



Universidad Nacional Autónoma de México

---

---

Facultad de Estudios Superiores Iztacala

Clonación y caracterización del gen *APOBEC3B*

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGO

P R E S E N T A:

RUÍZ CRUZ ALAIL ANTONIO

DIRECTOR DE TESIS:

Dr. FELIPE VACA PANIAGUA

México 2015



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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

Esta tesis la dedico principalmente a Patricia por ser el apoyo que hasta el día de hoy no me ha permitido desistir de cada locura que pasa por mi cabeza, te agradezco por ser la mejor persona que conozco y por ser la mamá más fabulosa del mundo, te quiero.

También quiero dedicar este trabajo a Luis por ser el mejor hermano y sin duda a mis abuelos Elena y Carmen por ser mis segundos padres, igualmente les dedico a todos los miembros de mi familia este trabajo, ya que sin saberlo han contribuido que todo esto se haya convertido en realidad. Gracias.

Y para concluir la dedico a ti Itzel Sarahí, por ser la inagotable fuente de mi inspiración: *Je t'aime d'ici à l'éternité.*

## **AGRADECIMIENTOS**

Quiero agradecer principalmente a mi tutor, el Dr. Felipe Vaca Paniagua por darme la oportunidad de formar parte de su equipo de trabajo, por el apoyo brindado para desarrollar este proyecto de tesis y sin duda por las enseñanzas que en un futuro me permitirán ser un mejor profesionalista.

Agradezco la asesoría técnica brindada por la Dra. Clara Estela Díaz Velásquez para la realización de este proyecto de tesis y por la ayuda ofrecida indistintamente a todos los integrantes del laboratorio.

También quiero agradecer por supuesto a cada uno de mis compañeros de la carrera con los que pasé momentos fabulosos en clases y en todas las prácticas de campo, especialmente agradezco a mis amigos Ivone, Sandra, Daniela y Carlos. A mis compañeros y amigos del laboratorio: Héctor, David, Sandra, Monse, Rosie, Mark, Fernanda, Maciel, Stef, Luis y en especial a Marco por su ayuda en cultivo celular.

Quiero agradecer a la Dra. Martha Ofelia Salcedo, Dra. Yolanda Irasema Chirino, Dr. Luis Ignacio Terrazas, Dr. Sergio Vaca y al Dr. Felipe Vaca por los comentarios atinados hechos durante la revisión de mi tesis.

Agradezco también a los doctores Gloria Paniagua, Irasema Chirino, Leticia Moreno, Ana Victoria Vega, Itzell Gallardo, Sergio Vaca, Eric Monroy, Maximiliano Ibarra y Rafael Villalobos por la ayuda siempre amablemente ofrecida, igualmente agradezco a todos los profesores de la carrera con los cuales tuve el honor de haber aprendido lo maravilloso que es este mundo al verlo bajo los ojos de la biología y por supuesto a la Universidad Nacional Autónoma de México por abrirme las puertas y brindarme la oportunidad de caminar hacia el éxito.

# INDICE

LISTA DE FIGURAS .....	1
INDICE DE TABLAS .....	2
ABREVIATURAS .....	3
RESUMEN.....	1
1.1 Cáncer .....	2
1.2 Cáncer de mama.....	4
1.2.1 Clasificación del cáncer de mama .....	6
1.2.2 Alteraciones moleculares en el cáncer de mama .....	9
1.2.3 Tratamientos contra el cáncer de mama.....	10
1.2.4 Firmas mutacionales y cáncer de mama .....	11
1.3 APOBEC.....	12
1.3.1 APOBEC3B .....	16
2. OBJETIVOS .....	18
3. MATERIALES Y MÉTODOS .....	19
3.1 Diseño de primers .....	19
3.2 Extracción de RNA total y síntesis de cDNA.....	19
3.3 PCR-Touchdown .....	20
3.4 Clonación de <i>A3B</i> .....	21
3.5 Tamizaje de colonias.....	22
3.6 Análisis de datos de secuenciación .....	23
3.7 Transfección de clonas en línea celular MDA-MB-231 .....	24
3.8 PCR en tiempo real .....	25
4 RESULTADOS .....	26
4.1 Diseño de primers dirigidos contra el ORF de <i>A3B</i> .....	26
4.2 PCR Touchdown de líneas celulares MDA-MB-231 y MCF-7 .....	27
4.3 Tamizaje de colonias.....	29
4.4 Análisis de secuencia obtenida por secuenciación Sanger para las clonas 3 y 5.....	30
4.5 Transfección de plásmidos aislados de las clonas C3 y C5 .....	36
4.6 PCR en tiempo real para evaluar los niveles de expresión de los transgenes.....	38
5. DISCUSIÓN.....	40
6. CONCLUSIONES.....	46
7. PERSPECTIVAS.....	47
8. REFERENCIAS .....	49

## LISTA DE FIGURAS

**Figura 1.** Incidencia y mortalidad del cáncer en la población mundial durante el año 2012

**Figura 2.** Incidencia y mortalidad del cáncer en la población femenina de México durante el año 2012

**Figura 3.** Estructura lobular de la mama

**Figura 4.** Firmas mutacionales alrededor del cáncer

**Figura 5.** Reacción de desaminación llevada a cabo por la familia de enzimas APOBEC y edición de Apolipoproteína B

**Figura 6.** Dominio de coordinación de zinc de las proteínas APOBEC

**Figura 7.** Locus de los miembros de la subfamilia APOBEC3

**Figura 8.** Primers utilizados para la clonación del gen *A3B*

**Figura 9.** PCR Touchdown

**Figura 10.** Productos de la amplificación de *APOBEC3B* por medio de PCR Touchdown

**Figura 11.** Tamizaje de colonias

Figura 12. Secuencias nucleotídica y de aminoácidos de la isoforma 1 del gen *APOBEC3B*

**Figura 13.** Alineamiento de secuencias obtenidas por secuenciación Sanger de las clonas C3 y C5 vs secuencia silvestre de *APOBEC3B*

**Figura 14.** Secuencias nucleotídica y de aminoácidos de la clona C3 que codifica para el gen *APOBEC3B*

**Figura 15.** Secuencias nucleotídica y de aminoácidos de la clona C5 que codifica para el gen *APOBEC3B*

## INDICE DE TABLAS

**Tabla 1.** Incidencia y mortalidad del cáncer en la población mundial durante el año 2012

**Tabla 2.** Subtipos moleculares de cáncer de mama e inmunofenotipos

**Tabla 3.** Miembros de la familia APOBEC

**Tabla 4.** Primers de *A3B* diseñados con el programa Primer3

**Tabla 5.** Primer-BLAST de oligonucleótidos dirigidos contra el gen *A3B*

**Tabla 6.** Análisis de formación de dímeros mediante programa Oligocalc para el primer antisentido

**Tabla 7.** Predicciones resultantes del análisis realizado con SIFT y PolyPhen-2 para las mutaciones presentes en la clona C3

**Tabla 8.** Predicciones resultantes del análisis realizado con SIFT y PolyPhen-2 para las mutaciones presentes en la clona C5

## ABREVIATURAS

°C	Grados Celcius
Aa	Aminoácido
ApoB	Apolipoproteína B
APOBEC	Apolipoprotein B mRNA-editing Enzyme Catalytic, polypeptide-like
cDNA	Ácido desoxirribonucleico complementario
DCIS	Carcinoma ductal <i>in situ</i>
DNA	Ácido desoxirribonucleico
EGFR	Receptor del factor de crecimiento epidermal
HA	Hemaglutinina
HER2	Receptor del factor de crecimiento epidérmico
IARC	Agencia Internacional para la Investigación del Cáncer
IDC	Carcinoma ductal invasivo
INEGI	Instituto Nacional de Estadística y Geografía
KDa	Kilodaltones
LB	Luria Bertani
LCIS	Carcinoma lobulillar <i>in situ</i>
LDL	Lipoproteínas de baja densidad
LINEs	Elementos nucleares intercalados largos
MgCl <sub>2</sub>	Cloruro de magnesio
mM	Milimolar
mRNA	Ácido ribonucleico mensajero
NCBI	Centro Nacional para Información



	Biotecnológica
OMS	Organización Mundial de la Salud
PCR	Reacción en cadena de la polimerasa
PCR-TDW	Reacción en cadena de la polimerasa variante Touchdown
PIK3CA	Fosfatidilinositol-4-5-bifosfato 3-cinasa, subunidad catalítica alfa
RE	Receptor de estrógenos
RNA	Ácido ribonucleico
RP	Receptor de progesterona
SINEs	Elementos cortos intercalados
T <sub>m</sub>	Temperatura de alineamiento
TP53	Gen supresor de tumores 53
UV	Ultra violeta
UVB	Ultra violeta B
VHB	Virus de la hepatitis B
VIH	Virus de la inmunodeficiencia humana
VPH	Virus del papiloma humano
ZCD	Dominio de coordinación de zinc
μL	Microlitro
μM	Micromolar

## RESUMEN

El cáncer es una enfermedad genética causada por la acumulación de mutaciones que se caracterizado por presentar un crecimiento incontrolado de células, así como un alto poder invasivo de tejido. Uno de los tipos de cáncer que más afectan a la población es el cáncer de mama, ya que en el 2012, fue el segundo cáncer más diagnosticado a nivel mundial y se ubicó como la primera causa de muerte por cáncer en mujeres. En México, el cáncer de mama ocupa el primer lugar en incidencia y mortalidad de neoplasias malignas en mujeres y nueva evidencia señala que está aumentando a un ritmo acelerado en el mundo. Actualmente mediante análisis genético-molecular de los genomas de células cancerígenas se ha encontrado que el gen *APOBEC3B* es la principal fuente de mutaciones en el cáncer de mama. Esta proteína con actividad desaminasa de citidina, se encuentran sobre-expresada en los tumores de mama, pulmón, riñón y las neoplasias de cabeza y cuello, y se asocian a un fenómeno de hipermutación localizada en diferentes regiones genómicas, conocida como *kataegis*, favoreciendo la generación de alteraciones genéticas que promueven procesos oncogénicos.

En este trabajo se clonó el gen *APOBEC3B* a partir de las líneas celulares MCF-7 y MDA-MB-231 utilizando un vector de clonación que permite su expresión en células de mamífero. De esta clonación se obtuvieron 2 clonas (nombradas C3 y C4) que fueron analizadas mediante secuenciación Sanger y se encontró que contenían 5 y 6 mutaciones respectivamente. El análisis de las mutaciones en la región codificante de estas clonas se hizo con los programas computacionales SIFT y PolyPhen-2 para inferir el impacto de estos cambios a nivel de proteína y se encontró que únicamente los cambios presentes en la clona C3 podrían afectar la función de desaminasa del gen *APOBEC3B* al ser expresado. Para evaluar los genes clonados a nivel de mensajero se transfectaron ambas clonas en la línea celular MDA-MB-231 y mediante PCR en tiempo real se evaluaron los niveles de mRNA, los resultados fueron que los niveles de expresión de las células transfectadas con el vector que contiene la secuencia del gen *A3B* son más altos en comparación con las células que no fueron tratadas.

# 1. INTRODUCCIÓN

## 1.1 Cáncer

El cáncer es una enfermedad genética causada por la acumulación de mutaciones somáticas (Burns *et al.*, 2013(1,2); Forte 2013) que se caracteriza por presentar un proceso de crecimiento y diseminación incontrolados de células, así como un alto poder invasivo de tejido, capaz de provocar metástasis en regiones anatómicamente distantes del tumor primario (O.M.S., 2013). Estos cambios genéticos alteran la homeostasis celular, debido a que frecuentemente se acumulan en genes supresores de tumores y oncogenes, afectando procesos celulares fundamentales como proliferación, diferenciación y apoptosis, favoreciendo así la supervivencia de las células mutantes (Scholzová *et al.*, 2007) y dependiendo del tejido en donde se encuentren darán lugar al desarrollo de diversos tipos de cáncer. Según los datos proporcionados por la Agencia Internacional para la Investigación del Cáncer (IARC por sus siglas en inglés) en 2012 la incidencia mundial de cáncer en el mundo alcanzó los 14 millones de casos, dentro de éstos, los de mayor incidencia fueron el cáncer de pulmón (13%), de mama (11.9%), colorectal (9.7%) y estómago (6.8%) y los de mayor mortalidad fueron pulmón (19.4%), hígado (9.1%), estómago (8.8%), colorectal (8.5%) y mama (6.4%) (GLOBOCAN, 2012) (Tabla 1; figura 1).

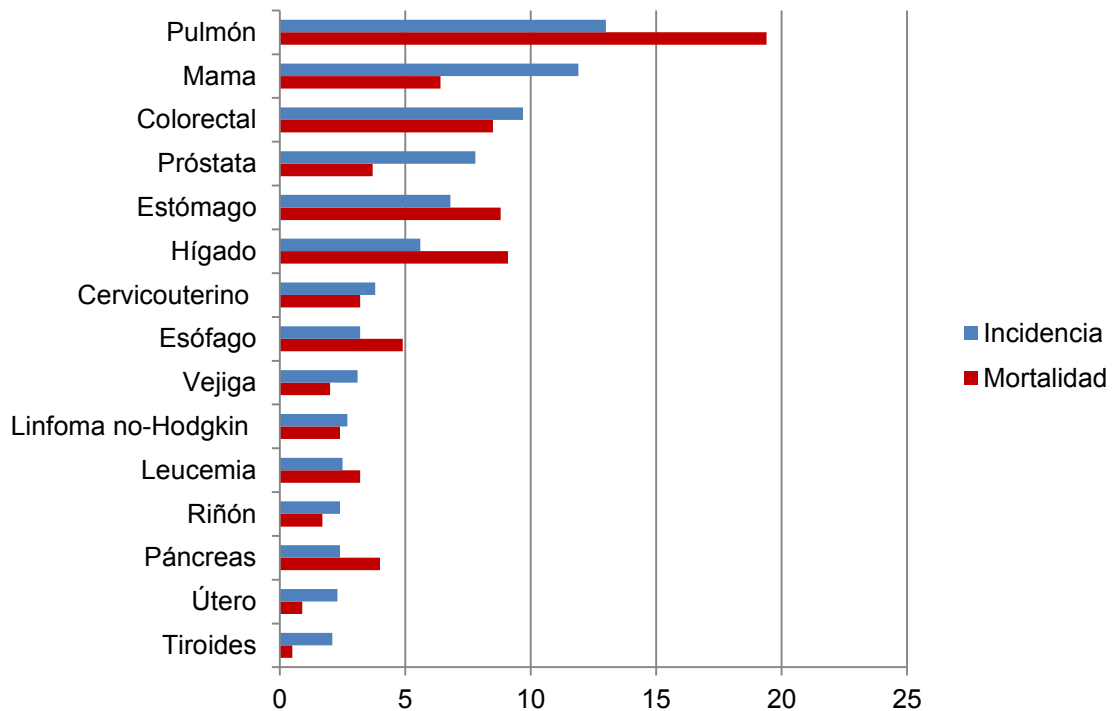
**Tabla 1.** Incidencia y mortalidad del cáncer en la población mundial durante el año 2012\*\*\*.

Cáncer	Incidencia*			Mortalidad*		
	Número	(%)	ASR** (W)	Número	(%)	ASR** (W)
Oral	300373	2.1	4	145353	1.8	1.9
Nasofaríngeo	86691	0.6	1.2	50831	0.6	0.7
Faringe (otros)	142387	1	1.9	96105	1.2	1.3
Esófago	455784	3.2	5.9	400169	4.9	5
Estómago	951594	6.8	12.1	723073	8.8	8.9
Colorectal	1360602	9.7	17.2	693933	8.5	8.4
Hígado	782451	5.6	10.1	745533	9.1	9.5
Vesícula biliar	178101	1.3	2.2	142823	1.7	1.7
Páncreas	337872	2.4	4.2	330391	4	4.1
Laringe	156877	1.1	2.1	83376	1	1.1
Pulmón	1824701	13	23.1	1589925	19.4	19.7
Melanoma de piel	232130	1.7	3	55488	0.7	0.7
Sarcoma Kaposi	44247	0.3	0.6	26974	0.3	0.3
Mama	1671149	11.9	43.1	521907	6.4	12.9
Cervicouterino	527624	3.8	14	265672	3.2	6.8
Útero	319605	2.3	8.3	76160	0.9	1.8
Ovario	238719	1.7	6.1	151917	1.9	3.8
Próstata	1094916	7.8	30.7	307481	3.7	7.8
Testículos	55266	0.4	1.5	10351	0.1	0.3
Riñón	337860	2.4	4.4	143406	1.7	1.8
Vejiga	429793	3.1	5.3	165084	2	1.9
Cerebro y SN	256213	1.8	3.4	189382	2.3	2.5
Tiroides	298102	2.1	4	39771	0.5	0.5
Linfoma de Hodgkin	65950	0.5	0.9	25469	0.3	0.3
Linfoma no-Hodgkin	385741	2.7	5.1	199670	2.4	2.5
Mieloma múltiple	114251	0.8	1.5	80019	1	1
Leucemia	351965	2.5	4.7	265471	3.2	3.4
Todos los cánceres excluyendo cáncer de piel no melanoma	14067894	100	182	8201575	100	102.4

\*Incidencia y mortalidad para todas las edades.

\*\*ASR (tasa estandarizada por edad) y proporciones por 100,000.

\*\*\*Datos tomados de [GLOBOCAN](http://globocan.iarc.fr), proyecto encargado de presentar estimaciones acerca de la incidencia, mortalidad y prevalencia de los principales tipos de cáncer a nivel internacional tomando en consideración a 184 países del mundo de la población adulta de más de 15 años. Poner la página de GLOBOCAN [http://globocan.iarc.fr/Pages/fact\\_sheets\\_population.aspx](http://globocan.iarc.fr/Pages/fact_sheets_population.aspx)

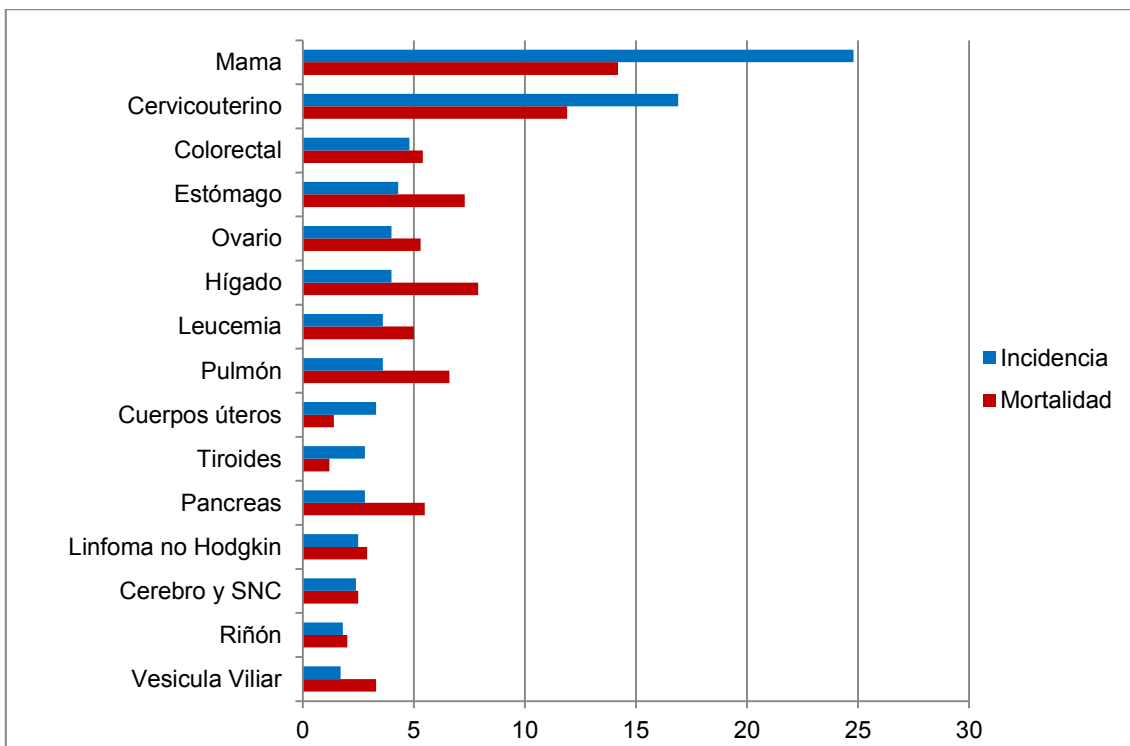


**Figura 1.** Incidencia y mortalidad del cáncer en la población mundial durante el año 2012. Imagen tomada y modificada de [GLOBOCAN](#)

## 1.2 Cáncer de mama

Dentro de los tipos de cáncer que más afectan a la población se encuentra el cáncer de mama. El cáncer de mama es considerado una enfermedad compleja y heterogénea debido al gran número de diferencias histopatológicas, moleculares y biológicas que presenta, así como la heterogeneidad clínica que muestra en relación a la respuesta al tratamiento. Lo anterior hace que sea difícil establecer una clasificación que permita distinguir grupos de pacientes que se beneficiarán con tratamientos específicos (Viale G., 2012; Cetin y Topcul, 2014). En el 2012 según el IARC, se estimó que el cáncer de mama fue el segundo cáncer más diagnosticado a nivel mundial con 1.7 millones de detecciones, después del pulmonar con 1.8 millones y aunque se puede presentar tanto en hombres como en mujeres (Giordano *et al.* 2002), es en la población femenina donde representa la enfermedad cancerígena más común (25.1%) y la causa de muerte más frecuente con 521,907 decesos en el año 2012 (GLOBOCAN, 2012). La evidencia

actual señala un aumento de nuevos casos en los países en vías de desarrollo donde el riesgo de mortalidad por la enfermedad es mayor. También en estas regiones económicamente menos favorecidas anualmente se presenta un aumento de aproximadamente 1.5% (Porter, 2008). En México, el cáncer de mama representa un grave problema de salud pública al posicionarse como el primer lugar en incidencia de neoplasias malignas en mujeres, ya que está representado por un 24.8% dentro de todos los cánceres (Arce *et al.* 2011, GLOBOCAN, 2012), además de representar la primera causa de mortalidad entre las mujeres mexicanas, atribuyéndole 5,680 (7.5%) muertes de las 43,053 que provocó el cáncer durante el año 2012 (GLOBOCAN, 2012; INEGI, 2013) (Figura 2). Entre los factores de riesgo de esta enfermedad se pueden mencionar principalmente factores hormonales como exposición a estrógenos y progesterona, factores genéticos como mutaciones hereditarias y factores reproductivos como número de embarazos, edad en la que ocurre el primer embarazo, lactancia materna, primera menstruación y menopausia, además de otros factores como antecedentes familiares, radiación ionizante sobre el pecho, y agentes ambientales (Key *et al.*, 2001; MacMahon, 2005; Hulka *et al.*, 2008 y Herbert, 2009).



**Figura 2.** Incidencia y mortalidad del cáncer en la población femenina de México durante el año 2012. Imagen tomada y modificada de GLOBOCAN

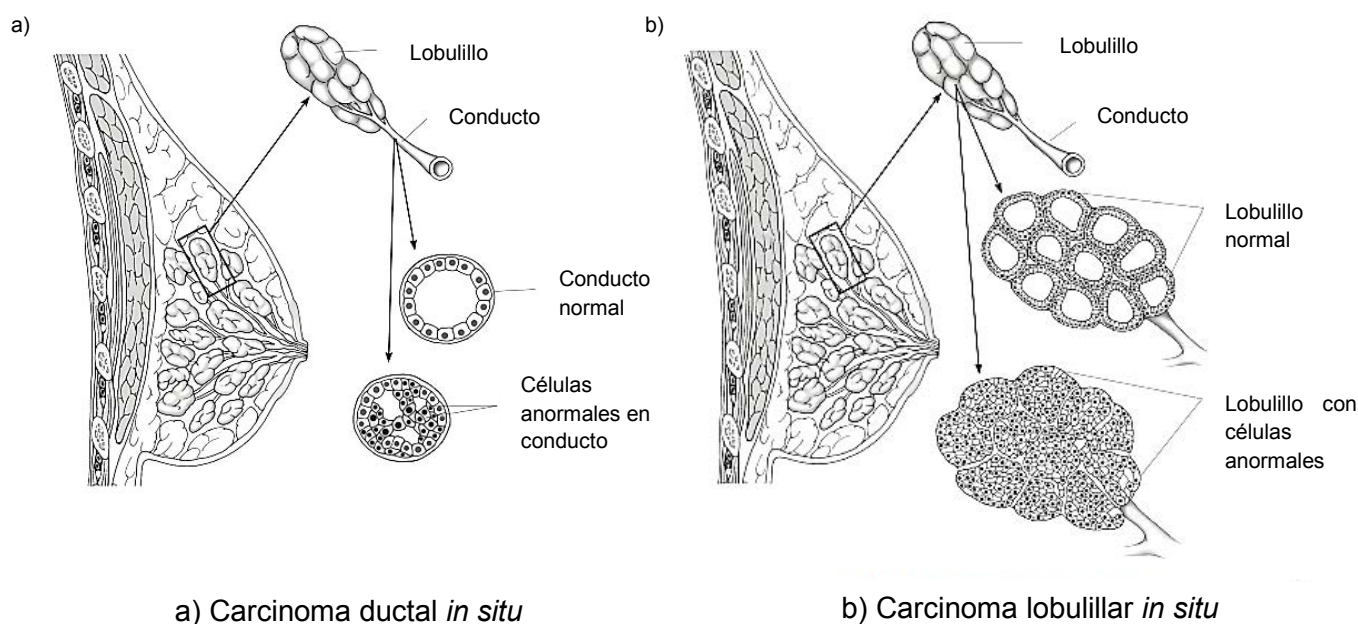
### 1.2.1 Clasificación del cáncer de mama

Los cánceres de mama se clasifican tradicionalmente mediante características histopatológicas, aunque durante la última década, se han empleado perfiles de expresión de genes por medio de técnicas moleculares en una clasificación denominada sistema “intrínseco” o molecular.

#### 1.2.1.1 Clasificación histológica

Los tumores de mama histológicamente se clasifican en tumores *in situ* e invasivos o infiltrantes (Malhotra *et al.*, 2010). Los tumores mamarios no invasivos son denominados comúnmente *in situ* (“en su sitio”), ya que no son capaces de invadir tejidos adyacentes dentro o fuera de la mama (Modesto *et al.*, 2014).

Dentro de los no invasivos se encuentran el carcinoma lobulillar *in situ* (LCIS) y el carcinoma ductal *in situ* (DCIS) (Figura 3). El carcinoma ductal *in situ* (Figura 3a) se considera el más importante de éstos ya que de todos los tumores de mama diagnosticados representa el 20% en la población Estadounidense y en México su incidencia es menor al 10% (Leonard y Swain, 2004; Modesto *et al.*, 2014). Los tumores mamarios invasivos se denominan así debido a que son capaces de causar metástasis en tejidos mamarios circundantes y en regiones distantes del organismo como los pulmones, el cerebro, el hígado o los huesos (Nguyen *et al.*, 2009; Don *et al.*, 2009). Los principales cánceres son el carcinoma ductal invasivo (IDC) y el carcinoma lobulillar invasivo (ILC). El IDC es el más frecuente, ya que se presenta del 75-80% en las neoplasias mamarias invasivas diagnosticadas, mientras que el ILC, se encuentra con el 15% de los diagnósticos (Nardone *et al.*, 2011).



**Figura 3.** Estructura lobular de la mama. Se muestran los carcinomas *in situ*. a) Carcinoma ductal *in situ*, se muestra un conducto normal y un conducto con células anormales. b) Carcinoma lobulillar *in situ*, lobulillo normal y lobulillo con células anormales. Imagen tomada y modificada de American Cancer Society <http://www.cancer.org/>



### 1.2.1.2 Clasificación molecular

La clasificación molecular del cáncer de mama se realiza mediante el denominado sistema "intrínseco", el cual está constituido por cinco subtipos de cáncer de mama (Tabla 2) (Perou *et al.*, 2000), cuyas características han sido descritas durante la última década mediante la utilización de técnicas de biología molecular. Los subtipos moleculares del cáncer de mama incluyen el "Luminal A" [positivo para receptor de estrógenos (RE) y/o positivo para receptor de progesterona (RP) y negativo para el receptor del factor de crecimiento epidérmico (HER2)], "Luminal B" [positivo para RE y/o positivo para RP y positivo para HER2], "HER2 sobreexpresado" [negativo para RE, negativo para RP y positivo para HER2], "Tipo Basal" [negativo para RE, negativo para RP, negativo para HER2, positivo para citoqueratinas 5/6 y/o positivo para HER1] y "Tumores de Mama tipo Normal" (Boyle, 2012). Más recientemente un nuevo subtipo denominado "tumores con baja Claudina" (Sabatier *et al.*, 2014) que ha sido caracterizado por tener propiedades similares a las de las células madre y por presentar una transición epitelio-mesenquima (Reis-Filho *et al.*, 2010).

**Tabla 2.** Subtipos moleculares de cáncer de mama e inmunofenotipos.

Expresión en inmunohistoquímica	Luminal A	Luminal B	HER2 sobreexpresado	Tipo basal o Triple negativo
RE	+	+	-	-
RP	+	+	-	-
HER2	-	+	+	-

Inmunofenotipos: RE: receptor de estrógenos; RP: receptor de progesterona; HER2: receptor del factor de crecimiento epidermal epidérmico

### 1.2.2 Alteraciones moleculares en el cáncer de mama

Actualmente mediante análisis genéticos y moleculares de los genomas de células cancerígenas se ha encontrado una gran diversidad de genes asociados al cáncer de mama que experimentan mutaciones, alteraciones epigenéticas y en sus niveles de expresión, favoreciendo así el desarrollo de esta patología (Millán, 2008; Hanahan y Weinberg, 2011). Entre los genes que presentan mutaciones con mayor frecuencia se encuentra *TP53* (gen supresor de tumores 53), el cual es un gen supresor de tumores que bajo situaciones de estrés puede poner fin a la progresión del ciclo celular. Este gen, también conocido como “guardián del genoma”, se encuentra alterado en aproximadamente 30% de los cánceres de mama (Hanahan y Weinberg, 2011; Alexandrov *et al.* 2013 y Dumay *et al.* 2013). El gen *BRCA1* es un supresor de tumores que participa en la reparación de rupturas en la doble cadena del DNA, es un gen que experimenta mutaciones germinales que predisponen a desarrollar tumores de mama frecuentemente antes de los 50 años (Venkitaraman, 2001 y Reis-Filho *et al.*, 2010). El gen *PIK3CA* (fosfatidilinositol-4-5-bifosfato 3-cinasa, subunidad catalítica alfa) se encuentra alterado del 18% al 40% de los cánceres de mama, promoviendo en este estado la generación segundos mensajeros, llevando a un aumento en la señalización celular hacia vías como Ras, Raf y MAP cinasas involucradas en el aumento del crecimiento y proliferación celular (Steven *et al.*, 2005; Schuab, 2011). Además la sobreexpresión de genes como *HER2*, miembro de una familia de receptores de tirosina cinasa unidas a la membrana (*HER*) se encuentra sobreexpresado de 15 a 20% en los cánceres de mama (Reis-Filho *et al.*, 2010) y *MYC* el cual es un oncogén que regula hasta un 15% de los genes humanos y está implicado en el crecimiento celular, proliferación, metabolismo, diferenciación y apoptosis (Xu J. *et al.* 2010), en el cáncer de mama se encuentra sobreexpresado en un 40% a nivel proteína (Chrzan *et al.*, 2001). Dentro de la literatura estos genes se proponen como parte de la patogénesis molecular del cáncer de mama, promoviendo la progresión tumoral, y algunos de ellos se proponen como blancos terapéuticos en la lucha contra esta enfermedad. (Rimawi *et al.*, 2011 y Qian *et al.* 2008).

### 1.2.3 Tratamientos contra el cáncer de mama

En la actualidad la quimioterapia continua siendo el tratamiento más utilizado para combatir el cáncer de mama, sin embargo, se están investigando tratamientos dirigidos contra blancos específicos que participan en el proceso de oncogénesis. Un ejemplo clásico de tratamiento dirigido en cáncer de mama utilizado ampliamente durante la última década es el relacionado con el oncogén *HER2*. En algunos subtipos de cáncer de mama el receptor *HER2* se encuentra sobreexpresado y en este estado se vuelve blanco del tratamiento con el anticuerpo monoclonal humanizado Trastuzumab (Weigel y Dowsett, 2010). *EGFR* otro miembro de la familia *HER* que se encuentra sobreexpresado hasta en un 60% en algunos tipos de cáncer de mama donde se vuelve susceptible al anticuerpo dirigido Cetuximab (Reis-Filho *et al.*, 2010). Además de tratamientos dirigidos contra una molécula en particular existen también tratamientos dirigidos contra procesos celulares fundamentales que se encuentran desregulados, como la reparación del DNA que comúnmente en cáncer de mama se relaciona con la alteración en la función de los genes *BRCA1* y *BRCA2*, que participan en el proceso de recombinación homóloga, encargado de la reparación de las rupturas en la doble cadena del DNA y donde el defecto en la función de estas proteínas conlleva a la acumulación de mutaciones y al mantenimiento de las células mutantes. Para combatir esta contrariedad hoy en día se utilizan como tratamiento las sales derivadas del platino como Cisplatino y Carboplatino, las cuales se intercalan en la doble cadena de DNA impidiendo su duplicación y provocando la entrada en apoptosis de las células (Dawood *et al.* 2010; Roberts y Gordenin, 2014).

La supervivencia de las células cancerígenas de mama así como su potencial metastásico también se ha convertido en tema de estudio para posibles tratamientos dirigidos contra la propia biología del tumor originado en la mama, esta supervivencia se ve favorecida en algunos tipos de cáncer mamario por los altos niveles de expresión de VEGF (factor de crecimiento endotelial vascular) por lo que uno de los tratamientos utilizados es el uso del anticuerpo Bevacizumab, que es un inhibidor de angiogénesis que impide la formación de nuevos vasos

sanguíneos evitando así que lleguen nutrientes a las células tumorales (Cetin y Topcul, 2014).

Esta búsqueda de blancos potenciales para el tratamiento se vuelve de vital importancia para poder proponer nuevas formas de combatir la enfermedad, aunque el trasfondo de la oncogénesis también es algo que se investiga en la actualidad y con el uso de herramientas de secuenciación se sabe ahora que agentes o fallos en mecanismos celulares podrían contribuir en gran medida en el desarrollo de esta enfermedad. Esto se sabe gracias a que estos agentes tanto endógenos como exógenos son capaces de causar mutaciones específicas, lo que se conoce como firmas mutacionales.

#### **1.2.4 Firmas mutacionales y cáncer de mama**

Las mutaciones somáticas pueden ser consecuencia de diversos factores tanto endógenos como exógenos (Alexandrov *et al.*, 2013). Los factores endógenos de daño al DNA incluyen errores en la maquinaria de replicación, estrés oxidativo, desaminación hidrolítica espontánea y daño por alquilación. Entre los factores exógenos de daño al DNA se encuentran las radiaciones UV y UVB, así como compuestos reactivos contenidos en el tabaco (Alexandrov *et al.*, 2013; Wilson, 2014). Durante la última década el análisis del genoma de células tumorales mediante el uso de técnicas de secuenciación masiva complementadas con el uso de herramientas bioinformáticas avanzadas han permitido identificar diversos agentes mutagénicos que generan cambios de nucleótidos específicos, algunos dependientes del contexto de la secuencia adyacente. A estos patrones únicos se les denominó "firmas mutacionales" (Figura 4) (Alexandrov *et al.* 2013). Estas firmas se encuentran como patrones repetitivos en diversos tipos de cáncer y permiten inferir el impacto y origen de los cambios a nivel genético de los tumores. En el caso del cáncer de mama se han identificado firmas atribuidas a la edad, mutaciones en *BRCA1/2* y a la familia de desaminasas de citidina APOBEC, entre otras (Alexandrov *et al.* 2013; Burns *et al.* 2013(a); Burns *et al.* 2013(b)).

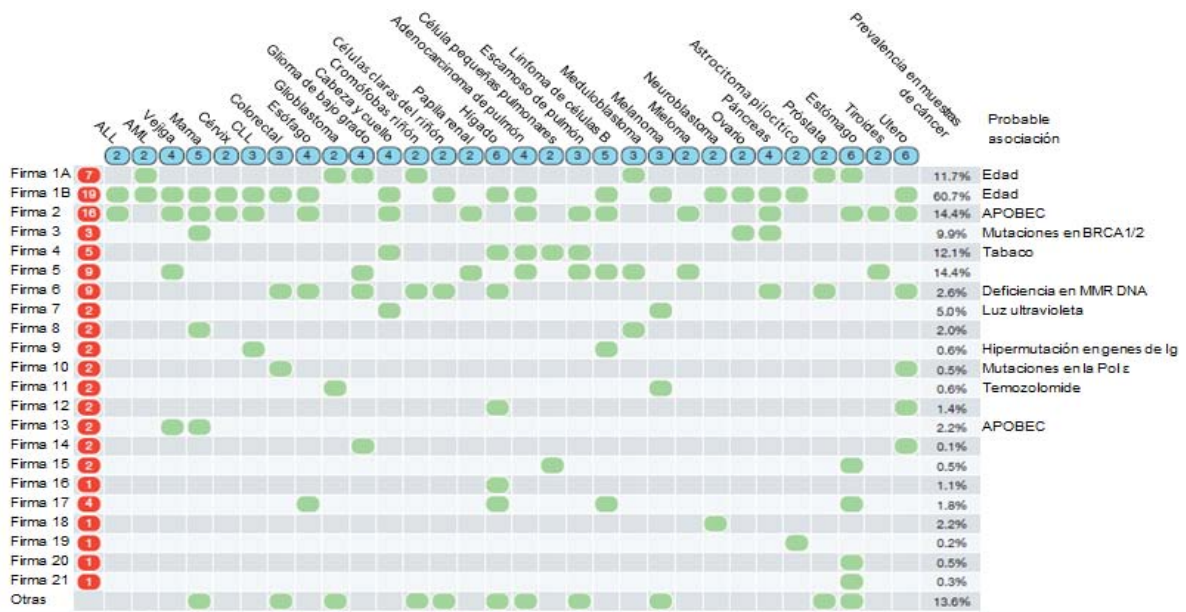


Figura 4. Firmas mutacionales alrededor del cáncer. Las más representativas son las causadas por la edad (Firmas 1A y 1B), APOBEC (Firmas 2 y 13) y tabaco (Firma 4). Verde: Firmas mutacionales presentes; Azul: Total de firmas validadas en cada tipo de cáncer; Rojo: Total de tipos de cáncer donde la firma es operativa. Imagen tomada y modificada de Alexandrov *et al.* 2013

### 1.3 APOBEC

APOBEC es una familia de proteínas que catalizan la desaminación de las citosinas generando como producto un uracilo que mediante distintos mecanismos de reparación puede establecerse en el genoma como mutaciones de timina o guanina (Figura 5a). Estas proteínas tienen un papel importante en diversos procesos biológicos, entre los que destacan: **(i)** brindar protección a la célula restringiendo la movilidad de elementos genéticos móviles como transposones y retrotransposones (Esnault *et al.*, 2005; Bogerd *et al.*, 2006; Koito y Ikeda, 2011; Smith *et al.*, 2012; Richardson *et al.*, 2014), **(ii)** actuar como factor antiviral llevando a cabo la restricción de virus patógenos (Harris y Liddament, 2004; Malim, 2009), **(iii)** tienen un rol importante en el sistema inmune de vertebrados (Moris A. *et al.*, 2014; Don-Marc y Petersen-Mahrt, 2014; Burns *et al.*, 2015), ya que inducen cambios genéticos en los genes de las inmunoglobulinas, incrementando la diversidad de anticuerpos, aumentando así la afinidad en las interacciones antígeno-anticuerpo y **(iv)** estudios recientes implican a miembros de esta familia como promotores en la mutagénesis endógena del cáncer (Don-Marc

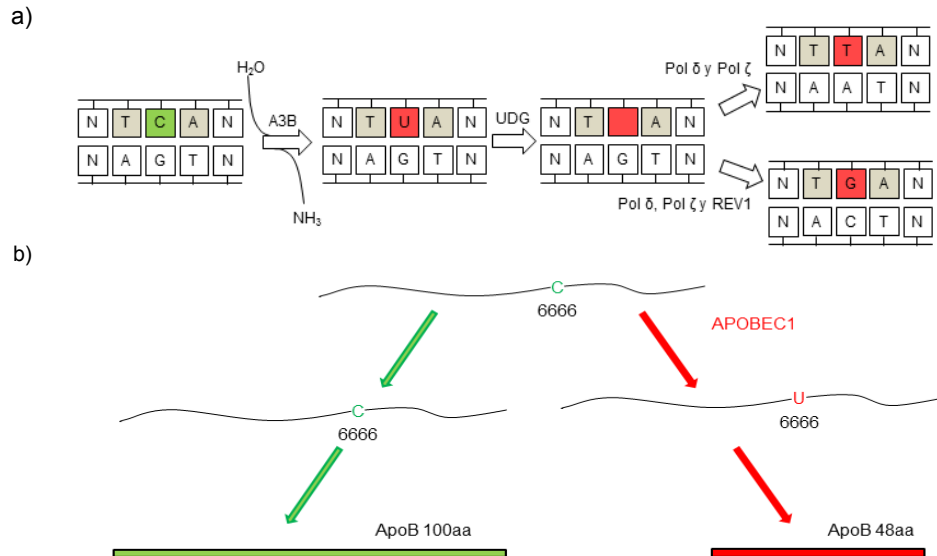
y Petersen-Mahrt, 2014; Swanton *et al.*, 2015). El primer miembro descrito y al cual se debe el nombre de esta familia es la proteína APOBEC1 (**A**polipoprotein **B** mRNA-editing **E**nzyme **C**atalytic polypeptide-like **1**), que fue identificada como una enzima editora del mRNA de la apolipoproteína B (ApoB) (Teng *et al.*, 1993REF). Habitualmente ApoB está compuesta por 100 aminoácidos, sin embargo cuando se lleva a cabo la edición por parte de APOBEC1 mediante una reacción de desaminación en la citocina ubicada en la posición 6666 del mensajero da lugar a un codón stop prematuro CAA→UAA, que genera una isoforma no funcional de la proteína de 48 aa (Figura 5b). Este cambio se traduce en una disminución en el transporte de colesterol en la célula debido a que la ApoB actúa como ligando de las lipoproteínas de baja densidad (LDL) que dentro de la partícula contienen los esteres de colesterol (Powell *et al.*, 1987; Teng *et al.*, 1993; Smith *et al.*, 2012; Moris *et al.*, 2014, y Don-Marc y Petersen-Mahrt, 2014). Hoy en día se conocen 11 miembros de la familia APOBEC, los cuales se encuentran localizados en distintos cromosomas dentro del genoma. Los miembros APOBEC1 y AID se localizan en el cromosoma 12, APOBEC2 en el cromosoma 6, APOBEC3 en el cromosoma 22 y APOBEC4 en el cromosoma 1 (Tabla 3) (Smith *et al.*, 2012; Moris *et al.*, 2014 y Prohaska *et al.*, 2014; Burns *et al.*, 2015).

**Tabla 3.** Miembros de la familia APOBEC

Gen	Localización cromosómica	Longitud	Peso molecular
<i>AID</i>	Cromosoma 12	198 aa	23.954 KDa
<i>APOBEC1</i>	Cromosoma 12	236 aa	28.192 KDa
<i>APOBEC2</i>	Cromosoma 6	224 aa	25.703 KDa
<i>APOBEC3A</i>	Cromosoma 22	199 aa	23.012 KDa
<i>APOBEC3B</i>	Cromosoma 22	382 aa	45.924 KDa
<i>APOBEC3C</i>	Cromosoma 22	190 aa	22.826 KDa
<i>APOBEC3DE</i>	Cromosoma 22	386 aa	46.598 KDa
<i>APOBEC3F</i>	Cromosoma 22	373 aa	45.020 KDa
<i>APOBEC3G</i>	Cromosoma 22	384 aa	46.408 KDa

<i>APOBEC3H</i>	Cromosoma 22	200 aa	23.531 KDa
<i>APOBEC4</i>	Cromosoma 1	367 aa	41.581 KDa

Se muestran los 11 miembros de la familia APOBEC. Para cada miembro se muestra su localización cromosómica, longitud en aminoácidos (aa) y peso molecular en Kilodaltones (KDa).

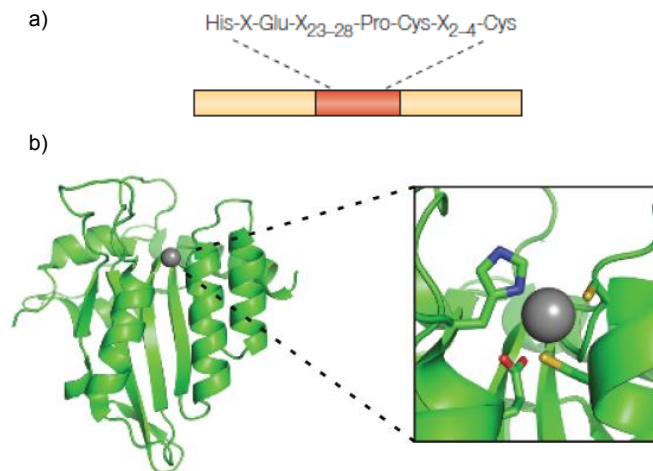


**Figura 5.** Reacción de desaminación llevada a cabo por la familia de enzimas APOBEC y edición de Apolipoproteína B. a) reacción de desaminación característica de la familia de enzimas APOBEC; b) Edición llevada a cabo en el mRNA de la Apolipoproteína B por APOBEC1, generando una isoforma disfuncional de 48 aa. Imagen b) modificada de Lackey, 2012.

Todos los miembros de esta familia dentro de su estructura contienen uno o dos sitios conservados que se denominan dominios de coordinación de zinc (ZCD), cuya secuencia proteica es **His/Cys-X-Glu-X<sub>23-28</sub>-Pro-Cys-X<sub>2-4</sub>-Cys** (Figura 6a). Estos sitios conservados a nivel de proteína tienen una conformación globular que está constituida por una lámina-β hidrofóbica ubicada en el centro del dominio que está rodeada por seis α-hélices, las hélices α-2 y α-3 de los motivos **H-X-E** y **P-C-X<sub>2-4</sub>-C** contienen las secuencias conservadas que definen al sitio activo (Harris y Liddament, 2004; Smith *et al.*, 2012). El dominio de coordinación de zinc se coordina con el átomo debido a los residuos de los aminoácidos cisteína e histidina del sitio activo (Figura 6b) y cuando se da la interacción con el material genético en presencia de una molécula de H<sub>2</sub>O se inicia la reacción de

desaminación de la citosina obteniendo como productos finales U y NH<sub>3</sub> (Figura 5a) (Prohaska *et al.*, 2014; Swanton *et al.*, 2015).

El mecanismo concreto mediante el cual estas proteínas endógenas causan daño al genoma de la célula es una controversia, pero se sugieren algunos mecanismos celulares en los que podrían estar implicadas como en la reparación de rupturas en la doble cadena del DNA, ya que como lo reportan Taylor y colaboradores, al inducir rupturas en el DNA aumenta el número de transiciones T→C (Taylor *et al.*, 2013), sin embargo se hipotetiza que estas proteínas podrían estar implicadas en otros procesos celulares donde el material genético pudiera estar expuesto a la acción de estos peptidos como en la transcripción, sin embargo no se ha confirmado. (Harris y Liddament, 2004; Smith *et al.*, 2012, Prohaska *et al.*, 2014 y Moris *et al.* 2014).

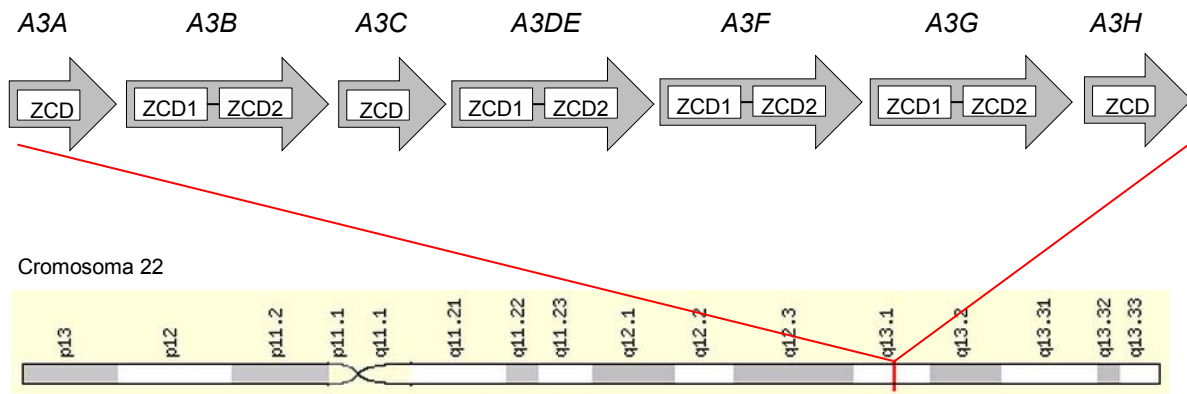


**Figura 6.** Dominio de coordinación de zinc de las proteínas APOBEC. a) Secuencia de aminoácidos que codifica para el sitio activo de las proteínas APOBEC. b) Modelo de la estructura cristalina del dominio de coordinación de zinc de la proteína APOBEC3G. En amarillo y azul el residuo de unión a zinc constituido por Cis-Cis-His, en rojo el ácido glutámico y la esfera gris representa un átomo de zinc. Imágenes tomadas y modificadas de Harris y Liddament, 2004 y Swanton *et al.*, 2015.



### 1.3.1 APOBEC3B

La subfamilia de proteínas APOBEC3 es la más numerosa dentro de la familia APOBEC. Hasta el momento se conocen 7 isoformas que guardan un grado de homología de hasta el 80% entre algunas isoformas y están presentes únicamente en mamíferos. La distribución de estas proteínas en las diferentes especies de mamíferos es variable, sin embargo en la especie *Homo sapiens*, se encuentran a los siete miembros, los cuales se localizan agrupados en *tandem* en el cromosoma 22 (Figura 7) (Dunham *et al.*, 1999; Harris y Liddament, 2004; Conticello *et al.*, 2005; Smith *et al.*, 2012).



**Figura 7.** Locus de los miembros de la subfamilia APOBEC3. Los miembros de la subfamilia A3 se localizan en *tandem* en el brazo largo del cromosoma 22, en el locus q13.1 y su principal característica es que dentro de su estructura contienen uno o dos dominios de coordinación de zinc (ZCD). ZCD1: dominio de coordinación de zinc en el extremo amino terminal; ZCD2: dominio de coordinación de zinc en el extremo carboxilo terminal.

La principal función de los miembros de la subfamilia APOBEC3 es ser la línea de defensa intracelular ante los elementos genéticos móviles, tales como virus, entre los que se pueden mencionar el del VIH-1, VHB, VPH. Las enzimas APOBEC3 también controlan a transposones, como los de tipo L1 pertenecientes a los elementos nucleares intercalados largos (LINEs) y los elementos Alu que pertenecen a los elementos cortos intercalados (SINEs) (Esnault *et al.*, 2005; Koito y Ikeda, 2011; Smith *et al.*, 2012; Don-Marc y Petersen-Mahrt, 2014; Avesson y Barry, 2014; Burns *et al.*, 2015). No obstante, durante los últimos años se sabe que además de esta importante función de defensa, APOBEC podría ser la principal fuente de mutación endógena en cáncer (Alexandrov *et al.* 2013; Burns *et*

*al.*, 2013b). Dentro de las enzimas APOBEC3, se sugiere a la isoforma *APOBEC3B* como el principal exponente en la mutagenesis endógena, ya que ha sido implicada en procesos de mutagenesis en tumores de vejiga, cervix, pulmón, cabeza y cuello, y mama donde su firma mutacional ha sido reportada (Burns *et al.*, 2013b; Roberts *et al.*, 2013).

La proteína APOBEC3B tiene localización estrictamente nuclear (Lackey *et al.*, 2013). En muestras de cáncer de pacientes con tumores de vejiga, cervical, pulmón, cabeza y cuello, y mama, esta enzima muestra un aumento considerable en los niveles de expresión a nivel de mRNA, en comparación con otros miembros de la familia APOBEC (Burns *et al.*, 2013b). También se ha asociado su nivel de expresión a un aumento en los niveles de uracilo genómico y en las transiciones C→T dentro del genoma. No obstante es en el cáncer de mamá donde se ha propuesto a esta enzima como la principal fuente de mutagenesis al mostrar patrones mutacionales repetitivos entre líneas celulares y muestras de pacientes. Estos patrones han sido descritos gracias a estudios realizados con metodologías de secuenciación masiva en donde se observó un aumento de transiciones C→T y transversiones C→G enriquecidas en el trinucleótido 5'-TCW (donde C es la base mutada). Dentro de los tumores con esta firma, también se determinó la presencia de un fenómeno de hipermutación localizada en regiones cortas del genoma ( $\geq 6$  mutaciones en  $\leq 1000$  pb), el cual fue definido como *Kataegis* (del griego tormenta) (Kuong y Loeb, 2013; Alexandrov *et al.* 2013; Harris, 2015; Burns *et al.*, 2013a; Burns *et al.*, 2013b; Roberts *et al.*, 2013; Taylor *et al.*, 2013; Roberts y Gordenin, 2014; Swanton *et al.*, 2015). Además algunos trabajos sugieren que esta enzima podría estar implicada en la inactivación por mutación de genes de importancia para la oncogenesis, por ejemplo Burns y colaboradores en el 2013 (Burns *et al.*, 2013a) proponen a *A3B* como la causante de la inactivación del supresor de tumores *TP53* en el cáncer de mama favoreciendo así “al rápida evolución del tumor”. También Sieuwerts y colaboradores en el 2014 proponen que altos niveles de mRNA de *A3B* podrían contribuir a la progresión del cáncer de mama en tumores RE+, en especial en el subtipo Luminal A (Sieuwerts *et al.*, 2014).

Sin embargo la asociación de *A3B* en otros procesos mutagénicos fundamentales sigue siendo poco clara, por lo que la justificación de esta tesis es la clonación del gen *APOBEC3B* y su posterior caracterización para contar con una herramienta molecular que permita evaluar en estudios posteriores la participación de *A3B* en procesos oncogénicos a nivel molecular.

## 2. OBJETIVOS

### General

- Clonar en un vector de expresión la secuencia codificante completa del el gen *APOBEC3B* fusionada a la secuencia del epítipo HA a partir de líneas celulares humanas de cáncer de mama

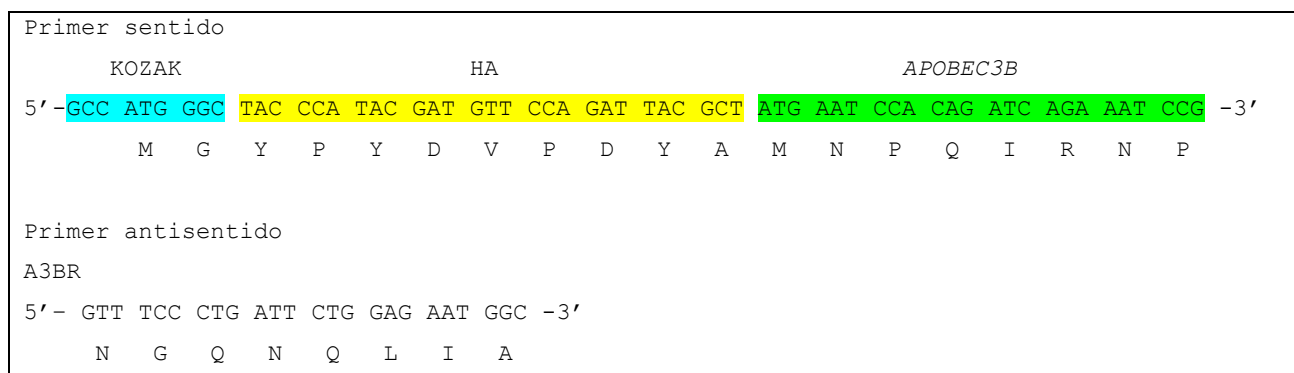
### Particulares

- Diseñar primers dirigidos contra el gen *APOBEC3B* con el epítipo artificial de hemaglutinina (HA) en el extremo N-terminal.
- Amplificar el marco abierto de lectura del gen *A3B* por medio de PCR Toehold.
- Ligar los productos de PCR del marco abierto de lectura del gen *APOBEC3B* en un vector de expresión en células eucariontes.
- Evaluar la orientación de los insertos dentro del vector utilizando los primers T7F y BGHR.
- Analizar el marco abierto de lectura de las clonas mediante secuenciación tipo Sanger.
- Transfectar los plásmidos con el gen *APOBEC3B* en células de cáncer de mama y medir sus niveles de expresión exógenos mediante PCR en tiempo real.

### 3. MATERIALES Y MÉTODOS

#### 3.1 Diseño de primers

Se diseñaron primers para amplificar de manera específica el gen *APOBEC3B* (*A3B*) humano a partir de su secuencia codificante correspondiente a la secuencia del exoma del gen *A3B* isoforma 1 (NM\_004900.4) utilizando el programa Primer3. Para el diseño de los primers se tomaron en cuenta los valores óptimos del porcentaje de GC (40 al 60%) y la  $T_m$  que se propone en un intervalo de 55 a 65°C, para verificar la correspondencia de los primers con el gen *A3B* se utilizó el programa Primer-BLAST. Posterior a ello se utilizaron los programas Oligocalc y PriDimerCheking para cotejar los valores de complementariedad y autocomplementariedad entre primers generados en el análisis por Primer3 y Primer-BLAST. Posteriormente se agregó una etiqueta molecular en el extremo 5' del primer sentido que codifica para el epítipo artificial de hemaglutinina (HA) para su análisis a nivel proteína en estudios posteriores y se introdujo una secuencia Kozak canónica en el sitio de inicio de la traducción del epítipo HA para aumentar su tasa de traducción (Figura 8).



**Figura 8.** Primers utilizados para la clonación del gen *A3B*. Primers sentido y antisentido. Se muestra el diseño de los primers utilizados para la clonación del gen *A3B*. El primer sentido se diseñó con la secuencia Kozak y la etiqueta molecular que codifica para la hemaglutinina.

#### 3.2 Extracción de RNA total y síntesis de cDNA

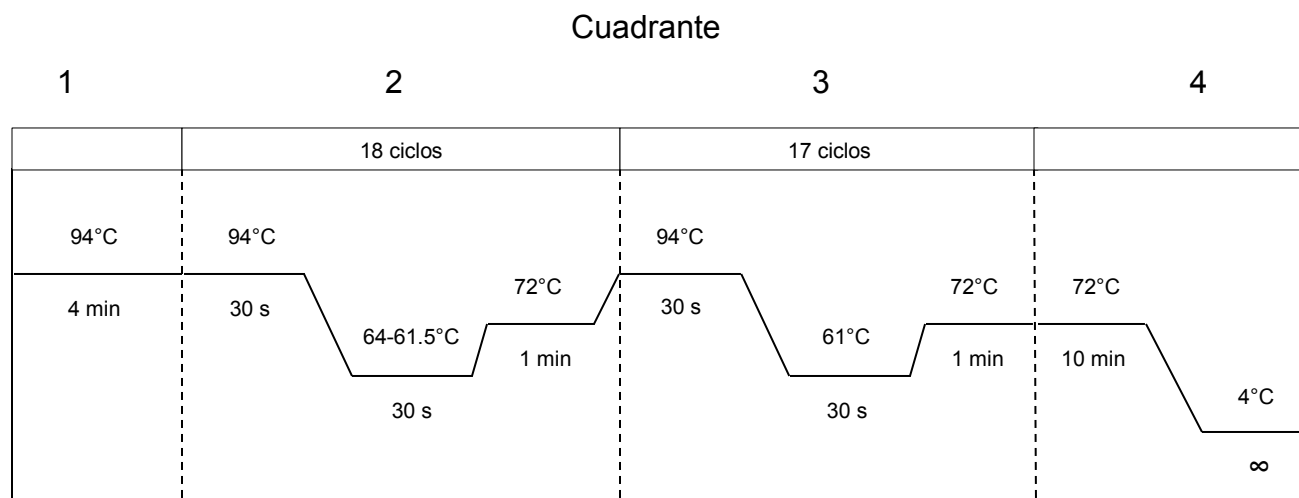
Se realizaron extracciones de RNA total a partir de  $1 \times 10^6$  células de las líneas MDA-MB-231 y MCF-7 utilizando 1 mL de TRIzol® Reagent (Life Technologies).

Después del lisado celular se agregó cloroformo y se centrifugó a 15,000 g por 15 minutos, se extrajo la fase acuosa y se le añadió isopropanol. Posteriormente se centrifugó a 12,000 g por 15 minutos y se removió el sobrenadante, se lavó la pastilla con etanol al 75 % y se resuspendió la pastilla en H<sub>2</sub>O libre de nucleasas. La síntesis de cDNA se llevó a cabo con el mix SuperScript® Enzyme Mix (Invitrogen), la mezcla se realizó agregando 1x de Reaction Mix (RNaseOUT™, random primers, MgCl<sub>2</sub> (1.5 mM) y dNTPs (200 μM)), 1x SuperScript® Enzyme (1U/μL retrotranscriptasa), H<sub>2</sub>O desionizada y 500 ng de RNA a un volumen final de 20 μL, el ciclaje para la síntesis de cDNA fue de 37°C durante 60 minutos y 95°C por 5 minutos. Las preparaciones de cDNA se almacenaron a -20°C.

### **3.3 PCR-Touchdown**

Una vez que se obtuvo el cDNA se llevó a cabo la reacción de PCR-Touchdown (TDW), cuyo fundamento se basa en utilizar temperaturas de alineamiento superiores a la temperatura de desnaturalización (T<sub>m</sub>) de los primers durante los primeros ciclos de PCR para aumentar la especificidad de la reacción y progresivamente va disminuyendo hasta llegar a la temperatura calculada para la hibridación de los oligos. Se utilizaron los primers específicos para *A3B*, el primer sentido fue 5'-GCC ATG GGC TAC CCA TAC GAT GTT CCA GAT TAC GCT ATG AAT CCA CAG ATC AGA AAT CCG-3' y la secuencia del antisentido fue 5'-GTT TCC CTG ATT CTG GAG AAT GGC-3'. Las condiciones de la reacción fueron 500 ng de cDNA, 0.4 μM de primers, 1.5 mM de MgCl<sub>2</sub>, 0.4 mM de dNTPs y 1 u/μL de Taq polimerasa con su respectivo buffer. La PCR-TDW inició con una fase de desnaturalización a 94°C por 5 minutos. A continuación se llevaron a cabo 35 ciclos con una temperatura de desnaturalización de 94°C por 30 segundos, el ciclo PCR-TDW se inició a 64°C y cada tres ciclos disminuyó 0.5°C hasta llegar a 61°C, todo por 30 segundos durante 18 ciclos (Figura 9 cuadrante 2). Posteriormente se llevaron a cabo 17 ciclos de alineamiento a 61°C, cada ciclo con una temperatura de elongación de 72° por 1 minuto (Figura 9 cuadrante 3).

Los productos de PCR fueron observados en geles de agarosa al 1.5% teñidos con bromuro de etidio (BrEt). Los productos de PCR se purificaron utilizando el Kit Gel/PCR Extraction Kit (Fast Gene™).



**Figura 9.** PCR Touchdown. Se representan las condiciones de temperatura y tiempo empleados en la PCR Touchdown (TDW). En el segundo cuadrante de la imagen se muestra el programa de PCR TDW utilizado para la amplificación de *APOBEC3B*. La PCR TDW se llevó a cabo con la disminución progresiva de la Tm cada 3 ciclos. El programa fue el siguiente: 64°C (ciclos 1-3), 63.5°C (ciclos 4-6), 63°C (ciclos 7-9), 62.5°C (ciclos 10-12), 62°C (ciclos 13-15) y 61.5°C (ciclos 16-18). Finalmente, se realizaron 17 ciclos de PCR a la temperatura de alineamiento 61°C (tercer cuadrante) y una elongación final de 10 minutos a 72°C (cuarto cuadrante).

### 3.4 Clonación de A3B

La clonación se realizó utilizando el vector pcDNA™3.1/V5-His TOPO® TA Expression Kit (Invitrogen), que contiene la secuencia del promotor de citomegalovirus (CMV), para aumentar la eficiencia de la transcripción en mamíferos. Este vector proporciona la ventaja de llevar a cabo la inserción del gen mediante un producto de PCR ya que sus extremos 5' son ricos en residuos de timina, y debido a que la Taq polimerasa tiene actividad de adenilil transferasa terminal en los extremos 3' se facilita la ligación, sin embargo la energía para formar el enlace fosfodiéster entre el vector y el producto de PCR la proporciona el residuo tirosil de la tirosina 274 de la topoisomerasa I del virus *Vaccinia* asociado

en el extremo 3' del vector. La reacción de ligación se llevó a cabo utilizando 30 ng del vector TOPO, 0.7  $\mu$ L de H<sub>2</sub>O esteril, 1  $\mu$ L de solución salina (1.2 M NaCl y 0.06 M de MgCl<sub>2</sub>) y 4  $\mu$ L de producto de PCR purificado. Se transformaron bacterias químicamente competentes provenientes de la cepa de *Escherichia coli* DH5 $\alpha$  añadiendo 2  $\mu$ L del producto de la ligación y el control positivo, el plásmido PUC19, a 25  $\mu$ L de *E. coli* DH5 $\alpha$ . Las células fueron transformadas mediante un shock térmico de 30 segundos a 42°C. Posteriormente se incubó las bacterias en medio SOC durante 60 minutos a 37°C e inmediatamente después se inocularon en cajas Petri con medio LB y 50  $\mu$ g de ampicilina como agente de selección durante 16 horas a 37°C.

### 3.5 Tamizaje de colonias

Debido a que la clonación antes mencionada no es direccional, se verificó la orientación correcta de *A3B* por medio de PCR de colonia (Woodman, 2008). Se llevó a cabo PCR a partir de las colonias obtenidas. Para ello se utilizaron los primers: T7F 5'-TAA TAC GAC TCA CTA TAG GG-3', BGHR 5'-TAG AAG GCA CAG TCG AGG-3' y A3BR 5'-GTT TCC CTG ATT CTG GAG AAT GGC-3'. Se realizaron las combinaciones T7F-BGHR y T7F-A3BR para generar productos que sirvieran como indicadores de la presencia del inserto (T7F-BGHR) la ligación en dirección correcta (T7F-A3BR) y de la inserción incorrecta (A3BR-BGHR). El programa de la PCR fue: 94°C por 30 segundos, 61°C por 30 segundos y 72°C por 1 minuto durante 35 ciclos y se le realizó electroforesis de agarosa con bromuro de etidio para corroborar la orientación correcta del marco abierto de lectura de los productos amplificados. De las colonias que dieron positivo para orientación correcta se realizaron preparaciones puras de plásmidos utilizando el kit GeneJet Plasmid Miniprep Kit (Thermo Scientific) y se secuenciaron por el método de Sanger, el cual tiene como principio la adición de una mezcla de desoxinucleótidos trifosfato (dNTPs), didesoxinucleótidos trifosfato (ddNTPs) y Taq Polimerasa. Los ddNTPs en su extremo 3' carecen de grupo OH, lo que provoca una terminación prematura de la elongación de la cadena por parte de la polimerasa, generando distintos tamaños de productos, los cuales tienen un

fluoróforo terminal 3' correspondiente a la última base añadida. El orden de los fragmentos y su marcaje con el fluoróforo terminal permite generar la secuencia de nucleótidos (Sanger *et al.*, 1977). La secuenciación se llevó a cabo en Laboratorio de Secuenciación Genómica de la Biodiversidad y de la Salud ubicado en el Instituto de Biología de la UNAM a cargo de la M.C. Laura Margarita Márquez Valdelamar.

### **3.6 Análisis de datos de secuenciación**

Una vez que se obtuvieron los datos de la secuenciación de las clonas C3 y C5 se llevó a cabo un análisis utilizando el programa A Plasmid Editor (ApE) que permite alinear la secuencia obtenida contra la secuencia esperada para poder observar si se generan cambios en los nucleótidos, además de que muestra el cromatograma correspondiente de la secuenciación permitiendo evaluar la calidad de la secuenciación. Debido a la presencia de mutaciones en ambas clonas se realizó un Blast en el programa Blastx del Centro Nacional para Información Biotecnológica (NCBI) para generar la secuencia de aminoácidos y calcular así el porcentaje de identidad que guarda la secuencia con otras proteínas a nivel de aminoácidos.

De las sustituciones de aminoácidos que se generaron por los cambios en los nucleótidos se analizó la repercusión funcional putativa con las herramientas computacionales SIFT y PolyPhen-2. SIFT (SIFT Human Protein) es un programa desarrollado por el Instituto Craig Venter<sup>TM</sup> que analiza únicamente la secuencia de aminoácidos de la proteína y predice el grado de tolerancia que puede tener ante la sustitución de algún aminoácido dependiendo del sitio donde se lleve a cabo el cambio y la naturaleza química del aminoácido que sustituye, por lo que puede decir si un cambio es tolerable (si la naturaleza química del aminoácido que sustituye es similar o en anteriores estudios no se ha observado que afecte la función de la proteína) o dañina (si la mutación se encuentra en un sitio activo, región conservada o predice que el aminoácido que sustituye puede alterar la conformación y la función por su naturaleza química) (Henikoff *et al.*, 2003; Hicks *et al.*, 2011). En contraste, PolyPhen-2 es una herramienta que mediante su



algoritmo computacional predice el posible impacto en sustituciones de aminoácido y el efecto que tendrá en la estructura y función de una proteína humana. El análisis se hace a partir de consideraciones como la secuencia, estructura y relaciones filogenéticas de la proteína humana, usando consideraciones físicas y comparativas directas de secuencias obtenidas de bases de datos de proteínas como UniProt y NCBI que son analizadas mediante algoritmos matemáticos y computacionales de programas como MAFFT, LEON, DAAP y MULTIZ, que en conjunto predicen que las variantes pueden ser benignas (muy probablemente la sustitución no tiene un efecto fenotípico), posiblemente dañina (predice que posiblemente la sustitución afecta la estructura o función de la proteína) o probablemente dañina (con alto porcentaje de confianza en la predicción, la sustitución afecta la estructura o función de la proteína). Para la predicción el programa utiliza dos parámetros diferentes que son HumDiv (Predictor probabilístico que evalúa alelos raros mediante el análisis de estudios relacionados con todos los genomas y análisis de selección natural) y HumVar (Predictor diagnóstico que valora el efecto de las mutaciones en enfermedades humanas) (Adzhubei *et al.*, 2010; Hicks *et al.*, 2011; Adzhubei *et al.*, 2013).

### **3.7 Transfección de clonas en línea celular MDA-MB-231**

Se utilizó la línea celular MDA-MB-231 de cáncer de mama cultivada en medio DMEM suplementado al 10% con suero fetal bovino (FBS) a condiciones de 37°C y 5% de CO<sub>2</sub>. Para la transfección se sembraron 125,000 células por pozo en placas de 24 con la finalidad de tener una confluencia del 70% al momento de la transfección. Las células fueron divididas en 5 grupos, cada uno por triplicado, los grupos fueron: A3B-C3, A3B-C5, PIREs2 (control positivo de transfección), control lipofectamina y control sin tratamiento y se les añadió una mezcla que contenía 2 µL de Lipofectamina<sup>2000</sup> (Invitrogen) y 600 ng de los plásmidos (A3B-C3, A3B-C5 y PIREs2) en medio OPTI-Mem en un volumen final de 250 µL por pozo, posteriormente se incubaron durante 6 horas a 37°C y 5% de CO<sub>2</sub>. Después de 6 h se sustituyó la mezcla de transfección por medio DMEM suplementado con 10% de FBS. 24 h después de la transfección las células fueron observadas al

microscopio de fluorescencia de la Unidad de Biomedicina de la FES-Iztacala para verificar que se hubiera llevado a cabo la transfección, usando como gen reportero la proteína verde fluorescente (eGFP) clonada en el vector PIREs2. Al corroborar que la transfección fue positiva se prosiguió con la extracción de RNA y se realizó una mezcla de 300 ng de RNA de cada una de las repeticiones de los tratamientos y se les sometió a un tratamiento con DNAsa (Thermo Scientific) como lo indica el fabricante, se prosiguió con la retrotranscripción como se ha descrito anteriormente.

### **3.8 PCR en tiempo real**

Para realizar la PCR en tiempo real (qPCR) se utilizaron los primers específicos para cuantificar los niveles de expresión endógena de *A3B*, el primer sentido fue 5'-GTG TAC AAT GAA GGT CAG C-3' y la secuencia del antisentido fue 5'-AGC ACA AGT AGG TCT GG-3'. Estos primers alínean con las regiones ubicadas en los exones 4 y 5 (respectivamente) del ORF de *APOBEC3B*. Para normalizar los niveles de expresión se utilizaron los primers correspondientes a los genes de mantenimiento *GAPDH* (sentido 5'-CCT TGG AGA AGG CTG GGG-3'; antisentido 5'-CAA AGT TGT CAT GGA TGA CC-3') y  $\beta$ -Actina (sentido 5'-AGA GCT ACG AGC TGC CTG AC-3'; antisentido 5'-AGC ACT GTG TTG GCG TAC AG-3'). La reacción de PCR en tiempo real se llevó a cabo utilizando 2.5 ng de cDNA, 10  $\mu$ L del Mix 2X Brilliant III Ultra-Fast SYBR<sup>®</sup> Green QPCR que contiene 1.5 mM de  $MgCl_2$ , 0.4 mM de dNTPs y 1 u/ $\mu$ L de polimerasa,  $H_2O$  y 0.4  $\mu$ M de primers de *A3B*, *GAPDH* y  $\beta$ -Actina. Para cada cDNA se le realizó qPCR por triplicado de *A3B* y los dos genes house keeping con su respectivo blanco, las condiciones de ciclaje fueron 95°C por 30 s, 55°C por 30 s y 72°C por 20 s durante 35 ciclos. La reacción se llevó a cabo en el termociclador Rotor-Gene Q (Quiagen) del Laboratorio de Análisis Clínicos de la CUSI de la FES-Iztacala a cargo de la Dra. Gloria Luz Paniagua Contreras. Los datos se analizaron mediante el cálculo de la  $\Delta\Delta Ct$ , la cual compara la expresión de los niveles de mRNA de un gen de interés en muestras con y sin tratamiento, donde cada muestra es comparada con un

control interno (Livak *et al.*, 2011). Los datos fueron analizados con una prueba de ANOVA de dos factores con una  $\alpha=0.05$ .

## 4 RESULTADOS

### 4.1 Diseño de primers dirigidos contra el ORF de A3B

Los primers diseñados para hibridar en el ORF del gen *A3B* se diseñaron con el programa Primer3 (Untergasser *et al.*, 2012). El primer sentido tuvo valores de 41.67% de GC y una  $T_m$  de 58.94, para el primer antisentido los valores fueron 50% de GC y una  $T_m$  de 61.70°C (Tabla 4). La correspondencia de los oligos con el gen *A3B* se analizó mediante el programa Primer-BLAST y los datos arrojados fueron que las secuencias diseñadas corresponden en su totalidad con las dos variantes del gen *A3B* (Tabla 5). Los valores de autocomplementariedad generados en los análisis anteriores se reanalizaron con el programa Oligocalc y según los datos proporcionados, las secuencias del primer sentido son lo suficientemente complementarias para que se dé la formación de tres dímeros, pero no para causar la formación de una horquilla (Tabla 6), mientras que para el caso del primer antisentido también es lo suficientemente complementaria para generar dos juegos de dímeros (Tabla 6), pero no una horquilla. Para el análisis de la posible complementariedad entre los primers sentido y antisentido se realizó un análisis con el programa PriDimerCheking el cual arrojó que no se da hibridación entre los primers (no se muestra).

**Tabla 4.** Primers de *A3B* diseñados con el programa Primer3

	Secuencia (5'→3')	Longitud	$T_m$	%GC	Auto complementariedad	3' Autocomplementariedad
<b>Primer Sentido</b>	ATG AAT CCA CAG ATC AGA AAT CCG	24	58.94	41.67	4.00	2.00
<b>Primer Antisentido</b>	GTT TCC CTG ATT CTG GAG AAT GGC	24	61.70	50.00	8.00	2.00

Los valores que proporciona el programa son la temperatura de alineamiento (óptimo 55-65°C), el % de GC (óptimo de 40 a 60%), autocomplementariedad entre cada primer y autocomplementariedad en el extremo 3' entre cada primer.

**Tabla 5.** Primer-BLAST de oligonucleótidos dirigidos contra el gen *A3B*

> <a href="#">NM_004900.4</a> Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B (APOBEC3B), transcript variant 1, mRNA			
Tamaño del producto: 1146			
Primer Sentido	1	ATGAATCCACAGATCAGAAATCCG	24
Templado	56	.....	79
Primer Antisentido	1	GTTTCCCTGATTCTGGAGAATGGC	24
Templado	1201	.....	1178
> <a href="#">NM_001270411.1</a> Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B (APOBEC3B), transcript variant 2, mRNA			
Tamaño del producto: 1071			
Primer Sentido	1	ATGAATCCACAGATCAGAAATCCG	24
Templado	56	.....	79
Primer Antisentido	1	GTTTCCCTGATTCTGGAGAATGGC	24
Templado	1126	.....	1103

El tamaño del producto para la variante 1 es de 1146pb y para la variante 2 es de 1071pb.

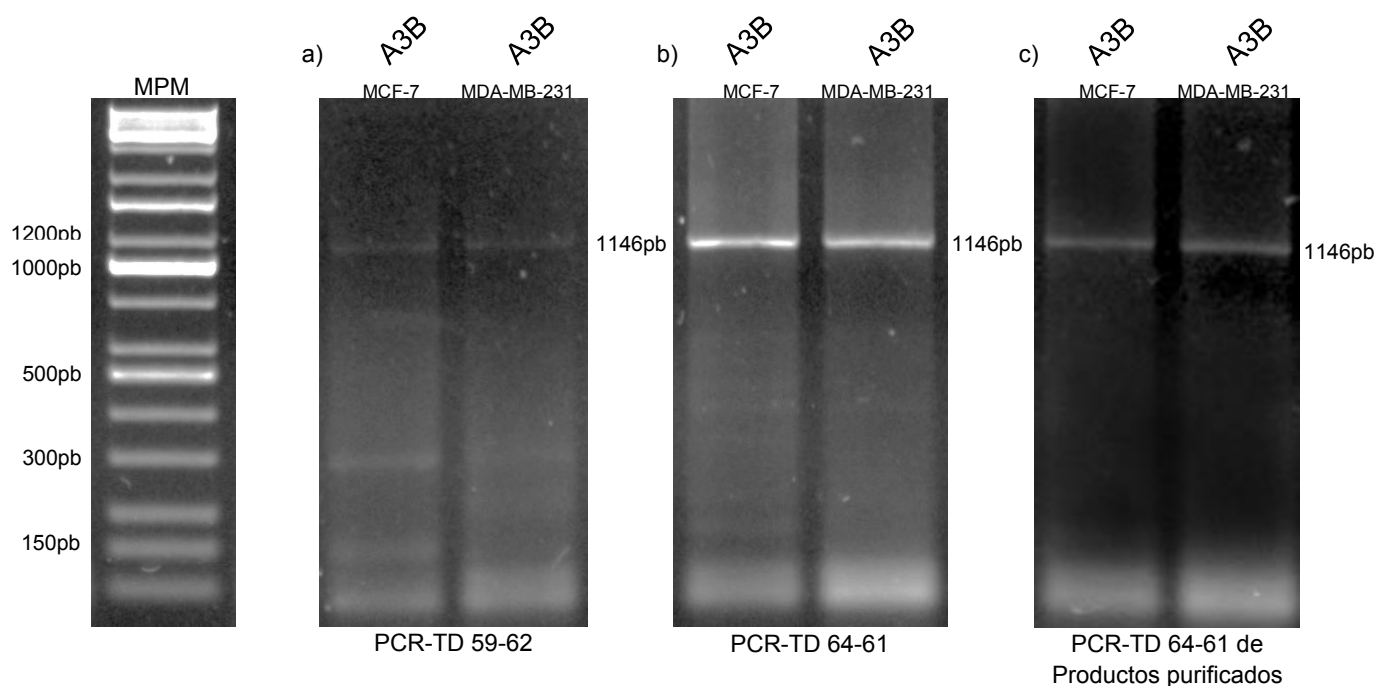
**Tabla 6.** Análisis de formación de dímeros mediante programa Oligocalc para los primers sentido y antisentido

Primer	Dímero	Tm
Sentido	5' X <sub>36</sub> ATGAATCCACAGATCAGAAATCCG 3' 3' GCCTAAAGACTAGACACCTAAGTA <sub>36</sub> X 5'	24°C
Sentido	5' GCCATGGGCTACCCATACGATGTT <sub>36</sub> X 3' 3' X <sub>36</sub> TTGTAGCATACCCATCGGGTACCG 5'	25°C
Sentido	5' GCCATGGGCTACCCATACGATGTT <sub>36</sub> X 3' 3' X <sub>36</sub> TTGTAGCATACCCATCGGGTACCG 5'	25°C
Antisentido	5' GTTTCCTGATTCTGGAGAATGGC 3' 3' CGGTAAGACCTTTAGTCCCTTTG 5'	24°C
Antisentido	5' GTTTCCTGATTCTGGAGAATGGC 3' 3' CGGTAAGACCTTTAGTCCCTTTG 5'	24°C

#### 4.2 PCR Touchdown de líneas celulares MDA-MB-231 y MCF-7

Se utilizó en primera instancia una PCR-Touchdown, empleando las temperaturas de alineamiento de primers 62 a 59°C. Sin embargo con estas condiciones se obtuvieron productos inespecíficos, adicionales al esperado de *A3B* de 1107pb (Figura 10a). Se aumentaron las temperaturas de alineamiento de primers, abarcando 64 a 61°C para favorecer la hibridación específica de los primers. La cantidad de productos inespecíficos disminuyó, y el rendimiento del producto del tamaño esperado aumentó considerablemente (Fig. 10b). Sin embargo la

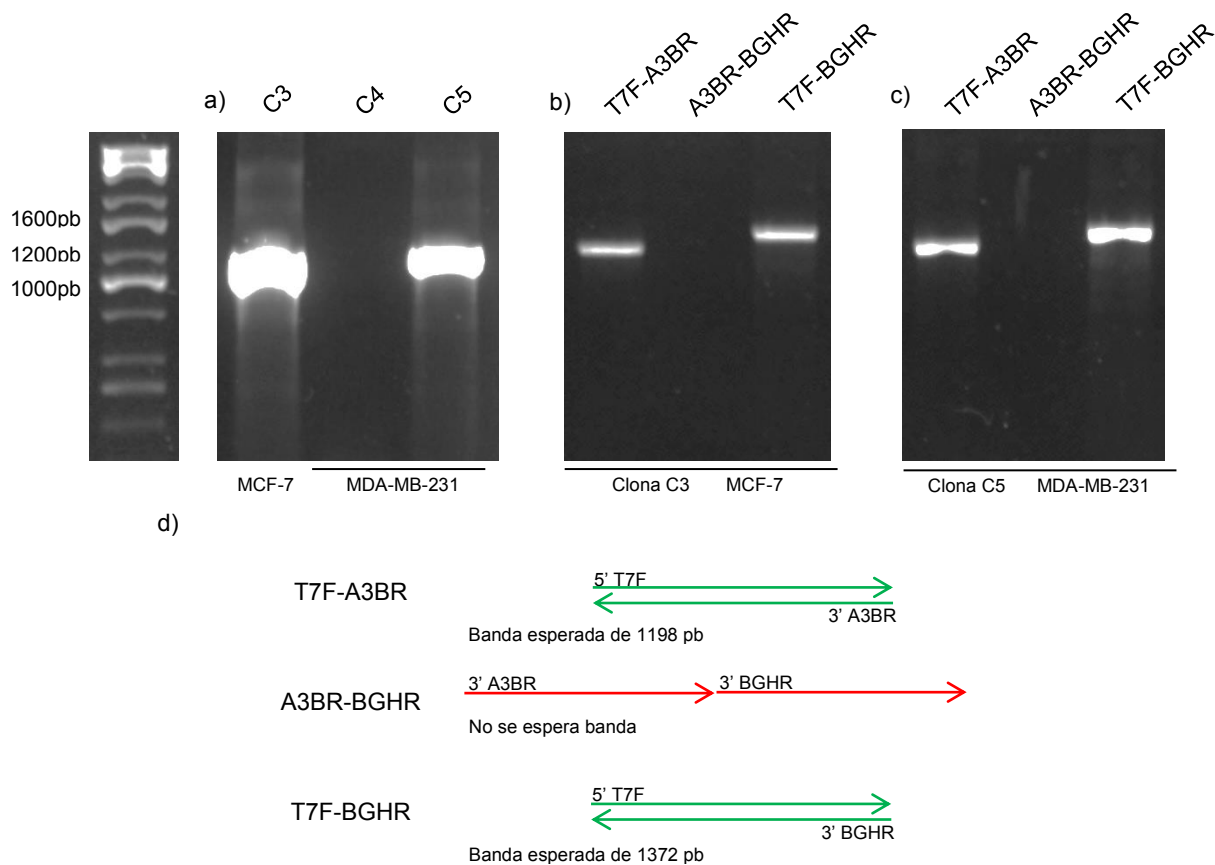
presencia de productos inespecíficos en una reacción de clonación interfiere en el proceso de inserción de la secuencia única esperada dentro del vector y la eficiencia de la clonación disminuye, por ello los productos que se obtuvieron de la PCR-TDW 64 a 61 fueron purificados y se les realizó una reamplificación volviendo a utilizar el programa de PCR TDW de 64 a 62°C, como se puede observar en la Figura 10c se tienen los productos de 1146 pb correspondientes a la secuencia esperada del gen *A3B*, sin embargo en todas las reacciones de PCR se observaron productos que corresponden a dímeros de primer menores a 150 pb. También se observaron dímeros de primers con tamaños menores a 150 pb.



**Figura 10.** Productos de la amplificación de *APOBEC3B* por medio de PCR Touchdown a partir de cDNA extraído de las líneas celulares MCF-7 y MDA-MB-231, así como reamplificación de productos purificados analizados en geles de agarosa al 1.5% teñidos con bromuro de etidio. a) Amplificación por PCR-TDW de 62 a 59°C. Para ambas líneas celulares se observa muy tenue la banda de 1146pb esperada para *A3B* y aunado a ello la presencia de productos inespecíficos. b) Aumento en las temperaturas de la PCR-TDW (64 a 62 °C), la cantidad de producto de *A3B* aumentó considerablemente, sin embargo aún se observa la presencia de productos inespecíficos y dímeros de primers. c) Reamplificación de los productos purificados de la PCR-TDW 64 a 61°C, se observa la banda de 1146 pb de *A3B* para ambas líneas celulares y los productos inespecíficos eliminados en su totalidad. Se observan dímeros de primers.

### 4.3 Tamizaje de colonias

Los productos purificados se ligaron en el vector pcDNA™3.1/V5-His TOPO® y con las ligaciones se transformaron bacterias competentes *E. coli* DH5α. La transformación con el vector mencionado les confiere resistencia a la ampicilina por lo que se inocularon en medio LB con el antibiótico. Las bacterias se incubaron por 16 horas y se obtuvieron 3 colonias de bacterias (denominadas clona 3, clona 4 y clona 5) a las cuales se les realizó una PCR de colonia utilizando el primer antisentido de *A3B* con el primer sentido T7 que flanquea en la región 5' del plásmido para verificar la orientación de los insertos. Los productos de PCR de las clonas 3 (MCF-7) y 5 (MDA-MB-231) son consistentes con la orientación correcta del inserto ya que tienen el tamaño esperado de 1198 pb, mientras que la clona 4 (MDA-MB-231) no generó ninguna banda (figura 11a). Para corroborar estos datos se llevó a cabo una nueva PCR de colonia para las colonias 3 y 5 para confirmar la dirección de los insertos, se usaron los siguientes juegos de primers *A3BR*-T7F (presencia de producto de PCR clonado), T7F-BGHR (amplificación si la orientación es correcta) y *A3BR*-BGHR (amplificación si la orientación es incorrecta) (Figura 11). Se observaron productos en las reacciones con los juegos de oligos T7F-*A3BR* y T7F-BGHR, obteniéndose bandas de aproximadamente 1198 pb y 1372 pb (respectivamente), consistentes con el tamaño esperado del inserto de *A3B* y la clonación en la orientación correcta. Las reacciones con primers *A3BR*-BGHR fueron negativas, confirmando la orientación correcta de los insertos (Figura 11bc). Se procedió con la secuenciación de las clonas C3 y C5.



**Figura 11.** Tamizaje de colonias obtenidas de la transformación con el vector pcDNA™3.1/V5-His TOPO® TA-A3B. a) PCR de colonia. Gel de agarosa al 1.5% teñido con bromuro de etidio. Las colonias C3 y C5 muestran la una banda de 1198 pb consistente con el tamaño esperado del inserto de A3B, mientras que el inserto de la colonia C4 es de alrededor de 100-150 pb. b) y c) Electroforesis en gel de agarosa al 1.5% con bromuro de etidio de PCR de colonia para las clonas C3 y C5. Se obtuvieron las bandas esperadas de 1198 y 1372 pb (carriles 1 y 3) indicativo de la correcta inserción de A3B. En el carril 2 no se generó banda. d) Modelo teórico de amplificación de los productos durante la PCR de colonia. Los pares de primers T7F-A3BR y T7F-BGHR se emplearon para verificar la presencia de inserto y la ligación en dirección correcta, respectivamente. El par de de primers A3BR-BGHR sólo amplifica productos clonados en dirección incorrecta; en las clonas con la dirección correcta no generan banda, ya que ambos elongan en la misma dirección. Se muestra la cantidad de bases para cada producto de PCR.

#### 4.4 Análisis de secuencia obtenida por secuenciación Sanger para las clonas 3 y 5

Se ha descrito que *APOBEC3B* presenta dos isoformas. En este trabajo se aisló solo la isoforma 1, la cual es la isoforma funcional más estudiada del gen A3B (Figura 12) (Burns *et al.*, 2013b; Forte, 2013; Kuong y Loeb, 2013; McDougle *et al.*, 2013). Los datos obtenidos de la secuenciación Sanger muestran que la clona C3 obtenida de la línea celular MCF-7 contiene 5 sustituciones de nucleótidos,

mientras que la colonia C5, cuyo DNA se obtuvo de la línea celular MDA-MB-231, presentó un total de 6 mutaciones (Figura 13). Para verificar el porcentaje de identidad de las clonas con la isoforma funcional del gen *APOBEC3B*, la cual consta de 1146 nucleótidos y 382 aminoácidos, se llevó a cabo un Blastx y se obtuvo que el porcentaje de identidad de ambas clonas con la isoforma 1 de *APOBEC3B* es del 99% (No se muestra). La clona C3 tiene una identidad compartida con la isoforma funcional de 378/382 aminoácidos, ya que presentó 4 sustituciones: K146T, R252H, C284R y D360N (Figura 14). La clona C5 guarda una identidad con *A3B* de 379/382, ya que presentó 3 sustituciones correspondientes a: K62E, V220A y D328G (Figura 15).

Para predecir el posible impacto de las sustituciones de aminoácidos presentes en ambas clonas, se llevó a cabo un análisis utilizando los algoritmos bioinformáticos SIFT y PolyPhen-2 (ver Materiales y Métodos). Para la clona C3, las sustituciones K146T y D360N, se clasificaron como cambios tolerados/benignos (SIFT/PolyPhen-2), para la posición R252H la predicción fue dada como tolerable/benigna, no obstante HumDiv lo predice como probablemente dañina y por último las herramientas computacionales predicen que la sustitución en la posición C284R tiene una alta probabilidad de ser dañina/probablemente dañina (Tabla 7). Para la clona C5 las sustituciones en posición K62E y D328G se predicen como tolerables/benignas, sin embargo para la sustitución ubicada en la posición V220A se predijo como dañina/posiblemente dañina (Tabla 8).



**Kozak** HA APOBEC3B

GCCATGGGCTACCCATACGATGTTCCAGATTACGCTATGAATCCACAGATCAGAAATCCGATGGAGCGGATGT  
 -M--G--Y--P--Y--D--V--P--D--Y--A--M--N--P--Q--I--R--N--P--M--E--R--M--  
 ATCGAGACACATTCTACGACAACTTTGAAAACGAACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTA  
 Y--R--D--T--F--Y--D--N--F--E--N--E--P--I--L--Y--G--R--S--Y--T--W--L--C--Y  
 TGAAGTAAAAATAAGAGGGGCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTCGAGGCCAGGTGTATTC  
 --E--V--K--I--K--R--G--R--S--N--L--L--W--D--T--G--V--F--R--G--Q--V--Y--F--  
 AAGCCTCAGTACCACGCAGAAATGTGCTTCTCTTGGTTCTGTGGCAACCAGCTGCCTGCTTACAAGTGT  
 -K--P--Q--Y--H--A--E--M--C--F--L--S--W--F--C--G--N--Q--L--P--A--Y--K--C--  
 TCCAGATCACCTGGTTTGTATCCTGGACCCCTGCCCGGACTGTGTGGGAAGCTGGCCGAATTCCTGTCTGA  
 F--Q--I--T--W--F--V--S--W--T--P--C--P--D--C--V--A--K--L--A--E--F--L--S--E  
 GCACCCCAATGTCACCCTGACCATCTCTGCCGCCCGCCTTACTACTACTGGGAAAGAGATTACCGAAGGGCG  
 --H--P--N--V--T--L--T--I--S--A--A--R--L--Y--Y--Y--W--E--R--D--Y--R--A--  
 CTCTGCAGGCTGAGTCAGGCAGGAGCCCGCTGAAGATCATGGACTATGAAGAATTTGCATACTGCTGGGAAA  
 -L--C--R--L--S--Q--A--G--A--R--V--K--I--M--D--Y--E--E--F--A--Y--C--W--E--  
 ACTTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACAAATTCGATGAAAATTATGCATTCTGCACCG  
 N--F--V--Y--N--E--G--Q--Q--F--M--P--W--Y--K--F--D--E--N--Y--A--F--L--H--R  
 CACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTCACTTTCAACTTTAATAATGACCCTTG  
 --T--L--K--E--I--L--R--Y--L--M--D--P--D--T--F--T--F--N--F--N--N--D--P--L--  
 GTCCTTCGACGGCCAGACCTACTTGTGCTATGAGGTGGAGCGCCTGGACAATGGCACCTGGGTCTGATGG  
 -V--L--R--R--R--Q--T--Y--L--C--Y--E--V--E--R--L--D--N--G--T--W--V--L--M--  
 ACCAGCACATGGGCTTTCTATGCAACGAGGCTAAGAATCTTCTGTGGCTTTTACGGCCGCATGCGGAGCT  
 D--Q--H--M--G--F--L--C--N--E--A--K--N--L--L--C--G--F--Y--G--R--H--A--E--L  
 GCGCTTCTTGGACCTGGTTCTTCTTTGCAGTTGGACCCGCCAGATCTACAGGGTCACCTGGTTTCATCTCC  
 --R--F--L--D--L--V--P--S--L--Q--L--D--P--A--Q--I--Y--R--V--T--W--F--I--S--  
 TGGAGCCCTGCTTCTCCTGGGGCTGTGCCGGGAAGTGCCTGCGTTCTTCAGGAGAACACACACGTGAGAC  
 -W--S--P--C--F--S--W--G--C--A--G--E--V--R--A--F--L--Q--E--N--T--H--V--R--  
 TGCGCATCTTCGCTGCCCGCATCTATGATTACGACCCCTATATAAGGAGGCGCTGCAAATGCTGCGGGATGC  
 L--R--I--F--A--A--R--I--Y--D--Y--D--P--L--Y--K--E--A--L--Q--M--L--R--D--A  
 TGGGGCCCAAGTCTCCATCATGACCTACGATGAGTTTGTGACTGCTGGGACACCTTTGTGTACCGCCAGGGA  
 --G--A--Q--V--S--I--M--T--Y--D--E--F--E--Y--C--W--D--T--F--V--Y--R--Q--G--  
 TGTCCCTTCCAGCCCTGGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGG**GCCATTCTCC**  
 -C--P--F--Q--P--W--D--G--L--E--E--H--S--Q--A--L--S--G--R--L--R--A--I--L--  
**AGAATCAGGGAAC**TGA  
 Q--N--Q--G--N--\*-

**Figura 12.** Secuencias nucleotídica y de aminoácidos de la isoforma 1 del gen *APOBEC3B*. Negro: secuencia nucleotídica; Rojo: secuencia aminoacidica; Gris y amarillo: dominio de coordinación de zinc; Amarillo: sitio activo; Verde: primers diseñados para clonación; Cursivas: epitopo HA.

A3B WT	<u>GCCATGGGC</u> <u>TACCCATACGATGTTCCAGATTACGCTATG</u> <u>AATCCACAGATCAGAAATCCGATGGAGCGGATGTATCGAG</u>	>79
A3B C3	GCCATGGGCTACCCATACGATGTTCCAGATTACGCTATGAATCCACAGATCAGAAATCCGATGGAGCGGATGTATCGAG	
A3B C5	GCCATGGGCTACCCATACGATGTTCCAGATTACGCTATGAATCCACAGATCAGAAATCCGATGGAGCGGATGTATCGAG	
A3B WT	<u>ACACATTCTACGACAACCTTTGAAAACGAACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTATGAAGTAAAAAT</u>	>158
A3B C3	ACACATTCTACGACAACCTTTGAAAACGAACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTATGAAGTAAAAAT	
A3B C5	ACACATTCTACGACAACCTTTGAAAACGAACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTATGAAGTAAAAAT	
A3B WT	<u>AAAGAGGGGCCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTCGAGGCCAGGTGTATTTCAAGCCTCAGTACCACGCA</u>	>237
A3B C3	AAAGAGGGGCCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTCGAGGCCAGGTGTATTTCAAGCCTCAGTACCACGCA	
A3B C5	AAAGAGGGGCCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTCGAGGCCAGGTGTATTTCAAGCCTCAGTACCACGCA	
A3B WT	<u>GAAATGTGCTTCTCTCTTGGTCTGTGGCAACCAGCTGCCTGCTTACAAGTGTTCAGATCACCTGGTTGTATCCT</u>	>316
A3B C3	GAAATGTGCTTCTCTCTTGGTCTGTGGCAACCAGCTGCCTGCTTACAAGTGTTCAGATCACCTGGTTGTATCCT	
A3B C5	GAAATGTGCTTCTCTCTTGGTCTGTGGCAACCAGCTGCCTGCTTACAAGTGTTCAGATCACCTGGTTGTATCCT	
A3B WT	<u>GGACCCCTGCCCGACTGTGTGGCGAAGCTGGCCGAATTCCTGTCTGAGCACCCCAATGTCACCCTGACCATCTCTGC</u>	>395
A3B C3	GGACCCCTGCCCGACTGTGTGGCGAAGCTGGCCGAATTCCTGTCTGAGCACCCCAATGTCACCCTGACCATCTCTGC	
A3B C5	GGACCCCTGCCCGACTGTGTGGCGAAGCTGGCCGAATTCCTGTCTGAGCACCCCAATGTCACCCTGACCATCTCTGC	
A3B WT	<u>CGCCCGCCTCTACTACTACTGGGAAAGAGATTACCGAAGGGCGCTCTGCAGGCTGAGTCAGGCAGGAGCCCGCGTGAAG</u>	>474
A3B C3	CGCCCGCCTCTACTACTACTGGGAAAGAGATTACCGAAGGGCGCTCTGCAGGCTGAGTCAGGCAGGAGCCCGCGTGAAG	
A3B C5	CGCCCGCCTCTACTACTACTGGGAAAGAGATTACCGAAGGGCGCTCTGCAGGCTGAGTCAGGCAGGAGCCCGCGTGAAG	
A3B WT	<u>ATCATGGACTATGAAGAATTTGCATACTGCTGGGAAAACCTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACA</u>	>553
A3B C3	ATCATGGACTATGAAGAATTTGCATACTGCTGGGAAAACCTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACA	
A3B C5	ATCATGGACTATGAAGAATTTGCATACTGCTGGGAAAACCTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACA	
A3B WT	<u>AATTCGATGAAAATTATGCATTCCTGCACCGCACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTAC</u>	>632
A3B C3	AATTCGATGAAAATTATGCATTCCTGCACCGCACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTAC	
A3B C5	AATTCGATGAAAATTATGCATTCCTGCACCGCACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTAC	
A3B WT	<u>TTTCAACTTTAATAATGACCCTTTGGTCTTCGACGGCGCCAGACCTACTTGTGCTATGAGGTGGAGCGCCTGGACAAT</u>	>711
A3B C3	TTTCAACTTTAATAATGACCCTTTGGTCTTCGACGGCGCCAGACCTACTTGTGCTATGAGGTGGAGCGCCTGGACAAT	
A3B C5	TTTCAACTTTAATAATGACCCTTTGGTCTTCGACGGCGCCAGACCTACTTGTGCTATGAGGTGGAGCGCCTGGACAAT	
A3B WT	<u>GGCACCTGGGTCTCTGATGGACCAGCACATGGGCTTTCTATGCAACGAGGCTAAGAATCTTCTCTGTGGCTTTTACGGCC</u>	>790
A3B C3	GGCACCTGGGTCTCTGATGGACCAGCACATGGGCTTTCTATGCAACGAGGCTAAGAATCTTCTCTGTGGCTTTTACGGCC	
A3B C5	GGCACCTGGGTCTCTGATGGACCAGCACATGGGCTTTCTATGCAACGAGGCTAAGAATCTTCTCTGTGGCTTTTACGGCC	
A3B WT	<u>GCCATGCGGAGCTGCGCTTCTTGGACCTGGTTCCTTCTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTT</u>	>869
A3B C3	GCCATGCGGAGCTGCGCTTCTTGGACCTGGTTCCTTCTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTT	
A3B C5	GCCATGCGGAGCTGCGCTTCTTGGACCTGGTTCCTTCTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTT	
A3B WT	<u>CATCTCCTGGAGCCCCTGCTTCTCCTGGGGCTGTGCCGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACGTGAGA</u>	>948
A3B C3	CATCTCCTGGAGCCCCTGCTTCTCCTGGGGCTGTGCCGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACGTGAGA	
A3B C5	CATCTCCTGGAGCCCCTGCTTCTCCTGGGGCTGTGCCGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACGTGAGA	
A3B WT	<u>CTGCGCATCTTCGCTGCCGCATCTATGATTACGACCCCCATATAAAGGAGGCGCTGCAAATGCTGCGGGATGCTGGGG</u>	>1027
A3B C3	CTGCGCATCTTCGCTGCCGCATCTATGATTACGACCCCCATATAAAGGAGGCGCTGCAAATGCTGCGGGATGCTGGGG	
A3B C5	CTGCGCATCTTCGCTGCCGCATCTATGATTACGACCCCCATATAAAGGAGGCGCTGCAAATGCTGCGGGATGCTGGGG	
A3B WT	<u>CCCAAGTCTCCATCATGACCTACGATGAGTTTGGAGTACTGCTGGGACACCTTTGTGTACCGCCAGGGATGTCCCTTCCA</u>	>1106
A3B C3	CCCAAGTCTCCATCATGACCTACGATGAGTTTGGAGTACTGCTGGGACACCTTTGTGTACCGCCAGGGATGTCCCTTCCA	
A3B C5	CCCAAGTCTCCATCATGACCTACGATGAGTTTGGAGTACTGCTGGGACACCTTTGTGTACCGCCAGGGATGTCCCTTCCA	
A3B WT	<u>GCCCTGGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAGGAAAAC</u>	>1182
A3B C3	GCCCTGGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAGGAAAAC	
A3B C5	GCCCTGGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAGGAAAAC	

**Figura 13.** Alineamiento de secuencias obtenidas por secuenciación Sanger de las clonas C3 y C5 vs secuencia silvestre de APOBEC3B. Gris: secuencia de Kozak; Amarillo: tag Hemaglutinina; Verde: codón de inicio; Rojo: mutaciones.

GCCATGGGCTACCCATACGATGTTCCAGATTACGCTATGAATCCACAGATCAGAAATCCGATGGAGCGGATGT  
 -M--N--P--Q--I--R--N--P--M--E--R--M--  
 ATCGAGACACATTCTACGACAACCTTTGAAAACGAACCCATCCTCTATGGTTCGGAGCTACACTTGGCTGTGCTA  
 Y--R--D--T--F--Y--D--N--F--E--N--E--P--I--L--Y--G--R--S--Y--T--W--L--C--Y  
 TGAAGTGAATAAAGAGGGGCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTTCGAGGCCAGGTGTATTTTC  
 --E--V--K--I--K--R--G--R--S--N--L--L--W--D--T--G--V--F--R--G--Q--V--Y--F--  
 AAGCCTCAGTACCACGCAGAAATGTGCTTCTCTTGGTTCTGTGGCAACCAGCTGCCTGCTTACAAGTGTT  
 -K--P--Q--Y--H--A--E--M--C--F--L--S--W--F--C--G--N--Q--L--P--A--Y--K--C--  
 TCCAGATCACCTGGTTTGTATCCTGGACCCCTGCCCGGACTGTGTGGCGAAGCTGGCCGAATTCTGTCTGA  
 F--Q--I--T--W--F--V--S--W--T--P--C--P--D--C--V--A--K--L--A--E--F--L--S--E  
 GCACCCCAATGTCACCCTGACCATCTCTGCCGCCGCTCTACTACTACTGGGAAAGAGATTACCGAAGGGCG  
 --H--P--N--V--T--L--T--I--S--A--A--R--L--Y--Y--Y--W--E--R--D--Y--R--R--A  
 CTCTGCAGGCTGAGTCAGGCAGGAGCCCGCTGAGATCATGGACTATGAAGAATTTCATACCTGCTGGGAAA  
 -L--C--R--L--S--Q--A--G--A--R--V--T--I--M--D--Y--E--E--F--A--Y--C--W--E--  
 ACTTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACAAATTCGATGAAAATTATGCATTCCTGCACCG  
 N--F--V--Y--N--E--G--Q--Q--F--M--P--W--Y--K--F--D--E--N--Y--A--F--L--H--R  
 CACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTCACCTTTCAACTTTAATAATGACCCTTTG  
 --T--L--K--E--I--L--R--Y--L--M--D--P--D--T--F--T--F--N--F--N--N--D--P--L  
 GTCCTTCGACGGCGCCAGACCTACTTGTGCTATGAGGTGGAGCGCCTGGACAATGGCACCTGGGTCTGTATGG  
 -V--L--R--R--R--Q--T--Y--L--C--Y--E--V--E--R--L--D--N--G--T--W--V--L--M--  
 ACCAGCACATGGGCTTCTATGCAACGAGGCTAAGAATCTTCTCTGTGGCTTTTACGGCCACCATGCGGAGCT  
 D--Q--H--M--G--F--L--C--N--E--A--K--N--L--L--C--G--F--Y--G--H--H--A--E--L  
 GCGCTTCGTTGGACCTGGTTCTTCTTTCAGTTGGACCCGCGCCAGATCTACAGGGTCACTTGGTTTCATCTCC  
 --R--F--L--D--L--V--P--S--L--Q--L--D--P--A--Q--I--Y--R--V--T--W--F--I--S  
 TGGAGCCCCGCTTCTCTGGGGCTGTGCCGGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACCTGAGAC  
 -W--S--P--R--F--S--W--G--C--A--G--E--V--R--A--F--L--Q--E--N--T--H--V--R--  
 TGCGCATCTTCGCTGCCCGCATCTATGATTACGACCCCTATATAAGGAGGCGCTGCAAAATGCTGCGGGATGC  
 L--R--I--F--A--A--R--I--Y--D--Y--D--P--L--Y--K--E--A--L--Q--M--L--R--D--A  
 TGGGGCCAAGTCTCCATCATGACCTACGATGAGTTTGTACTGTGGACACCTTGTGTACCCGAGGGA  
 --G--A--Q--V--S--I--M--T--Y--D--E--F--E--Y--C--W--D--T--F--V--Y--R--Q--G  
 TGTCCCTTCAGCCCTGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCC  
 -C--P--F--Q--P--W--N--G--L--E--E--H--S--Q--A--L--S--G--R--L--R--A--I--L--  
 AGAATCAGGGAAAC  
 Q--N--Q--G--N--

**Figura 14.** Secuencias nucleotídica y de aminoácidos de la clona C3 que codifica para el gen *APOBEC3B*... En Negro: secuencia de nucleótidos; Rojo: secuencia de aminoácidos; Gris y amarillo: dominio de coordinación de zinc; Amarillo: sitio activo; Rosa: mutaciones a nivel de nucleótido; Verde: mutaciones a nivel de aminoácido. \* Mutación generada en el sitio activo.

**Tabla 7.** Predicciones resultantes del análisis realizado con SIFT y PolyPhen-2 para las mutaciones presentes en la clona C3

Mutación	Cambio nucleotídico	Cambio aminoácido	Análisis SIFT	Análisis Polyphen-2	
				HumDiv	HumVar
1	A → C	K146T	Tolerable	Benigna	Benigna
2	G → A	R252H	Tolerable	Pos. dañina	Benigna
3	T → C	L260L	/	/	/
4*	T → C	C284R	Dañina	Prob. dañina	Prob. dañina
5	G → A	D360N	Tolerable	Benigna	Benigna

Pos. Dañina: Posiblemente dañina; Prob. Dañina: Probablemente dañina

GCCATGGGCTACCCATACGATGTTCCAGATTACGCTATGAATCCACAGATCAGAAATCCGATGGAGCGGATGT  
 -M--N--P--Q--I--R--N--P--M--E--R--M--  
 ATCGAGACACATTCTACGACAACTTTGAAAACGAACCCATCCTCTATGGTTCGGAGCTACACTTGGCTGTGCTA  
 Y--R--D--T--F--Y--D--N--F--E--N--E--P--I--L--Y--G--R--S--Y--T--W--L--C--Y  
 TGAAGTGAATAAAGAGGGGCGCTCAAATCTCCTTTGGGACACAGGGGTCTTTTCGAGGCCAGGTGATTTTC  
 --E--V--K--I--K--R--G--R--S--N--L--L--W--D--T--G--V--F--R--G--Q--V--Y--F--  
 GAGCCTCAGTACCACGCAGAAATGTGCTTCTCTTGGTTCTGTGGCAACCAGCTGCCTGCTTACAAGTGT  
 E--P--Q--Y--H--A--E--M--C--F--L--S--W--F--C--G--N--Q--L--P--A--Y--K--C--  
 TCCAGATCACCTGGTTTGTATCCTGGACCCCTGCCCGGACTGTGTGGCGAAGCTGGCCGAATTCCGTCTCGA  
 F--Q--I--T--W--F--V--S--W--T--P--C--P--D--C--V--A--K--L--A--E--F--L--S--E  
 GCACCCCAATGTCACCCTGACCATCTCGCCGCCGCTCTACTACTACTGGGAAAGAGATTACCGAAGGGCG  
 --H--P--N--V--T--L--T--I--S--A--A--R--L--Y--Y--Y--W--E--R--D--Y--R--R--A  
 CTCTGCAGGCTGAGTCAGGCAGGAGCCCGCTGAAGATCATGGACTATGAAGAATTGCATACTGCTGGGAAA  
 -L--C--R--L--S--Q--A--G--A--R--V--K--I--M--D--Y--E--E--F--A--Y--C--W--E--  
 ACTTTGTGTACAATGAAGGTCAGCAATTCATGCCTTGGTACAAATTCGATGAAAATTATGCATTCTGCACCG  
 N--F--V--Y--N--E--G--Q--Q--F--M--P--W--Y--K--F--D--E--N--Y--A--F--L--H--R  
 CACGCTAAAGGAGATTCTCAGATACCTGATGGATCCAGACACATTCACTTTCAACTTAAATAATGACCCTTTG  
 --T--L--K--E--I--L--R--Y--L--M--D--P--D--T--F--T--F--N--F--N--N--D--P--L  
 GTCCTTCGACGGCGCCAGACCTACTTGTGCTATGAGCGGAGCGCCTGGACAATGGCACCTGGGTCTGTATGG  
 -V--L--R--R--R--Q--T--Y--L--C--Y--E--A--E--R--L--D--N--G--T--W--V--L--M--  
 ACCAGCACATGGGCTTCTATGCAACGAGGCTAAGAATCTTCTCTGTGGCTTTTACGGCCGCATGCGGAGCT  
 D--Q--H--M--G--F--L--C--N--E--A--K--N--L--L--C--G--F--Y--G--R--H--A--E--L  
 GCGCTTCTGGACCTGGTTCTTCTTTCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTTTCATCTCC  
 --R--F--L--D--L--V--P--S--L--Q--L--D--P--A--Q--I--Y--R--V--T--W--F--I--S  
 TGGAGCCCTGCTTCTCTGGGGCTGTGCCGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACGTGAGAC  
 -W--S--P--C--F--S--W--G--C--A--G--E--V--R--A--F--L--Q--E--N--T--H--V--R--  
 TGCGCATCTTCGCTGCCCGCATCTATGATTATGACCCCTATATAAGGAGGCGCTGCAAAATGCTGCGGGGTGC  
 L--R--I--F--A--A--R--I--Y--D--Y--D--P--L--Y--K--E--A--L--Q--M--L--R--G--A  
 TGGGCCCCAAGTCTCCATCATGACCTACGATGAGTTTGACTACTGCTGGGACACCTTGTGTACCGCCAGGGA  
 --G--A--Q--V--S--I--M--T--Y--D--E--F--E--Y--C--W--D--T--F--V--Y--R--Q--G  
 TGTCCTTCCAGCCCTGGGATGGACTAGAGGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCC  
 -C--P--F--Q--P--W--D--G--L--E--E--H--S--Q--A--L--S--G--R--L--R--A--I--L--  
 AGAATCAGGGAAAC  
 Q--N--Q--G--N--

**Figura 15.** Secuencias nucleotídica y de aminoácidos de la clona C5 que codifica para el gen *APOBEC3B*. En Negro: secuencia de nucleótidos; Rojo: secuencia de aminoácidos; Gris y amarillo: dominio de coordinación de zinc; Amarillo: sitio activo; Rosa: mutaciones a nivel de nucleótido; Verde: mutaciones a nivel de aminoácido.

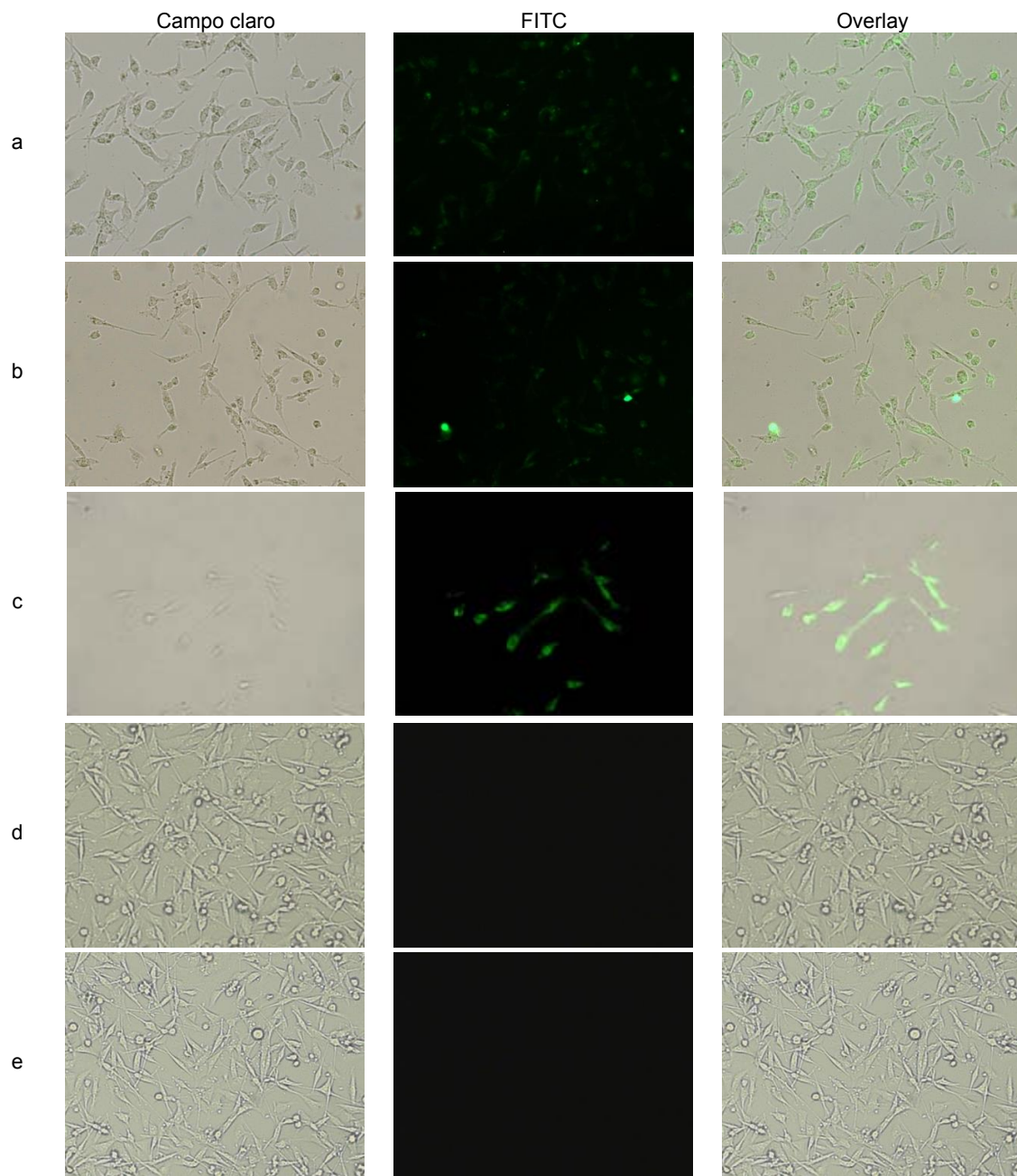
**Tabla 8.** Predicciones resultantes del análisis realizado con SIFT y PolyPhen-2 para las mutaciones presentes en la clona C5

Mutación	Cambio nucleotídico	Cambio aminoácido	Análisis SIFT	Análisis Polyphen-2	
				HumDiv	HumVar
1	A → G	K62E	Tolerable	Benigna	Benigna
2	T → C	S97S	/	/	/
3	T → C	S107S	/	/	/
4	T → C	V220A	Dañina	Pos. dañina	Pos. dañina
5	C → T	Y303Y	/	/	/
6	A → G	D328G	Tolerable	Benigna	Benigna

Pos. Dañina: Posiblemente dañina; Prob. Dañina: Probablemente dañina

#### **4.5 Transfección de plásmidos aislados de las clonas C3 y C5**

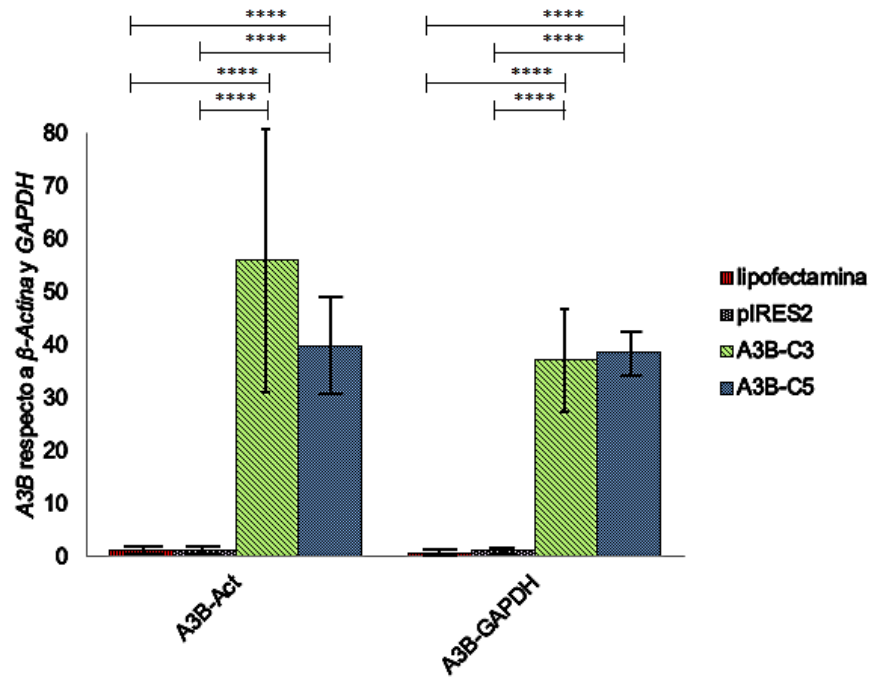
Se llevó a cabo una cotransfección de cada una de las clonas de *A3B* con el vector PIREs2-eGFP como control positivo de transfección en una proporción 10:1 en la línea celular humana de cáncer de mama MDA-MB-231. Como se puede observar en la Figura 16a/b, la intensidad de fluorescencia emitida en la cotransfección es menor en comparación con el tratamiento donde se transfectó el plásmido PIREs2-EGFP en la misma proporción que que las clonas 3 y 5 (600 ng), donde se observa una mayor intensidad de fluorescencia (Figura 16c). Igualmente, para los tratamientos expuestos a los complejos liposómicos (plásmidos más lipofectamina) se observó una disminución en la densidad celular (Figura 16a/b y c). En los controles lipofectamina sin plásmido y control sin transfección no se observó una disminución en la densidad celular. Con estos ensayos se confirmó la correcta entrada del plásmido *A3B* al citoplasma de las células MDA-MB-231, por lo que se prosiguió con la PCR en tiempo real para medir los niveles de expresión de *A3B* en las células transfectadas y no transfectadas.



**Figura 16.** Cotransfección de las clonas C3 y C5 con el plásmido pIRES2-eGFP en la línea celular humana de cáncer de mama MDA-MB-231. a) A3BC3 y b) A3BC5, cotransfección de las clonas 3 y 5 con el plásmido pIRES2-eGFP en proporción 1:10. c) transfección del plásmido pIRES2-eGFP en igual proporción que las clonas 3 y 5 (600 ng). d) Control únicamente con lipofectamina y e) Control sin transfección.

#### 4.6 PCR en tiempo real para evaluar los niveles de expresión de los transgenes

Se evaluaron los niveles de expresión de *APOBEC3B* endógeno y del transgen (*A3B-C3/5*) mediante PCR en tiempo real, empleando oligonucleótidos dirigidos contra la región ubicada entre los exones 3 y 4 del ORF del gen *A3B*. Los niveles de expresión se analizaron con el método  $2^{-\Delta\Delta C_t}$ , para comparar los niveles relativos de mRNA de *A3B* en relación a la expresión de los genes constitutivos  $\beta$ -*Actina* y *GAPDH*. La formula se aplicó comparando los niveles basales de *A3B*,  $\beta$ -*Actina* y *GAPDH* de las células sin tratamiento contra cada una de las células que fueron transfectadas con *A3B-C3* y *A3B-C5*. Se observó la sobreexpresión de 56 y 40 veces respecto al control de *A3B-C3* y *A3B-C5*, respectivamente, al normalizar contra  $\beta$ -*Actina*. Al normalizar contra *GAPDH* se observó una sobreexpresión de 37 y 38 veces, respecto al control para las clonas *A3B-C3* y *A3B-C5* (respectivamente). En las células donde no hubo transfección de los transgenes los niveles de expresión de *A3B* con respecto a ambos controles estuvo en un rango de 0.8 a 1.5 con respecto a la expresión de ambos genes constitutivos, por lo que al realizar los análisis estadísticos se encontró que existe una diferencia significativa en los niveles relativos de la expresión de *A3B* normalizados con los niveles de mRNA de  $\beta$ -*Actina* y *GAPDH* ( $n=3$ ;  $p<0.05$ ), entre las células que fueron transfectadas con las clonas 3 y 5, y las células no transfectadas con estos transgenes (tratamiento con PIREs2 y lipofectamina). Por lo anterior, se confirmó que los transgenes *A3B-C3* y *A3B-C5* se encuentran sobreexpresados en las células transfectadas.



**Figura 17.** Expresión relativa de *A3B* con respecto a  $\beta$ -Actina y *GAPDH*. Los datos se obtuvieron mediante PCR en tiempo real y fueron analizados con la fórmula  $2^{-\Delta\Delta C_t}$ . Se observa una diferencia significativa entre los niveles de expresión de *A3B* en las células transfectadas las clonas 3 y 5 en relación a las células transfectadas con el plásmido piRES2-eGFP y tratadas solo con lipofectamina.



## 5. DISCUSIÓN

En el presente trabajo se clonó la isoforma 1 del gen *APOBEC3B* a partir de las líneas celulares de cáncer de mama MCF-7 y MDA-MB-231. El contar con este reactivo biológico permitirá la realización de experimentos para estudiar la participación de *APOBEC3B* en procesos oncogénicos a nivel molecular en estudios posteriores.

El oligonucleótido sentido que se diseñó para la clonación de *A3B* tiene la secuencia de nucleótidos de un epítipo de reconocimiento HA que permite identificar a nivel de proteína el producto HA-*A3B*. Este epítipo no se expresa en células humanas y se empleó para identificar la proteína de fusión HA-*A3B* con anticuerpos anti-HA, debido a que actualmente no hay anticuerpos contra *A3B*. El epítipo artificial que codifica para la hemaglutinina (HA) es ampliamente utilizado en la detección de proteínas, además se tiene reportado su uso en estudios de clonación y detección de la proteína *A3G* con anticuerpos Anti-HA (Mehle *et al.*, 2004; Wang *et al.*, 2011; Ali *et al.*, 2013). Por otro lado, al primer sentido también se le añadió la secuencia Kozak para aumentar la eficiencia de la traducción de *A3B* (Figura 8). La secuencia de Kozak de 7 nucleótidos (5'-GCCATGG-3') está conservada evolutivamente y se ha demostrado que promueve el inicio de la traducción (Kozak, 1989). La importancia de que esta secuencia contenga guaninas en las posiciones -3 y +4, a partir del sitio de inicio de la traducción, radica en que esas posiciones están conservadas para guanina en los mRNAs de eucariontes y cambios en estos sitios reducen los niveles de traducción (Kozak, 1989).

Antes de comenzar con los ensayos de PCR se llevó a cabo un análisis computacional de los oligos y debido a que los resultados arrojados predecían la probable formación de dímeros entre la secuencia del primer antisentido (Tabla 4) por lo que se decidió utilizar una variante de la reacción en cadena de la polimerasa denominada PCR-TDW desarrollada por Don y colaboradores para

aumentar la especificidad de los primers con el molde disminuyendo así la formación de productos inespecíficos (Don *et al.*, 1991; Korbie y Mattick, 2008). Para los primers diseñados contra *A3B* la temperatura de desnaturalización fue establecida a 60°C, por lo que se decidió hacer la primera PCR-TDW 2°C por encima de su  $T_m$  (62°-59°). Sin embargo se amplificaron productos inespecíficos bajo estas condiciones, por lo que se realizó una nueva reacción empleando 4°C por encima de la  $T_m$  estimada (64-61°C) y en este último caso disminuyó casi en su totalidad la presencia de productos inespecíficos (Figura 10). Además un problema recurrente a lo largo de las reacciones de PCR fue la presencia de dímeros de primers, lo cual pudo deberse a la longitud del oligonucleótido sentido de 60 bases, la cual aumenta la posibilidad de (hetero) dímeros de primers (Tabla 6). Para asegurar que el molde de DNA utilizado para la clonación fuera únicamente la región codificante de *A3B* se realizó una purificación de las bandas del tamaño esperado. El DNA obtenido a partir de las bandas purificadas se reamplificó por PCR-TDW (Figura 10c) y se obtuvieron únicamente productos específicos, los cuales fueron empleados para llevar a cabo la ligación con el vector y proseguir con la transformación de las bacterias *E. coli* DH5 $\alpha$ .

La transformación de bacterias para el proceso de clonación depende en gran medida de la capacidad de competencia de las bacterias, la cual se define como la capacidad de una bacteria de internalizar material genético extracelular. Estas bacterias pueden ser sometidas a tratamientos *in vitro* tanto físicos como químicos, para generar esta competencia, y métodos como la electroporación y transformación por choque térmico se usan para transformar bacterias con DNA recombinante. Dependiendo del tipo de procedimiento, se alcanzan diferentes eficiencias de transformación. Por ejemplo la electroporación (método físico) alcanza una eficiencia de  $1 \times 10^9$  a  $1 \times 10^{10}$  transformantes por  $\mu\text{g}$  de plásmido (Dower *et al.*, 1988) y es considerada como de alta eficiencia. Por otro lado los métodos químicos (de los que provienen las bacterias usadas en este trabajo) alcanzan una eficiencia menor, que es de  $5 \times 10^6$  a  $2 \times 10^7$  transformantes por  $\mu\text{g}$  de plásmido (Cohen *et al.*, 1972). En este trabajo se obtuvo una eficiencia muy baja, ya que sólo se obtuvieron 3 colonias empleando 300 ng de plásmido (10

trasformantes por  $\mu\text{g}$  de plásmido. Las causas posibles de la baja eficiencia obtenida pueden incluir preparación inadecuada de las bacterias químicamente competentes, condiciones subóptimas de transformación (como tiempo tiempo de choque térmico insuficiente), baja eficiencia de las ligaciones empleadas en la transformación o debido a que la topoisomerasa I encargada de llevar a cabo la formación del enlace fosfodiéster tuviera poca actividad.

No obstante la baja eficiencia, se obtuvieron dos clonas (C3 y C5) con el inserto en la orientación adecuada, determinada por PCR de colonia (Figura 11). Las colonias C3 y C5 que resultaron positivas en la PCR de colonia fueron secuenciadas por el método de Sanger y se detectó que contienen 5 y 6 mutaciones, respectivamente (Figura 13). Las mutaciones en ambas clonas fueron clasificadas como silenciosas, variantes naturales y variantes no reportadas.

En la clona C3 únicamente se encontró una variante silenciosa que es la sustitución L260L (Tabla 7), mientras que en la clona C5 los cambios S97S y S107S que generan el aminoácido serina no están reportadas en ninguna enfermedad humana, sin embargo el cambio Y303Y que genera el aa tirosina se reporta como una variante presente en *A3B* codificada en los tumores de intestino delgado (Ensembl) por lo que podría ser una mutación propia de la línea celular MDA-MB-231, aunque dicha variante no está descrita en la base de datos COSMIC que contiene información sobre las mutaciones presentes en líneas celulares empleadas en investigación.

Las sustituciones de aminoácidos K146T de la clona C3 y K62E en la clona C5, fueron definidas como toleradas/benignas por SIFT y PolyPhen-2. Estas sustituciones anteriormente han sido reportadas por Collins y colaboradores en el 2004, después de que en un estudio analizaron la secuencia de mRNA del cromosoma 22, concluyeron que 546 genes codificaban para proteínas, posteriormente analizaron el ORF del 70% de los genes y encontraron entre otras estas variantes a las que nombraron como variante Lys-146 y variante Glu-62 (Collins *et al.*, 2004), actualmente son clasificadas como variantes naturales, ya que en la base de datos UniProt, la Lys-146 (VAR\_048722) se encuentra en una

proporción del 5% en la población mundial y Glu-62 (VAR\_048722; UniProt) en una proporción mayor al 60% y no se reporta su presencia en alguna enfermedad humana.

De las variantes no reportadas, para la clona C3 la sustitución D360N fue pronosticada por SIFT y PolyPhen-2 como tolerable/benigna y no hay reportes acerca de su participación en alguna enfermedad humana (PolyPhen-2). Por otro lado la sustitución R252H, SIFT arrojó como resultado que puede ser tolerable, no obstante PolyPhen-2 en su predictor probabilístico HumDiv dio como resultado que es posiblemente dañina debido a que en la historia evolutiva de A3B no se observa ninguna variante alélica que contenga esta sustitución. Igualmente, el predictor probabilístico HumVar de PolyPhen-2 le da una calificación de benigna debido a que no se ha reportado su presencia en enfermedades humanas. En lo que concierne al cambio C284R de esta clona C3, éste fue predicho por SIFT y PolyPhen-2 como dañino y probablemente dañino, respectivamente. Esto se debe a que se encuentra en una región sumamente conservada a lo largo de la historia evolutiva de la familia APOBEC, particularmente dentro del sitio activo donde esta cisteína participa directamente en la coordinación del átomo de zinc (Harris y Liddament, 2004; Swanton *et al.*, 2015). Por esta razón se espera que la capacidad de esta clona para coordinar el átomo de zinc sea ineficiente y por consiguiente su actividad catalítica esté comprometida.

Respecto a la clona C5, la sustitución nucleotídica T→C (V220A), fue identificada por SIFT y PolyPhen-2 como dañina y posiblemente dañina, respectivamente debido a que no es un cambio que esté presente en la historia natural de esta proteína y cambios en este aminoácido en particular se han reportado en cáncer de pulmón. Sin embargo el cambio de aminoácido que genera esta variante es mínimo, ya que no cambia el tipo (alifático) y la carga de la cadena lateral (no cargada). Por lo anterior, consideramos que las predicciones de SIFT y PolyPhen-2 como dañina/posiblemente dañina probablemente son incorrectas y se deben principalmente a la presencia de esta alteración en algunos tumores humanos. A este respecto, se conoce a través de muchos estudios que no todas las

mutaciones presentes en tumores tienen impacto funcional (mutaciones conductoras), sino que existen mutaciones que se acumulan debido a los procesos generadores de mutaciones, pero que no alteran la función proteica ni tienen relevancia en la neoplasia (Marx, 2014; McFarland *et al.*, 2014; Merid *et al.*, 2014)

Por último, en la clona C5 la sustitución D328G fue predicha como tolerable/benigna debido a que se da en una región variable de la proteína, por lo que no se ha reportado hasta el momento su participación en enfermedades humanas, además de que esta misma sustitución se encuentra presente en el gibón de mejillas blancas (*Nomascus leucogenys*).

La presencia de estas mutaciones no reportadas y de las variantes silenciosas podría deberse al uso de la enzima *Taq* polimerasa, proveniente de la bacteria extremofila *Thermus aquaticus*, para amplificar a *A3B*. Esta enzima permite la elongación de los fragmentos de DNA durante la PCR, pero carece de actividad de 5'→3' exonucleasa y tiene una tasa de error de 1 mutación en cada 9000 nucleótidos (Eom *et al.*, 1996). En este sentido, es posible que la *Taq* polimerasa haya incorporado las bases erróneas observadas durante las dos reacciones de PCR que se utilizaron. Los análisis realizados a ambas clonas mediante las herramientas computacionales sugieren que los cambios observados en la clona C5 probablemente no tienen consecuencias, sin embargo para la clona C3 se esperaba una repercusión importante en su función de desaminasa debido a la mutación C284R que se generó en el sitio activo.

Se eligió utilizar como reportero para la cotransfección la proteína verde fluorescente del plásmido pIRES2-eGFP debido a que comparte con el vector pcDNA™3.1/V5-His TOPO (utilizado para la clonación de *A3B*) la secuencia correspondiente al promotor de citomegalovirus (CMV), que permite altos niveles de expresión en células de mamíferos, así como una alta eficiencia de la señal de poliadenilación obtenida del virus SV-40. En la Figura 16c se puede observar que la señal emitida por la fluorescencia de la proteína verde fluorescente es considerable al ser transfectada en la misma proporción que las clonas 3 y 5 por lo que se infiere que los niveles de expresión del gen *APOBEC3B* en las células

transfectadas con estas clonas deben ser similares y proporcionales a las cantidades empleadas en la transfección, debido a que comparten características que permiten una alta eficiencia durante la traducción. Aunque en este trabajo no evaluamos la expresión de los transgenes de *APOBEC3B* a nivel de proteína, se realizó un perfil de expresión génica a nivel de mensajero, en donde se observó un aumento significativo entre los niveles relativos del gen *A3B* con respecto a los genes constitutivos  *$\beta$ -Actina* y *GAPDH* de las células sin transfectar y las células que fueron transfectadas con los transgenes *A3B-C3* y de *A3B-C5* (Figura 17). Aunque estos ensayos confirman la expresión de los transgenes a nivel de mRNA, es necesario evaluar la expresión a nivel de proteína empleando anticuerpos anti-HA, específicos contra las quimeras transgénicas construidas de *APOBEC3B* para conocer el efecto a nivel biológico de esta sobreexpresión inducida en células con y sin tratamiento para evaluar más a fondo la participación de estas proteínas en el proceso oncogénico.

Por lo descrito anteriormente, se cumplió el objetivo de este trabajo: clonar el gen *APOBEC3B* a partir de líneas celulares de cáncer de mama y caracterizarlo hasta nivel de mRNA, además de que se indujo su sobreexpresión a nivel del transcrito, por lo que las clonas generadas permitirán la realización de estudios futuros para evaluar el papel de *APOBEC3B* en diferentes procesos moleculares alterados en el cáncer.

## 6. CONCLUSIONES

Se clonó la secuencia codificante completa el gen *APOBEC3B* a partir de las líneas celulares MDA-MB-231 y MCF-7. Se obtuvieron 2 clonas, las cuales contienen el epítipo HA. Las secuencias codificantes presentaron variantes alélicas presentes en la población europea/amerindia/asiática y africana.

La clona C3 presentó sustituciones que probablemente afectan la función de *A3B* a nivel de proteína, mientras que para la clona C5 se predijo una sola variante que podría afectar la función de este péptido, sin embargo no se encuentra en un sitio activo y el cambio de aminoácido es neutro, por lo cual es importante concluir su caracterización a nivel funcional.

Se confirmó la expresión a nivel de mRNA de las clonas *A3B-C3* y *A3B-C5*, en células que fueron transfectadas. Estas clonas sobreexpresaron hasta 55 veces el gen *A3B* en comparación a los niveles basales del gen en células que no fueron sometidas al tratamiento.

## 7. PERSPECTIVAS

Se realizará la caracterización de las clonas obtenidas para evaluar la expresión a nivel de proteína y para determinar su actividad de desaminasa. Para ello se tiene planeado lo siguiente:

Para llevar a cabo el análisis de las clonas a nivel de proteína, se transfectarán las clonas de *A3B* en células de mamífero para que las moléculas de DNA recién internalizadas se traduzcan y permita tener un gran número de copias de la proteína. Se realizará un análisis de expresión mediante Western-Blot, ya que es una técnica abundantemente utilizada en biología celular y molecular para el estudio proteínas de interés y está basada en la utilización de un anticuerpo específico para llevar a cabo el aislamiento del péptido específico (Mahmood y Yang, 2012). Dado que actualmente no se tiene reportado en la literatura el uso un anticuerpo dirigido específicamente contra *A3B*, se utilizará un anticuerpo dirigido contra HA, epítipo localizado en el extremo N-terminal de la proteína traducida.

La función editora de *APOBEC3B* se evaluará mediante ensayos enfocados en la función de desaminación de *A3B*, como el reportado por Burns y colaboradores (Burns *et al.*, 2013 (1)), cuyo principio se basa en la transferencia de energía de Förster (FRET), que se logra mediante la utilización de una sonda constituida por una secuencia de DNA de cadena sencilla y las moléculas fluorescentes TAMRA y -6-FAM en sus extremos 5' y 3' (respectivamente), moléculas que se encuentran acopladas mediante sus emisiones de energía (Felekyan *et al.*, 2013). La secuencia de nucleótidos contiene el trinucleótido TCA, después de la desaminación generada por este péptido se generará una ruptura de la cadena sencilla seguida del desacoplamiento de -6-FAM y TAMRA por lo que únicamente es captada la señal de emisión de 492 nm perteneciente a -6-FAM. Este ensayo resulta el más adecuado, ya que es específico para la medir actividad de *APOBEC3B* como molécula editora del DNA y ya contamos con los reactivos necesarios para realizar estos análisis (Burns *et al.*, 2013a; Caval *et al.*, 2015).



Al concluir con estos estudios, ya habiendo evaluado la principal característica de este gen como editor del material genético se estudiará su participación en procesos oncogénicos particulares. Además de que se realizarán estudios de colocalización de estas proteínas con proteínas que participan en procesos celulares como las rupturas de doble cadena y la transcripción de las células, para estudiar si *A3B* a nivel de proteína está alterando estos procesos, causando así la acumulación de mutaciones (Alexandrov *et al.*, 2013).

## 8. REFERENCIAS

- Adzhubei I., Schmidt S., Peshkin L., Ramensky V., Gerasimova A., Bork P., Kondrashov A., y Sunyaev S. 2010. *A Method and Server for Predicting Damaging Missense Mutations*. *Nature Methods*. 7(4): 248–249 pp.
- Adzhubei I., Jordan D., Sunyaev S. 2013. *Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2*. *Curr Protoc Hum Genet*. Unit7.20
- Alexandrov L., Nik-Zainal S., Wedge D., Aparicio S., Behjati S., Biankin A., Bignell G., Bolli N., Borg A., Børresen-Dale A., Boyault S., Burkhardt B., Butler A., Caldas C., Davies H., Desmedt C., Eils R., Eyfjörd J., Foekens J., Greaves M., Hosoda F., Hutter B., Ilcic T., Imbeaud S., Imielinski M., Jäger N., Jones D., Jones D., Knappskog S., Kool M., Lakhani S., López-Otín C., Martin S., Munshi N., Nakamura H., Northcott P., Pajic M., Papaemmanuil E., Paradiso A., Pearson J., Puente X., Raine K., Ramakrishna M., Richardson A., Richter J., Rosenstiel P., Schlesner M., Schumacher T., Span P., Teague J., Totoki Y., Tutt A., Valdés-Mas R., van Buuren M., van 't Veer L., Vincent-Salomon A., Waddell N., Yates L., Zucman-Rossi J., Futreal P., McDermott U., Lichten P., Meyerson M., Grimmond S., Siebert R., Campo E., Shibata T., Pfister S., Campbell P. y Stratton M. 2013. *Signatures of mutational processes in human cancer*. *Nature*. 500(7463):415-21 pp.
- Ali S., Karki N., Bhattacharya C., Zhu R., MacDuff D., Stenglein M., Schumacher A., Demorest Z., Harris R. S., Matin A. and Aggarwal S. 2013. *APOBEC3 Inhibits DEAD-END Function to Regulate microRNA Activity*. *BMC Molecular Biology*. 14(16): 1-9 p.
- Arce C., Bargalló E., Villaseñor Y., Gamboa C., Lara F., Pérez V. y Villareal P. 2011. *Oncoguía: Cáncer de mama*. *Cancerología*. 6, 77-86 pp.
- Avesson L. y Barry G. 2014. *The Emerging Role of RNA and DNA Editing in Cancer*. *Biochimica et Biophysica Acta*. 1845: 308-316 pp.
- Bogerd H., Wiegand H., Doehle B., Lueders K. y Cullen B. 2006. *APOBEC3A and APOBEC3B are Potent Inhibitors of LTR-retrotransposon Function in Human Cells*. *Nucleic Acids Research*. 34(1): 89–95 pp.

- Boyle P. 2012. *Triple-negative breast cancer: epidemiological considerations and recommendations*. *Annals of Oncology*. 23, 7-12 pp.
- Burns M., Lackey L., Carpenter M., Rathore A., Land A., Leonard B., Refsland E., Kotandeniya D., Tretyakova N., Nikas J., Yee D., Temiz N., Donohue D., McDougle R., Brown W., Law E y Harris R. 2013. *APOBEC3B is an Enzymatic Source of Mutation in Breast Cancer*. *Nature*. 494: 366-370 pp.  
(a)
- Burns M., Leonard B. y Harris R. S. 2015. *APOBEC3B: Pathological Consequences of an Innate Immune DNA Mutator*. *Biomed J*. 38(2): 102-110pp.
- Burns M., Temiz N. y Harris R. 2013. *Evidence for APOBEC3B Mutagenesis in Multiple Human Cancers*. *Nature Genetics*. 45(9): 977-983 pp. (b)
- Caval V., Bouzidi M., Suspène R., Laude H., Dumargne M. C., Bashamboo A., Krey T., Vartanian J. P. y Wain-Hobson S. 2015. *Molecular Basis of the Attenuated Phenotype of Human APOBEC3B DNA Mutator Enzyme*. *Nucleic Acids Research*. doi: 10.1093/nar/gkv935
- Cetin I. y Topcul M. 2014. *Triple Negative Breast Cancer*. *Asian Pacific Journal of Cancer Prevention*. 15(6): 2427-2431 pp.
- Chrzan P., Skokowski J., Karmolinski A. y Pawelczyk T. 2001. *Amplification of c-myc gene and overexpression of c-Myc protein in breast cancer and adjacent non-neoplastic tissue*. *Biochem*. 34(7): 557-562 pp.
- Cohen S., Chang A. y Hsu L. 1972. *Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of Escherichia coli by R-Factor DNA*. *Proc. Natl. Acad. Sci. U.S.A.* 69(8): 2110–2114 pp.
- Collins J., Wright C., Edwards C., Davis M., Grinham J., Cole C., Goward M., Aguado B., Mallya M., Mokrab Y., Huckle E., Beare D. y Dunham I. 2004. *A genome Annotation-driven Approach to Cloning the Human ORFeome*. *Genome Biology*. 5(10): R84.1- R84.9 p.
- Coticello S., Thomas C., Petersen-Mahrt S. K. y Neuberger M. 2005. *Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases*. *Molecular Biology and Evolution*. 22(2):367-377 pp.
- Dawood S. 2010. *Triple-Negative Breast Cancer*. *Drugs*. 70(17): 2247-2258 pp.

- Don R. H., Cox P. T., Wainwright B., Baker K. y Mattick J. 1991. *'Touchdown' PCR to circumvent spurious priming during gene amplification* . Nucleic Acids Research. 19(14): 4008 pp.
- Don-Marc F. y Petersen-Mahrt S. 2014. *AID and APOBEC Deaminases: Balancing DNA Damage in Epigenetics and Immunity*. Epigenomics. 6(4): 427-443 pp.
- Dower W., Miller J. and Ragsdale C. 1988. *High Efficiency Transformation of E.coli by High Voltage Electroporation*. Nucleic Acids Research. 16(13): 6127-6145 pp.
- Dunham I., Hunt A., Collins J. E., Bruskiewich R., Beare D. M., Clamp M., Smink L. J., Ainscough R., Almeida J. P., Babbage A., Bagguley C., Bailey J., Barlow K., Bates K. N., Beasley O., Bird C. P., Blakey S., Bridgeman A. M., Buck D. , Burgess J., Burrill W. D., Burton J., Carder C., Carter N. P., Chen Y., Clark G., Clegg S. M., Copley V., Cole C. G., Collier R. E., Connor R. E., Conroy D., Corby N., Coville G. J., Cox A. V., Davis J., Dawson E., Dhami P. D., Dockree C., Dodsworth S. J., Durbin R. M., Ellington A., Evans K. L., Fey J. M., Fleming K., French L., Garner A. A., Gilbert J. G. R., Goward M. E., Grafham D., Griffrths M. N., Hall C., Hall R., Hall-Tamlyn G., Heathcott R. W., Ho S., Holmes S., Hunt S. E., Jones M. C., Kershaw J., Kimberley A., King A., Laird G. K., Langford C. F., Leversha M. A., Lloyd C., Lloyd D. M., Martyn I. D., Mashreghi-Mohammadi M., Matthews L., McCann O. T., McClay J., McLaren S., McMurray A. A., Milne S. A., Mortimore B. J., Odell C. N., Pavitt R., Pearce A. V., Pearson D., Phillimore B. J., Phillips S. H., Plumb R. W., Ramsay H., Ramsey Y., Rogers L., Ross M. T., Scott C. E., Sehra H. K., Skuce C. D., Smalley S., Smith M. L., Soderlund C., Spragon L., Steward C. A., Sulston J. E., Swann R. M., Vaudin M., Wall M., Wallis J. M., Whiteley M. N., Willey D., Williams L., Williams S., Williamson H., Wilmer T. E., Wilming L., Wright C. L., Hubbard T., Bentley D. R., Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S., Roe B. A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H. I., Lewis J., Lewis S., Lin S.P., Loh P., Malaj E., Nguyen T., Pan H., Phan S., Qi S.,

- Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z., Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P., Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohling T., Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J., Korf I., Bedell J. A., Hillier L., Mardis E., Waterston R., Wilson R., Emanuel B. S., Shaikh T., Kurahashi H., Saitta S., Budarf M. L., McDermid H. E., Johnson A., Wong A. C. C., Morrow B. E., Edelman L., Kim U. J., Shizuya H., Simon M. I., Dumanski J. P., Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C. E. y O'Brien K. P. 1999. *The DNA sequence of human chromosome 22*. 402, 489-495 pp.
- Eom S., Wang J. y Steitz T. 1996. Structure of Taq Polymerase with DNA at the Polymerase active site. *Nature*. 382, 278 – 281 pp.
- Esnault C., Heidmann O., Delebecque F., Dewannieux M., Ribet D., Hance A., Heidmann T. y Schwartz O. 2005. *APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses*. 433, 430-433 pp.
- Felekyan S., Sanabria H., Kalinin S., Kühnemuth R. y Seidel C. 2013. Chapter Two – Analyzing Förster Resonance Energy Transfer with Fluctuation Algorithms. *Methods in Enzymology*. 519, 39–85 pp.
- Forte Giusi Irma. 2013. *APOBEC3B Spreads Somatic Mutations in Breast Cancer Genome*. *Transl Cancer Res*. 2(2):97-99 pp.
- Giordano S., Buzdar A. y Hortobagyi G. 2002. *Breast cancer in men*. *Ann Intern Med*. 137(8): 678-687 pp.
- Hanahan D. y Weinberg R. 2011. *Hallmarks of Cancer: The Next Generation*. *Cell*. 144, 646-674 pp.
- Harris S. R. 2015. *Molecular Mechanism and Clinical Impact of APOBEC-catalyzed Mutagenesis in Breast Cancer*. *Breast Cancer Research*. 17(8): 1-10 p.
- Harris S. y Liddament. 2004. *Retroviral Restriction by APOBEC Proteins*. *Nature*. 4, 868-877 pp.
- Hebert JR. 2009. *The Biology and Treatment of Cancer: Understanding Cancer*. New Jersey. Eds. Pardee AB, Stein GS. John Wiley and Sons. 223-255 p.

- Henderson S. y Fenton T. 2015. *APOBEC3 Genes: Retroviral Restriction Factors to Cancer Drivers*. Cell Press. 21(5): 274-284 pp.
- Henikoff S. y Ng P. 2003. SIFT: Predicting Amino Acid Changes that Affect Protein Function. Nucleic Acids Research. 31(13): 3812–3814 pp.
- Hicks S., Wheeler D., Plon S., y Kimmel M. 2011. *Prediction of Missense Mutation Functionality Depends on Both the Algorithm and Sequence Alignment Employed*. Human Mutation. 32(6): 661–668pp.
- Hulka B. S. y Moorman G. 2008. *Reprint of Breast Cancer: Hormones and Other Risk Factors*. Journal of the Climateric Postmenopause. 61(1-2): 203-213 pp.
- Key T., Verkasalo P. y Banks E. 2001. *Epidemiology of breast cancer*. The Lancet Oncology. 2, 133-140 pp.
- Koito A. y Ikeda A. 2011. *Intrinsic Restriction Activity by AID/APOBEC Family of Enzymes Against the Mobility of Retroelements*. Mobile Genetic Elements. 1(3): 197-202 pp.
- Korbie D. y Mattick J. 2008. *Touchdown PCR for increased specificity and sensitivity in PCR amplification*. Nature Reviews. 3, 1452-1456 pp.
- Kozak M. 1989. *The Scanning Model for Translation: An Update*. The Journal of Cell Biology. 108, 229-241 pp
- Kuong K. y Loeb L. 2013. *APOBEC3B Mutagenesis in Cancer*. Nature Genetics. 45(9): 964-965 pp.
- Lackey L., Law E., Brown W. y Harris R. S. 2013. *Subcellular Localization of the APOBEC3 Proteins During Mitosis and Implications for Genomic DNA Deamination*. Cell Cycle. 12(5): 762-772 pp.
- Lackey Lela Lynn. 2012. *APOBEC3 Subcellular Localization and Genomic Editing*. Thesis Degree of Doctor of Philosophy, U.S.A. University of Minnesota.
- Leonard G. y Swain S. 2004. *Ductal Carcinoma In Situ, Complexities and Challenges*. Journal of the National Cancer Institute. 96(12): 906-920 pp.
- MacMahon Brian. 2005. *Epidemiology and the causes of Breast Cancer*. Int. J. Cancer. 118, 2373-2378 pp.
- Madsen P., Anant S., Rasmussen H., Gromov P., Vorum H., Dumanski J., Tommerup N., Collins J., Wright C., Dunham I., MacGinnitie A., Davidson

- N. y Celis J. 1999. *Psoriasis Upregulated Phorbol-1 Shares Structural but not Functional Similarity to the mRNA-Editing Protein APOBEC-1*. Journal of Investigative Dermatology. 113, 162–169
- Mahmood T. y Yang P. 2012. *Western Blot: Technique, Theory, and Trouble Shooting*. North American Journal of Medical Sciences. 4(9): 429–434 pp.
- Malhotra G.K., Zhao X., Band H. and Band V. 2010. *Histological, Molecular and Functional Subtypes of Breast Cancers*. Cancer Biology y Therapy. 10(10): 955-960 pp.
- Malim M. 2009. *APOBEC proteins and intrinsic resistance to HIV-1 infection*. Phil. Trans. Of the Royal Society. 364, 675–687 pp.
- Marx V. 2014. *Cancer Genomes: Discerning Drivers From Passengers*. Nature Methods. 11(4): 375-379pp.
- McFarland C., Mirny L., y Korolev K. 2014. *Tug-of-war Between Driver and Passenger Mutations in Cancer and other Adaptive Processes*. 111(42): 15138-15143 pp.
- Mehle A., Strack B., Ancuta P., Zhang C., McPike M. and Gabuzda D. 2004. *Vif Overcomes the Innate Antiviral Activity of APOBEC3G by Promoting Its Degradation in the Ubiquitin-Proteasome Pathway*. Journal of Biological Chemistry. 279, 7792-7798 pp.
- Merid S., Goranskaya D. and Alexeyenko A. 2014. *Distinguishing Between Driver and Passenger Mutations in Individual Cancer Genomes by Network Enrichment Analysis*. BMC Bioinformatic. 1-21p.
- Modesto A., Gandy C., Mery E., Filleron T., Massabeau C., Izar F., Charitansky H., Roché H. y de Lafontan B. 2014. *Breast Ductal carcinoma in situ with Microinvasion: Pathological review and clinical implications*. Cancer/Radiothérapie 18, 107–110 pp.
- Moris A., Murray S. y Carninaud S. 2014. *AID and APOBECs Span the Gap Between Innate and Adaptive Immunity*. Frontiers in Microbiology. 5(534): 1-13 p.
- Nardone A., Corvigno S., Brescia A., D'Andrea D., Limite G. y Veneziani M. 2011. *Long-term cultures of stem/progenitor cells from lobular and ductal breast carcinomas under non-adherent conditions*. Cytotechnology. 63, 67–80 pp.

- Nelson, J. A., Reynolds-Kohler, C., and Smith, B. A. (1987). Negative and Positive Regulation by a Short Segment in the 5'-Flanking Region of the Human Cytomegalovirus Major Immediate-Early Gene. *Molec. Cell. Biol.* 7, 4125-4129pp.
- Nguyen D., Bos P. y Massagué J. 2009. *Metastasis: from Dissemination to Organ-specific Colonization*. *Nature Reviews*. 9, 274-285 pp.
- Perou C., Sùrlie T., Eisen M., Rijn M., Jeffreyk S., Rees C., Pollack J., Ross D., Johnsen H., Akslen L., Flugel E., Pergamenschikov A., Cheryl Williams C., Zhu S., Lunning P., Bùrresen-Dale A., Brown P. y Botstein D. 2000. *Molecular Portraits of Human Breast Tumours*. *Nature*. 406, 747-752 pp.
- Porter P. 2008. "Westernizing Women Risk? Breast Cancer in Lower Income Countries". *The New England Journal of Medicine*. 358, 213-216 pp.
- Powell L., Wallis S., Pease R., Edwards Y., Knott T. y Scott J. 1987. *A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine*. *Cell*. 50(6): 831-840 pp.
- Prohaska K., Bennett R., Salter J. y Smith H. 2014. *The Multifaceted Roles of RNA Binding in APOBEC Cytidine Deaminase Functions*. John Wiley y Sons, Ltd. 5: 493-502 pp.
- Reis-Filho J., Foulkes W. y Smith I. 2010. *Triple-Negative Breast Cancer*. *The New England Journal of Medicine*. 63(20): 1938-1948 pp.
- Richardson S., Narvaiza I., Planegger R., Weitzman M. y Moran J. 2014. *APOBEC3A Deaminates Transiently Exposed Single-Strand DNA During LINE-1 Retrotransposition*. *Elife*. 10.7554/eLife.02008
- Roberts S. y Gordenin D. 2014. *Hypermutation in Human Cancer Genomes: Footprints and Mechanisms*. *Nature Reviews*. 14, 786-800 pp.
- Roberts S., Lawrence M., Klimczak L., Grimm S., Fargo D., Stojanov P., Kiezun A., Kryukov G.,<sup>4</sup> Carter S., Saksena G., Harris S., Shah R., Resnick M., Getz G. y Gordenin D. 2013. *An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers*. *Nature*. 45(9): 970-977 pp.
- Sabatier R., Finetti P., Guille A., Adelaide J., Chaffanet M., Viens P., Birnbaum D. y Bertucci F. 2014. *Claudin-low breast cancers: clinical, pathological,*



- molecular and prognostic characterization*. *Molecular Cancer*. 13(228): 1-14 p.
- Scholzová E., Malík R., Ševčík J. y Kleibl Z. 2007. *RNA regulation and cancer development*. *Cancer Letters*. 246, 12-23 pp.
- Sieuwerts A., Willis S., Burns M., Look M., Meijer-Van Galder M., Schlicker A., Heideman M., Jacobs H., Wessels L., Leyland-Jones B., Gray K., Foekens J., Harris R. S. y Martens J. 2014. *Elevated APOBEC3B Correlates with Poor Outcomes for Estrogen-Receptor-Positive Breast Cancer*. *Hormones and Cancer*. 5(6): 405-413 pp.
- Smith H., Bennett R., Kizilyer A., McDougall W. y Prohaska K. 2012. *Functions and Regulation of the APOBEC Family of Proteins*. *Seminars in Cell y Developmental Biology*. 23: 258-268 pp.
- Swanton C., McGranahan N., Starrett G. J. y Harris R. S. 2015. *APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity*. *Cancer Discovery*. 5(7): 704-712 pp.
- Taylor B., Nik-Zainal S., Ling Wu Y., Stebbings L., Raine K., Campbell P., Rada C., Stratton M. R. y Neuberger M. 2013. *DNA Deaminases Induce Break-associated Mutation Showers With Implication of APOBEC3B and 3A in Breast Cancer Kataegis*. *eLIFE*. DOI: 10.7554/eLife.00534.
- Teng B., Burant C. y Davidson N. 1993. *Molecular cloning of an apolipoprotein B messenger RNA editing protein*. *Science*. 260(5115): 1816–1819 pp.
- Untergasser A., Cutcutache I. Koressaar T., Ye J., Faircloth B., Remm M. y Rozen S. 2012. *Primer3—new Capabilities and Interfaces*. *Nucleic Acids Research*. 40(15): 1-12 p.
- Venkitaraman A. 2001. *Functions of BRCA1 and BRCA2 in the biological response to DNA damage*. *Journal of Cell Science* 114, 3591-3598 pp.
- Viale G. 2012. *The current state of breast cancer classification*. *Annals of Oncology*. 23, suplement 10.
- Wang Y., Shao Q., Yu X., Kong W., Hildreth J. E. K. y Liu B. 2011. *N-Terminal Hemagglutinin Tag Renders Lysine-Deficient APOBEC3G Resistant to HIV-1 Vif-Induced Degradation by Reduced Polyubiquitination*. *Journal of Virology*. 85(9): 4510–4519 pp.

- Weigel M. y Dowsett M. 2010. *Current and Emerging Biomarkers in Breast Cancer: prognosis and prediction*. Endocrine-Related Cancer. 17, 245–262 pp.
- Woodman M. 2008. *Direct PCR of intact bacteria (colony PCR)*. Curr. Protoc. Microbiol. Appendix 3D
- Xu J., Chen Y. y Olopade O. 2010. *MYC and Breast Cancer*. Genes Cancer. 1(6): 629–640 pp.