



Análisis de Genes Candidatos en Cáncer de Mama

Setiembre 2014

Lucía Brignoni Benzo

Orientador: Dr. Bernardo Bertoni

Departamento de Genética

Facultad de Medicina

Agradecimientos

A Bernardo, por recibirme en el laboratorio y ayudarme durante la realización de esta tesis, por la confianza y el apoyo diarios en el laboratorio. Por compartir también nuestros gustos en temas de gastronomía y ciencia ficción.

A mis compañeros de la isla, Moni, Sole, Vale, Sil, Tati y George, por hacer de nuestro lugar de trabajo un lugar donde da gusto venir todos los días, y por poder compartir también excelentes momentos fuera del mismo.

A todos los integrantes del Departamento de Genética de Facultad de Medicina, donde terminé de descubrir mi vocación por la enseñanza. A todos los alumnos que hacen que valga la pena el esfuerzo.

A mi abuela, la primera que me enseñó a aprender y con la que aprendí a enseñar.

A mis padres, por haberme inculcado el valor de la educación y haber estado siempre ahí cuando más los necesitaba. Por respetar siempre mis elecciones y darme siempre para adelante, poniendo siempre primero mi felicidad.

A mis hermanas, por haberme enseñado tanto, por haber sido un ejemplo para mí, cuando se suponía que tenía que ser al revés.

A mis amigas, Naty, por ser la mejor compañera durante las interminables horas de estudio para salvar tantos cursos y exámenes, Natu, por compartir conmigo la pasión por la genética y aguantarme siempre, Amy, por tener siempre los mejores apuntes y darme siempre para adelante, sin ustedes no hubiera llegado nunca hasta acá. A Vicky, por ser la mejor compañera de monografías posible. A Diego y Seba, que a pesar de abandonarnos a medio camino de la carrera, siguen siendo parte fundamental de mi vida. Pero por sobre todo, gracias a todos y cada uno de ustedes por ser mis amigos. A todos los compañeros de facultad que de una manera u otra me ayudaron a llegar hasta acá.

A Fede, no alcanzan las palabras para agradecerte todos estos años de aguante y apoyo incondicionales, por compartir tu vida conmigo, por estar siempre para festejar o llorar juntos. Te amo.

A todos, infinitas gracias por ayudarme a llegar hasta acá, y ojalá estén conmigo para compartir todo lo que se viene.

Abreviaturas

ADN: ácido desoxirribonucleico

ARN: ácido ribonucleico

ARNm: ARN mensajero

BCAC: Breast Cancer Association Consortium

BIC: Breast Cancer Information Core

IMC: índice de masa corporal por su sigla en inglés Body Mass Index

CI: Intervalo de confianza

COGS: Estudio Oncológico Colaborativo de Genes-Ambiente

DCIS: carcinoma ductal in situ por su nombre en inglés Ductal Carcinoma In Situ

df: Grados de libertad por su sigla en inglés degrees of freedom

DSB: rupturas de doble cadena por su sigla en inglés Double Strand Break

FRR: riesgo familiar relativo por su sigla en inglés Familial Relative Risk

GWAS: Genome Wide Association Study – Estudios de asociación genómicos

HR: Recombinación Homóloga

HW: Hardy-Weinberg

HWD: Desequilibrio de Hardy-Weinberg

HWE: Equilibrio de Hardy-Weinberg

IBC: Cáncer inflamatorio de mama por su nombre en inglés Inflammatory Breast Cancer

IDC: Carcinoma Ductal Invasivo por su nombre en inglés Invasive Ductal Carcinome

ILC: Carcinoma lobulillar Invasivo por su nombre en inglés Invasive Lobular Carcinome

kg: Kilogramos

LD: Desequilibrio de ligamiento

LOH: pérdida de heterocigosidad por su sigla en inglés Loss Of Heterozygosity

MAPK: Proteínas activadas por mitógenos por su sigla en inglés Mitogen-activated protein kinase

OR: Odds Ratios

SCF: Complejo que contiene las proteínas Skp, Cullin, y F-box

SNP: Polimorfismo de Nucleótido Simple por su sigla en inglés Single Nucleotide Polymorphism

Resumen

De todos los tipos de cáncer, el de mama es uno de los más comunes, y la causa principal de muertes relacionadas con cáncer en mujeres alrededor del mundo. Uruguay, junto con Argentina, es de los países de América Latina con mayores tasas de incidencia y mortalidad en cáncer de mama, sobrepasando el promedio de la región. Las tasas de incidencia se asemejan a las de países desarrollados, mientras que las tasas de mortalidad se acercan más a las de los países en vías de desarrollo.

Entre los años 2004-2008 se diagnosticaron 2258 casos de cáncer de mama dando cuenta del 35% de todos los casos de cáncer en el Uruguay. En el mismo período 729 mujeres murieron como consecuencia de cáncer de mama en nuestro país lo que representa un 20% de las muertes asociadas a esta enfermedad.

Se analizaron los datos obtenidos de una muestra de 205 mujeres afectadas por cáncer de mama esporádico – pacientes – y 216 mujeres sanas – controles para 27 SNPs de 8 genes candidatos (BRCA1, BRCA2, ESR1, ESR2, FBXO31, TP53, TUBB3 y VDR).

Debemos tener en cuenta el IMC, siendo el sobrepeso y la obesidad factores de riesgo para cáncer de mama en nuestra población.

El cáncer de mama esporádico muestra el mismo comportamiento que en otras poblaciones, no se encuentra asociación de los genes TP53, BRCA1 y BRCA2.

El SNP rs827423 está relacionado con cáncer de mama, así como una amplia región marcada por el haplotipo CACG de rs827423 rs6912184, rs2982712, y rs2273206, por lo tanto ESR1 está relacionado con cáncer de mama en nuestra población.

Nuestros resultados indican que VDR podría estar relacionado, sugiriendo nuevos candidatos para futuras investigaciones.

FBXO31 podría estar relacionado, debiendo estudiarse más este gen para poder obtener resultados concluyentes.

El mestizaje debe ser tenido en cuenta al realizar estudios poblacionales en el Uruguay.

Palabras clave: cáncer, mama, Uruguay, asociación, SNP, IMC, genes candidatos

Introducción

Cáncer de Mama

La organización mundial de la salud define al cáncer como la multiplicación rápida de células anormales que se extienden más allá de sus límites habituales y pueden invadir partes adyacentes del cuerpo o propagarse a otros órganos. Otros términos usados para referirse a esta enfermedad son tumores malignos y neoplasmas (OMS, 2013).

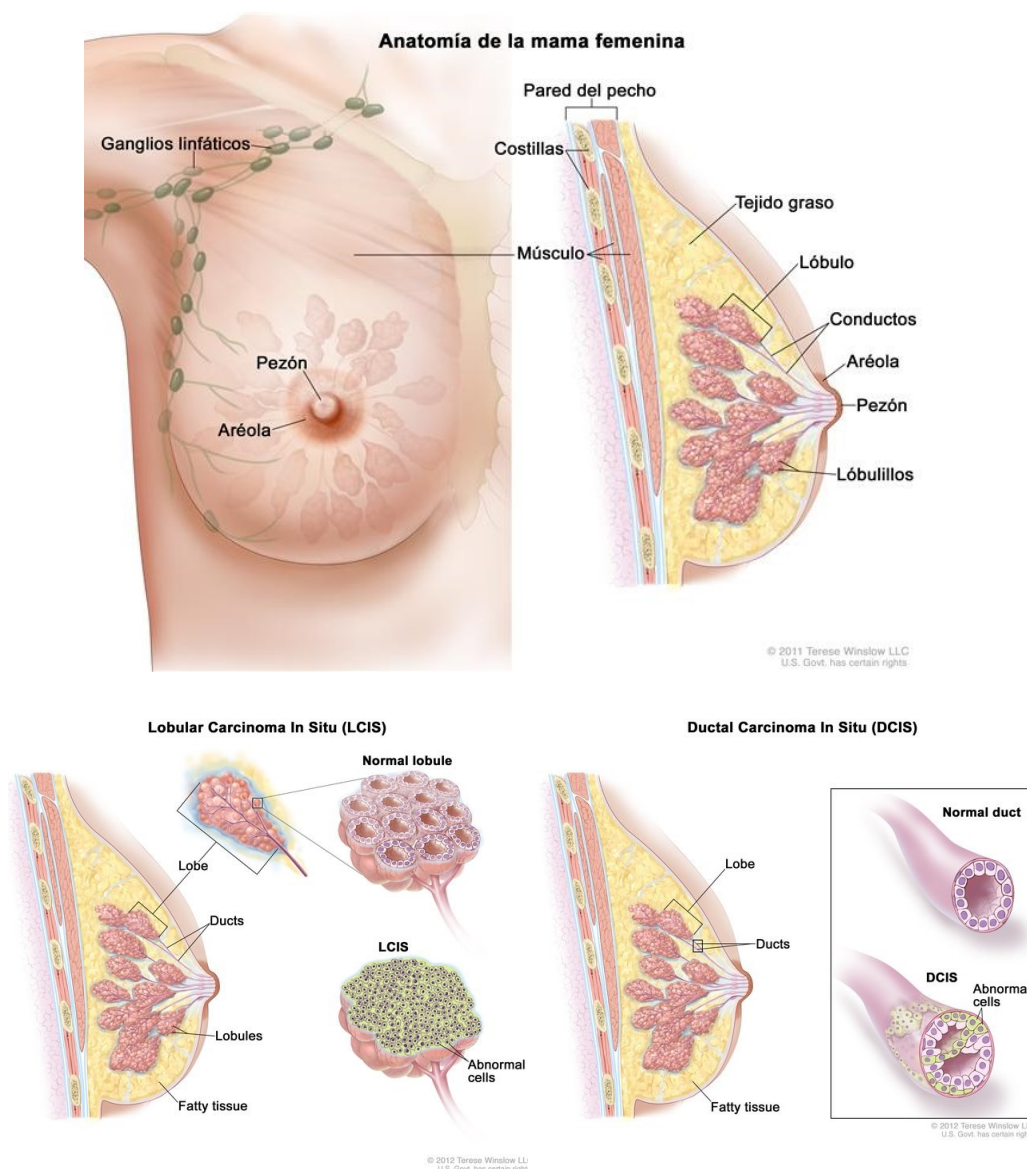


Figura 1. Anatomía de la mama femenina. En la parte superior se observa un diagrama de la anatomía de una mama normal. En la parte inferior se observan diagramas de los dos tipos de cáncer, lobulillar (a la izquierda) y ductal (a la derecha). Imágenes obtenidas del National Cancer Institute (www.cancer.gov)

Según el tejido donde se forma, podemos distinguir distintos tipos de cáncer de mama (Instituto Nacional del Cáncer, 2013), aunque algunos de ellos se presentan en pocas ocasiones. (Ver Figura 1)

En algunos casos, un solo tumor de mama puede ser una combinación de estos tipos o ser una mezcla de in situ y cáncer invasivo (American Cancer Society, 2013).

En principio se definen los distintos tipos como:

Carcinoma Ductal - El tipo más común de cáncer de mama es el carcinoma ductal. Este cáncer empieza en las células que revisten los conductos de la mama. Casi siete de cada diez mujeres con cáncer de mama tienen carcinoma ductal (Instituto Nacional del Cáncer, 2013).

Carcinoma ductal in situ - El carcinoma ductal in situ (ductal carcinoma in situ, DCIS; también conocido como carcinoma intraductal) se considera cáncer de mama no invasivo o preinvasivo. DCIS significa que las células que cubren los conductos han cambiado y lucen como células cancerosas. La diferencia entre el DCIS y el cáncer invasivo consiste en que las células no se han propagado (invadido) a través de las paredes de los conductos hacia el tejido que rodea la mama. El DCIS se considera un precáncer porque en algunos casos se puede convertir en un cáncer invasivo. Alrededor de uno de cinco casos nuevos de cáncer de mama serán DCIS (American Cancer Society, 2013).

Carcinoma ductal invasivo (o infiltrante) - El carcinoma ductal invasivo o infiltrante (IDC, sigla en inglés) es el tipo más común de cáncer de mama. Este cáncer comienza en un conducto lácteo de la mama, penetra a través de la pared del conducto y crece en el tejido adiposo de la mama. En este punto puede tener la capacidad de propagarse (metástasis) hacia otras partes del cuerpo a través del sistema linfático y el torrente sanguíneo. Aproximadamente ocho de cada diez de los cánceres invasivos de mama son carcinomas ductales infiltrantes.

Carcinoma lobulillar - El segundo tipo más común de cáncer de mama es el carcinoma lobulillar. Este cáncer empieza en un lobulillo de la mama. Casi una de cada diez mujeres con cáncer de mama tiene carcinoma lobulillar.

Carcinoma lobulillar in situ - El carcinoma lobulillar in situ no es un verdadero cáncer o precáncer.

Carcinoma lobulillar invasivo (o infiltrante) - El carcinoma lobulillar invasivo (invasive lobular carcinoma, ILC) comienza en las glándulas productoras de leche (lobulillos). Al igual que el IDC, se puede propagar (metástasis) a otras partes del cuerpo. De diez casos de cáncer invasivo de mama, aproximadamente uno es ILC (American Cancer Society, 2013).

A su vez, el cáncer de mama se puede presentar como una mezcla de los tipos de carcinoma ductal y lobulillar o como un tipo menos común de cáncer de mama. (Instituto Nacional del Cáncer, 2013)

Al considerar los tipos de cáncer de mama menos comunes encontramos otra vez diferentes tipos definidos como:

Cáncer inflamatorio de mama - Este tipo de cáncer invasivo de la mama representa aproximadamente del uno al tres por ciento de todos los cánceres de mama. No se presenta una sola protuberancia o tumor, sino que el cáncer inflamatorio del mama (inflammatory breast cancer, IBC) produce síntomas similares a los de una infección del mama (mastitis), provocados por la obstrucción de los vasos linfáticos por las células cancerosas.

Este tipo de cáncer de mama suele tener una mayor probabilidad de propagación y un peor pronóstico que el típico cáncer ductal invasivo o lobulillar invasivo.

Cáncer de mama triple negativo - Este término se usa para describir los cánceres de mama (usualmente carcinomas ductales invasivos), cuyas células carecen de receptores de estrógeno y receptores de progesterona, y no tienen un exceso de proteína HER2 en sus superficies. Los cánceres de mama con estas características tienden a presentarse con más frecuencia en mujeres más jóvenes y en mujeres afrodescendientes. El cáncer de mama triple negativo tiende a crecer y a propagarse más rápidamente que la mayoría de los otros tipos de cáncer de mama.

Enfermedad de Paget del pezón - Este tipo de cáncer de mama comienza en los conductos del mama y se propaga hacia la piel del pezón y después hacia la areola. Representa sólo alrededor del uno por ciento de todos los casos de cáncer de mama. La enfermedad de Paget está casi siempre asociada al carcinoma ductal in situ o el carcinoma ductal infiltrante.

Tumor filoides - Este tipo de tumor de mama se forma en el estroma (tejido conectivo) del mama, a diferencia de los carcinomas, que se forman en los conductos o en los lobulillos. Por lo general, estos tumores son benignos, pero en pocos casos pueden ser malignos.

Angiosarcoma - Esta forma de cáncer se origina en las células que cubren los vasos sanguíneos o los vasos linfáticos. En raras ocasiones, este cáncer se origina en los senos.

También es posible encontrar tipos especiales de carcinoma invasivo de la mama. Éstos son subtipos de carcinoma invasivo. Algunos pueden tener un mejor pronóstico que el carcinoma ductal infiltrante regular, como los carcinomas quístico, adenoide, adenoescamoso de bajo grado, medular, mucinoso, papilar y tubular.

Entre los subtipos con pronóstico igual o peor que el carcinoma ductal infiltrante regular se incluyen los carcinomas metaplásico, micropapilar y mixto (tiene características de ductal invasivo y lobulillar). Aunque en general, todos estos subtipos siguen siendo tratados como carcinoma ductal infiltrante regular (American Cancer Society, 2013).

Aunque muchas veces no es considerado, el cáncer de mama ocurre tanto en hombres como en mujeres, aunque el cáncer de mama masculino es raro representando menos del 1% de todos los casos de cáncer de mama en el Uruguay (Instituto Nacional del Cáncer, 2013).

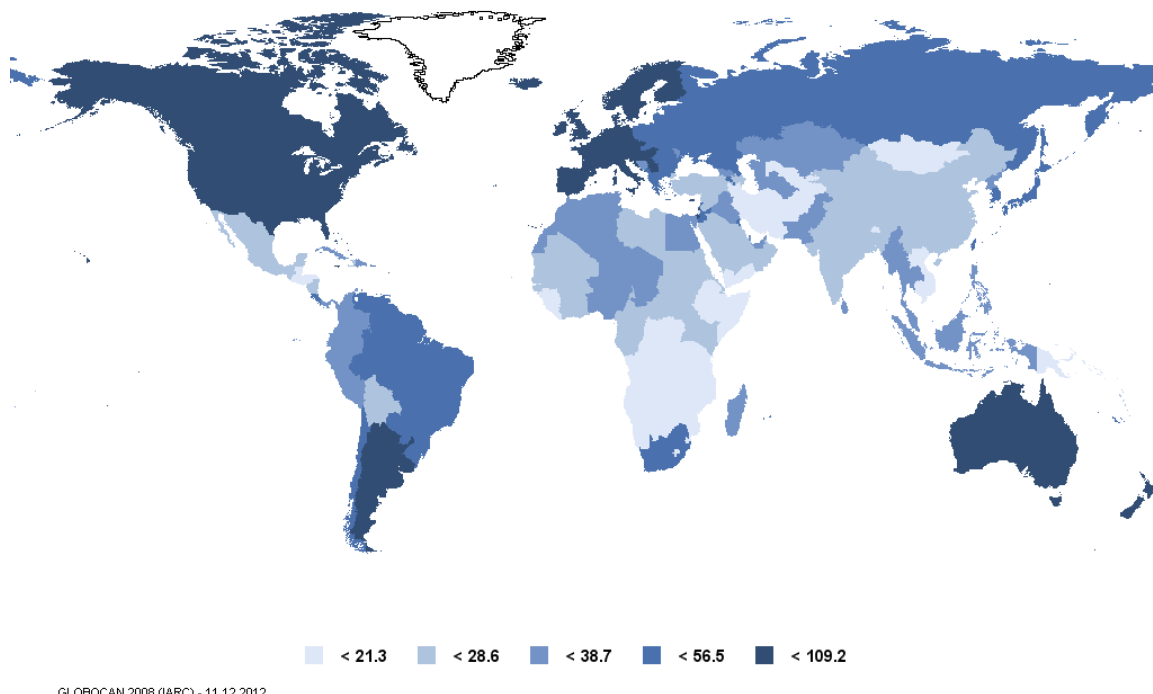


Figura 2. Tasas de incidencia estandarizadas por edad x 100000 para Cáncer de Mama en 2008 a nivel mundial. (IARC, 2008)

A nivel mundial, el cáncer de mama es el cáncer más frecuente entre mujeres con un número estimado de 1.38 millones de nuevos casos diagnosticados en 2008, lo cual representa el 23 % de todos los tipos de cáncer. Mientras que considerando hombres y mujeres es el segundo sobre el total de tipos de cáncer (10.9 %). En este momento es el cáncer más común tanto en regiones desarrolladas como en desarrollo, con alrededor de 690.000 nuevos casos estimados en cada región (Ferlay et al., 2010).

Las tasas de incidencia varían desde 19.3 cada 100.000 mujeres en África Oriental a 89.7 cada 100.000 mujeres en Europa Occidental, y son altas (mayores a 80 cada 100.000) en regiones desarrolladas del mundo (excepto Japón) y bajas (menores de 40 cada 100.000) en la mayoría de las regiones en desarrollo (Ferlay et al., 2010)Figura 2.

El rango de mortalidad a nivel global es mucho menor (aproximadamente 6-19 cada 100.000) debido a que la probabilidad de supervivencia es más favorable en regiones desarrolladas. Como resultado, el cáncer de mama es la quinta causa de muerte por cáncer a nivel global (458.000 muertes), pero entre mujeres es la causa más frecuente de muertes por cáncer ya sea en regiones en vías de desarrollo (12.7%), como en regiones desarrolladas, donde el estimado de 189.000 fallecidas es casi igual al número estimado de muertes por cáncer de pulmón (188.000 muertes) en el mismo período (Ferlay et al., 2010).

A nivel regional, Uruguay, junto con Argentina, es de los países de América Latina con mayores tasas de incidencia y mortalidad en cáncer de mama, sobrepasando el promedio de la región. Las tasas de incidencia se asemejan a las de países desarrollados, Figura 1, mientras que las tasas de mortalidad se acercan más a las de los países en vías de desarrollo, Figura 3.

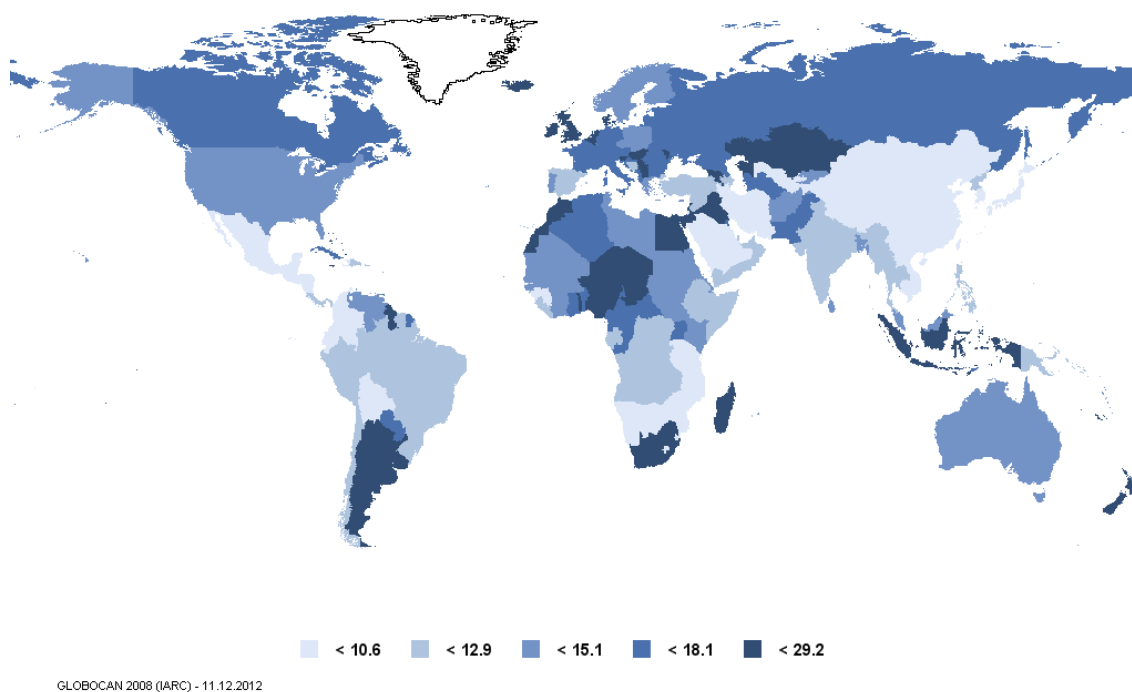


Figura 3. Tasas de mortalidad estandarizadas por edad x 100000 para Cáncer de Mama en 2008 a nivel mundial. (IARC, 2008)

Entre los años 2004-2008 se diagnosticaron 2258 casos de cáncer de mama dando cuenta del 35% de todos los casos de cáncer en el Uruguay (Figura 4). En el mismo período, 729 mujeres murieron como consecuencia de cáncer de mama en nuestro país lo que representa un 20% de las muertes asociadas a esta enfermedad (Figura 5) (IARC, 2008). Estos datos muestran la importancia de este tipo de cáncer en el Uruguay con un perfil de incidencia semejante al que se encuentra en las poblaciones de regiones desarrolladas pero con una tasa de mortalidad mayor aun que el promedio mundial para regiones en vías de desarrollo.

El cáncer de mama es una enfermedad multifactorial que involucra una compleja interacción entre factores genéticos y ambientales (Maguire et al., 2005), incluso ciertos factores evolutivos podrían estar involucrados. Un dato relevante que surge de lo señalado anteriormente es que las tasas de incidencia y mortalidad difieren entre distintos grupos raciales, siendo mayor la incidencia para mujeres eurodescendientes y mayor la mortalidad para las afroamericanas en la población de EUA (Gleason, Mdzinarishvili, & Sherman, 2012). A su vez, la prevalencia de cáncer de mama es mayor en poblaciones europeas que en poblaciones taiwanesas, y los patrones epidemiológicos son diferentes en los dos

grupos de pacientes (Kuo et al., 2012). La población uruguaya tiene contribución europea, africana y nativo americana, por lo que esperamos que no se comporte como ninguna de las poblaciones fundadoras, sino de una manera única, de acuerdo a sus componentes ancestrales mezcladas (Sans, 2000).

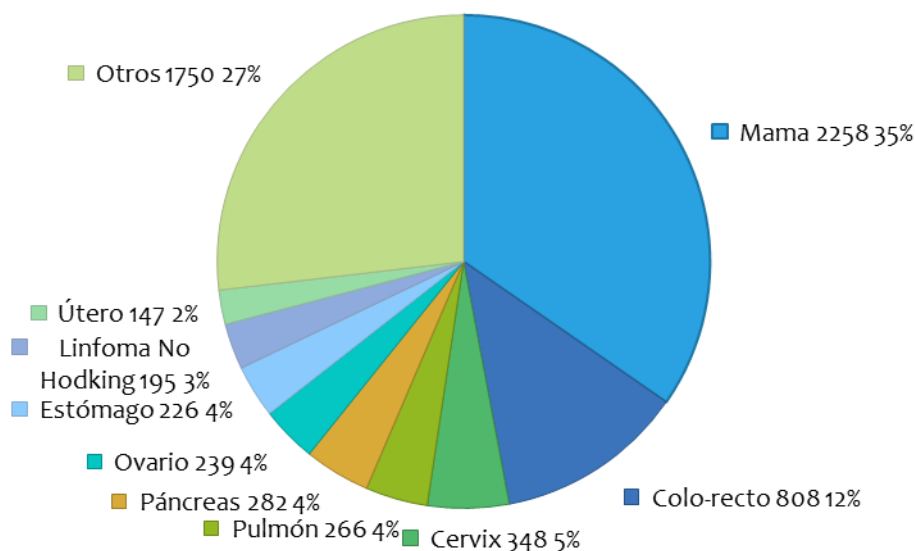


Figura 4. Número de casos de cáncer en mujeres en el Uruguay entre 2004-2008, adaptado de (IARC, 2008)

Diversos factores inherentes al estilo de vida, como no tener hijos o tenerlos a edades tardías, el uso de anticonceptivos orales, las terapias de reemplazo hormonal, la ingesta de alcohol, el hecho de haber amamantado poco o no haberlo hecho, y tener sobrepeso u obesidad están todos asociados con un riesgo aumentado de padecer cáncer de mama (Gleason et al., 2012).

Otro factor común a tener en cuenta en la mayoría de los tipos de cáncer es la edad, y está ampliamente aceptado el utilizar tasas ajustadas por edad. Para muchos tipos de cáncer, las tasas aumentan en edades intermedias y disminuyen para las edades más jóvenes y cerca de los 70-80 años (Gleason et al., 2012).

Sin embargo, contrario a otros tipos de cáncer, para cáncer de mama la curva de tasas específicas por edad tienen un único punto de inflexión cercano a los 50 años, por lo que se ha propuesto la hipótesis de que coincide con la edad de comienzo de la menopausia. Esto ha permitido a algunos investigadores asumir que el cáncer de mama es una enfermedad heterogénea, con distintas etiologías – fenotipos de cáncer de mama – y que las poblaciones pueden ser fraccionadas de acuerdo a su exposición a distintos factores de riesgo. Dicha caracterización continúa siendo un desafío. La idea de una población heterogénea no es nueva; por años los clínicos han dividido a los pacientes de cáncer de mama en dos

grupos, de acuerdo a la edad de diagnóstico, como si se tratara de dos enfermedades diferentes. En un grupo, denominado de comienzo precoz, se agrupan las pacientes diagnosticadas con cáncer de mama a edades cercanas a los cuarenta, mientras que en el segundo grupo, denominado de comienzo tardío, se agrupan a aquellas mujeres cuyo diagnóstico fue realizado a edades cercanas a los 70 años (Gleason et al., 2012).

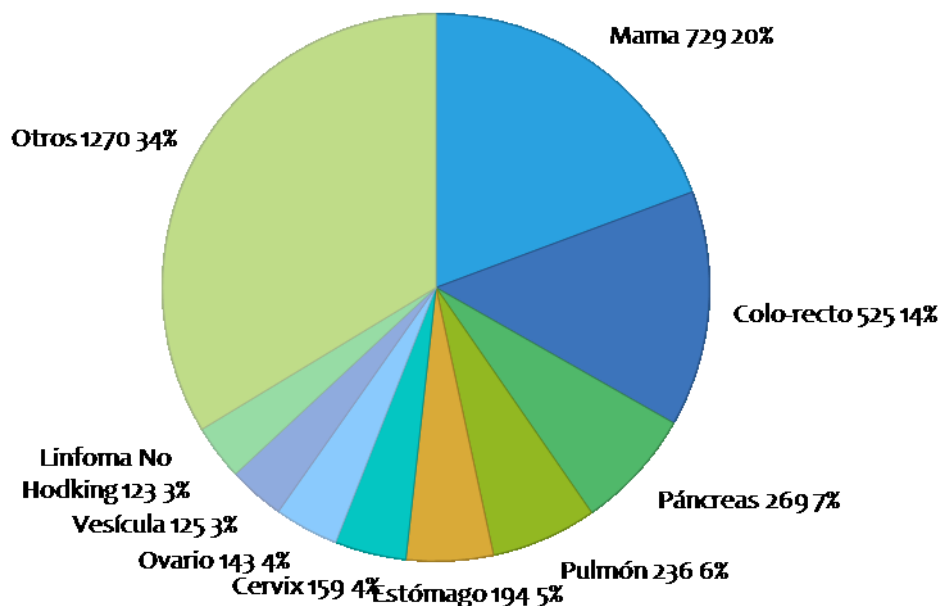


Figura 5. Número de muertes por cáncer en mujeres en el Uruguay entre 2004-2008, adaptado de Ferlay, 2010

La edad promedio de comienzo de la menopausia, alrededor de los 50 años, se utiliza como punto de corte y sirve como una barrera entre precoces y tardíos. La separación de la población de cáncer de mama en dos fracciones por un punto de corte de edad, sin embargo, podría reflejar solamente una tendencia general de los fenotipos de cáncer de mama que son prevalentes a edades extremas. Sin embargo, este acercamiento es una simplificación del problema, ya que no tiene en cuenta la estratificación por distintos marcadores de riesgo, pudiendo tener cada uno distintos patrones de envejecimiento. Datos recientes sugieren que ciertos factores de riesgo no modificables, como la raza, presentan distintos patrones al envejecer: las tasas en mujeres eurodescendientes son menores que para las afrodescendientes hasta aproximadamente los 40 años a partir de ese punto las curvas se cruzan y las tasas para éstas últimas son menores que para las primeras. Si bien este cruce ha sido documentado, se desconocen aún las razones y no ha sido considerado aún como parte de la heterogeneidad del cáncer de mama (Gleason et al., 2012).

Debido a su naturaleza compleja es razonable pensar que el riesgo aumentado de padecer cáncer de mama se debe a la contribución de diversos factores, tanto genéticos como ambientales, exógenos o en

particular endógenos, tales como los hormonales. Desde hace un tiempo se presta especial atención al efecto de las hormonas en la patogénesis del cáncer de mama. Se ha encontrado evidencia sustancial proveniente de estudios tanto epidemiológicos como experimentales que demuestran un rol crucial de las mismas en la etiología del cáncer de mama. La actividad ovárica y endócrina relacionadas con la reproducción están marcadas, al inicio y al final respectivamente, por la menarca y la menopausia. El tiempo que transcurre entre la menarca y la menopausia, años reproductivos de la mujer, el desarrollo y la función mamaria se ven directamente afectados por la producción de hormonas esteroideas por el ovario. La menarca temprana y la menopausia tardía son conocidas por aumentar el riesgo de las mujeres de desarrollar cáncer de mama lo que aumenta la exposición a estrógenos y progesterona. Se ha observado en mujeres postmenopáusicas con cáncer de mama mayores niveles circulantes de estrógenos y andrógenos, dando evidencia epidemiológica directa de la relación entre hormonas y cáncer de mama. El nexo entre los esteroides sexuales y el riesgo de cáncer pre-menopáusico aún no está claro. El estrógeno juega un rol central en el desarrollo normal de la glándula mamaria pero también está involucrado en la progresión del cáncer de mama. Estudios epidemiológicos han mostrado que la exposición a estrógeno es un factor de riesgo para el cáncer de mama, con una exposición prolongada o aumentada elevando el riesgo, mientras que una exposición reducida tiene un efecto protector. Se han reportado múltiples mecanismos relacionados con procesos hormonales involucrados con la aparición y la progresión del cáncer de mama. La respuesta de la célula a los estrógenos esta mediada por los receptores específicos de estrógeno, ESR1 y ESR2. Dichos receptores regulan la transcripción de genes blanco tanto por unión directa a los elementos de respuesta a estrógenos como por la interacción con otros factores de transcripción. Estudios experimentales han mostrado que los estrógenos contribuyen a la carcinogénesis de la mama ya que estimulan la proliferación celular mediada por el receptor de estrógeno-ESR1- por efectos genotóxicos de los metabolitos o la inducción de aneuploidía. Las variantes presentes en genes responsables del metabolismo de hormonas esteroideas o de su señalización son factores potenciales que podrían afectar el riesgo de cáncer en tejidos hormona-dependientes. Dado que los estrógenos son conocidos por afectar el desarrollo de cáncer de mama, es esencial analizar polimorfismos en los dos genes de receptores de estrógeno, ESR1 y ESR2 así como en sus proteínas co-reguladoras, ya que podrían afectar el riesgo de cáncer de mama. Los genes involucrados en la biosíntesis de hormonas y metabolismo son candidatos adecuados como genes que posiblemente estén involucrados con la susceptibilidad de cáncer de mama dado que una aumentada exposición hormonal es un factor de riesgo para el cáncer de mama (Collaborative Group on Hormonal Factors in Breast Cancer, 2012; Maguire et al., 2005; Süllner et al., 2012; Zheng et al., 2012).

La historia familiar es el principal factor de riesgo para cáncer de mama, indicando que los factores genéticos juegan un papel muy importante en el desarrollo de la enfermedad, aunque entre todos los casos de cáncer de mama, 5-10% son asignados a la predisposición hereditaria (Fanale et al., 2012) (Jiaxue et al., 2010).

Desde la década de los 90, como resultado de estudios de ligamiento en familias con casos de cáncer de mama, se conocen dos genes principales asociados con la susceptibilidad a cáncer de mama familiar, BRCA1 y BRCA2. Mutaciones en la línea germinal de estos genes ocurren con baja frecuencia en la población general pero confieren un alto riesgo de padecer cáncer de mama y ovario y bajo riesgo para otros tipos de cáncer (Fanale et al., 2012). Las mutaciones de la línea germinal de BRCA1 usualmente ocurren en un alelo, mientras que el otro alelo sano puede mutar o perderse durante el desarrollo del cáncer (hipótesis de Knudson) (Jiaxue et al., 2010). Por otra parte, mutaciones en el gen TP53 también aparecen asociadas con un alto riesgo de cáncer de mama. Sin embargo, estudios genético-epidemiológicos han estimado que las alteraciones en estos genes dan cuenta solamente del 15% del riesgo de cáncer familiar. Los análisis de ligamiento genético fallaron al tratar de identificar otros genes de alta penetrancia responsables de la susceptibilidad a esta enfermedad y la identificación de variantes raras asociadas con un riesgo moderado de enfermedad pueden explicar solo una pequeña proporción del riesgo asociado al cáncer familiar (Fanale et al., 2012).

Desde el descubrimiento de BRCA1 y BRCA2 los esfuerzos intensivos en descubrir otros genes de efecto mayor han sido ampliamente infructuosos. Todas las mutaciones conocidas hasta el momento dan cuenta solamente del 25% de todos los casos de cáncer de mama familiar. El fallo al identificar nuevos genes de efecto mayor, indica la existencia de genes con mutaciones de penetrancia moderada y de baja penetrancia, que combinados aumentan el riesgo de cáncer de mama, cada uno aportando un pequeño efecto al riesgo global. Es probable que el riesgo varíe entre estos genes raros con moderada penetrancia y la acumulación de varios alelos comunes de baja penetrancia. Teniendo en cuenta este modelo denominado poligénico un gran número de variantes de bajo riesgo que ocurren con alta frecuencia en las poblaciones podrían tener un efecto multiplicativo al determinar el riesgo general de la enfermedad. Este modelo poligénico de susceptibilidad de cáncer de mama es sobre el que se basan los estudios de asociación de polimorfismos de nucleótido simple -SNPs- de genes candidatos para identificar esos genes de baja penetrancia. (Fanale et al., 2012; Maguire et al., 2005)

Estudios recientes, utilizando el arreglo de Illumina iCOGS (Collaborative Oncological Gene-environment Study) específicamente designado para genotipar un gran número de individuos en estudios de asociación, detectaron 41 nuevos loci relacionados con el riesgo global de padecer cáncer de mama. De los 27 loci previamente publicados asociados con cáncer de mama, todos menos cuatro mostraron una clara asociación en los datos de COGS. A esto se agregan variantes comunes en 41 nuevos loci que fueron asociadas con cáncer de mama. Los SNPs en 4 loci (rs4245739 en 1q32.1, rs6678914 en 1q32.1, rs12710696 en 2p24.1, rs11075995 en 16q12.2) están específicamente relacionados con cáncer de mama ER-negativo en poblaciones europeas (García-Closas et al., 2013). Otra variante dentro del mismo locus (rs2290854, 1q32) también fue asociada con el riesgo de cáncer de mama en portadores de mutaciones en BRCA1 (Couch et al., 2013), mientras que un locus independiente (rs9348512 6p24) fue asociado

específicamente con el riesgo de cáncer de mama en portadores de mutaciones de BRCA2 (Gaudet et al., 2013).

El proyecto COGS incluyó el mapeo fino de más de 10 regiones genómicas seleccionadas para identificar variantes más fuertemente asociadas con la enfermedad, así como variantes asociadas adicionales que podrían contribuir a la heredabilidad. Dichos estudios identificaron tres loci independientes en la región previamente reportada 11q13 y dos loci de susceptibilidad en la región 5p15. Estos hallazgos aumentan el número de loci para susceptibilidad al cáncer de mama a 76 (Bahcall, 2013). En la Figura 6 se resumen los loci de susceptibilidad a cáncer de mama conocidos y las proporciones estimadas de riesgo familiar relativo (FRR) atribuibles a las diferentes variaciones genéticas dentro de cáncer de mama familiar. El resultado del trabajo citado, en concordancia con su enfoque, obtiene resultados de susceptibilidad en cáncer de mama familiar a nivel general.

La estrategia de genes candidatos se enfoca en las asociaciones entre variación genética dentro de genes de interés pre-especificados con fenotipos o estados de enfermedad. Esto contrasta con los estudios de asociación de genoma completo (como los de los trabajos citados) los cuales analizan el genoma entero para encontrar variación genética común.

En nuestro trabajo esperamos encontrar resultados acerca del cáncer de mama esporádico de manera específica para nuestra población.

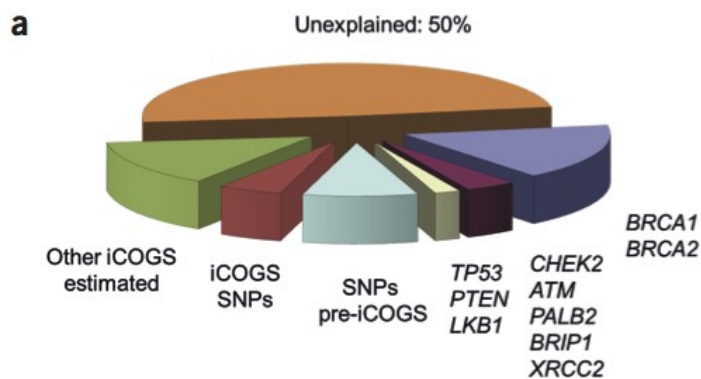


Figura 6. Proporciones estimadas del FRR atribuibles a los distintos tipos de variación genética. Los genes listados se refieren a: genes con mutaciones conocidas causantes de la enfermedad (BRCA1 y BRCA 2), otros genes que confieren alto riesgo (TP53, PTEN y LKB1), y genes de riesgo moderado (CHEK2, ATM, PALB2, BRIP1, XRCC2). Pre-iCOGS SNPs se refiere a la contribución de todos los SNPs asociados con este tipo de cáncer antes de la publicación citada, iCOGS SNPs refiere a la contribución de los SNPs establecida por el trabajo citado y Other iCOGS estimated refiere a la contribución estimada por el trabajo citado. (Bahcall, 2013)

Estudios de Asociación

El término estudio de asociación se refiere a estudios que consideran la relación entre la información de secuencias genéticas y un fenotipo. En el contexto de estudios de asociación, la información genética cruda caracterizada por la secuencia de ADN es el principal predictor en investigación, y el objetivo es entender como polimorfismos en las secuencias explican la variabilidad en un rasgo de la enfermedad. En este tipo de estudios, generalmente se trata al genotipo como la variable de predicción y al rasgo como la variable dependiente (Foulkes, 2009).

Investigaciones poblacionales versus familiares

Las investigaciones poblacionales se refieren a aquellas en la que los individuos no están relacionados, mientras que las familiares, como su nombre lo indica, involucra la colección de datos de múltiples individuos dentro de la misma unidad familiar. Las consideraciones estadísticas para ambos tipos de estudios son diferentes, dado que dentro de la familia es probable que sean más similares entre ellos que individuos provenientes de otras familias. Como resultado, la característica bajo investigación está más altamente correlacionada entre individuos de la misma familia. En los estudios poblacionales es fundamental que los individuos no estén relacionados, sin embargo otras formas de agrupamiento pueden existir. Por ejemplo, individuos que han sido reclutados a lo largo de múltiples hospitales, los pacientes del mismo hospital son más similares entre ellos que respecto a los otros. Esto puede surgir especialmente si las áreas donde se encuentran dichos hospitales incluyen diferentes estatus socioeconómicos o si los estándares de cuidado al paciente son claramente diferentes. Una segunda diferencia remarcable entre ambos estudios es la determinación de la fase alélica, definida como el alineamiento de nucleótidos en un único cromosoma homólogo. Mientras que en los estudios poblacionales se determina en forma estadística, en los estudios familiares se determina en forma analítica. Como resultado de estas diferencias los métodos para analizar los datos en ambos casos difieren (Foulkes, 2009).

En los estudios de asociación poblacionales la unidad fundamental de análisis es el SNP. Un SNP describe un único cambio de base que es variable a lo largo de la población general con una frecuencia de al menos 1%. (Foulkes, 2009)

Genes Candidatos

Otra forma de aproximarse a comprender el desarrollo de una enfermedad compleja como el cáncer es analizar genes que formen parte de vías metabólicas importantes vinculadas con el órgano afectado o profundizar sobre la estructura de genes conocidos.

En el cáncer de mama esporádico se puede pensar que genes de efecto mayor pueden llevar mutaciones que sin causar la pérdida de función de la proteína cambien la eficiencia de sus funciones. En este trabajo se analizan varios genes que podrían ser candidatos. En la siguiente parte describimos los genes considerados.

BRCA1

17q21

El gen BRCA1 fue uno de los primeros genes en ser descrito asociado a cáncer de mama familiar por Gayther y colaboradores en 1995, se expande más de 70 kb en el ADN genómico, contiene 24 exones y codifica una proteína de 1863 aminoácidos (Kuo et al., 2012).

El exón 11 es el mayor y codifica más del 60% de los aminoácidos de BRCA1. Aunque comparte muy limitada homología con otras proteínas conocidas, BRCA1 contiene dos dominios funcionales: un anillo N-terminal y un dominio C-terminal - BRCT. Similar a otros dominios de anillo, el anillo de BRCA1 contiene actividad E3 ubiquitin ligasa y facilita la ubiquitinación de proteínas. El dominio C-terminal es un dominio de unión a fosfoproteínas, que también es requerido para la translocación y acumulación de BRCA1 en sitios de daño al ADN (Jiaxue, et al. 2010). El mismo presenta más de 6000 alelos diferentes la mayoría de los cuales implican mutaciones de cambio de sentido, de acuerdo con la base de datos BIC - Breast Cancer Information Core (Kuo et al., 2012).



Figura 7. Contexto genómico del gen BRCA1. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

Aunque las mutaciones en BRCA1 están repartidas a lo largo del gen, muchas mutaciones asociadas al cáncer han sido encontradas en el dominio de anillo y BRCT de BRCA1, indicando que ambos son importantes en la supresión de la formación de cáncer de mama y ovario (Jiaxue, et al. 2010). Poder categorizar los cambios de sentido en mutaciones deletéreas, asociadas con el desarrollo de cáncer de mama, o como polimorfismos benignos resulta difícil y presenta un obstáculo significativo para los estudios genéticos (Kuo et al., 2012).

En el trabajo de Delgado y colaboradores (2011), realizado para familias con cáncer de mama en el Uruguay, detectaron dos mutaciones en BRCA1 asociadas a esta enfermedad (5583insT y 2687T>G,) las cuales no habían sido previamente reportadas. Estos datos aumentan la relevancia de BRCA1 como gen candidato en el presente estudio.

BRCA2

13q12.3

En los individuos portadores de mutaciones con pérdida de función en el gen supresor de tumores BRCA2 se observa una alta predisposición a una amplia variedad de tipos de cáncer. En las mujeres heterocigotas para un alelo mutado de BRCA2 se observa un riesgo elevado de desarrollar cáncer de mama y ovario. Adicionalmente, los cánceres de mama en hombres y de próstata están fuertemente asociados con las mutaciones en el gen BRCA2 (Barber et al., 2011).

En ambos sexos, la deficiencia de BRCA2 aumenta el riesgo de desarrollar cáncer de páncreas, estómago, vesícula, ducto biliar, y melanoma. BRCA2 juega un rol clave en el mantenimiento de la integridad genómica, particularmente a través de la regulación de la reparación por recombinación homóloga (RH) un proceso que también está controlado por otra proteína supresora de tumores, BRCA1. La RH es un proceso altamente libre de errores que restablece la secuencia original en el sitio de rupturas de doble cadena (*Double Strand Breaks DSBs*). Las DSBs surgen frecuentemente y pueden ser causadas por replicación celular normal así como por estrés exógeno o exposición a radiación ionizante (Barber et al., 2011).

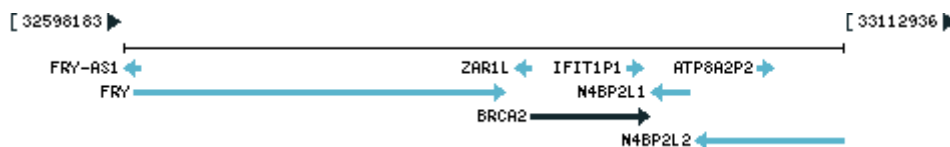


Figura 8. Contexto genómico del gen BRCA2. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

En ausencia de RH, por ejemplo por pérdida de función de BRCA2, las DSBs parecen ser reparadas por procesos más susceptibles de errores, que en última instancia lleva a la acumulación de re arreglos cromosómicos gruesos. Se piensa que los errores en los procesos de reparación de ADN por la ausencia de la función de BRCA2 albergan tumorigénesis (Barber et al., 2011).

En el trabajo de Delgado et al. (2011) previamente mencionado, también encuentran cinco variantes en BRCA2 asociadas con cáncer de mama familiar en nuestro país, (4359ins6, 5579insA, 3829insTdel35, 4088delA y 1617delAG), de las cuales 3829insTdel35 no había sido previamente reportada en la bibliografía. Todas las mutaciones generaron codones de terminación prematuros, y se encontraron con mayor frecuencia que aquellas reportadas para BRCA1. El hecho de contar con datos de nuestro país que relacionan a BRCA2 con cáncer de mama nos reafirman la elección de este gen como candidato para nuestro trabajo.

ESR1

6q25.1

El estrógeno ejerce su efecto biológico a través de receptores de estrógeno, el receptor de estrógeno alfa (ESR1) y el más recientemente descubierto receptor de estrógeno beta (ESR2). ESR1 y ESR2 pertenecen a la superfamilia de receptores nucleares de factor de transcripción ligando inducibles. Poseen alta similitud de secuencia en su ADN y dominios de unión a ligando pero divergen en el dominio de activación transcripcional, indicando que pueden tener como blanco un conjunto distinto de genes (Maguire et al., 2005).

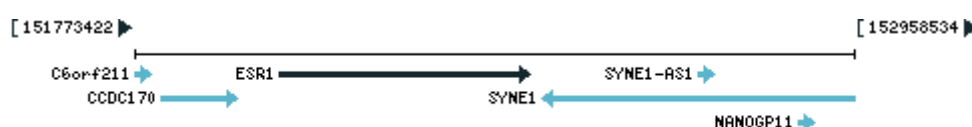


Figura 9. Contexto genómico del gen ESR1. Extraído del GenBank del National Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

El estrógeno estimula la proliferación celular y aumenta la frecuencia de mutaciones espontáneas, llevando a un fenotipo maligno. Las células mamarias responden a través de receptores de estrógeno – ERs – a través de un proceso bioquímico definido: luego de la unión al ligando los ERs son sometidos a un cambio de conformación que facilita la dimerización del receptor, la unión al ADN, el reclutamiento de cofactores y la modulación de la expresión de genes blanco. La terapia endocrina provee fuerte evidencia de que la atenuación de la actividad de ESR1 puede reducir el riesgo de cáncer de mama, y las mujeres con tumores ER positivos serían las que probablemente se beneficiarían más con estos tratamientos. Los estudios genéticos de ESR1, sin embargo, han tenido resultados contradictorios. Solo recientemente, a través de un gran estudio de asociación genético, se ha demostrado una pequeña pero significativa asociación de polimorfismos de ESR1 con el riesgo de cáncer de mama (Y. Li et al., 2011).

ESR2

14q23.2

Como en el caso con ESR1, también existen varias isoformas de ESR2, resultando del splicing alternativo del ESR2 wildtype. Los receptores de estrógeno de ambos homodímeros como heterodímeros y diferentes isoformas de ESR2 han sido sugeridos de tener un rol regulatorio en la actividad de ESR1. (Maguire et al., 2005)

Varios estudios han demostrado que ESR2 se expresa en tejido de mama normal y tumoral. (Iwao, Miyoshi, Egawa, Ikeda, & Noguchi, 2000; Leygue, Dotzlaw, Watson, & Murphy, 1999; Palmieri et al.,

2002; Zhao et al., 2003) También se ha mostrado que ESR2 está subexpresado en carcinogénesis de mama (Iwao et al., 2000; Leygue, Dotzlaw, Watson, & Murphy, 1998; Zhao et al., 2003)

El estudio llevado a cabo por Maguire y colaboradores (2005) identificó un haplotipo de ESR2 que se sugiere está asociado con un riesgo aumentado de cáncer de mama.



Figura 10. Contexto genómico del gen ESR2. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

Los resultados implican que a pesar de esto los SNPs examinados en este estudio no juegan un rol causal en la etiología del cáncer de mama, el gen ESR2 podría estar involucrado en la susceptibilidad al cáncer de mama. Debido al desequilibrio de ligamiento a lo largo del gen ESR2 es posible que estos SNPs estén en ligados a variantes causales todavía no identificadas en la cercanía del gen ESR2. Por lo tanto valdrá la pena investigar este locus en más detalle para intentar identificar nuevos alelos de susceptibilidad al cáncer de mama y entender un poco más la compleja etiología de la carcinogénesis de mama.

FBXO31

16q24.2

La progresión del ciclo celular está regulada de forma precisa; positivamente por ciclinas y kinasas dependientes de ciclinas (CDKs) y negativamente por inhibidores de CDKs, pertenecientes a las familias Ink y Kip. Durante un ciclo celular normal, la ciclina D1, un mediador esencial de la progresión de la mitosis, se acumula en la fase G1 y media en la activación transcripcional por señales mitogénicas. Mientras que en la fase S decrece debido a la degradación dirigida mediante el complejo multiproteico SCF (Skp, Cullin, F-box containing complex) (Jia & Sun, 2009).

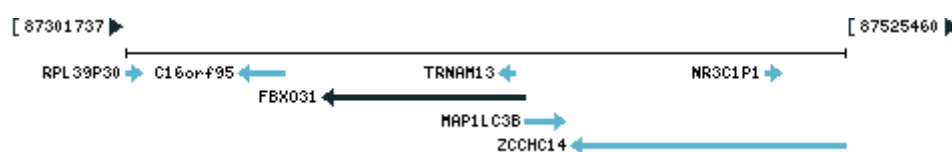


Figura 11. Contexto genómico del gen FBXO31. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

Las ligasas SCF son las mayores E3 ubiquitin ligasas encargadas de regular una variedad de procesos biológicos promoviendo la degradación temporal de diversos sustratos. Estructuralmente, Cullin-1 actúa como proteína de scaffold, donde su N-terminal se une al complejo Skp-F-box, y el C-terminal se une a RBX1, este último se une a E2 y cataliza la transferencia de ubiquitinas desde los sustratos a E2. La especificidad de sustrato del complejo SCF es, por otro lado, determinada por las proteínas F-box que reconocen los sustratos fosforilados para su unión y subsecuente degradación. Hasta el presente, tres proteínas de la familia F-box FBXO31, FBX4, y FBXW8, se han identificado por unir a la fosfociclina D1 para la degradación dirigida en respuesta al daño en el ADN y señales de MAPK (Proteínas activadas por mitógeno) (Jia & Sun, 2009).

Se ha encontrado un rol esencial de la proteína FBXO31 en la degradación dirigida de la ciclina D1 en respuesta al daño en el ADN en células de melanoma. FBXO31 fue inicialmente identificado en un ensayo de siARN como una proteína candidata requerida para que el gen oncogén BRAF induzca senescencia en fibroblastos primarios y melanocitos. La expresión ectópica de FBXO31 induce el arresto en G1 en células de melanoma como consecuencia de la reducción en la ciclina D1. (Jia & Sun, 2009)

En el trabajo de Kumar y colaboradores (2005), mostraron que FBXO31, puede inducir la senescencia celular en la línea de cáncer de mama MCF-7 y es el probable gen candidato de senescencia. FBXO31 tiene propiedades consistentes con un supresor de tumores, debido a que la expresión ectópica en dos líneas celulares de cáncer de mama inhibe el crecimiento de colonias in vitro e inhibe la proliferación en la línea celular MCF-7. Además, comparado con la expresión relativa en mama normal, los niveles de FBXO31 eran menores en líneas celulares y tumores primarios de mama. FBXO31 está regulado por el ciclo celular en las líneas MCF-10A y SKBR3 con expresión máxima desde G2 tardía hasta G1 temprana. La expresión ectópica de FBXO31 en la línea celular MDA-MB-468 resultó en la acumulación de células en la fase G1 del ciclo celular. Proponen que FBXO31 funciona como un supresor de tumores al formar complejos SCF cuyos objetivos son sustratos particulares, críticos para la normal ejecución del ciclo celular, los cuales son ubiquitinados y subsecuentemente degradados.

TP53

17p13.1

El gen supresor de tumores TP53 juega un rol crucial en el mantenimiento de la integridad genómica, ya que regula el arresto del ciclo celular, la apoptosis, la senescencia, la reparación del ADN y otros aspectos del crecimiento celular, y es un factor crítico en la inhibición de la tumorigénesis (Hu et al., 2010).

Designado como el guardián del genoma, es el gen más frecuentemente mutado en diferentes tipos de cáncer en humanos. En cáncer de mama esporádico, el gen TP53 está mutado en el 20-30% de los casos. Estas mutaciones pueden dañar la función de la proteína p53 como factor de transcripción y por lo tanto

muchas funciones celulares cruciales, como el control del ciclo celular, la reparación del ADN y la apoptosis, pueden estar alteradas (Trifa et al., 2010).

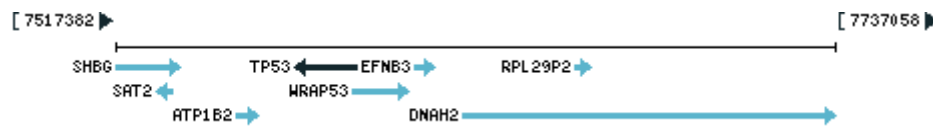


Figura 12. Contexto genómico del gen TP53. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

La frecuencia de mutaciones varía notablemente entre los diferentes tipos de tumores, en una gran proporción de los casos de cáncer de pulmón, gastrointestinal y de células gliales, y menos frecuentemente en melanomas y sarcomas. Dentro de las mutaciones de TP53 identificadas en cáncer, muchas codifican para proteínas mutantes con ganancia de función, adquiriendo la habilidad de cooperar con oncogenes activados, transformando morfológicamente células primarias. También se han descrito mutaciones más complejas como inserciones, deleciones, y mutaciones sin sentido. La frecuencia de mutaciones de TP53 es más alta cuando el cáncer de mama surge en un contexto de BRCA1 y BRCA2 mutados (Lo Nigro et al., 2012).

Además de las mutaciones, varios polimorfismos han sido identificados, tanto en exones como en intrones del gen TP53. El polimorfismo Arg72Pro (rs1042522) del exón 4 afecta la función apoptótica de la proteína p53 y por lo tanto juega un rol importante en cáncer en humanos. El codón 72 se localiza en una región rica en prolina que es necesaria para que la proteína p53 induzca apoptosis. Evidencia de estudios *in vitro* sugieren que estas dos formas de la proteína, Arg72 y Pro72, tienen actividades distintas y diferentes efectos en la regulación del crecimiento celular. Se ha reportado que la variante Arg/Arg es capaz de inducir apoptosis más eficientemente que el genotipo Pro/Pro (Trifa et al., 2010).

A su vez, la proteína con la variante Arg72 tiene incrementada la actividad apoptótica 15 veces en comparación con la variante Pro72. Estudios en células han reportado diferencias en la localización celular, la progresión del ciclo celular, la reparación del ADN, el arresto del crecimiento y la transactivación transcripcional para ambas variantes (Hu et al., 2010).

Por otro lado, el genotipo Pro/Pro parece inducir un mayor nivel de arresto de G1 que el genotipo Arg/Arg. La asociación entre el polimorfismo Arg72Pro y el riesgo de cáncer gástrico, cervical, mama y pulmón ha sido ampliamente reportada a través de meta-análisis y revisiones sistemáticas. Varios estudios han encontrado una asociación entre la variante Arg72 y un riesgo aumentado para cáncer de pulmón, colo-rectal, ovario y mama. Reportes contradictorios disponibles en la literatura implican la variante Pro72 en riesgo de cáncer nasofaríngeo, tiroides, pulmón y mama. Se ha sugerido una

asociación protectora para la variante heterocigota y el riesgo de cáncer de mama postmenopáusico. El polimorfismo de TP53 y su relación con la susceptibilidad de padecer cáncer de mama puede variar de una población a la otra, dependiendo del trasfondo genético y el nivel de exposición a factores ambientales (Trifa et al., 2010).

Un análisis utilizando un pool de datos incluyendo 10.618 pacientes y 8.743 controles realizado por el Breast Cancer Association Consortium (BCAC) no apoya el rol de este polimorfismo en el riesgo de desarrollar cáncer de mama. En el trabajo de Hu y colaboradores (2010), se realiza un meta-análisis de 37 estudios del codón 72 incluyendo 23.567 pacientes y 25.995 controles, y encuentran que este codón no está asociado con el riesgo de cáncer de mama, lo cual es consistente con la mayoría de los estudios publicados. Pero cuando realizaron un meta-análisis estratificado por subgrupos, encontraron una asociación significativa entre el codón y un riesgo disminuido de cáncer de mama para los estudios Mediterráneos. Interesantemente, las frecuencias de la variante Pro en dichos estudios son marcadamente más altas que en otros estudios Europeos. Más aún, notaron que la incidencia de cáncer en Europa del Sur (primariamente países Mediterráneos) es mucho menor que en Europa del Norte y del Oeste. Por lo cual, determinar si la alta frecuencia de Pro contribuye a la baja incidencia de cáncer de mama en las poblaciones Mediterráneas o si es el resultado de la interacción entre genes y aspectos ambientales requiere de una investigación más profunda (Hu et al., 2010).

TUBB3

16q24.3

La regulación génica específica del tipo celular es crucial para la diferenciación de tejidos y órganos, y también para el mantenimiento de varias funciones celulares en organismos adultos. La desregulación de estos tipos de especificidad celular podría ser una causa primaria de malformación de tejidos durante el desarrollo e incluso de malfuncionamiento de tejido en algunas enfermedades, incluyendo el cáncer, en la vida adulta. Los tumores usualmente muestran expresión génica anormal, y un número de tumores no neuronales como el de pulmón, mama y ovario revelan sobre expresión aberrante de β III-tubulina (TUBB3) (Shibazaki et al., 2012).



Figura 13. Contexto genómico del gen TUBB3. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

TUBB3 es una isoforma neural-específica de la tubulina, sin embargo, se encuentra ocasionalmente expresada ectópicamente en ciertas células cancerígenas no neuronales. La proteína TUBB3 es

usualmente utilizada como marcador diferencial de neuronas. Sin embargo parece que su expresión no es absolutamente neuro-específica, hay evidencia de que se expresa en bajos niveles en intestino grueso normal, fibroblastos y keratinocitos. Por lo tanto, es incierto si la expresión ectópica anormal antes mencionada de la proteína TUBB3 en ciertos tumores no neuronales es dependiente de remodelación de la cromatina o des-represión transcripcional y/o activación. Las tubulinas, incluyendo TUBB3, son un constituyente esencial del citoesqueleto celular, formando los microtúbulos que juegan roles cruciales en la división celular (como microtúbulos del huso), y transporte de organelos (como microtúbulos axónicos y dendríticos en neuronas). Aunque los microtúbulos son polímeros de α - y β -tubulinas, formando una estructura cilíndrica larga, esta estructura no es estática, sino que muestra plasticidad a través de la denominada inestabilidad dinámica. Los microtúbulos están involucrados en varios eventos celulares, su rol como componente del huso mitótico siendo particularmente crítico para el crecimiento normal de las células (división celular) y para la plasticidad neuronal (células no divisibles). La alta expresión de TUBB3 en neuronas maduras es la causa para considerarlo como un marcador confiable de diferenciación neuronal. Con respecto al involucramiento fisiológico de TUBB3 en células no neuronales, se expresa en la línea germinal masculina y es esencial para la viabilidad y fertilidad en *Drosophila melanogaster*. Sin embargo, poco se sabe sobre su rol en células de mamífero a excepción de las neuronas maduras (Shibazaki et al., 2012).

Varios tumores no-neuronales como los de ovario, mama y pulmón muestran expresión aberrante de TUBB3. Sin embargo, el detalle de los mecanismos que subyacen a la sobreexpresión de TUBB3 en células cancerígenas no neuronales son ampliamente desconocidos. Se encontró que la expresión de TUBB3 dependiente del ciclo celular está involucrada en el crecimiento celular en células no neuronales y que la expresión temporal de TUBB3 está directamente regulada por REST/NRSF durante el ciclo celular (Shibazaki et al., 2012).

El gen que codifica para MC1R se localiza en el extremo telomérico del cromosoma 16 en un locus que tiene una alta densidad génica. Consecuentemente, solo 2.5 Kb de ADN intergénico separan a MC1R del tramo corriente abajo posicionado en tándem del gen de la beta tubulina III TUBB3. Como muchos genes de receptores acoplados a proteína G, MC1R se expresa primariamente como un gen sin intrones. Debido a su señal de clivaje y poliadenilación inusual, que permite la transcripción continua con el locus de TUBB3, vía splicing alternativo resulta en la expresión de al menos dos ARNm quiméricos conteniendo tanto los siete dominios transmembrana de MC1R como una extensión C-terminal de TUBB3. Dichos transcriptos son humano-específicos y se traducen en proteínas que se localizan en la membrana plasmática y el retículo endoplásmico. Los niveles de expresión de ambos mensajeros, MC1R y MC1R-TUBB3 quimérico, pueden ser modulados al estimular las vías de señalización asociadas con la pigmentación de la dermis en respuesta a la radiación solar (Dalziel et al., 2010).

VDR

12q13.11

La vitamina D es una hormona esteroidea que ejerce la mayor parte de su actividad biológica al unirse a un receptor específico de alta afinidad, el receptor de vitamina D (VDR). VDR pertenece a la superfamilia de receptores nucleares para hormonas esteroideas y regula la expresión génica actuando como un factor de transcripción ligando activado (Lopes, Paredes, Costa, Ylstra, & Schmitt, 2012).

A pesar que la vitamina D esta clásicamente asociada con el rol fisiológico de la regulación del calcio y el transporte de fosfatos en el metabolismo óseo, varios estudios han demostrado su amplio rango de funciones para esta hormona, las cuales son particularmente importantes en el campo del cáncer.

Los mecanismos subyacentes a la acción protectora de la vitamina D en el desarrollo de cáncer son entendidos parcialmente, pero la evidencia muestra que la vitamina D participa en la regulación del crecimiento celular, apoptosis y diferenciación celular. Además ha sido implicada en la supresión de invasión de células cancerígenas, angiogénesis y metástasis.



Figura 14. Contexto genómico del gen VDR. Extraído del GenBank del Nacional Center for Biotechnology Information, disponible online <http://www.ncbi.nlm.nih.gov/gene>

La mayor parte de las acciones biológicas de la vitamina D están mediadas por el receptor de vitamina D y la síntesis y catabolismo de esta hormona están reguladas por las enzimas CYP27B1 y CYP24A1 (Lopes et al., 2012).

El rol de la vitamina D en salud y enfermedad está lejos de ser entendido y la literatura está llena de hallazgos contradictorios, en cáncer es particularmente controversial (Abbas et al., 2008; Barroso et al., 2008; McKay et al., 2009; Rollison et al., 2012).

Sin embargo, existe evidencia de que la vitamina D es capaz de modular varias características del cáncer. Las propiedades anticarcinogénicas de esta hormona incluyen la inhibición de proliferación celular, invasión, metástasis y angiogénesis, y la inducción de apoptosis y diferenciación (Lopes et al., 2012).

Dada la posible relación entre los genes candidatos y el cáncer de mama intentaremos dilucidar si los mismos influyen de alguna manera en el desarrollo de cáncer de mama esporádico en mujeres uruguayas.

Objetivos

Objetivo General

Analizar la posible relación entre los genes candidatos, BRCA1, BRCA2, ESR1, ESR2, FBXO31, TP53, TUBB3, VDR y el desarrollo de cáncer de mama esporádico en una muestra de mujeres uruguayas.

Objetivos Específicos

Analizar dos SNPs del gen BRCA1, cinco del gen BRCA2, cuatro del gen ESR1, tres del gen ESR2, tres del gen FBXO31, tres del gen TP53, dos del gen TUBB3, y cinco del gen VDR.

Determinar si existe asociación entre los SNPs de los genes candidatos y el cáncer de mama en nuestra población.

Analizar los haplotipos de los genes candidatos en relación al desarrollo de cáncer de mama.

Comprobar si existe alguna relación entre los datos epidemiológicos obtenidos y el estatus de salud o enfermedad.

Materiales y Métodos

Población de estudio

Como parte de un proyecto previo, se analizaron los datos obtenidos de una muestra de 205 mujeres afectadas por cáncer de mama esporádico – pacientes – y 216 mujeres sanas – controles. Todas las participantes firmaron un consentimiento informado, completaron las encuestas de datos epidemiológicos y se les extrajo una muestra de sangre periférica. Para este trabajo se seleccionaron las variables epidemiológicas: estatus de fumador, índice de masa corporal, edad de menarca y de menopausia.

El presente trabajo fue aprobado por el Comité de Ética de la Facultad de Medicina.

Selección de SNPs

Sin contar un par de excepciones se intentaron utilizar SNPs neutros, o descriptos con resultados contradictorios en la bibliografía consultada. Utilizamos la estrategia de genes candidatos, con la intención de observar que sucede en nuestra población más allá de confirmar o no los resultados observados en otras poblaciones, poder describir más ampliamente a la nuestra. Los mismos se seleccionaron para complementar un estudio previo donde se analizaron 120 SNPs candidatos de 89 genes asociados a cáncer de mama en poblaciones europeas.

Los 27 SNPs seleccionados se resumen en la Tabla 1.

Ensayo Genotípico

Se lograron genotipar 204 controles y 167 casos para los cuales se obtuvo el panel de SNPs completo mediante el ensayo de genotipado KASP contratando el servicio de KBiosciences, actualmente LGC Genomics.

<http://www.lgcgenomics.com/genotyping/kasp-genotyping-reagents/kasp-overview/>

Como control de calidad de la muestra se comprobó que no hubiera muestras con más de 3 SNPs sin genotipar (10% de los SNPs con los que estamos trabajando).

Tabla 1. Resumen de Genes Candidatos y SNPs seleccionados

SNP ID	CHR	Gen	BP	Región	Alelos
rs827423	6	ESR1	152156197	Intrónica	C/T
rs6912184	6	ESR1	152260206	Intrónica	G/A
rs2982712	6	ESR1	152358179	Intrónica	C/T
rs2273206	6	ESR1	152382311	Intrónica	T/G
rs2239182	12	VDR	48255411	Intrónica	G/A
rs2239179	12	VDR	48257766	Intrónica	G/A
rs2238136	12	VDR	48277713	Intrónica	A/G
rs4237855	12	VDR	48287203	Intrónica	A/G
rs7299460	12	VDR	48296268	Intrónica	T/C
rs9534174	13	BRCA2	32893791	Intrónica	A/G
rs144848	13	BRCA2	32906729	Missense	G/T
rs206081	13	BRCA2	32922136	Intrónica	A/G
rs11571787	13	BRCA2	32958466	Intrónica	G/T
rs4942505	13	BRCA2	32963707	Intrónica	C/T
rs1952585	14	ESR2	64773392	Intrónica	C/T
rs7154455	14	ESR2	64736660	Intrónica	C/G
rs3020449	14	ESR2	64773392	Intrónica	C/T
rs11649550	16	FBXO31	87366770	Intrónica	A/G
rs4843599	16	FBXO31	87383709	Intrónica	A/G
rs9937099	16	FBXO31	87398155	Intrónica	C/T
rs3212363	16	MC1R	89985441	5' UTR	T/A
rs7191944	16	TUBB3	89993878	Intrónica	G/C
rs1042522	17	TP53	7579472	Missense	C/G
rs2078486	17	TP53	7583083	Intrónica	A/G
rs8064946	17	TP53	7589311	Intrónica	C/G
rs8176193	17	BRCA1	41231516	Intrónica	A/G
rs8176092	17	BRCA1	41270229	Intrónica	C/A

Se muestra para cada SNP su código de identificación (ID), el cromosoma (CHR) y gen donde se encuentra, la distancia cromosómica en pares de bases (BP), en qué región del gen se encuentra v sus variantes alélicas siendo el primer alelo el mavoritario v el segundo el minoritario.

Análisis Estadísticos

Estudio de asociación de genes candidatos

Estos estudios generalmente apuntan a relacionar la información de secuencia genética derivada de individuos no relacionados con una medida de la progresión de enfermedad o estatus de enfermedad. Este tipo de estudios tienen un marcado potencial al centrarse directamente sobre las vías metabólicas importantes en ese tejido. Esto supone proponer una hipótesis que se prueba a diferencia de los estudios de asociación donde no hay hipótesis previa.

Es esencial tomar en cuenta las co-variantes como potenciales confusores y modificadores de la enfermedad y su progresión para obtener conclusiones clínicamente significativas (Foulkes, 2009).

Además de capturar información del genotipo y los rasgos, los estudios poblacionales generalmente involucran la colección de otra información en características específicas de los pacientes. La información adicional puede incluir índice de masa corporal (IMC), género, edad y estatus de fumador. La información adicional suele ser de variables que han sido previamente asociadas con los rasgos de interés, y pueden incluir factores ambientales, demográficos y clínicos. Si una variable, está en la ruta causal de la enfermedad, lo correcto es incluirla como co-variable (Foulkes, 2009).

Estos estudios pueden incluir tipificar 5–50 SNPs dentro de un gen (incluyendo secuencias codificantes y flanqueantes, las cuales pueden incluir sitios de splicing o regulatorios). El gen puede ser un candidato posicional que resulte de un estudio de asociación anterior, o un candidato funcional que esté basado, por ejemplo, en homología con un gen de función conocida en alguna especie modelo (Balding, 2006).

Los estudios de genes candidatos generalmente incluyen múltiples SNPs dentro de un mismo gen. La elección de SNPs depende de bloques de desequilibrio de ligamiento (LD) definidos. La premisa subyacente de estos estudios es que los SNPs bajo investigación capturan información sobre la variabilidad genética subyacente del gen en consideración, aunque el SNP pueda no servir como una variante causal de la enfermedad. Esto es, los SNPs que están siendo estudiados no son necesariamente funcionales. Considerando la situación actual, en la que queremos investigar la asociación entre un gen y una enfermedad, en un modelo simple, podemos asumir que un único SNP en la región funcional del gen resulta en la enfermedad. En general, la localización precisa de esta variante no se conoce. En cambio, podemos medir múltiples SNPs que se presumen ‘cerca’ de este sitio en el genoma. El término ‘cerca’ puede ser pensado como distancia física, aunque se deben considerar varias características para elegir SNPs apropiados (Foulkes, 2009).

Estos SNPs próximos son comúnmente denominados marcadores, dado que el genotipo observado en estas localizaciones tiende a estar asociados con el genotipo en el locus causante de la enfermedad. La

idea subyacente a este fenómeno es que, a lo largo del tiempo evolutivo (a lo largo de muchas generaciones de reproducción), el alelo de la enfermedad fue heredado junto con variantes en estos loci marcadores (Figura 15).

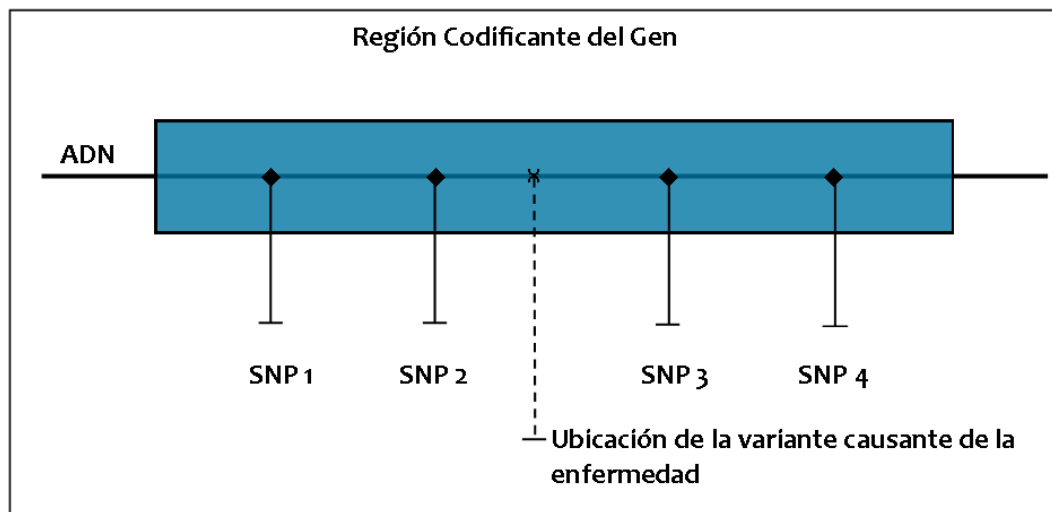


Figura 15. SNPs marcadores. Adaptado de Foulkes, 2009

Esto ocurre cuando la probabilidad de un evento de recombinación para las regiones de ADN entre el locus para la enfermedad y el locus para el marcador es pequeña. Por lo tanto, al capturar la variabilidad en estos loci intentaremos capturar la variabilidad en el locus de la enfermedad (Foulkes, 2009).

Los análisis preliminares de la calidad de los datos son de suma importancia y los datos deben ser chequeados minuciosamente, por ejemplo, para patrones inusuales de falta de datos. Analizar el equilibrio de Hardy-Weinberg (HWE) también puede ser de ayuda, como lo son los análisis que permiten seleccionar un buen subgrupo de los SNPs disponibles o inferir haplotipos desde genotipos (Balding, 2006).

Equilibrio de Hardy-Weinberg

Las desviaciones del HWE pueden deberse a endogamia, estratificación de la población o selección. Desvíos al HWE en la muestra de afectados pueden estar indicando asociación con alguna enfermedad, cuyas implicaciones son usualmente sub-explotadas.

Las desviaciones aparentes del equilibrio de HW pueden surgir en presencia de un polimorfismo común de delección o debido a una tendencia a asignar por omisión heterocigotas como homocigotas. Usualmente se realiza un análisis de HWE primariamente como un chequeo de la calidad de los datos y se descartan loci que, por ejemplo, sus controles se desvían del equilibrio con un nivel de significancia de $p = 10^{-3}$ o 10^{-4} . Dado que en muchos casos este desvío puede ser causado por un problema técnico

del laboratorio, la incapacidad para medir un alelo o hacerlo en forma parcial se ve reflejada en un desequilibrio de HW, por lo cual este análisis nos sirve como un control de calidad de nuestras muestras.

Sin embargo, la posibilidad de que una desviación del HWE sea debido a un polimorfismo de delección o a una duplicación de segmento que podría ser importante en las causas de enfermedad debería ser considerada antes de descartar algún loci.

El análisis de desviaciones del HWE puede ser llevado a cabo usando una prueba de Pearson de bondad de ajuste, usualmente conocida como prueba de χ^2 porque su estadístico tiene una distribución χ^2 aproximadamente nula. El análisis de Pearson es fácil de computar, pero la aproximación de χ^2 puede ser pobre cuando hay conteos genotípicos bajos, y es mejor usar un análisis exacto de Fisher el cual no depende de la aproximación de χ^2 (Balding, 2006).

Pruebas de asociación

El análisis más común de genotipos y de estatus de caso-control para un SNP individual es analizar la hipótesis nula de que no hay asociación entre las filas y las columnas de la matriz de 2x3 que contiene la cuenta de los tres genotipos (dos homocigotas y el heterocigota) entre casos y controles. Entre otros, se pueden utilizar el análisis de Pearson (2 df) o una prueba exacta de Fisher. Este último es preferible cuando algunas de las clases (genotipos) tienen un conteo reducido, es computacionalmente más demandante, pero se implementa en varios software. Para rasgos complejos, se considera que la contribución a la enfermedad de los SNPs individuales será, por lo general, aditiva, o sea, el riesgo para el heterocigota tendrá un valor intermedio entre los riesgos de ambos homocigotas. Los análisis de tipo general (Pearson y Fisher) tienen un alcance razonable a pesar de las desventajas subyacentes, pero si el riesgo asociado a un genotipo es aditivo no serán análisis tan poderosos como aquellos que si tienen en cuenta el riesgo aditivo. Una forma de aumentar el poder de detectar riesgos aditivos es contar alelos en vez de genotipos así cada individuo contribuye dos veces a una tabla de 2x2. Este procedimiento requiere asumir HWE para casos y controles combinados y da como resultado un riesgo estimado difícil de interpretar.

Es posible llevar a cabo pruebas de asociación entre una enfermedad y una variante más allá de las pruebas alélicas básicas (los cuales comparan las frecuencias de los alelos en casos versus controles)

Estas pruebas son:

- Prueba de Cochran-Armitage
- Prueba con modelo Genotípico (2df)
- Prueba con modelo de alelo dominante (1df)
- Acción de gen Recesivo (1df)

El Análisis de Cochran–Armitage es similar al de conteo de alelos, es más conservativo y no se basa en la asunción de HWE. La idea es analizar la hipótesis de pendiente cero para una línea que ajusta con los tres riesgos fenotípicos mejor estimados.

Una ventaja de la prueba de Cochran-Armitage es que no asume equilibrio de Hardy-Weinberg, ya que es el individuo y no el alelo la unidad de análisis (aunque los p valores empíricos basados en permutación también tienen esta propiedad). Es importante recordar que los SNPs que muestran una severa desviación de Hardy-Weinberg usualmente es probable que sean malos SNPs (problemas técnicos del laboratorio) o reflejan estratificación en la muestra, así que es mejor que sean excluidos en la mayoría de los casos (Balding, 2006).

En las pruebas de los modelos alélicos, genotípicas, dominante o recesiva se analiza la relación entre ambos alelos del SNP. La prueba genotípica contrasta los 3 genotipos entre controles y afectados en una tabla de 2 por 3. Los modelos alélico, dominante y recesivo son pruebas para determinar el peso relativo del alelo minoritario entre casos y controles. O sea, si D es el alelo minoritario, y d es el mayoritario:

- En un modelo Alélico: D vs d
- Modelo Dominante: (DD, Dd) versus dd
- Modelo Recesivo: DD versus (Dd, dd)

Con la opción de permutación especificada, se llevará a cabo una prueba de permutación basada en el resultado más significativo de las pruebas alélica, dominante y recesiva. Esto es, para cada SNP, el mejor resultado original será comparado contra la mejor de estas tres pruebas, para ese SNP, por cada replicado. En máxima permutación, esto también va a ser comparado con el mejor resultado de todos los SNPs para el valor empírico. Esto controla el hecho de que hayamos elegido la mejor de tres pruebas correlativas para cada SNP (Purcell, 2010)

Regresión logística

Para resultados de tipo caso-control este es un enfoque más avanzado. El estatus de caso-control no se distribuye de manera normal y las probabilidades predichas pueden caer fuera del rango 0-1. Estos problemas son superados en la regresión logística, en la cual la transformación logit (π) = $\log(\pi / (1 - \pi))$ se aplica a π_i , el riesgo de enfermedad del individuo i-ésimo. El valor logit (π_i) equivale a β_0 , β_1 o β_2 , de acuerdo al genotipo del individuo i (β_1 para heterocigotas). La prueba de radio de verosimilitud de este modelo general, contra la hipótesis nula $\beta_0 = \beta_1 = \beta_2$, tiene 2 df, y para tamaños de muestra grandes es equivalente al análisis de Pearson. La regresión logística ofrece una herramienta flexible que puede

inmediatamente ajustar múltiples SNPs, con la posibilidad de añadir interacciones complejas, epistáticos y ambientales, o co-variantes como el sexo o la edad de aparición de la enfermedad (Balding, 2006).

Métodos basados en haplotipos

Para el análisis de múltiples SNPs de una misma región cromosómica, una estrategia que se ha popularizado es usar haplotipos para tratar de capturar la correlación de los SNPs en regiones de baja recombinación. Los haplotipos pueden capturar los efectos combinados de variantes causales estrechamente vinculadas actuando en cis. Un problema que surge al trabajar con muestras de individuos independientes es que los haplotipos no son observados directamente, en vez de esto, deben ser inferidos y surge incertidumbre al inferir las fases alélicas. Sin embargo, cuando el LD entre marcadores es alto, el nivel de incertidumbre es usualmente bajo. Dadas las asignaciones de haplotipos, el análisis más simple involucra probar para la independencia de filas y columnas en una tabla de contingencia de $2 \times k$, siendo k el número de distintos haplotipos (Balding, 2006).

Acercamientos alternativos pueden estar basados en proporciones de haplotipos estimadas entre casos y controles, sin un haplotipo asignado explícitamente para cada individuo: la prueba de verosimilitud compara las multinomiales para casos y controles con las obtenidas al combinar casos y controles. Un problema con ambos acercamientos es que asumen HWE y que el riesgo para la enfermedad se asume como aditivo. Un enfoque diferente, que lleva a una prueba con menos grados de libertad, es buscar un exceso de haplotipos compartidos entre casos en relación a los controles (Balding, 2006).

La estimación de las frecuencias alélicas, las pruebas de equilibrio de Hardy-Weinberg, de asociación de Cochran-Armitage, Dominante, Recesiva, Genotípica, y Alélica, análisis estándar de caso/control, análisis de haplotipos y la regresión logística se realizaron utilizando el software Plink versión 1.07 (Purcell, 2010).

Las tablas de contingencia y el análisis de los datos epidemiológicos se realizaron utilizando R versión 2.15.1

Resultados

Epidemiología

En la Tabla 2 se observan los resultados del análisis de datos epidemiológicos seleccionados para casos y controles. En los casos de variables continuas se categorizaron las mismas utilizando rangos para poder realizar el análisis de contingencia.

Tabla 2. Distribución de variables seleccionadas en casos y controles

Variable	Controles	Casos	p valor	OR (95% CI)
IMC				
< 25	93 (44%)	57 (32%)	0.0424	
25 - 29.9	66 (31%)	63 (35%)		
> 30	53 (25%)	59 (33%)		
IMC pre menopáusicas				
< 25	37 (44%)	21 (40%)	0.4821	
25 - 29.9	25 (30%)	21 (40%)		
> 30	22 (26%)	11 (20%)		
IMC post menopáusicas				
< 25	56 (44%)	36 (29%)	0.0179	
25 - 29.9	41 (32%)	42 (33%)		
> 30	31 (24%)	48 (38%)		
Edad Menarca				
≤ 13	177 (82%)	163 (82%)	1	1.03 (0.61- 1.75)
> 13	39 (18%)	37 (18%)		
Menopausia				
≤ 50	93 (73%)	102 (72%)	0.8913	1.07 (0.61-1.90)
>50	34 (27%)	40 (28%)		
Fumador				
No	112 (53%)	116 (58%)	0.2755	0.79 (0.53 - 1.19)
Si	101 (47%)	83 (42%)		

En negrita se destacan los valores de p menores a 0.05. IMC índice de masa corporal - OR Odds Ratio - CI Intervalo de confianza. Los números de las diferentes variables no suman el total de la muestra reportada debido a falta de información.

IMC, índice de masa corporal, definido como el peso en kg sobre el cuadrado de la altura en m, se analizan los valores categorizados como < 25 , $25 - 29.9$ y > 30 , correspondientes a peso normal, sobrepeso y obesidad respectivamente. Estos son puntos de corte utilizados internacionalmente como se describe en el trabajo de Milne et al., 2010.

La edad de la menarca se divide en dos rangos, ≤ 13 y > 13 , tomando como referencia el trabajo de Modugno et al., 2005.

En el caso de la edad de comienzo de la menopausia, se distinguen aquellas con edades ≤ 50 o > 50 , de acuerdo a Gold et al., 2004.

Para el status de fumador se toma como fumador aquella persona que fumó de manera activa, en el pasado o en el presente, y como no fumador a aquellas que nunca fumaron de forma activa (Henríquez-Hernández et al., 2009; Modugno et al., 2005).

Se observan resultados significativos correspondientes al IMC para la muestra completa y las mujeres post-menopáusicas. Para la muestra completa, con un p valor de 0.04244, para los percentiles correspondientes a sobrepeso y obesidad hay un mayor porcentaje de casos, mientras que para el IMC correspondiente a peso normal el porcentaje es mayor para los controles. Al observar IMC en mujeres post-menopáusicas se acentúan las diferencias con un mayor porcentaje de sobrepeso en los casos con respecto a los controles.

Equilibrio de Hardy-Weinberg

En las tablas 3-8 se observan los resultados del análisis de equilibrio de Hardy-Weinberg para cada SNP teniendo en cuenta los casos (Casos) los controles (Controles) y la muestra completa (Todos). Cada tabla recoge los resultados de un gen, o de más de uno en caso de encontrarse ambos en el mismo cromosoma.

Como se observa en la Tabla 3 todos los SNPs de ESR1 se encuentran en equilibrio de HW.

Para el gen VDR, únicamente el rs4237855 se encuentra fuera del HWE, tanto para la muestra completa, los casos y los controles, con valores de p de $1.01E-08$, $2.31E-02$ y $1.71E-04$, respectivamente. En todos los casos se observan menos heterocigotas de los esperados (Tabla 4).

Tabla 3. Equilibrio de Hardy-Weinberg para los SNPs de ESR1

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
<i>rs827423</i>						
	Todos	C	T	0.5041	0.4892	0.5948
	Casos	C	T	0.5152	0.4982	0.755
	Controles	C	T	0.495	0.4763	0.6581
<i>rs6912184</i>						
	Todos	G	A	0.3068	0.3361	0.1176
	Pacientes	G	A	0.3091	0.3236	0.6288
	Controles	G	A	0.305	0.346	0.1016
<i>rs2982712</i>						
	Todos	C	T	0.4781	0.4832	0.8294
	Pacientes	C	T	0.497	0.4903	1
	Controles	C	T	0.4627	0.476	0.7668
<i>rs2273206</i>						
	Todos	T	G	0.1549	0.1565	0.7394
	Pacientes	T	G	0.1506	0.1494	1
	Controles	T	G	0.1584	0.1623	0.6596

MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) – Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – p valor

De los resultados resumidos en la

Tabla 5, todos los SNPs de BRCA2 se encuentran en equilibrio, con excepción de rs206081. Para este último se observan valores de p significativos para la muestra completa y los controles, lo que indica que ambas muestras se encuentran en desequilibrio de HW. En ambos casos se observa una disminución en el número de heterocigotas observado.

Tabla 4. Equilibrio de Hardy-Weinberg para los SNPs de VDR

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
rs2239182						
	Todos	G	A	0.4698	0.4983	0.2929
	Pacientes	G	A	0.4847	0.4997	0.7539
	Controles	G	A	0.4577	0.4964	0.2595
rs2239179						
	Todos	G	A	0.4609	0.4845	0.3827
	Pacientes	G	A	0.5	0.4912	0.8724
	Controles	G	A	0.43	0.4779	0.1823
rs2238136						
	Todos	A	G	0.3278	0.3512	0.2306
	Pacientes	A	G	0.3415	0.3494	0.8228
	Controles	A	G	0.3166	0.3527	0.159
rs4237855						
	Todos	A	G	0.3158	0.4983	1,01E-08
	Pacientes	A	G	0.3289	0.4997	2,31E-02
	Controles	A	G	0.3053	0.492	1,71E-04
rs7299460						
	Todos	T	C	0.4027	0.411	0.703
	Pacientes	T	C	0.3758	0.3967	0.5552
	Controles	T	C	0.425	0.422	1

En negrita se destacan los resultados con un p valor menor a 0.05. MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) – Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – p valor

En el caso de los SNPs de ESR2 observamos que rs1952585 se encuentra en equilibrio para todos los grupos de muestras. En el caso de rs7154455 los controles no cumplen con el HWE, observándose menos heterocigotas de los esperados. En el caso de rs3020449 tanto para la muestra completa como para los controles no se cumple el HWE y en ambos casos se observa disminución de heterocigotas respecto a lo esperado si se cumpliera con el equilibrio HW. (Tabla 6)

Tabla 5. Equilibrio de Hardy-Weinberg para los SNPs de BRCA2

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
<i>rs9534174</i>						
	Todos	A	G	0.4724	0.4991	0.343
	Pacientes	A	G	0.4785	0.5	0.6383
	Controles	A	G	0.4673	0.4975	0.3947
<i>rs144848</i>						
	Todos	G	T	0.3819	0.3948	0.5095
	Pacientes	G	T	0.4172	0.4205	1
	Controles	G	T	0.3532	0.3712	0.4537
<i>rs206081</i>						
	Todos	A	G	0.2548	0.3193	0.0002421
	Pacientes	A	G	0.2857	0.3185	0.2134
	Controles	A	G	0.23	0.32	0.0002146
<i>rs11571787</i>						
	Todos	G	T	0.3424	0.318	0.1885
	Pacientes	G	T	0.3473	0.3023	0.07163
	Controles	G	T	0.3383	0.3306	0.8339
<i>rs4942505</i>						
	Todos	C	T	0.4834	0.4995	0.5295
	Pacientes	C	T	0.497	0.4995	1
	Controles	C	T	0.4721	0.4995	0.4758

En *negrita* se destacan los resultados con un *p* valor menor a 0.05.
MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) - Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – *p* valor

Tabla 6. Equilibrio de Hardy-Weinberg para los SNPs de ESR2

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
<i>rs1952585</i>						
	Todos	C	T	0.2514	0.2625	0.4268
	Pacientes	C	T	0.2395	0.2546	0.3763
	Controles	C	T	0.2611	0.2689	0.6075
<i>rs7154455</i>						
	Todos	C	G	0.4153	0.4548	0.1072
	Pacientes	C	G	0.4573	0.4624	0.8671
	Controles	C	G	0.3812	0.4482	0.04056
<i>rs3020449</i>						
	Todos	C	T	0.4402	0.4967	0.02809
	Pacientes	C	T	0.4639	0.4959	0.4343
	Controles	C	T	0.4208	0.4972	0.03347

En *negrita* se destacan los resultados con un *p* valor menor a 0.05.
MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) – Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – *p* valor

Cuando observamos los resultados del análisis de HWE realizado con los SNPs de FBXO31 observamos que únicamente la muestra de casos para rs11649550 se encuentra fuera del equilibrio, con una representación menor de los heterocigotas (Tabla 7).

En el caso de rs3212363, SNP de MC1R, tanto la muestra completa como la muestra control se encuentran fuera del equilibrio, y se observa la misma tendencia que para los SNPs anteriores, en la que los heterocigotas observados se encuentran en menos proporción de la esperada (Tabla 7).

Para rs7191944, SNP de TUBB3, todos los sets de muestras se encuentran en equilibrio (Tabla 7).

Tabla 7. Equilibrio de Hardy-Weinberg para los SNPs de FBXO31, MC1R y TUBB3

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
<i>rs11649550</i>						
	Todos	A	G	0.3939	0.4268	0.1393
	Pacientes	A	G	0.325	0.4149	0.007363
	Controles	A	G	0.4495	0.4357	0.7452
<i>rs4843599</i>						
	Todos	A	G	0.2452	0.2393	0.8273
	Pacientes	A	G	0.2012	0.2188	0.284
	Controles	A	G	0.2808	0.2553	0.2662
<i>rs9937099</i>						
	Todos	C	T	0.518	0.4932	0.3932
	Pacientes	C	T	0.494	0.4912	1
	Controles	C	T	0.5385	0.4947	0.2488
<i>rs3212363</i>						
	Todos	T	A	0.3674	0.4198	0.01773
	Pacientes	T	A	0.4012	0.4424	0.2853
	Controles	T	A	0.34	0.3987	0.04955
<i>rs7191944</i>						
	Todos	G	C	0.4238	0.4448	0.4071
	Pacientes	G	C	0.4596	0.4574	1
	Controles	G	C	0.395	0.4334	0.2522

En negrita se destacan los resultados con un p valor menor a 0.05. MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) - Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – p valor

Tanto los SNPs de TP53 como los de BRCA1 se encuentran en equilibrio de Hardy Weinberg como se observa en la Tabla 8.

Tabla 8. Equilibrio de Hardy-Weinberg para los SNPs de TP53 y BRCA1

SNP	MUESTRA	A1	A2	O(HET)	E(HET)	P
<i>rs1042522</i>						
	Todos	C	G	0.4254	0.3871	0.07678
	Pacientes	C	G	0.4568	0.4012	0.1154
	Controles	C	G	0.4	0.375	0.4506
<i>rs2078486</i>						
	Todos	A	G	0.1703	0.1647	0.7539
	Pacientes	A	G	0.1916	0.1829	1
	Controles	A	G	0.1527	0.1493	1
<i>rs8064946</i>						
	Todos	C	G	0.276	0.2611	0.4206
	Pacientes	C	G	0.297	0.2936	1
	Controles	C	G	0.2587	0.2326	0.2161
<i>rs8176193</i>						
	Todos	A	G	0.4417	0.4575	0.5643
	Pacientes	A	G	0.4074	0.4524	0.2242
	Controles	A	G	0.4697	0.4614	0.8781
<i>rs8176092</i>						
	Todos	C	A	0.4345	0.4588	0.3032
	Pacientes	C	A	0.4085	0.4554	0.2288
	Controles	C	A	0.4564	0.4617	0.8773

MUESTRA – Indica si se analizaron todas las muestras (Todos), las de los casos (Pacientes) o las de controles (Controles) A1 – Alelo minoritario A2 – Alelo mayoritario O (HET) – Frecuencia de heterocigotas observada E (HET) – Frecuencia de heterocigotas esperada P – p valor

Frecuencias y asociaciones alélicas

Las frecuencias alélicas para cada SNP, el valor de p para la comparación entre casos y controles y el OR correspondiente se presentan en la Tabla 9.

Tabla 9. Análisis del modelo alélico por SNP para casos y controles

SNP	A1	F_P	F_C	A2	P	OR
rs827423	C	0.4697	0.3911	T	0.03571	1.379
rs6912184	G	0.203	0.2225	A	0.5864	0.8902
rs2982712	C	0.4303	0.3905	T	0.2906	1.179
rs2273206	T	0.08133	0.08911	G	0.7915	0.9049
rs2239182	G	0.4877	0.4577	A	0.4554	1.128
rs2239179	G	0.4335	0.395	A	0.3205	1.172
rs2238136	A	0.2256	0.2286	G	0.9294	0.9829
rs4237855	A	0.5132	0.4368	G	0.05393	1.359
rs7299460	T	0.2727	0.3025	C	0.4122	0.8647
rs9534174	A	0.4969	0.4648	G	0.4112	1.137
rs144848	G	0.3006	0.2463	T	0.1111	1.316
rs206081	A	0.1988	0.2	G	1	0.9922
rs11571787	G	0.1856	0.209	T	0.4583	0.8629
rs4942505	C	0.4848	0.4848	T	1	1
rs1952585	C	0.1497	0.1601	T	0.7599	0.9236
rs7154455	C	0.3628	0.3391	G	0.5332	1.11
rs3020449	C	0.4548	0.4629	T	0.8818	0.9681
rs11649550	A	0.2938	0.3207	G	0.4644	0.881
rs4843599	A	0.125	0.1502	G	0.3364	0.808
rs9937099	C	0.4337	0.4487	T	0.7073	0.941
rs3212363	T	0.3302	0.275	A	0.1211	1.3
rs7191944	G	0.354	0.3175	C	0.3034	1.178
rs1042522	C	0.2778	0.25	G	0.4446	1.154
rs2078486	A	0.1018	0.08128	G	0.3684	1.281
rs8064946	C	0.1788	0.1343	G	0.1011	1.403
rs8176193	A	0.3457	0.3611	G	0.6957	0.9347
rs8176092	C	0.3506	0.3615	A	0.8145	0.9535

En negrita se destacan los p valores menores a 0.05. SNP – Nombre del SNP A1 – Alelo minoritario F_P – Frecuencia de A1 en pacientes F_C – Frecuencia de A1 en controles A2 – Alelo mayoritario P – p valor OR – Odds Ratios

Como se desprende de la Tabla 9 solamente para rs827423 la diferencia entre las frecuencias del alelo minoritario entre casos y controles es significativa, con un p de 0.03571 y un OR de 1.379.

Pruebas de Asociación

En las tablas 10-15 se observan los valores de p obtenidos de las pruebas de asociación utilizando los modelos Genotípico (GENO), Cochran-Armitage (TREND), Dominante (DOM) y Recesivo (REC). En cada tabla se recogen los resultados de un gen, o de más de uno en caso de encontrarse en el mismo cromosoma.

Únicamente utilizando el modelo de Cochran-Armitage para la asociación de rs827423 obtenemos un p valor significativo, como se observa en la Tabla 10.

Tabla 10. Valores de probabilidad de las Pruebas de Asociación para SNPs de ESR1

MODELO	SNPs			
	rs827423	rs6912184	rs2982712	rs2273206
GENO	0.09522	0.714	0.5057	0.9531
TREND	0.02959	0.5401	0.2788	0.7087
DOM	0.07349	0.7446	0.273	0.7788
REC	0.09746	0.5086	0.5783	1

En negrita se destacan los resultados con un p valor menor a 0.05. MODELO – Indica la prueba utilizada siendo GENO - genotípico TREND - Cochran-Armitage DOM - dominante REC - recesivo

Tabla 11. Valores de probabilidad de las Pruebas de Asociación para SNPs de VDR

MODELO	SNPs				
	rs2239182	rs2239179	rs2238136	rs4237855	rs7299460
GENO	0.6775	0.3196	0.7804	0.2189	0.5791
TREND	0.4326	0.3096	0.9251	0.08915	0.3819
DOM	0.4183	0.1826	0.914	0.1147	0.3439
REC	0.7119	1	0.6666	0.2405	1

En negrita se destacan los resultados con un p valor menor a 0.05. MODELO – Indica la prueba utilizada siendo GENO - genotípico TREND - Cochran-Armitage DOM - dominante REC - recesivo

Tabla 12. Valores de probabilidad de las Pruebas de Asociación para SNPs de BRCA2

MODELO	SNPs				
	rs9534174	rs144848	rs206081	rs11571787	rs4942505
GENO	0.7108	0.2431	0.3347	0.2931	0.9066
TREND	0.4019	0.1063	0.9698	0.411	0.9984
DOM	0.4828	0.1131	0.6519	0.7455	0.8137
REC	0.6226	0.4437	0.3134	0.1196	0.8067

En negrita se destacan los resultados con un p valor menor a 0.05. MODELO – Indica la prueba utilizada siendo GENO - genotípico TREND - Cochran-Armitage DOM - dominante REC - recesivo

Tabla 13. Valores de probabilidad de las Pruebas de Asociación para SNPs de ESR2

MODELO	SNPs		
	rs1952585	rs7154455	rs3020449
GENO	0.9321	0.3423	0.6789
TREND	0.7036	0.5214	0.8363
DOM	0.7275	0.2462	0.8228
REC	1	0.7642	0.5407

En negrita se destacan los resultados con un p valor menor a 0.05.
 MODELO – Indica la prueba utilizada siendo GENO - genotípico
 TREND - Cochran-Armitage DOM - dominante REC - recesivo

Como se observa en las Tablas 11, 12 y 13, ninguno de los SNPs de VDR, BRCA2 y ESR2 arrojó resultados significativos para ninguno de los modelos de asociación utilizados.

Como se observa en la Tabla 14 la única prueba de asociación que resultó significativa fue la correspondiente a rs11649550 cuando se utiliza el modelo genotípico.

Tabla 14. Valores de probabilidad de las Pruebas de Asociación para SNPs de FBXO31, MC1R y TUBB3

MODELO	SNPs				
	rs11649550	rs4843599	rs9937099	rs3212363	rs7191944
GENO	0.04996	0.1316	0.6715	0.2694	0.4004
TREND	0.4544	0.3195	0.6784	0.1282	0.3119
DOM	0.1107	0.1889	0.4894	0.1133	0.2032
REC	0.3147	0.4137	0.8919	0.5111	1

En negrita se destacan los resultados con un p valor menor a 0.05. MODELO – Indica la prueba utilizada siendo GENO - genotípico TREND - Cochran-Armitage DOM - dominante REC - recesivo

Tabla 15. Valores de probabilidad de las Pruebas de Asociación para SNPs de TP53 y BRCA1

MODELO	SNPs				
	rs1042522	rs2078486	rs8064946	rs8176193	rs8176092
GENO	0.5222	0.6689	0.1094	0.4877	0.6646
TREND	0.3735	0.3248	0.08803	0.6719	0.7666
DOM	0.2923	0.3385	0.2045	0.3926	0.5219
REC	1	1	0.09464	0.7556	0.7609

En negrita se destacan los resultados con un p valor menor a 0.05. MODELO – Indica la prueba utilizada siendo GENO - genotípico TREND - Cochran-Armitage DOM - dominante REC - recesivo

En la Tabla 15 se observan los resultados de las pruebas de asociación para TP53 y BRCA1, de las cuales ninguna fue significativa.

Asociación Haplotípica

En las tablas 16-23 se observan los resultados del análisis de asociación para los haplotipos formados por dos, tres o cuatro SNPs a la vez, de un único gen o de dos genes que se encuentren a una distancia razonable para formar un haplotipo.

Analizando los haplotipos formados por los SNPs de ESR1 observamos que aquellos haplotipos asociados de manera significativa con cáncer de mama son tres. El formado por el alelo C del SNP rs827423 y el alelo A del SNP rs6912184, el haplotipo formado por los dos SNPs anteriores y el alelo C de rs2982712, y el haplotipo conteniendo a éstos y al alelo G de rs2273206, con valores de p decrecientes en significancia (0.03618, 0.04653 y 0.04979 respectivamente), ver Tabla 16.

En el caso de los SNPs de VDR un único haplotipo dio significativo, el formado por el alelo A de rs4237855 y el alelo C de rs4237855 como se observa en la Tabla 17.

Como se observa en las tablas 18 – 22 ningún otro haplotipo además de los previamente mencionados arrojó resultados significativos. Ninguno de los haplotipos formados por los SNPs de BRCA2, ESR2, FBXO31, TUBB3, MC1R y BRCA1 se encuentra asociado significativamente con cáncer de mama.

Tabla 16. Prueba de Asociación Haplotípica para SNPs de ESR1

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs827423/rs6912184</i>						
	CG	0.06942	0.06412	0.08079	1	0.7762
	TG	0.1361	0.1606	0.8454	1	0.3579
	CA	0.3999	0.3248	4.389	1	0.03618
	TA	0.3946	0.4505	2.287	1	0.1305
<i>rs6912184/rs2982712</i>						
	GC	0.1321	0.1327	0.0005758	1	0.9809
	AC	0.2974	0.2613	1.163	1	0.2808
	GT	0.07346	0.09207	0.8084	1	0.3686
	AT	0.4971	0.514	0.2041	1	0.6514
<i>rs2982712/rs2273206</i>						
	CT	0.07892	0.08934	0.2521	1	0.6156
	CG	0.3495	0.3034	1.74	1	0.1871
	TG	0.5716	0.6072	0.9442	1	0.3312

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs827423/rs6912184/rs2982712</i>						
	CGC	0.0497	0.04464	0.1046	1	0.7464
	TGC	0.082	0.08638	0.04522	1	0.8316
	CAC	0.1664	0.1155	3.962	1	0.04653
	TAC	0.1317	0.1437	0.2199	1	0.6391
	CGT	0.0188	0.01761	0.01455	1	0.904
	TGT	0.054	0.07436	1.241	1	0.2653
	CAT	0.2348	0.2106	0.6192	1	0.4314
	TAT	0.2626	0.3073	1.775	1	0.1827
<i>rs6912184/rs2982712/rs2273206</i>						
	GCT	0.03755	0.0465	0.3604	1	0.5483
	ACT	0.04192	0.04142	0.001129	1	0.9732
	GCG	0.09405	0.08414	0.222	1	0.6376
	ACG	0.257	0.2181	1.532	1	0.2158
	GTG	0.06929	0.09017	1.074	1	0.3
	ATG	0.5002	0.5196	0.2756	1	0.5996
<i>rs827423/rs6912184/rs2982712/rs2273206</i>						
	CGCT	0.02682	0.025	0.02393	1	0.8771
	TGCT	0.01258	0.02152	0.8464	1	0.3576
	CACT	0.02444	0.019	0.2564	1	0.6126
	TACT	0.01753	0.02256	0.2313	1	0.6306
	CGCG	0.02288	0.02056	0.04635	1	0.8295
	TGCG	0.07035	0.06448	0.09994	1	0.7519
	CACG	0.1437	0.09676	3.848	1	0.04979
	TACG	0.1129	0.1207	0.107	1	0.7436
	CGTG	0.02015	0.01823	0.03553	1	0.8505
	TGTG	0.04974	0.07209	1.56	1	0.2116
	CATG	0.2334	0.2092	0.6167	1	0.4323
	TATG	0.2656	0.3099	1.733	1	0.1881

En negrita se destacan los resultados con un p valor menor a 0.05. SNPS: SNPs en cada haplotipo
HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en
controles CHISQ: valor de χ^2 DF: grados de libertad P: p valor

Tabla 17. Prueba de Asociación Haplotípica para SNPs de VDR

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs2239182/rs2239179</i>						
	GG	0.4146	0.3858	0.6055	1	0.4365
	AG	0.01899	0.01269	0.4559	1	0.4996
	GA	0.07278	0.07614	0.02861	1	0.8657
	AA	0.4937	0.5254	0.7057	1	0.4009
<i>rs2239179/rs2238136</i>						
	GA	0.1003	0.08276	0.648	1	0.4208
	AA	0.1241	0.148	0.8396	1	0.3595
	GG	0.3292	0.3019	0.6008	1	0.4383
	AG	0.4465	0.4674	0.3054	1	0.5805
<i>rs2238136/rs4237855</i>						
	AA	0.03203	0.0241	0.3882	1	0.5333
	GA	0.4813	0.4114	3.291	1	0.06965
	AG	0.1946	0.2151	0.427	1	0.5135
	GG	0.292	0.3494	2.494	1	0.1143
<i>rs4237855/rs7299460</i>						
	AT	0.1196	0.1305	0.18	1	0.6713
	GT	0.1037	0.1422	2.254	1	0.1333
	AC	0.397	0.3106	5.468	1	0.01937
	GC	0.3796	0.4166	0.9493	1	0.3299
<i>rs2239182/rs2239179/rs2238136</i>						
	GGA	0.1047	0.08702	0.6565	1	0.4178
	AAA	0.1215	0.1414	0.6201	1	0.431
	GGG	0.3092	0.2906	0.2993	1	0.5843
	AGG	0.02036	0.01404	0.4363	1	0.5089
	GAG	0.07386	0.07675	0.02177	1	0.8827
	AAG	0.3704	0.3902	0.3019	1	0.5827
<i>rs2239179/rs2238136/rs4237855</i>						
	AAA	0.02161	0.01764	0.1479	1	0.7005
	GGA	0.2312	0.2007	0.9907	1	0.3196
	AGA	0.2521	0.2197	1.051	1	0.3052
	GAG	0.09536	0.08273	0.3542	1	0.5517
	AAG	0.1011	0.125	1.011	1	0.3146
	GGG	0.09698	0.1074	0.2117	1	0.6455
	AGG	0.2016	0.2468	2.087	1	0.1486
<i>rs2238136/rs4237855/rs7299460</i>						
	AAT	0.01314	0.009674	0.1964	1	0.6576
	GAT	0.1275	0.1368	0.1348	1	0.7135
	AGT	0.03184	0.0275	0.1189	1	0.7302
	GGT	0.09565	0.1269	1.774	1	0.1829
	AAC	0.01907	0.01231	0.548	1	0.4591
	GAC	0.3507	0.2847	3.661	1	0.05569
	AGC	0.1617	0.1794	0.3995	1	0.5273
	GGC	0.2004	0.2227	0.5388	1	0.4629

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
rs2239182/rs2239179/rs2238136/rs4237855						
	AAAA	0.01944	0.01748	0.03794	1	0.8456
	GGGA	0.2207	0.1861	1.325	1	0.2497
	AGGA	0.01521	0.01272	0.08074	1	0.7763
	GAGA	0.04673	0.02786	1.813	1	0.1781
	AAGA	0.2041	0.1943	0.1062	1	0.7445
	GGAG	0.09577	0.08336	0.3376	1	0.5612
	AAAG	0.09855	0.1229	1.061	1	0.303
	GGGG	0.08967	0.1025	0.3337	1	0.5635
	GAGG	0.02775	0.04851	2.026	1	0.1546
	AAGG	0.182	0.2042	0.5571	1	0.4554
rs2239179/rs2238136/rs4237855/rs7299460						
	GGAT	0.05521	0.06257	0.1711	1	0.6791
	AGAT	0.07401	0.07472	0.001302	1	0.9712
	AAGT	0.02502	0.02528	0.0004806	1	0.9825
	GGGT	0.04404	0.04252	0.009815	1	0.9211
	AGGT	0.05687	0.08756	2.425	1	0.1194
	AAAC	0.01339	0.01186	0.03325	1	0.8553
	GGAC	0.1837	0.1414	2.337	1	0.1263
	AGAC	0.1777	0.1457	1.344	1	0.2463
	GAGC	0.09255	0.08287	0.2074	1	0.6488
	AAGC	0.07519	0.1004	1.382	1	0.2398
	GGGC	0.04897	0.06202	0.5654	1	0.4521
	AGGC	0.1534	0.1631	0.1239	1	0.7249

En negrita se destacan los resultados con un p valor menor a 0.05. SNPS: SNPs en cada haplotipo
HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en
controles CHISQ: valor de X^2 DF: grados de libertad P: p valor

Tabla 18. Prueba de Asociación Haplotípica para SNPs de BRCA2

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs9534174/rs144848</i>						
	AG	0.2885	0.248	1.463	1	0.2264
	AT	0.2093	0.216	0.04667	1	0.829
	GT	0.5022	0.536	0.7972	1	0.3719
<i>rs144848/rs206081</i>						
	TA	0.1896	0.1947	0.02908	1	0.8646
	GG	0.2977	0.2506	1.96	1	0.1615
	TG	0.5127	0.5547	1.24	1	0.2654
<i>rs206081/rs11571787</i>						
	GG	0.1903	0.2094	0.4012	1	0.5265
	AT	0.1965	0.1992	0.007916	1	0.9291
	GT	0.6132	0.5914	0.3488	1	0.5548
<i>rs11571787/rs4942505</i>						
	TC	0.4848	0.4923	0.03929	1	0.8429
	GT	0.1879	0.2113	0.6122	1	0.434
	TT	0.3273	0.2964	0.7942	1	0.3728
<i>rs9534174/rs144848/rs206081</i>						
	ATA	0.1997	0.1958	0.01717	1	0.8958
	AGG	0.2898	0.2493	1.491	1	0.222
	ATG	0.01688	0.02133	0.1861	1	0.6661
	GTG	0.4937	0.5336	1.139	1	0.2859
<i>rs144848/rs206081/rs11571787</i>						
	TGG	0.1867	0.209	0.566	1	0.4518
	TAT	0.1971	0.2005	0.013	1	0.9092
	GGT	0.297	0.2453	2.479	1	0.1154
	TGT	0.3192	0.3452	0.557	1	0.4555
<i>rs206081/rs11571787/rs4942505</i>						
	ATC	0.1996	0.1996	1,80E-04	1	0.9997
	GTC	0.2887	0.2867	0.003481	1	0.9529
	GGT	0.1862	0.2113	0.7162	1	0.3974
	GTT	0.3255	0.3024	0.4517	1	0.5015
<i>rs9534174/rs144848/rs206081/rs11571787</i>						
	GTGG	0.1899	0.2137	0.6246	1	0.4293
	ATAT	0.1998	0.1958	0.01807	1	0.8931
	AGGT	0.2897	0.2493	1.49	1	0.2223
	ATGT	0.01684	0.02154	0.2065	1	0.6495
	GTGT	0.3037	0.3197	0.2114	1	0.6457
<i>rs144848/rs206081/rs11571787/rs4942505</i>						
	TATC	0.1979	0.2027	0.02645	1	0.8708
	GGTC	0.2177	0.1971	0.4689	1	0.4935
	TGTC	0.07026	0.08628	0.6373	1	0.4247
	TGGT	0.1886	0.2127	0.6538	1	0.4188
	GGTT	0.07843	0.05171	2.165	1	0.1412
	TGTT	0.2471	0.2494	0.005282	1	0.9421

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de X^2 DF: grados de libertad P: p valor

Tabla 19. Prueba de Asociación Haplotípica para SNPs de ESR2

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs1952585/rs7154455</i>						
	CC	0.1059	0.1038	0.008296	1	0.9274
	TC	0.2569	0.2345	0.4914	1	0.4833
	CG	0.04657	0.0579	0.4634	1	0.4961
	TG	0.5906	0.6038	0.1303	1	0.7181
<i>rs7154455/rs3020449</i>						
	CC	0.3571	0.3367	0.3293	1	0.5661
	GC	0.09627	0.1281	1.791	1	0.1808
	GT	0.5466	0.5352	0.09326	1	0.7601
<i>rs1952585/rs7154455/rs3020449</i>						
	CCC	0.1071	0.1063	0.001115	1	0.9734
	TCC	0.2544	0.23	0.5909	1	0.4421
	CGC	0.04177	0.0556	0.7406	1	0.3895
	TGC	0.05902	0.0719	0.4876	1	0.485
	TGT	0.5377	0.5361	0.001772	1	0.9664

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de X^2 DF: grados de libertad P: p valor

Tabla 20. Prueba de Asociación Haplotípica para SNPs de FBXO31

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs11649550/rs4843599</i>						
	AA	0.02033	0.0187	0.02471	1	0.8751
	GA	0.1094	0.131	0.7696	1	0.3803
	AG	0.2676	0.3011	0.961	1	0.3269
	GG	0.6026	0.5492	2.048	1	0.1524
<i>rs4843599/rs9937099</i>						
	AC	0.05172	0.05921	0.1889	1	0.6639
	GC	0.3839	0.3918	0.0472	1	0.828
	AT	0.07098	0.09285	1.115	1	0.291
	GT	0.4934	0.4561	0.9896	1	0.3198
<i>rs11649550/rs4843599/rs9937099</i>						
	GAC	0.04495	0.05196	0.1921	1	0.6612
	AGC	0.05755	0.06915	0.4087	1	0.5226
	GGC	0.3273	0.3259	0.001601	1	0.9681
	AAT	0.01252	0.01359	0.01607	1	0.8991
	GAT	0.06123	0.08196	1.159	1	0.2817
	AGT	0.2208	0.235	0.2062	1	0.6497
	GGT	0.2757	0.2225	2.766	1	0.0963

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de X^2 DF: grados de libertad P: p valor

Tabla 21. Prueba de Asociación Haplotípica para SNPs de MC1R y TUBB3

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs3212363/rs7191944</i>						
	TG	0.293	0.2576	1.106	1	0.2929
	AG	0.06369	0.06061	0.02873	1	0.8654
	TC	0.03185	0.01768	1.505	1	0.2199
	AC	0.6115	0.6641	2.111	1	0.1462

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de χ^2 DF: grados de libertad P: p valor

Tabla 22. Prueba de Asociación Haplotípica para SNPs de TP53

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs1042522/rs2078486</i>						
	CA	0.0838	0.05807	1.83	1	0.1761
	GA	0.02114	0.02193	0.005293	1	0.942
	CG	0.194	0.1919	0.00483	1	0.9446
	GG	0.7011	0.7281	0.6411	1	0.4233
<i>rs8176193/rs8176092</i>						
	AC	0.3428	0.3568	0.1499	1	0.6987
	GA	0.6572	0.6432	0.1499	1	0.6987
<i>rs1042522/rs2078486/rs8064946</i>						
	CAC	0.08152	0.0601	1.296	1	0.2549
	GAC	0.02027	0.02118	0.007461	1	0.9312
	CGC	0.04947	0.04026	0.3663	1	0.5451
	GGC	0.02568	0.01283	1.655	1	0.1983
	CGG	0.1437	0.15	0.05822	1	0.8093
	GGG	0.6794	0.7156	1.146	1	0.2844

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de χ^2 DF: grados de libertad P: p valor

Tabla 23. Prueba de Asociación Haplotípica para SNPs de BRCA1

SNPS	HAPLOTIPO	F_P	F_C	CHISQ	DF	P
<i>rs2078486/rs8064946</i>						
	AC	0.103	0.08	1.166	1	0.2802
	GC	0.07576	0.0525	1.657	1	0.1981
	GG	0.8212	0.8675	2.983	1	0.08415

SNPS: SNPs en cada haplotipo HAPLOTIPO: haplotipo analizado F_P: frecuencia del haplotipo en casos F_C: frecuencia del haplotipo en controles CHISQ: valor de χ^2 DF: grados de libertad P: p valor

Regresión logística

En la Tabla 24 se observa el resultado del análisis de asociación utilizando el modelo de regresión logística utilizando como co-variable los datos de IMC.

Tabla 24. Regresión logística utilizando el IMC como co-variable

SNP	A1	MUESTRA	NMISS	OR	STAT	P
<i>rs827423</i>						
	C	ADD	328	1.428	2.142	0.03222
	C	IMC	328	1.393	2.374	0.0176
<i>rs6912184</i>						
	G	ADD	327	0.8934	-0.5931	0.5531
	G	IMC	327	1.327	2.047	0.04066
<i>rs2982712</i>						
	C	ADD	327	1.162	0.9062	0.3648
	C	IMC	327	1.349	2.144	0.03201
<i>rs2273206</i>						
	T	ADD	330	0.9802	-0.07107	0.9433
	T	IMC	330	1.349	2.178	0.02941
<i>rs2239182</i>						
	G	ADD	327	1.064	0.3949	0.6929
	G	IMC	327	1.348	2.156	0.03105
<i>rs2239179</i>						
	G	ADD	323	1.21	1.185	0.236
	G	IMC	323	1.341	2.099	0.0358
<i>rs2238136</i>						
	A	ADD	325	1.09	0.4721	0.6368
	A	IMC	325	1.359	2.223	0.02622
<i>rs4237855</i>						
	A	ADD	308	1.195	1.247	0.2125
	A	IMC	308	1.394	2.327	0.01999
<i>rs7299460</i>						
	T	ADD	326	0.8168	-1.133	0.2571
	T	IMC	326	1.359	2.218	0.02656

SNP	A1	MUESTRA	NMISS	OR	STAT	P
<i>rs9534174</i>						
	A	ADD	324	1.094	0.5656	0.5717
	A	IMC	324	1.342	2.121	0.03388
<i>rs144848</i>						
	G	ADD	326	1.33	1.628	0.1036
	G	IMC	326	1.363	2.23	0.02576
<i>rs206081</i>						
	A	ADD	325	0.8755	-0.6832	0.4945
	A	IMC	325	1.394	2.378	0.01742
<i>rs11571787</i>						
	G	ADD	329	0.906	-0.4739	0.6356
	G	IMC	329	1.352	2.195	0.02815
<i>rs4942505</i>						
	C	ADD	323	0.9263	-0.4763	0.6338
	C	IMC	323	1.345	2.133	0.03294
<i>rs1952585</i>						
	C	ADD	331	0.8962	-0.4936	0.6216
	C	IMC	331	1.339	2.119	0.03407
<i>rs7154455</i>						
	C	ADD	328	1.023	0.1402	0.8885
	C	IMC	328	1.36	2.212	0.02699
<i>rs3020449</i>						
	C	ADD	329	0.9068	-0.6406	0.5218
	C	IMC	329	1.329	2.059	0.03951
<i>rs11649550</i>						
	A	ADD	322	0.9463	-0.3235	0.7464
	A	IMC	322	1.29	1.832	0.06701
<i>rs4843599</i>						
	A	ADD	330	0.7807	-1.042	0.2975
	A	IMC	330	1.344	2.125	0.03356
<i>rs9937099</i>						
	C	ADD	322	0.8364	-1.063	0.2877
	C	IMC	322	1.325	2.021	0.04327
<i>rs3212363</i>						
	T	ADD	327	1.227	1.228	0.2195
	T	IMC	327	1.304	1.897	0.05785

SNP	A1	MUESTRA	NMISS	OR	STAT	P
<i>rs7191944</i>						
	G	ADD	326	1.185	1.007	0.3139
	G	IMC	326	1.292	1.825	0.06804
<i>rs1042522</i>						
	C	ADD	324	1.257	1.193	0.2328
	C	IMC	324	1.336	2.086	0.03695
<i>rs2078486</i>						
	A	ADD	331	1.247	0.7632	0.4453
	A	IMC	331	1.343	2.141	0.03224
<i>rs8064946</i>						
	C	ADD	329	1.492	1.718	0.08585
	C	IMC	329	1.38	2.31	0.02086
<i>rs8176193</i>						
	A	ADD	325	0.9923	-0.0472	0.9624
	A	IMC	325	1.311	1.961	0.0499
<i>rs8176092</i>						
	C	ADD	321	1.006	0.03721	0.9703
	C	IMC	321	1.368	2.251	0.02437

En negrita se destacan los resultados con un p valor menor a 0.05. A1: alelo minoritario MUESTRA: Prueba utilizada ADD: modelo aditivo usando IMC como covariable IMC: relación entre la covariable y el SNP NMISS: número de alelos observados OR: Odds ratio STAT: Coeficiente t-estadístico P: p-valor asintótico para la distribución t

Discusión

No hay estudios previos en Uruguay en cáncer de mama esporádico a nivel genético. El presente trabajo refleja parte de los resultados obtenidos en el marco de un proyecto sobre ancestralidad y cáncer de mama esporádico e intenta ser una primera aproximación a la caracterización de esta enfermedad en nuestro país.

Si se han realizado trabajos en nuestro país en cáncer de mama familiar. En los trabajos sin publicar de Artagaveytia se analiza el rol de TP53 (comunicación personal), y Delgado y colaboradores (2011) encuentran nuevas variantes en los genes BRCA1 y BRCA2. En este último se analizan variantes directamente asociadas con cambios fenotípicos.

El hecho de que no hayamos encontrado asociación con cáncer de mama para los genes TP53, BRCA1 y BRCA2, genes candidatos más importantes en cáncer de mama familiar, estaría indicando que éstos no son relevantes en el desarrollo de cáncer de mama esporádico, o que nuestro enfoque (que intenta analizar la estructura de los genes candidatos y observar si existe relación de variantes neutras con la enfermedad) no nos está permitiendo encontrar la relación de estos genes con cáncer de mama esporádico.

Por otro lado, podríamos pensar que, el hecho de no hallar los mismos marcadores que en cáncer de mama familiar en nuestra población, estaríamos confirmando que nuestras pacientes realmente presenta la variante esporádica de esta enfermedad.

Para el análisis de datos epidemiológicos los resultados significativos son los correspondientes al IMC. Se puede observar que dentro de los pacientes se encuentra el mayor porcentaje para los rangos mayores, siendo el sobrepeso y la obesidad dos factores asociados al cáncer de mama. Esto se contradice con otros trabajos realizados en Uruguay que indican que estos factores no son determinantes para el desarrollo de cáncer de mama (Ronco et al., 2012). También se contradice con estudios realizados en mujeres latinas, donde se encontró que el índice de masa corporal estaba asociado con el cáncer de mama únicamente en mujeres pre-menopáusicas, siendo además, más bajo para los casos que para los controles (Fejerman et al., 2010). En nuestro caso, son las mujeres post-menopáusicas las que tienen asociado el IMC con el riesgo de cáncer de mama, siendo sus IMC más altos que para los controles, reafirmando el hecho de que el sobrepeso y la obesidad son factores de riesgo para esta enfermedad.

Considerando los resultados del análisis de HWE los genes ESR1, TP53 y BRCA1 tienen la totalidad de sus

SNPs en equilibrio. Por este motivo, consideramos que tenemos buenos SNPs para estos genes y que el resto de los análisis realizados que asumen HWE no pierden validez.

En el caso de rs7154455 y rs3020449 (SNPs de ESR2) y rs3212363 (SNP de TUBB3) las muestras de controles se desvían levemente del equilibrio y por lo tanto no necesitamos descartarlos como SNPs, aunque no podemos asumir equilibrio para los mismos (Balding, 2006).

El SNP rs206081 (BRCA2) presenta un desvío severo del HWE para las muestras total y de controles, por lo que dicho SNP quizás debería ser descartado en futuros análisis. Dado que igualmente no encontramos relación alguna de BRCA2 a lo largo de los distintos análisis realizados (incluyo en aquellos modelos que no asumen HWE) el mismo no fue eliminado del presente trabajo.

Para rs4237855 (SNP de VDR) y rs11649550 (SNP de FBXO31), las muestras de casos se desvían severamente del equilibrio, lo cual podría estar indicándonos que está operando algún proceso de selección o que podríamos estar viendo alguna relación con el cáncer de mama y que no deberíamos descartar estos SNP a priori (Balding, 2006). Esto podría estar dándonos alguna indicación sobre la relación de FBXO31 y VDR con el cáncer de mama (Balding, 2006).

En estos casos, en que observamos una muestra control en equilibrio pero un desvío del mismo en los casos, debemos referirnos a los resultados de los análisis utilizando los distintos modelos de asociación para comprobar si estamos ante un indicio importante.

En todos los casos de HWD se observa una disminución de los heterocigotas respecto a lo esperado, lo cual podría considerarse como error de genotipado, asignando como homocigotas a parte de la muestra de heterocigotas.

La meta de los estudios de asociación poblacionales es identificar patrones de polimorfismos que varían sistemáticamente entre individuos con diferentes estados de enfermedad y que por lo tanto podrían representar los efectos de alelos protectores o que aumentan el riesgo (Balding, 2006).

Dependiendo de su localización a lo largo del genoma los SNPs pueden alterar la función proteica, si están localizados en regiones exónicas, o modificar los niveles de expresión de encontrarse en los promotores. Mediante estudios de GWAS se han detectado un gran número de polimorfismos génicos postulados como responsables de influenciar el desarrollo de varias enfermedades, entre ellas la susceptibilidad al cáncer de mama (Süllner et al., 2012).

La premisa primordial de los estudios de asociación es que existe variabilidad en las secuencias de ADN a lo largo de los individuos que captura la información de una enfermedad (Foulkes, 2009).

Las regiones de ADN dentro de un gen presentan variabilidad genética si los alelos para dicha región varían a lo largo de la población. Las regiones conservadas, por otro lado, no exhiben variabilidad en la

población. Las regiones altamente conservadas de ADN son menos relevantes en el contexto de estudios de asociación dado que no serán capaces de capturar la variabilidad en los rasgos de la enfermedad. Estudiar regiones altamente conservadas sería equivalente a, en una investigación epidemiológica tradicional, reclutar únicamente fumadores, y luego tratar de valorar el impacto de fumar en el riesgo de cáncer.

En este estudio encontramos una asociación de ESR1 con el cáncer de mama, asociación que es ampliamente discutida en la bibliografía (Dunning et al., 2009; Fejerman et al., 2012; Garcia-Closas et al., 2013; Jeon et al., 2010; N. Li, Dong, Hu, Shen, & Dai, 2010; Modugno et al., 2005; Tapper et al., 2008). Distintas variantes han sido directamente relacionadas con el cáncer de mama (Dunning et al., 2009; N. Li et al., 2010; Tapper et al., 2008; Wang et al., 2007), sin embargo no existen estudios de este tipo en nuestro país sobre la relación de este gen y cáncer de mama esporádico.

El SNP rs827423 del gen ESR1 presenta una diferencia significativa entre las frecuencias para el alelo minoritario entre casos y controles, con un p de 0.03571 y un OR de 1.379. Este valor lo podemos interpretar como que, el riesgo para un portador del alelo C para este SNP es 1.379 veces más que para quien porta el alelo T. Dado que el SNP se encuentra en la región intrónica del gen podría estar influenciando un sitio de regulación del mismo, o nos podría estar informando sobre otra variable que afecta directamente la función de la proteína y que se encuentra en desequilibrio de ligamiento con la misma. No existe bibliografía para este SNP.

Si existen trabajos que estudian la asociación de rs2273206, único de los SNPs de ESR1 analizados en este trabajo del que hay bibliografía disponible, que indican que ESR1 podría estar relacionado con mayor riesgo de ser diagnosticado en etapas más avanzadas del cáncer de próstata (Chae et al., 2009) y asociado a esquizofrenia (Weickert et al., 2008).

Al comparar nuestros datos de frecuencias alélicas con los disponibles en el HapMap (The International HapMap Consortium, 2010) para las poblaciones europea (CEU), africana (YRI) y mestizada (MEX) observamos que para el SNP rs827423 la frecuencia para nuestra población, en pacientes (0.4697), se encuentra entre los valores de las poblaciones europea (0.447) y africana (0.372). En el caso de los controles (0.3911), el valor se encuentra entre los de las poblaciones europea y mexicana (0.644). De ambos valores, el que más se acerca a la frecuencia observada para europeos es el de pacientes. Esto coincide con el hecho de que un mayor componente de ancestría europeo aumenta el riesgo de padecer cáncer de mama. Podríamos estar observando un efecto protector relacionado con la ancestría indígena. Dado que para el resto de las variantes de ESR1 estudiadas, los valores mantienen esta tendencia (pacientes con frecuencias más cercanas a las europeas y controles con frecuencias más cercanas a las indígenas) no podemos ser concluyentes en este punto. Estos resultados nos permiten afirmar que no podemos asumir que la población uruguaya es solamente europea, debemos tener en

cuenta el mestizaje propio de nuestra población. Por este motivo, es importante realizar estudios de este tipo en mujeres uruguayas en vez de extrapolar resultados de otras investigaciones en lo que refiere a pronóstico y tratamiento de esta enfermedad.

Dentro de los resultados significativos para los análisis de asociación, se encuentra el correspondiente al SNP rs827423, utilizando el método de Cochran-Armitage, cuyo p valor es más significativo que para el modelo alélico, por lo que podríamos decir que el primero describe mejor la asociación del alelo C con el cáncer de mama.

En el estudio de genes candidatos no sabemos a priori que mutaciones pueden estar involucradas. Pueden ser mutaciones ya descritas o pueden ser nuevas, propias de nuestra población. Es por esto que optamos por utilizar SNPs conocidos, frecuentes y neutros. La intención al utilizar haplotipos es intentar describir algún tipo de secuencia que pueda estar asociado al cáncer de mama, esperando recuperar información que se puede perder al utilizar SNPs de manera independiente.

Al observar lo que sucede con los haplotipos de ESR1 encontramos que los formados por el alelo C de rs827423 y el alelo A de rs6912184, por estos y el alelo C de rs2982712, y el que contiene a los tres anteriores y el alelo G de rs2273206, son todos significativos. En todos los casos la frecuencia del haplotipo es más alta en pacientes que en controles. Dado que los haplotipos que contienen al alelo C de rs827423, dieron una asociación significativa con cáncer de mama, continuamos infiriendo una asociación de este SNP con cáncer de mama más que de algún haplotipo en particular. Considerando, que el haplotipo queda contenido entre el SNP que nos da significativo, y el que se ha asociado a otras patologías en la bibliografía, y que en total se extiende 200Kb, debemos concluir que la región en sí, más que el haplotipo, tiene relación con el cáncer de mama. Probablemente estos SNPs se encuentren en LD con otras variantes que si afectan el desarrollo de cáncer de mama. De todos los SNPs citados en la bibliografía que se encuentran en esta región, se destacan: rs3020314, asociado con un aumento en la susceptibilidad en cáncer de mama, de manera dominante, en poblaciones europeas con un OR de 1.05 para su alelo C (Dunning et al., 2009); rs2234693 y rs1801132 asociados con riesgo de cáncer de mama en un meta análisis de 10300 pacientes y 16620 controles (N. Li et al., 2010); el alelo C de rs3020410 asociado con un riesgo aumentado de recaída en pacientes con cáncer de mama (Tapper et al., 2008); y un rol protector del SNP rs9340799 en un estudio realizado en 393 pacientes y 1183 controles entre mujeres caucásicas mayores de 65 años (Wang et al., 2007). Estos datos sugieren que el estudio de los SNPs presentes en esta región sería un paso posterior a este trabajo.

Dado que el SNP que da significativo a lo largo de todos los análisis es uno correspondiente a ESR1, receptor de estrógeno, íntimamente relacionado con el metabolismo del colesterol, a su vez relacionado con el sobrepeso y la obesidad, nuestros resultados epidemiológicos y genéticos son concordantes. Dado que al corregir los datos por IMC, la asociación de dicho SNP sigue siendo

significativa, estamos confirmando que es realmente ESR1, y no la combinación de los factores genéticos (ESR1) y epidemiológicos (IMC) lo que influye sobre el desarrollo de cáncer de mama esporádico.

El otro gen donde encontramos un resultado significativo para la asociación es rs11649550, aunque el p valor es apenas significativo para el modelo genotípico, seguimos infiriendo una relación entre FBXO31 y cáncer de mama para nuestra población.

El gen FBXO31 ha sido estudiado en cáncer de mama esporádico (Kumar et al., 2005), en melanoma (Santra, Wajapeyee, & Green, 2009) y en tejido neuronal (Vadhvani, Schwedhelm-Domeyer, Mukherjee, & Stegmüller, 2013). En todos los casos se ha propuesto a FBXO31 como candidato supresor de tumores y se ha observado variación de su expresión en líneas celulares tumorales, debido a su regulación del ciclo celular, la síntesis de ADN y el crecimiento de las células en cultivo. En el estudio más reciente, (Vadhvani et al., 2013), observaron también su interacción con el complejo centrosomal relacionado con su importancia en el control de la morfogénesis neuronal y la identidad axonal, así como su requerimientos para el crecimiento de las dendritas y la migración de las neuronas en la corteza cerebelar en desarrollo. Todo esto lo convierte en un buen candidato para cáncer de mama.

Es interesante observar que la inclusión de IMC como covariante no modificó sustancialmente la significancia de los SNPs en un modelo aditivo. Si bien es conocido el efecto de la obesidad en el desarrollo de cáncer de mama (Ronco et al., 2012), no parece ser importante con respecto a los genes estudiados.

Para dos SNPs del gen VDR, rs4237855 y rs7299460, el haplotipo formado por sus alelos A y C respectivamente, es significativo. Esto nos está dando la idea de que si bien, cada uno por separado no es capaz de asociarse a la enfermedad, el haplotipo formado por ambos logra detectar una asociación a la misma. Estos SNPs se encuentran en intrones consecutivos del gen, a una distancia de 9 Kb y entre ambos se ubican cientos de SNPs. Si bien la mayoría de estos SNPs son neutros, uno de ellos el rs4760658, ha sido asociado con cáncer de mama en mujeres chinas. (Dorjgocho et al., 2011) El estudio de esta región quizás arroje más información respecto a la relación de VDR con el cáncer de mama. Lo que sucede con VDR es un punto a favor de la estrategia de utilización de haplotipos. La secuenciación de esta región nos brindaría más información al respecto.

Si bien en los estudios de este tipo es conveniente tener un alto número de casos y controles para brindar soporte estadístico a los resultados, hay que tomar en cuenta el tamaño de la población estudiada y la incidencia de la enfermedad, ambos factores que impiden aumentar el número de muestras obtenidas. Sin embargo, consideramos que por la misma razón el N utilizado en el presente trabajo es suficientemente representativo de nuestra población.

Conclusiones

El IMC debe ser tenido en cuenta en todo lo referido a cáncer de mama en el Uruguay. El sobrepeso y la obesidad son factores de riesgo para esta enfermedad en nuestro país.

El cáncer de mama esporádico muestra el mismo comportamiento que en otras poblaciones, no se encuentra asociación de los genes TP53, BRCA1 y BRCA2.

El SNP rs827423 está relacionado con cáncer de mama, así como una amplia región marcada por el haplotipo CACG de rs827423 rs6912184, rs2982712, y rs2273206, por lo tanto ESR1 está relacionado con cáncer de mama en nuestra población.

Nuestros resultados indican que VDR podría estar relacionado, sugiriendo nuevos candidatos para futuras investigaciones.

FBXO31 podría estar relacionado, debiendo estudiarse más este gen para poder obtener resultados concluyentes.

El mestizaje debe ser tenido en cuenta al realizar estudios poblacionales en el Uruguay.

Perspectivas

A partir de los resultados obtenidos en el presente trabajo consideramos a futuro realizar otros análisis para continuar caracterizando la población uruguaya en lo que refiere a cáncer de mama.

En primer lugar, la corrección de nuestros datos utilizando información sobre la ancestría de las participantes de este estudio nos permitirá obtener resultados más confiables y la confirmación de que el grado de mestizaje de un individuo realmente afecta sus características genéticas.

Por otra parte, el análisis de otras variantes epidemiológicas nos brindará más información al momento de considerar factores de riesgo y comprender mejor aquellos factores que llevan a un desenlace de la enfermedad.

Un estudio más profundo de aquellos genes que parecen estar afectando de alguna manera el riesgo de IMC con cáncer de mama nos podrá permitir un mejor entendimiento del proceso subyacente a esta interferencia.

El análisis de los SNPs de ESR1 que se encuentran asociados con cáncer de mama en la bibliografía, contenidos en la región formada por el haplotipo hallado en este estudio nos permitirá comprender mejor la relación hallada en el presente trabajo.

Un análisis más en detalle de la región de VDR que nos da asociada con cáncer de mama puede arrojar otras variantes relacionadas con éste.

Deberíamos hacer un análisis más extenso del gen FBXO31 y terminar de dilucidar si la relación entre este y el cáncer es algo real o estamos observando una asociación espúrea.

En todos los casos, la obtención de más muestras, tanto de pacientes como de controles, nos permitirá confirmar o corroborar nuestros datos.

Bibliografía

- Abbas, S., Nieters, A., Linseisen, J., Slinger, T., Kropp, S., Mutschelknauss, E. J., ... Chang-claude, J. (2008). Vitamin D receptor gene polymorphisms and haplotypes and postmenopausal breast cancer risk. *Breast Cancer Research*, 10(2), 1–11. doi:10.1186/bcr1994
- American Cancer Society. (2013). Cancer de Seno. *Tipos de Cancer de Seno*. Retrieved February 05, 2014, from <http://www.cancer.org/espanol/cancer/cancerdeseno/guiadetallada/cancer-de-seno-what-is-breast-cancer-types>
- Bahcall, O. (2013). Common variation and heritability estimates for breast, ovarian and prostate cancers. *Nature Genetics - iCOGS*. doi:10.1038/ngicogs.1
- Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews. Genetics*, 7(10), 781–91. doi:10.1038/nrg1916
- Barber, L. J., Rosa Rosa, J. M., Kozarewa, I., Fenwick, K., Assiotis, I., Mitsopoulos, C., ... Ashworth, A. (2011). Comprehensive genomic analysis of a BRCA2 deficient human pancreatic cancer. *PloS One*, 6(7), e21639. doi:10.1371/journal.pone.0021639
- Barroso, E., Fernandez, L. P., Milne, R. L., Pita, G., Sendagorta, E., Floristan, U., ... Ribas, G. (2008). Genetic analysis of the vitamin D receptor gene in two epithelial cancers: melanoma and breast cancer case-control studies. *BMC Cancer*, 8, 385. doi:10.1186/1471-2407-8-385
- Chae, Y. K., Huang, H.-Y., Strickland, P., Hoffman, S. C., & Helzlsouer, K. (2009). Genetic polymorphisms of estrogen receptors alpha and beta and the risk of developing prostate cancer. *PloS One*, 4(8), e6523. doi:10.1371/journal.pone.0006523
- Collaborative Group on Hormonal Factors in Breast Cancer. (2012). Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *The Lancet Oncology*, 13(11), 1141–1151. doi:10.1016/S1470-2045(12)70425-4
- Couch, F. J., Wang, X., McGuffog, L., Lee, A., Olswold, C., Kuchenbaecker, K. B., ... Antoniou, A. C. (2013). Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genetics*, 9(3), e1003212. doi:10.1371/journal.pgen.1003212
- Dalziel, M., Kolesnichenko, M., das Neves, R. P., Iborra, F., Goding, C., & Furger, A. (2010). {alpha}-MSH regulates intergenic splicing of MC1R and TUBB3 in human melanocytes. *Nucleic Acids Research*, 1–15. doi:10.1093/nar/gkq1125
- Delgado, L., Fernández, G., Grotiuz, G., Cataldi, S., González, a, Lluveras, N., ... Muse, I. M. (2011). BRCA1 and BRCA2 germline mutations in Uruguayan breast and breast-ovarian cancer families. Identification of novel mutations and unclassified variants. *Breast Cancer Research and Treatment*, 128(1), 211–8. doi:10.1007/s10549-010-1320-2
- Dorjgocho, T., Delahanty, R., Lu, W., Long, J., Cai, Q., Zheng, Y., ... Shu, X. O. (2011). Common genetic variants in the vitamin D pathway including genome-wide associated variants are not associated with breast cancer risk among Chinese women. *Cancer Epidemiol Biomarkers Prev*, 20(10), 2313–2316. doi:10.1158/1055-9965.EPI-11-0704

- Dunning, A. M., Healey, C. S., Baynes, C., Maia, A.-T., Scollen, S., Vega, A., ... Chenevix-Trench, G. (2009). Association of ESR1 gene tagging SNPs with breast cancer risk. *Human Molecular Genetics*, 18(6), 1131–9. doi:10.1093/hmg/ddn429
- Fanale, D., Amodeo, V., Corsini, L. R., Rizzo, S., Bazan, V., & Russo, A. (2012). Breast cancer genome-wide association studies: there is strength in numbers. *Oncogene*, 31(17), 2121–8. doi:10.1038/onc.2011.408
- Fejerman, L., Chen, G. K., Eng, C., Huntsman, S., Hu, D., Williams, A., ... Burchard, E. G. (2012). Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Human Molecular Genetics*, 21(8), 1907–1917. doi:10.1093/hmg/ddr617
- Fejerman, L., Romieu, I., John, E. M., Lazcano-ponce, E., Beckman, K. B., Pérez-stable, E. J., ... Torres-mejía, G. (2010). European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol Biomarkers Prev*, 19(655), 1074–1082. doi:10.1158/1055-9965.EPI-09-1193.
- Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C., & Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer. Journal International Du Cancer*, 127(12), 2893–2917. doi:10.1002/ijc.25516
- Foulkes, A. (2009). *Applied statistical genetics with R: for population-based association studies*. (R. Parmigiani, G. Kurt, & H. Giovanni, Eds.). Springer. doi:10.1007/978-0-387-89554-3
- Garcia-Closas, M., Couch, F. J., Lindstrom, S., Michailidou, K., Schmidt, M. K., Brook, M. N., ... Kraft, P. (2013). Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nature Genetics*, 45(4), 392–8, 398e1–2. doi:10.1038/ng.2561
- Gaudet, M. M., Kuchenbaecker, K. B., Vijai, J., Klein, R. J., Kirchhoff, T., McGuffog, L., ... Offit, K. (2013). Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk. *PLoS Genetics*, 9(3), e1003173. doi:10.1371/journal.pgen.1003173
- Gayther, S. A., Warren, W., Mazoyer, S., Russell, P. A., Harrington, P. A., Chiano, M., ... Ponder, B. A. (1995). Germline mutations of the BRCA1 gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation. *Nature Genetics*, 11(4), 428–33. doi:10.1038/ng1295-428
- Gleason, M. X., Mdzinarishvili, T., & Sherman, S. (2012). Breast cancer incidence in black and white women stratified by estrogen and progesterone receptor statuses. *PloS One*, 7(11), e49359. doi:10.1371/journal.pone.0049359
- Gold, B., Kalush, F., Bergeron, J., Scott, K., Mitra, N., Wilson, K., ... Offit, K. (2004). Estrogen receptor genotypes and haplotypes associated with breast cancer risk. *Cancer Research*, 64(24), 8891–8900. doi:10.1158/0008-5472.CAN-04-1256
- Henríquez-Hernández, L. A., Murias Rosales, A., Hernández González, A., Cabrera de León, A., Díaz-Chico, B. N., Mori De Santiago, M., & Fernández Pérez, L. (2009). Gene polymorphisms in TYMS, MTHFR, p53 and MDR1 as risk factors for breast cancer: A case-control study. *Oncology Reports*, 22, 1425–1433. doi:10.3892/or
- Hu, Z., Li, X., Yuan, R., Ring, B. Z., & Su, L. (2010). Three common TP53 polymorphisms in susceptibility to breast cancer, evidence from meta-analysis. *Breast Cancer Research and Treatment*, 120, 705–714. doi:10.1007/s10549-009-0488-9

- IARC (International Agency for Research on Cancer). (2008). Online Analysis. GLOBOCAN 2008: Cancer Incidence and Mortality Worldwide. Retrieved December 11, 2012, from <http://globocan.iarc.fr/Default.aspx>
- Instituto Nacional del Cáncer. (2013). Tipos de cancer de seno. *Lo que usted necesita saber sobre el cáncer de seno*. Retrieved September 24, 2013, from <http://www.cancer.gov/espanol/tipos/necesita-saber/seno>
- Iwao, K., Miyoshi, Y., Egawa, C., Ikeda, N., & Noguchi, S. (2000). Quantitative analysis of estrogen receptor-beta mRNA and its variants in human breast cancers. *International Journal of Cancer. Journal International Du Cancer*, 88(5), 733–6. doi:10.1002/1097-0215(20001201)88:5<733::AID-IJC8>3.0.CO;2-M
- Jeon, S., Choi, J.-Y., Lee, K.-M., Park, S. K., Yoo, K.-Y., Noh, D.-Y., ... Kang, D. (2010). Combined genetic effect of CDK7 and ESR1 polymorphisms on breast cancer. *Breast Cancer Research and Treatment*, 121(3), 737–42. doi:10.1007/s10549-009-0640-6
- Jia, L., & Sun, Y. (2009). F-box proteins FBXO31 and FBX4 in regulation of cyclin D1 degradation upon DNA damage. *Pigment Cell Melanoma Research*, 22(5), 518–519. doi:10.1111/j.1755-148X.2009.00611.x.F-box
- Jiaxue, W., Lin-Yu, L., & Xiaochun, Y. (2010). The role of BRCA1 in DNA damage response. *Protein Cell*, 1(2), 117–123. doi:10.1007/s13238-010-0010-5.The
- Kumar, R., Neilsen, P. M., Crawford, J., McKirdy, R., Lee, J., Powell, J. a, ... Callen, D. F. (2005). FBXO31 is the chromosome 16q24.3 senescence gene, a candidate breast tumor suppressor, and a component of an SCF complex. *Cancer Research*, 65(24), 11304–13. doi:10.1158/0008-5472.CAN-05-0936
- Kuo, W.-H., Lin, P.-H., Huang, A.-C., Chien, Y.-H., Liu, T.-P., Lu, Y.-S., ... Chang, K.-J. (2012). Multimodel assessment of BRCA1 mutations in Taiwanese (ethnic Chinese) women with early-onset, bilateral or familial breast cancer. *Journal of Human Genetics*, 57(2), 130–8. doi:10.1038/jhg.2011.142
- Leygue, E., Dotzlaw, H., Watson, P. H., & Murphy, L. C. (1998). Altered Estrogen Receptor α and β Messenger RNA Expression during Human Breast Tumorigenesis Advances in Brief Messenger RNA Expression during Human, 3197–3201.
- Leygue, E., Dotzlaw, H., Watson, P. H., & Murphy, L. C. (1999). Expression of Estrogen Receptor β 1, β 2, and β 5 Messenger RNAs in Human Breast Tissue. *Cancer Research*, 59, 1175–1179.
- Li, N., Dong, J., Hu, Z., Shen, H., & Dai, M. (2010). Potentially functional polymorphisms in ESR1 and breast cancer risk: a meta-analysis. *Breast Cancer Research and Treatment*, 121(1), 177–84. doi:10.1007/s10549-009-0532-9
- Li, Y., Li, Y., Wedrén, S., Li, G., Charn, T. H., Desai, K. V., ... Liu, J. (2011). Genetic variation of ESR1 and its co-activator PPARGC1B is synergistic in augmenting the risk of estrogen receptor-positive breast cancer. *Breast Cancer Research : BCR*, 13(1), R10. doi:10.1186/bcr2817
- Lo Nigro, C., Vivenza, D., Monteverde, M., Lattanzio, L., Gojis, O., Garrone, O., ... Crook, T. (2012). High frequency of complex TP53 mutations in CNS metastases from breast cancer. *British Journal of Cancer*, 106(2), 397–404. doi:10.1038/bjc.2011.464

- Lopes, N., Paredes, J., Costa, J. L., Ylstra, B., & Schmitt, F. (2012). Vitamin D and the mammary gland: a review on its role in normal development and breast cancer. *Breast Cancer Research : BCR*, 14(3), 211. doi:10.1186/bcr3178
- Maguire, P., Margolin, S., Skoglund, J., Sun, X.-F., Gustafsson, J.-A., Børresen-Dale, A.-L., & Lindblom, A. (2005). Estrogen receptor beta (ESR2) polymorphisms in familial and sporadic breast cancer. *Breast Cancer Research and Treatment*, 94(2), 145–52. doi:10.1007/s10549-005-7697-7
- McKay, J. D., McCullough, M. L., Ziegler, R. G., Kraft, P., Saltzman, B. S., Riboli, E., ... Thun, M. J. (2009). Vitamin D receptor polymorphisms and breast cancer risk: results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. *Cancer Epidemiology, Biomarkers & Prevention*, 18(1), 297–305. doi:10.1158/1055-9965.EPI-08-0539
- Milne, R. L., Gaudet, M. M., Spurdle, A. B., Fasching, P. a, Couch, F. J., Benítez, J., ... Chang-Claude, J. (2010). Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. *Breast Cancer Research : BCR*, 12(6), R110. doi:10.1186/bcr2797
- Modugno, F., Zmuda, J. M., Potter, D., Cai, C., Ziv, E., Cummings, S. R., ... Cauley, J. a. (2005). Association of estrogen receptor alpha polymorphisms with breast cancer risk in older Caucasian women. *International Journal of Cancer*, 116(6), 984–991. doi:10.1002/ijc.21105
- OMS. (2013). Cáncer. Nota descriptiva N° 297. Retrieved September 24, 2013, from <http://www.who.int/mediacentre/factsheets/fs297/es/index.html>
- Palmieri, C., Cheng, G. J., Saji, S., M, Z.-H., Wärr, A., Van Noorden, S., ... Gustafsson, J.-A. (2002). Estrogen receptor beta in breast cancer. *Endocrine-Related Cancer*, 9, 1–13.
- Purcell, S. (2010). PLINK (1.07) Documentation. Retrieved from <http://pngu.mgh.harvard.edu/~purcell/plink/pdf.shtml>
- Rollison, D. E., Cole, A. L., Tung, K., Slattery, M. L., Baumgartner, K. B., Byers, T., ... Giuliano, A. R. (2012). Vitamin D intake, vitamin D receptor polymorphisms, and breast cancer risk among women living in the southwestern U.S. *Breast Cancer Research and Treatment*, 132(2), 683–691. doi:10.1007/s10549-011-1885-4
- Ronco, A. L., De Stefani, E., Deneo-Pellegrini, H., & Quarneti, A. (2012). Diabetes, overweight and risk of postmenopausal breast cancer: a case-control study in Uruguay. *Asian Pacific Journal of Cancer Prevention : APJCP*, 13(1), 139–46. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22502657>
- Santra, M. K., Wajapeyee, N., & Green, M. R. (2009). F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage. *Nature*, 459(7247), 722–725. doi:10.1038/nature08011
- Sanz, M. (2000). Admixture Studies in Latin America: From the 20th to the 21st Century. *Human Biology*, 72(1), 155–177.
- Shibazaki, M., Maesawa, C., Akasaka, K., Kasai, S., Yasuhira, S., Kanno, K., ... Mori, N. (2012). Transcriptional and post-transcriptional regulation of β III-tubulin protein expression in relation with cell cycle-dependent regulation of tumor cells. *International Journal of Oncology*, 40(3), 695–702. doi:10.3892/ijo.2011.1291

- Süllner, J., Lattrich, C., Häring, J., Görse, R., Ortmann, O., & Treeck, O. (2012). A polymorphism in the nuclear receptor coactivator 7 gene and breast cancer susceptibility. *Oncology Letters*, 3(1), 131–134. doi:10.3892/ol.2011.421
- Tapper, W., Hammond, V., Gerty, S., Ennis, S., Simmonds, P., Collins, A., & Eccles, D. (2008). The influence of genetic variation in 30 selected genes on the clinical characteristics of early onset breast cancer. *Breast Cancer Research : BCR*, 10(6), R108. doi:10.1186/bcr2213
- The International HapMap Consortium. (2010). International HapMap Project. Retrieved from www.hapmap.org
- Trifa, F., Karray-Chouayekh, S., Mabrouk, I., Baccouche, S., Khabir, A., Sellami-boudawara, T., ... Mokdad-gargouri, R. (2010). Haplotype analysis of p53 polymorphisms: Arg72Pro, Ins16bp and G13964C in Tunisian patients with familial or sporadic breast cancer. *Cancer Epidemiology*, 34(2), 184–8. doi:10.1016/j.canep.2010.02.007
- Vadhvani, M., Schwedhelm-Domeyer, N., Mukherjee, C., & Stegmüller, J. (2013). The centrosomal E3 ubiquitin ligase FBXO31-SCF regulates neuronal morphogenesis and migration. *PloS One*, 8(2), e57530. doi:10.1371/journal.pone.0057530
- Wang, J., Higuchi, R., Modugno, F., Li, J., Umblas, N., Lee, J., ... Rhee, B. (2007). Estrogen receptor alpha haplotypes and breast cancer risk in older Caucasian women. *Breast Cancer Research and Treatment*, 106(2), 273–80. doi:10.1007/s10549-007-9497-8
- Weickert, C. S., Miranda-Angulo, A. L., Wong, J., Perlman, W. R., Ward, S. E., Radhakrishna, V., ... Kleinman, J. E. (2008). Variants in the estrogen receptor alpha gene and its mRNA contribute to risk for schizophrenia. *Human Molecular Genetics*, 17(15), 2293–309. doi:10.1093/hmg/ddn130
- Zhao, C., Lam, E. W.-F., Sunter, A., Enmark, E., De Bella, M. T., Coombes, R. C., ... Dahlman-Wright, K. (2003). Expression of estrogen receptor beta isoforms in normal breast epithelial cells and breast cancer: regulation by methylation. *Oncogene*, 22(48), 7600–6. doi:10.1038/sj.onc.1207100
- Zheng, Y., Huo, D., Zhang, J., Yoshimatsu, T. F., Niu, Q., & Olopade, O. I. (2012). Microsatellites in the Estrogen Receptor (ESR1, ESR2) and Androgen Receptor (AR) Genes and Breast Cancer Risk in African American and Nigerian Women. *PloS One*, 7(7), e40494. doi:10.1371/journal.pone.0040494

Recursos electrónicos

PLINK

Package: PLINK 1.07

Author: Shaun Purcell

URL: <http://pngu.mgh.harvard.edu/purcell/plink/>

R Project

<http://www.r-project.org/>

Base de datos del National Center for Biotechnology Information

<http://www.ncbi.nlm.nih.gov>

LCG Genomics – KASP Genotyping Assay

<http://www.lgcgenomics.com/genotyping/kasp-genotyping-reagents/kasp-overview/>

HapMap

<http://www.hapmap.org>

Contenido

Agradecimientos.....	2
Abreviaturas.....	3
Resumen	5
Introducción.....	6
Cáncer de Mama.....	6
Estudios de Asociación.....	16
Investigaciones poblacionales versus familiares	16
Genes Candidatos.....	16
BRCA1	17
BRCA2.....	18
ESR1	19
ESR2.....	19
FBXO31	20
TP53.....	21
TUBB3.....	23
VDR.....	25
Objetivos	26
Objetivo General.....	26
Objetivos Específicos.....	26
Materiales y Métodos.....	27
Población de estudio.....	27
Selección de SNPs.....	27
Ensayo Genotípico.....	27
Análisis Estadísticos.....	29
Estudio de asociación de genes candidatos.....	29
Equilibrio de Hardy-Weinberg.....	30
Pruebas de asociación.....	31
Resultados	34
Epidemiología	34
Equilibrio de Hardy-Weinberg.....	35
Frecuencias y asociaciones alélicas	42
Pruebas de Asociación	43
Asociación Haplotípica	45

Regresión logística	52
Discusión	55
Conclusiones	60
Perspectivas	61
Bibliografía	62
Recursos electrónicos	67