



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

**“Clonación de 5 exones de variantes de splicing
potencialmente patogénicas de pacientes con cáncer de
mama hereditario”**

TESIS

Que para obtener el título de:

Bióloga

PRESENTA

Cristal Ramos Madrigal

DIRECTOR DE TESIS

Dr. Felipe Vaca Paniagua

Los Reyes Iztacala, Tlalnepantla, Estado de México, 2019





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dedicatoria

A mis padres, por ser en cada aspecto de mi vida el mejor ejemplo. A mi mamá por todo su cariño y atención, por escucharme siempre que lo necesité. A mi papá por su apoyo y por poner siempre como prioridad mi salud y bienestar.

A mis hermanos: Jazmín, Robin y Fany, por ser mi inspiración y fortaleza. A mis primos: Ivan, Andrea y Lupita por ser mis mejores amigos y por estar conmigo a la distancia todos estos años que he estado lejos.

A Fernando, por ser mi compañero constante y apoyo incondicional, fortaleciendo cada día los hilos que nos unen desde hace ya algunos años.

A todos ustedes les dedico mi esfuerzo y sus frutos. Los amo desde y hasta siempre.

Agradecimientos

Agradezco a la UNAM por la oportunidad de realizar mis estudios en tan bella universidad y a la FES Iztacala por todos estos años de crecimiento académico.

A mi tutor, el Dr. Felipe Vaca Paniagua, por su invitación a formar parte de su laboratorio y por darme la oportunidad de realizar mi tesis bajo su tutoría y dirección. Por apoyarme en los aspectos académicos así como deportivos y estar siempre pendiente de mi desarrollo.

A mis asesoras, la Dra. Clara y la Dra. Fabiola por su inmensa paciencia y por guiarme a lo largo de todo el proyecto.

A la Dra. Irasema Chirino, por su disposición y siempre instructivo asesoramiento.

Al Dr. Terrazas por siempre brindarme un espacio en su ocupada agenda, para revisar mi proyecto en cada punto de control a lo largo de su realización.

A todos mis amigos, que durante todo este proceso, supieron sacarme una sonrisa en días tristes. A Fernanda, Ana, Ernesto, Fernando, Héctor, Monse, Jimena, Luis y Sebastián. Han sido una segunda familia durante estos últimos dos años. Los quiero.

Y a mi entrenador el profesor Flavio Camacho por no dejar que me rindiera, por ser estricto en los momentos precisos y por ser flexible en los momentos que lo requirieron.

Investigación realizada gracias al Programa de Apoyo a Proyectos de Investigación Tecnológica PAPIIT de la UNAM IA205519.

*"Somos polvo de estrellas que piensa acerca de las estrellas.
Somos la forma en la que el universo se piensa a sí mismo"*

Carl Edward Sagan 1934- 1996

Índice

Lista de figuras.....	I
Índice de tablas.....	III
Glosario de abreviaturas.....	IV
1. Resumen.....	1
2. Abstract.....	2
3. Introducción.....	3
3.1 Definición de cáncer.....	3
3.1.1 Estadísticas.....	3
3.1.2 Global.....	3
3.1.3 México.....	3
3.2 Cáncer de mama.....	5
3.2.1 Clasificación histológica del cáncer de mama.....	5
3.2.2 Epidemiología.....	6
3.2.3 Factores de riesgo de cáncer de mama.....	6
3.2.3.1 Factores de riesgo genético.....	8
3.2.4 Alteraciones moleculares en el cáncer de mama.....	8
3.3 Maduración del pre-ARNm.....	10
3.3.1 Corte y empalme o <i>splicing</i>	10
3.4 Métodos para analizar la funcionalidad de variantes genéticas de sitios de corte y empalme.....	12
4. Antecedentes.....	13
5. Genes con mutaciones en los sitios canónicos de <i>splicing</i>	13
6. Justificación.....	18
7. Hipótesis.....	18
8. Objetivos.....	18
8.1 General.....	18
8.2 Particulares.....	18
9. Material y métodos.....	19
9.1 Análisis del efecto <i>in silico</i> de las 5 variantes de <i>splicing</i>	19
9.2 Diseño de oligonucleótidos.....	19
9.3 Amplificación de exones de interés por PCR.....	21
9.4 Clonación direccionada en el vector de expresión pET01.....	22
9.5 Tamizaje de colonias.....	23
9.6 Secuenciación Sanger.....	23
10. Resultados.....	24
10.1 Predicción del efecto <i>in silico</i> para cada variante.....	25
10.2 Diseño de los oligonucleótidos.....	26
10.3 Amplificación de los exones por PCR de punto final.....	28
10.4 Clonación direccionada en el vector de expresión pET01.....	29
10.5 Tamizaje de colonias.....	29
10.6 Análisis de la secuencia obtenida por secuenciación Sanger de cada variante wt y mt.....	35
11. Discusión.....	37
12. Conclusiones.....	39
13. Perspectivas.....	39
14. Referencias.....	40

Lista de figuras

Figura 1. Incidencia y mortalidad del cáncer en la población mundial en el año 2018. Imagen tomada y modificada de GLOBOCAN.

Figura 2. Incidencia de los diferentes tipos de cáncer en la población Mexicana, 2018. A) Mujeres B) Hombres. Tomada y modificada de GLOBOCAN (Bray *et al*, 2018).

Figura 3. Morfología y estructura de glándula mamaria; tipos principales de cáncer de mama. Imagen tomada y modificada de American Cancer Society: <http://www.cáncer.org>

Figura 4. Mecanismo de empalme de pre-ARNm.

Figura 5. Secuencias de reconocimiento del splicing. Tomado y modificado de Padgett, 2012.

Figura 6. Proceso secuencial de reparación global del ADN por escisión de nucleótidos. Elaboración propia.

Figura 7. Mecanismo de reparación de ruptura de ADN de doble cadena por recombinación homóloga.

Figura 8. Vía de reparación de apareamientos erróneos (MMR).

Figura 9. Esquema de la vía de señalización fosfatidil inositol 3 kinasa.

Figura 10. Mapa de restricción del plásmido con las enzimas *XhoI* y *BamHI* con el programa NEBcutter V2.0 de New England Biolabs.

Figura 11. Construcción de plásmido recombinante para el exón 5 de *ERCC3* con ayuda del programa Vector NTI® Express Designer.

Figura 12. PCR de punto final, se representan las temperaturas empleadas para cada paso y el tiempo de duración de éstas.

Figura 13. Mapa del vector pET01 tomada y modificada del kit Exontrap de MoBiTec.

Figura 14. Representación de los sitios de unión de oligonucleótidos 2Forward y 3Reverse del kit MoBiTec, empleados para la PCR en colonia así como de los fragmentos que genera en ausencia y presencia del inserto correspondiente.

Figura 15. Linealización del plásmido con *BamHI*-HF. Se observan las bandas de 4177 pb correspondientes a la longitud de pET01. Digestión con la enzima *PstI*-HF: Resultado obtenido (A) y resultado esperado (B), obtenido con el programa NEBcutter de NEB.

Figura 16. Sitios de hibridación de los oligonucleótidos, flanqueando el sitio aceptor/donador mutado.

Figura 17. Estandarización temperaturas de hibridación de los oligonucleótidos para cada variante de splicing. Para *ERCC3 MSH6* y *PTEN* se determinó como Tm óptima la de 55 °C. Se mejoraron las condiciones de hibridación al cambiar la Tm para *FANCB* (52 Y 54°C) y *RECQL4* (60 y 62°C), se determinó como óptima a 52°C y 60°C respectivamente. Geles de agarosa al 1%.

Figura 18. Exones amplificados por PCR de punto final, a partir de ADN genómico wt y mt (A) y a partir de ADN de biblioteca (B). Los productos de PCR corresponden al tamaño esperado del

exón amplificado para cada variante y se observan bandas únicas que indican la especificidad de los mismos.

Figura 19. Placas de la variante *PTEN* con sus respectivos controles positivo y negativo y sus colonias bacterianas a 22 horas de cultivo, contadas y numeradas.

Figura 20. Resultado de la PCR en colonia para *FANCB wt y mt y ERCC3 wt*. 5/5 colonias analizadas wt y 4/5 colonias analizadas mt para *FANCB* y 5/5 colonias analizadas *ERCC3 wt* resultaron positivas para este análisis, con un tamaño de banda de 1163 y 1096 pb respectivamente. Gel de agarosa al 1%.

Figura 21. Resultado de la PCR en colonia para *ERCC3 mt*. 15/15 colonias analizadas resultaron positivas para este análisis, con un tamaño de banda de 1096 pb. Gel de agarosa al 1%.

Figura 22. Resultado de la PCR en colonia para *MSH6 y RECQL4 wt y mt*. 3/5 colonias analizadas wt y 4/5 colonias analizadas mt de *MSH6* y 2/3 colonias analizadas wt y 3/5, analizadas mt para *RECQL4* resultaron positivas para este análisis, con un tamaño de banda de 1129 y 1070 pb respectivamente. Gel de agarosa al 1%.

Figura 23. Resultado de la PCR en colonia para *PTEN wt y mt*. 10/10 colonias analizadas wt y 8/10 colonias analizadas mt resultaron positivas para este análisis, con un tamaño de banda de 1376 pb. Gel de agarosa al 1%.

Figura 24. Resultado del análisis de restricción de *FANCB wt y mt* de las minipreps positivas. Se observa la construcción linealizada con *MfeI*-HF (4501 pb) en las muestras 1, 2, 3, 4 wt y 1, 2, 3 y 5 mt lo que indica la presencia del inserto *FANCB* en éstas.

Figura 25. Resultado del análisis de restricción de *ERCC3 wt y mt* de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4477 pb) y con *Bam*HI-HF y *Xho*I, la liberación del inserto de 300 pb *ERCC3* en la muestra 1 y 2 wt y 1, 2, 3 y 5 mt, lo que indica la presencia del inserto *ERCC3* en éstas.

Figura 26. Resultado del análisis de restricción de *MSH6 wt y mt* de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4467 pb), la liberación del inserto *MSH6* de 333 pb con *Bam*HI-HF y *Xho*I y la presencia de los fragmentos de 3,252 pb y 1200 pb tras la digestión con *Nsi*I, en las muestras 2, 3 y 4 wt y 1, 2, 4 y 5 mt, lo que indica la presencia del inserto *MSH6* en éstas.

Figura 27. Resultado del análisis de restricción de *PTEN wt y mt* de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4714 pb), y la liberación del inserto *PTEN* de 580 pb con *Bam*HI-HF y *Xho*I, en las muestras 1, 2, 3, 4 y 5 wt y 1, 3, 4, 5 y 6 mt, lo que indica la presencia del inserto *PTEN* en éstas.

Figura 28. Resultado del análisis de restricción de *RECQL4 wt y mt* de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4408 pb) y con *Bam*HI-HF y *Xho*I, la liberación del inserto de 274 pb *RECQL4* en la muestra 1 y 2 wt y 1, 3 y 4 mt, lo que indica la presencia del inserto *RECQL4* en éstas.

Figura 29. Resultado de la secuenciación Sanger. Electroferogramas wt y mt de cada variante.

Índice de tablas

Tabla 1. Probables factores de riesgo para el cáncer de mama establecidos.....	7
Tabla 2. Características de oligonucleótidos para el Exón 5 del gen <i>ERCC3</i>	20
Tabla 3. Resultado del análisis <i>in silico</i> de las mutaciones en los sitios aceptor/donador para cada variante (exones en mayúsculas e intrones en minúsculas, mutación en rojo).....	25
Tabla 4. Oligonucleótidos diseñados para la amplificación de los exones de interés.....	26
Tabla 5. Conteo de colonias obtenidas de cada transformación y cultivo en medio selectivo con ampicilina para cada variante <i>wt</i> y <i>mt</i>	30

Glosario de abreviaturas

INEGI: Instituto Nacional de Estadística y Geografía

NIH: National Institutes of Health

SHCMO: Síndrome hereditario de cáncer de mama y ovario

CM: Cáncer de mama

OMS: Organización Mundial de la Salud

VUS: Variante de significado clínico incierto

wt: variante silvestre

mt: variante mutante

RS virus: El virus del sarcoma de Rous

LTR: Repetición terminal larga (LTR)

ng: nanogramos

FA: Anemia de Fanconi

1. Resumen

El cáncer de mama (CM) es una neoplasia caracterizada por el crecimiento anormal de las células que forman el tejido mamario. Esta enfermedad ocupa el primer lugar en incidencia y mortalidad en cáncer de mujeres a nivel mundial y cada año va en aumento. De los factores de riesgo, la predisposición genética es responsable del 5-10% de todos los casos de CM, pero se ha observado que además está relacionada al desarrollo temprano de la enfermedad y a características clínicas de mayor agresividad.

De los diferentes tipos de mutaciones, estudios recientes han demostrado la relevancia de mutaciones que afectan el proceso de corte y empalme en varios tipos de cáncer. Se ha propuesto que un 60% de las mutaciones que causan enfermedades genéticas alteran el proceso de splicing, por lo que en los portadores de estas alteraciones a nivel germinal, éste es un mecanismo relevante en la susceptibilidad a CM hereditario.

La interpretación de las numerosas variantes de secuencia de significado biológico y clínico desconocido que se encuentran en los estudios genéticos de pacientes, representa un desafío importante en el diagnóstico molecular y la susceptibilidad al cáncer. Una fracción de variantes de significado clínico incierto (VUS) puede ser patogénica porque afecta el empalme de ARNm. En un estudio reciente realizado en nuestro laboratorio, analizaron un panel de 143 genes en 327 pacientes mexicanas con historial heredo-familiar de Síndrome Hereditario de Cáncer de Mama y Ovario (SHCMO) en búsqueda de nuevos alelos patogénicos que contribuyan al desarrollo de dicho síndrome, obteniendo como resultado un 10.8% con variantes VUS. De éstas variantes se seleccionaron 5 potencialmente patogénicas, presentes en sitios canónicos de splicing en los genes: *ERCC3*, *FANCB*, *MSH6*, *PTEN* y *RECQL4*, para el presente trabajo.

Se realizó la clonación de los 5 exones completos en un vector minigen (pET01). En todos los casos se clonó el segmento genómico que abarca el exón con la mutación de interés en el sitio aceptor o donador así como un segmento en ambos extremos 3' y 5' del exón con algunas bases de los dos intrones aledaños. Lo anterior se realizó por PCR a partir de ADN genómico de cada paciente mutante. Como control se empleó ADN de individuos sin cada mutación (genotipo silvestre). Los genotipos de los exones mutantes y silvestres clonados se verificaron por secuenciación tipo Sanger para verificar que no se hubiesen generado mutaciones adicionales que interfirieran con la correcta interpretación del efecto en el corte y empalme en ensayos posteriores. Con estas construcciones se analizará en un futuro trabajo el procesamiento de los 5 exones mutantes en modelos experimentales de células tumorales humanas para evaluar el impacto a nivel funcional de dichas variantes genéticas mediante el ensayo de minigen.

2. Abstract

Breast cancer is a neoplasm characterized by abnormal growth of the cells that form the breast tissue. This disease occupies the first place in incidence and mortality in cancer of women worldwide and every year it is increasing. Of the risk factors, the genetic predisposition is responsible for 5-10% of all cases of breast cancer, but it has been observed that it is also related to the early development of the disease and to more aggressive clinical characteristics.

Of the different types of mutations, recent studies have shown the relevance of mutations that affect the splicing process in several types of cancer. It has been proposed that 60% of the mutations that cause genetic diseases alter the splicing process, so that in carriers of these alterations at the germinal level, this is a relevant mechanism in the susceptibility to hereditary breast cancer.

The interpretation of the numerous sequence variants of unknown biological and clinical significance found in the genetic studies of patients represents an important challenge in the molecular diagnosis and the susceptibility to cancer. A fraction of variants of uncertain clinical significance (VUS) may be pathogenic because it affects the splicing of mRNA. In a recent study conducted in our laboratory, they analyzed a panel of 143 genes in 327 Mexican patients with inherited family history of Hereditary Breast and Ovarian Cancer Syndrome (SHCMO) in search of new pathogenic alleles that contribute to the development of this syndrome, obtaining as a result 10.8% with VUS variants. Of these variants, 5 potentially pathogenic were selected, present in canonical sites of splicing in the genes: ERCC3, FANCB, MSH6, PTEN and RECQL4, for the present work.

Cloning of the 5 complete exons was performed in a minigene vector (pET01). In all cases, the genomic segment covering the exon with the mutation of interest at the acceptor or donor site was cloned, as well as a segment at both ends 3' and 5' of the exon with some bases of the two surrounding introns. The above was performed by PCR from genomic DNA of each mutant patient. As control, DNA from individuals without each mutation (wild genotype) was used. The genotypes of mutant exons and cloned wild-type exons were verified by Sanger-type sequencing to verify that no additional mutations had been generated that interfere with the correct interpretation of the effect in the splicing in subsequent trials. With these constructions, the processing of the 5 mutant exons in experimental models of human tumor cells will be analyzed in a future work to evaluate the impact at the functional level of said genetic variants through the minigene assay.

3. INTRODUCCIÓN

3.1. Definición de cáncer

Cáncer es el nombre común que recibe un conjunto de enfermedades relacionadas en las que se observa un proceso descontrolado en la división celular, resultado de dos procesos sucesivos: la proliferación de un grupo de células, denominado tumor o neoplasia, y la capacidad invasiva que les permite colonizar y proliferar en otros tejidos u órganos (metástasis) (Muñoz, 2004).

El cáncer es causado por la acumulación y selección sucesiva de alteraciones genéticas y epigenéticas, que desregulan en las células tumorales, los mecanismos de supervivencia, replicación, apoptosis, proliferación y el ciclo celular (Willingham *et al*, 2004). Estas alteraciones pueden ser provocadas por distintos agentes carcinógenos, como la radiación ionizante, ultravioleta, productos químicos procedentes de la industria, del humo del tabaco y de la contaminación en general, o de agentes infecciosos como el virus del papiloma humano o el virus de la hepatitis B, así como anomalías genéticas adquiridas durante la replicación normal del ADN (Instituto Nacional de Cancerología, 2004).

3.1.1. Estadísticas

3.1.1.1. Global

El cáncer es una de las principales causas de morbilidad y mortalidad en el mundo. En 2018 se registraron más de 18 millones de nuevos casos y ocasionó más de 9 millones de defunciones (Fig. 1) (Bray *et al*, 2018). Casi una de cada seis defunciones en el mundo se debe a esta enfermedad (OMS, 2017).

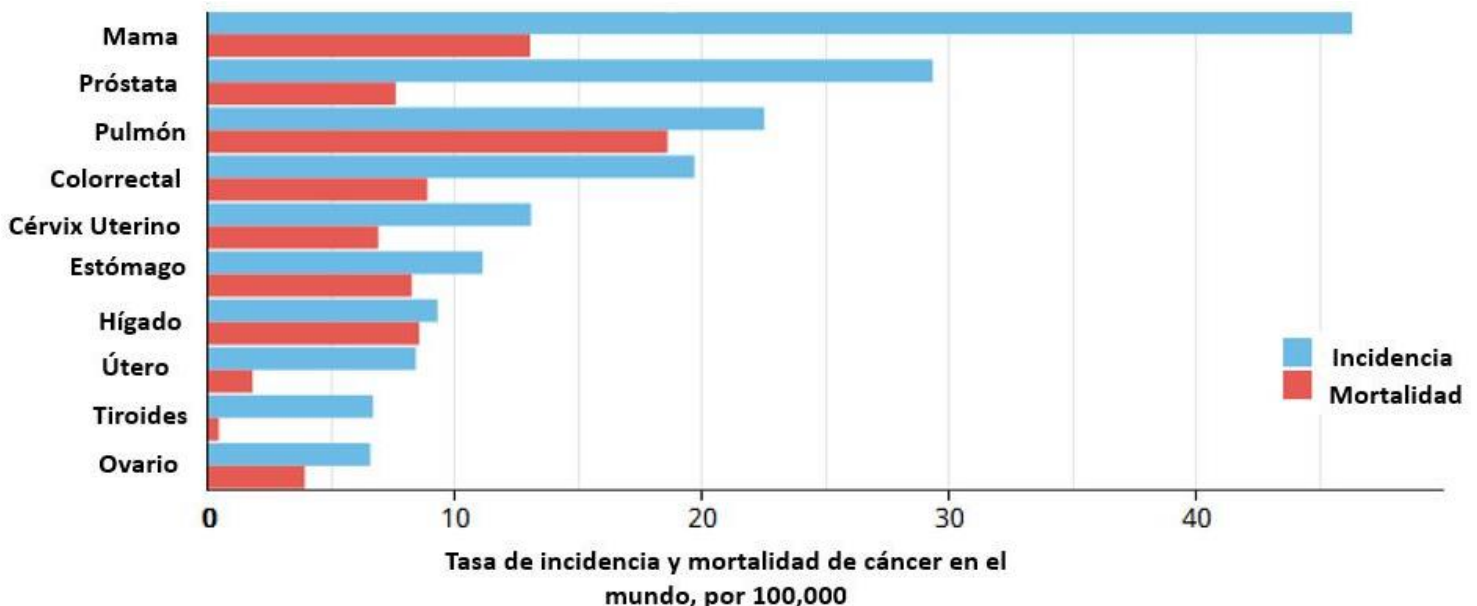


Figura 1. Incidencia y mortalidad del cáncer en la población mundial en el año 2018. Imagen tomada y modificada de GLOBOCAN.

3.1.1.2. México

El Instituto Nacional de Estadística y Geografía (INEGI) en 2015, señaló que en México los tumores malignos ocupan el tercer lugar en causa de defunciones con un 13%, precedido de las

enfermedades del sistema circulatorio con un 25.5% y las enfermedades endócrinas, nutricionales y metabólicas con un 17.5%.

Los principales tipos de cáncer en hombres mexicanos son: próstata (29.3%), colorrectal (9.1%), testículo (5.4%), pulmón (5.3), estómago (4.6%), leucemia (4.1), hígado (4%) y otros (38.2%); mientras que los más recurrentes en mujeres son: mama (26%), tiroides (9.7%), cérvix (7.5%), útero (6.9%), colorrectal (6.8%), ovario (4.5%), hígado (3.6%) y otros (35%) (Bray *et al*, 2018).

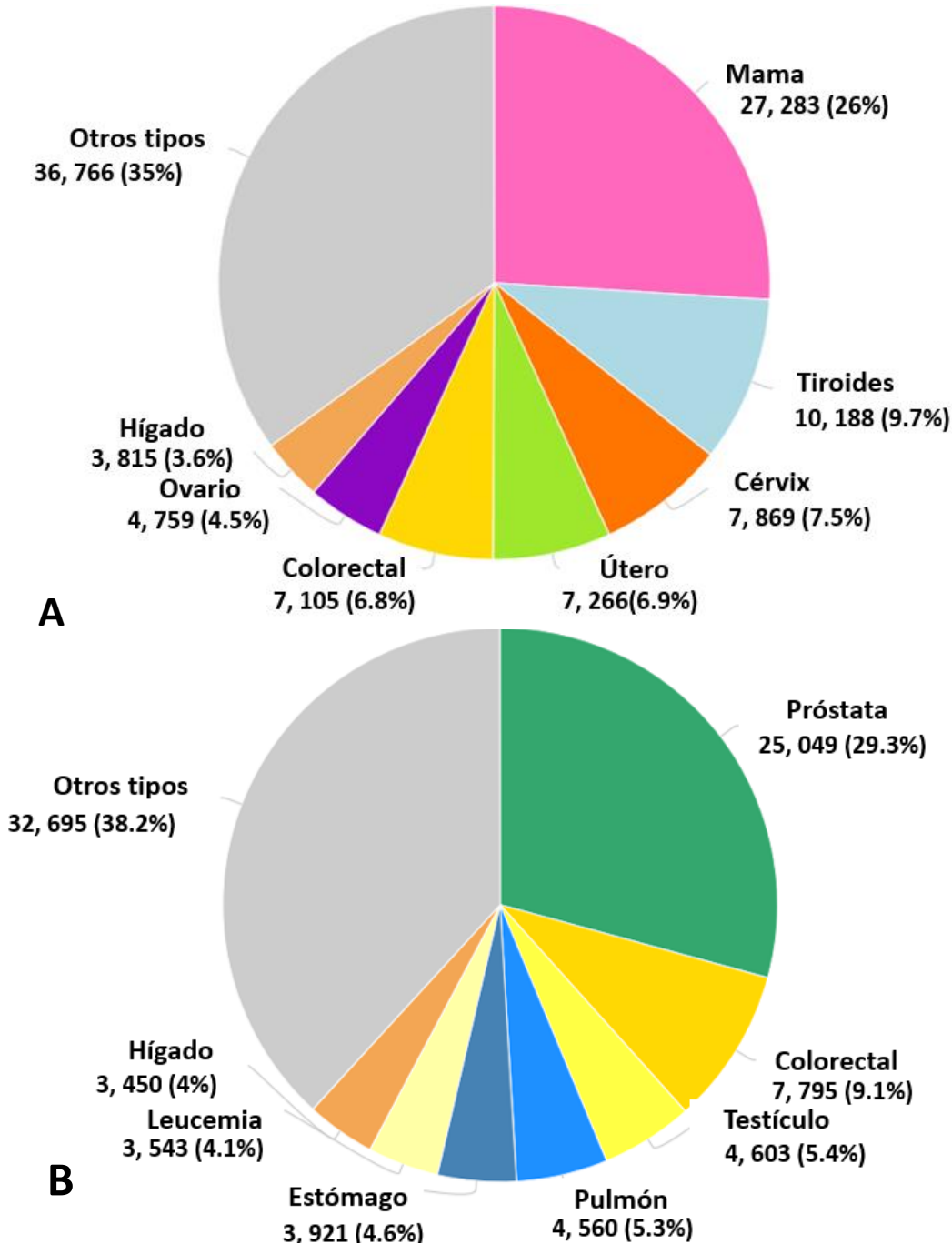


Figura 2. Incidencia de los diferentes tipos de cáncer en la población Mexicana, 2018. A) Mujeres B) Hombres. Tomada y modificada de GLOBOCAN.

De los diferentes tipos de cáncer, la organización mundial de la salud (OMS) considera al CM como uno de los padecimientos predominantes a nivel mundial, y como podemos observar en las figuras 1 y 2, el CM figura como el de mayor incidencia a nivel mundial para ambos sexos, incluso por encima de cánceres como el de pulmón y de próstata (OMS, 2012). En México, no es la excepción, el CM ocupa el primer lugar en incidencia en mujeres con un 26% al 2018 (Fig. 2A).

3.2. Cáncer de mama

La Sociedad Americana de Cáncer menciona que este cáncer se origina cuando las células en la región mamaria comienzan a crecer en forma descontrolada. Estas células normalmente forman un tumor que a menudo se puede observar en una radiografía o se puede palpar como una protuberancia (masa o bulto). El tumor es maligno (cáncer) si las células pueden crecer penetrando (invadiendo) los tejidos circundantes o propagándose (metástasis) a áreas distantes del cuerpo. También hay tumores benignos o no cancerosos de seno, los cuales no se propagan fuera de los senos y no ponen la vida en peligro. Sin embargo, algunas protuberancias benignas en los senos pueden aumentar el riesgo de padecer cáncer de seno.

3.2.1. Clasificación histológica del cáncer de mama

El tipo específico de células afectadas determina el tipo de CM. La mayoría de los cánceres de mama son carcinomas, tumores que se originan a partir de las células epiteliales mamarias. El tipo de CM más común es el carcinoma ductal, que empieza en el revestimiento de los conductos galactóforo, tubos delgados que llevan leche desde los lobulillos de la mama hasta el pezón). Otro tipo de CM es el carcinoma lobulillar, que empieza en los lobulillos (glándulas productoras de leche) (NIH, 2017: <https://www.cancer.gov/espanol/cancer/causas-prevencion/genetica>).

La mayoría de los cánceres de mama son a menudo un tipo de carcinoma llamado adenocarcinoma que se origina en las células del tejido glandular. Los adenocarcinomas del seno se originan en los conductos galactóforos o en los lobulillos (Dillon *et al*, 2010).

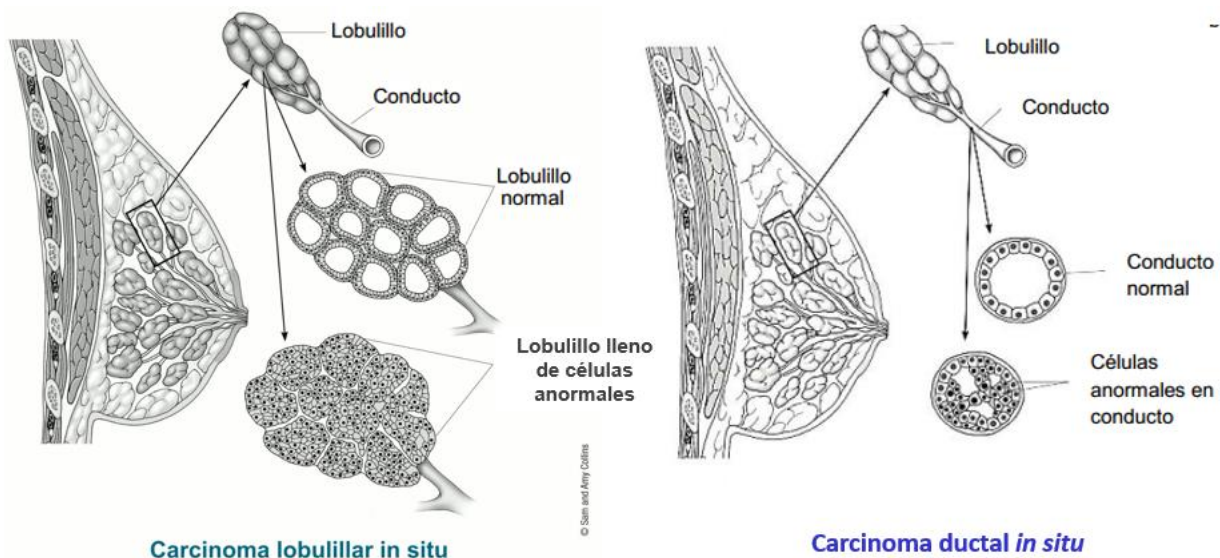


Figura 3. Morfología y estructura de glándula mamaria; tipos principales de cáncer de mama. Imagen tomada y modificada de [American Cancer Society: http://www.cancer.org](http://www.cancer.org).

Cuando los cánceres de seno se han propagado hacia el tejido mamario circundante se conocen como cánceres de seno invasivos. Hay muchos tipos diferentes de cáncer de seno invasivo, pero los más comunes se llaman carcinoma ductal infiltrante y carcinoma lobulillar invasivo. Por otro lado, existen otros tipos de cáncer de seno menos comunes como el cáncer de seno inflamatorio, un tipo infrecuente de cáncer de seno invasivo que es responsable de aproximadamente 1% a 5% de todos los casos de cáncer del seno, la enfermedad de Paget del pezón que representa sólo alrededor de 1% a 3%, los tumores filodes los cuales son tumores del seno que se originan en el tejido conectivo (estroma) del seno, los sarcomas de seno que constituyen menos del 1% de todos los cánceres de seno y los angiosarcomas se originan en las células que revisten los vasos sanguíneos o los vasos linfáticos y puede involucrar el tejido mamario o la piel del seno (Dillon *et al*, 2010).

3.2.2. Epidemiología

En 2018 se diagnosticaron con adenocarcinoma mamario 5.5 millones de mujeres a nivel mundial, cifra que incluso podría resultar subestimada. El cáncer de mama es el más comúnmente diagnosticado en mujeres (24.2%, es decir casi 1 de cada 4 de todos los nuevos casos de cáncer diagnosticado en mujeres en todo el mundo son cáncer de mama) (Bray *et al*, 2018).

En México, el INEGI en el año 2014 destacó que la principal causa de morbilidad en mujeres de 20 años o más se debe a esta neoplasia y es la segunda causa de mortalidad en las mujeres comprendidas en este intervalo de edad.

3.2.3. Factores de riesgo de cáncer de mama

Edad.

El riesgo de padecer cáncer de mama aumenta con la edad y en algunos entornos de altos recursos con poblaciones que tienen una esperanza de vida larga el riesgo para toda la vida podría llegar a 1 de cada 8 mujeres. La repercusión de la edad sobre la incidencia del cáncer de mama en los países de ingresos medios bajos está menos estudiada, pero se vuelve más importante a medida que la esperanza de vida mejora. (OMS, 2017).

Factores reproductivos. Los eventos relacionados a la salud reproductiva de la mujer, provocan cambios fisiológicos que pueden influir en el riesgo de padecer cáncer de mama (Sauter y Daly, 2010), como el aumento de las hormonas estrógeno y progesterona que al unirse a sus receptores, estimulan la proliferación en las células mamarias así es probable que los estrógenos, un grupo de hormonas sexuales femeninas, sean carcinógenos humanos endógenos. Aunque estas hormonas tienen funciones fisiológicas esenciales tanto en mujeres como en hombres, la exposición por largo tiempo o las concentraciones elevadas de estas hormonas están relacionadas con un mayor riesgo de cáncer de mama. Aumentos en la exposición pueden ser causados por empezar la menstruación a una edad temprana (<12 años), por llegar a la menopausia a una edad tardía (>54 años), tener más años cuando sucede el primer embarazo (40 años) (McPherson, *et al*, 2000) y nunca haber dado a luz (Kapoor, 2015). Por otro lado, la lactancia disminuye el riesgo a padecer CM en mujeres que han tenido varios hijos (Sauter y Daly, 2010) ya que aumenta la proporción de células diferenciadas en la mama, que en modelos animales, se ha demostrado, protege a las células expuestas a carcinógenos de la transformación maligna (Russo *et al*, 1982)

Factores ambientales y localización geográfica. Los factores ambientales se aprecian en la diferente incidencia y prevalencia de comunidades genéticamente homogéneas pero que se desarrollan en regiones geográficamente distintas (Peralta, 2007).). Un factor importante ligado

a la localización geográfica es la exposición a contaminantes. En un trabajo reciente realizado en México, se estudió la relación entre la presencia de aguas subterráneas contaminadas, contaminación química en suelos e industrias en 120 municipios y la incidencia de CM en mujeres, observándose una mayor incidencia de CM en las poblaciones cercanas a dichas zonas de contaminación (Castrezana, 2017).

A pesar de esto, se ha observado una diferencia significativa de 5 a 10 veces en las tasas de incidencia y mortalidad del cáncer de mama entre las áreas de bajo riesgo (Lejano Oriente, África y América del Sur) y de alto riesgo (América del Norte y Europa del Norte). La gran variación de la incidencia del cáncer de mama en diferentes regiones del mundo o dentro de ellas puede atribuirse a las diferencias genéticas entre poblaciones (Parkin *et al*, 1999), es decir a la raza o grupo étnico. Esta incidencia es mayor en mujeres blancas seguidas por mujeres de raza negra e hispánica y menor proporción en mujeres asiáticas. Estas disparidades podrían también reflejar las diferencias en el estilo de vida de estos grupos étnicos (Sauter y Daly, 2010).

Dieta y estilo de vida

Numerosos estudios epidemiológicos han observado por ejemplo, que realizar actividad física regular se asocia a una reducción en el riesgo de CM. De 44 estudios realizados en el mundo, 32 de éstos mostraron una reducción de riesgo promedio de entre 30–40% de desarrollar esta neoplasia (Friedenreich y Orenstein, 2002). Estudios realizados en mujeres mexicanas han mostrado un mayor riesgo de desarrollar CM en aquellas mujeres con pobre consumo de frutas y vegetales, consumo excesivo de carnes rojas y de grasas saturadas (Romieu *et al*, 2004) así como un consumo alto de carbohidratos (Romieu *et al*, 2009).

Tabla 1. Probables factores de riesgo para el cáncer de mama establecidos

Factor	Riesgo relativo	Grupos de alto riesgo
Edad	>10	mayor
Localización geográfica	5	Países desarrollados
Edad de la menarca	3	Menarca antes de los 11 años
Edad de la menopausia	2	Menopausia después de los 54 años
Edad al primer embarazo	3	Primer niño a los 40 años
Historia familiar	>/=2	Cáncer de mama de primer grado relativo, en la juventud
Enfermedad benigna previa	4-5	Hiperplasia atípica
Cáncer en el otro seno	>4	
Grupo socioeconómico	2	Grupos I y II
Dieta	1.5	Alto consumo de grasas saturadas
Peso		
Premenstrual	0.7	Índice de masa corporal >35
Postmenstrual	2	Índice de masa corporal >35
Consumo de alcohol	1.3	Consumo excesivo
Exposición a radiación ionizante	3	Exposición anormal en mujeres jóvenes después de los 10 años.

Hormonas exógenas		
Anticonceptivos orales	1.24	Uso común
Terapia de suplementos hormonales	1.35	Uso por ≥ 10 años
Dietilestilbestrol	2	Uso durante el embarazo

Tomado y modificado de K. McPherson, et al, 2000.

3.2.3.1 Factores de riesgo genéticos

De los factores de riesgo, la predisposición genética es responsable del 5-10% de todos los casos de cáncer de mama, y se ha observado que está relacionada al desarrollo temprano de la enfermedad (Ellisen y Haber, 1998).

La susceptibilidad al cáncer de mama tiene un tipo de herencia autosómica dominante con penetrancia incompleta, esto significa que ésta puede ser transmitida indistintamente del sexo del individuo y que algunos miembros de la familia pueden transmitir el gen anormal sin desarrollar el cáncer (McPherson *et al*, 2000).

Aproximadamente 20 a 30% de las mujeres con diagnóstico de cáncer de mama tiene al menos un familiar con la enfermedad (Slattery *et al*, 1993). El riesgo incrementa con el número de parientes afectados, una mujer que tiene una familiar en primer grado (madre, hermana, hija) tiene 2 o más veces riesgo de presentar la enfermedad antes de los 50 años y en edad más temprana que las parientes afectadas (Nelson, 2008). En un estudio en México con 115 mujeres, el antecedente heredofamiliar de cáncer de mama se asoció a la neoplasia con una estimación de riesgo de 3 veces más, siendo para el antecedente de cáncer de mama en familiar de segundo grado, con un riesgo de 4 veces más en el grupo de casos (Muñoz, 2010).

Existen genes implicados a la predisposición a cáncer de mama como *BRCA1* y *BRCA2* de los que ya se conoce el riesgo asociado a portar dichas mutaciones; para *BRCA1* es de 46-87%, y de un 38-84% para *BRCA2* (Antoniou *et al*, 2003). Sin embargo un alto porcentaje de los casos no se encuentran asociados a los genes *BRCA*.

3.2.4. Alteraciones moleculares en el cáncer de mama

Los mecanismos responsables de mantener y reparar el ADN pueden estar afectados debido a alteraciones moleculares o mutaciones (Valko *et al*, 2004). Los genes *BRCA1* o *BRCA2* tienen una alta relevancia en el riesgo a padecer cáncer de mama. Mujeres portadoras de una mutación en éstos, tienen un riesgo de un 50% a los 50 años y superior al 80% a lo largo de la vida. Existen algunos trabajos, sin embargo, que han encontrado mutaciones patogénicas en otros genes relacionados al desarrollo de cáncer hereditario, lo cual demuestra que *BRCA1* y *BRCA2* no son los únicos genes que participan en la epidemiología molecular del cáncer de mama (Quezada-Urban *et al*, 2018).

Se han descrito otros genes de alta penetrancia asociados al cáncer de mama como *TP53* (síndrome de Li-Fraumeni), *PTEN* (Síndrome de Cowden) y *STK11* (Síndrome de Peutz-Jeghers) (Chirivella y Garcés, 2018).

El gen *TP53* es el responsable del síndrome Li-Fraumeni con una herencia autosómica dominante. Es una enfermedad muy rara y se asocia a una alta probabilidad de desarrollar cáncer a lo largo de la vida (Chirivella y Garcés, 2018). Este gen se conoce como el “guardián del

genoma” ya que juega un papel importante en el control del ciclo celular y la apoptosis (González *et al*, 2009).

El gen *PTEN* es un gen supresor tumoral con penetrancia alta, aproximadamente del 80% (Nelen *et al*, 2001).

El gen *STK11* codifica una enzima multifuncional del tipo serina treonina importante en la transducción de señales por segundos mensajeros. Se hereda de forma autosómica dominante y cuando está mutado es responsable del síndrome de Peutz-Jeghers, con un riesgo de cáncer de mama entre el 30-54%. El riesgo depende de la edad y se ha descrito del 8% a los 40 años, del 13% a los 50 años, del 31% a los 60 años y del 45% a los 70 años (Harle *et al*, 2006).

Existen mutaciones en otros genes que son más frecuentes en la población y producen un riesgo menor de cáncer de mama como aquellas en los genes *CDH1*, *ATM*, *PALB2* y *CHEK2*. En un análisis retrospectivo de 337 mujeres que cumplían los criterios de la NCCN para el estudio de los genes *BRCA1/2*, en las que se utilizó un panel multigénico y se detectaron mutaciones en 25 de ellas. Las más frecuentes fueron en *PALB2* (23%), *CHEK2* (15%) y *ATM* (15%) (Kapoor *et al*, 2015).

El gen *ATM* que debido a la frecuencia estimada de los heterocigotos en la población general, sus mutaciones podrían asociarse a un 5% de los casos de CM (Díez *et al*, 2006).

El gen *cadherina 1 (CDH1)* es un gen supresor tumoral que requiere de la inactivación del segundo alelo por metilación, mutación somática o pérdida de heterocigosidad. Se hereda de forma autosómica dominante y el riesgo de desarrollar cáncer de mama asociada a este gen es de 39-52% (Pharoah *et al*, 2001).

El gen *CHEK2 (cell cycle checkpoint kinase 2)* es otro de los genes identificados que aumenta el riesgo de cáncer de mama. En un estudio en mujeres con una fuerte historia familiar de cáncer de mama y ovario detectó variantes en *CHEK2* en el 5% (Walsh *et al*, 2006). El riesgo acumulado de cáncer de mama en mujeres con mutación en *CHEK2* e historia familiar de cáncer de mama se ha estimado entre el 28 y el 37% (Easton *et al*, 2015).

El gen *PALB2 (partner and localizer of BRCA2)* es un gen relacionado con la anemia de fanconi. Las mutaciones en este gen aumentan el riesgo de cáncer de mama. El riesgo de cáncer de mama asociado a mutaciones en este gen es de un 14% a los 50 años y de 35% a los 70 años de edad (Antoniou *et al*, 2014).

A nivel de nucleótido, estas mutaciones pueden ser por sustitución, adición o delección, lo que altera la fisiología celular induciendo a la transformación de la misma (Dey *et al*, 2004). De los diferentes tipos de mutaciones, estudios recientes han demostrado la relevancia de mutaciones que afectan el proceso de corte y empalme en varios tipos de cáncer (Bonnal *et al*, 2012). Se ha propuesto que un 60% de las mutaciones que causan enfermedades genéticas alteran el proceso de splicing (López-Bigas *et al* 2005), por lo que es un mecanismo relevante en la susceptibilidad a cáncer de mama hereditario.

Las alteraciones en el splicing confieren ventajas selectivas a las células tumorales, cuando ocurren en genes supresores de tumores. Las mutaciones en los sitios de splicing alteran el procesamiento del ARNm y pueden encontrarse en las secuencias reguladoras del splicing o en los factores involucrados en dicha regulación (Yoshida *et al*, 2011 y Brooks *et al*, 2014).

3.3. Maduración del pre-ARNm

La información genética, antes de dar lugar a proteínas, genera moléculas de ARN mediante la transcripción de los genes y la maduración del ARNm (Eyras, 2016) que en células eucariontes consiste en tres etapas:

1. En una primera etapa, se lleva a cabo la adición de la estructura denominada “caperuza”, que es un nucleótido modificado de guanina, la 7-metilguanosina trifosfato. El extremo 5´ de todos los ARNm posee inicialmente un trifosfato derivado del primer trifosfato de nucleósido incorporado en el sitio donde inicia la síntesis de ARN. Una vez que se forma el extremo 5´ mediante la acción de la ARN polimerasa II, varias enzimas actúan para modificar ese extremo de la molécula. En el primer paso, una fosfohidrolasa elimina el último de los tres fosfatos convirtiendo el extremo terminal 5´ en difosfato. A continuación una guanilil transferasa añade un residuo de guanina en orientación invertida mediante una unión 5´ a 5´. En un último paso, a los grupos metilo se les añade tanto el casquete de guanina terminal, como el nucleótido situado en el extremo del ARN (Devlin, 2004). Esta caperuza es necesaria para el proceso normal de traducción del ARN y para mantener su estabilidad; esto es fundamental para el reconocimiento y el acceso apropiado del ribosoma (Herveg, 2006).
2. En un segundo proceso, la enzima poliA polimerasa añade al extremo 3' una secuencia larga de poliadenilato, de unos 50-250 residuos de adenosina, lo que se conoce como cola poliA. En éste, se requiere el reconocimiento de una secuencia hexámera “AAUAAA” por parte de una nucleasa que desdobra el pre-ARN situado río debajo de este sitio. Posteriormente del desdoblamiento, la polimerasa poliA añade las adenosinas. Ésta “cola” de poliadenilato protege al ARNm frente a la degradación en el núcleo y aumenta su vida media en el citosol, de modo que se pueda sintetizar mayor cantidad de proteína (Herveg, 2006).
3. En un tercer proceso, se lleva a cabo el splicing o corte y empalme, proceso de eliminación de secuencias no codificantes (“intrones”) y de empalme de las secuencias codificantes (“exones”). Para llevar a cabo este proceso de corte y empalme, existe en el núcleo una maquinaria molecular muy compleja, que se conoce con el nombre de “spliceosoma” un complejo conformado de ribonucleoproteínas (ARNnp) llamadas: U1, U2, U4, U5 y U6. (Wahl *et al*, 2009).

3.3.1 Corte y empalme o *splicing*

El corte y empalme requiere la identificación precisa de los límites entre las regiones intrónicas y exónicas. Los pre-ARNm de los vertebrados contienen la información necesaria para el reconocimiento de un intrón por parte de la maquinaria de splicing, a través de pequeñas secuencias altamente conservadas en la evolución. En el comienzo del intrón, se encuentra el sitio de splicing 5´ o donador; mientras que el extremo opuesto, son necesarias tres secuencias: el sitio de splicing 3´ o aceptor; el punto de ramificación, localizado a 18-40 nucleótidos río arriba del sitio aceptor y el tracto de polipirimidina, situado entre el punto de ramificación y el sitio aceptor (véase figura 4) (Ast, 2004).

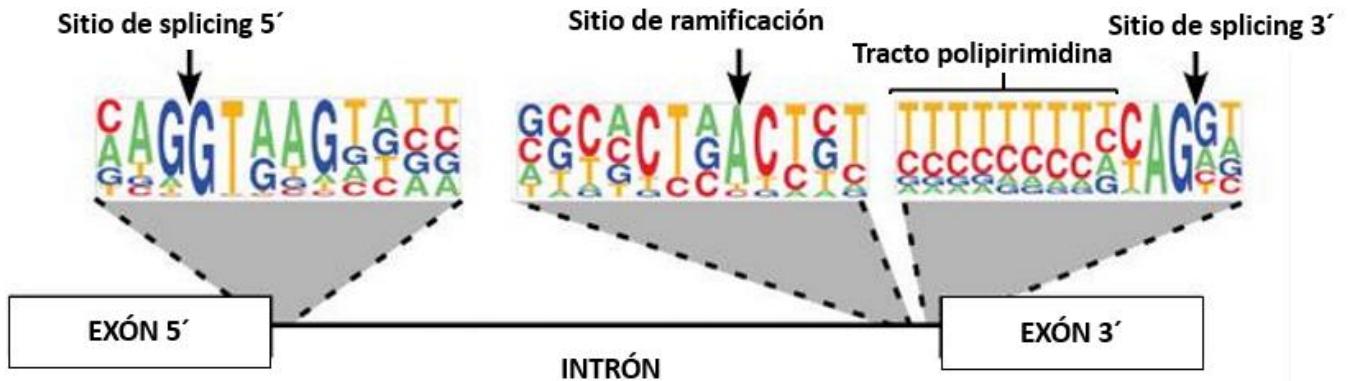


Figura 4. Representación gráfica de las secuencias de reconocimiento del splicing. Las bases están ordenadas por frecuencia de arriba abajo y el tamaño indica el grado de conservación evolutiva. Tomado y modificado de Padgett, 2012.

El proceso de corte y empalme puede verse en la figura 5 y se desarrolla en el siguiente orden: Primero U1 se une a la secuencia consenso GU del extremo 5' (1), después U2 se une al sitio de ramificación e hidroliza ATP generando un doblez en el intrón, causado por la interacción entre U1 y U2 (2). Posteriormente U5, U4 y U6 ARNnp trimerizan (3), U5 se une al exón 5' y U6 a U2 (4). Después U4 se disocia de U6 y éste último se une al extremo 5' del sitio de corte. A continuación U6 y U2 catalizan la primera reacción de transesterificación, cortando el extremo 5' del intrón y como resultado se forma una estructura en lazo característica, denominada lariat. (5) Enseguida realizan la segunda reacción de transesterificación, cortando el extremo 3' del intrón y liberando el lazo. Finalmente los exones son empalmados, lo que conlleva gasto de ATP y el complejo se disocia (6).

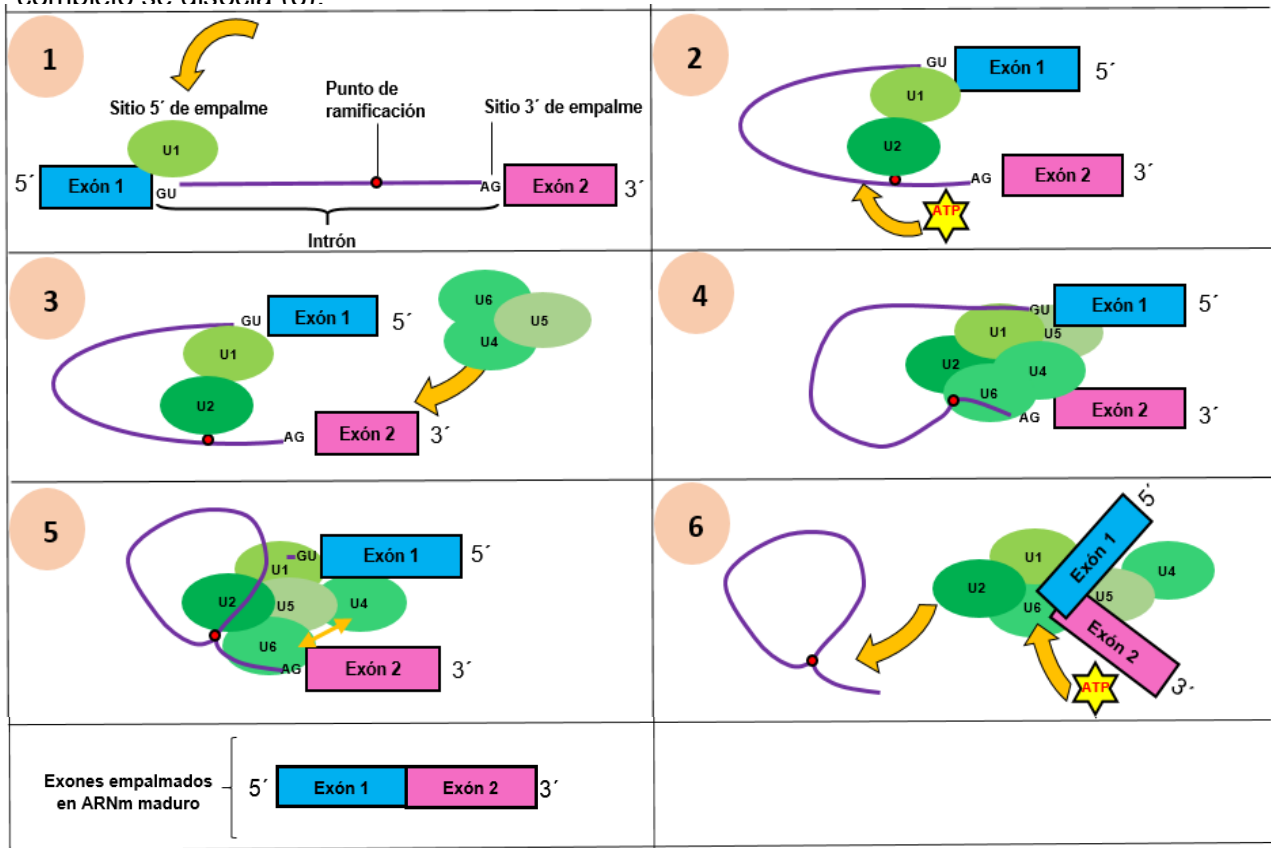


Figura 5. Proceso secuencial de corte y empalme. Se observan las ARNnp que conforman la maquinaria del spliceosoma y los eventos secuenciales de su ensamblaje, para realizar el corte y empalme del pre-ARNm. Elaboración propia.

3.4. Métodos para analizar la funcionalidad de variantes genéticas de sitios de corte y empalme

Existen varios estudios realizados que evalúan la funcionalidad del proceso de corte y empalme en genes implicados en el desarrollo de algunas enfermedades. En estos trabajos se utilizan minigenes híbridos. Un minigen es una construcción artificial, un fragmento de gen mínimo que incluye al menos un exón y las regiones intrónicas de control necesarias para que se exprese de la misma manera que un fragmento de gen tipo *wt*, en un plásmido o vector de expresión, que permite introducir las variantes detectadas en pacientes y analizar su efecto en el splicing (Velasco, 2014).

En este ensayo, los minigenes híbridos son transfectados en células de cultivo y el ARNm obtenido tras 48 horas es procesado por retrotranscripción para obtener el ADNc que permite la comparación del tamaño de producto del alelo mutante (*mt*) con el silvestre (*wt*) en un gel de agarosa. Esta estrategia se utiliza para investigar patrones de corte y empalme *in vivo* e *in vitro*, y mientras que los métodos tradicionales necesitan ARNm del paciente, cuya obtención resulta a veces complicada, los investigadores han diseñado estos minigenes que no lo necesitan.

En 2003, Baralle y colaboradores, identificaron una mutación que altera el corte y empalme del gen *NF1* utilizando muestras de ADN genómico y un ensayo de minigén. Probaron que la mutación en cuestión (exón 3 +5 G> C) afectó dramáticamente el procesamiento de pre-ARNm, causando un salto aberrante del exón 3 que lo eliminaba del ARNm.

En 2007, Pepe y colaboradores realizaron ensayos de splicing en células Y1, para comprobar si la mutación c655G>A podría alterar los mecanismos fisiológicos reguladores del exón 5 favoreciendo su exclusión. Descubrieron bajos niveles de ARNm sin el exón 5, que podían expresarse aun en presencia de la mutación, explicando así el fenotipo de déficit parcial observado en la paciente.

Hernández-Moro y colaboradores en 2013, realizaron un ensayo funcional con un minigén que contenía los exones 16-17 de *BRCA1*. Se probaron *in silico* con los programas NNSplice y Human splicing finder y se analizaron funcionalmente 34 mutaciones de variantes de splicing previamente reportadas en pacientes con cáncer de mama (CM) para identificar patrones anómalos de splicing. Encontraron que 13 de las variantes seleccionadas presentaban splicing aberrante y concluyeron que los resultados bioinformáticos no son suficientemente precisos para clasificar una mutación y una validación funcional es estrictamente requerida.

En otro estudio realizado por Kohmoto y colaboradores en 2015, por medio de un ensayo de minigen determinaron el efecto de una mutación en el corte y empalme del gen *COL11A1* en un paciente con síndrome de Styler. Esta mutación en el sitio donador de splicing, generó la pérdida del exón 3 en el ARNm, lo cual sugiere que esta mutación fue la causa del síndrome de Styler en este paciente.

En conjunto, estos ensayos funcionales han demostrado los efectos a nivel funcional de mutaciones en sitios canónicos de splicing, los cuales incluyen: (i) la pérdida del exón flanqueado por el sitio aceptor o donador mutado, (ii) la retención del intrón en cuestión o (iii) el correcto empalme de los exones. Este tipo de ensayos ayudan en la interpretación del impacto de estos cambios para poder determinar su importancia biológica y patológica (Baralle, 2005).

4. Antecedentes

Recientemente en nuestro laboratorio Quezada-Urban y colaboradores, realizaron un estudio de 143 genes de predisposición al cáncer en 327 pacientes mexicanas con historial heredo-familiar de SHCMO (síndrome hereditario de cáncer de mama y ovario), con el objetivo de detectar nuevos alelos patogénicos que posiblemente contribuyan al desarrollo de esta enfermedad. Se detectó, en el grupo de pacientes con cáncer, un total de 47 (16.6%) pacientes con variantes patogénicas, 72 (24%) con VUS (variantes de significado incierto) y 181 (60%) pacientes con resultado negativo (Quezada-Urban *et al*, 2018).

Dentro del grupo de las VUS, se detectaron 3 variantes con un cambio de nucleótido único en los genes *ERCC3*, *RECQL4*, *PTEN*, una inserción en el gen *FANCB* y una delección en el gen *MSH6* que afectaría el sitio donador o aceptor canónico de splicing, con las que se procedió a la realización de este trabajo.

5. Genes con mutaciones en los sitios canónicos de splicing

ERCC3

El gen ***ERCC3*** codifica para la proteína llamada XPB, la cual es una helicasa de ADN dependiente de ATP que funciona en la reparación de la escisión de nucleótidos, como miembro del factor de transcripción TFIIH. Las mutaciones en este gen se asocian con Xeroderma pigmentosum B, síndrome de Cockayne y tricotodistrofia (Weeda *et al*, 1990).

La reparación por escisión de nucleótidos es una vía que se usa para eliminar y reemplazar bases dañadas. La reparación por escisión de nucleótidos detecta y corrige tipos de daño que distorsionan la doble hélice del ADN, causado por ejemplo por la exposición a las sustancias químicas del humo de cigarrillos (Dexheimer, 2013). También se emplea para reparar daños causados por la radiación UV. Esta radiación puede causar que la citosina y la timina reaccionen con bases vecinas y se formen enlaces que distorsionan la doble hélice, llamados dímeros de pirimidinas (Goodsell, 2007).

El proceso de reparación por escisión de nucleótidos se observa en la figura 6 y se lleva a cabo en el siguiente orden: una vez que el UV induce el daño en el ADN, la proteína XPC reconoce la región dañada, sobre la cual se une otra proteína XPA (1). Ésta última es necesaria para reclutar a un gran complejo multiprotéico formado por TFIIH que es un factor de transcripción que incluye además a XPB y XPD (2). Este complejo tiene actividad helicasa, el cual desenrolla el ADN en la vecindad de la lesión y la abre en una gran burbuja, al cual se une RPA, una proteína de unión a cadena sencilla de ADN (3). Posteriormente se unen dos proteínas más XPF y XPG que son endonucleasas que cortan en sus extremos 5' y 3' de la región dañada (4). La región dañada se desprende (5) y finalmente las ADN polimerasas δ y ϵ reemplazan el hueco con los nucleótidos correspondientes y una ADN ligasa sella la brecha en el esqueleto de la cadena (6) (Dexheimer, 2013).

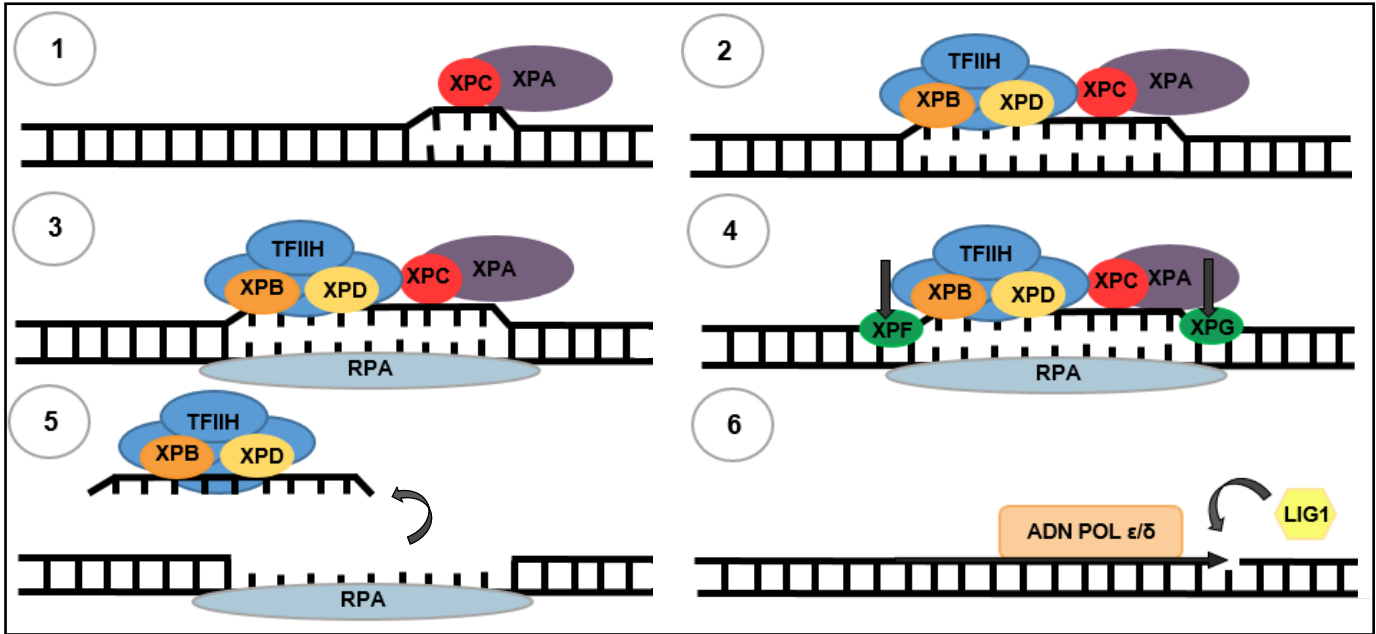


Figura 6. Proceso secuencial de reparación global del ADN por escisión de nucleótidos. Elaboración propia.

FANCB

El gen **FANCB**, codifica para un miembro del grupo de complementación B de anemia de Fanconi (FA). La anemia de Fanconi (FA) es una enfermedad autosómica recesiva caracterizada por anomalías congénitas, hemopoyesis defectuosa y un alto riesgo a desarrollar leucemia mieloide aguda y ciertos tumores (Tischkowitz y Hodgson, 2003). El complejo FA está conformado por las proteínas: *FANCA*, *FANCC*, *FANCG*, *FANCB*, *FANCD2*, *FANCE*, *FANCF*, *FANCT*, *FANCL* y *FANCM*, las cuales están involucradas en un proceso celular conocido como la ruta de Fanconi, implicado en la reparación de lesiones de ADN de doble cadena (Kato *et al*, 2015 y Tischkowitz y Hodgson, 2003).

El reconocimiento de las rupturas de doble cadena por *FANCM* y las proteínas asociadas conduce al reclutamiento del complejo FA y el reclutamiento de nucleasas y polimerasas que se requieren para el proceso de reparación en la que participan los complejos *BRCA1* y *BRCA2* reguladores de la recombinación homóloga (Deans y West, 2011).

En la recombinación homóloga, se utiliza la información del cromosoma homólogo que coincide con la del dañado (o de una cromátida hermana si el ADN se ha copiado) para reparar la fragmentación. En este proceso se acercan los dos cromosomas homólogos y se utiliza la región sin daños del homólogo o la cromátida como molde para sustituir la región dañada del cromosoma roto (Kimball, 2015).

En la figura 7 se ejemplifica el proceso de reparación por recombinación homóloga, el cual se lleva a cabo en el siguiente orden: Primero los extremos del ADN roto son escindidos por nucleasas especializadas para formar extremos 3' protuyentes de una sola hebra (1-2). El siguiente paso es el intercambio de bandas durante el cual uno de los extremos de una sola hebra 3' de la molécula de ADN dañada se introducen en el la plantilla dúplex y busca secuencias homólogas a través del emparejamiento de bases (3). Una vez que se establece un par de bases

estable, la ADN polimerasa sintetiza la hebra dañada utilizando la información proporcionado por la molécula de la plantilla en buen estado, restaurando así el daño ADN (4). Posteriormente, se libera el ADN templado y se termina de sintetizar las hebras dañadas (5) y finalmente, una ligasa sella los extremos recién sintetizados para completar el proceso de reparación (6).

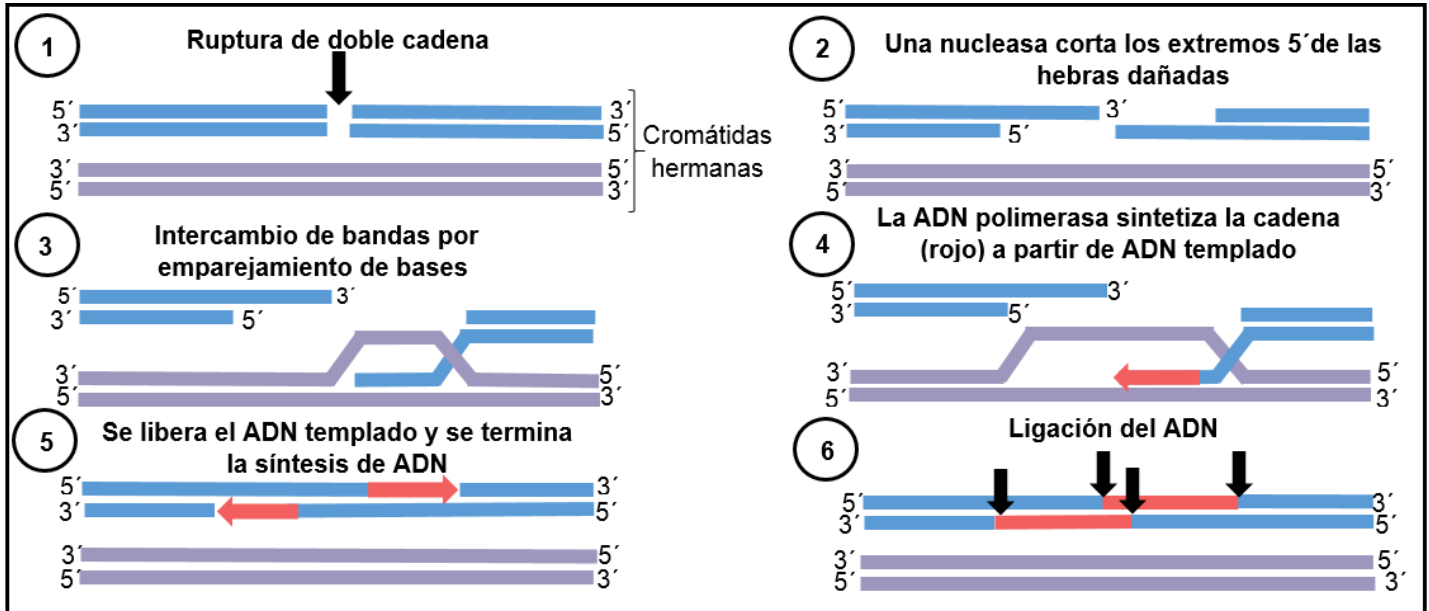


Figura 7. Mecanismo de reparación de ruptura de ADN de doble cadena por recombinación homóloga. Elaboración

MSH6

El gen ***MSH6*** es un gen que codifica para la proteína humana GTBP (G/T binding protein) también llamada p160 o hMSH6, de reparación de errores de ADN por apareamientos erróneos. Defectos en hMSH6 se asocian con el cáncer colorrectal hereditario no polipósico atípico (Dulak, et al, 2013). La reparación de apareamientos erróneos del ADN se ejemplifica en la figura 8 y se lleva a cabo de la siguiente manera: la proteína MSH6 se une con otra proteína llamada MSH2 para formar un complejo proteico. Este complejo identifica ubicaciones en el ADN donde se han cometido errores durante la replicación del ADN. Otro grupo de proteínas, el complejo de proteínas MLH1-PMS2, luego repara los errores (a). En un caso alternativo, se genera la detención del ciclo celular para asegurar que las células no se dividan en condiciones de daño al ADN, permitiendo la reparación de éste o la eventual destrucción de la célula mediante apoptosis, en caso de no ser reparado dicho daño (b).

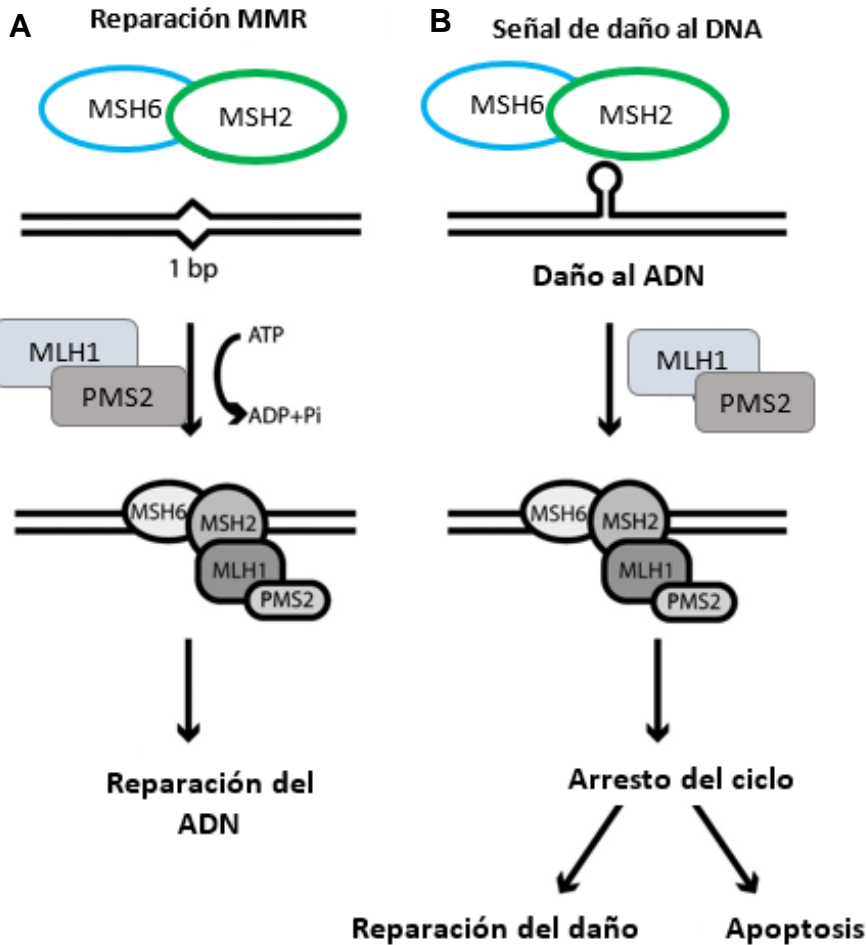


Figura 8. Proceso de reparación de apareamientos erróneos del ADN.

PTEN

El gen ***PTEN*** codifica una proteína fosfatidilinositol-3,4,5-trifosfato 3-fosfatasa, que regula negativamente los niveles intracelulares de fosfatidilinositol-3,4,5-trisfosfato en las células y funciona como un supresor de tumores mediante la regulación negativa de la vía AKT / PKB (Leslie y Downes, 2004).

La vía de señalización de la fosfatidilinositol-3-quinasa (PI3K) se muestra en la figura 9. Esta vía es estimulada fisiológicamente como consecuencia de la activación de receptores de membrana tirosina quinasa, los cuales se autofosforilan y fosforilan el sustrato del receptor de insulina (IRS); éste último, a la vez, fosforila la subunidad p85 de la PI3K. La fosforilación de la subunidad p85 conduce a un cambio conformacional de dicha proteína que conduce a la unión de la subunidad catalítica (p110). La PI3K activa, fosforila el fosfatidil inositol 3,4 difosfato (PIP2) convirtiéndolo en el segundo mensajero fosfatidil inositol 3,4,5 trifosfato (PIP3), el cual, corriente abajo, conduce a la activación de la proteína Akt. La Akt tiene múltiples blancos, como GSK3 (kinasa sintasa de glicógeno) y NFκβ (factor nuclear) responsables de los efectos de la activación normal, que conducen a una respuesta proliferativa y antiapoptótica relacionada con el desarrollo de múltiples

tipos de cáncer (4-5). PTEN codifica la proteína fosfatidilinositol-3,4,5-trifosfato 3-fosfatasa que inactiva la PI3K actuando como regulador negativo de esta vía (Pinzón *et al*, 2009).

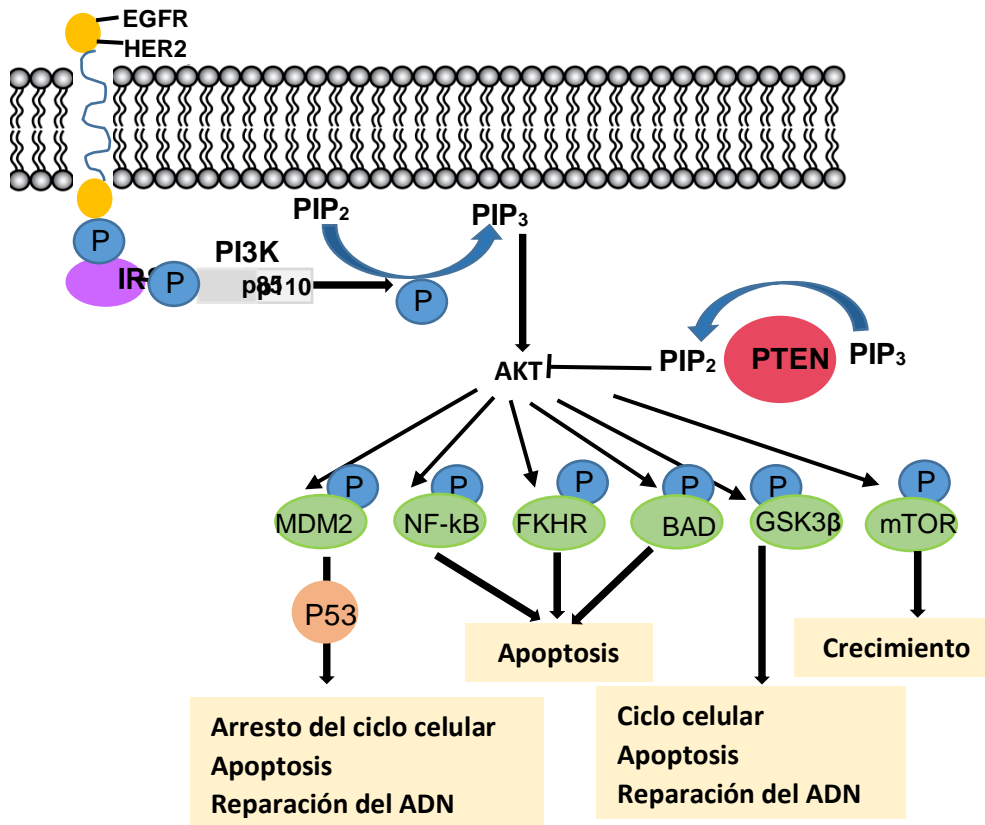


Figura 9. Esquema de la vía de señalización fosfatidil inositol 3 quinasa. Elaboración propia.

RECQL4

La proteína codificada por el gen **RECQL4**, es una helicasa de ADN que pertenece a la familia de la helicasa RecQ. Las helicinas de ADN desenrollan ADN de doble hebra en ADN de una sola hebra y pueden modular la segregación cromosómica. Este gen se expresa predominantemente en el timo y los testículos. Las mutaciones en este gen están asociadas con síndromes de Rothmund-Thomson, RAPADILINO (RA: radial ray defect; PA: patellar aplasia, arched or cleft palate; DI: diarrhea, dislocated joints; LI: little size (short stature), limb malformation; NO: nose slender and normal intelligence) y Baller-Gerold (Siitonen *et al*, 2009).

El síndrome de Rothmund-Thomson (RTS), un raro trastorno genético caracterizado por fragilidad cromosómica, anomalías del desarrollo y predisposición a los cánceres, incluido el sarcoma osteogénico (Kitao *et al* 1999, Vennos y James, 1995). Distintas mutaciones de **RECQL4** también ocurren en el síndrome de RAPADILINO, donde se asocian con malformaciones esqueléticas pero no predisposición al cáncer (Siitonen *et al*, 2003). La homocigosis de alelos mutantes del homólogo murino de **RECQL4** da como resultado una letalidad embrionaria temprana acompañada de una proliferación celular defectuosa (Ichikawa *et al*, 2002), lo que sugiere un papel indispensable en el crecimiento celular normal. Una mutación hipomórfica que elimina un solo exón en el gen murino causa retraso del crecimiento, anomalías del desarrollo e impedimento

de la división celular *in vitro* (Hoki *et al*, 2003), mientras que una deleción del gen que trunca la mitad carboxilo-terminal de la proteína codificada desencadena anomalías centroméricas, aneuploidía y predisposición al cáncer (Mann *et al*, 2005). Por lo tanto, las correlaciones entre el genotipo y el fenotipo en diferentes alelos mutantes de *RECQL4* son complejas e inciertas, y proporcionan poca información sobre las funciones biológicas de la proteína codificada (Manesh *et al*, 2005).

6. JUSTIFICACIÓN

Ya que uno de los principales factores de riesgo para el CM es la susceptibilidad genética, es muy importante definir el impacto de las variantes VUS. En nuestro laboratorio se detectaron 5 variantes cuya mutación afecta potencialmente el sitio donador o aceptor canónico de splicing. Las alteraciones en este proceso, han demostrado ser un importante mecanismo de patogenicidad en la gran mayoría de enfermedades de etiología genética (Baralle *et al*, 2009), por lo que su conocimiento resulta de gran importancia para el entendimiento de los mecanismos moleculares del cáncer. Por lo anterior, en el presente trabajo se realizará la confección de los plásmidos de estas 5 variantes, correspondientes a los genes: *FANCB*, *ERCC3*, *RECQL4*, *PTEN* y *MSH6* potencialmente patogénicas, lo que permitirá realizar estudios funcionales posteriores para evaluar su posible carácter patogénico.

7. HIPÓTESIS

Debido a que en el presente trabajo se realizará la clonación de los exones que tienen las variantes de splicing mencionadas, no cuenta con una hipótesis experimental a probar, ya que se centrará en la confección de los plásmidos, que contengan los exones con mutaciones en los sitios canónicos de splicing de pacientes con cáncer de mama hereditario.

8. OBJETIVOS

8.1. General:

Clonar los exones de 5 variantes de splicing potencialmente patogénicas en el vector de expresión pET01.

8.2. Específicos:

1. Determinar por un análisis *in silico* con ayuda de los programas human splicing finder, NNSplice, NetGene2 y MaxEntScan el posible efecto de la mutación en los sitios donador/aceptor de splicing para cada variante.
2. Diseñar oligonucleótidos dirigidos a cada exón flanqueando la mutación en cuestión, con sus respectivos sitios de restricción, que permitan la amplificación y clonación direccionada dentro del vector pET01.
3. Realizar la clonación direccionada del exón del gen blanco flanqueado por los sitios donador/aceptor que contenga la mutación.
4. Verificar la clonación por análisis de restricción.
5. Verificar la secuencia del inserto por secuenciación tipo Sanger.

9. Material y métodos

Obtención del vector para clonación

El plásmido pET01 del kit Exontrap de MoBiTec se empleó para transformar bacterias de *E. coli* de la cepa DH5 α . Se cultivaron en medio sólido las bacterias transformadas y de éstas, se sembraron dos colonias en medio líquido.

Posteriormente se extrajo el plásmido por el método de lisis alcalina y se purificó por medio de minipreps manuales (Sambrook and Rusell, 2001. Modificado de L. Servín).

Este método explota las diferencias en las propiedades de desnaturalización y renaturalización entre el ADN plasmídico y el ADN cromosómico. Primero se realiza la resuspensión de los pellet, en una solución amortiguadora TE* (50mM Tris-HCl), después la lisis de las bacterias con una solución detergente de Dodecilsulfato de sodio (SDS 1%) y una base fuerte, el hidróxido de sodio (NaOH 0.2N). La alcalinización con NaOH en presencia de un detergente fuertemente aniónico, provoca la lisis celular, la desnaturalización del ADN cromosómico y de las proteínas, y la liberación de los plásmidos. Los plásmidos se ven menos afectados por su pequeño tamaño y estructura superenrollada. En seguida se realiza la neutralización con acetato de sodio (3M pH 4.8), en este paso se renaturaliza el ADN plasmídico. Posteriormente, la neutralización del medio en presencia de una concentración alta de sal, provoca que se cierre covalentemente el ADN plasmídico y la precipitación de las proteínas (por el tratamiento con el detergente y la insolubilidad de la sal potásica del dodecilsulfato) y la del ADN cromosómico (por reasociaciones aleatorias intracatenarias). Después se lleva a cabo la separación de fases con fenol: cloroformo alcohol isoamílico. Los agregados insolubles de proteínas y ADN cromosómico se separan por centrifugación del ADN plasmídico que queda en el sobrenadante y conserva mayoritariamente su estructura nativa. La precipitación con etanol absoluto y acetato de sodio ya presente en la muestra, los lavados con etanol al 70% y resuspensión en agua desionizada (dH₂O) así como un tratamiento adicional con RNAsa A para eliminar este componente de la muestra.

Análisis de restricción

El plásmido pET01 se cortó con dos enzimas de restricción, una de corte único (*Bam*HI-HF) y otra con tres sitios de restricción dentro del plásmido (*Pst*I-HF), con la finalidad de corroborar su identidad. Los mapas de restricción se observaron en gel de agarosa al 0.8% teñido con bromuro de etidio.

9. 1. Análisis del efecto *in silico* de las 5 variantes de splicing

7.1.1. Utilizando los programas *Human splicing finder*, *NNSplice*, *NetGene2* y *MaxEnt* se determinó el posible efecto de la mutación en los sitios donador/aceptor de splicing mutados para cada variante: *ERCC3* con una mutación puntual en el exón 5 correspondiente al sitio donador, *FANCB* con una inserción en el exón 5 que afecta el sitio aceptor, *MSH6* con una deleción que afecta el sitio aceptor correspondiente al exón 9, *RECQL4* con una mutación puntual en el sitio donador que afecta al exón 17 y *PTEN* con una mutación puntual en el exón 6 que afecta el sitio aceptor.

9.2. Diseño de oligonucleótidos

9.2.1. Por medio de la base de datos de Ensembl Genome Browser GRCh37 se obtuvo la secuencia de los genes de interés, el exón con la mutación en el sitio canónico de splicing fue seleccionado y con el programa Primer3 se procedió a diseñar los oligonucleótidos dirigidos hacia cada una de las secuencias exónicas que queden flanqueadas por los sitios

donador/aceptor mutado con el programa considerando como valores óptimos los proporcionados por el programa. Posteriormente, con el programa PrimerBlast, se corroboró el reconocimiento con la secuencia deseada para cada gen en cuestión. Por último, se añadieron además las secuencias de reconocimiento de enzimas de restricción *XhoI* y *BamHI* en el extremo 5' de cada oligonucleótido.

Tabla 2. Características de oligonucleótidos, sitios de restricción añadidos y valores óptimos proporcionados por el programa Primer3.

Diana	Sitio de restricción	Tm °C	longitud	% G/C
<i>XhoI</i>	<u>CTCGAG</u>	55-65°C	18-25 pb	40-60%
<i>BamHI</i>	<u>GGATCC</u>			

9.2.2. Los sitios de restricción se verificaron por medio de un mapa de restricción, analizando los patrones de corte con el programa NEBcutter V2.0 de New England Biolabs y Ape 2.0.49 (Fig.9).

9.2.3. Las enzimas de restricción compatibles con el vector de clonación pET01, se seleccionaron con ayuda del programa Vector NTI® Express Designer y se realizaron las construcciones del plásmido recombinante. Se realizó la construcción para cada exón de los 5 genes en estudio, para los que fueron compatibles las mismas enzimas *XhoI* y *BamHI* (Fig. 10).

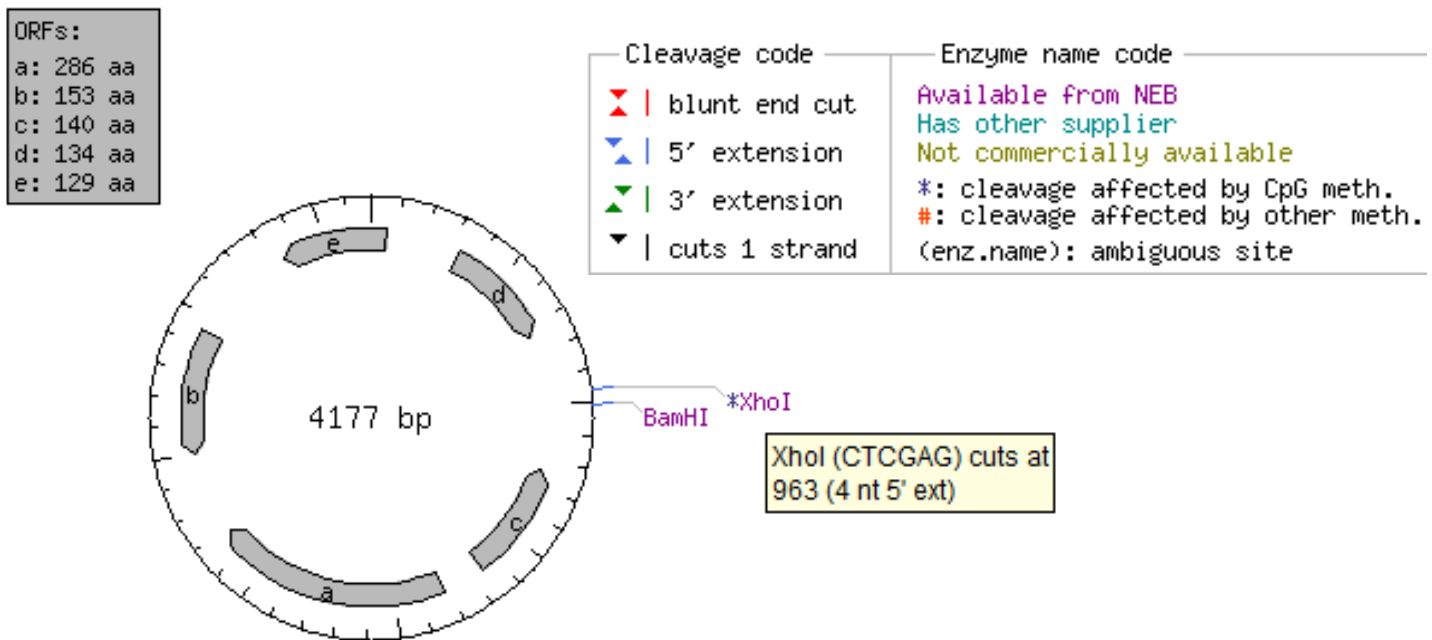


Figura 10. Mapa de restricción del plásmido con las enzimas *XhoI* y *BamHI* obtenido con el programa NEBcutter V2.0 de New England Biolabs. Se muestran las enzimas de restricción compatibles con el vector pET01 y sus sitios de restricción.

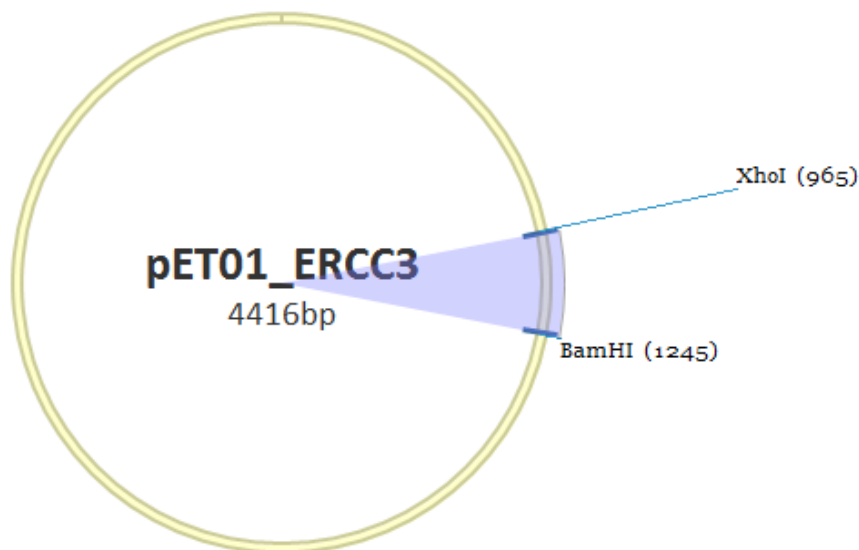


Figura 11. Construcción de plásmido recombinante para el exón 5 de *ERCC3* con ayuda del programa Vector NTI® Express Designer. Las enzimas de restricción compatibles fueron las mismas para todos los insertos (*XhoI* y *BamHI*).

9.3. Amplificación de los 5 exones de interés por PCR

Para realizar la amplificación de los insertos (conformados por el exón flanqueado por la mutación con regiones de sus intrones adyacentes) se empleó ADN genómico extraído de sangre de los pacientes con la mutación, así como ADN silvestre anteriormente verificado por Sanger. La única excepción fue el gen *ERCC3* mutante, del cual no se contó con suficiente ADN genómico o sangre de la paciente por lo que se amplificó a partir de ADN de la biblioteca del trabajo de Quezada-Urban *et al* (2018) previamente publicado del que deriva este proyecto. La enzima empleada para la amplificación por PCR fue 2X KAPA Taq ReadyMix de KAPA BIOSYSTEMS, empleando 1.5 mM de MgCl₂ y 0.2 mM de cada dNTP.

El exón de interés para cada gen fue amplificado por PCR en una reacción estándar de punto final en el equipo T100™ Thermal Cycler de BIO-RAD. Previamente se estandarizaron las temperaturas de hibridación para cada par de oligonucleótidos de cada variante. Las condiciones de reacción fueron: 20 ng de ADN genómico mutante y silvestre de cada variante, 0.2 μM de primers y 1x concentración final de 2XKAPA Ready Mix (1.5mM MgCl₂ a 1x).

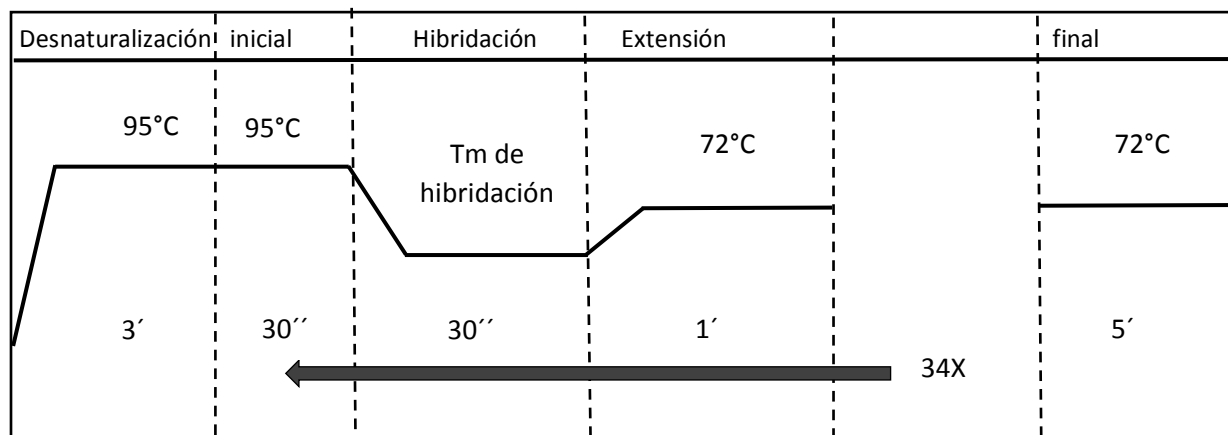


Figura 12. PCR de punto final empleada en la amplificación de los exones silvestres y mutantes. Se representan las temperaturas para cada paso y el tiempo de duración de éstas.

9.3.1. Purificación

Se realizó la purificación de los productos obtenidos por PCR, mediante columna de afinidad con el kit NucleoSpin® Gel and PCR Clean-up de Macherey-Nagel, siguiendo el protocolo recomendado.

9.4. Clonación direccionada en el vector de expresión pET01

Se empleó el vector plásmidico pET01 del kit Exontrap de MoBiTec, el cual tiene características que lo hacen adecuado para este trabajo, como la presencia de dos exones propios separados por un intrón de 600 pb con un sitio de clonación múltiple, con sitios de restricción para enzimas (MCS) un promotor que permite su expresión en células humanas (RS Virus LTR), un origen de replicación procarionte (ColE1), un origen de replicación eucarionte (SV40 origin) y un gen que codifica para β -lactamasa, el cual le confiere resistencia a la ampicilina y permite la selección de las colonias recombinantes (Fig. 13).

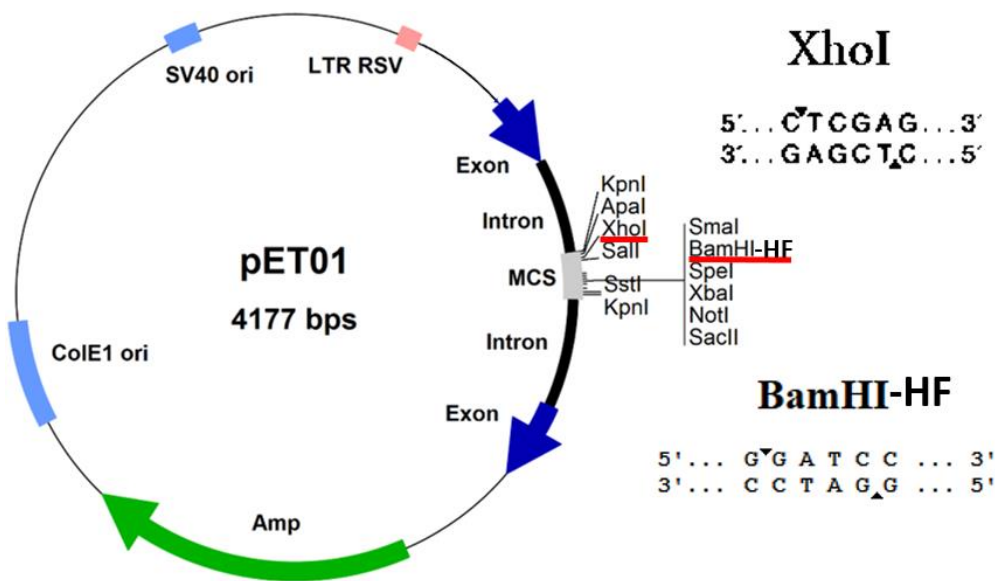


Figura 13. Mapa del vector pET01. Se muestran los sitios de corte y reconocimiento de cada enzima seleccionada para la clonación direccionada. Imagen tomada y modificada del kit Exontrap de MoBiTec.

9.4.1. Tanto el vector como los productos del PCR fueron cortados con las enzimas *XhoI* y *BamHI-HF* seleccionadas.

9.4.2. Los productos de la doble digestión de los insertos fueron repurificados por medio del kit PCR Clean-up de Macherey-Nagel.

9.4.3. El vector fue desfosforilado con Fosfatasa alcalina FastAP de Thermo Scientific.

9.4.4. El vector obtenido de la doble digestión y desfosforilación, fue purificado a partir de gel de agarosa de bajo punto de fusión (LMT) con el protocolo de Fenol cloroformo alcohol isoamílico.

9.4.5. Los productos de PCR fueron clonados dentro del vector de clonación pET01. Se realizó la ligación direccional con las proporciones 3:1 del exón en 200-500 ng del vector pET01 con Ligasa T4 de Thermo scientific siguiendo el protocolo recomendado para ligación de extremos cohesivos (Sticky-end Ligation). Se realizaron dos controles: un control negativo con el plásmido

digerido con ambas enzimas, desfosforilado y ligado en ausencia del inserto, y un control positivo con plásmido cortado con una sola enzima y purificado.

9.4.6. El producto de la ligación fue utilizado para la transformación de bacterias de la cepa DH5 α de *E. coli*. Se utilizaron 3 uL del producto de la ligación anterior en 50 uL de bacterias calcio competentes se dejaron 15 minutos en hielo y posteriormente se sometieron a shock térmico a 42 °C por 40 segundos, se les añadió 200 uL de medio LB y se dejaron recuperar a 37°C por 30 minutos. Posterior a esto se sembró en placas de cultivo el contenido total de los viales en medio LB sólido con ampicilina (100 ug/mL) para seleccionar las colonias recombinantes y se dejó crecer aproximadamente 22 horas.

9.4.7. El vector fue amplificado por medio del cultivo en medio líquido de las respectivas colonias bacterianas seleccionadas, en presencia de ampicilina.

9.5. Tamizaje de colonias

9.5.1. Se realizó una PCR en colonia a partir del medio líquido obtenido por cultivo, con los oligonucleótidos 2'Forward y 3'Reverse del kit Exontrap de MoBiTec, los cuales flanquean el sitio polilinker donde es clonado el exón y tienen la siguiente secuencia: 5'-GAGGGATCCGCTTCCTGGCCC-3' y 5'-CTCCCGGGCCACCTCCAGTGCC-3'. De estar ausente el inserto, al igual que el control positivo (plásmido sin inserto) genera un fragmento de 839 pb (Figura 13), mientras que de estar presente el inserto, para cada variante se genera un fragmento de: 1,096 pb para la variante *ERCC3*, 1,163 pb para la variante *FANCB*; 1,129 pb para la variante *MSH6*; 1,376 pb para la variante *PTEN* y 1,070pb para la variante *RECQL4*.

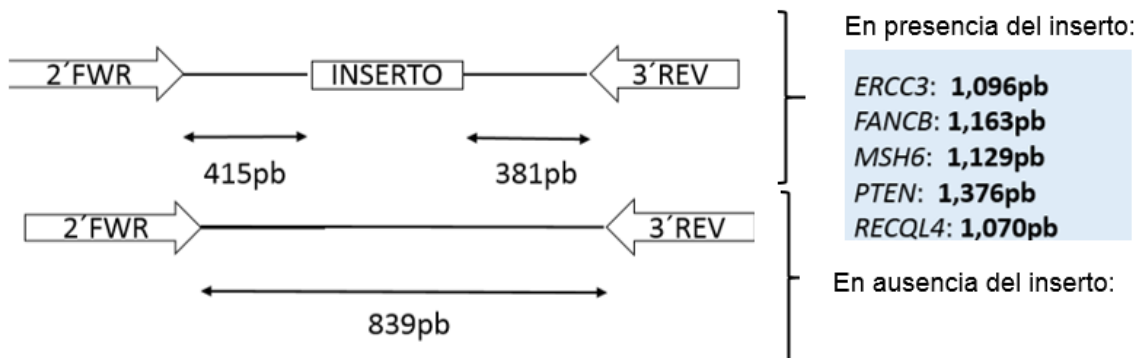


Figura 14. Representación de los sitios de unión de oligonucleótidos 2Forward y 3Reverse del kit MoBiTec, empleados para la PCR en colonia así como de los fragmentos que genera en ausencia y presencia del inserto correspondiente.

Se realizó la extracción del plásmido por el método de lisis alcalina del protocolo de Sambrook and Rusell, 2001, de las muestras que resultaron positivas para el PCR en colonia.

9.5.2 Análisis de restricción

Las construcciones purificadas, fueron cuantificadas con el NanoPhotometer® NP80 y se digirió ~2 ug de cada una con las enzimas *Bam*HI-HF y *Xho*I para liberar el fragmento clonado, adicional a esto, se seleccionaron enzimas de restricción con cortes específicos dentro de cada exón clonado, para *FANCB* con la enzima *Mfe*I-HF con un sitio de restricción dentro del fragmento *FANCB* que sólo linealiza la construcción en presencia de este inserto, y las construcciones *MSH6*

fueron digeridas adicionalmente con la enzima *NsiI* la cual, además de un sitio de restricción en el vector, posee un sitio de restricción dentro del inserto *MSH6*, generando dos fragmentos; uno de 3,252 pb y otro de 1,200 pb.

9.6. Secuenciación Sanger

9.6.1. Se realizó la secuenciación tipo Sanger para corroborar la presencia o la ausencia de la mutación para cada caso wt y mt de cada variante, además de asegurar que en la secuencia de interés no se haya introducido ningún cambio nucleotídico extra. Las construcciones fueron enviadas a Macrogen Corea para su secuenciación Sanger.

10. Resultados

Obtención del vector para clonación

Análisis de restricción de pET01 con las enzimas *Bam*HI-HF y *Pst*I-HF.

Después de la transformación bacteriana con el plásmido pET01 del kit MoBiTec, se extrajo el ADN plasmídico de dos cultivos bacterianos por el método de minipreps manual y se realizó un análisis de restricción a las muestras con dos enzimas: *Bam*HI-HF de corte único dentro del vector, para linealizarlo y una de tres cortes *Pst*I-HF. Las digestiones se observaron en gel de agarosa y cómo podemos observar para ambas muestras con *Bam*HI-HF, el plásmido linealizado correspondiente a 4177pb y con *Pst*I el patrón de bandeado esperado de 2271 pb, 1369 pb y 537 pb, confirmaron la identidad del plásmido de interés, para poder continuar con el procedimiento (figura 15).

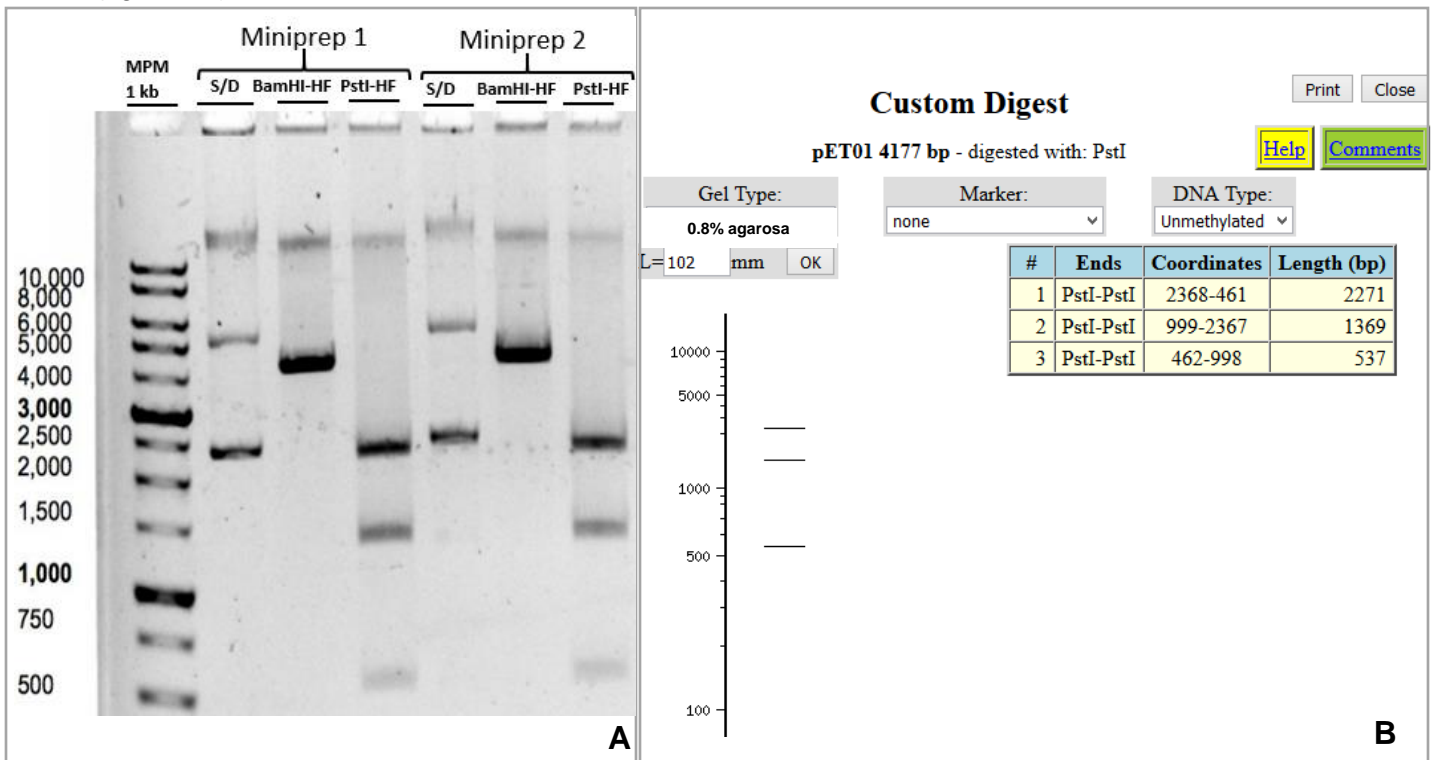


Figura 15. Linealización del plásmido con *Bam*HI-HF. Se observan las bandas de 4177 pb correspondientes a la longitud de pET01. Digestión con la enzima *Pst*I-HF: Resultado obtenido (A) y resultado esperado (B), obtenido con el programa NEBcutter de NEB.

10.1. Predicción del efecto *in silico* para cada variante

Se realizó una análisis *in silico* de las variantes de splicing por medio de 4 programas bioinformáticos: human splicing finder, NNSplice, NetGene2, MaxEnt Scan. El programa human splicing finder, predijo un efecto en el corte y empalme para las variantes en los genes *ERCC3* con una mutación puntual en el exón 5 correspondiente al sitio donador y *PTEN* con una mutación puntual en el exón 6 que afecta el sitio aceptor. Por otro lado, los programas NNSplice y NetGene2, no reconocen los sitios canónicos mutados de las variantes de estos mismos genes. Y por último, el programa MaxEnt predijo para las variantes *ERCC3*, *PTEN* y *RECQL4* (con una mutación puntual en el sitio donador que afecta al exón 17), una disminución de la fuerza de reconocimiento de los sitios canónicos de splicing (score) que podría afectar el corte y empalme.

Para este análisis, ninguno de los programas utilizados predijo un efecto de las mutaciones en las variantes *FANCB* (con una inserción en el exón 5 que afecta el sitio aceptor) ni *MSH6* (con una delección que afecta el sitio aceptor correspondiente al exón 9).

Tabla 3. Resultado del análisis *in silico* de las mutaciones en los sitios aceptor/donador para cada variante. Secuencia (exones en mayúsculas e intrones en minúsculas, mutación en rojo).

Gen	Programa	Efecto	Valores wt	Valores mt
<i>ERCC3</i> (Exón 5 >chr2:128047264- 128047264) (reverse complement)	HSF (score)	si	73.67	46.84
	NNSplice		NA	NA
	Netgene2 (confidence)		-----	-----
	MaxEnt		6.35	-1.83
SECUENCIA			ATCTGCCg ^t atgtggactcctgg	ATCTGCCa ^t atgtggactcctgg
<i>FANCB</i> (Exón5 >chrX: 14861029- 14891684).	HSF (score)	no	NA	NA
	NNSplice		0.89	0.89
	Netgene2(confidence)		0.76	0.76
	MaxEnt		-11.13	-11.13
SECUENCIA			tatttatt----agAGTGAACCA	tatttatt ^t atagAGTGAACCA
<i>MSH6</i> (Exón 9 >chr2:48009721- 48034592).	HSF (score)	no	NA	NA
	NNSplice		0.98	0.98
	Netgene2 (confidence)		-----	-----
	MaxEnt		6.25	6.25
SECUENCIA			TTATTTCGg ^t aactaactaac	TTATTTCGg----taactaac
<i>PTEN</i> (Exón 6 >chr10:89623145- 89728582) (reverse complement)	HSF (score)	si	90.29	61.34
	NNSplice		0.91	-----
	Netgene2 (confidence)		0.0	-----
	MaxEnt		-16.46	-23.79
SECUENCIA			tttctgtccacca ^g GGAGTAA	tttctgtccacca ^t GGAGTAA
<i>RECQL4</i> (Exón 17 >chr8:145736617- 145743260 (reverse complement)	HSF (score)	no	97.83	70.99
	NNSplice		1.0	-----
	Netgene2 (confidence)		0.0	-----
	MaxEnt		9.88	1.38
SECUENCIA			CCCACAG ^g taagcagccc	CCCACAG ^t taagcagccc

10.2. Diseño de oligonucleótidos para clonación

Se diseñaron los oligonucleótidos para cada variante tomando los valores óptimos proporcionados por el programa Primer3: un 40-60% de G/C, una Tm entre 55-65 °C, una longitud de 18-25 pb, una autocomplementariedad ≤ 8 y una autocomplementariedad 3' ≤ 3 . Se corroboró su correspondencia con el gen para cada caso con el programa Primer Blast y se añadieron además las secuencias de reconocimiento de enzimas de restricción *XhoI* y *BamHI* en el extremo 5' de cada oligo como se muestra en la siguiente tabla. Además de esto, se añadieron 4-5 pb adicionales en el extremo 5' de cada oligo, asegurando así que las enzimas de restricción (que son endonucleasas) puedan cortar en el sitio de reconocimiento; *BamHI* requiere al menos 3 pb en el extremo del sitio de restricción y *XhoI* requiere al menos 4 pb. (Véase tabla 3).

Por otro lado, los sitios probables para la selección de los oligonucleótidos eran de alta cantidad de SNPs, estas regiones son altamente variables, por lo que se hizo lo posible por evitarlas (Véase tabla 4).

Tabla 4. Oligonucleótidos diseñados para la amplificación de los exones de interés.

Gen	Diana	Oligonucleótidos	Tm °C	longitud	%G/C	Auto-complementariedad		Tamaño del producto
							3'	
ERCC3	<i>XhoI</i>	AGACCCTCGAGTTGAATGAGCAGTG	64	25	52	6	3	300 pb
	<i>BamHI</i>	TGCTGGATCCAAGACACCTGAGAG	63	24	54	8	1	
FANCB	<i>XhoI</i>	CACACTCGAGAAACTGTTCTGAC	59	23	47	6	3	367 pb
	<i>BamHI</i>	ATTTGGATCCAGGTGCCAAAAG	60	23	43	8	3	
MSH6	<i>XhoI</i>	GCACTTCTCGAGCTAGCACATGTATC	63	26	50	6	2	333 pb
	<i>BamHI</i>	TTCATGGATCCTCATCCCTTCCCCTT	65	26	50	6	0	
PTEN	<i>XhoI</i>	GCTACGCTCGAGTTACCATAGCAATTT	63	27	44	6	3	580 pb
	<i>BamHI</i>	CAAATGGATCCACTCCTCTGGTGGAC	64	26	53	6	2	
RECQL4	<i>XhoI</i>	GGAACCTCGAGTAAGCCACAGGGGT	67	25	60	6	1	274 pb
	<i>BamHI</i>	GAAAAGGATCCATGTGGCCAACAGC	64	25	52	6	2	

Exón
Intrón

SNIPS EN EXONES
Snips en intrones

Oligonucleótidos
Mutación

Gen	Sitio de hibridación
<i>ERCC3</i>	tttcagaccatctaattgaatgagcagtggaatctccaggaggtacctgctggtcagtagaccctcttcctcccctgcttcag ATACTTCGTTGAAAGTTGCCACCCTGATGTAATCCAGCATCTTCTCCAGGACCCCGTGATCCGAGAA TGCCGCTTAAGAACTCTGAAGGGGAGGCCACTGAGCTCATCACAGAGACTTTCACAAGCAAATCT GCCgtatgtggactcctgggccaccCttgggtggggcaggcatttaggatttcagttcctcaggtgtcttaaccagcact
<i>FANCB</i>	tttatacactagggaaactgttctgactggatatactttaattgattttgttagtttgaattttcattgttgcaaaattcagt cttagtattttttttatttatttttagAGTGAACCATCAGATTGCATGAAGATGACTTATTTGAAGACAAAC AAGAGAATCGTACCTGGTGGTCCACCTCTAGAAACAGGACTAAAgttagtgtctgtttgcttTttgctttt gattagtgttattctaaatgaAggaaagaacatTgaAgaaggtccttaaagtataagcaactgggtgcttgggtaaaatgg tatttCaaaggatctttatcttttggca.ctgtacaaat
<i>MSH6</i>	AGggcactctctgtagcatgtatcgctaataatcttctttaagGCATGCATGGTAGAAAATGAATGTGAAG ACCCAGCCAGGAGACTATTACGTTCTCTATAAATCATTAAAGGAGCTTGCCTAAAAGCTATG GCTTTAATGCAGCAAGGCTTGCTAATCTCCAGAGGAAGTTATTCAAAGGGACATAGAAAAGCA AGAGAATTTGAGAAGATGAATCAGTCACTACGATTATTTGgtaactaactataatggaattataactaa ctgaccttaagttcaagaacagtaaaagggaagggtgatgactatgaaaaaac
<i>PTEN</i>	tggtacgaccagttaccatagcaatttagtgaataactataatggaacattttttcaatttgcttctctttTttctgtccac caGGAGTAACTATTCCAGTCAGAGGCGCTATGTGTATTATTATAGCTACCTGTTAAAGAATCATC TGGATTATAGACCAGTGGCACTGTTGTTTACAAGATGATGTTTGAAGACTATTCCAATGTTCACTG GCGGAACCTGCAgtaagtgttgaaattctcatccttccatgtattggAacagttttctaacatattctagaagttacataa aaatttagaaagaaattaccacattgaaatttatgaggagactatatttctgaagcatttGaacaataattagctttgtgtt caactcattgggctaaagaagccaaaagcaatgggttttaattgtagtgaagccaaattatattatgaagaaAtattctgtgtt ataaccaccaataacagcccaattctgactagatgatggaagaacctgtcccatcagaggtccagcatgaggtccagcagaggt ccaccagaggagttcagcaattgct
<i>RECQL4</i>	AggaacctgggtaagccacaggggtgtggagggctgtcccgcgtcgctgagccctgctctgccccagCCATCGAGA CTTTGCTGTGCTACCTGGAGCTGCACCCACCACTGGCTGGAGCTGCTGGCGACCCTATACCC ATTGCCGTCTGAACTGCCCTGGGGCCCTGCCAGCTCCAGGCCCTGGCCACAGgtaagcagccctg ccaagtggagacgaggttgagaatcagggctgtggccacatgctccttttc

Figura 16. Sitios de hibridación de los oligonucleótidos, flanqueando el sitio aceptor/donador mutado.

Estandarización de las temperaturas de hibridación

Previo a la amplificación, se llevó a cabo una estandarización de las temperaturas óptimas de hibridación para cada uno de los pares de oligonucleótidos, se tomaron en cuenta las temperaturas de los certificados de análisis y se redujo 5 °C, comenzando para todos los oligonucleótidos con una temperatura de 55°C y modificando la temperatura a 52 y 54°C para *FANCB* y a 60 y 62 °C para *RECQL4* para obtener bandas más específicas y claras. Todos los exones se amplificaron de manera específica, obteniendo únicamente las regiones de interés correspondientes al tamaño esperado (figura 17).

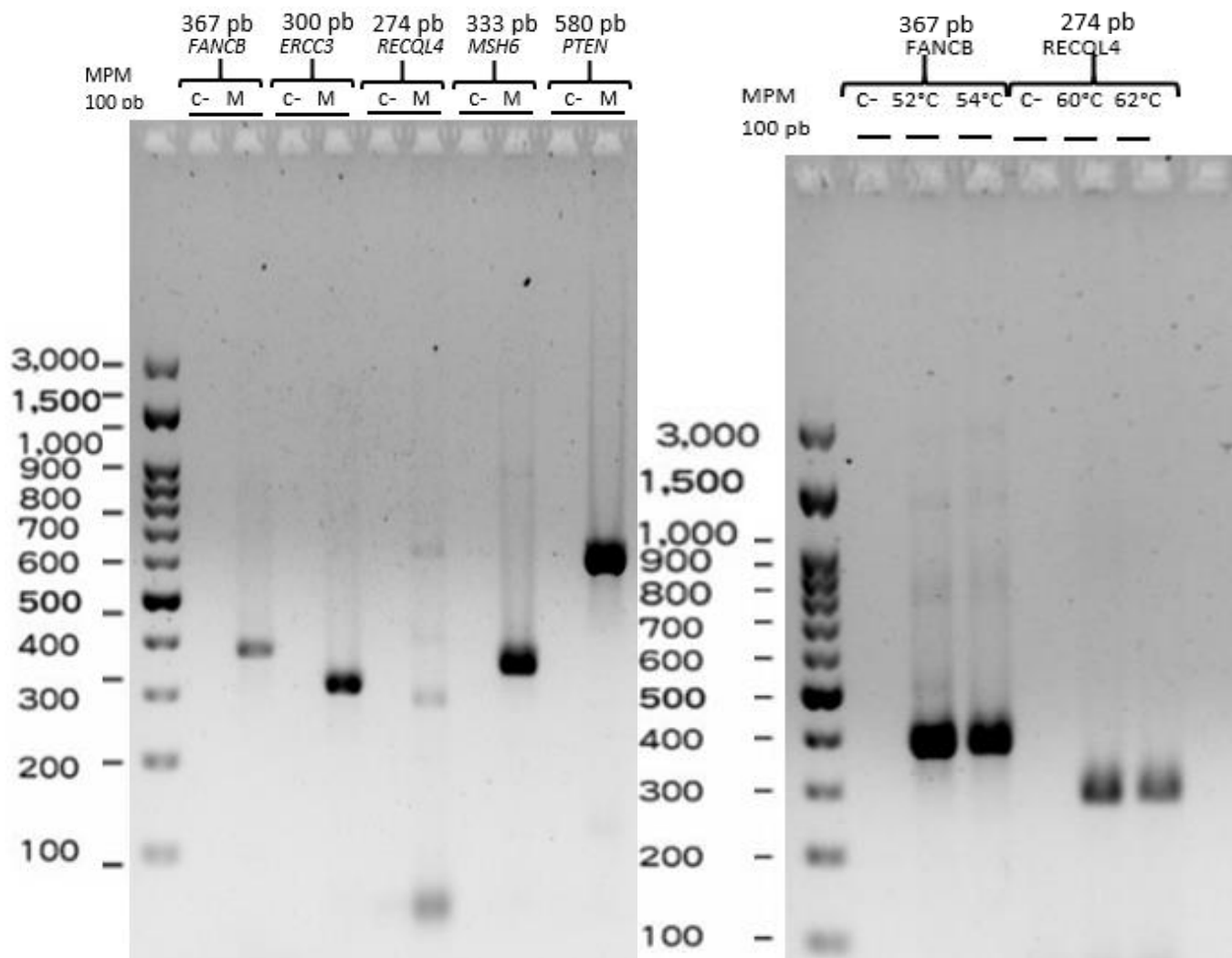


Figura 17. Estandarización temperaturas de hibridación de los oligonucleótidos para cada variante de splicing. Para *ERCC3* *MSH6* y *PTEN* se determinó como T_m óptima la de 55 °C. Se mejoraron las condiciones de hibridación al cambiar la T_m para *FANCB* (52 Y 54°C) y *RECQL4* (60 y 62°C), se determinó como óptima a 52°C y 60°C respectivamente. Geles de agarosa al 1%.

10.3. Amplificación de los exones por PCR de punto final

Para la amplificación de los insertos, se empleó como ADN templado, ADN genómico wt y mt para cada variante. Lá única excepción fue gen *ERCC3* mutante, para el cual se amplificó a partir de ADN de la biblioteca del trabajo de Quezada-Urban *et al* (2018) previamente publicado, del que deriva este proyecto.

En la figura 18 se puede observar que lo productos de PCR obtenidos, corresponden al tamaño esperado del exón amplificado para cada variante y se observan bandas únicas, que indican la especificidad de los mismos. Las bandas para las variantes wt se muestran más claras, esto pudo deberse a que la concentración no era la adecuada y esto se resolvió duplicando la cantidad de ADN templado para la muestra.

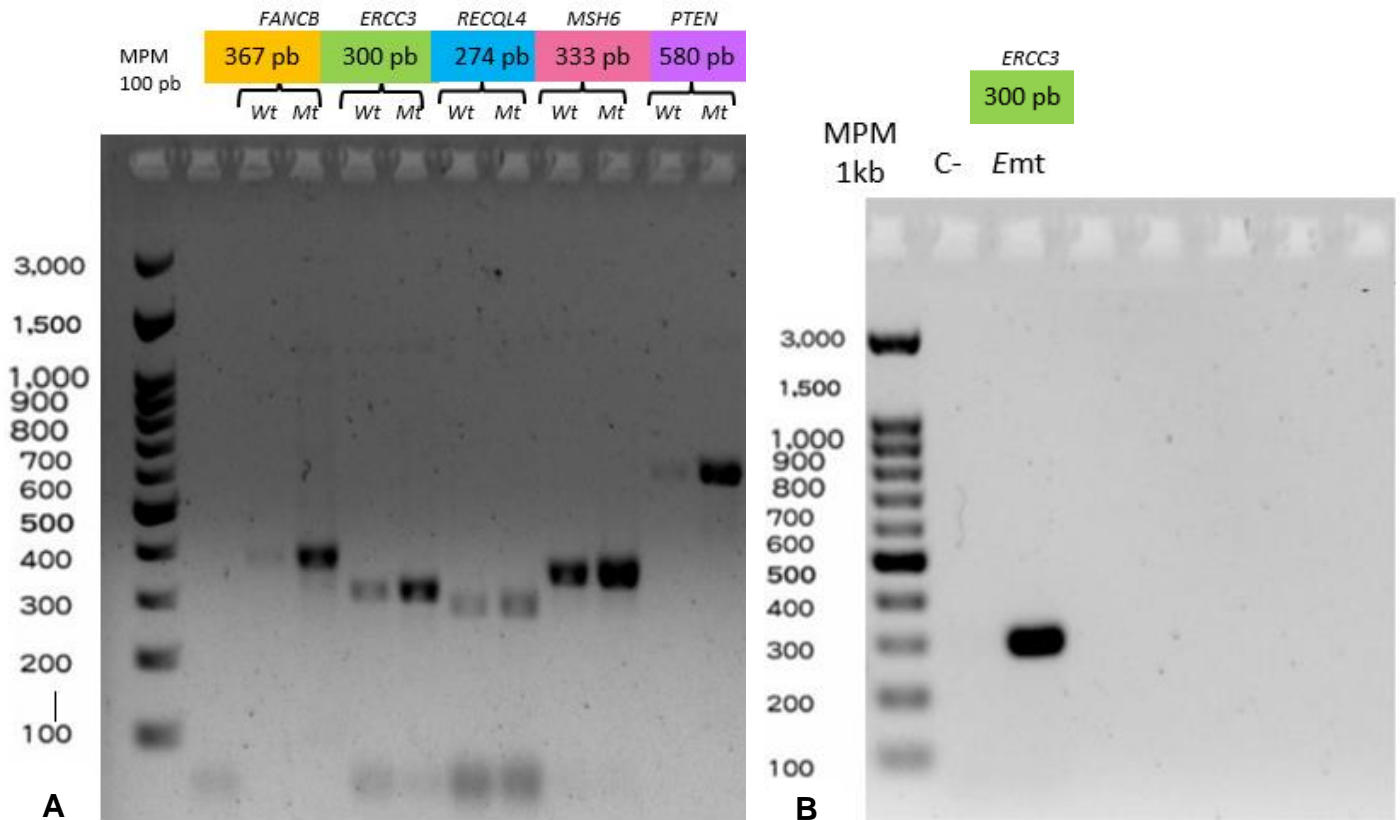


Figura 18. Exones amplificados por PCR de punto final, a partir de ADN genómico wt y mt (A) y a partir de ADN de biblioteca (B).

10.4. Clonación direccionada en el vector de expresión pET01

Tanto los productos de la amplificación como el vector, fueron digeridos con ambas enzimas seleccionadas (*Bam*HI-HF y *Xho*I), por poseer un solo sitio de corte dentro del vector y ningún sitio de corte dentro de ninguno de los insertos a clonar, estas además presentan cortes cohesivos los cuales permiten una clonación direccionada del inserto en el vector, es decir, que el inserto sólo pueda ser clonado en una sola dirección dada por la incompatibilidad de los cortes de ambas enzimas.

10.5. Tamizaje de colonias

Los productos de la digestión fueron purificados y ligados en una proporción 3:1 con el vector pET01 de MoBiTec. Con los productos de ligación, se transformaron bacterias de la cepa DH5 α de *E. coli* y se seleccionaron las bacterias recombinantes en medio LB sólido con ampicilina en una concentración de 100 μ g/mL. Las bacterias se incubaron por 22 horas y se obtuvieron los siguientes resultados.

Tabla 5. Conteo de colonias obtenidas de cada transformación y cultivo en medio selectivo con ampicilina para cada variante *wt* y *mt*.

Placa	# Colonias	Expansión clonal
C-	0-3	
C+	126-200	
<i>ERCC3 wt</i>	7	5
<i>ERCC3 mt</i>	11	5
<i>FANCB wt</i>	25	5
<i>FANCB mt</i>	11	5
<i>MSH6 wt</i>	25	5
<i>MSH6 mt</i>	28	5
<i>PTEN wt</i>	28	10
<i>PTEN mt</i>	25	10
<i>RECQL4 wt</i>	3	3
<i>RECQL4 mt</i>	15	5

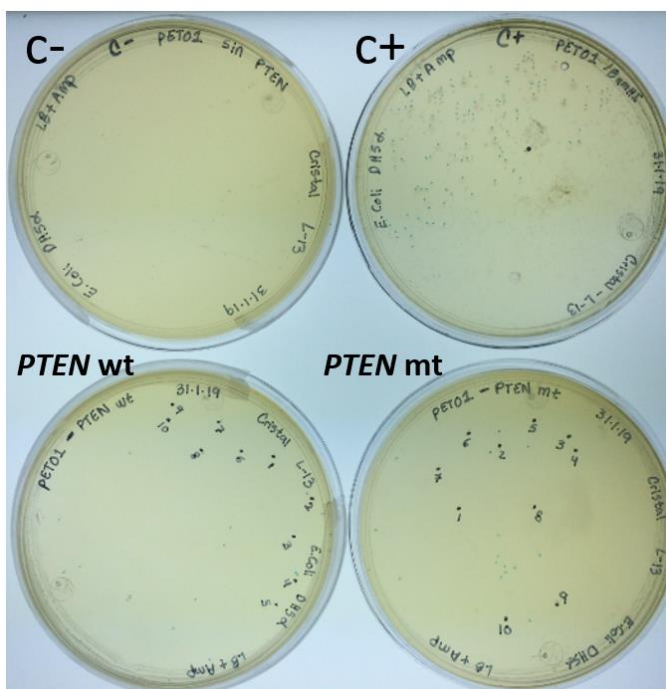


Figura 19. Placas de la variante *PTEN* con sus respectivos controles positivo y negativo y sus colonias bacterianas a 22 horas de cultivo, contadas y numeradas.

Del total de colonias obtenidas tras 22 horas de incubación, se cultivaron en medio líquido selectivo con ampicilina un número considerable, tomando en cuenta que las mutaciones descritas son heterocigas, la probabilidad de encontrar la variante *mt* en las colonias analizadas es del 50%, por lo cual, se trató de analizar un mayor número de las clonaciones mutantes, esperando así obtener al final del análisis al menos una de estas, para cada variante.

PCR en colonia

A partir del medio de cultivo líquido, se realizó una PCR en colonia. Este es un método muy práctico para analizar la presencia o ausencia de un inserto clonado dentro de un plásmido. Es un método que utiliza bacterias de cultivo vivas en lugar de ADN templado en una reacción normal de PCR. En este análisis pudimos descartar las clonaciones vacías antes de realizar las extracciones plasmídicas para análisis posteriores, como se observa en las figuras 20-23, es evidente la diferencia de los productos obtenidos cuando se encuentra presente el inserto, que corresponde a una banda de mayor tamaño que va de 1,070 pb en el caso del inserto más pequeño (*RECQL4*) hasta 1,376 pb para el inserto de mayor tamaño (*PTEN*) a diferencia del plásmido vacío (control positivo) correspondiente a 839 pb notablemente debajo de éstas.

En el caso del PCR en colonia para *FANCB* se observa una banda adicional que parece coincidir con el producto producido en el control positivo correspondiente al plásmido vacío (839 pb). Lo anterior podría deberse a la presencia de contaminación en las muestras con plásmido vacío. Sin embargo la ausencia de estas bandas en los productos de las otras variantes, contravienen esta hipótesis, ya que fueron preparados con el mismo "Master Mix" y en las mismas condiciones. Otra posibilidad es que esta banda adicional se deba a que los oligonucleótidos empleados para esta

PCR hibriden con alguna otra secuencia dentro del genoma bacteriano y que esta otra banda inespecífica coincida con el tamaño esperado para el plásmido vacío.

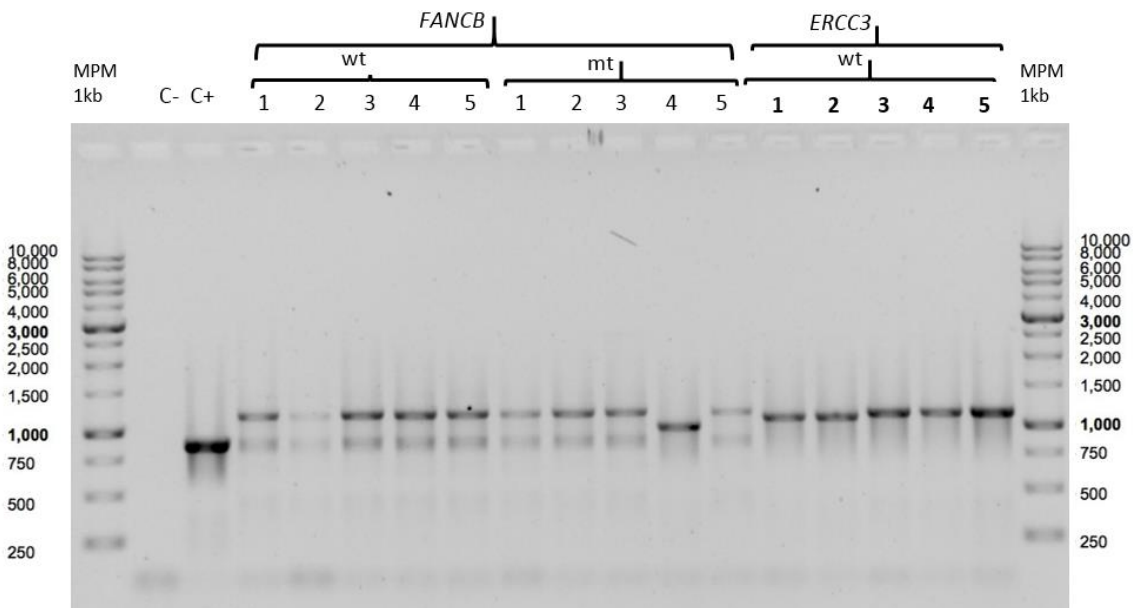


Figura 20. Resultado de la PCR en colonia para *FANCB* wt y mt y *ERCC3* wt. 5/5 colonias analizadas wt y 4/5 colonias analizadas mt para *FANCB* y 5/5 colonias analizadas *ERCC3* wt resultaron positivas para este análisis, con un tamaño de banda de 1163 y 1096 pb respectivamente. Gel de agarosa al 1%.

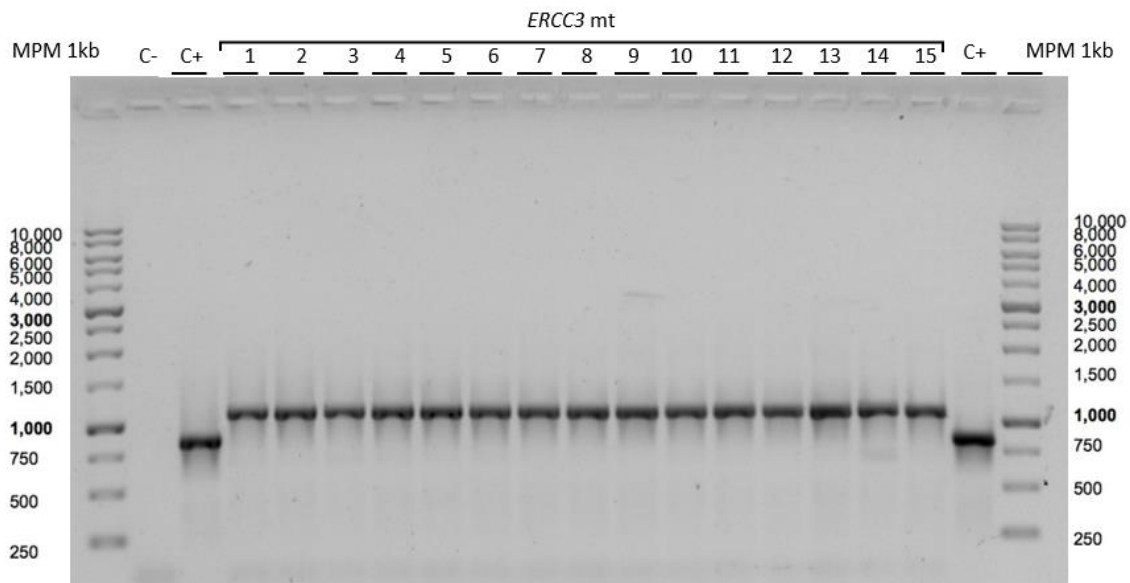


Figura 21. Resultado de la PCR en colonia para *ERCC3* mt. 15/15 colonias analizadas resultaron positivas para este análisis, con un tamaño de banda de 1096 pb. Gel de agarosa al 1%.

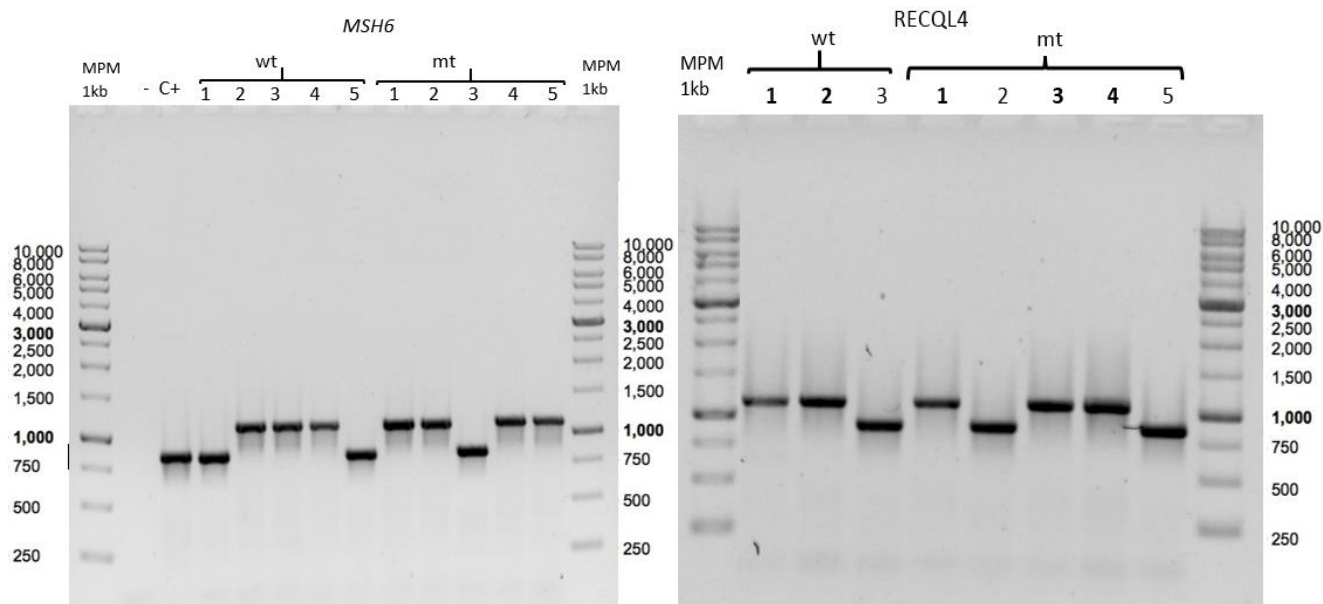


Figura 22. Resultado de la PCR en colonia para *MSH6* y *RECQL4* wt y mt. 3/5 colonias analizadas wt y 4/5 colonias analizadas mt de *MSH6* y 2/3 colonias analizadas wt y 3/5, analizadas mt para *RECQL4* resultaron positivas para este análisis, con un tamaño de banda de 1129 y 1070 pb respectivamente. Gel de agarosa al 1%.

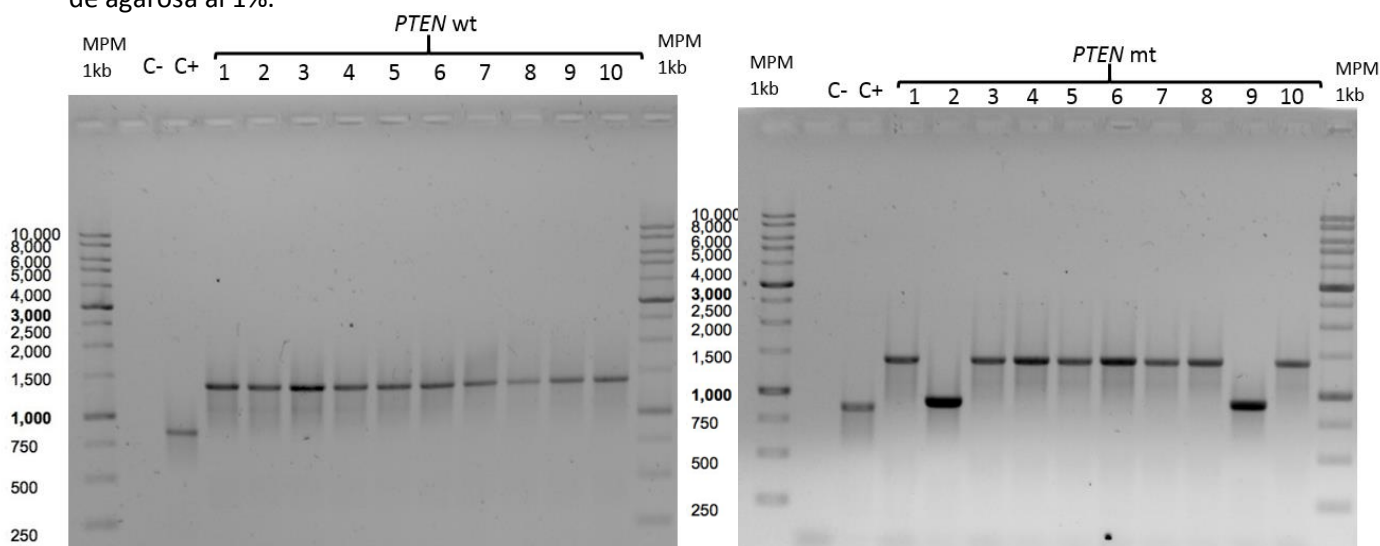


Figura 23. Resultado de la PCR en colonia para *PTEN* wt y mt. 10/10 colonias analizadas wt y 8/10 colonias analizadas mt resultaron positivas para este análisis, con un tamaño de banda de 1376 pb. Gel de agarosa al 1%.

Análisis de restricción

La presencia de sitios de restricción para algunas enzimas dentro de los exones de las diferentes variantes fue aprovechada para poder determinar con mayor certeza la presencia o ausencia del inserto en cuestión. Para este análisis sólo en el caso de *FANCB* y *MSH6* se contaban con enzimas con sitios dentro de estos exones. Las muestras de *FANCB* wt y mt de las minipreps obtenidas fueron digeridas con *Bam*HI-HF para linealizar la construcción, con *Bam*HI-HF y *Xho*I para liberar el inserto clonado analizando así su tamaño y con *Mfe*I-HF, esta última con un sitio

de restricción único en el inserto *FANCB* para linealizar la construcción solamente en presencia del inserto *FANCB* dentro de la secuencia. Las muestras de *MSH6* wt y mt fueron digeridas de igual manera con *Bam*HI-HF, con *Bam*HI-HF y *Xho*I y con una enzima adicional *Nsi*I, con dos sitios de restricción: uno dentro del inserto *MSH6* y otro en el vector, generando un patrón de bandeado visualizable de 3,252 pb y 1200 pb. El resto de las construcciones fueron analizadas únicamente con las enzimas *Bam*HI-HF para linealizar y *Bam*HI-HF/*Xho*I para liberar el fragmento.

Este análisis permitió filtrar aún más las clonaciones obtenidas, determinando cuales no tenían el inserto por medio del PCR en colonia y cuales no contenían el inserto de interés por medio del análisis de restricción.

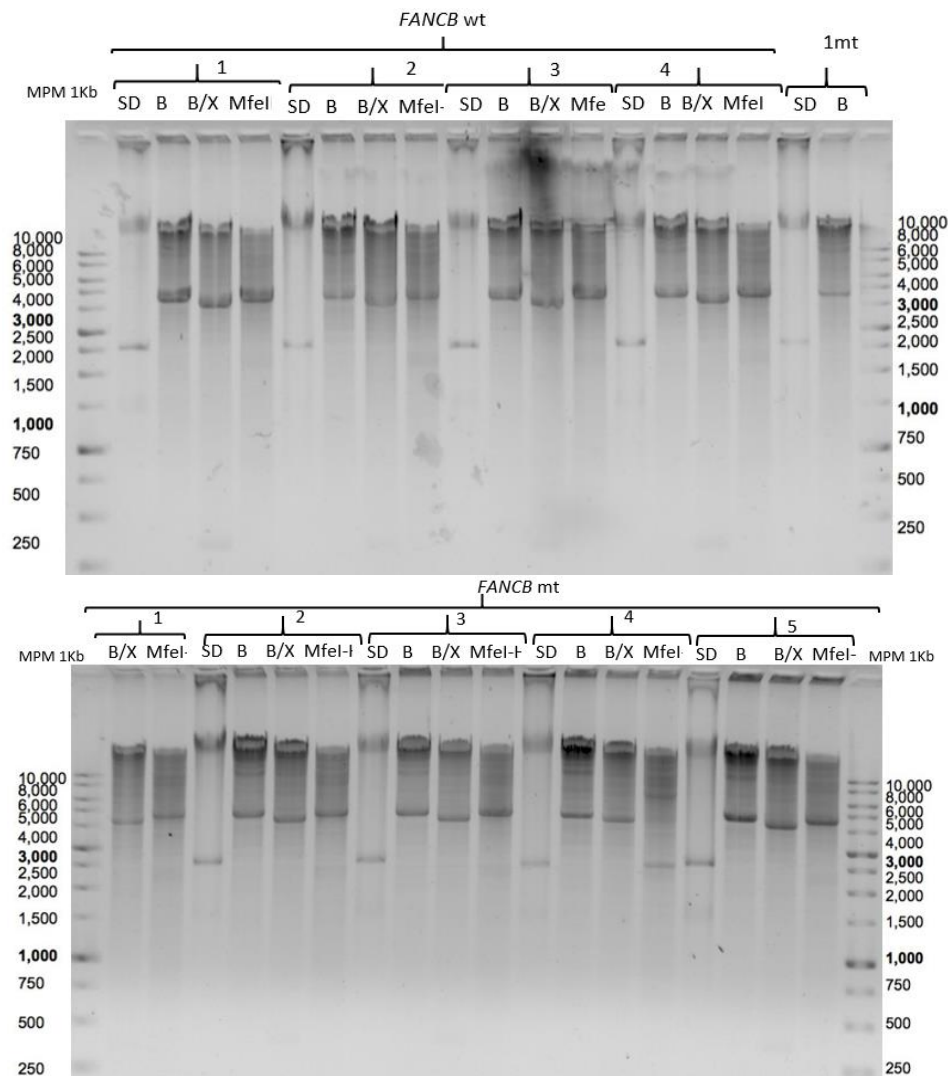


Figura 24. Resultado del análisis de restricción de *FANCB* wt y mt de las minipreps positivas. Se observa la construcción linealizada con *Mfe*I-HF (4501 pb) en las muestras 1, 2, 3, 4 wt y 1, 2, 3 y 5 mt lo que indica la presencia del inserto *FANCB* en éstas.

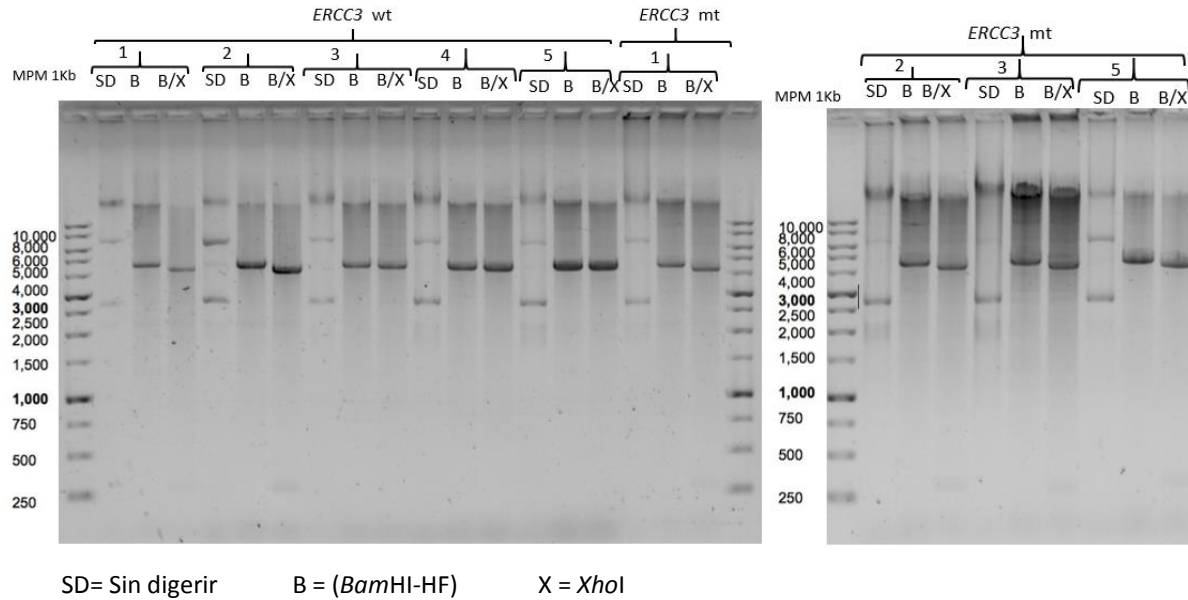


Figura 25. Resultado del análisis de restricción de *ERCC3* wt y mt de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4477 pb) y con *Bam*HI-HF y *Soy*, la liberación del inserto de 300 pb *ERCC3* en la muestra 1 y 2 wt y 1, 2, 3 y 5 mt, lo que indica la presencia del inserto *ERCC3* en éstas.

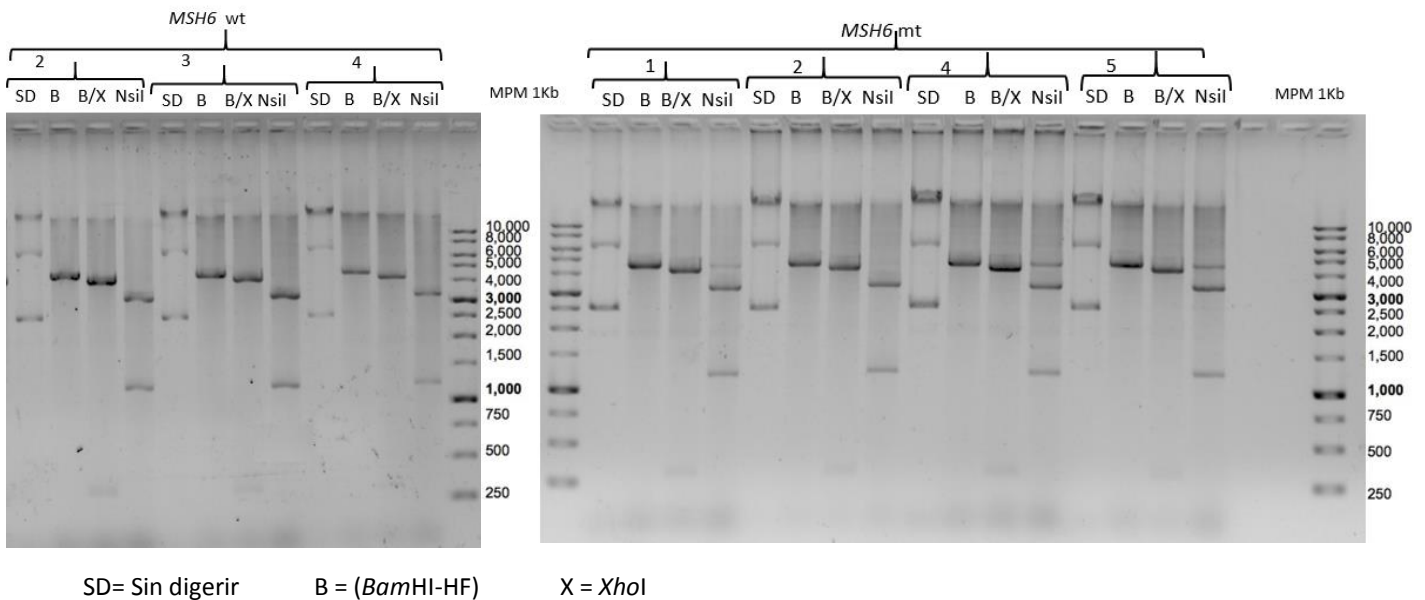


Figura 26. Resultado del análisis de restricción de *MSH6* wt y mt de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4467 pb), la liberación del inserto *MSH6* de 333 pb con *Bam*HI-HF y *Xho*I y la presencia de los fragmentos de 3,252 pb y 1200 pb tras la digestión con *Nsil*, en las muestras 2, 3 y 4 wt y 1, 2, 4 y 5 mt, lo que indica la presencia del inserto *MSH6* en éstas.

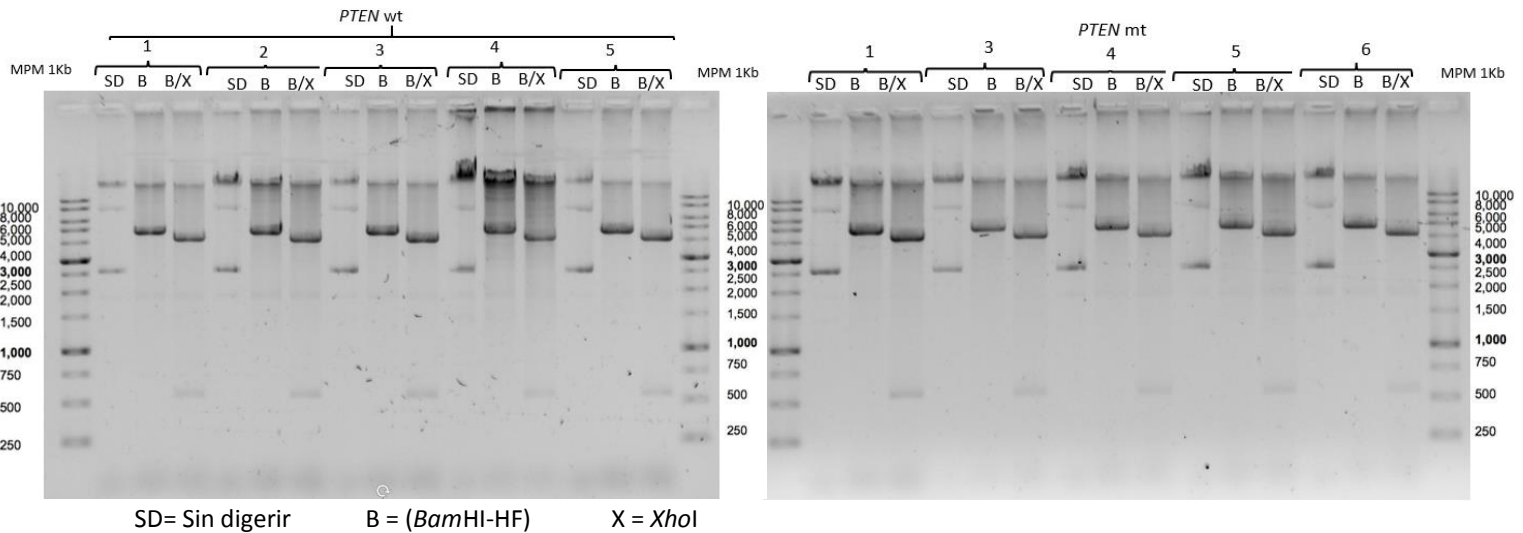


Figura 27. Resultado del análisis de restricción de *PTEN* wt y mt de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4714 pb), y la liberación del inserto *PTEN* de 580 pb con *Bam*HI-HF y *Xho*I, en las muestras 1, 2, 3, 4 y 5 wt y 1, 3, 4, 5 y 6 mt, lo que indica la presencia del inserto *PTEN* en éstas.

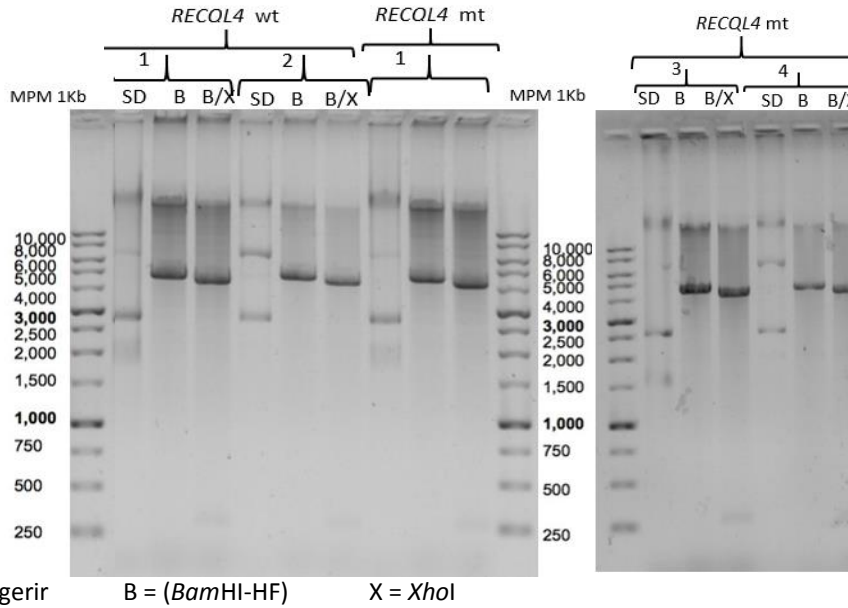


Figura 28. Resultado del análisis de restricción de *RECQL4* wt y mt de las minipreps positivas. Se observa la construcción linealizada con *Bam*HI-HF (4408 pb) y con *Bam*HI-HF y *Xho*I, la liberación del inserto de 274 pb *RECQL4* en la muestra 1 y 2 wt y 1, 3 y 4 mt, lo que indica la presencia del inserto *RECQL4* en éstas.

Análisis de la secuencia obtenida por secuenciación Sanger de cada variante wt y mt

Fueron seleccionadas por medio de estos últimos análisis 48 clonas de las 5 variantes de splicing en sus alelos *wt* y *mt*, las cuales se secuenciaron en la empresa MacroGen. Se analizaron los resultados de la secuenciación y se eliminaron las clonaciones que ganaron mutaciones adicionales en la secuencia durante el procedimiento, probablemente introducidas por la polimerasa durante la amplificación, ya que ésta (2X KAPA Taq ReadyMix de KAPA BIOSYSTEMS) posee una tasa de aproximadamente 1 error por 2.2×10^5 nucleótidos incorporados. Se seleccionaron aquellas, en su variante *wt* y *mt* que correspondieran a los datos

11. Discusión

El cáncer de mama se ha convertido en una de las neoplasias predominantes a nivel mundial y cada año el número de casos aumenta. Actualmente la predisposición genética es responsable del 10% del total de los casos de cáncer de mama, y de este porcentaje, existen genes de los que se desconoce el nivel de riesgo que confieren sus mutaciones al desarrollo de esta neoplasia. Del estudio realizado en 2018 por Quezada-Urban y colaboradores se tomaron 5 variantes VUS reportadas, presentes en pacientes con SHCMO correspondientes a mutaciones en sitios canónicos de splicing en los genes *ERCC3*, *FANCB*, *MSH6*, *PTEN* y *RECQL4*, para clonar los exones flanqueados por el sitio aceptor/donador mutado para cada variante en el vector de expresión pET01. Lo anterior con la finalidad de obtener las construcciones de los minigenes que puedan ser empleados en análisis funcionales de corte y empalme posteriores y así definir la patogenicidad a nivel funcional de dichas mutaciones.

Análisis *in silico*

Existen diversos programas bioinformáticos para predecir el efecto de mutaciones en el corte y empalme. De entre los programas que evalúan las mutaciones en sitios canónicos de splicing (aceptor y donador) están los programas seleccionados para el análisis *in silico* de este estudio: Human Splicing Finder, MaxEntScan, NetGene2 y NNSplice. El programa human splicing finder, es uno de los programas de predicción de splicing más completos hasta el momento; integra matrices de detección de elementos intrónicos y exónicos, matrices de detección de secuencias de reconocimiento para las secuencias reguladoras ya existentes, además del reconocimiento de sitios de splicing con un valor de probabilidad asignado (Desmet et al, 2009). MaxEntScan se basa en el enfoque para modelar las secuencias de motivos de secuencia corta, como los involucrados en el empalme de ARN, que simultáneamente cuenta con dependencias entre las posiciones no adyacentes y adyacentes. Este método se basa en el "Principio de máxima entropía" y generaliza la mayoría de los modelos probabilísticos anteriores de motivos de secuencia (Yeo et al, 2004). Por otro lado NetGene2 y NNSplice se desarrollaron solo para análisis de sitios canónicos de corte y empalme así como posiciones vecinas (Jian et al, 2014) (véase anexo 1).

Los resultados de este análisis arrojaron un efecto negativo en el corte y empalme para las variantes *ERCC3*, *PTEN* y *RECQL4*, mientras que para las variantes en los genes *FANCB* y *MSH6*, no mostraron ningún impacto predicho. A pesar que existen múltiples programas bioinformáticos para evaluar las mutaciones en los sitios de splicing, se sabe que los algoritmos no son exactos en todos los casos, por lo que se requieren además de estas predicciones *in silico*, el uso de análisis funcionales que simulen las condiciones bajo las cuales se llevan a cabo estos procesos biológicos, para así poder validar la patogenicidad de estas variables de splicing a nivel funcional. Aunque los eventos más comunes del corte y empalme son la omisión de exón y la selección del sitio de empalme alternativo, hay estudios recientes que demuestran la aparición de diferentes proporciones de isoformas que además incluyen la delección de algunos nucleótidos o incluso la presencia de cierta proporción de la isoforma silvestre (wt) (Acedo et al, 2012).

Amplificación

La enzima empleada para la amplificación por PCR fue 2X KAPA Taq ReadyMix de KAPA BIOSYSTEMS, carece de actividad 3' a 5' exonucleasa (proofreading) por lo que tiene una tasa de aproximadamente 1 error por 2.2×10^5 nucleótidos incorporados. Aunque existen enzimas de

alta fidelidad para este tipo de análisis, como *Pfu* (de *Pyrococcus furiosus*), la cual posee actividad *proofreading* con tasa de error de 1×10^6 es decir 10 veces menor que la Taq polimerasa empleada, para este análisis se corroboró la ausencia de mutaciones adicionales en el exón clonado y sus secuencias adyacentes por medio de secuenciación Sanger. Así, se pudo confirmar que no hubieron errores de polimerasa presentes en las clonas.

Clonación direccionada

La transformación bacteriana para el proceso de clonación depende de la capacidad de competencia de las células bacterianas, es decir, la capacidad de incluir ADN exógeno, estas bacterias pueden ser sometidas a tratamientos que les permiten adquirir esta capacidad. Para este trabajo se emplearon bacterias calcio-competentes, las cuales son sometidas a métodos químicos empleando cloruro de calcio (método reportado por Mandel e Higa en 1970). Éste método de calcio-competencia permite alcanzar una eficiencia de 5×10^6 transformantes por μg de plásmido (Cohen *et al*, 1972). En este trabajo se obtuvo una eficiencia baja, ya que se obtuvieron un mínimo de 3 colonias y un máximo de 28 (véase tabla 6). Lo anterior se puede deber a que la cantidad de plásmido empleado para la ligación fue de tan solo 50-70 ng y a que la totalidad de plásmido digerido no se haya ligado al 100%. Sin embargo, aunque se obtuvieron pocas clonas, se pudieron tener construcciones de todos los genes en sus genotipos silvestres y mutantes.

Secuenciación Sanger

Por medio de la secuenciación Sanger, se pudo validar la secuencia del inserto de cada variante y descartar aquellas clonaciones con mutaciones adicionales. Esto fue importante, ya que estas mutaciones adicionales podrían alterar el resultado del ensayo de minigen posterior. Además, fue indispensable la verificación debido a que las mutaciones presentes en las pacientes son heterocigotas, lo cual implica que teóricamente sólo el 50% de las clonaciones del alelo mutante tendrían la mutación. Debido a esto, se puede afirmar que se cumplió el objetivo del presente proyecto: clonar los exones de 5 variantes de splicing potencialmente patogénicas de los genes *ERCC3*, *FANCB*, *MSH6*, *PTEN* y *RECQL4* en un vector de expresión, así como verificar la secuencia del inserto por secuenciación tipo Sanger. De esta manera las clonas obtenidas están listas para la realización de los estudios funcionales pertinentes, los cuales podrán determinar el posible carácter patogénico de estas variantes VUS, y de esta manera poder contribuir a la su correcta clasificación.

Interpretación de las variantes de significado clínico incierto

En 2013 se estableció un grupo de trabajo formado por ACMG, Asociación para Patología Molecular (AMP), que representan a los directores de laboratorios clínicos, con el objetivo de desarrollar una recomendación para el uso de terminología estándar para clasificar variantes genéticas utilizando la evidencia disponible, ponderada de acuerdo con un sistema desarrollado a través de la opinión de expertos, el consenso del grupo de trabajo y la opinión de la comunidad.

Se proporcionaron criterios para la clasificación de las variantes como patogénicas, probablemente patogénicas, benignas o probablemente benignas y si una variante no cumple criterios de uso de cualquiera de estos conjuntos (patógenos o benignos), o la evidencia de benigno y patógeno es confluyente, la variante por defecto se clasifica como de significado clínico incierto. Las variantes seleccionadas para este trabajo tienen una clasificación de variantes de significado clínico incierto (*ERCC3*, *MSH6*, *PTEN* y *RECQL4*), a excepción de la variante *FANCB* que es clasificada en el ClinVar como benigna/probable benigna. Sin embargo los artículos que

son considerados para esta clasificación no realizaron análisis funcionales, un elemento fundamental para la clasificación de estas variantes VUS.

Determinar el efecto en el corte y empalme, nos permitirá contribuir a la clasificación o reclasificación de las variantes estudiadas en el presente trabajo, lo cual, a largo plazo, podría permitir un mejor asesoramiento a las personas que por análisis genético se les detectara estas mutaciones en dichos genes.

Determinar el efecto en el corte y empalme de estas cinco variantes, nos permitirá contribuir a la clasificación o reclasificación de las variantes estudiadas en el presente trabajo, lo cual, a largo plazo, podría permitir un mejor asesoramiento a las personas que por análisis genético se les detectara estas mutaciones en dichos genes.

12. Conclusiones

1. El análisis *in silico* predijo un efecto negativo en el corte y empalme para las variantes *ERCC3*, *PTEN* y *RECQL4*, mientras que para las variantes en los genes *FANCB* y *MSH6*, no.
2. Se clonaron los exones de las cinco variantes de splicing del presente trabajo tanto *wt* como *mt* en el vector de expresión pET01.
3. Se verificaron por análisis de restricción la clonaciones de las 5 variantes.
4. Se verificaron por secuenciación Sanger las clonaciones de 48 clonas, obteniendo un resultado positivo para al menos 1 clonación de las variantes: *FANCB*, *ERCC3*, *MSH6*, *PTEN* y *RECQL4*, tanto *wt* como *mt*.
5. Las construcciones realizadas están listas para los análisis funcionales de estas variantes.

13. Perspectivas

Este trabajo aporta las construcciones de los vectores de expresión de las 5 variantes de splicing en los genes *ERCC3*, *FANCB*, *MSH6*, *PTEN* y *RECQL4*, para ser probados a través de un ensayo de minigen. Estos ensayos permitirán establecer si estas mutaciones tienen algún impacto negativo en el corte y empalme y a su vez, podrán contribuir a su clasificación como variantes benignas o patogénicas.

14. Referencias

- Acedo, A., Sanz, D. J., Durán, M., Infante, M., Pérez-Cabornero, L., Miner, C., & Velasco, E. A. (2012). Comprehensive splicing functional analysis of ADN variants of the BRCA2 gene by hybrid minigenes. *Breast Cancer Research*, 14(3), R87.
- Antoniou A, Pharoah P.D.P., Narod S., Risch H. A., Eyfjord JE, Hopper J.L., Loman N., Olsson H., Johansson O., Borg A., Pasini B., Radice P., Manoukian S. Eccles D. M., Tang N., Olah E., Anton-Culvier H., Warner E., Lubinski J., Gronwald J., Tulinius H., Thorlacius S., Eerola H., Nevanlinna H., Syrjakoski K., Kallioniemi O. P., Thomson D., Evans C., Peto J., Lalloo R. Evans D. G. y Easton D. F. 2003. Average Risks of Breast and Ovarian Cancer Associated with BRCA1 or BRCA2 Mutations Detected in Case Series Unselected for Family History: A Combined Analysis of 22 Studies. *Am J Hum Genet*; 72(5):1117–30 pp.
- Antoniu A. C. 2014. Breast-cancer risk in families with mutations in PALB2. *N Engl J Med*, 371, 497-506pp.
- Ast, G. 2004. "How did alternative splicing evolve?" *Nat Rev Genet* 5(10): 773-782 pp.
- Baralle, M., Baralle, D., De Conti, L., Mattocks, C., Whittaker, J., Knezevich, A., & Baralle, F. E. (2003). Identification of a mutation that perturbs NF1 agene splicing using genomic ADN samples and a minigene assay. *Journal of medical genetics*, 40(3), 220-222.
- Baralle, D., and Baralle, M. (2005) Splicing in action: assessing disease causing sequence changes. *J. Med. Genet.* 42, 737–748pp.
- Baralle, D., A. Lucassen and Buratti E. 2009. "Missed threads. The impact of pre-mRNA splicing defects on clinical practice." *EMBO Rep* 10 (8): 810-816 pp.
- Bonnal S., Vigevani L., Valcárcel J. 2012. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov.* (11):847-59 pp.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A Global cáncer statistics 2018: GLOBOCAN estimates of incidence and mortality woldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018 Nov; 68 (6):394-424. doi: 10.3322/caac.21492. Epub 2018 Sep 12.
- Brooks A., Choi S. P., Waal L., Sharifnia T. Imielinski M. Saskena G. Pedamayu C. S. Sivachenco A., Rosembreg M., Chmielecki J., Lawrence M. S., De Luca D. S., Getz G. y Meyerson M. 2014. A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. 9(1): 1-10pp.
- Castrezana Campos, María del Rocío. (2017). Geografía del cáncer de mama en México. *Investigaciones geográficas*, (93)<https://dx.doi.org/10.14350/rig.56879>
- Chirivella González I., Garcés Honrubia V. Cáncer de mama hereditario, más allá de *BRCA1/2*. 2018. *Rev Génética médica y genómica*.Valencia, España.
- Cohen S., Chang A. y Hsu L. 1972. Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of Escherichia coli by R-Factor ADN. *Proc. Natl. Acad. Sci. U.S.A.* 69(8):2110-2114 pp.
- Deans, A. J., & West, S. C. (2011). ADN interstrand crosslink repair and cancer. *Nature reviews. Cancer*, 11(7), 467–480pp.
- Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37-67pp.
- Dey P. 2004. Aneuploidy and malignancy: an unsolved equation. *J Clin Pathol*; 57: 1245-9.

- Dexheimer, T. S. (2013). DNA repair pathways and mechanisms. In L. A. Mathews, S. M. Cabarcas, E. M. Hurt (Eds.), DNA repair of cancer stem cells (pp. 19–32). Dordrecht, Netherlands: Springer. doi:10.1007/978-94-007-4590-2_2
- Devlin t. m. 2004. bioquímica 4a edición. reverté Barcelona. 1240pp.
- Díez O., Gutiérrez-Enríquez S., Ramón Cajal T. 2006. Genes de susceptibilidad al cáncer de mama. Medicina Clínica. Volume 126, Issue 8. 304-310pp.
- Dillon DA, Guidi AJ, Schnitt SJ. Pathology of invasive breast cancer. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 4th ed. Philadelphia, Pa: Lippincott-Williams & Wilkins; 2010: 374–407.
- Dulak, A. M., Stojanov, P., Peng, S., Lawrence, M. S., = Fox, C., Stewart, C., Bandla, S., Imamura, Y., Schumacher, S. E., Shefler, E., McKenna, A., Cibulskis, K., Sivachenko, A., Carter, S. L., Saksena, G., Voet, D., Ramos, A. H., Auclair, D., Thompson, K., Sougnez, C., Onofrio, R. C., Guiducci, C., Beroukhim, R., Zhou, D., Lin, L., Lin, J., Reddy, R., Chang, A., Luketich, J. D., Pennathur, A., Ogino, S., Golub, T. R., Gabriel, S. B., Lander, E. S., Beer, D. G., Godfrey, T. E., Getz, G. & Bass, A. J. (2013). «Exome and whole genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity.». Nature Genet. 45 (5).
- Easton. 2015 Gene-panel sequencing and the prediction of breast cancer RISK. N Engl J Med, 372, 2243-57pp.
- Ellisen L. W. y Haber D. A. 1998. Hereditary Breast Cancer. Annual Review of Medicine 49:1, 425-436 pp.
- Eyras E. 2016. El papel del Splicing en cáncer. Consultado el 29 de agosto de 2017 en: <https://revistageneticamedica.com/2016/06/07/splicing-cancer/>
- Friedenreich CM, Orenstein MR. Physical activity and cancer prevention: etiologic evidence and biological mechanisms. *J Nutr.* Nov 2002; 132(11 Suppl):3456S–3464S.
- González KD. 2009. Beyond Li Fraumeni syndrome: clinical characteristics of families with p53 germline mutations. *J Clin Oncol*, 27, 1250-6pp.
- Goodsell, David. Cellular Landscapes in Watercolor. *Journal of Biocommunication*, [S.I.], v. 40, n. 1, feb. 2016. ISSN 0094-2499. Available at: <<https://firstmonday.org/ojs/index.php/jbc/article/view/6627/5596>>. Date accessed: 12 July 2019. doi:<https://doi.org/10.5210/jbc.v40i1.6627>.
- Herveg, J. (2006). *Study on Legal and Regulatory aspects of eHealth "Legally eHealth" : deliverable 4 : aspects of competition and trade law*. Namur: CRID.
- Hoki Y., R. Araki, A. Fujimori, T. Ohhata, H. Koseki, R. Fukumura, M. Nakamura, H. Takahashi, Y. Noda, S. Kito, M. Abe Growth retardation and skin abnormalities of the Recq4-deficient mouse. *Hum. Mol. Genet.*, 12 (2003), pp. 2293-2299pp.
- Hernandez-Moro Hernández-Moro, Cristina; Curiel-García, Álvaro; Díez-Gómez, Beatriz; Acedo, Alberto; Infante, Mar; Durán, Mercedes; Velasco, Eladio. 2013. Splicing functional assays of a BRCA1 minigene with exons 15-19.
- Ichikawa, T. Noda, Y. Furuichi. 2002. Preparation of the gene targeted knockout mice for human premature aging diseases, Werner syndrome, and Rothmund-Thomson syndrome caused by the mutation of ADN helicases. *Nippon Yakurigaku Zasshi*, 119, 219-226 pp.
- Instituto Nacional de Cancerología, 2004. El Cáncer. Aspectos básicos sobre su biología, clínica, prevención, diagnóstico y tratamiento. Ministerio de Protección Social. República de Colombia. Consultado el 20 de enero de 2019. En: <http://cancer.gov.co/documentos/Carillas/Elcancer.pdf>

- INEGI (Instituto Nacional de Estadística y Geografía) (2014), “Estadísticas a propósito del Día mundial de la lucha contra el cáncer de mama 19 de octubre”, Instituto Nacional de Estadística, Geografía e Informática, México.
- INEGI. Estadísticas de mortalidad 2015. Base de datos [<http://www.inegi.org.mx/saladeprensa/aproposito/2015/mama0.pdf>]
- Jian X, Boerwinkle E, Liu X. *In silico tools for splicing defect prediction: a survey from the viewpoint of end users. Genet Med* 2014; 16:497–503pp.
- Kapoor NS et al. 2015. Multigene Panel Testing Detects Equal Rates of Pathogenic BRCA1/2 mutations and has higher diagnostic yield compared to limited BRCA1/2 Analysis alone in patients at risk for hereditary breast cancer. *Ann Surg oncol*, 22, 3282-8pp.
- Kato Y., Alavattam K. G., Sin HS, Meetei A. R., Pang Q., Andreassen P. R. 2015. FANCB is essential in the male germline and regulates H3K9 methylation on the sex chromosomes during meiosis. *Hum Mol Genet.*; 24: 5234–5249 pp.
- Kimball, J. W. (31 de octubre de 2015). DNA repair (Reparación del ADN). En *Kimball's biology pages*. Consultado en <http://www.biology-pages.info/D/DNArepair.html#DSBs>.
- Kitao S., A. Shimamoto, M. Goto, R.W. Miller, W.A. Smithson, N.M. Lindor, Y. Furuichi Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat. Genet.*, 22 (1999), pp. 82-84pp.
- Kohmoto T., Naruto T., Watanabe M., Kobayashi H., Okamoto N., Masuda K., Imoto I y Okamoto N. 2015. A novel COL11A1 mutation affecting splicing in a patient with Stickler syndrome. *Human Genome Variation*. Osaka Japon. 4 pp.
- Leslie N. y Downes C. P. 2004. PTEN function: how normal cells control it and tumour cells lose it. Consultado el 05 de octubre de 2017 en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1133909/>
- López-Bigas N, Audit B, Ouzounis C, Parra G, Guigo R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett* 2005; 579:1900–3.
- Mandel M and Higa A (1970) Calcium-dependent bacteriophage DNA infection. *J Mol Biol* 53(1): 159–162.
- Mahesh N.Sangrithi, Juan A.Bernal, MarkMadine, AnnaPhilpott, JoonLee, William G.Dunphy, Ashok R.Venkitaraman. Initiation of ADN Replication Requires the RECQL4 Protein Mutated in Rothmund-Thomson Syndrome. 2005. Volume 121, Issue 6. 887-898pp.
- McPherson, K., Steel, C., & Dixon, J. M. (2000). Breast cancer—epidemiology, risk factors, and genetics. *Bmj*, 321(7261), 624-628.
- Mann, J. R., Backlund, M. G., & DuBois, R. N. (2005). Mechanisms of disease: Inflammatory mediators and cancer prevention. *Nature Reviews Clinical Oncology*, 2(4), 202.
- Muñoz S. V. L. 2010. Factores de riesgo asociados con el cáncer de mama en pacientes con tamizaje de mastografía casos y controles. Tesis posdoctoral. IMSS. H. Veracruz, Ver. 30pp.
- Muñoz A. Cáncer: genes y nuevas terapias.Ed Hélice. 2004. p 6-16.2.
- Nelen MR. 2001. Breast cancer genetics: what we know and what we need. *Nat Med*, 7, 552-6pp.
- Nelson R. AACR FCPR 2008: Family History Increases Breast Cancer Risk. *Medscape Medical News* 2008 Consultado 8 diciembre 2008 Disponible en: <http://www.medscape.com/viewarticle/584292>

- NIH. 2017. Genética del cáncer. Consultado el 20 noviembre del 2018 en: <https://www.cancer.gov/espanol/cancer/causas-prevencion/genetica>
- OMS. 2017. Cáncer. Consultado el 29 de Agosto del 2017 en: <http://www.who.int/mediacentre/factsheets/fs297/es/>
- OMS (Organización Mundial de la Salud) (2012), “Cáncer”, Nota descriptiva No. 297, 2012, Organización Mundial de la Salud, México. Actualizado a febrero de 2015. [<http://www.who.int/mediacentre/factsheets/fs297/es/index.html>]
- Padgett R. A. (2012). New connections between splicing and human disease. *Trends in genetics : TIG*, 28(4), 147–154. doi:10.1016/j.tig.2012.01.001
- Parkin D. M., Pisani P., Ferlay J., 1999. Global cancer statistics, *CA Cancer J. Clin.*, 49: 33-64, 1.
- Pepe C. M., Saraco N. I., Baquedano M. S., Guercio G., Vaiani E., Berensztein E., Rivarola M. A. y Belgorosky A. 2007. El splicing alternativo del exon 5 de la citocromo p450 aromatasasa podría ser un mecanismo de regulación de la producción de estrógenos en humanos. *MEDICINA - Volumen 67 - N° 4*. Buenos Aires, Argentina. 369-373pp.
- Peralta M. O., 2007. Cáncer de Mama. *Epidemiología y Factores de Riesgo. Cuad Med Soc (Chile)* 2007, 47 (1): 18-30pp.
- Pharoah PD et al. 2001. Incidence of gastric cancer and breast cancer in CDH1 (E-Cadherin mutation carrier) from hereditary diffuse gastric cancer families. *Gastroenterology*, 121, 1348-53pp.
- Pinzón, C., & Serrano, M., & Sanabria, M. (2009). Papel de la vía fosfatidilinositol 3 kinasa (PI3K/Akt) en humanos. *Revista Ciencias de la Salud*, 7 (2), 47-66.
- Quezada Urban, R., Díaz Velásquez, C., Gitler, R., Rojo Castillo, M., Sirota Toporek, M., Figueroa Morales, A., & Delgado Enciso, I. (2018). Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility. *Cancers*, 10(10), 361.
- Romieu I, Lazcano E, Sanchez LM, Willett W, Hernandez M. Carbohydrates and the Risk of Breast Cancer among Mexican Women. *Cancer Epidemiology, Biomarkers & Prevention* 2004; 13(8):1283–9.
- Romieu I, Lajous M. The role of obesity, physical activity and dietary factors on the risk for breast cancer: Mexican experience. *Salud Publica Mex* 2009; 51 suppl 2:S172- S180
- Russo J, Tay Lk, Russo IH. Differentiation of the mammary gland and susceptibility to carcinogenesis. *Breast Cancer Res Treat.* 1982;2(1):5-73.
- Sambrook J.F. and D.W. Russell, ed., Cold Spring Harbor Laboratory Press, 2001, 2100 pp.
- Sauter E. R. y Daly M. B. 2010. Breast Cancer Risk Reduction and Early Detection. Consultado el 29 de agosto de 2017 en: <https://link.springer.com/book/10.1007/978-0387-87583-5>
- Siitonen H. A., Sotkasiira J., Biervliet M, Benmansour A, Capri Y, Cormier-Daire V., Crandall B., Hannula-Jouppi K, Hennekam R., Herzog D., Keymolen K., Lipsanen-Nyman M., Miny P., Plon S. E., Riedl S., Sarkar A., Vargas F. R., Verloes A, Wang LL, Kääriäinen H, Kestilä M. *Eur J. Hum Genet.* 2009. The mutation spectrum in RECQL4 diseases. (2):151-8 pp.
- Siitonen H. A., O. Kopra, H. Kaariainen, H. Haravuori, R.M. Winter, A.M. Saamanen, L. Peltonen, M. Kestilä Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL4 diseases. *Hum. Mol. Genet.*, 12 (2003), pp. 2837-2844

- Slattery ML, Kerber RA: A comprehensive evaluation of family history and breast cancer risk. The Utah Population Database. *Jama* 270:1563-8, 1993.
- Tischkowitz MD, Hodgson SV. Fanconi anaemia. *Journal of Medical Genetics* 2003; 40:1-10pp.
- Valko M, Izakovic M, Mazur M, et al. Role of oxygen radicals in ADN damage and cancer incidence. *Mol Cell Biochem* 2004; 266: 37-56.
- Velasco, E., Acedo, A., & Díez-Gómez, B. (2014). Plásmido pSAD para ensayos funcionales de splicing.
- Vennos E. M., W.D. James. Rothmund-Thomson síndrome *Dermatol. Clin.*, 13 (1995), pp. 143-150pp.
- Wahl M. C., Will C. L. y Lührmann R. 2009. The spliceosome: design principles of a dynamic RNP machine *Cell*. Consultado el 28 de agosto del 2017 en: https://www.institutoroche.es/biotecnologia/76/splicing_alternativos_y_enfermedad.
- Walsh T, Casadei S, Coats KH, et al. Spectrum of Mutations in *BRCA1*, *BRCA2*, *CHEK2*, and *TP53* in Families at High Risk of Breast Cancer. *JAMA*. 2006;295(12):1379–1388. doi:10.1001/jama.295.12.1379
- Willingham AT, Deveraux QL, Hampton GM, et al. RNAi HTS: exploring cancer by systematic loss-of-function. *Oncogene* 2004; 23: 8392-400.
- Weeda, G., van Ham, R. C., Vermeulen, W., Bootsma, D., van der Eb, A. J., & Hoeijmakers, J. H. (1990). A presumed ADN helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell*, 62(4), 777-791.
- Yeo, G., S. Hoon, B. Venkatesh, and C. Burge, 2004, Nov)Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc. Natl. Acad. Sci. U.S.A.* 101: 15700–15705.
- Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, Chalkidis G, Suzuki Y, Shiosaka M, Kawahata R, Yamaguchi T, Otsu M, Obara N, Sakata-Yanagimoto M, Ishiyama K, Mori H, Nolte F. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64 (2011). 10.1038/nature10496pmid:21909114 doi: 10.1038/nature10496.

ANEXO 1

Programas bioinformáticos para el análisis del corte y empalme. Consultado de Jian *et al*, 2014.

Programas bioinformáticos	Formato de entrada	Formato de salida	Interpretación
Human Splicing Finder www.umd.be/HSF	Secuencias únicas $\leq 5,000$ pb	Puntuación (0 ~ 100)	Una puntuación más alta implica un mejor sitio de empalme potencial
MaxEntScan http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html	Únicas/múltiples secuencias (5': 9 pb (-3~+6); 3': 23pb (-20~+3))	Puntuación máxima de entropía	Una puntuación más alta implica una mayor probabilidad de que la secuencia sea un verdadero sitio de empalme
NetGene2 www.cbs.dtu.dk/services/NetGene2	Secuencias únicas (200 pb < longitud < 80,000 pb)	Puntuación de confianza (0 ~ 1)	Una puntuación más alta implica un mejor sitio de empalme potencial
NNSplice www.fruitfly.org/seq_tools/splice.html	Únicas/múltiples secuencias	Puntuación (0~1)	Una puntuación más alta implica una alta confianza de que la secuencia sea un verdadero sitio de empalme