

## **MANEJO DE LAS VARIANTES GENÉTICAS SIN CLASIFICAR EN UNA CONSULTA DE CONSEJO GENÉTICO.**

### **ANA MARÍA SÁNCHEZ DE ABAJO**

*Servicio de Bioquímica Clínica. Hospital Universitario Insular de Gran Canaria. Las Palmas de Gran Canaria.*

### **MIGUEL DE LA HOYA MANTECÓN**

*Unidad de Cáncer Familiar. Laboratorio de Oncología Molecular del Hospital Universitario Clínico San Carlos. Madrid.*

## **1. DESCRIPCIÓN DEL CASO CLÍNICO.**

Paciente (II-3) diagnosticada a los 43 años de un carcinoma ductal infiltrante de mama izquierda grado 3, receptor de estrógeno (RE) +, receptor de progesterona (RP) +++ y HER2 -. Hermana (II-2) diagnosticada a los 49 años de un carcinoma ductal infiltrante de mama izquierda, pobremente diferenciado, con RE +++, RP +++ y HER2 -, estadio pT1, N3, M0. Madre (I-8) diagnosticada de cáncer de cérvix a los 60 años. Dos tíos paternos (I-1, I-4) fallecidos de cáncer de próstata e hígado, respectivamente (Figura 1).

Basándonos en los antecedentes familiares y personales de la paciente sospechamos que podríamos estar ante un síndrome de cáncer de mama y ovario hereditario (HBOC) (OMIM: 604370) y se decide hacer un test genético BRCA (análisis de la secuencia codificante completa y regiones intrónicas adyacentes de los genes *BRCA1* (OMIM: 113705) y *BRCA2* (OMIM: 600185)).

De acuerdo a nuestra experiencia, el número de variantes genéticas (por definición, cambios respecto a una secuencia de referencia dada) que podemos esperar en un test genético BRCA en población española oscila entre 5 y 36, con una media en torno a 20. Por supuesto, estas cifras son orientativas y en cualquier caso dependiente del tamaño de la región diana (número de nucleótidos) que se incluya en el test BRCA. En general, todos los protocolos actuales incluyen la secuencia codificante completa de ambos genes (5589 nucleótidos en el caso de *BRCA1* y 10254 en el caso de *BRCA2*), aunque puede variar de forma sustancial la cantidad de secuencia intrónica que analizan (y en consecuencia, la proporción de variantes intrónicas que identifican).

---

En nuestro caso clínico, el test genético BRCA ha identificado las 30 variantes genéticas que se muestran en la figura 1, y nos planteamos la siguiente pregunta: ¿estamos ante un test BRCA informativo? Para poder contestar a esta pregunta debemos saber si hemos identificado alguna variante genética patogénica, y para ello debemos adentrarnos en el campo de la clasificación clínica de variantes genéticas.

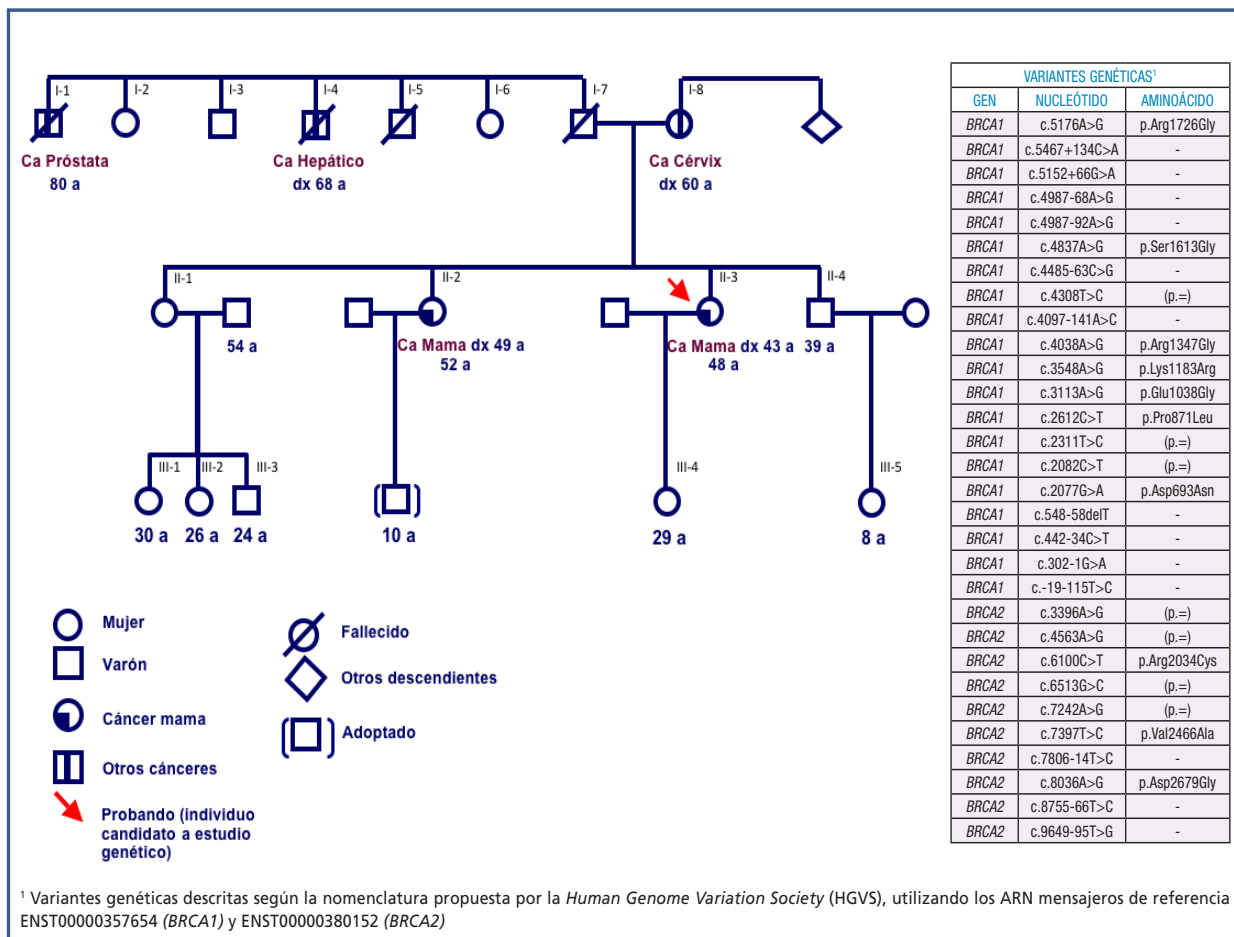


Figura 1: Árbol genealógico del caso clínico.

## 2. TEST GENÉTICO Y CÁNCER FAMILIAR: ALGUNAS DEFINICIONES Y CONSENSOS BÁSICOS.

Se estima que los síndromes hereditarios de cáncer pueden explicar un 5-10 % de todos los casos de cáncer. Estos síndromes se suelen caracterizar por edades de diagnóstico más tempranas, múltiples casos de cáncer en un individuo (y/o bilateralidad) y antecedentes familiares. Se han descrito más de 100 síndromes hereditarios de cáncer, aunque muchos son extraordinariamente raros en la población. La mayoría están causados por mutaciones germinales que causan pérdida de función en genes supresores de tumores. Algunos ejemplos clásicos son el Síndrome de Li-Fraumeni (OMIM: 151623) causado por mutaciones germinales en el gen *TP53* (OMIM: 191170), la Poliposis Adenomatosa Familiar

(OMIM: 175100) causada por mutaciones en el gen *APC* (OMIM: 611731), el Síndrome de Lynch (OMIM: 114500) causado mayoritariamente por mutaciones en los genes *MLH1* (OMIM: 120436) y *MSH2* (OMIM: 120435) y el propio Síndrome de Cáncer de Mama y Ovario Hereditario diagnosticado en nuestro caso clínico (causado por mutaciones en los genes supresores de tumor *BRCA1* y *BRCA2*).

Según el último consenso publicado de la Sociedad Americana de Oncología Clínica (ASCO), únicamente se recomienda ofrecer un test genético cuando los resultados del mismo ayuden en el diagnóstico o influyan en el manejo médico o quirúrgico del individuo o de sus familiares a riesgo. Desde un punto de vista clínico un test genético sólo tiene dos resultados posibles: informativo y no informativo.

Un test informativo es aquel que identifica una variante genética capaz de explicar el síndrome presente en la familia. En la jerga utilizada habitualmente en consejo genético, estas variantes de alta penetrancia se suelen designar indistintamente como variantes (o mutaciones) patogénicas, variantes (o mutaciones) clínicamente relevantes, y variantes (o mutaciones) informativas. No obstante, en la actualidad se desaconseja el uso del término mutación, que es poco preciso y puede llevar a equívocos. En cualquier caso, estas variantes genéticas relacionadas con el síndrome son informativas porque permiten individualizar las estimaciones de riesgo y por tanto, el consiguiente asesoramiento genético (el manejo clínico de las familias). El ejemplo paradigmático de variante informativa en genes supresores de tumores es una mutación sin sentido (*nonsense*), o una pequeña inserción/delección (*Indel*), que introduce un codón de parada prematuro (PTC, del inglés *pre-mature termination codon*) y en consecuencia podemos inferir que inactiva el gen supresor sin necesidad de información adicional. La inmensa mayoría de mutaciones patogénicas descritas en *BRCA1* y *BRCA2* son de este tipo, y cuando clasificamos otro tipo de variantes genéticas (por ejemplo una variante no sinónima, en inglés *missense*) como patogénicas estamos asumiendo que confieren un riesgo de desarrollar cáncer equivalente (y por tanto que son variantes igualmente informativas).

Un test no informativo es aquel que no identifica ninguna variante genética capaz de explicar el síndrome observado en la familia. De este modo, el resultado del test genético no modifica el manejo clínico del probando y/o sus familiares. Dentro de esta categoría se incluyen tanto los test que sólo identifican variantes no relacionadas con el síndrome, como aquellos test que identifican variantes cuya asociación con el síndrome se desconoce (no está ni confirmada ni descartada). En este último caso hablamos de variantes de significado clínico incierto o variantes sin clasificar (VSC) (en la literatura científica lo más frecuente es utilizar el acrónimo VUS, del inglés *Variant of Uncertain Significance*).

Los resultados VSC además de no informativos, añaden un alto grado de incertidumbre al resultado del test genético. En este caso se sugiere que la variante identificada sea objeto de investigación con el propósito de dilucidar su relevancia.

### **3- VARIANTES GENÉTICAS SIN CLASIFICAR Y CONSEJO GENÉTICO.**

Una variante sin clasificar es un cambio en la secuencia génica para el que carecemos (en el momento de redactar nuestro informe de laboratorio) de datos que nos permitan concluir (o excluir) su relevancia clínica.

Las variantes sin clasificar son un problema importante asociado a los test genéticos. En esencia, un test genético que concluye con la identificación de una variante sin clasificar es un test no informativo exactamente igual que cualquier otro test donde no se identifican mutaciones patogénicas. Sin embargo, la identificación de variantes sin clasificar nunca es deseable y añade complicaciones adicionales a un proceso de asesoramiento genético ya de por sí complejo.

Por un lado, es difícil transmitir de forma adecuada la información al portador sin incrementar innecesariamente el estrés emocional. Es frecuente que el portador concluya de forma errónea que la "variante sin clasificar" es con toda probabilidad la causante del cáncer en la familia (la realidad es bien distinta, lo más probable es que se trate simplemente de una variante genética rara sin implicaciones clínicas).

Por otro lado, si queremos contribuir a "clasificar clínicamente la variante", es probable que necesitemos involucrar a familiares para intentar hacer estudios de cosegregación, para recabar información clínica adicional, informes de anatomía patológica, solicitar nueva extracción de sangre para aislar ARN, etc. Todo ello muchas veces no contribuye más que a extender el estrés emocional a otros familiares sin aportar (en la mayoría de los casos) ningún beneficio concreto a la familia.

Finalmente, a todo ello hay que sumar un problema importante de falta de estandarización en la clasificación clínica de variantes. En general, distintas variantes genéticas se asocian con distintos niveles de evidencia apoyando o descartando su potencial relevancia clínica. Además, las guías o documentos consenso disponibles en la actualidad están lejos de resolver todos los casos y su uso no está particularmente extendido. El resultado final es que una misma variante genética pueda asociarse a clasificaciones clínicas distintas (en algunos casos abiertamente contradictorias) en diferentes bases de datos y/o informes de laboratorio.

### **4. CLASIFICACIÓN DE VARIANTES GENÉTICAS IDENTIFICADAS TRAS REALIZAR UN TEST GENÉTICO.**

Entre el 5 % y el 20 % (según la población analizada) de los individuos a los que se les realiza un test genético BRCA son portadores de una variante sin clasificar. El porcentaje puede ser mucho mayor en el caso de síndromes hereditarios más raros, donde disponemos de mucha menos información.

De este modo, la clasificación clínica de variantes se convierte en una actividad laboriosa pero indispensable si se pretende ofrecer un consejo genético de calidad. Como ya hemos comentado, el proceso de clasificación clínica de variantes genéticas está lejos de haberse estandarizado. No obstante, existen diversas estrategias que se pueden utilizar con el objetivo de clasificar estas variantes. De especial relevancia es la calidad de la información que cada una de las estrategias puede aportar y la cuantificación estadística de esa información. De manera general, las evidencias se pueden clasificar en directas e indirectas. Las directas son aquellas que miden una asociación entre la variante y la enfermedad. Las indirectas, por su parte, se basan en el efecto observado o predicho de la variante sobre aspectos funcionales del gen en cuestión, que se asumen como determinantes de la patogénesis.

## **4.1 Evidencias directas:**

### **4.1.1- Análisis de la cosegregación**

Es la evidencia más directa y menos ambigua. Para este análisis se estudian familiares del probando, tanto afectos como sanos, y se observa si la variante segrega con la enfermedad o se hereda de forma independiente. Estas aproximaciones resultan útiles sobre todo para establecer la no patogenicidad de una variante. Se realiza asumiendo para la variante sin clasificar unos valores de penetrancia y frecuencia, que se infieren de los valores de aquellas variantes patogénicas conocidas. Como resultado se obtiene una medida estadística de *LOD score*. Un *LOD score* >3 demuestra una evidencia fuerte de cosegregación. Aunque la cosegregación es un método robusto, su poder estadístico depende del número de individuos de una misma familia analizados. Desafortunadamente, la mayoría de las familias en el contexto del consejo genético no poseen el tamaño suficiente para este tipo de análisis y además puede ser complejo el obtener muestras de múltiples individuos afectos (estas limitaciones se pueden compensar en parte integrando datos de distintas familias portadoras de una misma variante). También hay que considerar la posible existencia de fenocopias en las familias (casos esporádicos integrados dentro de un contexto familiar), un problema especialmente relevante en el caso de genes que confieren susceptibilidad a desarrollar cánceres de alta incidencia en la población general, como pueden ser el cáncer de mama o el cáncer colorrectal. Por otra parte, aunque se confirme la cosegregación, no se puede descartar que la variante sea únicamente un marcador ligado a una alteración patogénica no detectada.

### **4.1.2- Estudios caso-control**

Se basa en la comparación de la frecuencia de la variante de interés en una serie de casos y de controles. Metodológicamente es un método robusto que ofrece como resultado la medida estadística de riesgo relativo. Aunque es un método muy útil para estimar el riesgo conferido por variantes genéticas comunes (frecuencia >1 %), la mayoría de las VSC

son individualmente muy raras (frecuencia  $<0,1\%$ ) y el tamaño muestral requerido para obtener un resultado significativo es tan grande que hasta fechas recientes este tipo de estudios resultaba inabordable. En la práctica, es difícil que un estudio caso-control pueda establecer el papel patogénico de una VSC, pero es relativamente sencillo que pueda descartarlo. En concreto, se considera que una VSC con frecuencia del 1 % o superior en controles es muy improbable que se trate de una variante patogénica de penetrancia elevada. En este sentido, resultan particularmente útiles los datos del proyecto internacional 1000 Genomes (<http://www.1000genomes.org>), que proporcionan la frecuencia de numerosas variantes genéticas en 5 poblaciones humanas control de origen europeo, americano, asiático (dos poblaciones) y africano. También es común el uso de las frecuencias proporcionadas por los proyectos internacionales *Exome Sequencing Project* (ESP) (<http://evs.gs.washington.edu/EVS/>), y *Exome Aggregation Consortium* (ExAC) (<http://exac.broadinstitute.org/>), aunque es importante destacar que los datos proporcionados por ESP y ExAC no pueden considerarse controles estrictos, pues incluyen cohortes de pacientes con diferentes patologías.

#### 4.1.3- Historia personal y familiar

Se basa en el hecho de que las variantes patogénicas tienden a aparecer de manera más frecuente en familias con una mayor incidencia de casos de cáncer relacionados con el síndrome. El análisis de una variante de significado incierto en familias con distintos grados de historia familiar permite analizar su distribución y compararla con la distribución de las variantes patogénicas de la misma población. La ventaja de esta aproximación es que únicamente requiere el análisis del genotipo de un individuo por familia. La desventaja es que requiere información detallada de todas las familias analizadas en una población y no sólo de aquellas portadoras de la variante sin clasificar. Dentro de la información de utilidad que se puede obtener de las familias, se pueden incluir tanto los datos que se recogen habitualmente en la práctica clínica (edades de diagnóstico y tipos de cáncer) como los datos anatomopatológicos. Por ejemplo, la expresión del RE, RP y HER2 en cáncer de mama puede ser útil para clasificar variantes en *BRCA1*.

#### 4.1.4- Co-ocurrencia

En algunos síndromes de cáncer hereditario, como el que nos atañe en este caso clínico, la inactivación de las dos copias (materna y paterna) del gen o genes responsables de la enfermedad causa letalidad embrionaria. Experimentos en ratones han demostrado que la delección de los dos alelos del gen *BRCA1* o *BRCA2* es letal. En base a esta letalidad embrionaria, clásicamente se asumió que no existían individuos portadores de mutaciones bialélicas en estos genes. Recientemente se han descrito humanos portadores bialélicos de mutaciones patogénicas en *BRCA1* (dos casos) y *BRCA2* (unas pocas decenas) que presentan numerosas anomalías congénitas y manifestaciones clínicas características del sín-

drome recesivo Anemia de Fanconi (OMIM: 227650). Estos hallazgos muestran la posibilidad de existencia de mutaciones bialélicas en ambos genes, aunque produciendo fenotipos mucho más extremos y fácilmente distinguibles del síndrome de mama y ovario hereditario.

Traducido en términos prácticos: en ausencia de fenotipos extremos la demostración de la presencia de una variante patogénica en una copia del gen y de una variante sin clasificar en la otra copia del gen (fase en *trans*) proporciona una evidencia muy fuerte en contra de la patogenicidad de la variante sin clasificar. Como desventaja, para demostrar la fase de dos variantes a menudo se requiere genotipar a múltiples individuos de una familia.

## 4.2 Evidencias Indirectas

### 4.2.1- Análisis *in silico* y experimental del *splicing*

El procesamiento correcto de los intrones es un requisito imprescindible para la síntesis de una proteína. Variantes genéticas que producen defectos en el *splicing* han demostrado ser la causa de múltiples enfermedades hereditarias.

El *splicing* es un proceso complejo en el que intervienen por una parte un complejo riboproteico conocido como spliceosoma y por otra parte multitud de señales presentes en la secuencia de ARN mensajero (ARNm). Dentro de las señales presentes en el ARNm, el sitio 5' (o donador), el sitio 3' (o aceptor) y el sitio de ramificación definen los límites entre el exón y el intrón y están implicados directamente en la interacción con el spliceosoma y en las reacciones de procesamiento del ARNm. Estas señales, a pesar de su papel clave, están caracterizadas por un grado de conservación moderado. Las únicas posiciones conservadas completamente son los primeros/últimos 2 nucleótidos del intrón. Estas secuencias se conocen como sitios canónicos o sitios consenso de *splicing*. Otro tipo de señales conocidas como potenciadores e inhibidores aparecen distribuidas a lo largo del ARNm y actúan como reguladores del *splicing*, permitiendo por ejemplo la existencia de fenómenos de *splicing* alternativo.

El análisis del efecto sobre el *splicing* de una variante puede aportar evidencias en la clasificación de una VSC. Sus ventajas son:

- i) La existencia de programas que predicen la probabilidad con la que un cambio de secuencia alterará los patrones normales de *splicing* (en general, estos programas predicen adecuadamente la existencia de alteraciones relacionadas con los donadores y aceptores de *splicing*, pero no así las alteraciones en señales reguladoras. Se puede encontrar información adicional sobre este tipo de herramientas en el sitio web <http://www.ngri.org.uk/Manchester/page/splice-site-prediction-tools>).
- ii) La facilidad del diseño experimental en el análisis del *splicing*.
- iii) La metodología es esencialmente idéntica independientemente del gen que estemos analizando.

iiii) Su utilidad no está restringida en modo alguno al análisis de variantes intrónicas. Las variantes exónicas también tienen el potencial de afectar al *splicing*, independientemente de que se traten de cambios sinónimos, cambios *missense*, o incluso cambios *nonsense* e *Indels*.

No obstante, el análisis del *splicing* no está exento de ciertas desventajas, como son:

- i) La necesidad de disponer de una muestra de ARN (la extracción de ARN no suele estar incorporada en la rutina de los laboratorios que realizan test genéticos).
- ii) La posibilidad de detectar patrones de alteración en el *splicing* cuya relevancia clínica no es clara.
- iii) La necesidad de unos criterios claros en la clasificación de los resultados y de estandarización de los métodos de análisis.

La detección de una aberración clara en el *splicing* permite clasificar con mucha confianza a una variante genética como patogénica. Por el contrario, la detección de un patrón de *splicing* normal o no concluyente no aporta información que permita descartar el papel patogénico de la variante. Nótese que los programas *in silico* pueden predecir adecuadamente una alteración de *splicing*, pero sólo un análisis experimental permite determinar la naturaleza exacta de la alteración, y por tanto su probabilidad de que sea patogénica.

#### **4.2.2- Análisis *in silico* y experimental del efecto de los cambios de aminoácido (variantes *missense*)**

Evidentemente, un cambio de aminoácido puede inactivar la función de la proteína codificada por el gen de interés y, de hecho, muchas de las VSC descritas en las bases de datos y en la literatura científica son cambios de este tipo. Sin embargo, a diferencia de las mutaciones *nonsense*, *Indels*, o de *splicing*, las variantes *missense* patogénicas no suelen distribuirse al azar a lo largo del gen, sino que se concentran en determinadas regiones que codifican para dominios funcionalmente relevantes. Así, por ejemplo, las variantes *missense* patogénicas en *TP53* y *BRCA2* se concentran en sus respectivos dominios de unión a DNA, las de *MLH1* se concentran en el dominio ATPasa, y las de *BRCA1* se localizan en los dominios RING y BRCT localizados respectivamente en los extremos N-terminal y C-terminal de la proteína.

De este modo, la manera de evaluar *in silico* la relevancia de los cambios de aminoácido se basa en la evidencia de que las posiciones que se mantienen conservadas en una proteína a lo largo de la evolución son relevantes para su función y en que la mayoría de cambios de aminoácido patogénicos conocidos se localizan en estas posiciones conservadas. La herramienta disponible para analizar la conservación de residuos es el alineamiento múltiple de secuencias proteicas. Una vez identificados los residuos conservados, se debe evaluar el efecto del cambio de aminoácido concreto en esa posición. Esta tarea no es sencilla



y se han desarrollado una serie de algoritmos que realizan el alineamiento múltiple de secuencias (obteniendo las secuencias del gen de interés de las bases de datos disponibles o partiendo de un alineamiento múltiple aportado por el usuario) y analizan el efecto de los cambios de aminoácido. Algunos ejemplos dentro de la variedad de algoritmos disponibles son SIFT, Panther y align-GVGD. Estos análisis se pueden complementar con una predicción del efecto de un cambio de aminoácido sobre la estructura de la proteína. Así, el algoritmo PolyPhen utiliza métodos basados en alineamientos múltiples, a los cuales se les añade información de la estructura cristalina de la proteína.

La calidad del alineamiento múltiple utilizado como referencia es clave en la calidad de los resultados de estos algoritmos (dependiendo de la variabilidad de la proteína en cuestión, la profundidad requerida –número de especies incluidas en el análisis- será diferente). Por ello align-GVGD incluye una serie de alineamientos múltiples revisados y de alta calidad para los genes de susceptibilidad al cáncer más relevantes. Por otra parte, align-GVGD ha sido calibrado con un conjunto de 70.000 individuos testados para mutaciones en *BRCA1/2*, lo cual ha permitido establecer una serie de grados en la predicción del efecto de los cambios de aminoácido. De esta manera cada grado correlaciona con una medida estadística de una probabilidad de patogenicidad.

Para obtener información complementaria sobre este tipo de análisis, se recomienda consultar la página <http://www.ngrl.org.uk/Manchester/page/missense-prediction-tool-catalogue>. En cualquier caso, ningún método de predicción *in silico* ha demostrado ser suficientemente robusto para clasificar por sí solo una variante *missense*, en particular si dicha clasificación se pretende trasladar a la clínica.

#### **4.2.3- Ensayos funcionales para la caracterización de variantes *missense*.**

A diferencia de los análisis de *splicing*, la aproximación experimental a los ensayos funcionales es muy dependiente de la actividad biológica codificada por el gen de interés. En el caso de *BRCA1* y *BRCA2*, su carácter multifuncional ha impulsado el desarrollo de múltiples ensayos *in vitro* diferentes, que a menudo analizan de manera separada la función de cada uno de los dominios funcionales. Aún no está claro cuáles son las funciones que específicamente están implicadas en la susceptibilidad al cáncer. Por ello, la capacidad de discriminar variantes patogénicas y no patogénicas debe ser validada para cada ensayo. Algunos ensayos que han sido validados para *BRCA1* son el test de activación de la transcripción y el test de actividad E3 ubiquitina ligasa. En *BRCA2* algunos test validados son el de reparación mediada por recombinación homóloga y el de amplificación del centrosoma.

Independientemente del gen de interés, los ensayos funcionales tienen, en el momento presente, dos limitaciones que dificultan su aplicación diagnóstica:

- i) muchas veces el resultado del ensayo funcional detecta pérdidas parciales de función que no está claro como traducir en niveles de riesgo y,
- ii) ante un resultado negativo (el ensayo funcional no detecta pérdida de función) es difícil saber si la variante realmente no afecta a la función, o por contra el ensayo funcional no interrogaba la función biológica afectada por la variante.

### 4.3 Modelo probabilístico multifactorial

Teniendo en cuenta las ventajas y desventajas de cada una de las evidencias (directas e indirectas) anteriormente descritas, se ha propuesto el uso de un modelo probabilístico multifactorial para la clasificación clínica de VSC. Esta aproximación se ha aplicado con éxito a la clasificación de variantes en *BRCA1*, *BRCA2*, *MLH1* y *MSH2*, aunque en principio podría aplicarse a VSC identificadas en cualquier gen causante de síndromes hereditarios. Este modelo bayesiano integra las múltiples evidencias para crear un valor de la probabilidad de la patogenicidad de una variante sin clasificar. Este modelo parte de una probabilidad *a priori* sobre la cual se integran evidencias para dar lugar a una probabilidad *a posteriori*. La clasificación de variantes se considera dinámica, en el sentido de que a medida que avanzan los conocimientos científicos la relevancia que adquieren ciertas evidencias puede variar. En lo que se refiere al cálculo de la probabilidad *a priori*, el modelo más frecuentemente utilizado hasta la fecha considera la probabilidad derivada de Align-GVGD, lo que implica limitar la aplicabilidad del modelo multifactorial a cambios *missense*. Sin embargo, modelos más recientes tienen en cuenta también la probabilidad *a priori* de que la variante afecte al *splicing*, lo que permite extender su aplicabilidad a variantes sinónimas e intrónicas (además de estimar mejor la probabilidad *a priori* en el caso de las variantes *missense*).

En el caso de *BRCA1* y *BRCA2*, las evidencias que se utilizan habitualmente para construir la probabilidad *a posteriori* son: el análisis de la cosegregación, la co-ocurrencia, la historia familiar y personal y los datos anatomopatológicos del tumor. Por su parte, aunque los análisis de *splicing* y funcionales pueden proporcionar datos relevantes en lo que se refiere a la clasificación de variantes, hasta la fecha no se han podido integrar en el modelo multifactorial.

En teoría, con el objetivo de establecer una clasificación objetiva y estandarizada, se debería aplicar el modelo probabilístico multifactorial a cada una de las nuevas variantes encontradas en *BRCA1/2*, independientemente del tipo que se trate. Sin embargo, el modelo es laborioso y aplicarlo a cada una de las variantes nuevas restaría eficacia y retrasaría la aplicación del consejo genético en muchos individuos. Por ello, únicamente se aplica el modelo multifactorial en aquellas variantes nuevas que presentan dificultad en su clasificación, como es el caso de los cambios de aminoácido, los cambios silenciosos, los cambios en intrones y las inserciones/deleciones de un pequeño número de aminoácidos.

Por su parte, aquellas variantes que producen un desplazamiento del marco de lectura del ARNm, aquellas que introducen un codón de parada prematuro o las localizadas en las posiciones canónicas de los sitios de *splicing* se consideran de fácil clasificación, permitiéndose clasificarlas de manera directa, como patogénicas sin requerir la aplicación del modelo multifactorial.

Sin embargo, hay que tener ciertas precauciones al asumir esta clasificación directa, puesto que podría llevar a errores. Especialmente hay que tener cautela con:

- i) codones de parada prematuros en el último exón puesto que pueden no implicar una pérdida de función de la proteína y,
- ii) variantes localizadas en posiciones canónicas de sitios de *splicing*. El hecho de que estas posiciones sean las únicas conservadas sugiere que cualquier cambio producirá aberraciones en el procesado de los intrones. No obstante, el resultado concreto de la aberración del *splicing* influirá en la clasificación de la variante. Mientras que aquellas aberraciones del *splicing* que producen un cambio en la fase de lectura del ARNm se clasificarán como patogénicas con un alto grado de evidencia, aquellas que produzcan la delección/inserción de un pequeño número de aminoácidos pasan a considerarse variantes de difícil clasificación. En la base de datos DBASS (<http://www.dbass.org.uk/>) aparecen recopilados ejemplos de variantes que alteran el *splicing* junto con el efecto concreto que producen. Se puede observar la incidencia tanto de aberraciones que producen cambios en la fase de lectura como de aquellas que no la alteran. Por ello, hay que tener especial precaución a la hora de clasificar de manera directa variantes situadas en sitios canónicos de *splicing*.

El grupo de trabajo de variantes sin clasificar de la Agencia Internacional para la Investigación del Cáncer (más conocida por sus siglas IARC) sugiere que, tras la aplicación del modelo multifactorial, las variantes se clasifiquen en base a su probabilidad de ser patogénica en 5 categorías (Tabla 1). Para cada clase se hace una recomendación con respecto a los puntos más importantes relacionados con el consejo genético, como son:

- i) La utilidad de dicha variante como test predictivo para la individualización del riesgo de cáncer y la posibilidad de extender el test al resto de los miembros de la familia.
- ii) La recomendación de realizar estudios de investigación posteriores, incluyendo el análisis de miembros adicionales de la familia.
- iii) Las recomendaciones a tener en cuenta en referencia a la susceptibilidad al cáncer en los individuos portadores de dicha variante.

Clase	Definición Cuantitativa (P)	Definición Cualitativa	¿Incluir la variante en el informe de laboratorio?	¿Test Genético predictivo en familiares?	¿Test Genético con fines científicos en familiares?	¿Seguimiento?
Clase-5	>0.99	Patogénica	Sí	Sí	No	Medidas de seguimiento y/o profilaxis correspondientes a individuos de alto riesgo
Clase-4	0.95-0.99	Probablemente Patogénica	Sí	Sí	Deseable	Medidas de seguimiento y/o profilaxis correspondientes a individuos de alto riesgo
Clase-3	0.05-0.949	Significado Clínico Desconocido	Sí	No	Deseable	Test genético irrelevante. Medidas basadas en la historia familiar y/o personal y en otros factores de riesgo
Clase-2	0.001-0.049	Probablemente No Patogénica	Sin recomendación específica	No	Deseable	Test genético irrelevante. Medidas basadas en la historia familiar y/o personal y en otros factores de riesgo
Clase-1	<0.001	No Patogénica	No	No	No	Test genético irrelevante. Medidas basadas en la historia familiar y/o personal y en otros factores de riesgo

**Tabla 1:** Clasificación clínica de variantes genéticas y su traslación a la clínica: y su traslación a la clínica.

(Tabla adaptada de Eccles et al, 2015).

## 5. BASES DE DATOS CON CLASIFICACIÓN CLÍNICA DE VARIANTES.

Cuando realizamos un test genético, lo más probable es que la inmensa mayoría de variantes que identifiquemos ya se hayan descrito en una o varias bases de datos asociadas a clasificaciones clínicas más o menos rigurosas. En esta situación, nuestra labor principal será evaluar la información contenida en dichas bases de datos. Para ello, debemos plantearnos las siguientes preguntas: i) ¿la base de datos contiene una única clasificación clínica por variante?, ii) ¿dicha clasificación está respaldada por un panel de expertos?, y iii) ¿se aporta la información detallada sobre la cual se sustenta dicha clasificación?. El último punto es esencial, ya que nos permitirá valorar si estamos de acuerdo o no con la clasificación propuesta.

Por ejemplo, la base de datos ClinVar es una herramienta on-line del NCBI (*National Center for Biotechnology Information*) norteamericano ([www.ncbi.nlm.nih.gov/clinvar/](http://www.ncbi.nlm.nih.gov/clinvar/)) que agrega distintas fuentes de información sobre variaciones genéticas y su relación con las enfermedades humanas. Cada variante recibe una clasificación clínica (*clinical significance*) mediante un sistema de 5 categorías, que en este caso concreto se expresa en la forma de texto: *pathogenic*, *likely-pathogenic*, *uncertain significance*, *likely-benign*, *benign*). Además, la clasificación clínica va acompañada de una referencia al nivel de confianza de dicha clasificación, expresado mediante un código de estrellas. Una estrella implica que la clasificación clínica está respaldada por un único laboratorio (*submitter*), o que distintos laboratorios discrepan en torno a la clasificación clínica de la variante (en este caso en la *clinical significance* se indican las discrepancias entre laboratorios). Dos estrellas implican que la clasificación clínica está respaldada por varios laboratorios que de forma independiente han llegado a la misma conclusión. Tres estrellas implican que un panel de expertos designado por ClinVar ha revisado la clasificación clínica. El nivel máximo de confianza, 4 estrellas, implica que la clasificación clínica está universalmente reconocida e incorporada en guías de manejo clínico.

Existen numerosas bases de datos, muchas de ellas específicas de genes concretos. Por ejemplo, la base de datos del Insight (<http://insight-group.org/variants/database>) que proporciona la clasificación clínica de numerosas variantes genéticas en los genes reparadores *MLH1*, *MSH2*, *MSH6* y *PMS2*, o la base de datos BRCAShare (<http://www.umd.be/>), que proporciona la clasificación clínica de numerosas variantes genéticas en *BRCA1* y *BRCA2*.

Aunque el código de estrellas es exclusivo de ClinVar, es fundamental valorar siempre de forma individual la calidad de la anotación clínica indicada en una base de datos antes de trasladarla a la clínica. Por esta misma razón, si una base de datos no aporta información suficiente para poder valorar adecuadamente la calidad de una anotación clínica, esa anotación no debería ser trasladada a la clínica.

## 6- CONSENSOS BÁSICOS PARA INFORMAR VARIANTES EN GENES DE SUSCEPTIBILIDAD AL CÁNCER.

A pesar de la falta de estandarización en el proceso de clasificación clínica de variantes, existen aspectos relacionados que hoy en día sí gozan de amplio consenso entre los profesionales, en particular en lo que se refiere a la comunicación y manejo de variantes relacionadas con la susceptibilidad genética al cáncer (Tabla 1). Así, se recomienda:

- Clasificar clínicamente las variantes genéticas de acuerdo al sistema de 5 categorías propuesto inicialmente por un panel de expertos de la IARC. Según este esquema, las variantes Clase-1 no tienen relevancia clínica (probabilidad de ser patogénicas,  $p < 0,001$ ), las variantes Clase-5 son patogénicas ( $p > 0,99$ ), y las variantes Clase-2, -3 y -4 se encuentran en situaciones intermedias. Las variantes Clase-3 engloban un rango de probabilidades muy amplio ( $0,05 \leq p \leq 0,99$ ) y, estrictamente hablando, equivalen a las variantes sin clasificar.
- No incluir en el informe genético las variantes Clase-1.
- Considerar las variantes Clase-4 y -5 equivalentes desde el punto de vista del manejo clínico de las familias. Dicho de otro modo, tanto los estudios que identifican variantes Clase-4 como los estudios que identifican variantes Clase-5 se consideran igualmente informativos.
- Considerar las variantes Clase-1,-2 y -3 equivalentes desde el punto de vista del manejo clínico de las familias. Dicho de otro modo, los estudios que identifican variantes Clase-3 se consideran no informativos, exactamente igual que los informes que identifican variantes Clase-1 y -2.
- Las variantes Clase-2, -3 y -4 merecen estudios de investigación para poder llegar a asignarles en el futuro una clasificación clínica definitiva. Lógicamente, en el caso de las variantes Clase-2 se espera que se pueda llegar a alcanzar la certeza de que no sean patogénicas (reclasificarlas como Clase-1), y en el caso de las variantes Clase-4 se espera que se pueda llegar a alcanzar la certeza de que sean patogénicas (reclasificarlas como Clase-5).

No obstante (y hay que tenerlo presente) es posible, aunque improbable, que una variante Clase-2 se termine reclasificando como Clase-5 y una variante Clase-4 termine reclasificándose como Clase-1.

## 7. RESOLUCIÓN DEL CASO CLÍNICO

Tras realizar el estudio de los genes *BRCA1* y *BRCA2* por secuenciación masiva se identificaron en el individuo probando 30 variantes genéticas (20 en *BRCA1* y 10 en *BRCA2*). El

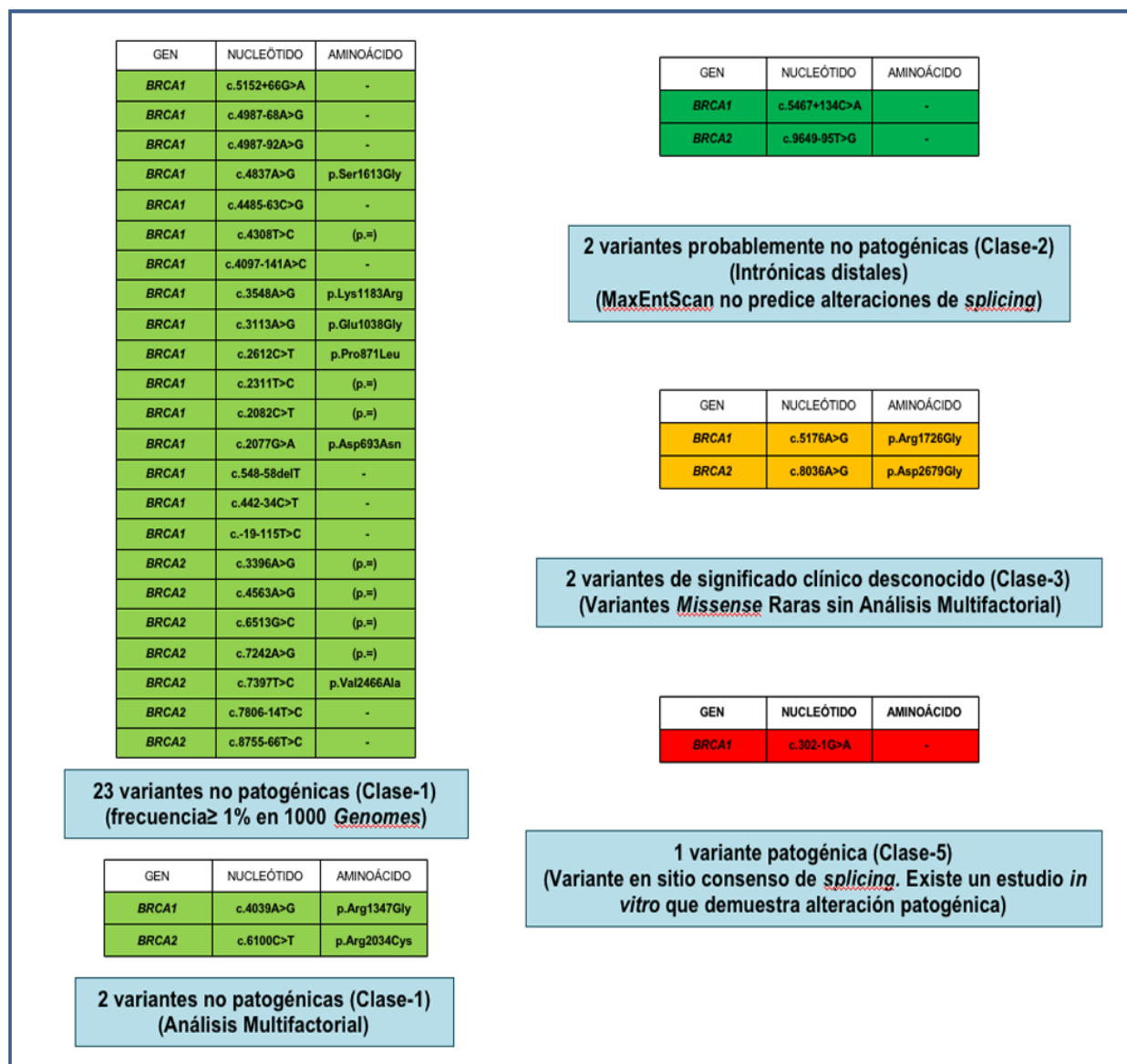


Figura 2: Clasificación clínica de variantes genéticas identificadas en el individuo probando del caso.

De las 30 variantes, 23 se clasificaron inmediatamente como no patológicas (Clase-1) basándonos en el hecho de que su frecuencia en la población era superior al 1 % ( $>0,01$ ). La regla del 1 % sólo debe aplicarse en poblaciones suficientemente grandes y heterogéneas (*admixture population*). Para fomentar la sistematización y evitar errores relevantes recomendamos utilizar la regla del 1 % consultando fundamentalmente los datos generados

por el consorcio internacional 1000 *Genomes*. Esta aproximación es de enorme utilidad para la identificación de variantes no patogénicas. En concreto, la frecuencia global descrita en 1000 *Genomes* para estas 23 variantes osciló entre 0,02 (*BRCA2* c.7397T>C) y 0,48 (*BRCA2* c.8755-66T>C).

Se identificaron dos variantes *missense*, *BRCA1* c.4039A>G (p.Arg1347Gly) y *BRCA2* c.6100C>T (p.Arg2034Cys) muy infrecuentes en la población. Tras consultar los datos del proyecto 1000 *Genomes* obtuvimos frecuencias globales de 0,0009 y 0,0018 respectivamente (valores muy por debajo del 0.01 recomendado). Sin embargo, las dos variantes *missense* cuentan con análisis multifactorial, cuyo resultado (<http://hci-exlovd.hci.utah.edu>) justificó en ambos casos su clasificación como no patogénicas (Clase-1).

Se identificaron dos variantes intrónicas, *BRCA1* c.5467+134C>A y *BRCA2* c.9649-95T>G, que no aparecían descritas en el proyecto 1000 *Genomes*. La falta absoluta de información en lo que se refería a estas variantes podía justificar ser informadas como variantes de significado clínico desconocido (Clase-3). *A priori* estas variantes intrónicas profundas no afectan a los sitios consenso de *splicing*. Sin embargo, esto no excluye que dichas variantes puedan afectar al *splicing*. Se recomienda realizar estudios *in silico* con el objetivo de predecir el efecto de las variantes a nivel de donadores y aceptores de *splicing* (posible creación de sitios *de novo*). Para ello, existen numerosos programas *on line* que predicen con alta sensibilidad y especificidad tal efecto. Estudios recientes sugieren que el programa MaxEntScan (accesible en diversos sitios web, por ejemplo en <http://www.umd.be/HSF3/>) es la mejor herramienta *in silico* para las predicciones de *splicing*. Utilizamos este programa para analizar estas dos variantes intrónicas no identificando la creación/destrucción de ningún sitio aceptor/donador de *splicing*. Por tanto, se consideraron las dos variantes probablemente no patogénicas (Clase-2).

Otras dos variantes *BRCA1* c.5176A>G (p.Arg1726Gly) y *BRCA2* c.8036A>G (p.Asp2679Gly) fueron cambios *missense* raros (sin datos de frecuencia en 1000 *Genomes*) localizados en dominios clínicamente relevantes de las proteínas BRCA1 y BRCA2. Sin embargo, ninguna de las dos variantes cuenta con análisis multifactorial, por lo que se recomendó informarlas como variantes de significado clínico desconocido (Clase-3).

Se identificó una variante intrónica en sitio consenso de *splicing* *BRCA1*c.302-1G>A (IVS6-1G>A). Como ya hemos mencionado, las posiciones intrónicas IVS±1,2 están extraordinariamente conservadas. Por ello, la probabilidad de que una variante IVS±1,2 altere el *splicing* es muy elevada. En la literatura científica identificamos un artículo que describe un ensayo *in vitro* de *splicing* que nos permitió clasificar a esta variante como patogénica (Clase-5).

En definitiva, la identificación de la variante *BRCA1*c.302-1G>A (IVS6-1G>A) nos permitió concluir que el test genético BRCA realizado en el probando de nuestro caso clínico es informativo. A partir de dicho resultado pudimos ofrecer el estudio de la variante pato-

génica al resto de familiares sanos y enfermos (mayores de 18 años) con el objetivo de ofrecer una estimación precisa del riesgo y un seguimiento clínico individualizado a cada uno de ellos.

## RESUMEN

La información proporcionada por un test genético BRCA (*BRCA1* y *BRCA2*) puede ser de enorme utilidad en el manejo de familias con sospecha de un síndrome de cáncer de mama y ovario hereditario. Para ello, el test genético debe identificar una variante genética patogénica en alguno de los dos genes. Dado que *BRCA1* y *BRCA2* se consideran genes supresores, se asume que cualquier variante genética que implique inactivación de uno de estos genes es patogénica.

Hay variantes genéticas que son reconocibles inmediatamente como patogénicas basándose exclusivamente en la naturaleza del cambio, sin necesidad de datos complementarios. Ejemplos típicos son las sustituciones y pequeñas inserciones/delecciones que introducen codones de parada prematuros. Muchas otras variantes (en realidad, la mayoría) se consideran no patogénicas por la sencilla razón de tener una frecuencia demasiado elevada en la población general. Finalmente, hay un tercer grupo de variantes que sin ser obvio que inactiven el gen (basándose exclusivamente en la naturaleza del cambio), son demasiado raras en la población general como para poder excluir un papel patogénico. En ausencia de datos complementarios, solemos describirlas como variantes de significado clínico incierto o variantes sin clasificar.

Lo cierto es que distintas variantes genéticas se asocian con distintos niveles de evidencia apoyando (o descartando) su potencial relevancia clínica, y la falta de guías y/o documentos consenso en que apoyarse dan como resultado un problema importante de falta de estandarización en la clasificación clínica de variantes.

## BIBLIOGRAFIA

**Couch FJ, Rasmussen LJ, Hofstra R, et al;** IARC Unclassified Genetic Variants Working Group. Assessment of functional effects of unclassified genetic variants. *Hum Mutat.* 2008;29(11):1314-26.

**Eccles DM, Mitchell G, Monteiro AN, et al;** ENIGMA Clinical Working Group. BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann Oncol.* 2015;26(10):2057-65.

**Jian X, Boerwinkle E, Liu X.** In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res.* 2014;42(22):13534-44.



**Lindor NM, Guidugli L, Wang X, et al.** A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Hum Mutat.* 2012;33(1):8-21.

**Padgett RA.** New connections between splicing and human disease. *Trends Genet.* 2012;28(4):147-54.

**Plon SE, Eccles DM, Easton D, et al;** IARC Unclassified Genetic Variants Working Group. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat.* 2008;29(11):1282-91.

**Richards S, Aziz N, Bale S, et al;** ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-24.

**Rosenthal ET, Bowles KR, Pruss D, et al.** Exceptions to the rule: Case studies in the prediction of pathogenicity for genetic variants in hereditary cancer genes. *Clin Genet.* 2015;88(6):533-41

**Ruiz de Garabay Ponce, G.** (2014). Síndromes hereditarios de cáncer de mama familiar: variantes de significado clínico incierto y consejo genético. (Tesis doctoral). Departamento de Bioquímica y Biología Molecular II. Facultad de Farmacia. Universidad Complutense de Madrid.

**Thompson BA, Spurdle AB, Plazzer JP, et al;** InSiGHT. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet.* 2014;46(2):107-15.

**Spurdle AB, Couch FJ, Hogervorst FB, et al;** IARC Unclassified Genetic Variants Working Group. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat.* 2008;29(11):1304-13.

---

## GENÉTICA MOLECULAR APLICADA AL DIAGNÓSTICO DE ENFERMEDADES HEREDITARIAS

Pilar Carrasco Salas (*presidenta*), Ana M<sup>a</sup> Sánchez de Abajo (*coordinadora*), Atocha Romero, Begoña Ezquieta, María Concepción Alonso, Cristina Torreira Banzas, Jesús Molano, María Santamaría González, Orland Diez, Josep Oriola Ambrós, Raquel Rodríguez, Hada Macher, Ana Cuesta.

### ACTIVIDADES FORMATIVAS DEL COMITÉ DE EDUCACIÓN

D. Balsells, B. Battikhi (*Residente*), R. Deulofeu, M. Gassó, N. Giménez, J.A. Lillo, A. Merino, A. Moreno, A. Peña (*Residente*), N. Rico, M. Rodríguez (*Presidente*), MC. Villà.

ISSN 1887-6463 – Mayo 2016 (recibido para publicación Marzo 2016).