) gene polymorphisms with prostate cancer patients in the Turkish population." DNA Cell Biol **26**(2): 100-107.
- Parekh, D. J., D. P. Ankerst, et al. (2007). "Biomarkers for prostate cancer detection." J Urol **178**(6): 2252-2259.
- Park, J. Y., Y. Huang, et al. (2009). "Single nucleotide polymorphisms in DNA repair genes and prostate cancer risk." Methods Mol Biol **471**: 361-385.
- Parliament, M. B. and D. Murray (2010). "Single nucleotide polymorphisms of DNA repair genes as predictors of radioresponse." Semin Radiat Oncol **20**(4): 232-240.
- Partin, A. W., M. K. Brawer, et al. (1998). "Prospective evaluation of percent free-PSA and complexed-PSA for early detection of prostate cancer." Prostate Cancer Prostatic Dis **1**(4): 197-203.
- Partin, A. W., M. W. Kattan, et al. (1997). "Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update." JAMA **277**(18): 1445-1451.
- Partin, A. W., L. A. Mangold, et al. (2001). "Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium." Urology **58**(6): 843-848.
- Partin, A. W., L. Van Neste, et al. (2014). "Clinical validation of an epigenetic assay to predict negative histopathological results in repeat prostate biopsies." J Urol **192**(4): 1081-1087.
- Partin, A. W., J. Yoo, et al. (1993). "The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer." J Urol **150**(1): 110-114.
- Patnala, R., J. Clements, et al. (2013). "Candidate gene association studies: a comprehensive guide to useful in silico tools." BMC Genet **14**: 39.
- Paz-y-Mino, C., T. Witte, et al. (2009). "Association among polymorphisms in the steroid 5alpha-reductase type II (SRD5A2) gene, prostate cancer risk, and pathologic characteristics of prostate tumors in an Ecuadorian population." Cancer Genet Cytogenet **189**(2): 71-76.
- Pegg, A. E. (1990). "Mammalian O6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents." Cancer Res **50**(19): 6119-6129.
- Pepe, M. S., R. Etzioni, et al. (2001). "Phases of biomarker development for early detection of cancer." J Natl Cancer Inst **93**(14): 1054-1061.
- Pepe, P. and F. Aragona (2011). "PCA3 score vs PSA free/total accuracy in prostate cancer diagnosis at repeat saturation biopsy." Anticancer Res **31**(12): 4445-4449.
- Pino-Yanes, M., A. Corrales, et al. (2011). "North African influences and potential bias in case-control association studies in the Spanish population." PLoS One **6**(3): e18389.
- Pollack, A., D. Cowen, et al. (2003). "Molecular markers of outcome after radiotherapy in patients with prostate carcinoma: Ki-67, bcl-2, bax, and bcl-x." Cancer **97**(7): 1630-1638.
- Pollak, M., W. Beamer, et al. (1998). "Insulin-like growth factors and prostate cancer." Cancer Metastasis Rev **17**(4): 383-390.
- Quaglia, A., S. Parodi, et al. (2003). "Differences in the epidemic rise and decrease of prostate cancer among geographical areas in Southern Europe. an analysis of differential trends in incidence and mortality in France, Italy and Spain." Eur J Cancer **39**(5): 654-665.
- Rando, J. C., V. M. Cabrera, et al. (1999). "Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands." Ann Hum Genet **63**(Pt 5): 413-428.
- Reich, D. E. and D. B. Goldstein (2001). "Detecting association in a case-control study while correcting for population stratification." Genet Epidemiol **20**(1): 4-16.
- Rini, B. I. and E. J. Small (2002). "Hormone-refractory Prostate Cancer." Curr Treat Options Oncol **3**(5): 437-446.

- Ritchev, J. D., W. Y. Huang, et al. (2005). "Genetic variants of DNA repair genes and prostate cancer: a population-based study." Cancer Epidemiol Biomarkers Prev **14**(7): 1703-1709.
- Rivera-Calzada, A., L. Spagnolo, et al. (2007). "Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs." EMBO Rep **8**(1): 56-62.
- Roach, M., 3rd (1993). "Re: The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer." J Urol **150**(6): 1923-1924.
- Roach, M., 3rd, C. Marquez, et al. (1994). "Predicting the risk of lymph node involvement using the pre-treatment prostate specific antigen and Gleason score in men with clinically localized prostate cancer." Int J Radiat Oncol Biol Phys **28**(1): 33-37.
- Ronen, A. and B. W. Glickman (2001). "Human DNA repair genes." Environ Mol Mutagen **37**(3): 241-283.
- Roses, A. D. (2000). "Pharmacogenetics and the practice of medicine." Nature **405**(6788): 857-865.
- Ross, A. E., Z. Feng, et al. (2012). "Effect of treatment with 5-alpha reductase inhibitors on progression in monitored men with favourable-risk prostate cancer." BJU Int **110**(5): 651-657.
- Rybicki, B. A., D. V. Conti, et al. (2004). "DNA repair gene XRCC1 and XPD polymorphisms and risk of prostate cancer." Cancer Epidemiol Biomarkers Prev **13**(1): 23-29.
- Rycaj, K., E. J. Cho, et al. (2016). "Longitudinal tracking of subpopulation dynamics and molecular changes during LNCaP cell castration and identification of inhibitors that could target the PSA-/lo castration-resistant cells." Oncotarget **7**(12): 14220-14240.
- Ryu, S. J., H. J. An, et al. (2008). "On the role of major vault protein in the resistance of senescent human diploid fibroblasts to apoptosis." Cell Death Differ **15**(11): 1673-1680.
- Sacco, A. G. and F. P. Worden (2016). "Molecularly targeted therapy for the treatment of head and neck cancer: a review of the ErbB family inhibitors." Onco Targets Ther **9**: 1927-1943.
- Sai, K. and Y. Saito (2011). "Ethnic differences in the metabolism, toxicology and efficacy of three anticancer drugs." Expert Opin Drug Metab Toxicol **7**(8): 967-988.
- Santivasi, W. L. and F. Xia (2014). "Ionizing radiation-induced DNA damage, response, and repair." Antioxid Redox Signal **21**(2): 251-259.
- Scariano, J. K., E. Treat, et al. (2008). "The SRD5A2 V89L polymorphism is associated with severity of disease in men with early onset prostate cancer." Prostate **68**(16): 1798-1805.
- Schaid, D. J. (2004). "The complex genetic epidemiology of prostate cancer." Hum Mol Genet **13 Spec No 1**: R103-121.
- Scheffer, G. L., P. L. Wijngaard, et al. (1995). "The drug resistance-related protein LRP is the human major vault protein." Nat Med **1**(6): 578-582.
- Schleutker, J. (2012). "Polymorphisms in androgen signaling pathway predisposing to prostate cancer." Mol Cell Endocrinol **360**(1-2): 25-37.
- Schmid, H. P. (1995). "Tumour markers in patients on deferred treatment: prostate specific antigen doubling times." Cancer Surv **23**: 157-167.
- Schreiber, V., J. C. Ame, et al. (2002). "Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1." J Biol Chem **277**(25): 23028-23036.
- Schultz, N., E. Lopez, et al. (2003). "Poly(ADP-ribose) polymerase (PARP-1) has a controlling role in homologous recombination." Nucleic Acids Res **31**(17): 4959-4964.
- Seng, K. C. and C. K. Seng (2008). "The success of the genome-wide association approach: a brief story of a long struggle." Eur J Hum Genet **16**(5): 554-564.

- Setiawan, V. W., F. R. Schumacher, et al. (2007). "CYP17 genetic variation and risk of breast and prostate cancer from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium (BPC3)." Cancer Epidemiol Biomarkers Prev **16**(11): 2237-2246.
- Setlur, S. R., C. X. Chen, et al. (2010). "Genetic variation of genes involved in dihydrotestosterone metabolism and the risk of prostate cancer." Cancer Epidemiol Biomarkers Prev **19**(1): 229-239.
- Shah, R. B., R. Mehra, et al. (2004). "Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program." Cancer Res **64**(24): 9209-9216.
- Sharpless, N. E., D. O. Ferguson, et al. (2001). "Impaired nonhomologous end-joining provokes soft tissue sarcomas harboring chromosomal translocations, amplifications, and deletions." Mol Cell **8**(6): 1187-1196.
- Shastry, B. S. (2009). "SNPs: impact on gene function and phenotype." Methods Mol Biol **578**: 3-22.
- Shibata, A., M. I. Garcia, et al. (2002). "Polymorphisms in the androgen receptor and type II 5 alpha-reductase genes and prostate cancer prognosis." Prostate **52**(4): 269-278.
- Shiloh, Y. and G. Rotman (1996). "Ataxia-telangiectasia and the ATM gene: linking neurodegeneration, immunodeficiency, and cancer to cell cycle checkpoints." J Clin Immunol **16**(5): 254-260.
- Shimamoto, Y., T. Sumizawa, et al. (2006). "Direct activation of the human major vault protein gene by DNA-damaging agents." Oncol Rep **15**(3): 645-652.
- Shipitsin, M., C. Small, et al. (2014). "Automated quantitative multiplex immunofluorescence in situ imaging identifies phospho-S6 and phospho-PRAS40 as predictive protein biomarkers for prostate cancer lethality." Proteome Sci **12**: 40.
- Shrivastav, M., L. P. De Haro, et al. (2008). "Regulation of DNA double-strand break repair pathway choice." Cell Res **18**(1): 134-147.
- Signorelli, S. S., V. Barresi, et al. (2008). "Polymorphisms of steroid 5-alpha-reductase type I (SRD5A1) gene are associated to peripheral arterial disease." J Endocrinol Invest **31**(12): 1092-1097.
- Sobin, L. H., M. Gospodarowicz, et al. (2009). TNM classification of malignant tumors. UICC International Union Against Cancer, Wiley-Blackwell.
- Stanczyk, F. Z., C. G. Azen, et al. (2013). "Effect of finasteride on serum levels of androstenedione, testosterone and their 5alpha-reduced metabolites in men at risk for prostate cancer." J Steroid Biochem Mol Biol **138**: 10-16.
- Suga, T., M. Iwakawa, et al. (2008). "Influence of multiple genetic polymorphisms on genitourinary morbidity after carbon ion radiotherapy for prostate cancer." Int J Radiat Oncol Biol Phys **72**(3): 808-813.
- Tanteles, G. A., R. J. Murray, et al. (2012). "Variation in telangiectasia predisposing genes is associated with overall radiation toxicity." Int J Radiat Oncol Biol Phys **84**(4): 1031-1036.
- Thomas, C. and A. V. Tulin (2013). "Poly-ADP-ribose polymerase: machinery for nuclear processes." Mol Aspects Med **34**(6): 1124-1137.
- Thompson, I. M., P. J. Goodman, et al. (2003). "The influence of finasteride on the development of prostate cancer." N Engl J Med **349**(3): 215-224.
- Thompson, L. H. and M. G. West (2000). "XRCC1 keeps DNA from getting stranded." Mutat Res **459**(1): 1-18.
- Tong, W. M., U. Cortes, et al. (2001). "Poly(ADP-ribose) polymerase: a guardian angel protecting the genome and suppressing tumorigenesis." Biochim Biophys Acta **1552**(1): 27-37.
- Trock, B. J., M. J. Brozman, et al. (2012). "Evaluation of GSTP1 and APC methylation as indicators for repeat biopsy in a high-risk cohort of men with negative initial prostate biopsies." BJU Int **110**(1): 56-62.

- Tseng, R. C., F. J. Hsieh, et al. (2009). "Lung cancer susceptibility and prognosis associated with polymorphisms in the nonhomologous end-joining pathway genes: a multiple genotype-phenotype study." *Cancer* **115**(13): 2939-2948.
- Van den Broeck, T., S. Joniau, et al. (2014). "The role of single nucleotide polymorphisms in predicting prostate cancer risk and therapeutic decision making." *Biomed Res Int* **2014**: 627510.
- van Zon, A., M. H. Mossink, et al. (2003). "The vault complex." *Cell Mol Life Sci* **60**(9): 1828-1837.
- Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." *Science* **291**(5507): 1304-1351.
- Verma, R., V. Gupta, et al. (2015). "Significance of p53 and ki-67 expression in prostate cancer." *Urol Ann* **7**(4): 488-493.
- Vickers, A. J., A. Gupta, et al. (2011). "A panel of kallikrein marker predicts prostate cancer in a large, population-based cohort followed for 15 years without screening." *Cancer Epidemiol Biomarkers Prev* **20**(2): 255-261.
- Wacholder, S., N. Rothman, et al. (2000). "Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias." *J Natl Cancer Inst* **92**(14): 1151-1158.
- Walsh, P. C. (2010). "Chemoprevention of prostate cancer." *N Engl J Med* **362**(13): 1237-1238.
- Wang, F., Y. F. Zou, et al. (2011). "CYP17 gene polymorphisms and prostate cancer risk: a meta-analysis based on 38 independent studies." *Prostate* **71**(11): 1167-1177.
- Wang, M., W. Wu, et al. (2006). "PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways." *Nucleic Acids Res* **34**(21): 6170-6182.
- Wang, W. Y., B. J. Barratt, et al. (2005). "Genome-wide association studies: theoretical and practical concerns." *Nat Rev Genet* **6**(2): 109-118.
- Wegiel, B., A. Bjartell, et al. (2005). "A role for cyclin A1 in mediating the autocrine expression of vascular endothelial growth factor in prostate cancer." *Oncogene* **24**(42): 6385-6393.
- West, C., B. S. Rosenstein, et al. (2010). "Establishment of a Radiogenomics Consortium." *Int J Radiat Oncol Biol Phys* **76**(5): 1295-1296.
- West, C. M. and G. C. Barnett (2011). "Genetics and genomics of radiotherapy toxicity: towards prediction." *Genome Med* **3**(8): 52.
- West, C. M., A. M. Dunning, et al. (2012). "Genome-wide association studies and prediction of normal tissue toxicity." *Semin Radiat Oncol* **22**(2): 91-99.
- West, S. C. (2003). "Molecular views of recombination proteins and their control." *Nat Rev Mol Cell Biol* **4**(6): 435-445.
- Wikstrom, P., P. Stattin, et al. (1998). "Transforming growth factor beta1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer." *Prostate* **37**(1): 19-29.
- Wilt, T. J., R. MacDonald, et al. (2008). "Five-alpha-reductase Inhibitors for prostate cancer prevention." *Cochrane Database Syst Rev*(2): CD007091.
- Woelfelschneider, A., O. Popanda, et al. (2008). "A distinct ERCC1 haplotype is associated with mRNA expression levels in prostate cancer patients." *Carcinogenesis* **29**(9): 1758-1764.
- Wood, R. D. (1996). "DNA repair in eukaryotes." *Annu Rev Biochem* **65**: 135-167.
- Wood, R. D., M. Mitchell, et al. (2001). "Human DNA repair genes." *Science* **291**(5507): 1284-1289.
- Wright, J. L., E. M. Kwon, et al. (2010). "CYP17 polymorphisms and prostate cancer outcomes." *Prostate* **70**(10): 1094-1101.
- Xu, H., P. Zou, et al. (2013). "Association between the XRCC6 Promoter rs2267437 polymorphism and cancer risk: evidence based on the current literature." *Genet Test Mol Biomarkers* **17**(8): 607-614.

- Xu, J., S. L. Zheng, et al. (2010). "Inherited genetic variant predisposes to aggressive but not indolent prostate cancer." *Proc Natl Acad Sci U S A* **107**(5): 2136-2140.
- Xu, Z., L. X. Hua, et al. (2007). "Relationship between XRCC1 polymorphisms and susceptibility to prostate cancer in men from Han, Southern China." *Asian J Androl* **9**(3): 331-338.
- Yang, L. P. (2011). "Abiraterone acetate: in metastatic castration-resistant prostate cancer." *Drugs* **71**(15): 2067-2077.
- Yarden, Y. (2001). "Biology of HER2 and its importance in breast cancer." *Oncology* **61 Suppl 2**: 1-13.
- Yasuda, S. U., L. Zhang, et al. (2008). "The role of ethnicity in variability in response to drugs: focus on clinical pharmacology studies." *Clin Pharmacol Ther* **84**(3): 417-423.
- Yu, Z., J. Chen, et al. (1999). "Human DNA repair systems: an overview." *Environ Mol Mutagen* **33**(1): 3-20.
- Zamble, D. B. and S. J. Lippard (1995). "Cisplatin and DNA repair in cancer chemotherapy." *Trends Biochem Sci* **20**(10): 435-439.
- Zelevsky, M. J., T. Hollister, et al. (2000). "Five-year biochemical outcome and toxicity with transperineal CT-planned permanent I-125 prostate implantation for patients with localized prostate cancer." *Int J Radiat Oncol Biol Phys* **47**(5): 1261-1266.
- Zelevsky, M. J., M. W. Kattan, et al. (2007). "Pretreatment nomogram predicting ten-year biochemical outcome of three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for prostate cancer." *Urology* **70**(2): 283-287.
- Zelevsky, M. J., X. Pei, et al. (2011). "Dose escalation for prostate cancer radiotherapy: predictors of long-term biochemical tumor control and distant metastases-free survival outcomes." *Eur Urol* **60**(6): 1133-1139.
- Zhang, J. (2014). "Poly (ADP-ribose) polymerase inhibitor: an evolving paradigm in the treatment of prostate cancer." *Asian J Androl* **16**(3): 401-406.
- Zhang, L., M. Yang, et al. (2010). "Association of TGF-beta1 and XPD polymorphisms with severe acute radiation-induced esophageal toxicity in locally advanced lung cancer patients treated with radiotherapy." *Radiother Oncol* **97**(1): 19-25.
- Zhao, L., N. Yu, et al. (2014). "Tissue biomarkers for prognosis of prostate cancer: a systematic review and meta-analysis." *Cancer Epidemiol Biomarkers Prev* **23**(6): 1047-1054.
- Zhu, Y., M. R. Spitz, et al. (2004). "An evolutionary perspective on single-nucleotide polymorphism screening in molecular cancer epidemiology." *Cancer Res* **64**(6): 2251-2257.

12.- Anexo: artículos generados a partir de la presente Tesis Doctoral.

Polymorphisms in DNA-Repair Genes in a Cohort of Prostate Cancer Patients from Different Areas in Spain: Heterogeneity between Populations as a Confounding Factor in Association Studies

Luis Alberto Henríquez-Hernández^{1,2,3*}, Almudena Valenciano², Palmira Foro-Arnalot⁴, María Jesús Álvarez-Cubero^{5,6}, José Manuel Cozar⁷, José Francisco Suárez-Novo⁸, Manel Castells-Esteve⁸, Adriana Ayala-Gil⁹, Pablo Fernández-Gonzalo⁹, Montse Ferrer¹⁰, Ferrán Guedea¹¹, Gemma Sancho-Pardo¹², Jordi Craven-Bartle¹², María José Ortiz-Gordillo¹³, Patricia Cabrera-Roldán¹³, Estefanía Herrera-Ramos¹⁴, Pedro C. Lara^{1,2,3}

1 Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain, **2** Instituto Canario de Investigación del Cáncer, Las Palmas, Spain, **3** Clinical Science Department, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain, **4** Institut d'Oncologia Radiooncològica, Hospital de la Esperanza, Parc de Salut Mar, Barcelona, Spain, **5** Laboratory of Genetic Identification, Legal Medicine and Toxicology Department, Facultad de Medicina, Universidad de Granada, Granada, Spain, **6** GENYO, Pfizer-University of Granada-Andalusian Government Centre for Genomics and Oncological Research, Granada, Spain, **7** Department of Urology, Hospital Universitario Virgen de las Nieves, Granada, Spain, **8** Department of Urology, Hospital Universitario de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain, **9** Radiation Oncology Department, Onkologikoa, Galdakao, Spain, **10** Health Services Research Group, Institut de Recerca Hospital del MarIMM, Barcelona, Spain, **11** Department of Radiation Oncology, Institut Català d'Oncologia ICO, L'Hospitalet de Llobregat, Barcelona, Spain, **12** Radiation Oncology Department, Hospital de la Santa Creu i Sant Pius, Barcelona, Spain, **13** Radiation Oncology Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain, **14** Immunology Department, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain

Abstract

Background: Differences in the distribution of genotypes between individuals of the same ethnicity are an important confounder factor commonly undervalued in typical association studies conducted in radiogenomics.

Objective: To evaluate the genotypic distribution of SNPs in a wide set of Spanish prostate cancer patients for determine the homogeneity of the population and to disclose potential bias.

Design, Setting, and Participants: A total of 601 prostate cancer patients from Andalusia, Basque Country, Canary and Catalonia were genotyped for 10 SNPs located in 6 different genes associated to DNA repair: XRCC1 (rs25487, rs25489, rs1799782), ERCC2 (rs13181), ERCC1 (rs11615), LIG4 (rs1805388, rs1805386), ATM (rs17503908, rs1800057) and P53 (rs1042522). The SNP genotyping was made in a Biotrove OpenArray[®] NT Cycler.

Outcome Measurements and Statistical Analysis: Comparisons of genotypic and allelic frequencies among populations, as well as haplotype analyses were determined using the web-based environment SNPator. Principal component analysis was made using the SnpMatrix and XSNPMatrix classes and methods implemented as an R package. Non-supervised hierarchical cluster of SNP was made using MultiExperiment Viewer.

Results and Limitations: We observed that genotype distribution of 4 out 10 SNPs was statistically different among the studied populations, showing the greatest differences between Andalusia and Catalonia. These observations were confirmed in cluster analysis, principal component analysis and in the differential distribution of haplotypes among the populations. Because tumor characteristics have not been taken into account, it is possible that some polymorphisms may influence tumor characteristics in the same way that it may pose a risk factor for other disease characteristics.

Conclusion: Differences in distribution of genotypes within different populations of the same ethnicity could be an important confounding factor responsible for the lack of validation of SNPs associated with radiation-induced toxicity, especially when extensive meta-analysis with subjects from different countries are carried out.

Citation: Henríquez-Hernández LA, Valenciano A, Foro-Arnalot P, Álvarez-Cubero MJ, Cozar JM, et al. (2013) Polymorphisms in DNA-Repair Genes in a Cohort of Prostate Cancer Patients from Different Areas in Spain: Heterogeneity between Populations as a Confounding Factor in Association Studies. PLoS ONE 8(7): e69735. doi:10.1371/journal.pone.0069735

Editor: Amanda Ewart Toland, Ohio State University Medical Center, United States of America

Received: May 2, 2013; **Accepted:** June 12, 2013; **Published:** July 23, 2013

Copyright: © 2013 Henríquez-Hernández et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was subsidized by a grant from the Instituto de Salud Carlos III (Ministerio de Economía y Competitividad from Spain), ID: PI12/01867. Almudena Valenciano has a grant from the Instituto Canario de Investigación del Cáncer (ICC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: henrique@icc.upgccs

Introduction

Genetic polymorphisms are variants of the genome that appear by mutations in some individuals, are transmitted to offspring and acquire some frequency (at least 1%) in the population after many generations. Polymorphisms are the basis of evolution and those that are consolidated may be silent or provide benefits to individuals, but can also be involved in disease development [1]. The most frequent polymorphisms are single nucleotide polymorphisms (SNPs). The ethnic origin of a population determines the distribution of genotypes in a population, and has not to be equal to others. Moreover, differences observed within populations of the same ethnic origin suggest that race is not a sufficient factor to ensure the homogeneity of the sample. In that sense, it is known the presence of several significant axes of stratification, most prominently in a northern-south-eastern trend, but also along an east-west axis, among the genotype distribution of European population [2]. In the case of Spain, although populations inhabiting the Iberian Peninsula show a substantial genetic homogeneity [3], there are findings suggesting that Northwest African influences existing among the Spanish populations and these differences might increase the risk for false positives in genetic epidemiology studies [4].

Radiation therapy (RT) is an effective treatment offered to patients with localized prostate cancer as a viable alternative to surgery [5]. Although both therapies showed comparable results in terms of survival [6], the main differences between them are related to adverse effects. Tumour control by RT requires the use of maximum dose that can be delivered while maintaining a tolerance risk of normal tissue toxicity, being clinical toxicity the factor limiting the efficacy of the treatment [7]. The role of genetics in the response of normal tissues to RT is widely accepted by the scientific community, and it would help to explain why patients treated with RT experience a large variation in normal tissue toxicity, even when similar doses and schedules are administered [8]. Radiation causes the loss of structure and function of most biologic molecules, including DNA. The individual DNA repair capacity consists of several mechanisms (nucleotide and base excision repair, homologous recombination, non-homologous endjoining, mismatch repair and telomere metabolism) and the individual capacity to repair damaged DNA may modify the response of tumour tissue and normal tissue to radiation [9]. Thus, studies of candidate genes have been focused on genes mainly involved in DNA damage recognition and repair (eg, ATM, XRCC1, XPD, ERCC1, LIG4, and TP53 among others), and also in free radical scavenging (eg, SOD2), or anti-inflammatory response (eg, TGFB1).

The association between SNPs and radiation toxicity has been deeply explored [10] and numerous consortia have been formed to identify common genetic variations associated with the development of radiation toxicity [11]. Although promising, the overall results failed at the validation stage [12] and today, the development of a SNP signature associated to the prediction of toxicity is still far away. Although this lack of association could be explained by different reasons (presence of confounding factors, insufficient sample size, and lack of consensus in the methodology

used in terms of genotyping, statistics, and even in the grading of radiation toxicity) [13], the heterogeneity of the studied populations is a factor whose effect has been commonly underestimated.

With all those assumptions in mind, we designed a study aimed to evaluate the genotypic distribution of 10 SNPs in 6 different genes involved in DNA repair and classically associated to radiation-induced toxicity, in a wide set of Spanish prostate cancer patients, to determine the homogeneity of the population and to disclose potential undervalued confounders in the association between SNPs and radiation toxicity.

Materials and Methods

1. Patients

A total of 601 patients with non-metastatic localized prostate cancer (PCa) were included in the study. Geographical distribution of patients was as follows (Table 1): 91 (15.14%) from Andalusia, 51 (8.48%) from Basque Country, 238 (39.60%) from Canary and 221 (36.77%) from Catalonia. All patients were from Spanish origin and all of them received written informed consent before sample collection. All participants provided their written informed consent to participate in this study. The study was approved by the Research and Ethics Committee of each institution participant in the study: Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria), Hospital de la Esperanza, Parc de Salut Mar (Barcelona), Hospital Universitario Virgen de las Nieves (Granada), Hospital Universitari de Bellvitge (L'Hospitalet de Llobregat), Onkologikoa (Guipuzcoa), Institut Català d'Oncologia (L'Hospitalet de Llobregat), Hospital de la Santa Creu i Sant Pau (Barcelona) and Hospital Universitario Virgen del Rocío (Sevilla).

2. DNA Isolation and Quantification

All the blood samples were sent to the Hospital Universitario de Gran Canaria Dr. Negrín for DNA extraction and subsequent analyses. DNA was isolated from 300 µl of whole-blood in an iPrep purification system (Applied Biosystems, Foster City, CA) using the iPrep™ PureLink™ gDNA Blood Kit (Applied Biosystems). DNA was quantified and the quality of samples was determined in a NanoDrop 2000 (Thermo Scientific, Wilmington, DE).

Table 1. Regional ancestry of study participants.

Regional ancestry	n	(%)	No. of hospitals
Andalusia	91	(15.14)	2
Basque Country	51	(8.48)	1
Canary	238	(39.60)	1
Catalonia	221	(36.77)	4
Total	601	(100)	8

doi:10.1371/journal.pone.0069735.t001

Table 2. Description of SNPs included in the study and analyzed by OpenArray.

Gene name	Symbol	Assay ID	SNP ID	Alleles	Chr	Position
X-ray repair complementing defective repair in Chinese hamster cells 1						
	XRCC1	C_622564_10	rs25487	C/T	19q13	44055726
	XRCC1	C_622570_10	rs25489	C/T	19q13	44056412
	XRCC1	C_11463404_10	rs1799782	A/G	19q13	44057574
Excision repair cross-complementing rodent repair deficiency, complementation group 2						
	ERCC2/XPD	C_3145033_10	rs13181	G/T	19q13	45854019
Excision repair cross-complementing rodent repair deficiency, complementation group 1						
	ERCC1	C_2532959_1_	rs11615	A/G	19q13	45923653
Ligase IV						
	LIG4	C_11427969_20	rs1805388	A/G	13q33	108863591
	LIG4	C_11427968_10	rs1805386	A/G	13q33	108861913
Ataxia telangiectasia mutated						
	ATM	C_33307908_10	rs17503908	G/T	11q22	108215397
	ATM	C_45273750_10	rs1800057	C/G	11q22	108143456
Tumour protein P53						
	P53	C_2403545_10	rs1042522	C/G	17p13	7579472

Abbreviations: Chr, chromosome; C, cytosine; T, thymine; A, adenine; G, guanine. All the assays were commercially available at Applied Biosystems (see Assay ID). doi:10.1371/journal.pone.0069735.t002

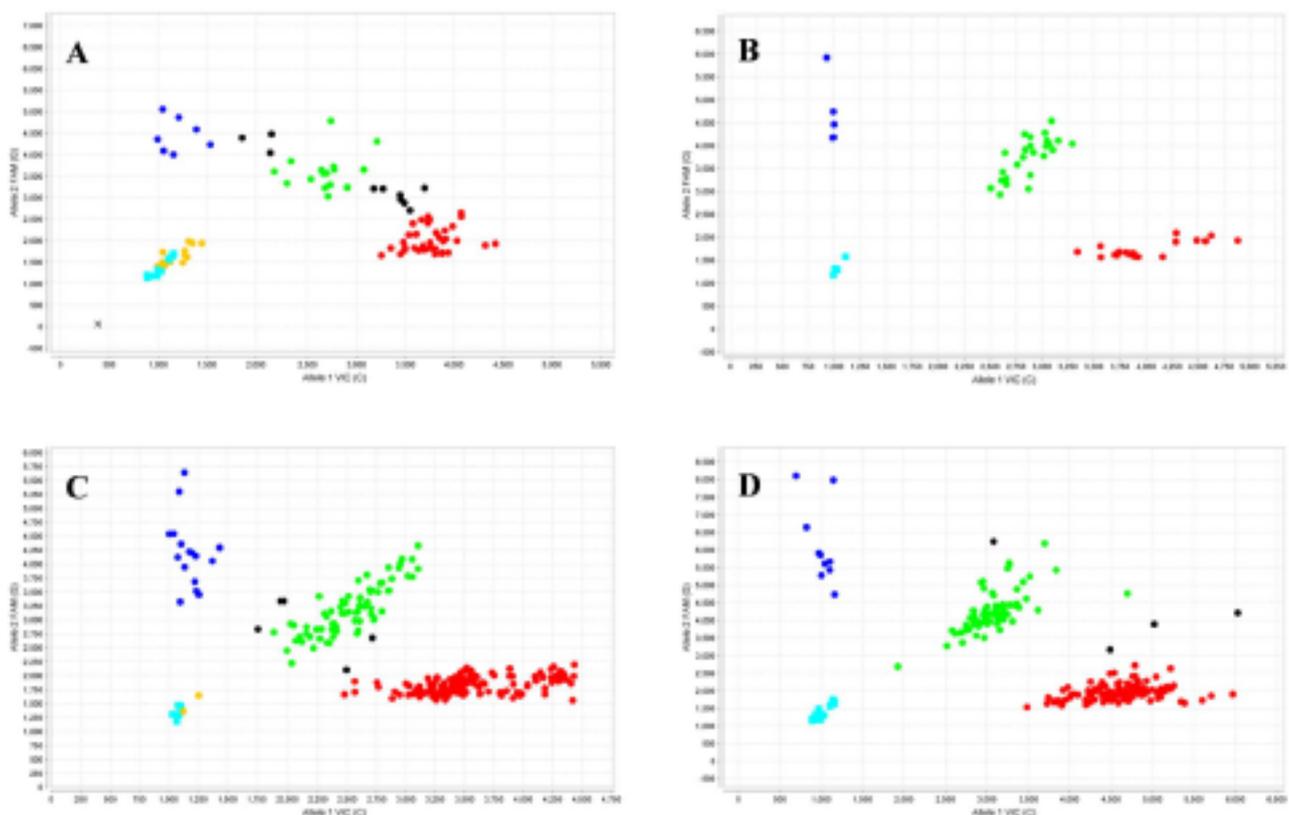


Figure 1. Scatter plots showing genotyping of SNP rs1042522 in (A) Andalusia, (B) Basque Country, (C) Canary, and (D) Catalonia using a Biotrove OpenArray[®] NT Cycler. Each graph depicts a scattered plot of one allele (FAM probe) versus the other allele (VIC probe). Those samples that are homozygous appear in blue or red; heterozygous samples contain fluorescence from both probes and appear in green. The NTCs appear in light-blue squares and represent the background fluorescence from samples with no template DNA. Samples non-determined appear as black points and samples not amplified appear as orange points. The scatter plots were obtained from the TaqMan Genotyper software version 1.0.1. doi:10.1371/journal.pone.0069735.g001

3. Genes and SNPs

A total of 10 SNPs in 6 different key genes involved in DNA repair were studied: X-ray repair cross-complementing protein 1 (XRCC1) [14,15], excision repair cross-complementing rodent repair deficiency, complementation group 2 (ERCC2) [16], excision repair cross-complementing rodent repair deficiency, complementation group 1 (ERCC1) [17], ligase IV (LIG4) [18], ataxia telangiectasia mutated (ATM) [19], and tumour protein p53 (TP53) [20]. Because RT acts producing DNA damage and genetic variation in DNA repair and damage response modify the susceptibility to radiotherapy, these SNPs have been classically associated to radiation-induced toxicity in several tumor types. Description of SNPs is contained in Table 2.

4. Genotyping

The SNP genotyping was made in a Biotrove OpenArray® NT Cycler (Applied Biosystems). DNA for OpenArray (OA) was diluted at a concentration of 50 ng/μl and a ratio of A260/A280

and A260/230 of 1.7–1.9. A total of 300 ng of genomic DNA was used. A final amount of 150 ng was incorporated into the array with the autoloader and genotyped according to the manufacturer's recommendations. A non-template control (NTC) consisting of DNase-free double-distilled water was introduced within each assay. When the DNA and master mix were transferred, the loaded OA plate was filled with an immersion fluid and sealed with glue. The multiplex TaqMan assay reactions were carried out in a Dual Flat Block (384-well) GeneAmp PCR System 9700 (Applied Biosystems) with the following PCR cycle: an initial step at 93°C for 10 minutes followed by 50 cycles of 45 seconds at 95°C, 13 seconds at 94°C and 2 minutes, 14 seconds at 53°C; followed by a final step during 2 minutes at 25°C and holding at 4°C.

The fluorescence results were read using the OpenArray® SNP Genotyping Analysis software version 1.0.5. (Applied Biosystems). The genotyping analysis was made with the TaqMan Genotyper software version 1.0.1. (available at: <http://www.invitrogen.com/>

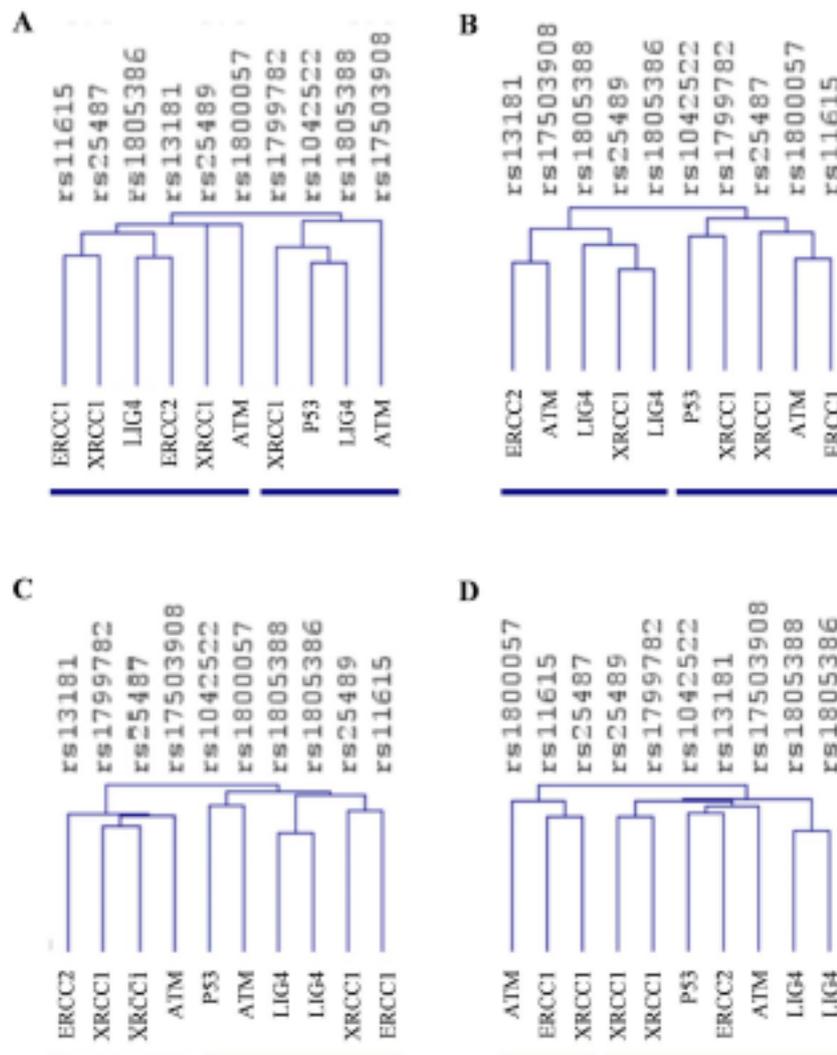


Figure 2. Non-supervised hierarchical clustering of SNPs in prostate cancer patients from (A) Andalusia, (B) Basque Country, (C) Canary and (D) Catalonia. Clustering was made using Euclidean distance correlation and average linkage, and was processed and displayed with MultiExperiment Viewer (<http://www.tigr.org>). The dendrogram shows clustering of SNPs. The gene symbol was added to identify each SNP. Lines below each panel shows the two main clusters generated.
doi:10.1371/journal.pone.0069735.g002

Table 3. Genotype and allelic frequencies of gene polymorphisms in this study.

	Call rate	Genotypes			HWE	Alleles	
XRCC1							
rs25487		CC	CT	TT		C	T
Andalusia	0.79	0.49	0.34	0.18	ns	0.65	0.35
Basque Country	0.80	0.44	0.51	0.05	ns	0.70	0.30
Canary	0.95	0.48	0.41	0.11	ns	0.68	0.32
Catalonia	0.83	0.36	0.55	0.09	*	0.63	0.37
P value					0.012		
rs25489		CC	CT	TT		C	T
Andalusia	0.91	0.81	0.19	0.00	ns	0.90	0.10
Basque Country	1.00	0.86	0.14	0.00	ns	0.93	0.07
Canary	0.97	0.87	0.13	0.00	ns	0.93	0.07
Catalonia	0.98	0.90	0.09	0.01	ns	0.95	0.05
P value					0.178		
rs1799782		AA	AG	GG		A	G
Andalusia	0.85	0.00	0.09	0.91	ns	0.05	0.95
Basque Country	1.00	0.00	0.08	0.92	ns	0.04	0.96
Canary	0.99	0.01	0.12	0.87	ns	0.07	0.93
Catalonia	0.98	0.01	0.11	0.88	ns	0.06	0.94
P value					0.936		
ERCC2							
rs13181		GG	GT	TT		G	T
Andalusia	0.74	0.19	0.15	0.66	*	0.27	0.73
Basque Country	1.00	0.06	0.37	0.57	ns	0.25	0.75
Canary	0.98	0.11	0.45	0.44	ns	0.33	0.67
Catalonia	0.97	0.09	0.53	0.38	ns	0.35	0.65
P value					0.0001		
ERCC1							
rs11615		AA	AG	GG		A	G
Andalusia	0.70	0.58	0.20	0.22	*	0.68	0.32
Basque Country	1.00	0.43	0.43	0.14	ns	0.65	0.35
Canary	0.98	0.43	0.41	0.16	ns	0.63	0.37
Catalonia	0.99	0.32	0.52	0.16	ns	0.58	0.42
P value					0.001		
LIG4							
rs1805388		AA	AG	GG		A	G
Andalusia	0.74	0.06	0.12	0.82	ns	0.12	0.88
Basque Country	0.98	0.04	0.38	0.58	ns	0.23	0.77
Canary	0.99	0.03	0.25	0.72	ns	0.15	0.85
Catalonia	0.99	0.05	0.22	0.73	ns	0.16	0.84
P value					0.051		
rs1805386		AA	AG	GG		A	G
Andalusia	0.85	0.78	0.16	0.06	ns	0.85	0.15
Basque Country	0.98	0.84	0.16	0.00	ns	0.92	0.08
Canary	0.96	0.73	0.25	0.02	ns	0.85	0.15
Catalonia	0.98	0.66	0.28	0.06	ns	0.80	0.20
P value					0.029		
ATM							
rs17503908		GG	GT	TT		G	T

Table 3. Cont.

	Call rate	Genotypes			HWE	Alleles	
Andalusia	0.81	0.03	0.08	0.89	ns	0.07	0.93
Basque Country	0.98	0.00	0.20	0.80	ns	0.10	0.90
Canary	0.99	0.01	0.20	0.79	ns	0.10	0.90
Catalonia	0.98	0.01	0.17	0.82	ns	0.09	0.91
P value					0.088		
rs1800057		CC	CG	GG		C	G
Andalusia	0.80	1.00	0.00	0.00	*	1.00	0.00
Basque Country	1.00	0.94	0.06	0.00	ns	0.97	0.03
Canary	0.97	0.95	0.04	0.01	ns	0.97	0.03
Catalonia	0.99	0.92	0.08	0.00	ns	0.96	0.04
P value					0.186		
TP53							
rs1042522		CC	CG	GG		C	G
Andalusia	0.68	0.63	0.26	0.11	ns	0.76	0.24
Basque Country	1.00	0.41	0.49	0.10	ns	0.66	0.34
Canary	0.97	0.61	0.32	0.07	ns	0.77	0.23
Catalonia	0.98	0.60	0.35	0.05	ns	0.78	0.22
P value					0.059		

Statistical differences among genotypes and Hardy-Weinberg equilibrium (HWE) are shown. Abbreviations: ns, non-significant. Differences in the genotype distribution were assessed by χ^2 test. Populations showing no HWE were indicated with an asterisk ($P < 0.01$).

doi:10.1371/journal.pone.0069735.t003

site/us/en/home/Globa1/Forms/taqman-genotyper-software-download-reg.html) using autocalling as the call method. The quality value of the data points genotype was determined by a threshold above 0.95. Genotyping analysis was done for each population separately (Figure 1).

5. Statistical Analysis

Genotype and allelic frequencies were determined using the web-based environment SNPator (SNP Analysis To Results, from the Spain's National Genotyping Center and the National Institute for Bioinformatic) [21]. Relative excess heterozygosity was determined to check compatibility of genotype frequencies with Hardy-Weinberg equilibrium (HWE). Thus, p-values from the standard exact HWE lack of fit test were calculated using SNPator. Comparisons of genotypic and allelic frequencies among populations, as well as haplotype analyses were also done in SNPator.

Principal component analysis (PCA) was made using the SnpMatrix and XSnMatrix classes and methods [22], implemented as an R package and available from Bioconductor (as of version 2.11; <http://bioconductor.org>). It consists in the transformation of the set of original variables in another set of variables – principal components – obtained as a linear combination of those. The new variables retain all the information, but most of the principal components have so small variability that can be ignored. Thus, few components (generally 3 or less) can represent and explain reasonably the set of objects of the sample without loss of information. PCA reduces the complexity of the data and permits the graphical representation of the variables.

Non-supervised hierarchical clustering [23] of SNP in each population was made using MultiExperiment Viewer (available at: www.igb.org). Clustering was made using Euclidean distance correlation and average linkage. To success perform the clusters,

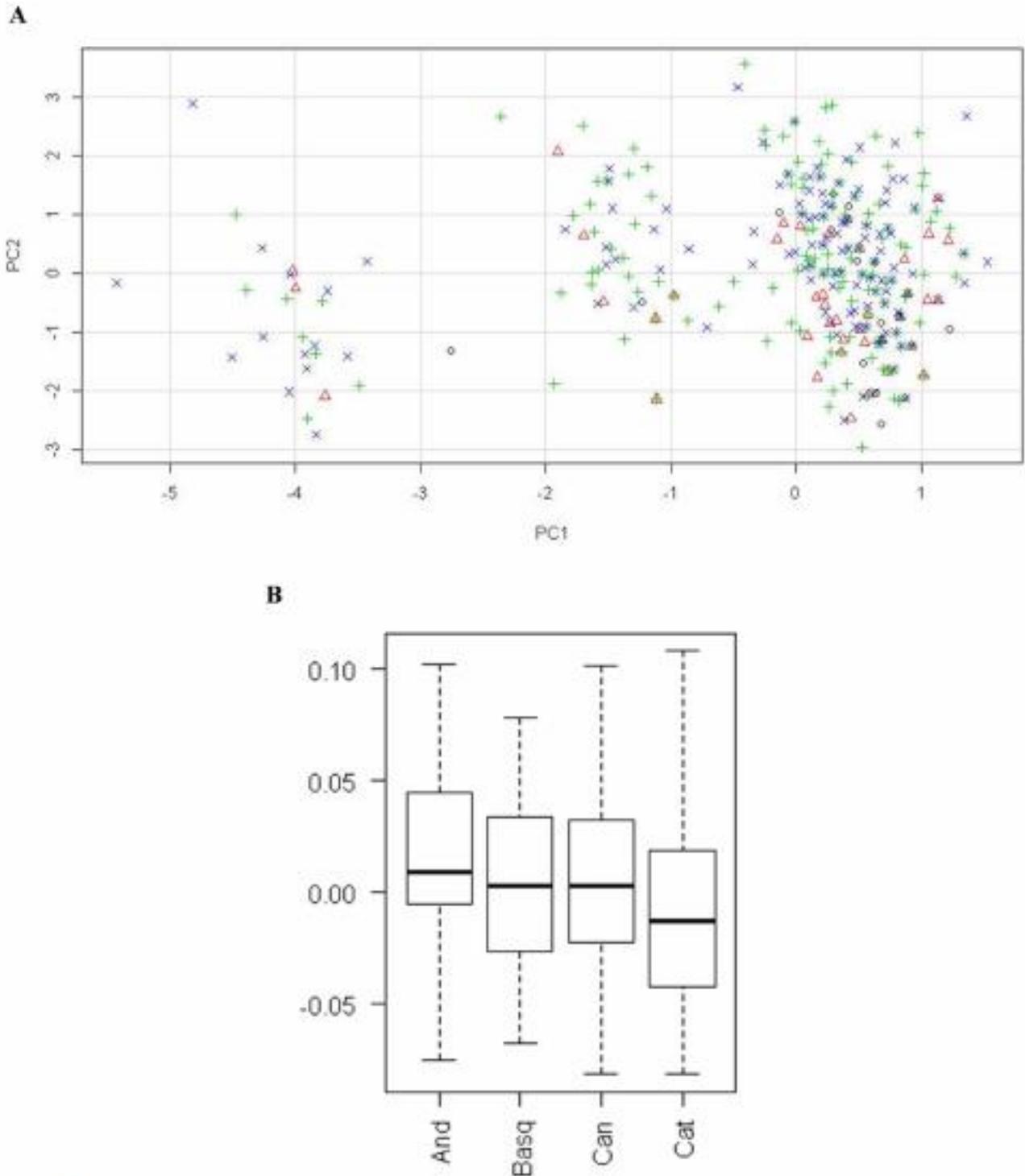


Figure 3. Plot of the top two principal components from the analysis of populations (A), and boxplot of component 1 among the different populations (B). Symbols in plot A: \circ (black), Andalusia; Δ (red), Basque Country; $+$ (green), Canary; \times (blue), Catalonia. Abbreviations in plot B: And, Andalusia; Basq, Basque Country; Can, Canary; Cat, Catalonia. doi:10.1371/journal.pone.0069735.g003

wild homozygous was encoded as -1 , heterozygous as 0 and mutated homozygous as 1 .

All additional statistical analyses were performed using PASW Statistics 15 (IBM Corporation, Armonk, NY, USA).

Results

All the genotyped samples met quality criteria as stated above, and all samples were genotyped with the same batch of material at

Table 4. Comparison among populations of allelic and genotypic frequencies.

Comparison	Allelic frequencies	Genotypic frequencies
Can vs. And	–	rs13181
Can vs. Basq	–	–
Can vs. Cat	–	–
And vs. Basq	–	rs1805388 rs13181
And vs. Cat	–	rs25487 rs13181 rs11615
Basq vs. Cat	rs1805386	–

SNPs differentially distributed are shown.

Abbreviations: Can, Canary; And, Andalusia; Basq, Basque Country; Cat, Catalonia. The analyses of the genotypic frequencies were performed including the three possible genotypes. Differences were significant with *p* values <0.01. doi:10.1371/journal.pone.0069735.t004

the same time. A total of 601 PCA patients were genotyped for 10 SNPs. Of the 6,010 possible determinations, 94.36% were successfully genotyped. The call rates among populations were (median (range)): 79.5% (68.1–91.2%) for Andalusia, 100% (80.4–100%) for Basque Country, 97.7% (94.5–99.2%) for Canary, and 97.9% (83.3–99.1%) for Catalonia.

The genotypic and allelic frequencies are shown in Table 3. A relative excess of heterozygosity, indicating a deviation from HWE, was observed in 4 SNPs from 2 different populations: rs25487 (XRCC1) in subjects from Catalonia and rs13181 (ERCC2), rs11615 (ERCC1) and rs180057 (ATM) in subjects from Andalusia (Table 3). The genotype distribution was different between the study populations in 4 of the 10 SNPs: rs25487, rs13181, rs11615, and rs1805386 (LIG4) (χ^2 test, Table 3), showing a differential distribution of genotypes among populations. A non-supervised hierarchical cluster was performed trying to visualize the differences in the genotype distributions among the four populations. Thus, as shown in Figure 2, polymorphisms were distributed into two main clusters, each one with different number

and identity of SNPs, suggesting heterogeneity among populations. Moreover, the web-based tool SNPator was used to compare populations individually one against one. Differences in genotypic distributions were mainly present between Andalusia and the other populations (χ^2 test, Table 4). According to that result, the populations from Catalonia and Andalusia showed the greatest differences, with 3 SNPs (rs25487, rs13181 and rs11615) differentially distributed among the PCA patients from both populations.

Principal component analysis (PCA) was done trying to identify global differences among populations. Components 1 and 2 were responsible for the 15.3% and 14.3% of the variance, respectively. When both components were plotted, the main components seemed not to discriminate between populations (Figure 3A). However, when components were analyzed separately, the first one could distinguish between the populations of Andalusia and Catalonia (Figure 3B), corroborating the results observed in Table 4 and clearly showing the differences in the distribution of genotypes between the analyzed populations.

Finally, haplotype analysis was performed in SNPator. As shown in Table 5, the three most frequent haplotypes were different among populations. Thus, for SNPs in chromosome 11 (those located in ATM gene), the haplotype GG was absent in the Andalusian population. For SNPs in chromosome 13 (those located in LIG4 gene), haplotypes GG and AA showed a different distribution among the populations. In the case of SNPs in chromosome 19 (those located in XRCC1, ERCC2 and ERCC1 genes), haplotype CCGGG was present only in PCA patients from Canary and Catalonia, while haplotype CCGTG was present only in PCA patients from Andalusia and Basque Country. The fact that the most frequent haplotypes were equal in all populations suggests a similarity between individuals of the same ethnicity.

Discussion

Radiogenomics is the study of genetic variants, primarily single nucleotide polymorphisms (SNPs), associated with the development of radiotherapy toxicity, in an attempt to find an assay capable of predicting which cancer patients are most likely to develop adverse effects after RT [9]. The prediction of normal tissue toxicity would allow the adjusting of radiation doses individually for each patient, especially when higher radiation

Table 5. Analysis of the three most frequent haplotypes (%) in chromosomes 11 (ATM gene), 13 (Lig4 gene) and 19 (XRCC1, ERCC2 and ERCC1 genes) among the studied populations.

Haplotype/chromosome	Andalusia	fr	Basque Country	fr	Canary	fr	Catalonia	fr
Hap 1	CT	93.18	CT	89.82	CT	89.15	CT	91.27
Hap 2	CG	6.80	CG	7.24	CG	8.23	CG	4.87
Hap 3	GT	0.02	GG	2.74	GG	2.39	GG	3.86
Haplotype/chromosome 13								
Hap 1	AG	76.08	AG	69.81	AG	70.17	AG	64.14
Hap 2	GG	11.89	AA	22.27	AA	15.22	GG	19.60
Hap 3	AA	10.53	GG	7.11	GG	14.49	AA	16.23
Haplotype/chromosome 19								
Hap 1	CCGTA	3252	CCGTA	31.19	CCGTA	2737	CCGTA	26.11
Hap 2	TCGTA	1486	CCGTG	14.81	TCGTA	1703	TCGTA	14.31
Hap 3	CCGTG	1137	TCGTA	14.29	CCGGG	1232	CCGGG	12.37

Abbreviations: chr, chromosome; fr, frequency; Hap, haplotype. Haplotypes in chr 11 is shaped with locus rs180057 and rs1750908, respectively; haplotypes in chr 13 is shaped with locus rs1805386 and rs1805388, respectively; haplotypes in chr 19 is shaped with locus rs25487, rs25488, rs1799782, rs13181, and rs11615, respectively. doi:10.1371/journal.pone.0069735.t005

dose levels are associated with improved biochemical control outcomes and reduction in distant metastases in PCa patients [24]. The role of genetics in radiation-toxicity has been proved [25]. In that sense, genetics seem to contribute to explain the high inter-individual variability observed between cases, even when patients are similar and are treated with the same treatment schedule [26]. However, although it has been published a lot of bibliography reporting the predictive role of some SNPs in normal tissue toxicity, the validation studies have failed, calling into question the utility of SNPs as a tool for predicting radiation-induced toxicity [12].

Population association between genotype at a particular locus and a binary trait (such as presence/absence of radiation-induced toxicity) can arise in three ways [27]: i) the locus may be causally related to the disease (different alleles carrying different risks), ii) the locus may not itself be causal (but may be sufficiently close to a causal locus as to be in linkage disequilibrium with it), or iii) the association may be due to confounding by population stratification or admixture. Confounding may act to create population association in the absence of a causal link or obscure a causal relationship. Thus, it is important to exclude spurious association by appropriate design and/or analysis of studies, taken into account that biases that result from systematic error (such as selection biases or biases in measuring outcomes) persist as the sample size increases. Confounding would arise if the population contained several ethnic groups, if allele frequencies at the locus of interest differed between groups, and if disease frequency also differed between groups for reasons quite unrelated to the locus of interest. It is known that ethnicity influences the applicability of pharmacogenetics [28].

Canary population, as well as the rest of populations included in this study, is considered as Caucasian. However, the natural history of, for example, Canary and Basque Country, are different. Thus, while Canary population has influence from Northwest Africa migration and European colonisation [29], Basques have a different origin [30]. However, in a recent published paper, 30 individuals from 10 different populations from Spain (Canary population was not included in that study) were genotyped for 120 SNPs, concluded that the studied populations were genotypically similar [3]. None of the SNPs considered in the present study were included in this previous article. We found that genotype distribution of 4 SNPs was different among populations from Andalusia, Basque Country, Canary and Catalonia. We compared our findings with the largest cohort of PCa patients analyzed in Spain [31]. A total of 698 Galician PCa patients were screened for 14 SNPs located in the ATM, ERCC2, LIG4, MLH1 and XRCC3 genes. Three of these SNPs were included in our multicenter study: rs1805388 (LIG4), rs1805386 (LIG4) and rs1800057 (ATM). Genotypic distributions of rs1805388 and rs1805386 were significantly different among Galician and the populations included in the present study (χ^2 test, $p=0.001$ and $p=0.007$, respectively), highlighting the variability between populations of the same ethnicity (Caucasians) from the same country in depending of each SNP. According to our results, Andalusia was the population differentially distributed, showing the greatest disparity with Catalan (results observed in χ^2 analyses and PCA). Differences among populations were also evident in haplotype analysis and subsequent distribution. Those results suggest that each SNP need to be considered individually, trying to find possible confounding variables that would be crucial for the interpretation of results. In case-control studies, which is the usual type of design in studies for discovering associations between SNPs and radiation toxicity, the fundamental assumption is that these two series of subjects (controls and cases) may be used to provide

unbiased estimates of the corresponding distributions among affected and unaffected members of some underlying population [27]. This fundamental assumption may not be met in practice, leading to biased findings that fall into two broad classes: selection bias caused by inappropriate sampling of cases and controls, and information bias caused by differential measurement errors in cases and controls. When the confounding variable is detected in the study, the classical method in epidemiology is by stratification of the analysis by the potentially confounding variable and testing for association between factors of interest (i.e. genotype) and disease within strata (i.e. grades of radiation-induced toxicity). Concern over the presence of bias from population stratification in genetic case-control studies should be alleviated by proper design and analysis of case-control studies, evaluation of the likelihood of major bias in a given study [32] and, if needed, methods for correction [33].

The present study has some limitations that should be noted. First, all subjects were PCa patients and the genotype frequency may be different if it is compared with a population of healthy subjects. However, in studies designed to evaluate possible associations between SNPs and radiation toxicity, controls are patients with null-low grade of toxicity and cases are patients with high grade of toxicity, but all subjects are cancer patients. Thus, this limitation could be considered as an advantage because it mimics the standard design of such studies. Second, the number of subjects from the different population varies widely. However, the fact that the main differences were not found in the population with the smallest number of patients (Basque Country, with 51 PCa) suggests that this limitation may not be decisive in the interpretation of results. Moreover, if heterogeneity among populations is considered a systematic bias, it is independent of sample size. Third, to blind the analysis, no clinical data on patients were available, that is, there are not data about TNM staging, tumor grade, biochemical failure, or Gleason Score. In that sense, it is possible that some polymorphism may influence tumor characteristics in the same way that it may pose a risk factor for other disease characteristics [34,35]. In the other hand, some advantages should be highlighted: i) it includes a number of subjects sufficient to have reliable data on the distribution of these 10 SNPs in the PCa populations studied (especially for Canary and Catalonia); ii) all subjects were male, then avoiding the possible bias generated by the gender; and iii) all the determinations (6,010 in total) were performed with the same methodology (OpenArray, Applied Biosystems), with the same batch of chips and by the same investigator, thus minimizing biases from technical origin.

Conclusions

Differences in distribution of genotypes within different populations of the same ethnicity could be an important confounding factor responsible for the lack of validation of those SNPs associated with radiation-induced toxicity, especially when extensive meta-analysis with subjects from different countries are carried out [36]. Our results suggest that equality between people (especially among those considered as control) should be checked before proceeding with any further analysis.

Acknowledgments

We thank the technical support from the Immunology Department (Hospital Universitario de Gran Canaria Dr. Negrete) staff: Nereida González-Quevedo and Yanira Florido-Ortega. Special thank to Eduardo Salazar Villaverde for the assistance in the figures preparation.

Author Contributions

Conceived and designed the experiments: LAHH PCL. Performed the experiments: LAHH AV EHR. Analyzed the data: LAHH AV.

References

- Guttmacher AF, Collins FS (2002) Genomic medicine—a primer. *N Engl J Med* 347: 1512–1520.
- Baecher M, McEvoy B, Pearson LN, Quillen EE, Sarkisian T, et al. (2007) Measuring European population stratification with microarray genotype data. *Am J Hum Genet* 80: 948–956.
- Layouni H, Calafell F, Betranpetit J (2010) A genome-wide survey does not show the genetic distinctiveness of Balearic. *Hum Genet* 127: 455–458.
- Pino-Yanes M, Comrak A, Baskias S, Hernandez A, Garcia L, et al. (2011) North African influences and potential bias in case-control association studies in the Spanish population. *PLoS One* 6: e18389.
- Becarda S, de Cobelli O, Gucco C, Poyes-Galeti T, Valdagni R, et al. (2005) Cancer of the prostate. *Crit Rev Oncol Hematol* 56: 379–396.
- Kapelian PA, Potton L, Khumria D, Cieki JP, Reddy CA, et al. (2004) Radical prostatectomy, external beam radiotherapy <72 Gy, external beam radiotherapy >= 72 Gy, permanent seed implantation, or combination of seed/external beam radiotherapy for stage T1-T2 prostate cancer. *Int J Radiat Oncol Biol Phys* 58: 25–33.
- Johanson S, Svensson H, Denksamp J (2000) Timescale of evolution of late radiation injury after postoperative radiotherapy of breast cancer patients. *Int J Radiat Oncol Biol Phys* 48: 745–750.
- Bentzen SM, Overgaard J (1994) Patient-to-Patient Variability in the Expression of Radiation-Induced Normal Tissue Injury. *Semin Radiat Oncol* 4: 68–80.
- Hernandez-Hernandez LA, Bordon E, Pinar B, Lloret M, Rodriguez-Gallego C, et al. (2012) Prediction of normal tissue toxicity as part of the individualized treatment with radiotherapy in oncology patients. *Surg Oncol* 21: 201–206.
- West CM, Bamert GC (2011) Genetics and genomics of radiotherapy toxicity: towards prediction. *Genome Med* 3: 52.
- West C, Rosenstein BS, Almer J, Acri D, Burnett G, et al. (2010) Establishment of a Radiogenomics Consortium. *Int J Radiat Oncol Biol Phys* 76: 1295–1296.
- Bamert GC, Coles CE, Elliott RM, Baynes C, Luccarini C, et al. (2012) Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study. *Lancet Oncol* 13: 65–77.
- West CM, Dunning AM, Rosenstein BS (2012) Genome-wide association studies and prediction of normal tissue toxicity. *Semin Radiat Oncol* 22: 91–99.
- Langenlehner T, Renner W, Genger A, Hofmann G, Thurner EM, et al. (2011) Association between single nucleotide polymorphisms in the gene for XRCC1 and radiation-induced late toxicity in prostate cancer patients. *Radiat Oncol* 98: 387–393.
- Mangoni M, Bitanti S, Carozzi F, Sani G, Bini G, et al. (2011) Association between genetic polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT genes and radiosensitivity in breast cancer patients. *Int J Radiat Oncol Biol Phys* 81: 52–58.
- Zhang L, Yang M, Bin J, Wu C, et al. (2010) Association of TGF-beta1 and XPD polymorphisms with severe acute radiation-induced esophageal toxicity in locally advanced lung cancer patients treated with radiotherapy. *Radiat Oncol* 97: 19–25.
- Biggart M, Symkowski DE, Wood RD (1993) Co-connection of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro. *EMBO J* 12: 3685–3692.
- Damraju S, Murray D, Daifur J, Ganesang D, Myrberg S, et al. (2006) Association of DNA repair and steroid metabolism gene polymorphisms with clinical late toxicity in patients treated with conformal radiotherapy for prostate cancer. *Clin Cancer Res* 12: 2545–2554.
- Tantlele GA, Murray RJ, Mills J, Rawell J, Ghakshori P, et al. (2012) Variation in telangiectasia predisposing genes is associated with overall radiation toxicity. *Int J Radiat Oncol Biol Phys* 84: 1031–1036.
- Chang-Claude J, Antoniou GB, Lila C, Krupp S, Helmhold I, et al. (2009) Genetic polymorphisms in DNA repair and damage response genes and late normal tissue complications of radiotherapy for breast cancer. *Br J Cancer* 100: 1680–1686.
- Morillo-Suarez C, Alegre J, Sangro R, Gazave E, de Gid R, et al. (2008) SNP analysis to results (SNPator): a web-based environment oriented to statistical genomics analyses upon SNP data. *Bioinformatics* 24: 1643–1644.
- Clayton D (2012) mpSNP, SepMatrix and XSepMatrix classes and methods. R package version 1.8.1.
- Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95: 14683–14688.
- Zelefsky MJ, Pei X, Chou JF, Schocher M, Kolmeier M, et al. (2011) Dose escalation for prostate cancer radiotherapy: predictors of long-term biochemical tumor control and distant metastases-free survival outcomes. *Eur Urol* 60: 1133–1139.
- Andreassen CN, Almer J (2009) Genetic variants and normal tissue toxicity after radiotherapy: a systematic review. *Radiat Oncol* 92: 290–309.
- Buchholz TA (1999) Finding our sensitive patients. *Int J Radiat Oncol Biol Phys* 45: 547–548.
- Clayton D (2001) Population Association. *Handbook of statistical genetics*: 519–540.
- Sai K, Saini Y (2011) Ethnic differences in the metabolism, toxicology and efficacy of three anticancer drugs. *Expert Opin Drug Metab Toxicol* 7: 967–988.
- Rando JC, Cabrera VM, Larruga JM, Hernandez M, Gonzalez AM, et al. (1999) Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann Hum Genet* 63: 413–428.
- Calafell F, Betranpetit J (1994) Principal component analysis of gene frequencies and the origin of Balearics. *Am J Phys Anthropol* 93: 201–215.
- Fachal I, Gomez-Guzman A, Peleteiro P, Carballo A, Calvo-Grejo P, et al. (2012) Association of a XRCC3 polymorphism and mean mean dose with the risk of acute radio-induced gastrointestinal toxicity in prostate cancer patients. *Radiat Oncol* 105: 321–328.
- Wacholder S, Rothman N, Caporaso N (2000) Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J Natl Cancer Inst* 92: 1151–1158.
- Reich DE, Goldstein DB (2001) Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol* 20: 4–16.
- Hernandez-Hernandez LA, Munias-Rosales A, Gonzalez-Hernandez A, de Leon AC, Diaz-Chico N, et al. (2010) Distribution of TYMS, MTHFR, p53 and MDR1 gene polymorphisms in patients with breast cancer treated with neoadjuvant chemotherapy. *Cancer Epidemiol* 34: 634–638.
- Hernandez-Hernandez LA, Perez LF, Hernandez AG, de Leon AC, Diaz-Chico N, et al. (2010) TYMS, MTHFR, p53 and MDR1 gene polymorphisms in breast cancer patients treated with adjuvant therapy. *Cancer Epidemiol* 34: 490–493.
- Bamert GC, Elliott RM, Almer J, Andreassen CN, Abdelhay O, et al. (2012) Individual patient data meta-analysis shows no association between the SNP rs1800469 in TGF-beta and late radiotherapy toxicity. *Radiat Oncol* 105: 289–295.

RESEARCH NOTE

Intraethnic variation in steroid-5-alpha-reductase polymorphisms in prostate cancer patients: a potential factor implicated in 5-alpha-reductase inhibitor treatment

LUIS ALBERTO HENRÍQUEZ-HERNÁNDEZ^{1,2,3*}, ALMUDENA VALENCIANO², PALMIRA FORO-ARNALOT⁴, MARÍA JESÚS ÁLVAREZ-CUBERO^{5,6}, JOSÉ MANUEL COZAR⁷, JOSÉ FRANCISCO SUÁREZ-NOVO⁸, MANEL CASTELLS-ESTEVE⁹, PABLO FERNÁNDEZ-GONZALO⁹, BELÉN DE-PAULA-CARRANZA⁹, MONTSE FERRER¹⁰, FERRÁN GUEDEA¹¹, GEMMA SANCHO-PARDO¹², JORDI CRAVEN-BARTLE¹², MARÍA JOSÉ ORTIZ-GORDILLO¹³, PATRICIA CABRERA-ROLDÁN¹³, ESTEFANÍA HERRERA-RAMOS^{14,15}, CARLOS RODRÍGUEZ-GALLEGO^{14,15} and PEDRO C. LARA^{1,2,3}

¹Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas 35010, Spain

²Instituto Canario de Investigación del Cáncer, Las Palmas 38204, Spain

³Clinical Sciences Department, Universidad de Las Palmas de Gran Canaria, Las Palmas 35016, Spain

⁴Institut d'Oncologia Radioteràpica, Hospital de la Esperanza, Parc de Salut Mar, Barcelona 08003, Spain

⁵Laboratory of Genetic Identification, Legal Medicine and Toxicology Department, Facultad de Medicina, Universidad de Granada, Granada 18012, Spain

⁶GENYO, Pfizer-University of Granada-Andalusian Government Centre for Genomics and Oncological Research, Granada 18016, Spain

⁷Department of Urology, Hospital Universitario Virgen de las Nieves, Granada 18014, Spain

⁸Department of Urology, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat 08907, Spain

⁹Radiation Oncology Department, Onkologikoa, Guipuzcoa 20014, Spain

¹⁰Health Services Research Group, Institut de Recerca Hospital del Mar (IMIM), Barcelona 08003, Spain

¹¹Department of Radiation Oncology, Institut Català d'Oncologia (ICO), L'Hospitalet de Llobregat 08907, Spain

¹²Radiation Oncology Department, Hospital de la Santa Creu i Sant Pau, Barcelona 08026, Spain

¹³Radiation Oncology Department, Hospital Universitario Virgen del Rocío, Sevilla 41013, Spain

¹⁴Department of Immunology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas 35010, Spain

¹⁵Department of Medical and Surgical Sciences, Universidad de Las Palmas de Gran Canaria, Las Palmas 35016, Spain

[Henríquez-Hernández L. A., Valenciano A., Foro-Arnalot P., Álvarez-Cubero M. J., Cozar J. M., Suárez-Novo J. F., Castells-Estève M., Fernández-Gonzalo P., De-Paula-Carranza B., Ferrer M., Guedea F., Sancho-Pardo G., Craven-Bartle J., Ortiz-Gordillo M. J., Cabrera-Roldán P., Herrera-Ramos E., Rodríguez-Gallego C. and Lara P. C. 2015 Intraethnic variation in steroid-5-alpha-reductase polymorphisms in prostate cancer patients: a potential factor implicated in 5-alpha-reductase inhibitor treatment. *J. Genet.* **94**, 335–341]

Introduction

Steroid-5-alpha-reductase inhibitors are commonly used in the treatment of benign prostatic enlargement and male alopecia with controversial results. Ethnic diversity is an important factor, accounting for interindividual variation in drug responsiveness. The aim of this study was to evaluate the genotypic distribution of 22 SNPs in steroid-5-alpha-reductase alpha polypeptides 1 and 2 in a set of 601 prostate cancer patients from four different Spanish regions. The genotyping was done in a Biotrove OpenArray NT Cycler.

Twelve SNPs of 22 analysed were differentially distributed among the different regions. Since all subjects were Caucasian of Spanish origin, this result showed differences in genotype distribution among subject from the same ethnic origin. These observations were confirmed in cluster analysis, principal component analysis and in the differential distribution of haplotypes among the populations. Differences in distribution of genotypes within different populations of the same ethnicity could be an important factor responsible for the wide variety of pharmacological responses to steroid-5-alpha-reductaseinhibitors.

Prostate cancer (PCa) is a hormone-dependent tumour which needs androgens for disease initiation and progression

*For correspondence. E-mail: lhenriquez@dcc.ulpgc.es.

Keywords. 5-alpha-reductase polymorphism; 5-alpha-reductase inhibitors; OpenArray; prostate cancer; pharmacoethnicity.

(Huggins 1967). Although, testosterone (T) is the most abundant serum androgen, dihydrotestosterone (DHT) is the main prostate androgen. T is converted to DHT by the enzyme steroid-5- α -reductase (SRD5A) in the prostate, testes, hair follicles and adrenal glands (McConnell 1995). Since the discovery of the implications of SRD5A2 deficiency, it has been an increasing interest in the development of steroid-5- α -reductase inhibitors (5 α -RI). Among several compounds, only two drugs have been FDA approved for clinical use: finasteride, which inhibits SRD5A2 and was approved for the treatment of benign prostatic enlargement (BPE) and male alopecia; and dutasteride, which inhibits SRD5A1 and SRD5A2 (approved for the treatment of BPE) (Edwards and Moore 2002; Andriole and Kirby 2003). Nowadays, dutasteride and finasteride do not prevent CaP, but merely temporarily shrink tumours that have a low potential for being lethal (Walsh 2010).

The enzymatic activity of SRD5A is highly modulated by single-nucleotide polymorphisms (SNPs) located in these genes (Giwercman *et al.* 2005; Cussenot *et al.* 2007). These genetic variations in key genes belonging to the androgen pathway seem to be important for the prostate response to androgens (Lindstrom *et al.* 2006). We have recently reported that differences in the distribution of genotypes within different populations of the same ethnicity could be an important confounding factor in association studies (Henríquez-Hernández *et al.* 2013a).

We designed a study aimed to evaluate the genotypic distribution of 22 SNPs in SRD5A1 and SRD5A2 in a set of Spanish prostate cancer patients, to determine the homogeneity of the population and to disclose potential bias associated with the failure of 5 α -RI when prescribed.

Materials and methods

A total of 601 patients with nonmetastatic localized prostate cancer were included in this study. Geographical distribution of patients was as follows (Henríquez-Hernández *et al.* 2013a): 91 (15.14%) from Andalusia, 51 (8.48%) from Basque Country, 238 (39.60%) from Canary and 221 (36.77%) from Catalonia. All patients were of Spanish origin and all of them received written informed consent before sample collection. Ethnicity of patients was established until the second parental generation. This study was approved by the Research and Ethics Committee of each institution participant in the study.

DNA was isolated from 300 μ L of whole-blood in an iPrep™ purification instrument using the iPrep™ PureLink™ gDNA Blood kit (Invitrogen, Life Technologies, Carlsbad, USA). DNA integrity was determined by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, USA).

SNPs were selected using data of individuals with European ancestry (CEU) from the HapMap Project (available at: <http://www.hapmap.org>). Pairwise linkage disequilibrium (LD) tagging was achieved with Haploview ver. 4.2 software (free downloaded from <http://www.broadinstitute.org>

scientific-community/science/programs/medical-and-population-genetics/haploview/haploview) (Barrett *et al.* 2005). The R^2 given by Haploview was >0.95 for all genes. Description of SNPs is provided in table 1 in electronic supplementary material at <http://www.ias.ac.in/jgenet/>.

The SNP genotyping was performed in a Biotrove OpenArray® NT Cyclor (Applied Biosystems, Foster City, USA) following the instructions of manufacturer (Henríquez-Hernández *et al.* 2013b). The fluorescence results were read using the OpenArray® SNP genotyping analysis software ver. 1.0.5 (Applied Biosystems). The genotyping analysis was performed using TaqMan Genotyper software ver. 1.0.1 (available at: <http://www.invitrogen.com/site/us/en/home/Global/forms/taqman-genotyper-software-download-reg.html>) using autocalling as the call method. The quality value of the data points was determined by a threshold above 0.95. Genotyping analysis was done for each population separately and it was done automatically without any manual assignment of genotypes as previously reported (Henríquez-Hernández *et al.* 2013a). All the genotyped samples met the quality criteria and all samples were genotyped with the same batch of material and at the same time.

Genotype and allelic frequencies were determined using the web-based environment SNPator (SNP Analysis To Results, from the Spain's National Genotyping Center and the National Institute for Bioinformatics) (Morcillo-Suarez *et al.* 2008).

Principal component analysis (PCA) was done using SnpMatrix and XSnMatrix classes and methods (Clayton 2012), implemented as an R package and available from Bioconductor (as of ver. 2.11; <http://bioconductor.org>). Non-supervised hierarchical clustering of SNP in each population was done using MultiExperiment Viewer ver. 4.9 (available at: <http://www.tigr.org>).

Results

A total of 601 PCa patients were genotyped for 22 SNPs; 10 SNPs located in *SRD5A1* gene and 12 SNPs located in *SRD5A2* gene. Of the 13,222 possible determinations, 94.31% were successfully genotyped.

The genotypic and allelic frequencies are shown in table 1. The genotype distribution was different among the study populations in 12 of the 22 SNPs: rs166050, rs501999, rs518673, rs3822430, rs8192120, rs39848 (*SRD5A1*); and rs2208532, rs12470143, rs2281546, rs3754838, rs523349, rs9332975 (*SRD5A2*) (χ^2 test, table 1), showing a differential distribution of genotypes among populations.

A nonsupervised hierarchical cluster was performed to visualize the genotype distributions among the four populations. For *SRD5A1*, polymorphisms were distributed into three main clusters, each one with different numbers and identities of SNPs, suggesting heterogeneity among populations (figure 1 in electronic supplementary material, left panel). For *SRD5A2*, SNPs were distributed into two main clusters, very similar for Basque, Canary and Catalan populations,

Pharmacogenetics of 5- α -reductase inhibitors

Table 1. Genotype and allelic frequencies of *SRD5A1* and *SRD5A2* gene polymorphisms among the populations.

	Call rate	Genotype			Allele		Functional consequence	MAF
<i>SRD5A1</i>								
rs166050		AA	AG	GG	A	G	Intron variation	0.12
Andalusia	0.75	0.65	0.21	0.14	0.75	0.25		
Basque Country	0.98	0.52	0.42	0.06	0.73	0.27		
Canary	0.99	0.62	0.33	0.05	0.79	0.21		
Catalonia	0.94	0.57	0.37	0.06	0.76	0.24		
<i>P</i> value				0.025				
rs501999		CC	CT	TT	C	T	Intron variation	0.49
Andalusia	0.70	0.48	0.20	0.32	0.59	0.41		
Basque Country	1.00	0.29	0.49	0.22	0.54	0.46		
Canary	0.99	0.29	0.49	0.22	0.53	0.47		
Catalonia	0.98	0.30	0.43	0.27	0.51	0.49		
<i>P</i> value				0.003				
rs518673		AA	AG	GG	A	G	Intron variation	0.32
Andalusia	0.75	0.19	0.19	0.62	0.29	0.71		
Basque Country	1.00	0.16	0.37	0.47	0.34	0.66		
Canary	0.96	0.07	0.42	0.51	0.28	0.72		
Catalonia	0.96	0.10	0.40	0.50	0.30	0.70		
<i>P</i> value				0.006				
rs3822430		AA	AG	GG	A	G	Synonymous codon	0.30
Andalusia	0.77	0.37	0.24	0.39	0.49	0.51		
Basque Country	1.00	0.27	0.61	0.12	0.58	0.42		
Canary	0.97	0.33	0.51	0.16	0.58	0.42		
Catalonia	0.97	0.37	0.47	0.16	0.60	0.40		
<i>P</i> value				<0.001				
rs500182		GG	GT	TT	G	T	Intron variation	0.20
Andalusia	0.81	0.01	0.14	0.85	0.08	0.92		
Basque Country	1.00	0.02	0.18	0.80	0.11	0.89		
Canary	0.98	0.01	0.17	0.82	0.09	0.91		
Catalonia	0.99	0.01	0.18	0.81	0.10	0.90		
<i>P</i> value				0.961				
rs8192120		AA	AC	CC	A	C	Intron variation	0.43
Andalusia	0.80	0.14	0.22	0.64	0.25	0.75		
Basque Country	0.99	0.08	0.42	0.50	0.29	0.71		
Canary	0.97	0.10	0.42	0.48	0.31	0.69		
Catalonia	0.98	0.10	0.48	0.42	0.34	0.66		
<i>P</i> value				0.015				
rs4702378		CC	CT	TT	C	T	Intron variation	0.24
Andalusia	0.82	0.07	0.23	0.70	0.18	0.82		
Basque Country	1.00	0.06	0.20	0.74	0.16	0.84		
Canary	0.99	0.04	0.33	0.63	0.21	0.79		
Catalonia	0.99	0.06	0.33	0.61	0.23	0.77		
<i>P</i> value				0.304				
rs1691053		AA	AG	GG	A	G	NA	0.15
Andalusia	0.87	0.90	0.09	0.01	0.95	0.05		
Basque Country	1.00	0.90	0.08	0.02	0.94	0.06		
Canary	0.98	0.86	0.13	0.01	0.93	0.07		
Catalonia	0.99	0.85	0.15	0.00	0.93	0.07		
<i>P</i> value				0.341				
rs39848		CC	CT	TT	C	T	Downstream variant	0.43
Andalusia	0.70	0.32	0.22	0.46	0.43	0.57		
Basque Country	0.96	0.22	0.43	0.35	0.44	0.56		
Canary	0.97	0.11	0.50	0.39	0.36	0.64		
Catalonia	0.94	0.16	0.45	0.39	0.39	0.61		
<i>P</i> value				<0.001				
rs3797179		AA	AG	GG	A	G	Intron variation	0.12
Andalusia	0.70	0.00	0.17	0.83	0.09	0.91		
Basque Country	0.96	0.00	0.29	0.71	0.14	0.86		
Canary	0.99	0.05	0.25	0.70	0.17	0.83		
Catalonia	0.96	0.03	0.26	0.71	0.17	0.83		
<i>P</i> value				0.213				

Table 1 (contd)

	Call rate	Genotype			Allele		Functional consequence	MAF
<i>SRD5A2</i>								
rs2208532		AA	AG	GG	A	G	Intron variation	0.44
Andalusia	0.70	0.45	0.28	0.27	0.59	0.41		
Basque Country	0.99	0.20	0.42	0.38	0.41	0.59		
Canary	0.99	0.34	0.46	0.20	0.57	0.43		
Catalonia	0.92	0.28	0.54	0.18	0.55	0.45		
<i>P</i> value				0.001				
rs12470143		CC	CT	TT	C	T	Intron variation	0.37
Andalusia	0.81	0.45	0.24	0.31	0.57	0.43		
Basque Country	1.00	0.26	0.53	0.21	0.52	0.48		
Canary	0.99	0.32	0.51	0.17	0.57	0.43		
Catalonia	0.98	0.29	0.54	0.17	0.56	0.44		
<i>P</i> value				0.001				
rs2281546		GG	GT	TT	G	T	Intron variation	0.17
Andalusia	0.83	0.09	0.07	0.85	0.12	0.88		
Basque Country	1.00	0.00	0.27	0.73	0.14	0.86		
Canary	0.99	0.03	0.22	0.75	0.14	0.86		
Catalonia	0.99	0.02	0.25	0.73	0.14	0.86		
<i>P</i> value				0.001				
rs3754838		CC	CT	TT	C	T	NA	0.12
Andalusia	0.78	0.10	0.11	0.79	0.16	0.84		
Basque Country	1.00	0.00	0.23	0.77	0.12	0.88		
Canary	0.96	0.03	0.20	0.77	0.12	0.88		
Catalonia	0.98	0.02	0.20	0.78	0.12	0.88		
<i>P</i> value				0.010				
rs4952222		AA	AC	CC	A	C	Intron variation	0.04
Andalusia	0.85	0.00	0.01	0.99	0.01	0.99		
Basque Country	1.00	0.00	0.00	1.00	0.00	1.00		
Canary	0.99	0.00	0.00	1.00	0.00	1.00		
Catalonia	0.97	0.00	0.00	1.00	0.00	1.00		
<i>P</i> value				0.086				
rs7562326		CC	CT	TT	C	T	Intron variation	0.14
Andalusia	0.80	0.05	0.12	0.83	0.12	0.88		
Basque Country	1.00	0.00	0.23	0.77	0.12	0.88		
Canary	0.98	0.03	0.20	0.77	0.13	0.87		
Catalonia	0.97	0.02	0.21	0.77	0.12	0.88		
<i>P</i> value				0.321				
rs2300702		CC	CG	GG	C	G	Intron variation	0.45
Andalusia	0.81	0.28	0.34	0.38	0.45	0.55		
Basque Country	0.96	0.12	0.49	0.39	0.37	0.63		
Canary	0.99	0.21	0.47	0.32	0.44	0.56		
Catalonia	0.94	0.16	0.50	0.34	0.41	0.59		
<i>P</i> value				0.127				
rs4952197		AA	AG	GG	A	G	Intron variation	0.25
Andalusia	0.80	0.12	0.19	0.69	0.22	0.78		
Basque Country	0.99	0.04	0.28	0.68	0.18	0.82		
Canary	0.99	0.06	0.31	0.63	0.22	0.78		
Catalonia	0.97	0.06	0.36	0.58	0.24	0.76		
<i>P</i> value				0.117				
rs676033		CC	CT	TT	C	T	NA	0.35
Andalusia	0.81	0.57	0.31	0.12	0.72	0.28		
Basque Country	0.99	0.60	0.32	0.08	0.76	0.24		
Canary	0.99	0.50	0.39	0.11	0.70	0.30		
Catalonia	0.98	0.50	0.43	0.07	0.71	0.29		
<i>P</i> value				0.402				
rs523349		CC	CG	GG	C	G	Missense	0.35
Andalusia	0.78	0.55	0.27	0.18	0.68	0.32		
Basque Country	0.92	0.64	0.32	0.04	0.80	0.20		
Canary	0.71	0.18	0.52	0.30	0.45	0.55		
Catalonia	0.89	0.52	0.40	0.08	0.72	0.28		
<i>P</i> value				<0.001				

Table 1 (contd)

	Call rate	Genotype			Allele		Functional consequence	MAF
		CC	CT	TT	C	T		
rs9332975		CC	CT	TT	C	T	utr variant	0.14
Andalusia	0.77	0.10	0.07	0.83	0.14	0.86		
Basque Country	1.00	0.00	0.26	0.74	0.13	0.87		
Canary	0.99	0.03	0.19	0.78	0.13	0.87		
Catalonia	0.98	0.02	0.20	0.78	0.12	0.88		
<i>P</i> value				0.002				
rs7594951		CC	CT	TT	C	T	Intron variation	0.11
Andalusia	0.75	0.84	0.10	0.06	0.89	0.11		
Basque Country	1.00	0.76	0.24	0.00	0.88	0.12		
Canary	0.99	0.77	0.20	0.03	0.88	0.12		
Catalonia	0.98	0.78	0.20	0.02	0.88	0.12		
<i>P</i> value				0.193				

MAF, minor allele frequency; NA, not available. Functional consequence and MAF are available at: <http://www.ncbi.nlm.nih.gov/projects/SNP/>. Differences in the genotype distribution were assessed by χ^2 test.

but clearly different from Andalusian population (figure 1 in electronic supplementary material, right panel). PCA was performed to identify global differences among populations. Components 1, 2 and 3 were responsible for the 43.7%, 22.6% and 15.3% of the variance for *SRD5A1*, respectively (cumulative percentage: 81.6%); and 48.8%, 33.2% and 9.1% of the variance for *SRD5A2* gene, respectively (cumulative percentage: 91.1%). The first component distinguished between the populations for both genes clearly showed the differences in the distribution of genotypes between the analysed populations (figure 1).

Haplotype analysis was performed in SNPator. For *SRD5A1* (chromosome 5), the haplotype GCCTATGGCA was only present among the Andalusian subjects. For *SRD5A2* (chromosome 2), there is a greater homogeneity in the distribution of haplotypes. The fact that the most frequent haplotypes were equal in all populations suggests a similarity between individuals of the same ethnicity.

Discussion

We have recently reported differences in the distribution of genotypes within different populations of the same ethnicity (Henriquez-Hernandez *et al.* 2013a), highlighting the importance of ethnic-specific differences in drug responsiveness (Yasuda *et al.* 2008; Ma and Lu 2011). In the present paper, we observed that genotype distribution of 12 of 22 SNPs were statistically different among the studied populations. The Spanish population is considered as Caucasian. However, the natural history of each subpopulation is different, and there is a high complexity in the Mediterranean migration processes that impact in the different population sources on the genetic composition of the Spanish population (Ambrosio *et al.* 2010). While Andalusian population seems to be originated by migrations from Arabian Peninsula, Fertile Crescent, Balkan region and North Africa (Ambrosio *et al.* 2010), Canary population has been influenced from Northwest Africa migration and European colonization (Rando *et al.* 1999). Differences

among populations were also evident in haplotype analysis, suggesting that each gene need to be considered individually to find possible confounding variables that would be crucial for the interpretation of results. Neither ethnicity nor genetic endowment of patients is taken into account in clinical trials.

5 α -RI arisen as a promise treatment for PCa prevention, as specific blocker of the enzyme responsible for the synthesis of DHT. We suggest here that SNPs located in *SRD5A1* and *SRD5A2* may be relevant factors to be considered when assessing the response to these drugs. Although, both genes encode for key enzymes in the conversion of testosterone into DHT and represent attractive targets for preventing prostate cancer development, they are expressed differentially in the prostate tissue. Thus, under normal physiologic conditions, SRD5A2 is preferentially expressed over SRD5A1 in the prostate (Thomas *et al.* 2005; Titus *et al.* 2005). In prostate cancer cells, however, the balance in expression of these genes shift toward predominant expression of SRD5A1 (Titus *et al.* 2005; Stanbrough *et al.* 2006), supporting a role for both enzymes in DHT bioavailability and carcinogenesis. However, the importance of androgens in early cancer initiation is emphasized by the fact that finasteride, a 5-AR type 2 inhibitor, and dutasteride, a dual 5-AR inhibitor targeting both 5-AR type 1 and type 2 enzymes, have been shown that the risk of prostate cancer incidence reduce by almost 23% (Thompson *et al.* 2003; Andriole *et al.* 2010). Since SRD5A1 and SRD5A2 are differentially expressed and are blocked by different drugs, the role of SNPs in each gene should be studied intensely taking these biological peculiarities in to account. However, the fact that rs523349—a SNP clearly involved in the success of chemoprevention of CaP using 5- α -reductase inhibitors (Cussenot *et al.* 2007)—appeared differentially distributed in our study, reinforce our findings.

The present study has some limitations that should be noted. First, to blind the analysis, no clinical data of patients were available, i.e., there are no data about TNM staging, tumour grade, biochemical failure, or Gleason Score. Second, the study was focussed on SRD5A1 and SRD5A2,

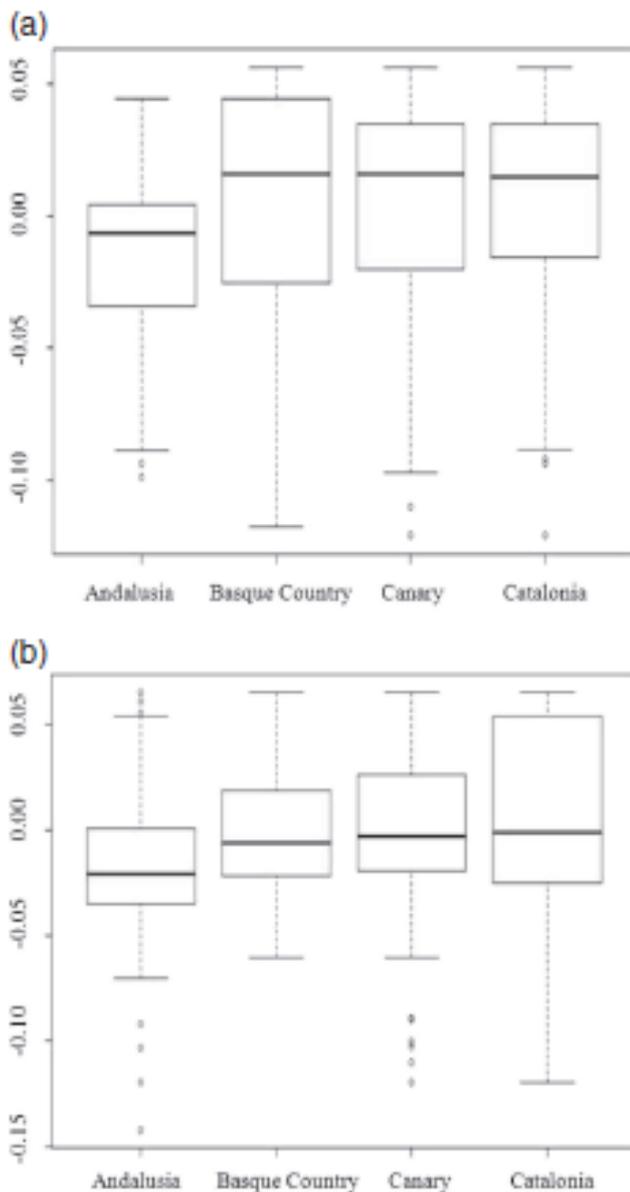


Figure 1. Box plot of component 1 among the different populations after PCA for (a) SRD5A1 and (b) SRD5A2.

ignoring possible associations with other genes and polymorphisms. Thus, the prostate cancer risk and aggressiveness conferred by certain SNPs located in 3 beta-hydroxysteroid dehydrogenase type II (HSD3B2) seem to be modified by SNPs located in SRD5A2 (Neslund-Dudas *et al.* 2007). In the same way, an association between SNPs in CYP17A1 (a gene encoding a key enzyme in the synthesis of androgens) and SRD5A2 has been suggested (Onen *et al.* 2007). Third, the number of subjects from different population varies widely. However, the fact that the main differences were not found in the population with the smallest number of patients (Basque Country, with 51 PCa) suggests that this limitation may not be decisive in the interpretation of results.

On the other hand, some advantages should be highlighted: (i) it includes a number of subjects sufficient to have

reliable data on the distribution of these 22 SNPs in the PCA populations studied (especially for Canary and Catalonia); (ii) all subjects were male (which is the most interesting population since the 5 α -RI are contraindicated in women), then avoiding the possible bias generated by the gender; and (iii) all the determinations (13,222 in total) were performed with the same methodology, with the same batch of chips and by the same investigator, thus minimizing biases from technical origin. All the analyses were automated and there was no manual assignment of genotypes in any case.

Conclusions

Differences in distribution of genotypes in SRD5A1 and SRD5A2 within different populations of the same ethnicity could be an important factor responsible for the failure of treatments based on 5 α -RI. Our results suggest that genotypic endowment should be taken into account (especially in clinical trials including people from different ethnics) in assessing the effectiveness of these drugs.

Acknowledgements

This work was subsidized by the grant from the Instituto de Salud Carlos III (Ministerio de Economía y Competitividad from Spain), ID: PI12/01867. Almudena Valenciano has a grant from the Instituto Canario de Investigación del Cáncer (ICIC). The founders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank the technical support from the Immunology Department (Hospital Universitario de Gran Canaria, Dr Negrín) staff: Nereida González-Quevedo and Yanira Florido-Ortega.

References

- Ambrosio B., Hernandez C., Novelletto A., Dugoujon J. M., Rodriguez J. N., Cuesta P. *et al.* 2010 Searching the peopling of the Iberian Peninsula from the perspective of two andalusian subpopulations: a study based on Y-chromosome haplogroups J and E. *Coll. Antropol.* **34**, 1215–1228.
- Andriole G. L. and Kirby R. 2003 Safety and tolerability of the dual 5 α -reductase inhibitor dutasteride in the treatment of benign prostatic hyperplasia. *Eur. Urol.* **44**, 82–88.
- Andriole G. L., Bostwick D. G., Brawley O. W., Gomella L. G., Marberger M., Montorsi F. *et al.* 2010 Effect of dutasteride on the risk of prostate cancer. *N. Engl. J. Med.* **362**, 1192–1202.
- Barrett J. C., Fry B., Maller J. and Daly M. J. 2005 Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265.
- Clayton D. 2012 snpStats: SnpMatrix and XSNpMatrix classes and methods. R package version 1.8.1. Available at: <http://www.bioconductor.org/packages/release/bioc/html/snpStats.html>.
- Cussenot O., Azzouzi A. R., Nicolaiew N., Mangin P., Cormier L., Fournier G. *et al.* 2007 Low-activity V89L variant in SRD5A2 is associated with aggressive prostate cancer risk: an explanation for the adverse effects observed in chemoprevention trials using 5 α -reductase inhibitors. *Eur. Urol.* **52**, 1082–1087.
- Edwards J. E. and Moore R. A. 2002 Finasteride in the treatment of clinical benign prostatic hyperplasia: a systematic review of randomised trials. *BMC Urol.* **2**, 14.

- Giwerzman Y. L., Abrahamsson P. A., Giwerzman A., Gadaleanu V. and Ahlgren G. 2005 The 5 α -reductase type II A49T and V89L high-activity allelic variants are more common in men with prostate cancer compared with the general population. *Eur. Urol.* **48**, 679–685.
- Henriquez-Hernandez L. A., Valenciano A., Foro-Arnalot P., Alvarez-Cubero M. J., Cozar J. M., Suarez-Novo J. F. et al. 2013a Polymorphisms in DNA-repair genes in a cohort of prostate cancer patients from different areas in Spain: heterogeneity between populations as a confounding factor in association studies. *PLoS One* **8**, e69735.
- Henriquez-Hernandez L. A., Valenciano A., Herrera-Ramos E., Lloret M., Riveros-Perez A. and Lara P. C. 2013b High-throughput genotyping system as a robust and useful tool in oncology: experience from a single institution. *Biologicals* **41**, 424–429.
- Huggins C. 1967 Endocrine-induced regression of cancers. *Cancer Res.* **27**, 1925–1930.
- Lindstrom S., Wiklund F., Adami H. O., Balter K. A., Adolfsson J. and Gronberg H. 2006 Germ-line genetic variation in the key androgen-regulating genes androgen receptor, cytochrome P450, and steroid-5- α -reductase type 2 is important for prostate cancer development. *Cancer Res.* **66**, 11077–11083.
- Ma Q. and Lu A. Y. 2011 Pharmacogenetics, pharmacogenomics, and individualized medicine. *Pharmacol. Rev.* **63**, 437–459.
- McConnell J. D. 1995 Prostatic growth: new insights into hormonal regulation. *Br. J. Urol.* **76** suppl 1, 5–10.
- Morcillo-Suarez C., Alegre J., Sangros R., Gazave E, de Cid R., Milne R. et al. 2008 SNP analysis to results (SNPator): a web-based environment oriented to statistical genomic analyses upon SNP data. *Bioinformatics* **24**, 1643–1644.
- Neslund-Dudas C., Bock C. H., Monaghan K., Nock N. L., Yang J. J., Rundle A. et al. 2007 SRD5A2 and HSD3B2 polymorphisms are associated with prostate cancer risk and aggressiveness. *Prostate* **67**, 1654–1663.
- Onen I. H., Ekmekci A., Eroglu M., Polat F. and Biri H. 2007 The association of 5 α -reductase II (SRD5A2) and 17 hydroxylase (CYP17) gene polymorphisms with prostate cancer patients in the Turkish population. *DNA Cell Biol.* **26**, 100–107.
- Rando J. C., Cabrera V. M., Larraga J. M., Hernandez M., Gonzalez A. M., Pinto F. et al. 1999 Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann. Hum. Genet.* **63**, 413–428.
- Stanbrough M., Bubley G. J., Ross K., Golub T. R., Rubin M. A., Penning T. M. et al. 2006 Increased expression of genes converting adrenal androgens to testosterone in androgen-independent prostate cancer. *Cancer Res.* **66**, 2815–2825.
- Thomas L. N., Lazier C. B., Gupta R., Norman R. W., Troyer D. A., O'Brien S. P. et al. 2005 Differential alterations in 5 α -reductase type 1 and type 2 levels during development and progression of prostate cancer. *Prostate* **63**, 231–239.
- Thompson I. M., Goodman P. J., Tangen C. M., Lucia M. S., Miller G. J., Ford L. G. et al. 2003 The influence of finasteride on the development of prostate cancer. *N. Engl. J. Med.* **349**, 215–224.
- Titus M. A., Gregory C. W., Ford 3rd O. H., Schell M. J., Maygarden S. J. and Mohler J. L. 2005 Steroid 5 α -reductase isozymes I and II in recurrent prostate cancer. *Clin. Cancer Res.* **11**, 4365–4371.
- Walsh P. C. 2010 Chemoprevention of prostate cancer. *N. Engl. J. Med.* **362**, 1237–1238.
- Yasuda S. U., Zhang L. and Huang S. M. 2008 The role of ethnicity in variability in response to drugs: focus on clinical pharmacology studies. *Clin. Pharmacol. Ther.* **84**, 417–423.

Received 23 June 2014, in revised form 8 January 2015; accepted 20 January 2015

Unedited version published online: 23 January 2015

Final version published online: 1 June 2015

RESEARCH ARTICLE

Open Access

Single nucleotide polymorphisms in DNA repair genes as risk factors associated to prostate cancer progression

Luis Alberto Henríquez-Hernández^{1,2,3*}, Almudena Valenciano², Palmira Foró-Arnalot⁴, María Jesús Álvarez-Cubero^{5,6}, José Manuel Cozar⁷, José Francisco Suárez-Novó⁸, Manel Castells-Estevé⁸, Pablo Fernández-Gonzalo⁹, Belén De-Paula-Carranza⁹, Montse Ferrer¹⁰, Ferrán Guedea¹¹, Gemma Sancho-Pardo¹², Jordi Craven-Bartle¹², María José Ortiz-Gordillo¹³, Patricia Cabrera-Roldán¹³, Estefanía Herrera-Ramos^{14,15}, Carlos Rodríguez-Gallego^{14,15}, Juan Ignacio Rodríguez-Melcón^{1,2} and Pedro C Lara^{1,2,3}

Abstract

Background: Besides serum levels of PSA, there is a lack of prostate cancer specific biomarkers. It is need to develop new biological markers associated with the tumor behavior which would be valuable to better individualize treatment. The aim of this study was to elucidate the relationship between single nucleotide polymorphisms (SNPs) in genes involved in DNA repair and prostate cancer progression.

Methods: A total of 494 prostate cancer patients from a Spanish multicenter study were genotyped for 10 SNPs in *XRCC1*, *ERCC2*, *ERCC1*, *LIG4*, *ATM* and *TP53* genes. The SNP genotyping was made in a Biotrove OpenArray® NT Cycler. Clinical tumor stage, diagnostic PSA serum levels, and Gleason score at diagnosis were obtained for all participants. Genotypic and allelic frequencies were determined using the web-based environment SNPator.

Results: SNPs rs11615 (*ERCC1*) and rs17503908 (*ATM*) appeared as risk factors for prostate cancer aggressiveness. Patients wild homozygous for these SNPs (AA and TT, respectively) were at higher risk for developing cT2b – cT4 (OR = 2.21 (confidence interval (CI) 95% 1.47 – 3.31), $p < 0.001$) and Gleason scores ≥ 7 (OR = 2.22 (CI 95% 1.38 – 3.57), $p < 0.001$), respectively. Moreover, those patients wild homozygous for both SNPs had the greatest risk of presenting D'Amico high-risk tumors (OR = 2.57 (CI 95% 1.28 – 5.16)).

Conclusions: Genetic variants at DNA repair genes are associated with prostate cancer progression, and would be taken into account when assessing the malignancy of prostate cancer.

Keywords: Single nucleotide polymorphism, *ERCC1*, *ATM*, Prostate cancer, OpenArray, DNA repair, Spanish cohort

Background

Prostate cancer (PCa) is a complex disease highly influenced by hormonal and genetic factors which would condition the tumor behavior. Tumor staging, tumor grading in terms of Gleason score and diagnostic prostate specific antigen (PSA) serum levels are clinically used to classify patients into different prognostic risk groups which will

condition treatment decisions. However, it is estimated that 293 men have to be screened and 12 men have to be treated to avoid one death related to PCa. Although recent advances in genomic research have made possible to identify new biomarkers for PCa, results are inconclusive [1] and it seems to be a need for new biomarkers of tumor behavior.

DNA is constantly damaged by endogenous oxygen free radicals and exogenous chemicals; thus, different repair pathways are available to reverse the different types of DNA damage [2]. Defects in these DNA repair pathways

* Correspondence: henriquez@dcc.upgccs

¹Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, C/Baranco de La Ballena s/n, CP 35010 Las Palmas de Gran Canaria, Las Palmas, Spain

²Instituto Canario de Investigación del Cáncer, Las Palmas, Spain

Full list of author information is available at the end of the article

may increase persistent mutations in daughter cell generations, genomic instability, and ultimately a more aggressive disease [3]. DNA repair genes play a major role in the DNA mismatch repair pathway, which includes base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR) and double strand break repair (DSBR); and are essential for maintaining the integrity of the genome [4]. Genetic variations in genes involved in DNA repair would confer susceptibility to the tumor, and would be associated to disease aggressiveness through the alteration of DNA repair pathways [5], which could induce tumor transformation and acquisition of oncologic properties. Single nucleotide polymorphisms (SNPs) are defined as inherited mutations that are present in more than 1% of the population. Given that there are millions of SNPs in the entire human genome, a major difficulty is to choose target SNPs that are most likely to affect phenotypic functions and ultimately contribute to disease development. Candidate gene studies are focused on the selection of genes that have been previously related to a disease, and thus come with prior knowledge about gene function. Among dozens of genes directly involved in DNA repair in humans [6], six of them have been previously studied in a wide series of Spanish PCa patients based on its relevance in the mechanism of the disease [7]: X-ray repair cross-complementing protein 1 (*XRCC1*), excision repair cross-complementing rodent repair deficiency, complementation group 2 (*ERCC2*), excision repair cross-complementing rodent repair deficiency, complementation group 1 (*ERCC1*), ligase IV (*LIG4*), ataxia telangiectasia mutated (*ATM*), and tumor protein p53 (*TP53*).

The ethnic origin of the studied population is a key factor in gene-association studies. In that sense, the literature is full of genetic variances that are risk factors for certain diseases among subjects of an ethnic origin but are not valid among subjects from other ethnicities [8,9]. Moreover, we have previously reported that differences in distribution of genotypes within different populations of the same ethnicity could be an important confounding factor in gene-association studies [7]. In that sense, since cohorts are often multi-ethnic, the STROGAR guidelines recommend to report whether ethnicity was controlled for in reporting of genotype-phenotype association, and encourage the use of cohorts from ethnically uniform populations [10].

We hypothesize here that genetic variations in DNA repair genes would confer different behavior to PCa cells and would result in a distinct clinical phenotype. Thus, the aim of the present study was to elucidate the relationship between 10 SNPs located in 6 different genes involved in DNA repair that have been classically associated to PCa risk [3], and tumor aggressiveness in a wide set of Spanish PCa patients.

Methods

Patients

A total of 601 patients with non-metastatic localized prostate cancer (PCa) from 4 different regions of Spain (15.14% from Andalusia, 8.48% from Basque Country, 39.60% from Canary Islands and 36.77% from Catalonia) were included in the study [7]. We have previously shown that differences in the distribution of genotypes within different populations of the same ethnicity are an important confounding factor in genetics epidemiology. In that sense, Andalusian subjects showed the greatest differences [7]. Thus, to homogenize the sample and minimize bias, we excluded this subset of patients from the analyses. A total of 494 PCa were included in the present study. All patients provided written informed consent before sample collection. The study was approved by the Research and Ethics Committee of each institution participant in the study: Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria), Hospital de la Esperanza - Parc de Salut Mar (Barcelona), Hospital Universitario Virgen de las Nieves (Granada), Hospital Universitari de Bellvitge (L'Hospitalet de Llobregat), Onkologikoa (Guipuzcoa), Institut Català d'Oncologia (L'Hospitalet de Llobregat), Hospital de la Santa Creu i Sant Pau (Barcelona) and Hospital Universitario Virgen del Rocío (Sevilla).

Clinical tumor size (cT), diagnostic PSA serum levels, and Gleason score [11] were recruited for all PCa patients. Clinical tumor size was assessed by digital rectal examination (DRE) followed by transrectal ultrasonography (TRUS) and magnetic resonance imaging (MRI); PSA serum levels were assessed by chemiluminescence in an Architect i2000 analyzer (Abbott Laboratories, IL, USA); Gleason score was determined in the biopsy specimen by a pathologist. Subjects were categorized into three risk-based recurrence groups according to D'Amico classification [12]: low, intermediate, and high risk. After collecting demographic and clinical data, a blood sample was taken after the signature of informed consent. All samples were sent by courier to the Hospital Universitario de Gran Canaria, for DNA isolation and genotype analyses as follows.

DNA isolation and quantification

DNA was obtained from blood samples at the Hospital Universitario de Gran Canaria Dr. Negrín. DNA was isolated from 300 μ l of whole-blood in an iPrep[™] Purification Instrument using the iPrep[™] PureLink[™] gDNA Blood Kit (Invitrogen, by Life Technologies, Carlsbad, CA), and its integrity was determined by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE).

Genes and SNPs

A total of 10 SNPs in 6 different genes involved in DNA repair were studied (Table 1): *XRCC1* (involved in base excision repair [6]), rs25487, rs25489, rs1799782; *ERCC2*

Table 1 Description of clinical variables

Clinical	n	(%)
Clinical tumor size (cT)		
cT1a – cT2a	270	(54.7)
cT2b – cT2c	141	(25.8)
cT3 – cT4	66	(13.4)
NA	17	(3.4)
Initial PSA (ng/mL)		
<10	306	(61.9)
10 – 19.99	108	(20.9)
>20	79	(16.0)
NA	6	(1.2)
Gleason score		
<7	226	(45.7)
7	195	(39.5)
>7	71	(14.4)
NA	2	(0.4)
D'Amico group		
Low	120	(24.3)
Intermediate	184	(37.2)
High	173	(35.0)
NA	17	(3.4)

Abbreviations: PSA prostate specific antigen, NA not available.

(involved in nucleotide excision repair [6]), rs13181; *ERCC1* (involved in nucleotide excision repair [6]), rs11615; *LIG4* (involved in double-strand break repair [6]), rs1805388, rs1805386; *ATM* (involved in double-strand break repair [6]), rs17503908, rs1800057; and *TP53* (involved in double-strand break repair [6]), rs1042522.

Genotyping

The SNP genotyping was performed in a Biotrove OpenArray[®] NT Cyclor (Applied Biosystems, Foster City, CA) [13]. DNA samples loaded in the OpenArray (OA) had a A260/A280 and A260/230 ratios of 1.7–1.9, and were adjusted to 50 ng/μL. A total of 300 ng of genomic DNA was used. A final amount of 150 ng was incorporated into the array with the autoloader, and was genotyped according to the manufacturer's recommendations. A non-template control (NTC) consisting of DNase-free double-distilled water was introduced within each assay. When the DNA and master mix were transferred, the loaded OA plate was filled with an immersion fluid and sealed with glue. The multiplex TaqMan assay reactions were carried out in a Dual Flat Block (384-well) GeneAmp PCR System 9700 (Applied Biosystems) with the following PCR cycle: an initial step at 93°C for 10 minutes followed by 50 cycles of 45 seconds at 95°C, 13 seconds at 94°C and 2:14 minutes at 53°C; followed by a final step during 2 minutes at 25°C and holding at 4°C.

The fluorescence was read using the OpenArray[®] SNP Genotyping Analysis software version 1.0.5. (Applied Biosystems). The genotyping analysis was made with the TaqMan Genotyper software version 1.0.1. (available at: <http://www.invitrogen.com>) using autocalling as the call method. The quality value of the data points was determined by a threshold above 0.95. Genotype analysis was performed with the same batch of chips and by the same investigator, as previously reported [7].

Statistical analysis

Genotype and allelic frequencies were determined using the web-based environment SNPator (SNP Analysis To Results, from the Spain's National Genotyping Centre and the National Institute for Bioinformatics) [14]. Relative excess of heterozygosity was determined to check compatibility of genotype frequencies with Hardy-Weinberg equilibrium (HWE). Thus, p-values from the standard exact HWE lack of fit test were calculated using SNPator.

Comparisons of genotypic and allelic frequencies were also done in SNPator.

All additional statistical analyses were performed using PASW Statistics 15 (IBM Corporation, Armonk, NY, USA).

Results

The majority of PCa patients were cT1a – cT2a (54.7%), PSA < 10 ng/mL (61.9%), and Gleason score < 7 (45.7%). Subsequently, a total of 120 patients (24.3%) were classified as low risk tumors according to D'Amico classification. We did not observed clinical differences among different regions of Spain (data not shown). Distribution of clinical variables is detailed in Table 1.

All the genotyped samples met the quality criteria stated above. A total of 494 PCa patients were genotyped for 10 SNPs. Of the 4,940 possible determinations, 97.17% were successfully genotyped. The genotypic and allelic frequencies are shown in Table 2. Minor allele frequencies (MAF) were similar to those reported in the literature. All SNPs were in HWE.

Among the 10 analyzed SNPs, rs11615 (minor allele frequency (MAF) = 0.39) and rs17503908 (MAF = 0.09), located in *ERCC1* and *ATM* respectively, were significantly different distributed among PCa patients according to the clinical variables (Additional file 1). Thus, rs11615 was significantly associated to the clinical tumor size (χ^2 test, $p = 0.002$) while rs17503908 was associated to the Gleason score (χ^2 test, $p = 0.005$). Concerning to rs11615, we observed that among the 259 patients diagnosed as cT1a – cT2a, 175 carried the G allele (67.57%). In the other hand, among the 66 patients diagnosed as cT3 – cT4, 31 carried the G allele (46.97%) (Additional file 2). With respect to rs17503908, 169 of the 224 patients (75.45%) scored with Gleason < 7 were genotyped

Table 2 Genotypic and allelic frequencies among Spanish prostate cancer patients

Gene/SNP	Function	chr	n	Genotypic frequencies				Allelic frequencies				MAF [#]	Functional consequence [#]		
XRCC1	BER	19q13													
rs25487*			436	CC	0.42	CT	0.48	TT	0.10	C	0.66	T	0.34	0.26	Missense
rs25489			483	CC	0.88	CT	0.12	TT	0.00	C	0.94	T	0.06	0.06	Missense
rs1799782*			487	AA	0.00	AG	0.12	GG	0.88	A	0.06	G	0.94	0.13	Missense
ERCC2	NER	19q13													
rs13181			482	GG	0.10	GT	0.47	TT	0.43	G	0.34	T	0.66	0.24	Missense
ERCC1	NER	19q13													
rs11615			488	AA	0.39	AG	0.46	GG	0.16	A	0.61	G	0.39	0.36	Synonymous codon
UG4	DSBR	13q23													
rs1805388 [†]			488	AA	0.04	AG	0.25	GG	0.71	A	0.17	G	0.83	0.15	Missense
rs1805386 [‡]			480	AA	0.70	AG	0.26	GG	0.04	A	0.88	G	0.12	0.10	Synonymous codon
ATM	DSBR	11q22													
rs17503908			486	GG	0.00	GT	0.18	TT	0.81	G	0.09	T	0.91	0.06	Intron variant
rs1800057			486	CC	0.94	CG	0.06	GG	0.00	C	0.97	G	0.03	0.02	Missense
TP53	DSBR	17p13													
rs1042522			484	CC	0.59	CG	0.35	GG	0.06	C	0.76	G	0.24	0.39	Missense

Abbreviations: BER base excision repair, NER nucleotide excision repair, DSBR double-strand break repair, chr chromosome, MAF minor allele frequency.

#Information available at: <http://www.ncbi.nlm.nih.gov/projects/SNP/>.

*SNPs in perfect linkage disequilibrium.

†SNPs in perfect linkage disequilibrium.

as TT, while 59 of the 70 patients (84.29%) scored with Gleason >7 were genotyped as TT (Additional file 2).

We explored the specific role of the SNPs rs11615 and rs17503908 in relation to the associated clinical variables. For this, we conducted the analysis according to various genetic models: recessive, dominant, homozygote, and heterozygote models (Table 3). We observed that patients carrying the GG + AG genotypes for rs11615 were at lower risk for develop cT2b- cT4 tumors than those AA PCa patients (OR = 0.50, (95% Confidence Interval (CI) 0.35 – 0.73), $p < 0.0001$). This result was also observed in the heterozygote model (Table 3). Similar trend was observed for rs17503908. Those patients carrying the GG + GT genotypes were at lower risk for develop Gleason scores 7 – 10 than those TT PCa patients (OR = 0.48, (95% CI 0.30 – 0.76), $p = 0.002$), and this trend was also observed in the heterozygote model (Table 3). A and T are

the ancestral alleles for rs11615 and rs17503908, respectively. According to our results, wild homozygous genotypes were associated to poor prognosis factors and these AA-rs11615 and TT-rs17503908 PCa patients were at higher risk for develop cT2b – cT4 (OR = 2.21 (CI 95% 1.47 – 3.31), $p < 0.001$) and Gleason scores ≥ 7 (OR = 2.22 (CI 95% 1.38 – 3.57), $p < 0.001$), respectively (data not shown).

We studied the role of rs11615 and rs17503908 in the context of D'Amico risk groups, a classification that estimate the biologic aggressiveness of prostate cancers by grouping them into different risk categories which reflect the risk of cancer growth and spread. Our results showed that those patients carrying the TT genotype for the rs17503908 SNP had an increased risk of developing D'Amico high-risk tumors (OR = 1.69, (CI 95% 1.01 – 2.83), $p = 0.044$). This trend was not observed for rs11615 SNP

Table 3 Univariate analysis for polymorphisms rs11615 (ERCC1) and rs17503908 (ATM) and clinical variables

Clinical variable	n	SNP	Recessive model		Dominant model		Homozygote		Heterozygote	
			GG vs. AG + AA	P	GG + AG vs. AA	P	GG vs. AA	P	AG vs. AA	P
cT, 1a-2a/2b-4	270/207	rs11615	1.02 (0.62 – 1.68)	0.928	0.50 (0.35 – 0.73)	<0.0001	0.67 (0.39 – 1.16)	0.152	0.45 (0.30 – 0.70)	<0.0001
			GG vs. GT + TT	P	GG + GT vs. TT	P	GG vs. TT	P	GT vs. TT	P
			OR (95% CI)		OR (95% CI)		OR (95% CI)		OR (95% CI)	
Gleason, <7/ ≥ 7	226/266	rs17503908	NA	–	0.48 (0.30 – 0.76)	0.002	NA	–	0.45 (0.28 – 0.72)	0.001

Abbreviations: OR odds ratio, CI confidence interval, NA not applicable (due to the limited sample size: only 2 subjects genotyped as GG-rs17503908).

Statistical test: binary logistic regression (Reference category for SNP rs11615: AA. Reference category for SNP rs17503908: TT).

(Figure 1). Nonetheless, those PCa patients carrying the AA-TT genotypes for rs11615 and rs17503908 respectively, had the greatest risk of developing D'Amico high-risk tumors (OR = 2.57 (CI 95% 1.28 – 5.16), $p = 0.008$) (Figure 1). Similar result was obtained when the series was dichotomized in low vs. intermediate-high D'Amico risk groups (OR = 1.98 (CI 95% 1.02 – 3.83), data not shown). These results are in line with previous results showed above, and suggest that these specific genotypes are associated to poor prognosis factors (Additional file 3).

No other polymorphism showed significant associations with clinical variables.

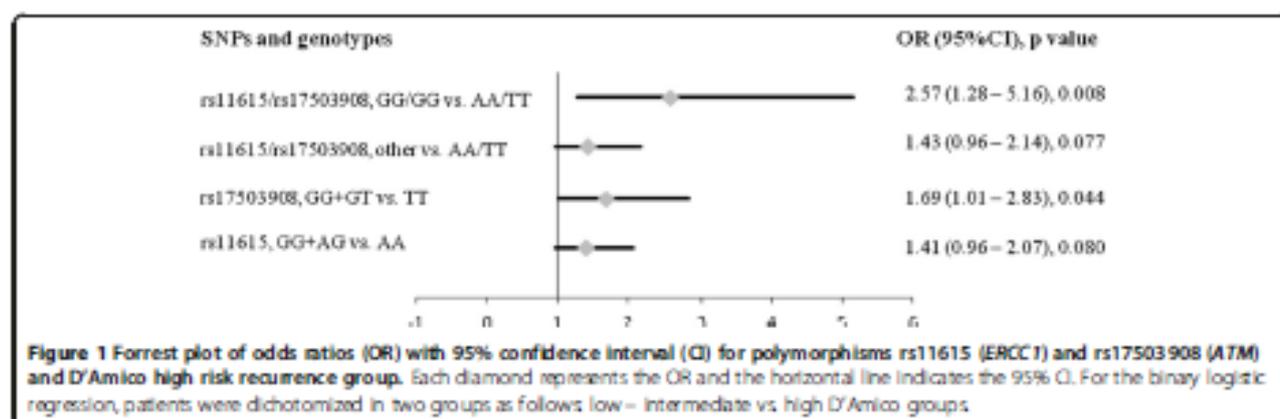
Discussion

DNA damage occurs very frequently and leads to gene deletions, amplifications, rearrangements, and translocation resulting in the alteration of cell homeostasis and tumor behaviour [2]. In the present study, we investigated the association between 10 SNPs located in DNA repair genes and the aggressiveness of prostate cancer in a wide set of Spanish PCa patients, following a candidate-gene approach based in a previously published study [7]. We observed a strong association of genotypes AA-rs11615 and TT-rs17503908 in the development of clinical variables of worst prognosis. These genetic variations would influence the nucleotide excision repair and double-strand break repair mechanisms of DNA, possibly favoring genomic instability and the development of more aggressive cell phenotypes that would cause the appearance of tumors of poor prognosis (i.e. D'Amico high-risk tumors).

There has been an increasing interest on the role of SNPs in the development and progression of PCa. In that sense, SNPs in DNA repair genes have been deeply explored, especially in the prediction of radiation-induced toxicity [15]. *ERCC1* is encoded in chromosome 19q13 and it is involved in nucleotide excision repair, forming with *XPF*, a free, nuclear flap structure-specific endonuclease [6]. In the context of PCa, it has been reported that polymorphisms affecting *ERCC1* may predispose prostate

epithelial cells to malignant transformation [16], but there is a lack of information about the role of that gene in disease aggressiveness. Genetic variants at chromosome 19q13 have been evaluated among 7,370 PCa cases, and no association with tumor aggressiveness was observed [1]. Although this study includes a big series, these patients came from different countries, and the ethnic origin was not considered as a confounding factor, especially when differences observed within populations of the same ethnic origin suggest that race is not a sufficient factor to ensure the homogeneity of the sample [17], as we have also previously published [7]. In the present study, we have observed that PCa patients carrying AA-rs11615 genotype were at higher risk for develop bigger tumors. Allele G, in combination with rs3212986 (also located in *ERCC1*), has been associated to low *ERCC1* expression, resulting in reduced DNA repair and better chemotherapy/radiotherapy response [18]. By extension, A allele would be required to maintain normal levels of *ERCC1*, thus conditioning tumor malignancy and response to treatment. In that sense, G allele, which is not the ancestral allele, would confer clinical advantage in terms of tumor size.

ATM is encoded in chromosome 11q22. In response to double-strand breaks (DSBs), *ATM* phosphorylates a variety of proteins involved in DSB repair and cell-cycle control [6]. *ATM/ATR* inactivation is a crucial step in promoting androgen-induced genomic instability and prostate carcinogenesis [19], and some missense variants of the *ATM* gene have been shown to confer a moderate increased risk of prostate cancer. Genetic variants at *ATM* have been associated with radiation-induced toxicity [20,21], although none of the previously reported associations were confirmed in a validated study [22]. A similar trend has been observed in the context of tumor aggressiveness [23]. We have observed that PCa patients carrying TT-rs17503908 genotype were at higher risk for develop high-grade tumors. Thus, G allele, which is not the ancestral allele, would confer clinical advantage in terms of Gleason score. Although this is a novel result



in PCa, a large variety of distinct *ATM* mutations and variants exist among breast cancer patients, and some of them can contribute to the etiology and progression of the malignancy [24].

Since it seems that wild genotypes represent a risk factor associated with tumor malignancy, we evaluated the combined role of both genotypes in relation to risk groups established by D'Amico. We observed that patients carrying both wild homozygous genotypes (AA + TT) were at higher risk of develop D'Amico high-risk tumors. *ERCC1* and *ATM* are encoded in different chromosomes, thus, there is not conserved combination of SNPs. However, genotypes frequencies for AA-rs11615 and TT-rs17503908 were 0.39 and 0.81, respectively; and a total of 148 PCa patients carried both genotypes (29.9% of the total series). Consequently, a combined analysis would give an idea about the role of both polymorphisms as predictors of tumor malignancy. Nonetheless, it has to be taken into account that functional consequences of rs11615 and rs17503908 are not missense; that is, there is not amino acid substitution in the translated protein. However, cumulative evidence suggest that synonymous mutations are also important, and there is a rapidly growing list of synonymous mutations that lead to human diseases [25,26]. Synonymous SNPs could affect protein function altering RNA secondary structures, affecting RNA stability, and subsequently reducing protein expression [27] and possibly affecting protein folding and function [28]. Therefore, it is possible that both polymorphisms can be important in determining the tumor characteristics of prostate cancer.

Candidate gene association studies are often criticized for their lack of validation. Thus, a replication study using an independent and random cohort of PCa patients should be considered in near future [29]. The present study has some weakness that need to be highlighted: i) although 494 PCa patients seem sufficient to obtain statistically reliable results, it is possible that some results may be of stochastic nature, especially for those SNPs with lower MAF; ii) other factors associated to prostate cancer (i.e. age, familiar aggregation, toxic habits or some kind of diets) have not been taken into account in the present study; iii) the observations should be confirmed in an independent cohort. Despite the above, the present study provides a number of advantages that contribute to their credibility: i) it is a multicenter study that provides patients from different areas of a country, minimizing the bias of studies performed in one hospital; ii) all subjects are from Spanish origin, and it has been reduced the possible influence of intra ethnic variations [7]; iii) all the determinations (4,940 in total) were performed with the same methodology (OpenArray, Applied Biosystems), with the same batch of chips and by the same investigator, thus minimizing biases from technical origin.

Conclusions

We found that genetic variants at DNA repair genes are associated with clinical variables of poor prognosis for prostate cancer. Prospective studies are required to validate our results.

Additional files

Additional file 1: Significant associations between clinical variables and SNPs.

Additional file 2: Genotype distribution of rs11615 across the different clinical tumor sizes (A) and distribution of rs17503908 across the different Gleason scores (B).

Additional file 3: Combine genotype distribution of SNPs rs11615 and rs17503908 according to the D' Amico risk -group. The genotype distribution was statistically different (χ^2 test, $p = 0.017$).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Substantial contributions to conception and design: LAHH and PCL. Acquisition of data: PFA, MJAC, JMC, JFSN, MCE, PRG, BDPC, MF, FG, GSP, JCB, MIOG, PCR, JRM. Analysis and interpretation of data: LAHH, AV, EHR and CRG. Drafting the article or revising it critically for important intellectual content: LAHH and PCL. Final approval of the version to be published: PCL. All authors read and approved the final manuscript.

Authors' information

LAHH and PCL are members of the Radiogenomics Consortium, an international consortium aimed to expand knowledge of the genetic basis for differences in radiosensitivity and to develop assays to help predict the susceptibility of cancer patients for the development of adverse effects resulting from radiotherapy.

Acknowledgments

We thank the technical support from the Immunology Department (Hospital Universitario de Gran Canaria Dr. Negrín) staff: Nerida González-Quevedo and Yanira Florido-Ortega. This work was subsidized by a grant from the Instituto de Salud Carlos III (Ministerio de Economía y Competitividad from Spain), ID: P11/2011/67. Almudena Valenciano has a grant from the Instituto Canario de Investigación del Cáncer (ICIC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author details

¹Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, C/Barranco de La Ballena s/n, CP 35010 Las Palmas de Gran Canaria, Las Palmas, Spain. ²Instituto Canario de Investigación del Cáncer, Las Palmas, Spain. ³Clinical Sciences Department, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain. ⁴Institut d'Oncologia Radioteràpica, Hospital de la Esperanza, Parc de Salut Mar, Universitat Pompeu Fabra, Barcelona, Spain. ⁵Laboratory of Genetic Identification, Legal Medicine and Toxicology Department, Facultad de Medicina, Universidad de Granada, Granada, Spain. ⁶GENYO, Pfizer-University of Granada-Andalusian Government Centre for Genomics and Oncological Research, Granada, Spain. ⁷Department of Urology, Hospital Universitario Virgen de las Nieves, Granada, Spain. ⁸Department of Urology, Hospital Universitario de Bellvitge, L'Hospitalet de Llobregat, Spain. ⁹Radiation Oncology Department, Onkologikoa, Guipuzkoa, Spain. ¹⁰Health Services Research Group, Institut de Recerca Hospital del Mar (IMM) Barcelona, Spain. ¹¹Department of Radiation Oncology, Institut Català d'Oncologia (ICO), L'Hospitalet de Llobregat, Spain. ¹²Radiation Oncology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. ¹³Radiation Oncology Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain. ¹⁴Department of Immunology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain. ¹⁵Department of Medical and Surgical Sciences, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain.

Received: 24 September 2014 Accepted: 17 December 2014
Published online: 24 December 2014

References

- Kote-Jaral Z, Easton DF, Stanbridge JL, Ostrander EA, Schleutker J, Ingles SA, Schaid D, Thibodeau S, Dörk T, Neal D, Donovan J, Hamdy F, Cox A, Maier C, Vogel W, Guy M, Muir K, Lophatananon A, Kadda MA, Spurdle A, Steingra S, John EM, Giles G, Hopper J, Chappuis PQ, Hutter P, Foulkes WD, Hamel N, Salinas CA, Koopmeiners JS, et al. Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL Consortium. *Cancer Epidemiol Biomarkers Prev* 2008, 17:2052–2061.
- Wood RD, Mitchell M, Sgouros J, Lindahl T: Human DNA repair genes. *Science* 2001, 291:1284–1289.
- Park JY, Huang Y, Sellers TA: Single nucleotide polymorphisms in DNA repair genes and prostate cancer risk. *Methods Mol Biol* 2009, 471:361–385.
- Yu Z, Chen J, Ford BN, Backley ME, Glickman BW: Human DNA repair systems: an overview. *Environ Mol Mutagen* 1999, 33:3–20.
- Zhu Y, Spitz MR, Amos CI, Lin J, Schabath MB, Wu X: An evolutionary perspective on single-nucleotide polymorphism screening in molecular cancer epidemiology. *Cancer Res* 2004, 64:2251–2257.
- Ronen A, Glickman BW: Human DNA repair genes. *Environ Mol Mutagen* 2001, 37:241–283.
- Henríquez-Hernández LA, Valendano A, Foro-Amalot P, Alvarez-Cubero MJ, Cazar JM, Suárez-Navo JF, Castells-Estève M, Ayala-Gil A, Ferrández-González P, Ferrer M, Guadaño F, Sancho-Pardo G, Craven-Bartle J, Ortiz-Gordillo MJ, Cabrera-Roldán P, Herrera-Ramos E, Lara PC: Polymorphisms in DNA-repair genes in a cohort of prostate cancer patients from different areas in Spain: heterogeneity between populations as a confounding factor in association studies. *PLoS One* 2013, 8:e69735.
- Gwercman C, Gwercman A, Pedersen HS, Toft G, Lundin K, Bonde JP, Lundberg Gwercman Y: Polymorphisms in genes regulating androgen activity among prostate cancer low-risk Inuit men and high-risk Scandinavians. *Int J Androl* 2008, 31:25–30.
- Correa-Cerro L, Wöhr G, Häussler J, Berthou P, Drelon E, Mangin P, Fournier G, Cusanotto O, Kraus P, Just W, Pils T, Canté JM, Vogel W: (CAG)_nCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. *Eur J Hum Genet* 1999, 7:357–362.
- Kerns SL, de Ruyscher D, Andreasen CN, Azria D, Barnett GC, Chang-Claude J, Davidson S, Deasy JO, Dunning AM, Ostrer H, Rosenzweig BS, West CM, Bertzen SM: STROGAR - Strengthening the reporting of genetic association studies in radiogenomics. *Radiat Oncol* 2014, 110:182–188.
- Epstein JI, Allbrock WC Jr, Amin MB, Egevad LL: The 2005 International society of urological pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol* 2005, 29:1228–1242.
- D'Amico AV, Whittington R, Malkowicz SB, Wu YH, Chen MH, Hurwitz M, Kantoff PW, Tomaszewski JE, Renshaw AA, Wein A, Richie JP: Utilizing predictions of early prostate-specific antigen failure to optimize patient selection for adjuvant systemic therapy trials. *J Clin Oncol* 2000, 18:3240–3246.
- Henríquez-Hernández LA, Valendano A, Herrera-Ramos E, Lloret M, Riveros-Perez A, Lara PC: High-throughput genotyping system as a robust and useful tool in oncology: Experience from a single institution. *Biologicals* 2013, 41:424–429.
- Mordillo-Suarez C, Alegre J, Sangros R, Gazave E, de Cid R, Milne R, Amigo J, Ferrer-Admetlla A, Moreno-Estrada A, Gardner M, Casals F, Pérez-Lezaola A, Comas D, Bosch E, Calafell F, Bertrand J, Navarro A: SNP analysis to results (SNPator): a web-based environment oriented to statistical genomics analyses upon SNP data. *Bioinformatics* 2008, 24:1643–1644.
- Van den Broeck T, Jorjau S, Clincemallo L, Helten C, Prekovic S, Spars L, Tosco L, Van Poppel H, Claessens F: The role of single nucleotide polymorphisms in predicting prostate cancer risk and therapeutic decision making. *Bioméd Res Int* 2014, 2014:627510.
- Márka DJ, Yao V, Hara DS, Gregg J, Robinson AR, Niedemhöfer LJ, Panwani AV, Maier C, Bach D: Deficiency of DNA repair nuclease ERCC1-XPF promotes prostate cancer progression in a tissue recombination model. *Prostate* 2012, 72:1214–1222.
- Bauchet M, McEvoy B, Pearson LN, Quinn EE, Sarkisian T, Hovhannesian K, Dola R, Badley DG, Shriver MD: Measuring European population stratification with microarray genotype data. *Am J Hum Genet* 2007, 80:948–956.
- Woolfshneider A, Popanda O, Ulla C, Lindeisen J, Mayer C, Colebi O, Debus J, Bartsch H, Chang-Claude J, Schmezer P: A distinct ERCC1 haplotype is associated with mRNA expression levels in prostate cancer patients. *Carcinogenesis* 2008, 29:1758–1764.
- Chiu YT, Liu J, Tang K, Wong YC, Khanna KK, Ling MT: Inactivation of ATM/ATR DNA damage checkpoint promotes androgen induced chromosomal instability in prostate epithelial cells. *PLoS One* 2012, 7:e31108.
- Cesaretti JA, Stock RG, Atencio DP, Peters SA, Peters CA, Burr RJ, Stone NN, Rosenzweig BS: A genetically determined dose-volume histogram predicts for rectal bleeding among patients treated with prostate brachytherapy. *Int J Radiat Oncol Biol Phys* 2007, 68:1410–1416.
- Cesaretti JA, Stock RG, Lehrer S, Atencio DA, Bernstein JL, Stone NN, Wallenstein S, Green S, Loeb K, Kälmbauer M, Smith M, Rosenzweig BS: ATM sequence variants are predictive of adverse radiotherapy response among patients treated for prostate cancer. *Int J Radiat Oncol Biol Phys* 2005, 61:196–202.
- Barnett GC, Coles CE, Elliot RM, Baynes C, Luccarini C, Conroy D, Wilkinson JS, Tyrer J, Mira V, Pato R, Gulliford SL, Sydes MR, Hall E, Bertzen SM, Dearnaley DP, Burnett NG, Pharaoh PD, Dunning AM, West CM: Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study. *Lancet Oncol* 2012, 13:65–77.
- Browning JE, Li H, Shinohara ET, Cai Q, Chen H, Coutney R, Cao C, Zheng W, Lu B: ATM polymorphism IVS62 + 60G > A is not associated with disease aggressiveness in prostate cancer. *Urology* 2006, 67:1320–1323.
- Dörk T, Bendix R, Bremer M, Rades D, Köpfer K, Nickle M, Skarwan B, Hector A, Yamini P, Steinhilber D, Weise S, Stührmann M, Kretzschmar JH: Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. *Cancer Res* 2001, 61:7608–7615.
- Chamary JV, Parmley JL, Hurst LD: Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat Rev Genet* 2006, 7:98–108.
- Cartegni L, Chew SL, Krainer AR: Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002, 3:285–298.
- Nackley AG, Shabalina SA, Tchivileva E, Satterfield K, Korzhynskiy O, Makarov SS, Makarov W, Diatchenko LI: Human cacchyl-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 2006, 314:1930–1933.
- Kimchi-Sarafy C, Oh JM, Kim W, Saura JE, Calcagno AM, Ambudkar SV, Gottesman MM: A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* 2007, 315:525–528.
- Patwa R, Clements J, Batra J: Candidate gene association studies: a comprehensive guide to useful in silico tools. *BMC Genet* 2013, 14:39.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



ORIGINAL ARTICLE

Association between single-nucleotide polymorphisms in DNA double-strand break repair genes and prostate cancer aggressiveness in the Spanish population

LA Henríquez-Hernández^{1,2,3}, A Valenciano², P Foro-Arnalot⁴, MJ Álvarez-Cubero^{5,6}, JM Cozar⁷, JF Suárez-Navo⁸, M Castells-Esteve⁹, P Fernández-González⁹, B De-Paula-Carranza⁹, M Ferrer¹⁰, F Guedea¹¹, G Sancho-Pardo¹², J Craven-Bartle¹², MJ Ortiz-Gordillo¹³, P Cabrera-Roldán¹³, JI Rodríguez-Melcón^{1,2}, E Herrera-Ramos^{14,15}, C Rodríguez-Gallego¹⁴ and PC Lara^{1,2,3}

BACKGROUND: Novel predictors of prognosis and treatment response for prostate cancer (PCa) are required to better individualize treatment. Single-nucleotide polymorphisms (SNPs) in four genes directly (*XRCC5* (X-ray repair complementing defective repair in Chinese hamster cells 5) and *XRCC6* (X-ray repair complementing defective repair in Chinese hamster cells 6)) or indirectly (*PAP1* and major vault protein (*MVP*)) involved in non-homologous end joining were examined in 494 Spanish PCa patients.

METHODS: A total of 22 SNPs were genotyped in a Biotrove OpenArray NT Cyder. Clinical tumor stage, diagnostic PSA serum levels and Gleason score at diagnosis were obtained for all participants. Genotypic and allelic frequencies were determined using the web-based environment SNPator.

RESULTS: (*XRCC6*) rs2267437 appeared as a risk factor for developing more aggressive PCa tumors. Those patients carrying the GG genotype were at higher risk of developing bigger tumors (odds ratio (OR) = 2.04, 95% confidence interval (CI) 1.26–3.29, $P = 0.004$), present higher diagnostic PSA levels (OR = 2.12, 95% CI 1.19–3.78, $P = 0.011$), higher Gleason score (OR = 1.65, 95% CI 1.01–2.68, $P = 0.044$) and D’Amico higher risk tumors (OR = 2.38, 95% CI 1.24–4.58, $P = 0.009$) than those patients carrying the CC/CG genotypes. Those patients carrying the (*MVP*) rs3815824 TT genotype were at higher risk of presenting higher diagnostic PSA levels (OR = 4.74, 95% CI 1.40–16.07, $P = 0.013$) than those patients carrying the CC genotype. When both SNPs were analyzed in combination, those patients carrying the risk genotypes were at higher risk of developing D’Amico higher risk tumors (OR = 3.33, 95% CI 1.56–7.17, $P = 0.002$).

CONCLUSIONS: We believe that for the first time, genetic variants at *XRCC6* and *MVP* genes are associated with risk of more aggressive disease, and would be taken into account when assessing the malignancy of PCa.

Prostate Cancer and Prostatic Diseases (2016) 19, 28–34; doi:10.1038/pcan.2015.63; published online 12 January 2016

INTRODUCTION

Prostate cancer (PCa) is a complex disease conditioned by hormonal and genetic factors. Tumor staging, tumor grading in terms of Gleason score and diagnostic PSA serum levels are used in the daily clinical practice to classify patients into different prognostic risk groups that condition treatment decisions. However, these clinical prognostic factors only explain a proportion of the patient-to-patient variation observed in the clinical outcome,¹ and it seems that better predictors of individualized prognosis and treatment response are urgently needed.

DNA double-strand break (DSB) is the most detrimental form of DNA damage, leading cells to chromosomal breakage and rearrangements.² DSBs can have exogenous (that is, exposure to ionizing radiation) and endogenous (that is, due to reactive

oxygen species) origin, and can also be generated during somatic recombination, DNA replication and when DNA single-strand breaks are encountered.² Unrepaired or misrepaired DSBs would cause cell death, genomic instability and/or oncogenic transformation.³ In eukaryotic cells, DSBs are repaired by two different pathways: homologous recombination (HR) and non-homologous end joining (NHEJ).⁴ Recently, single-nucleotide polymorphisms (SNPs)—defined as inherited mutations that are present in > 1% of the population—in HR genes have been proposed as novel factors associated with PCa aggressiveness.⁵ Because of the complexity of mammalian genomes, searching of homologous sequences for HR is not efficient. Thus, NHEJ has a predominant role in DSB repair.⁶ This fact makes sense when is taken into account that most of the mammalian genome consists of non-

¹Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain; ²Instituto Canario de Investigación del Cáncer, Las Palmas, Spain; ³Clinical Sciences Department, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain; ⁴Radiation Oncology Department, Institut d’Oncologia Radioteràpica, Hospital de la Esperanza, Parc de Salut Mar, Barcelona, Spain; ⁵Laboratory of Genetic Identification, Legal Medicine and Toxicology Department, Facultad de Medicina, Universidad de Granada, Granada, Spain; ⁶GENYO, Pfizer-University of Granada-Andalusian Government Centre for Genomics and Oncological Research, Granada, Spain; ⁷Department of Urology, Hospital Universitario Virgen de las Nieves, Granada, Spain; ⁸Department of Urology, Hospital Universitario de Bellmuga, L’Hospitalet de Llobregat, Barcelona, Spain; ⁹Radiation Oncology Department, Oncologia, Gipuzkoa, Spain; ¹⁰Health Services Research Group, Institut de Recerca Hospital del Mar (IMM), Barcelona, Spain; ¹¹Department of Radiation Oncology, Institut Català d’Oncologia (ICO), L’Hospitalet de Llobregat, Barcelona, Spain; ¹²Radiation Oncology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; ¹³Radiation Oncology Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain; ¹⁴Department of Immunology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain and ¹⁵Department of Medical and Surgical Sciences, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain. Correspondence: Dr LA Henríquez-Hernández. E-mail: lah@henriquezgipuzkoa.es

Received 22 September 2015; revised 27 October 2015; accepted 24 November 2015; published online 12 January 2016

coding sequences.⁷ Thus, common variants in genes involved in DNA DSB repair are good candidates for low-penetrance cancer biomarkers.⁸

The NHEJ repair mechanism starts with the recruitment of Ku70/Ku80 proteins to the DNA ends. Ku70 is encoded by the X-ray repair complementing defective repair in Chinese hamster cells 6 (*XRCC6*) gene, whereas Ku80 is encoded by the X-ray repair complementing defective repair in Chinese hamster cells 5 (*XRCC5*) gene. Although other proteins are involved in end ligation (that is, DNA protein kinase (DNA-PK), *XRCC4* or ligase IV (*LIG4*)), the well function of the Ku70/Ku80 complex is critical when NHEJ starts.⁹ To our knowledge, studies concerning polymorphic variants of NHEJ genes have not been done in PCa.

Vaults are ribonucleoprotein particles with a hollow barrel-like structure.¹⁰ It is composed of three proteins: the major vault protein (MVP), the vault poly(adenosine diphosphate-ribose) polymerase also known as VPARP, and telomerase-associated protein-1, together with small untranslated RNA (vRNA). Classically, MVP has been associated with the development of drug resistance.¹¹ However, in recent years there has been increasing interest in MVP, especially as MVP transcription and protein levels are increased in response to DNA-damaging agents including ionizing radiation.¹² In addition to its implication in the regulation of several cellular processes including transport mechanisms, signal transduction and immune responses,¹³ the interaction with partners such as PTEN or BCL-2 suggests a role in the cell fate.^{14,15} Furthermore, MVP has been recently linked to DSB repair machinery due to an association with Ku70/Ku80 expression.¹⁶ Thus, SNPs in MVP would have a role in PCa behavior, due to the role in the resistance to treatment, the regulation of apoptosis or the association with DNA damage repair mechanisms.¹³

Poly(ADP-ribose) synthase 1 is a protein encoded by the gene *PARP1*, which is associated with differentiation, proliferation and tumor transformation.¹⁷ *PARP1* seems to have a role in DNA DSB repair through an interaction with *BRCA1/2* and other genes.¹⁸ Moreover, *PARP1* and *Ku* compete for repair of DNA DSBs by different NHEJ pathways.¹⁹ In the last years, *PARP1* inhibitors have shown promising activity in ovarian and breast cancers patients with *BRCA1/2* mutation,²⁰ and also in PCa patients.²¹ The fact that *PARP1* is a crossroad between HR and NHEJ and it is also of interest as a novel therapeutic target makes this gene a good candidate associated with tumor aggressiveness. Nowadays, the role of single-nucleotide variations in *PARP1* has not been studied in the context of PCa.

The aim of this study was to test the hypothesis that genetic variants in genes involved directly (*XRCC6* and *XRCC5*) or indirectly (*MVP* and *PARP1*) in DNA DSB repair confer increased risk of more aggressive PCa. To our knowledge, this is the first time that SNPs located in these genes are studied in PCa patients.

MATERIALS AND METHODS

Patients

A total of 601 patients with non-metastatic localized PCa from four different regions of Spain (15.14% from Andalusia, 8.48% from Basque Country, 39.60% from Canary Islands and 36.77% from Catalonia) were included in the study between April 2003 and December 2012, in the context of a multicenter study aimed to evaluate quality of life and clinical outcome of PCa patients.²² All patients provided written informed consent before sample collection. The study was approved by the Research and Ethics Committee of each institution participant in the study: Hospital Universitario de Gran Canaria Dr Negrín (Las Palmas de Gran Canaria), Hospital de la Esperanza, Parc de Salut Mar (Barcelona), Hospital Universitario Virgen de las Nieves (Granada), Hospital Universitario de Bellvitte (L'Hospitalet de Llobregat), Onkologikoa (Guipuzcoa), Institut Català d'Oncologia (L'Hospitalet de Llobregat), Hospital de la Santa Creu i Sant Pau (Barcelona) and Hospital Universitario Virgen del Rocío (Sevilla). Details of the series are described previously.^{23,24}

Clinical tumor size (cT), diagnostic PSA serum levels and Gleason score²⁵ were recruited for all PCa patients. cT was assessed by digital rectal examination followed by transrectal ultrasonography and magnetic resonance imaging; PSA serum levels were assessed by chemiluminescence in an Architect i2000 analyzer (Abbott Laboratories, Lake Bluff, IL, USA); Gleason score was determined in the biopsy specimen by a specialized genitourinary pathologist from each institution. Subjects were categorized into three risk-based recurrence groups according to D'Amico classification:²⁶ low, intermediate and high risks. After collecting demographic and clinical data, a blood sample was taken after the signature of informed consent. All samples were sent by courier to the Hospital Universitario de Gran Canaria Dr Negrín for DNA isolation and genotype analyses.

DNA isolation and quantification

DNA was isolated from 300 μ l of whole-blood in an iPrep™ Purification Instrument using the iPrep™ PureLink™ gDNA Blood Kit (Invitrogen, by Life Technologies, Carlsbad, CA, USA), and its integrity was determined by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE, USA).

Selection criteria for SNPs

SNPs were selected using data of individuals with European ancestry (CEU) from the HapMap Project (available at: www.hapmap.org). Pairwise linkage disequilibrium tagging was achieved with the Haploview v4.2 software.²⁷ This strategy is based on the haplotype blocks formed by clustered SNPs.²⁸ By selecting a modest number of SNPs within a block, it is possible to recover up to 95% of the heterozygosity present at a locus. The R^2 given by Haploview was > 0.95 for all genes.

Thus, a total of 22 SNPs in 4 different genes involved in DNA DSB repair were studied (Supplementary Table S1): *XRCC6*, rs5751131, rs2267437, rs11912946, rs7291732, rs881092; *XRCC5*, rs1685458, rs9288516, rs1051677, rs1051685; *PARP1*, rs8679, rs3219123, rs1805410, rs3219062, rs1805414, rs1805404, rs3219027; and *MVP*, rs9923649, rs12149514, rs4788186, rs3815824, rs35916172, rs3764944.

Genotyping

The SNP genotyping was performed in a Biotrove OpenArray NT Cycler (Applied Biosystems, Foster City, CA, USA).²⁹ DNA samples loaded in the OpenArray (OA) had a A260/A280 and A260/230 ratio of 1.7–1.9, and were adjusted to 50 ng μ l⁻¹. A total of 300 ng of genomic DNA was used. A final amount of 150 ng was incorporated into the array with the autoloader, and was genotyped according to the manufacturer's recommendations. A non-template control consisting of DNase-free double-distilled water was introduced within each assay. When the DNA and master mix were transferred, the loaded OA plate was filled with an immersion fluid and sealed with glue. The multiplex TaqMan assay reactions were carried out in a Dual Flat Block (384-well) GeneAmp PCR System 9700 (Applied Biosystems) with the following PCR cycle: an initial step at 93 °C for 10 min followed by 50 cycles of 45 s at 95 °C, 13 s at 94 °C and 2:14 min at 53 °C; followed by a final step during 2 min at 25 °C and holding at 4 °C.

The fluorescence was read using the OpenArray SNP Genotyping Analysis software version 1.0.5. (Applied Biosystems). The genotyping analysis was made with the TaqMan Genotyper software version 1.0.1 (available at: <http://www.invitrogen.com>) using autotyping as the call method. The quality value of the data points was determined by a threshold above 0.95. Genotype analysis was performed with the same batch of chips and by the same investigator, as previously reported.^{5,23,24}

About 10% of the samples were randomly selected for confirmation, and the results were all consistent.

Statistical analysis

Genotype and allelic frequencies were determined using the web-based environment SNPator (SNP Analysis To Results, from the Spain's National Genotyping Centre and the National Institute for Bioinformatics).³⁰ Relative excess of heterozygosity was determined to check compatibility of genotype frequencies with Hardy-Weinberg equilibrium. Thus, P -values from the standard exact Hardy-Weinberg equilibrium lack of fit test were calculated using SNPator. Comparisons of genotypic and allelic frequencies were also done in SNPator. All additional statistical analyses were performed using PASW Statistics 15 (IBM, Armonk, NY, USA) as follows: descriptive analyses were conducted for clinical variables, categorical variables were represented as absolute numbers and percentages,

comparisons between groups were performed using chi-squared test (categorical variables), the associations between clinical variables and genetic variables were analyzed using a linear additive model employing logistic regressions. Different genetic models were taken into account.

RESULTS

Intra-ethnic variations of SNPs

All the genotyped samples met the quality criteria stated above. A total of 601 PCa patients were genotyped for 22 SNPs. Of the 13 222 possible determinations, 12 608 were successfully genotyped (95.0%).

We have previously shown that differences in the distribution of genotypes within different populations of the same ethnicity are an important confounding factor in genetics epidemiology.^{23,24} Thus, we first analyzed the genotype distribution among subjects from different Spanish regions, highlighting that all PCa patients were from the Caucasian origin. The genotypic and allelic frequencies are shown in Supplementary Table S2. We observed that 9 of the 22 analyzed SNPs were differentially distributed (*XRCC6* rs5751131, rs2267437, rs7291732, rs881092; *XRCC5* rs1051677, (*PARP1*) rs3219123, rs1805410, rs1805414; and (*MVP*) rs4788186. As expected, Andalusian population showed a genotypic and allelic distribution clearly different from those reported among subjects from Basque Country, Canary Islands and Catalonia. Thus, we excluded Andalusian patients from subsequent analyses to homogenize the sample and minimize bias. We have previously followed this strategy of analysis.⁵

Clinical variables associated with individual SNPs

A total of 494 PCa were included in the subsequent association studies. Distribution of clinical variables is detailed in Table 1. The majority of PCa patients were cT1a–cT2a (54.7%), PSA < 10 ng ml⁻¹ (61.9%) and Gleason score < 7 (45.7%). Subsequently, a total of 120 patients (61.5%) were classified as low-intermediate risk tumors according to D’Amico classification.

Among the 22 analyzed SNPs, (*XRCC6*) rs2267437, (*XRCC5*) rs1051685, (*PARP1*) rs3219123 and (*MVP*) rs3815824, were significantly different distributed among PCa patients according to the clinical variables (Table 2). Genotype distribution was consistent with Hardy–Weinberg equilibrium (data not shown).

We observed that rs2267437 was significantly associated with the cT (χ^2 test, $P=0.004$). Thus, among the 249 PCa patients with cT1a–cT2a tumors, 37 carried the GG genotype (14.8%), while among the 183 PCa patients with cT2b–cT4 tumors, 48 carried the GG genotype (26.2%). Similar results were observed regarding to Gleason score and D’Amico risk groups (χ^2 test, $P=0.037$ and $P=0.015$, respectively) (Table 2). In both cases, GG carriers showed more aggressive tumor characteristics; thus, GG genotype was detected in 12 out of 110 low-risk tumors and in 73 out of 323 intermediate–high risk tumors (10.9 vs 22.6%, respectively). Univariate analyses showed that GG genotype for (*XRCC6*) rs2267437 appeared as a risk factor for developing more aggressive PCa tumors (Table 3). The fact that recessive and homozygote models showed significant results reinforces the role of GG genotype as a factor of worst tumor characteristics. In that sense, those PCa patients carrying the GG genotype were at higher risk of developing bigger tumors (odds ratio (OR)=2.04, 95% confidence interval (CI) 1.26–3.29, $P=0.004$), present higher diagnostic PSA levels (OR=2.12, 95% CI 1.19–3.78, $P=0.011$), higher Gleason score (OR=1.65, 95% CI 1.01–2.68, $P=0.044$) and D’Amico higher risk tumors (OR=2.38, 95% CI 1.24–4.58, $P=0.009$) than those patients carrying the CC/CG genotypes (Table 3).

We observed that levels of diagnostic PSA were associated with three SNPs located in three different genes (Table 2). Among 404 PCa patients showing diagnostic PSA levels < 20 ng ml⁻¹, 78

Table 1. Description of clinical variables (n=494)

Variable	n	(%)
Clinical tumor size (cT)		
cT1a–cT2a	270	(54.7)
cT2b–cT2c	141	(25.8)
cT3–cT4	66	(13.4)
A	17	(3.4)
Initial PSA (ng ml⁻¹)		
< 10	306	(61.9)
10–19.99	103	(20.9)
> 20	79	(16.0)
NA	6	(1.2)
Gleason score		
< 7	226	(45.7)
7	195	(39.5)
> 7	71	(14.4)
NA	2	(0.4)
D’Amico risk group		
Low	120	(24.3)
Intermediate	184	(37.2)
High	173	(35.0)
NA	17	(3.4)

Abbreviation: NA, not available.

carried the (*XRCC5*) rs1051685 G allele (19.3%), while 8 out of 79 subjects with initial PSA levels ≥ 20 ng ml⁻¹ carried the G allele (10.1%) (χ^2 test, $P=0.007$). This result suggests that G allele appeared as a protective factor, a fact reinforced in univariate analysis: those PCa patients carrying the AG genotype were at lower risk of present higher diagnostic PSA levels (OR=0.388, 95% CI 0.16–0.93, $P=0.033$) than those patients carrying the AA genotype (Heterozygote model, data not shown). On the contrary, those PCa patients carrying the (*MVP*) rs3815824 TT genotype were at higher risk of presenting higher diagnostic PSA levels (OR=4.74, 95% CI 1.40–16.07, $P=0.013$) than those patients carrying the CC genotype (Homozygous model, data not shown). Finally, although (*PARP1*) rs3219123 was significantly associated with diagnostic PSA levels (χ^2 test, $P=0.041$; Table 2), this SNP did not appear as a risk factor for that clinical variable in univariate analysis (data not shown).

Combined analyses of multiple SNPs

As it is plausible that the genes considered in the present study contribute to PCa behavior—due to a direct or indirect interaction with NHEJ pathway—we estimated the combined effect of the significantly associated genes (Figure 1). We considered (*XRCC6*) rs2267437 as the basics for these analyses, because its high impact in tumor characteristics (as previously shown in Tables 2 and 3). Although the combination with (*XRCC5*) was not significant (OR=2.11, 95% CI 0.91–4.94), the combined role of (*XRCC6*) rs2267437 with (*PARP1*) rs3219123 or (*MVP*) rs3815824 was risk factor associated with the D’Amico risk group. Thus, PCa patients carrying genotypes different from those considered as reference (CC+CC and CC+GG for rs2267437+rs3815824 and rs2267437+rs3219123, respectively) were at higher risk of developing intermediate–high D’Amico risk groups (Figure 1). Interestingly, the combination of these three SNPs did not showed a significant increment in the risk, suggesting that (*XRCC6*) rs2267437 and (*MVP*) rs3815824 have a predominant role in PCa tumor behavior (OR=3.33, 95% CI 1.56–7.17, $P=0.002$).

DISCUSSION

Although polymorphic alleles of NHEJ genes would predispose carriers to a high risk of developing more aggressive tumors, to date, no direct clinical evidence has been obtained in patients with PCa. We have previously published that common genetic variations in genes involved in HR (such as *XRCC1*, *ERCC1* or *ERCC2*, among others) are associated with clinical variables of poor prognosis for PCa.⁵ In the present study, we focused in the study

of classical genes involved in NHEJ (*XRCC6* and *XRCC5*) as well as other genes associated with DSB DNA repair (*PARP1* and *MVP*). To our knowledge, this is the first comprehensive study that analyzes the possible role of these four genes and its influence in PCa behavior. Although it is a candidate gene study, we used an SNP tagging approach that reduces the bias inherent to the selection of well-published SNPs and allows the inclusion of novel genetic variants with possible new biological roles.

A total of 601 PCa patients from 4 different Spanish regions were included in the study. Intra-ethnic variations of genotype and allelic frequencies are considered as an important confounding factor frequently obviated. Confounding would arise if the population contained several ethnic groups, if allele frequencies at the locus of interest differed between groups and if disease frequency also differed between groups for reasons quite unrelated to the locus of interest.³¹ The phenomenon was previously observed in our series,^{23,24} a fact explained by differences in the natural history of the populations.^{32,33} As expected, the Andalusian population showed significant differences in genotype and allelic distribution, and subsequently, this subset of PCa patients was excluded from the association analyses to avoid the intra-ethnic bias.

Ku70/80 is a protein complex encoded by *XRCC6* and *XRCC5* genes, which is involved in starting the NHEJ process. Specifically, ku70/80 is responsible for the recognition of the DNA ends at the site where the DSB lesion takes place.^{9,24,25} As ku70/80 is a complex formed by two different monomers, the functionality of the complex could be compromised if there are structural changes in the proteins, and it is well known that genetic polymorphisms can compromise not only the binding of a protein complex, but also the strength of the bounds between proteins and DNA.³⁴ *XRCC6* rs2267437 appeared as a risk factor associated with clinical variables of poor prognosis for PCa in terms of tumor size, diagnostic PSA level, Gleason score and even D'Amico risk groups. The rs2267437 polymorphism has been investigated as a risk factor for cancer risk with inconsistent results. However, a recent systematic review and meta-analysis concluded that rs2267437 SNP is associated with a significant increase in risk of breast cancer, renal cell carcinoma and hepatocellular carcinoma.³⁶ Nonetheless, this study did not include any PCa patient and no associations with clinical variables were analyzed. The role of rs2267437—specifically the GG genotype—in cancer seems to be established in other meta-analysis.³⁷ Our results reinforce these previous results and postulate rs2267437 SNP as a novel genetic factor for PCa behavior.

The role of *XRCC5* rs1051685 has been previously evaluated in myeloma patients. Those patients carrying the GG genotype were at lower risk of develop myeloma.³⁸ This result was proven among Spanish multiple myeloma patients³⁹ and agree with those reported in the present study in PCa patients. Although it is expectable to find a link between SNPs in *XRCC5* and *XRCC6*, when

Table 2. Significant associations between clinical variables and SNPs

SNP	Genotypes	Clinical tumor size (cT)			P ^a
		cT1a-cT2a	cT2b-cT2c	cT3-cT4	
rs2267437 (<i>XRCC6</i>)	CC	105 (24.3)	41 (9.5)	29 (6.7)	0.004
	CG	107 (24.8)	45 (10.4)	20 (4.6)	
	GG	37 (8.6)	38 (8.8)	10 (2.3)	
		Diagnostic PSA (ng ml ⁻¹)			
		< 10	10–19.99	≥ 20	
rs1051685 (<i>XRCC5</i>)	AA	253 (52.4)	73 (15.1)	71 (14.7)	0.007
	AG	46 (9.5)	25 (5.2)	6 (1.2)	
	GG	3 (0.6)	4 (0.8)	2 (0.4)	
rs3219123 (<i>PARP1</i>)	AA	1 (0.2)	0 (0.0)	1 (0.2)	0.041
	AG	29 (6.0)	18 (3.7)	4 (0.8)	
	GG	274 (56.6)	84 (17.4)	73 (15.1)	
rs3815824 (<i>MVP</i>)	TT	4 (0.8)	2 (0.4)	5 (1.0)	0.025
	TC	61 (12.7)	29 (6.0)	18 (3.8)	
	CC	238 (49.6)	69 (14.4)	54 (11.3)	
		Gleason score			
		< 7	7	> 7	
rs2267437 (<i>XRCC6</i>)	CC	91 (20.4)	58 (13.0)	31 (7.0)	0.037
	CG	82 (18.4)	78 (17.5)	19 (4.3)	
	GG	31 (7.0)	41 (9.2)	14 (3.1)	
		D'Amico risk group			
		Low	Intermediate	High	
rs2267437 (<i>XRCC6</i>)	CC	52 (12.0)	58 (13.4)	64 (14.8)	0.015
	CG	46 (10.6)	76 (17.6)	52 (12.0)	
	GG	12 (2.8)	34 (7.9)	39 (9.2)	

Abbreviations: MVP, major vault protein; SNP, single-nucleotide polymorphism; XRCC5, X-ray repair complementing defective repair in Chinese hamster cells 5; XRCC6, X-ray repair complementing defective repair in Chinese hamster cells 6. The distribution of patients is expressed in absolute numbers and percentages (in brackets). ^aChi-square test.

Table 3. Univariate analysis for polymorphism rs2267437 (*XRCC6*) and clinical variables

Clinical variable	n	Recessive model		Dominant model		Homozygote model		Heterozygote model	
		GG vs CG+CC		CG+GG vs CC		CC vs GG		CG vs CC	
		OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
cT, 1a–2a/2b–4	270/207	2.04 (1.26–3.29)	0.004	1.18 (0.80–1.74)	0.413	1.95 (1.15–3.29)	0.013	0.91 (0.59–1.40)	0.673
PSA, < 20/≥ 20	409/79	2.12 (1.17–3.78)	0.011	1.30 (0.75–2.21)	0.355	2.07 (1.08–3.98)	0.029	0.96 (0.52–1.77)	0.892
Gleason, < 7/≥ 7	226/266	1.65 (1.01–2.68)	0.044	1.37 (0.94–2.01)	0.100	1.81 (1.07–3.08)	0.027	1.21 (0.80–1.83)	0.368
Risk group, low/int-high	120/357	2.38 (1.24–4.58)	0.009	1.48 (0.95–2.29)	0.080	2.59 (1.30–5.18)	0.007	1.19 (0.74–1.89)	0.475

Abbreviations: CI, confidence interval; Int, intermediate; OR, odds ratio; XRCC6, X-ray repair complementing defective repair in Chinese hamster cells 6. Statistical test: binary logistic regression (Reference category: CC).

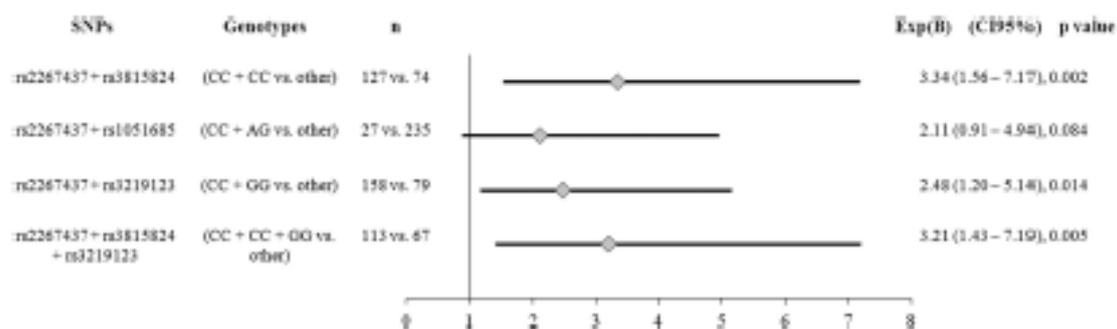


Figure 1. Forrest plot of odds ratios (ORs) with 95% confidence interval (CI) for polymorphisms rs2267437 (*XRCC6*), rs3815824 (*MVP*), rs1051685 (*XRCC5*) and rs3219123 (*PARP1*) analyzed in combination and D'Amico risk recurrence group. Each diamond represents the OR, and the horizontal line indicates the 95% CI. For the binary logistic regression, patients were dichotomized into two groups as follows: low vs intermediate-high D'Amico groups. SNP, single-nucleotide polymorphism.

rs2267437 and rs1051685 were analyzed together we did not observe such association, possibly due to the small amount of patients carrying the 'protective' genotype ($n=27$). The minor allele frequency for rs1051685 was around 10%, and this result agrees with those reported specifically in the European population (data available at <http://www.ncbi.nlm.nih.gov/projects/SNP/>). With the data available, the potential role of rs1051685 polymorphism needs to be explored deeply.

We did not find previously published studies aimed to explore the association between genetic variants in *PARP1* and PCa. Other SNPs in *PARP1* seem to be associated with glioma⁴⁰ and gastrointestinal cancers,⁴¹ but nothing is reported in relation to rs3219123 or PCa. In our series, we observed an association between this SNP and diagnostic PSA levels, but univariate analysis was not statistically significant. Even when rs3219123 was analyzed in combination with rs2267437, the results were less significant than those obtained with rs2267437 alone, suggesting a poor role of (*PARP1*) rs3219123 in the biology of PCa.

There is only one previous publication studying the role of genetic variants in *MVP*. It is an article aimed to explore the role of SNPs in multidrug resistant genes in patients with mesial temporal lobe epilepsy with hippocampal sclerosis.⁴² The authors did not observe any role of different SNPs in *MVP*—including rs3815824—in this type of patients. In that sense, *MVP* is an unknown gene which interest has increased as its association with Ku70/Ku80 (ref. 16) and its role in radiation response^{43–45} in the clinical setting. We reported here for the first time that PCa patients carrying the TT genotype were at higher risk of developing worst prostate tumors in terms of diagnostic PSA level (a classical prognosis factor for this type of cancer). Moreover, when combined with (*XRCC6*) rs2267437, these SNPs together increased the risk of a higher D'Amico score at diagnosis. The mechanisms behind this association are unknown, but support the existing hypothesis about a potential link between *MVP* and NHEJ.¹⁶

The functional consequences of the SNPs reported in the present study are difficult to address as rs2267437 and rs3815824 are intron variations which are not transduced into protein. However, introns are removed and exons are joined during a complex post-transcriptional modification of RNA named splicing, and the consequences of SNPs located in intron regions on cis regulatory elements which drive the splicing process are unknown but would influence the spliceosome.⁴⁶ Moreover, SNPs located in non-coding regions have been associated with human diseases and in any case, are needed to a better understanding of their impact on gene function and health of an individual.⁴⁷ Finally, SNPs located in intron regions are considered in *in silico* tools such as 'F-SNP: a collection of functional SNPs, specifically prioritized for disease association studies' (available at:

<http://compbio.cs.queensu.ca/F-SNP/>). This fact suggests the biological relevance of this type of genetic variations in human diseases.^{48,49} In that sense, functional studies are required to disclose the biological consequences of rs2267437 and rs3815824 in PCa.

The present study provides a number of advantages that contribute to their credibility. First, the high confidence of the genotyping system and the clinical end points considered (of total routine in daily clinical practice) minimize the possibility of inferring results because of the subjectivity of the observations. Second, the present study has been performed in a series of Spanish PCa patients who was homogeneous regarding to intra-ethnic variability, a factor that has been shown to be important in association studies with polymorphisms. And third, all the determinations (13 222 in total) were performed with the same methodology (OpenArray, Applied Biosystems), with the same batch of chips and by the same investigator, thus minimizing biases from technical origin. On the other hand, the present study has some weaknesses that need to be highlighted: (i) although 601 PCa patients seem sufficient to obtain statistically reliable results, it is possible that some results may be of stochastic nature (type I error), especially for those SNPs with lower minor allele frequency and taken into account that the association studies were performed in a series of 494 PCa patients; (ii) other factors associated with PCa (that is, age, familiar aggregation, toxic habits, prostate volume, testosterone level, other DNA repair genes or some kind of diets) have not been taken into account in the present study; (iii) in the present study, we did not account for multiple testing (that is, Bonferroni correction), thus limiting the strength of some statistical analyses; and (iv) functional studies to disclose the biological consequences of the SNPs are required to understand the real role of these genetic variations in PCa.

Taken together, our results suggest that genetic variants in DNA DSB repair genes and other genes close related to the NHEJ could be of relevance in PCa. An inability of cells to respond properly to, or to repair, DNA damage leads to genetic instability, which in turn may enhance the rate of cancer development or condition the aggressiveness of the tumor. Anyhow, to better understand the biological consequences of these genetic variations, other genes involved in DNA repair (such as *p53*, *ATM*, *BRCA1* or *BRCA2*) have to be taken into account, especially when there are accumulating evidence connecting deficiencies in cellular responses to DNA DSBs with tumorigenesis.²

CONCLUSIONS

In summary, our study shows for the first time that (*XRCC6*) rs2267437 and (*MVP*) rs3815824 polymorphisms are associated with clinical variables of more aggressive PCa. The present results

have to be considered as hypothesis generating, and additional studies with a larger separate cohort of men should be done to validate the results as well as their combined effects. Moreover, functional studies are required to understand the role of both SNPs in PCA.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the technical support from the Immunology Department (Hospital Universitario de Gran Canaria Dr Negrín) staff: Nereida González-Quevedo and Yareta Florido-Ortega. This work was subsidized by grants from the Instituto de Salud Carlos III (Ministerio de Economía y Competitividad from Spain), ID: P12/01867 and P13/00412; and Instituto Canario de Investigación del Cáncer (ICIC-GR-F-1411). Almudena Valenciano has a grant from the Instituto Canario de Investigación del Cáncer (ICIC). The founders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

REFERENCES

1. Fraser M, Berlin A, Bristow RG, van der Kwast T. Genomic pathological, and clinical heterogeneity as drivers of personalized medicine in prostate cancer. *Urol Oncol* 2015; **33**: 85–94.
2. Khanna KK, Jackson SP. DNA double-strand breaks: signaling repair and the cancer connection. *Nat Genet* 2001; **27**: 247–254.
3. Sharpless NE, Ferguson DO, O'Hagan RC, Castellon DH, Lee C, Fanzal PA et al. Impaired nonhomologous end-joining provokes soft tissue sarcomas harboring chromosomal translocations, amplifications, and deletions. *Mol Cell* 2001; **8**: 1187–1196.
4. Shivastav M, De Hao LP, Nickliff JA. Regulation of DNA double-strand break repair pathway choice. *Cell Res* 2008; **18**: 134–147.
5. Henríquez-Hernández LA, Valenciano A, Foro-Amalot P, Alvarez-Cubero MJ, Cozar JM, Suarez-Navo JF et al. Single nucleotide polymorphisms in DNA repair genes as risk factors associated to prostate cancer progression. *BMC Med Genet* 2014; **15**: 143.
6. Wiest SC. Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* 2003; **4**: 435–445.
7. Strachan T, Read AP. Human Molecular Genetics. BIOS Scientific Publishers; Wiley-Liss Oxford, New York, 1996, p xv, 597 pp.
8. Tsang RC, Hsieh FJ, Shih CM, Hsu HS, Chen CY, Wang YC. Lung cancer susceptibility and prognosis associated with polymorphisms in the nonhomologous end-joining pathway genes: a multiple genotype-phenotype study. *Cancer* 2008; **115**: 2930–2948.
9. Basling CH, Alt FW. The cellular response to general and programmed DNA double strand breaks. *DNA Repair (Amst)* 2004; **3**: 781–796.
10. Kedesha NL, Heuser JE, Chugani DC, Rome LH. Vault II. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry. *J Cell Biol* 1991; **112**: 225–235.
11. Schaffer GL, Wijngaard PL, Flans MJ, Izquierdo MA, Slovak ML, Pinedo HM et al. The drug resistance-related protein LRP is the human major vault protein. *Nat Med* 1995; **1**: 578–582.
12. Shimamoto Y, Sumitawa T, Haraguchi M, Gotanda T, Jueng HC, Furukawa T et al. Direct activation of the human major vault protein gene by DNA-damaging agents. *Oncol Rep* 2006; **15**: 645–652.
13. Lara PC, Pruschy M, Zimmermann M, Henríquez-Hernández LA. MVP and vaults: a role in the radiation response. *Radiat Oncol* 2011; **6**: 148.
14. Chung JH, Ginn-Weise ME, Eng C. Phosphatase and tensin homologue deleted on chromosome 10 (PTEN) has nuclear localization signal-like sequences for nuclear import mediated by major vault protein. *Cancer Res* 2005; **65**: 4108–4116.
15. Ryu SJ, An HI, Oh YS, Choi HR, Ha MK, Park SC. On the role of major vault protein in the resistance of senescent human diploid fibroblasts to apoptosis. *Cell Death Differ* 2008; **15**: 1673–1680.
16. Lloret M, Lara PC, Bordon E, Fontes F, Riera A, Pinar B et al. Major vault protein may affect nonhomologous end-joining repair and apoptosis through Ku70/80 and bax downregulation in cervical carcinoma tumors. *Int J Radiat Oncol Biol Phys* 2009; **73**: 976–979.
17. Thomas C, Tullin AV. Poly-ADP-ribose polymerase: machinery for nuclear processes. *Mol Aspects Med* 2013; **34**: 1124–1137.
18. Schultz N, Lopez E, Saleh-Gohari N, Helleday T. Poly(ADP-ribose) polymerase (PARP-1) has a controlling role in homologous recombination. *Nucleic Acids Res* 2003; **31**: 4959–4964.
19. Wang M, Wu W, Resnik B, Zhang L, Wang H, Bakis G. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* 2006; **34**: 6170–6182.
20. Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergal-Roelvink M et al. Inhibition of poly (ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* 2009; **361**: 123–134.
21. Zhang J. Poly (ADP-ribose) polymerase inhibitor: an evolving paradigm in the treatment of prostate cancer. *Asian J Androl* 2014; **16**: 401–406.
22. Ferrer M, Suarez JF, Guedea F, Fernández P, Macías V, Marín et al. Health-related quality of life 2 years after treatment with radical prostatectomy, prostate brachytherapy, or external beam radiotherapy in patients with clinically localized prostate cancer. *Int J Radiat Oncol Biol Phys* 2008; **72**: 421–432.
23. Henríquez-Hernández LA, Valenciano A, Foro-Amalot P, Alvarez-Cubero MJ, Cozar JM, Suarez-Navo JF et al. Intraethnic variation in steroid-5-alpha-reductase polymorphisms in prostate cancer patients: a potential factor implicated in 5-alpha-reductase inhibitor treatment. *J Genet* 2015; **94**: 335–341.
24. Henríquez-Hernández LA, Valenciano A, Foro-Amalot P, Alvarez-Cubero MJ, Cozar JM, Suarez-Navo JF et al. Polymorphisms in DNA-repair genes in a cohort of prostate cancer patients from different areas in Spain: heterogeneity between populations as a confounding factor in association studies. *PLoS One* 2013; **8**: e69735.
25. Epstein JI, Allbrook WC Jr, Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostate Carcinoma. *Am J Surg Pathol* 2005; **29**: 1228–1242.
26. D'Amico AV, Whittington R, Malkowicz SB, Wu YH, Chen MH, Hurwitz M et al. Utilizing predictions of early prostate-specific antigen failure to optimize patient selection for adjuvant systemic therapy trials. *J Clin Oncol* 2000; **18**: 3240–3246.
27. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
28. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenfeld B et al. The structure of haplotype blocks in the human genome. *Science* 2002; **296**: 2225–2229.
29. Henríquez-Hernández LA, Valenciano A, Herrera-Ramos E, Lloret M, Rivera-Perez A, Lara PC. High-throughput genotyping system as a robust and useful tool in oncology: experience from a single institution. *BioRxiv* 2013; **41**: 424–429.
30. Morcillo-Suarez C, Alegre J, Sangro R, Garza E, de Cid R, Milne R et al. SNP analysis to results (SNPto): a web-based environment oriented to statistical genomics analyses upon SNP data. *Bioinformatics* 2008; **24**: 1643–1644.
31. Clayton D. Population association. Balding DJ, Bishop M, Grinnings C (eds). *Handbook of Statistical Genetics*. Wiley-Chichester, NY, 2001, pp 519–540.
32. Rando JC, Cabrera VM, Larruga JM, Hernandez M, González AM, Pinto F et al. Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann Hum Genet* 1999; **63**: 413–428.
33. Calafell F, Bertoniotti J. Principal component analysis of gene frequencies and the origin of Basques. *Am J Phys Anthropol* 1994; **93**: 201–215.
34. Gomes BC, Silva SN, Azevedo AP, Marita I, Gil OM, Faria TC et al. The role of common variants of non-homologous end-joining repair genes XRCC4, LIG4 and Ku80 in thyroid cancer risk. *Oncol Rep* 2010; **24**: 1079–1085.
35. Rivera-Calzada A, Spagnolo L, Peat LH, Libeca O. Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO Rep* 2007; **8**: 56–62.
36. Jia J, Ren J, Yan D, Xiao L, Sun R. Association between the XRCC6 polymorphisms and cancer risks: a systematic review and meta-analysis. *Medicine (Baltimore)* 2015; **94**: e283.
37. Xu H, Zou P, Chen P, Zhao L, Zhao P, Lu A. Association between the XRCC6 Promoter rs2267437 polymorphism and cancer risk: evidence based on the current literature. *Genet Test Mol Biomarkers* 2013; **17**: 607–614.
38. Hayden PJ, Tewari P, Morris DW, Staines A, Crowley D, Neters et al. Variation in DNA repair genes XRCC3, XRCC4, XRCC5 and susceptibility to myeloma. *Hum Mol Genet* 2007; **16**: 3117–3127.
39. Cibella MT, de Lencastre CF, Navarro A, Diaz T, Fuster D, Tovar N et al. Impact on response and survival of DNA repair single nucleotide polymorphisms in relapsed or refractory multiple myeloma patients treated with thalidomide. *Leuk Res* 2011; **35**: 1178–1183.
40. Adif Falmidih M, Schwartzbaum J, Frumento P, Feychting M. Association between DNA repair gene polymorphisms and risk of glioma: a systematic review and meta-analysis. *Neuro Oncol* 2014; **16**: 807–814.
41. Hu Y, Zhou M, Li K, Zhang K, Kang X, Zhang Y et al. Two DNA repair gene polymorphisms on the risk of gastrointestinal cancers: a meta-analysis. *Tumour Biol* 2014; **35**: 1715–1725.
42. Balan S, Radhab SK, Sathyan S, Vijal J, Banerjee M, Radhakrishnan K. Major vault protein (MVP) gene polymorphisms and drug resistance in mesial temporal lobe epilepsy with hippocampal sclerosis. *Gene* 2013; **526**: 440–453.
43. Henríquez-Hernández LA, Moreno M, Rey A, Lloret M, Lara PC. MVP expression in the prediction of clinical outcome of locally advanced oral squamous cell carcinoma patients treated with radiotherapy. *Radiat Oncol* 2012; **7**: 147.

- 44 Henríquez-Hernández LA, Libret M, Pinar B, Bordon E, Rey A, Lubrano et al. BCL-2, in combination with MVP and IGF-1R expression, improves prediction of clinical outcome in complete response cervical carcinoma patients treated by radiochemotherapy. *Gynecol Oncol* 2011; **122**: 585–589.
- 45 Lloret M, Lana PC, Bordon E, Rey A, Falcon O, Apolinario RM et al. MVP expression is related to IGF1-R in cervical carcinoma patients treated by radiochemotherapy. *Gynecol Oncol* 2008; **110**: 304–307.
- 46 Jensen CJ, Oldfield BJ, Rubio JP. Splicing, cis genetic variation and disease. *Biochem Soc Trans* 2009; **37**: 1311–1315.
- 47 Shastry BS. SNPs impact on gene function and phenotype. *Methods Mol Biol* 2009; **578**: 3–22.
- 48 Lee PH, Shatkay H. Ranking single nucleotide polymorphisms by potential deleterious effects. *AMIA Annu Symp Proc* 2008, 667–671.
- 49 Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res* 2008; **36**: D820–D824.

Supplementary Information accompanies the paper on the Prostate Cancer and Prostatic Diseases website (<http://www.nature.com/pcan>)



Original article

Genetic variations in genes involved in testosterone metabolism are associated with prostate cancer progression: A Spanish multicenter study

Luis Alberto Henríquez-Hernández, Ph.D.^{a,h,c,+}, Almudena Valenciano, B.D.^b,
 Palmira Foro-Arnalot, Ph.D., M.D.^d, María Jesús Álvarez-Cubero, Ph.D.^{e,f},
 José Manuel Cozar, Ph.D., M.D.^g, José Francisco Suárez-Novo, M.D.^h,
 Manel Castells-Esteve, M.D.^h, Pablo Fernández-Gonzalo, M.D.ⁱ,
 Belén De-Paula-Carranza, M.D.ⁱ, Montse Ferrer, Ph.D., M.D.^j, Ferrán Guedea, Ph.D., M.D.^k,
 Gemma Sancho-Pardo, Ph.D., M.D.^l, Jordi Craven-Bartle, Ph.D., M.D.^l,
 María José Ortiz-Gordillo, Ph.D., M.D.^m, Patricia Cabrera-Roldán, Ph.D., M.D.^m,
 Juan Ignacio Rodríguez-Melcón, M.D.^{a,b}, Estefanía Herrera-Ramos, B.D.^{n,o},
 Carlos Rodríguez-Gallego, Ph.D.ⁿ, Pedro C. Lara, Ph.D., M.D.^{a,b,c}

^a Radiation Oncology Department, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain

^b Instituto Canario de Investigación del Cáncer, Las Palmas, Spain

^c Clinical Sciences Department, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain

^d Institut d'Oncologia Radioteràpica, Hospital de la Esperança, Parc de Salut Mar, Barcelona, Spain

^e Laboratory of Genetic Identification, Legal Medicine and Toxicology Department, Facultad de Medicina, Universidad de Granada, Granada, Spain

^f GENYO, Pfizer-University of Granada-Andalusian Government Centre for Genomics and Oncological Research, Granada, Spain

^g Department of Urology, Hospital Universitario Virgen de las Nieves, Granada, Spain

^h Department of Urology, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Spain

ⁱ Radiation Oncology Department, Onkologikoa, Galdakao, Spain

^j Health Services Research Group, Institut de Recerca Hospital del Mar (IMIM), Barcelona, Spain

^k Department of Radiation Oncology, Institut Català d'Oncologia (ICO), L'Hospitalet de Llobregat, Spain

^l Radiation Oncology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

^m Radiation Oncology Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain

ⁿ Department of Immunology, Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas, Spain

^o Department of Medical and Surgical Sciences, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain

Received 25 February 2015; received in revised form 20 March 2015; accepted 8 April 2015

Abstract

Background: Prostate cancer (PCa) is an androgen-dependent disease. Nonetheless, the role of single nucleotide polymorphisms (SNPs) in genes encoding androgen metabolism remains an unexplored area.

Purpose: To investigate the role of germline variations in cytochrome P450 17A1 (*CYP17A1*) and steroid-5 α -reductase, α -polypeptides 1 and 2 (*SRD5A1* and *SRD5A2*) genes in PCa.

Patients and methods: In total, 494 consecutive Spanish patients diagnosed with nonmetastatic localized PCa were included in this multicenter study and were genotyped for 32 SNPs in *SRD5A1*, *SRD5A2*, and *CYP17A1* genes using a Biotrove OpenArray NT Cycler. Clinical data were available. Genotypic and allelic frequencies, as well as haplotype analyses, were determined using the web-based environment SNPator. All additional statistical analyses comparing clinical data and SNPs were performed using PASW Statistics 15.

This work was subsidized by Grants from the Instituto de Salud Carlos III, Madrid, Spain (Ministerio de Economía y Competitividad, Spain), ID: PI12/01867 and PI13/00412; and Instituto Canario de Investigación del Cáncer, Canary Islands, Spain (ID: ICIC-GR-F-14/11). Almudena Valenciano received a Grant from the Instituto Canario de Investigación del Cáncer (ICIC).

* Corresponding author. Tel.: +34-92-845-0927; fax: +34-92-844-9805.

E-mail address: luis.henriquez@ulpgc.es (L.A. Henríquez-Hernández).

<http://dx.doi.org/10.1016/j.urolonc.2015.04.003>

1078-1439/© 2015 Elsevier Inc. All rights reserved.

Results: The call rate obtained (determined as the percentage of successful determinations) was 97.3% of detection. A total of 2 SNPs in *SRD5A1*—rs3822430 and rs1691053—were associated with prostate-specific antigen level at diagnosis. Moreover, G carriers for both SNPs were at higher risk of presenting initial prostate-specific antigen levels >20 ng/ml (Exp(B) = 2.812, 95% CI: 1.397–5.657, $P = 0.004$) than those who are AA-AA carriers. Haplotype analyses showed that patients with PCa nonhomozygous for the haplotype GCTTG TAGTA were at an elevated risk of presenting bigger clinical tumor size (Exp(B) = 3.823, 95% CI: 1.280–11.416, $P = 0.016$), and higher Gleason score (Exp(B) = 2.808, 95% CI: 1.134–6.953, $P = 0.026$).

Conclusions: SNPs in *SRD5A1* seem to affect the clinical characteristics of Spanish patients with PCa. © 2015 Elsevier Inc. All rights reserved.

Keywords: SNP; *SRD5A1*; *SRD5A2*; *CYP17A1*; Prostate cancer; OpenArray

1. Introduction

Prostate cancer (PCa) is an androgen-dependent disease [1]. The synthesis of testosterone is mediated by the enzyme 17-hydroxylase/17,20-lyase, encoded by *CYP17A1* gene. Dihydrotestosterone (DHT) is considered the active metabolite of testosterone. Testosterone is converted into DHT through the 5 α -reductase pathway, thus conditioning tumor availability of the hormone and subsequent tumor development and progression [2].

Single nucleotide polymorphisms (SNPs) are simple genetic variations in the DNA sequence, which modify the efficacy of encoded enzymes [3]. Genetic alterations in the testosterone metabolism pathway are expected to alter hormonal homeostasis and likely influence PCa development and progression.

The association between *CYP17A1* gene polymorphisms and PCa is controversial [4–6]. Although it has been suggested that some polymorphisms in *CYP17A1* are associated with PCa risk [7] and survival [8], it is not clear how SNPs in this gene affect clinical variables of PCa such as tumor size, prostate-specific antigen (PSA) levels, or Gleason score. In any case, *CYP17A1* is considered an important factor for PCa progression and has become a new therapeutic target. Thus, abiraterone acetate is a new drug very active in the treatment of certain PCa through specific *CYP17A1* blockade [9], although the role of SNPs in *CYP17A1* in response to abiraterone acetate has not been explored.

The 5 α -reductase enzymes are encoded by the steroid-5 α -reductase, α -polypeptide 1 (*SRD5A1*) and steroid-5 α -reductase, α -polypeptide 2 (*SRD5A2*) genes. The role of *SRD5A1* and *SRD5A2* polymorphisms in PCa has been little studied, although it has been shown positive associations of several *SRD5A1* and *SRD5A2* variations as independent predictors of PCa outcome in terms of biochemical recurrence [10]. Anyhow, *SRD5A1/A2* are also therapeutic targets, and the specific blockade of these enzymes with dutasteride or finasteride has shown to decrease the production of DHT using 5 α -reductase inhibitors, thus, decreasing the incidence of clinically localized PCa [11,12].

At a clinical level, tumor size, Gleason score, and pretreatment serum level of PSA are the most important prognostic factors [13]. Nonetheless, although these are key

factors that affect the outcome, the heterogeneity in clinical behavior requires the expansion of knowledge about the disease [14].

As the presence of genetic variations in the androgen biosynthesis and metabolism genes may alter hormone bioavailability, we hypothesized here that certain polymorphisms in the testosterone synthesis pathway may be important to define the biological characteristics of the disease, possibly influencing the classic prognostic factors of PCa. For the first time, we conducted a study to explore the association between germline variations in *SRD5A1/2* and *CYP17A1* and PCa progression in a cohort of Spanish white patients with PCa.

2. Materials and methods

2.1. Patients

A total of 601 patients with nonmetastatic localized PCa were initially included in the study. The patients were recruited from 4 different regions of Spain, as has been previously published [15]. As genotypic and allelic frequencies vary among subjects from these regions, and Andalusian patients showed the greatest differences [15,16], we excluded these subset of patients from further analyses to avoid bias. Patients who were initially operated were also excluded to form a homogenous group of patients with PCa. Thus, 494 subjects were included in the present study. All patients were of Spanish origin and provided written informed consent before blood sample collection. The study was approved by the Research and Ethics Committee of each institution that participated in the study: Hospital Universitario de Gran Canaria Dr. Negrín (Las Palmas de Gran Canaria), Hospital de la Esperanza. Parc de Salut Mar (Barcelona), Hospital Universitari de Bellvitge (L'Hospitalet de Llobregat), Orkologikoa (Guipuzcoa), Institut Català d'Oncologia (L'Hospitalet de Llobregat), and Hospital de la Santa Creu i Sant Pau (Barcelona).

Clinical tumor size (cT), initial PSA value, and Gleason score were analyzed for all PCa. cT was assessed by digital rectal examination, followed by transrectal ultrasonography and magnetic resonance imaging; PSA serum levels were assessed by chemiluminescence in an Architect i2000

analyzer (Abbott Laboratories, IL); and Gleason score was determined in the biopsy specimen by a pathologist. After collecting demographic and clinical data, a blood sample was taken after signature on the written informed consent form.

2.2. DNA isolation and quantification

All the blood samples were sent to the Hospital Universitario de Gran Canaria Dr. Negrín for DNA extraction and genotyping analyses. DNA was isolated from 300 µl of whole blood in an iPrepPurification Instrument using the iPrep PureLink gDNA Blood Kit (Invitrogen, by Life Technologies, Carlsbad, CA). DNA integrity was determined by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE).

2.3. Selection criteria for SNPs

SNPs were selected using data of individuals with European ancestry (CEU - Utah residents with ancestry from Northern and Western Europe) from the HapMap Project (available at: www.hapmap.org). Pairwise linkage disequilibrium tagging was achieved with Haploview v4.2 software (free downloaded from <http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>) [17]. This strategy is based on the haplotype blocks formed by clustered SNPs [18]. By selecting a modest number of SNPs within a block, it is possible to recover up to 95% of the heterozygosity present at a locus. Haploview tool allows to select tagged SNPs based on haplotype blocks identified using the existing genotype data from the HapMap Project. The R^2 given by Haploview was >0.95 for all genes. Thus, 32 SNPs located in 3 genes (*SRD5A1*, 10 SNPs; *SRD5A2*, 12 SNPs; and *CYP17A1*, 10 SNPs) were studied (Table 1).

2.4. Genotyping

The SNP genotyping was performed in a Biotrove OpenArray NT Cycler (Applied Biosystems, Foster City, CA) [19]. DNA samples loaded in OpenArray had A260/A280 and A260/230 ratios of 1.7 to 1.9 and were adjusted to 50 ng/µl. Overall, 300 ng of genomic DNA was used. A final amount of 150 ng was incorporated into the array with the autoloader and genotyped according to the manufacturer's recommendations. A nontemplate control consisting of DNase-free double-distilled water was introduced within each assay. When the DNA and master mix were transferred, the loaded OpenArray plate was filled with an immersion fluid and sealed with glue. The multiplex TaqMan assay reactions were carried out in a Dual Flat Block (384-well) GeneAmp PCR System 9700 (Applied Biosystems) with the following PCR cycle: an initial step at 93°C for 10 minutes followed by 50 cycles of 45 seconds

Table 1
Genotypic and allelic frequencies among Spanish patients with prostate cancer

Gene/SNP	n	Genotypic frequencies						Allelic frequencies					
<i>SRD5A1</i>													
rs166050	479	AA	0.59	AG	0.36	GG	0.05	A	0.77	G	0.23		
rs501999	488	CC	0.28	CT	0.47	TT	0.24	C	0.52	T	0.48		
rs518673	476	AA	0.09	AG	0.41	GG	0.50	A	0.29	G	0.71		
rs3822430	481	AA	0.33	AG	0.51	GG	0.16	A	0.59	G	0.41		
rs500182	489	GG	0.01	GT	0.18	TT	0.82	G	0.09	T	0.91		
rs8192120	483	AA	0.10	AC	0.45	CC	0.46	A	0.32	C	0.68		
rs4702378	490	CC	0.06	CT	0.31	TT	0.63	C	0.21	T	0.79		
rs1691053	488	AA	0.87	AG	0.13	GG	0.00	A	0.93	G	0.07		
rs39848	472	CC	0.14	CT	0.47	TT	0.39	C	0.38	T	0.62		
rs3797179	482	AA	0.04	AG	0.26	GG	0.70	A	0.17	G	0.83		
<i>SRD5A2</i>													
rs2208532	475	AA	0.30	AG	0.49	GG	0.21	A	0.54	G	0.46		
rs12470143	488	CC	0.30	CT	0.52	TT	0.17	C	0.57	T	0.43		
rs2281546	490	GG	0.02	GT	0.24	TT	0.74	G	0.14	T	0.86		
rs3754838	480	CC	0.02	CT	0.20	TT	0.78	C	0.12	T	0.88		
rs4952222	490	AA	0.00	AC	0.00	CC	1.00	A	0.00	C	1.00		
rs7562326	483	CC	0.02	CT	0.21	TT	0.77	C	0.13	T	0.87		
rs2300702	477	CC	0.18	CG	0.49	GG	0.33	C	0.42	G	0.58		
rs4952197	485	AA	0.06	AG	0.33	GG	0.61	A	0.22	G	0.78		
rs676033	487	CC	0.51	CT	0.40	TT	0.09	C	0.71	T	0.29		
rs523349	398	CC	0.40	CG	0.43	GG	0.17	C	0.62	G	0.38		
rs9332975	487	CC	0.02	CT	0.20	TT	0.78	C	0.12	T	0.88		
rs7594951	489	CC	0.78	CT	0.20	TT	0.02	C	0.88	T	0.12		
<i>CYP17A1</i>													
rs3781287	485	GG	0.21	GT	0.47	TT	0.32	G	0.45	T	0.55		
rs1004467	486	AA	0.79	AG	0.20	GG	0.02	A	0.89	G	0.11		
rs743572	465	AA	0.32	AG	0.48	GG	0.20	A	0.56	G	0.44		
rs10883782	477	AA	0.65	AG	0.32	GG	0.04	A	0.81	G	0.20		
rs619824	487	AA	0.22	AC	0.50	CC	0.28	A	0.47	C	0.53		
rs2486758	482	CC	0.04	CT	0.29	TT	0.67	C	0.18	T	0.82		
rs17115100	490	GG	0.80	GT	0.19	TT	0.01	G	0.89	T	0.11		
rs4919686	492	AA	0.48	AC	0.41	CC	0.10	A	0.69	C	0.31		
rs10786712	491	CC	0.33	CT	0.47	TT	0.20	C	0.57	T	0.43		
rs6163	474	AA	0.19	AC	0.48	CC	0.32	A	0.43	C	0.57		

at 95°C, 13 seconds at 94°C and 2:14 minutes at 53°C; followed by a final step during 2 minutes at 25°C and holding at 4°C.

The fluorescence results were read using the OpenArray SNP Genotyping Analysis software version 1.0.5. (Applied Biosystems). The genotyping analysis was made with the TaqMan Genotyper software version 1.0.1. (available at: <http://www.invitrogen.com/site/us/en/home/Global/forms/taqman-genotyper-software-download-reg.html>) using auto-calling as the call method. The quality value of the data points was determined by a threshold above 0.95.

2.5. Statistical analysis

Genotype and allelic frequencies were determined using the web-based environment SNPator (SNP Analysis To Results, from the Spain's National Genotyping Center and the National Institute for Bioinformatics) [20]. Relative

Table 2
Description of clinical variables

Clinical	n	(%)
Clinical tumor size (cT)		
cT1a–cT2a	270	54.7
cT2b–cT2c	141	25.8
cT3–cT4	66	13.4
NA	17	3.4
Initial PSA, ng/ml		
< 10	306	61.9
10–19.99	103	20.9
> 20	79	16.0
NA	6	1.2
Gleason score		
< 7	226	45.7
7	195	39.5
> 7	71	14.4
NA	2	0.4

NA = not available.

excess heterozygosity was determined to check compatibility of genotype frequencies with Hardy-Weinberg equilibrium (HWE). Thus, *P* values from the standard exact HWE lack of fit test were calculated using SNPator. Comparisons of genotypic and allelic frequencies, as well as haplotype analyses were also done in SNPator. Haplotype analysis was also performed using the SNPator tool.

All additional statistical analyses comparing clinical data and SNPs were performed using PASW Statistics 15 (IBM Corporation, Armonk, NY).

3. Results

All the genotyped samples met the aforementioned quality criteria, and all samples were genotyped with the same batch of material and at the same time. A total of 494 patients with PCa were genotyped for 32 SNPs. Of the 15,808 possible determinations, 97.33% were successfully genotyped. The genotypic and allelic frequencies are shown in Table 1. All SNPs were in HWE.

Distribution of clinical variables is detailed in Table 2. Most patients with PCa were staged cT1a–cT2a (54.7%), had PSA < 10 ng/ml (61.9%), and had Gleason score < 7 (45.7%). We did not observe different proportion of aggressiveness of PCa among different regions of Spain.

Among the 32 SNPs studied, only 2 of them (rs3822430 [minor allele frequency (MAF) = 0.41] and rs1691053 (MAF = 0.07), both located in *SRD5A1*) were significantly differently distributed among patients with PCa according to the initial PSA level (Table 3). Thus, among the 157 patients with PCa showing initial PSA levels < 10 ng/ml, 92 (58.6%) carried the genotype rs3822430—AA, whereas among the 180 PCa patients showing initial PSA levels \geq 10 ng/ml, 115 (63.9%) carried the G allele for this SNP

(chi-square test, *P* = 0.015). A similar result was observed for the SNP rs1691053 (*P* = 0.008). We explored the combined role of both SNPs in relation to the initial PSA level (Table 3), and we observed that patients with PCa who carried the G allele showed highest levels of initial PSA than those who were AA-AA carriers (73% vs. 27%, *P* = 0.002). Moreover, G carriers for SNPs rs3822430 and rs1691053 were at higher risk of presenting initial PSA levels > 20 ng/ml (Exp(*B*) = 2.812, 95% CI: 1.397–5.657, *P* = 0.004; binary logistic regression test) than those who are AA-AA carriers (data not shown). No other polymorphism showed significant associations with clinical variables.

We explored the role of the 3 most frequent haplotypes detected in our series in chromosomes (chr) 2 (encoding *SRD5A2* gene), 5 (encoding *SRD5A1* gene), and 10 (encoding *CYP17A1* gene) (Table 4). For chr 2, 70.9% of patients with PCa presented one of the 3 most frequent haplotypes; for chr 5, 54.9% of patients with PCa presented one of the 3 most frequent haplotypes; and for chr 10, 62% of patients with PCa presented one of the 3 most frequent haplotypes. We observed that among 23 patients homozygous for the haplotype GCTTGAGTA (located in chr 5), 19 (82.6%) were cT1a–cT2a, compared with the 55.4% (251 of 453) observed among patients with PCa nonhomozygous for this haplotype (chi-square test, *P* = 0.018, data not shown). A similar result was observed for Gleason score < 7 (69.6% vs. 44.9%, *P* = 0.032). Moreover, as shown in Table 5, we observed that patients with PCa nonhomozygous for the haplotype GCTTGAGTA were at higher risk of presenting bigger cT (Exp(*B*) = 3.823, 95% CI: 1.280–11.416, *P* = 0.016). In the same way, patients with PCa nonhomozygous for the haplotype GCTTGAGTA were at an elevated risk of showing higher Gleason score (Exp(*B*) = 2.808, 95% CI: 1.134–6.953, *P* = 0.026). No other haplotypes in chr 5 or in chromosomes 2 and 10 were significantly associated with the clinical variables considered in the study.

Table 3
Significant associations between initial PSA level and SNPs

SNP	Genotypes	Initial PSA level, ng/ml			<i>P</i> value ^a
		< 10	10–19.99	> 20	
rs3822430	AA	92	44	21	0.015
	AG	44	44	14	
	GG	21	50	7	
rs1691053	AA	272	89	99	0.008
	AG	30	13	18	
	GG	0	1	0	
rs3822430–rs1691053	AA + AA	80	38	10	0.002
	G carriers	213	64	66	

^aChi-square test.

Table 4
Frequency of detection (%) of the 3 most frequent haplotypes located in chromosomes 2 (*SRD5A2*), chromosome 5 (*SRD5A1*), and chromosome 10 (*CYP17A1*) in our series

	chr 2	Freq	chr 5	Freq	chr 10	Freq
Hap1	TTTGGACTCCTC	32.8	GCTTGTAGTA	21.6	AAGCAGTAGT	28.1
Hap2	TTCACGCCTCGT	21.0	ACCTATGGCA	18.2	CGGAATCCAT	17.9
Hap3	TTCGGGCTCCTC	11.1	GATTACAGTA	15.1	CAGAATCCAC	16.0

chr = Chromosome; freq = frequency; hap = haplotype.

Haplotypes in chromosome 2 are formed by the SNPs: rs9332975, rs2281546, rs12470143, rs4952197, rs2300702, rs2208532, rs7594951, rs7562326, rs4952222, rs523349, rs3754838, and rs676033.

Haplotypes in chromosome 5 are formed by the SNPs: rs518673, rs8192120, rs501999, rs500182, rs166050, rs4702378, rs3822430, rs3797179, rs39848, and rs1691053.

Haplotypes in chromosome 10 are formed by the SNPs: rs619824, rs10883782, rs17115100, rs4919686, rs1004467, rs3781287, rs10786712, rs6163, rs743572, and rs2486758.

4. Discussion

The role of 32 SNPs located in 3 key genes involved in testosterone metabolism (*SRD5A1*, *SRD5A2*, and *CYP17A1*) has been explored in a cohort of Spanish patients with PCa, on the assumption that interindividual levels of testosterone and DHT may be influenced by germline polymorphisms in those genes, and this fact could determine clinical characteristics of the tumor, a phenomenon previously observed in relation to SNPs located in genes involved in DNA repair [21].

We have observed in our series that certain SNPs in *SRD5A1* seem to be associated with clinical variables of PCa such as initial PSA levels. Thus, G carriers for SNPs rs3822430 and rs1691053 were at higher risk of presenting initial PSA levels >20 ng/ml. This finding may have an important influence in daily clinical practice. PSA levels decrease in patients with PCa treated with 5 α -reductase inhibitors [22]. Moreover, it has been suggested that DHT serum level may serve as a potential diagnostic marker of intraprostatic 5 α -reductase activity during treatment of patients with 5 α -reductase inhibitors [23]. Thus, the effect of 5 α -reductase inhibitors on PSA and DHT levels would be highly influenced by the presence of certain genotypes in *SRD5A1*. Interestingly, there are no studies exploring the role of SNPs in *SRD5A1* in relation to the response to 5 α -reductase inhibitors. It is known that specific SNPs

located in *SRD5A1* are associated with low enzymatic activity [24], but in any case, the knowledge of *SRD5A1* polymorphisms in PCa is very limited and contradictory. In a recent study, men with AG or GG vs. AA genotypes in rs1691053 were at higher risk of developing PCa [25], which supports our results and suggests a role for this specific SNP in PCa. With respect to rs3822430, it has been studied in the context of biochemical failure, with no significant associations reported [10]. The role of SNPs in *SRD5A1* in PCa seems to be reinforced by the fact that a specific haplotype in that gene (GCTTGTAGTA) was strongly associated with tumor size and Gleason score. The consequences this specific haplotype may have in the function of the enzyme have to be explored. Recently, a systematic analysis of both constitutional and somatic (PCa) variants of steroid 5 α -reductase type II indicated significant pharmacogenetic variation for both finasteride and dutasteride response and allowed to map areas of the wild-type enzyme that are responsible for the time-dependent inhibition for either (or both) enzyme inhibitor(s) [26]. SNPs are the basics of pharmacogenetics, and it is possible that individual SNPs or specific haplotypes can explain this heterogeneity. However, we did not observe any SNP or haplotype in *SRD5A2* associated with any clinical variable in our series. Nonetheless, other authors have published a role of SNPs rs4952197 or rs523349 (both included in our analyses) in biochemical recurrence [10]. In fact, rs523349 is one of the best characterized polymorphisms in *SRD5A2* gene, and it has been associated with poor prognosis in patients with PCa in terms of PSA failure [27] or extracapsular invasion [28]. Shibata et al. and Jaffe et al. published their results involving smaller cohorts of patients than our series, and Audet-Walsh et al. included subjects from different ethnic origins. We have recently published that there is an important influence of ethnics in the analysis of genetic variations [15,16]. Together with differences in the genotype determination techniques or in the strategies for SNP analyzes, these could be plausible reasons to explain the differences observed between our results and those reported by others.

Table 5
Haplotypes significantly associated with clinical variables

Haplotype	chr	Exp(B)	95% CI	P value
cT1–T2a vs. cT2b–T4				
Homozygous for GCTTGTAGTA ^a	5	1		
Other haplotypes		3.823	(1.280–11.416)	0.016
Gleason score (<7 vs. \geq 7)				
Homozygous for GCTTGTAGTA ^a	5	1		
Other haplotypes		2.808	(1.134–6.953)	0.026

chr = Chromosome.

^aReference category (binary logistic regression test).

In our series, none of the polymorphisms in *CYP17A1* were statistically associated with the clinical variables considered. An association between the risk of disease development and SNPs rs743572, rs619824, and rs2486758 has been reported [5,7]. However, the results are contradictory and conditioned by the ethnic group [2]. An association with PCa survival has been recently reported for rs10883783 polymorphism after a median follow-up of 13.2 years [8]. The role of this gene in PCa appears to be limited to development, thereby, not altering the biological characteristics of the tumor once it has occurred. This hypothesis seems to be reinforced by the fact that abiraterone acetate shows the best results when the testicular production of testosterone is inhibited by hormonal treatment, otherwise, the specific inhibition of *CYP17A1* is not fruitful [27].

Functional polymorphisms in *SRD5A1* gene may affect the production rate of dehydrotestosterone and the subsequent local exposure to androgens of androgen-responsive cells of the prostate, thus influencing biological characteristics of the tumor assessable by tumor size, Gleason score, or diagnostic PSA level. How these SNPs and haplotype may influence the function of the enzyme and the levels of testosterone is a question that must be answered. In any case, our results should be considered at a clinical level and would influence the therapeutic management of patients with PCa, particularly in relation to hormone treatment. Thus, patients with PCa who carry certain polymorphisms or haplotypes associated with higher levels of diagnostic PSA (such as G carriers for SNPs rs3822430 and rs1691053) should have a proper hormone-deprivation treatment to ensure the control of PSA level and minimize the risk of early biochemical relapse.

The present study has some weaknesses that need to be highlighted: (1) although 494 patients with PCa seem sufficient to obtain statistically reliable results, it is possible that some results may be of stochastic nature, especially for those SNPs with lower MAF; (2) other factors associated with PCa (i.e., age, familiar aggregation, toxic habits, or some kind of diets) have not been taken into account in the present study; (3) as it is a multicenter study, clinical end points were obtained by different physicians (including pathologists, urologists, and radiation oncologists), and there may be an interobserver bias in the determination of any of these parameters; and (4) observations should be confirmed in an independent cohort of patients including pathology data from patients who have been treated with radical prostatectomy. However, the present study provides a number of advantages that contribute to their credibility. First, the high confidence of the genotyping system and the clinical end points considered (of total routine in daily clinical practice) minimize the possibility of inferring results owing to the subjectivity of the observations. Second, the present study has been performed in a series of Spanish patients with PCa, which was homogeneous regarding ethnicity, a factor that has been shown to be

important in association studies with polymorphisms. Lastly, all the determinations (15,808 in total) were performed with the same methodology (OpenArray, Applied Biosystems), with the same batch of chips and by the same investigator, thus minimizing biases from technical origin.

5. Conclusions

This study reveals the importance of SNPs in relation to the biological characteristics of PCa at diagnosis, defining novel actions of *SRD5A1* gene in a cohort of Spanish patients with PCa. Future experiments are warranted to explore the role of the genotypes in the follow-up and prognosis.

Acknowledgment

We thank the Immunology Department (Hospital Universitario de Gran Canaria Dr. Negrín) staff for technical assistance: Nereida González-Quevedo and Yanira Florido-Ortega.

References

- [1] Haggins C. Endocrine-induced regression of cancers. *Cancer Res* 1967;11:1925–30.
- [2] Hamada A, Danesi R, Price DK, et al. Association of a *CYP17* polymorphism with overall survival in Caucasian patients with androgen-independent prostate cancer. *Urology* 2007;2:217–20.
- [3] Roses AD. Pharmacogenetics and the practice of medicine. *Nature* 2000;6788:857–65.
- [4] Setlwan VW, Schumacher FR, Haiman CA, et al. *CYP17* genetic variation and risk of breast and prostate cancer from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium (BPC3). *Cancer Epidemiol Biomarkers Prev* 2007;11:2237–46.
- [5] Mononen N, Schleutker J. Polymorphisms in genes involved in androgen pathways as risk factors for prostate cancer. *J Urol* 2009;4:1541–9.
- [6] Lindstrom S, Adami HO, Balter KA, et al. Inherited variation in hormone-regulating genes and prostate cancer survival. *Clin Cancer Res* 2007;17:5156–61.
- [7] Wang F, Zou YF, Feng XL, et al. *CYP17* gene polymorphisms and prostate cancer risk: a meta-analysis based on 38 independent studies. *Prostate* 2011;11:1167–77.
- [8] Wright JL, Kwon EM, Lin DW, et al. *CYP17* polymorphisms and prostate cancer outcomes. *Prostate* 2010;10:1094–101.
- [9] Yang LP. Abiraterone acetate: in metastatic castration-resistant prostate cancer. *Drugs* 2011;15:2067–77.
- [10] Audet-Walsh E, Bellemare J, Nadeau G, et al. *SRD5A* polymorphisms and biochemical failure after radical prostatectomy. *Eur Urol* 2011;6:1226–34.
- [11] Andriole G, Bostwick D, Brawley O, et al. Chemoprevention of prostate cancer in men at high risk: rationale and design of the reduction by dutasteride of prostate cancer events (REDUCE) trial. *J Urol* 2004;4(Pt 1):1334–7.
- [12] Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med* 2003;32:15–24.
- [13] D'Amico AV, Whittington R, Malkowicz SB, et al. Utilizing predictions of early prostate-specific antigen failure to optimize

- patient selection for adjuvant systemic therapy trials. *J Clin Oncol* 2000;18:3240–6.
- [14] Hoffman RM. Randomized trial results did not resolve controversies surrounding prostate cancer screening. *Curr Opin Urol* 2010;3:189–193.
- [15] Henríquez-Hernández LA, Valenciano A, Fozo-Amalot P, et al. Polymorphisms in DNA-repair genes in a cohort of prostate cancer patients from different areas in Spain: heterogeneity between populations as a confounding factor in association studies. *PLoS One* 2013;7:e69735.
- [16] Henríquez-Hernández LA, Valenciano A, Fozo-Amalot P, et al. Intra-ethnic variation in steroid-5-alpha-reductase polymorphisms in prostate cancer patients: a potential factor implicated in 5-alpha-reductase inhibitor treatment. *J Genet* 2014;[in press].
- [17] Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5.
- [18] Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–9.
- [19] Henríquez-Hernández LA, Valenciano A, Herrera-Ramos E, et al. High-throughput genotyping system as a robust and useful tool in oncology: experience from a single institution. *Biologicals* 2013;6:424–9.
- [20] Morcillo-Suarez C, Alegre J, Sangros R, et al. SNP analysis to results (SNPator): a web-based environment oriented to statistical genomics analyses upon SNP data. *Bioinformatics* 2008;14:1643–4.
- [21] Henríquez-Hernández L, Valenciano A, Fozo-Amalot P, et al. Single nucleotide polymorphisms in DNA repair genes as risk factors associated to prostate cancer progression. *BMC Med Genet* 2014;15:143.
- [22] Ross AE, Feng Z, Pierozzi PM, et al. Effect of treatment with 5-alpha reductase inhibitors on progression in monitored men with favourable-risk prostate cancer. *BJU Int* 2012;5:651–7.
- [23] Stanczyk FZ, Azen CG, Pike MC. Effect of finasteride on serum levels of androstenedione, testosterone and their 5alpha-reduced metabolites in men at risk for prostate cancer. *J Steroid Biochem Mol Biol* 2013;10–6.
- [24] Signorelli SS, Barresi V, Musso N, et al. Polymorphisms of steroid 5-alpha-reductase type I (SRD5A1) gene are associated to peripheral arterial disease. *J Endocrinol Invest* 2008;12:1092–7.
- [25] Setur SR, Chen CX, Hossain RR, et al. Genetic variation of genes involved in dihydrotestosterone metabolism and the risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2010;1:229–39.
- [26] Makridakis N, Reichardt JK. Pharmacogenetic analysis of human steroid 5 alpha reductase type II: comparison of finasteride and dutasteride. *J Mol Endocrinol* 2005;3:617–23.
- [27] Shibata A, Garcia MI, Cheng I, et al. Polymorphisms in the androgen receptor and type II 5 alpha-reductase genes and prostate cancer prognosis. *Prostate* 2002;4:269–78.
- [28] Jaffe JM, Malkowicz SB, Walker AH, et al. Association of SRD5A2 genotype and pathological characteristics of prostate tumors. *Cancer Res* 2000;6:1626–30.

