



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

IDENTIFICACIÓN DE MUTACIONES EN GENES MITOCONDRIALES DE LA CADENA
RESPIRATORIA EN CÁNCER DE MAMA

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Maestro en Ciencias

PRESENTA:

CARLOS JHOVANI PÉREZ AMADO

TUTOR PRINCIPAL

SILVIA JIMÉNEZ MORALES
[Instituto Nacional de Medicina Genómica](#)

MIEMBROS DEL COMITÉ TUTOR

ALFREDO HIDALGO MIRANDA
[Instituto Nacional de Medicina Genómica](#)

ELIZABETH LANGLEY McCARRON
[Instituto Nacional de Cancerología](#)

Ciudad de México. Agosto, 2019



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.

PMDCB/1245/2019

Pérez Amado Carlos Jhovani
Estudiante de Maestría en Ciencias Bioquímicas
P r e s e n t e

Los miembros del Subcomité Académico en reunión ordinaria del día 13 de mayo del presente año, conocieron su solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **MAESTRO EN CIENCIAS**, con la réplica de la tesis "**Identificación de mutaciones en genes mitocondriales de la cadena respiratoria en cáncer de mama**", dirigida por el/la Dr(a). **Jiménez Morales Silvia**.

De su análisis se acordó nombrar el siguiente jurado integrado por los doctores:

PRESIDENTE	Pérez Martínez Xochitl
VOCAL	Resendis Antonio Osbaldo
VOCAL	Guardado Estrada José Mariano
VOCAL	Tovar Romero Hugo Antonio
SECRETARIO	González Halphen Diego

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

Atentamente
"POR MI RAZA, HABLARÁ EL ESPÍRITU"
Cd. Universitaria, Cd. Mx., a 13 de mayo de 2019.
COORDINADORA



Dra. ANA BRÍGIDA CLORINDA ARIAS ÁLVAREZ

Dedicado a mi familia, es especial a mis padres:
Bruna Amado y Juan Pérez, quienes no han dejado de
apoyarme ni un solo momento.

Dedicado a Angel Ponce, a quien agradezco su compañía
en la mayor parte de este trayecto y el haber estado a mi
lado todo el tiempo que el destino lo permitió.

AGRADECIMIENTOS

Esta tesis fue realizada bajo la asesoría de la Dra. Silvia Jiménez Morales en el Laboratorio de Genómica del Cáncer perteneciente al Instituto Nacional de Medicina Genómica. Se contó con el apoyo del comité tutor integrado por el Dr. Alfredo Hidalgo Miranda y la Dra. Elizabeth Langley McCarron.

Se agradece al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo económico proporcionado durante los estudios de posgrado, así como al Programa de Apoyo a Estudiantes de Posgrado (PAEP) por los recursos brindados para la asistencia al congreso internacional ESHG 2018.

A la Dra. Silvia Jiménez Morales, agradezco profundamente su supervisión, asesoría, enseñanzas y la confianza depositada en mí para el desarrollo de este proyecto y alcanzar los objetivos propuestos.

Al Dr. Alfredo Hidalgo Miranda, jefe del Laboratorio de Genómica del Cáncer del INMEGEN, por sus asesorías y enseñanzas que contribuyeron tanto al proyecto como a mi formación académica.

A la Unidad de Secuenciación (USec) del INMEGEN, agradezco su apoyo técnico en la implementación del proceso de secuenciación masiva.

A la Unidad de Servicios Bioinformáticos del INMEGEN, en especial al Dr. Hugo Tovar y a la Dra. Laura Gómez por su apoyo y asesoría en el análisis de los datos genómicos; además de adentrarme en el campo de la bioinformática.

Al Instituto de Enfermedades de la Mama (FUCAM), por su apoyo y atenciones durante la recolección de las muestras. Así como a las pacientes, quienes desinteresadamente accedieron a donar sus muestras y que sin ellas este proyecto no hubiera podido ser llevado a cabo.

Finalmente, a todos los integrantes del Laboratorio de Genómica del Cáncer, por todas las experiencias vividas y que contribuyeron tanto a mi formación académica, profesional y personal, haciendo un lugar de trabajo ameno y productivo. Hago mención especial para Amalinalli Ruíz y Fredy Beltrán, quiénes me brindaron su compañerismo, amistad y confianza.

ÍNDICE

RESUMEN	1
ABSTRACT	2
1 INTRODUCCIÓN	3
1.1 <i>MITOCONDRIA</i>	4
1.1.1 SISTEMA DE FOSFORILACIÓN OXIDATIVA.....	4
1.1.1.1 Complejo I: NADH-ubiquinona oxidorreductasa	6
1.1.1.2 Complejo II: Succinato-ubiquinona deshidrogenasa.....	6
1.1.1.3 Complejo III: Ubiquinona-citocromo c oxidorreductasa	6
1.1.1.4 Complejo IV: Citocromo c oxidasa	6
1.1.1.5 Complejo V: F ₀ F ₁ -ATP sintasa	7
1.1.2 GENOMA MITOCONDRIAL	7
1.1.2.1 Alta tasa de mutación en el genoma mitocondrial	9
1.1.2.2 Haplogrupos.....	9
1.1.2.3 Heteroplasmia y homoplasmia	10
1.1.2.4 Expresión umbral	10
1.2 <i>MITOCONDRIA Y CÁNCER</i>	11
1.3 <i>CÁNCER DE MAMA</i>	14
1.3.1 EPIDEMIOLOGÍA DEL CÁNCER DE MAMA	15
1.3.2 CLASIFICACIÓN MOLECULAR DEL CÁNCER DE MAMA	15
1.3.2.1 Subtipo luminal A.....	16
1.3.2.2 Subtipo luminal B.....	16
1.3.2.3 Subtipo positivo a HER2	16
1.3.2.4 Subtipo basal.....	17
1.3.3 VARIANTES MITOCONDRIALES EN CÁNCER DE MAMA	18
1.3.4 HAPLOGRUPOS EN CÁNCER DE MAMA	19
2 JUSTIFICACIÓN	21
3 OBJETIVOS	22
3.1 <i>OBJETIVO GENERAL</i>	22
3.2 <i>OBJETIVOS PARTICULARES</i>	22
4 ESTRATEGIA EXPERIMENTAL	23
4.1 <i>POBLACIÓN DE ESTUDIO</i>	23
4.2 <i>OBTENCIÓN DE MUESTRA</i>	23
4.3 <i>ANÁLISIS MOLECULAR</i>	24
4.3.1 EXTRACCIÓN DE ÁCIDOS NUCLEICOS.....	24
4.3.1.1 Obtención de DNA de sangre periférica	24
4.3.1.2 Obtención de DNA de tejido tumoral	24
4.3.2 EVALUACIÓN DE PUREZA E INTEGRIDAD DE MUESTRAS.....	25
4.3.3 ENRIQUECIMIENTO DEL DNA MITOCONDRIAL MEDIANTE PCR LARGA	25
4.3.4 PURIFICACIÓN DE PRODUCTOS DE 8.5 Kb	26
4.3.5 POOL DE AMPLICONES	27

4.3.6	FRAGMENTACIÓN DE LOS AMPLICONES	27
4.3.7	SECUENCIACIÓN DEL DNA MITOCONDRIAL MEDIANTE PLATAFORMA MiSeq Illumina	27
4.4	ANÁLISIS DE DATOS	28
4.4.1	ANÁLISIS DE CALIDAD	28
4.4.2	ALINEAMIENTO DE LAS LECTURAS.....	28
4.4.3	PRE-PROCESAMIENTO DE ALINEAMIENTOS	29
4.4.3.1	Marcado de duplicados.....	29
4.4.3.2	Indexación.....	29
4.4.3.3	Recalibración de calidad de bases	29
4.4.4	IDENTIFICACIÓN DE VARIANTES EN TUMORES.....	29
4.4.5	DETECCIÓN DE HETEROPLASMIA.....	29
4.4.6	IDENTIFICACIÓN DE MUTACIONES SOMÁTICAS	30
4.4.7	ANOTACIÓN Y PREDICCIÓN DEL EFECTO FUNCIONAL	30
4.4.8	IDENTIFICACIÓN DE MUTACIONES SOMÁTICAS NUEVAS Y ASOCIADAS A CÁNCER	30
4.4.9	IDENTIFICACIÓN DE PERFILES MUTACIONALES POR SUBTIPO MOLECULAR	30
4.4.10	ANÁLISIS DE HAPLOGRUPOS	31
4.4.11	ANÁLISIS ESTADÍSTICOS.....	31
5	RESULTADOS	33
5.1	<i>CARACTERÍSTICAS CLÍNICAS DE LA POBLACIÓN DE ESTUDIO</i>	33
5.2	<i>NÚMERO Y DISTRIBUCIÓN DE LAS VARIANTES EN EL mtDNA DE MUJERES CON CÁNCER DE MAMA</i>	34
5.3	<i>ANÁLISIS DE HETEROPLASMIA</i>	36
5.4	<i>ANÁLISIS DE MUTACIONES EN TUMORES DE CÁNCER DE MAMA</i>	39
5.5	<i>MUTACIONES SOMÁTICAS NUEVAS</i>	39
5.6	<i>ANÁLISIS DEL EFECTO FUNCIONAL DE LAS MUTACIONES SOMÁTICAS</i>	42
5.7	<i>MUTACIONES SOMÁTICAS DEL mtDNA REPORTADAS COMO ASOCIADAS A OTROS TIPOS DE CÁNCER</i>	43
5.8	<i>DISTRIBUCIÓN DE MUTACIONES POR SUBTIPOS MOLECULARES DE CÁNCER DE MAMA</i>	44
5.9	<i>DISTRIBUCIÓN DE HAPLOGRUPOS</i>	46
6	DISCUSIÓN	48
7	CONCLUSIONES	57
8	PERSPECTIVAS	58
9	REFERENCIAS	59
10	PRODUCCIÓN CIENTÍFICA	65
10.1	<i>REUNIONES CIENTÍFICAS</i>	65
10.2	<i>PUBLICACIONES EN CÁNCER DE MAMA</i>	65
10.3	<i>PUBLICACIONES EN OTROS TEMAS</i>	66

ÍNDICE DE TABLAS

Tabla 1. Clasificación inmunohistoquímica de los tumores de cáncer de mama.	17
Tabla 2. Reactivos para la PCR larga.....	26
Tabla 3. Condiciones de termociclado para PCR larga.....	26
Tabla 4. Condiciones de sonicación para la fragmentación del mtDNA.	27
Tabla 5. Distribución de los subtipos moleculares en la población de estudio.	33
Tabla 6. Mutaciones somáticas más frecuentes en tumores de mama analizados.....	42
Tabla 7. Clasificación funcional de las mutaciones somáticas identificadas en cáncer de mama.	43
Tabla 8. Mutaciones somáticas previamente relacionadas con cáncer.....	44
Tabla 9. Número de mutaciones somáticas por subtipo molecular.	45
Tabla 10. Número de mutaciones por individuo y subtipo molecular.....	45
Tabla 11. Distribución de los haplogrupos por subtipo molecular de cáncer de mama.....	47

ÍNDICE DE FIGURAS

Figura 1. Compartimentalización mitocondrial y sistema de fosforilación oxidativa (OXPHOs).....	5
Figura 2. Mapa del genoma mitocondrial.....	8
Figura 3. Expansión de mutaciones en mtDNA.....	11
Figura 4. Efecto Warburg.....	13
Figura 5. Alteraciones en mtDNA en cáncer de mama.....	20
Figura 6. Análisis bioinformático de variantes mitocondriales.....	32
Figura 7. Distribución genómica de las variantes mitocondriales en cáncer de mama.....	35
Figura 8. Frecuencia alélica de las variantes mitocondriales.....	37
Figura 9. Distribución del grado de heteroplasmia de las variantes mitocondriales identificadas.....	38
Figura 10. Distribución genómica de las mutaciones somáticas mitocondriales.....	40
Figura 11. Niveles de heteroplasmia de las mutaciones somáticas.....	41
Figura 12. Distribución de mutaciones en los subtipos moleculares de cáncer de mama.....	46
Figura 13. Distribución de haplogrupos en cáncer de mama.....	47

ABREVIACIONES

ADP: Adenosin difosfato	SNV: Variante de un solo nucleótido.
ATP: Adenosin trifosfato	tRNA: RNA de transferencia
BWA: Burrows-Wheeler-Aligner	VCF: Variant Calling Format
CO: Citocromo oxidasa	VEP: Variant Effect Predictor
CTCF: Factor de unión a sitios CCCTC	
ETC: Cadena transportadora de electrones	
CYB: Citocromo b	
DNA: Acido desoxirribonucleico	
Drp: Proteína tipo dinamina	
FA: Fracción alélica	
FADH₂: Flavin adenín dinucleótido	
FMN: Flavin mononucleótido	
GATK: Genome analysis tool kit	
HER2: Receptor del factor de crecimiento epidérmico humano 2	
HIF1-α: Factor 1 alfa inducible por hipoxia	
HVR: Región hipervariable	
INDEL: Inserción y delección	
IT_H: Sitio de inicio de la transcripción de la cadena pesada	
IT_L: Sitio de inicio de la transcripción de la cadena ligera	
mRNA: RNA mensajero	
mtDNA: DNA mitocondrial	
mTFA: Factor de transcripción mitocondrial	
NADH₂: dinucleótido de nicotinamida	
ND: NADH deshidrogenasa	
nDNA: DNA nuclear	
NGS: Next generation sequencing	
O_H: Origen de replicación de la cadena pesada	
O_L: Origen de replicación de la cadena ligera	
OXPHOs: Sistema de fosforilación oxidativa	
PCR: Reacción en cadena de la polimerasa	
PDH: Piruvato deshidrogenasa	
rCRS: Secuencia revisada de referencia de Cambridge	
RE: Receptor de estrógenos	
RNA: Ácido ribonucleico	
ROS: Especies Reactivas de Oxígeno	
RP: Receptor de progesterona	
rRNA: RNA ribosomal	
SNP: Polimorfismo de un solo nucleótido	

RESUMEN

Introducción. El genoma mitocondrial codifica para 13 proteínas esenciales de los complejos enzimáticos que conforman el sistema de fosforilación oxidativa (OXPHOs). Actualmente se sabe que variaciones en la secuencia del DNA mitocondrial (mtDNA) se asocian al desarrollo de diversos padecimientos humanos e incluso, se han sugerido como biomarcadores de diagnóstico y pronóstico. El objetivo de este trabajo fue identificar mutaciones en genes codificantes del sistema OXPHOs en tumores de mama. **Material y métodos.** En este estudio se secuenció el DNA mitocondrial (mtDNA) completo de muestras pareadas (tumor-sangre periférica) de 92 pacientes mexicanas con diagnóstico clínico de cáncer de mama, con edad promedio de 53.83 ± 11.5 años en un intervalo de 33 a 92 años. **Resultados.** Se identificaron un total de 707 variantes en las muestras analizadas, con diferente grado de heteroplasmia. De todas las variantes encontradas, 140 fueron detectadas sólo en los tumores (mutaciones somáticas) de mama; de los cuales, 87 (62.1%) se localizaron en regiones codificantes y 53 (37.9%) en regiones no codificantes (regiones regulatorias). La mayor parte de las mutaciones somáticas mostraron un patrón homoplásmico. El gen *ND5* fue el que presentó una mayor frecuencia de mutaciones somáticas, lo que podría implicar alteraciones en la función del Complejo I. Las mutaciones más frecuentes en las pacientes fueron la T9950C (7.6%), C16278T (5.4%) y A11002T (4.3%). Al realizar el análisis de la distribución con respecto al subtipo molecular de cáncer de mama no se identificó un perfil característico para cada uno de ellos. **Conclusiones.** La tendencia hacia el estado homoplásmico de las mutaciones somáticas y la ausencia de un perfil relacionado con la biología del subtipo tumoral del cáncer de mama, nos sugiere que las mutaciones mitocondriales podrían ser independientes del subtipo y conferir ventajas adaptativas que le permite sobrevivir a las células tumorales. Sin embargo, se requieren estudios funcionales que describan los mecanismos que conllevan a esta adaptación.

ABSTRACT

Introduction. The mitochondrial genome encodes for 13 essential proteins of the oxidative phosphorylation system (OXPHOs). It is well known that mutations in the mitochondrial genome (mtDNA) are associated with several human diseases, even so, has been suggested as biomarkers of diagnosis and prognosis. Our aim was to identify mtDNA mutations in the OXPHOs encoded proteins in breast tumors. **Material and methods.** We sequenced the complete mitochondrial DNA (mtDNA) from paired samples (tumor-peripheral blood) of 92 Mexican women with clinical diagnosis of breast cancer. Age patients ranged from 33 to 92 (mean age of 53.83 ± 11.5) years old. **Results.** A total of 707 variants were identified, which exhibited different degrees of heteroplasmy. Overall variants, 140 were detected only in breast tumors (somatic mutations); of which, 87 (62.1%) were located in coding regions and 53 (37.9%) in non-coding regions (regulatory regions). Mostly of somatic mutations have a homoplasmic pattern. The *ND5* gene had the highest frequency of somatic mutations, which could affect the Complex I function. The most frequent somatic mutations were T9950C (7.6%), C16278T (5.4%) and A11002T (4.3%). No mutational profile was identified as associated with the molecular subtype of breast cancer. **Conclusions.** The tendency toward the homoplasmic state of somatic mutations and the absence of a profile associated with the tumor subtype of breast cancer, suggests that mitochondrial mutations could be independent of the subtype and that some mtDNA mutations might confer adaptive advantages that allow the tumor progression. However, functional studies describing these mechanisms are required.

1 INTRODUCCIÓN

La mitocondria es un organelo que reside en el citoplasma de las células y que está implicado en la producción de energía, señalización y muerte celular programada. Posee su propio genoma que se encarga de codificar elementos esenciales para el funcionamiento del sistema de fosforilación oxidativa (OXPHOs) y cubrir así el 97% del ATP requerido por la célula. Recientemente se ha descrito que las mutaciones en el genoma mitocondrial pueden generar alteraciones en la función del organelo y ello conducir o contribuir al desarrollo de enfermedades metabólicas, neurodegenerativas y cáncer, graves problemas de salud pública en el mundo. Como ejemplo, el cáncer de mama, el cual se ha convertido en el tipo de cáncer más frecuente en población femenina, cuya incidencia y mortalidad va en aumento tanto en el mundo como en México. El conocimiento de la contribución de las mutaciones en el DNA mitocondrial (mtDNA) en el desarrollo carcinogénico de la glándula mamaria aún es escaso; mismo que se ve limitado porque se ha observado que estas mutaciones varían en tipo, naturaleza y frecuencia de una población a otra. El objetivo de este estudio es describir el panorama mutacional del mtDNA de tumores de mama de pacientes mexicanas, inferir *in silico* su efecto funcional e identificar si existe un perfil mutacional en los diferentes subtipos de cáncer de mama descritos hasta la fecha.

1.1 MITOCONDRIA

La mitocondria es un organelo altamente dinámico, involucrado en procesos fundamentales, como la síntesis del adenosín trifosfato (ATP) necesario para cubrir los requerimientos energéticos celulares [1]. Este organelo es semiautónomo ya que su biogénesis depende tanto del genoma nuclear, como de su propio genoma, el cual tiene una replicación y expresión independiente del ciclo celular [2, 3]. Se estima que existen entre 1000 y 2000 mitocondrias por célula, dependiendo del requerimiento energético de cada una de ellas [4]. Además de sintetizar la mayor parte del ATP celular por el sistema OXPHOs, la mitocondria también es un importante centro metabólico y de señalización implicado en la regulación de la muerte celular y en la generación de especies reactivas de oxígeno (ROS) [3, 5-7].

1.1.1 SISTEMA DE FOSFORILACIÓN OXIDATIVA

El sistema OXPHOs tiene lugar en la mitocondria y se encarga de generar el ATP necesario para la célula a través de la oxidación de NADH_2 y FADH_2 provenientes del metabolismo intermediario [6, 8]. Es un sistema constituido por cinco complejos multienzimáticos embebidos en la membrana mitocondrial interna: NADH_2 -ubiquinona oxidoreductasa (Complejo I), Succinato-ubiquinona deshidrogenasa (Complejo II), Ubiquinona-citocromo *c* oxidoreductasa (Complejo III), Citocromo *c* reductasa (Complejo IV) y la F_1F_0 -ATP sintasa (Complejo V).

Los complejos I-IV son los complejos respiratorios que conforman la Cadena Transportadora de Electrones (ETC, por sus siglas en inglés *electron transport chain*) y se encargan de transportar los electrones acoplando la energía con el bombeo de protones (H^+) desde la matriz mitocondrial hacia el espacio intermembranal para generar el gradiente electroquímico que es empleado por el complejo V que lleva a cabo la síntesis de ATP. Otros componentes del sistema OXPHOs son la ubiquinona y el citocromo *c*, que son moléculas encargadas de transportar los electrones de un complejo a otro (Fig. 1) [9, 10].

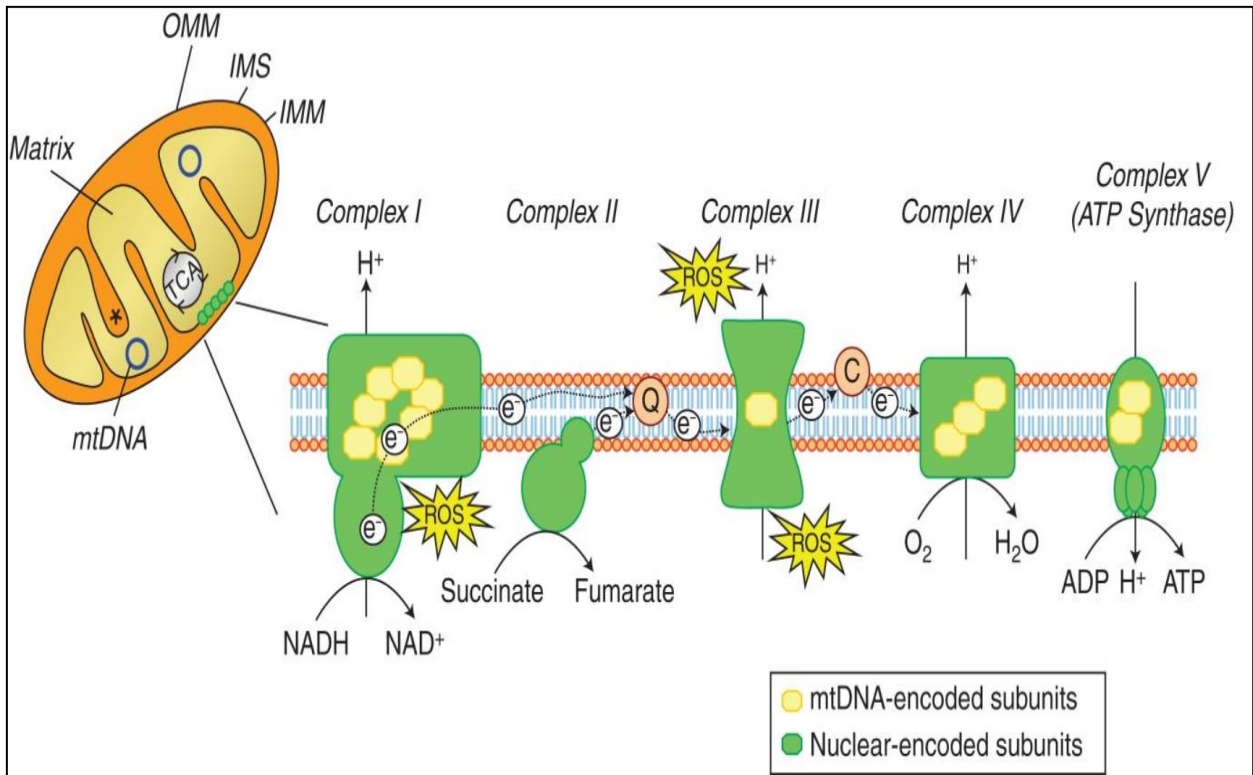


Figura 1. Compartimentalización mitocondrial y sistema de fosforilación oxidativa (OXPHOS). Se esquematiza la estructura de la mitocondria, la cual se divide en cuatro compartimientos: membrana externa (OMM del inglés outer mitochondrial membrane), membrana interna (IMM del inglés inner mitochondrial membrane), espacio intermembranal (IMS, del inglés inner membrane space) y matriz, donde se encuentra el DNA y enzimas de las vías metabólicas. En la IMM se localiza el sistema OXPHOS que está conformado por cinco complejos (I-V). Los electrones provenientes del metabolismo entran al complejo I a través de la oxidación del NADH_2 y al complejo II a través de la oxidación del succinato. Los electrones son transportados por la ubiquinona (Q) y llevados al complejo III donde son cedidos al citocromo c (C) y transportados al complejo IV donde se lleva a cabo la reducción del O_2 a H_2O . El transporte electrónico genera un gradiente electroquímico a través del bombeo de H^+ que es utilizado por el complejo V para sintetizar ATP. Durante el transporte electrónico se generan especies reactivas de oxígeno (ROS) principalmente en los complejos I y III. En los complejos enzimáticos se señalan las subunidades que son codificadas por el DNA mitocondrial [11].

1.1.1.1 Complejo I: NADH-ubiquinona oxidorreductasa

El complejo I es el más grande de la ETC. En mamíferos, está constituido por aproximadamente 45 subunidades, de las cuales, siete (genes *ND1*, *ND2*, *ND3*, *ND4L*, *ND4*, *ND5* y *ND6*) son codificadas por el mtDNA [10, 12, 13]. Tiene la función principal de oxidar el NADH proveniente del metabolismo y transferir los electrones hacia la ubiquinona, acoplado esta reacción al bombeo de 4H^+ desde la matriz mitocondrial hacia el espacio intermembranal [8, 13, 14].

1.1.1.2 Complejo II: Succinato-ubiquinona deshidrogenasa

Es un complejo transmembranal de únicamente 4 subunidades, las cuales son exclusivamente codificadas por el DNA nuclear (nDNA) [15]. Su función principal es catalizar la oxidación de succinato a fumarato, una reacción que también forma parte del ciclo de los ácidos tricarbónicos, y transferir los electrones a la ubiquinona. El complejo II es la segunda puerta de entrada de los electrones a la CTE y además el único complejo que no es capaz de bombear H^+ [8].

1.1.1.3 Complejo III: Ubiquinona-citocromo *c* oxidorreductasa

También conocido como complejo *bc₁*, es el complejo donde convergen los electrones provenientes del complejo I y II [8]. Está compuesto por 11 subunidades y funciona como un dímero. La subunidad proteica que contiene dos hemos *b* (codificada por el gen *CYB*) es crucial en la transferencia de los electrones y es codificada por el mtDNA [10]. La función principal del complejo III es la de oxidar el ubiquinol y transferir los electrones al citocromo *c*; además de translocar 4H^+ hacia el espacio intermembranal que contribuyen a la generación del gradiente electroquímico mediante el ciclo Q [8].

1.1.1.4 Complejo IV: Citocromo *c* oxidasa

El complejo IV es el último complejo respiratorio y tiene la función de catalizar la transferencia de los electrones provenientes del citocromo *c* al O_2 para reducirlo a H_2O , promoviendo la translocación de 2H^+ que contribuyen a la formación del gradiente electroquímico [16]. Este complejo ejerce su actividad en forma dimérica y tres de sus 13 subunidades (*COI*, *COII* y *COIII*) están codificadas en el mtDNA y corresponden a aquellas que contienen los centros catalíticos y los canales de bombeo de protones [8, 17].

1.1.1.5 Complejo V: F₀F₁-ATP sintasa

Es el último complejo del sistema OXPHOs que puede actuar de forma dimérica o monomérica [18, 19]. Cada monómero está constituido por alrededor de 17 subunidades, de las cuales a y A6L (genes *ATP6* y *ATP8*, respectivamente) están codificadas en el mtDNA y son importantes en mantener la estabilidad del complejo y de promover su dimerización [18, 20, 21].

Este complejo tiene la función principal de catalizar la síntesis de 2.5 moléculas de ATP empleando la energía proveniente del gradiente electroquímico de 10H⁺ generados por los complejos respiratorios [19, 21, 22]. Además, este complejo también participa en el modelaje de la arquitectura mitocondrial a partir de su forma dimerizada, que constituye a la formación de las crestas que compartimentalizan y hacen más eficiente el sistema OXPHOs [21, 22].

1.1.2 GENOMA MITOCONDRIAL

El mtDNA es una molécula duplohelicoidal circular que consta de 16569 pb. Tiene dos cadenas: la cadena H (heavy) rica en guaninas y la cadena L (light) rica en citosinas. Codifica para 37 genes que corresponden a 2 rRNA encontrados en los ribosomas mitocondriales (12S rRNA y 16S rRNA), 22 tRNAs utilizados para traducir los mRNA mitocondriales y 13 subunidades proteicas que forman parte de cuatro de los cinco complejos enzimáticos del sistema OXPHOs. (Fig. 2) [4, 10, 23]. La molécula de mtDNA tiene una organización extremadamente compacta, se compartimentaliza en nucleoides (asociación de mtDNA con proteínas), tiene genes sobrelapados, carece de intrones y, exceptuando la región D-Loop, sólo 87 pb forman parte de regiones intergénicas [3, 24]. La región D-Loop, bucle de desplazamiento o región control, es una región no codificante de alrededor de 1.1 kb de longitud [25] que se extiende de la posición 16024 a la 576 y contiene dos regiones hipervariables (HVR, del inglés Hypervariable Region) cortas, HVR1 y HVR2 [2, 26, 27]. Esta región se encuentra involucrada en la regulación de la transcripción y replicación de la molécula de mtDNA.

El genoma mitocondrial también se caracteriza por tener un gran número de copias dentro de una misma célula, estimándose que pueden existir de 1 a 10 moléculas por mitocondria, lo que, aunado a la alta tasa de mutaciones en este genoma, se da lugar al fenómeno de heteroplasmia. Así mismo, la falta de recombinación favorece la segregación de las mutaciones en haplogrupos [2, 26, 27].

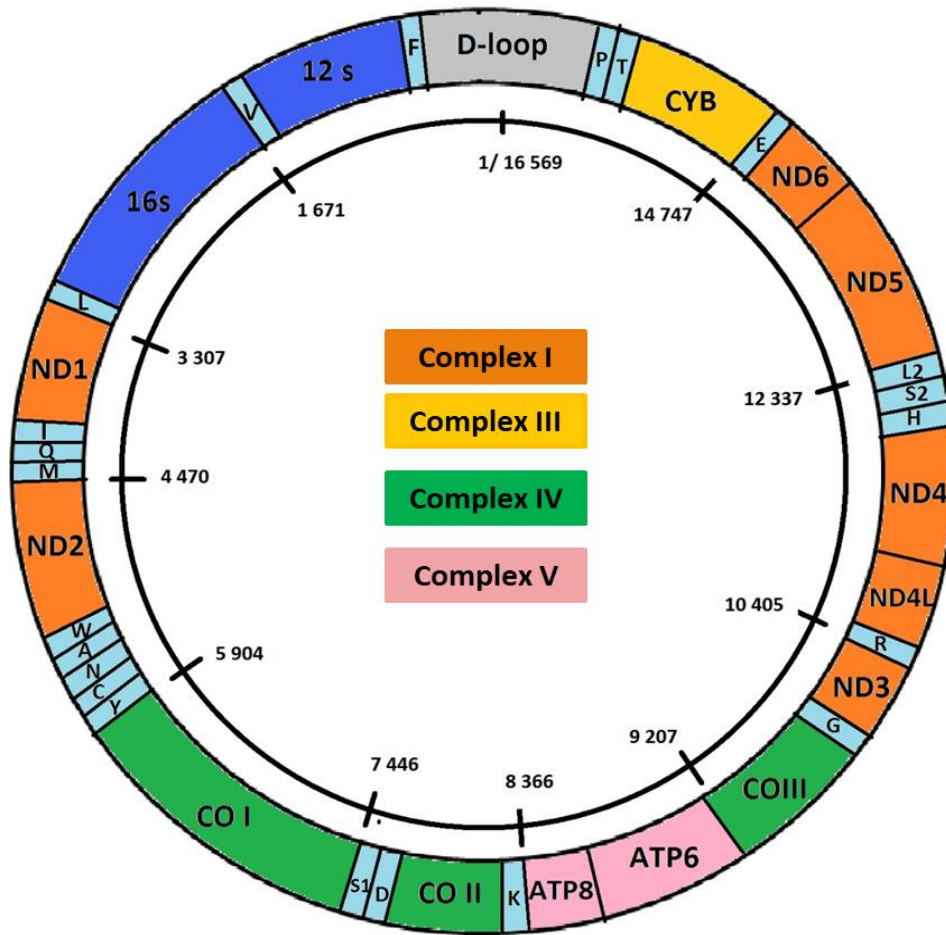


Figura 2. Mapa del genoma mitocondrial. Se esquematiza la organización de los genes codificados por el genoma mitocondrial. Los 13 genes codificantes se muestran coloreados según el complejo enzimático al que pertenecen; rRNA: azul oscuro, tRNA: azul claro; región no codificante (D-Loop): gris.

1.1.2.1 Alta tasa de mutación en el genoma mitocondrial

El mtDNA tiene una organización diferente y una amplia variación entre especies e incluso entre individuos de una misma especie [28]. Es una molécula susceptible a presentar daños que modifican su estructura; de hecho, la tasa de mutación espontánea estimada es de 10 a 20 veces más que la del nDNA (2.7×10^{-5} versus 1.0×10^{-8} sustituciones/sitio/año, en mamíferos) [10, 29-32]. La elevada tasa de mutación observada en el mtDNA se explica por el dinámico estado bioenergético de la mitocondria, lo que produce una alta concentración de ROS que dañan a la molécula desnuda del DNA (protegida débilmente por proteínas no histonas) y por falta de mecanismos de reparación eficientes [2, 24]. Se ha reportado que los daños o mutaciones en el genoma mitocondrial contribuyen a la disminución en la capacidad respiratoria de los tejidos, a un estado de envejecimiento celular y a la aparición de enfermedades crónicas como diabetes, entidades neurodegenerativas, cáncer, etc [25].

1.1.2.2 Haplogrupos

Las mutaciones que aparecen en el mtDNA pueden ser inocuas o patogénicas. Las mutaciones inocuas se fijan en la población como polimorfismos produciendo genotipos diferentes que dan lugar a una nueva línea mitocondrial y que, por la falta de recombinación de este genoma, se segregan en grupos (haplogrupos) a través de las generaciones. De hecho, son útiles para el estudio de fenómenos migratorios de las poblaciones humanas, conocer el origen étnico de un individuo, evaluar la capacidad de las poblaciones para adaptarse al ambiente. Más aún, han sido asociados al desarrollo de distintas patologías humanas, incluyendo el cáncer [33, 34].

Los haplogrupos están organizados en familias y son nombrados con letras mayúsculas (de A hasta Z), siendo A, B y C, los más comunes en poblaciones de ancestría amerindia. Estos tres haplogrupos representan más del 84% de todos los que se encuentran en la población mestiza de México. [35-37].

1.1.2.3 Heteroplasmia y homoplasmia

Las moléculas del mtDNA en cada célula pueden o no tener la misma secuencia nucleotídica, cuando ésta es la misma en todas las células se conoce como homoplasmia, pero cuando existen diferencias en una o más bases, se define como heteroplasmia. [10, 38]. La heteroplasmia se produce por la adquisición de mutaciones generadas durante la replicación, por el daño a la molécula del mtDNA debido a la presencia de ROS o por herencia. La heteroplasmia afecta a todas las células de un individuo, pero solo las del ovocito pueden transmitirse a la descendencia que habitualmente manifiesta niveles variados de la misma debido a la segregación aleatoria de las mitocondrias durante la ovogénesis [3, 10]. Dependiendo de la proporción del alelo mutado, en los tejidos podemos encontrar tres genotipos diferentes de acuerdo con la proporción del alelo mutado: homoplásmico normal (proporción de mutación <5%), homoplásmico mutante (proporción de mutación > 95%) y heteroplásmico (proporción de mutación de 5-95%). En consecuencia, el fenotipo de una célula con heteroplasmia dependerá del porcentaje y naturaleza del DNA mutado que contenga [38].

1.1.2.4 Expresión umbral

Las células son capaces de tolerar un porcentaje mínimo de mtDNA mutado, mismo que, si es superado, puede dar lugar a la expresión de una condición patológica [3, 26]. Cuando el número de moléculas de mtDNA mutadas es pequeño, el mtDNA normal es capaz de mantener la función mitocondrial (Fig. 3); de lo contrario, cuando el número de copias de mtDNA mutado sobrepasa un umbral (hasta ahora desconocido), la producción de ATP decrece por debajo del mínimo necesario para la función celular, llevando al desarrollo de diferentes patologías, incluyendo el cáncer [5, 10, 26].

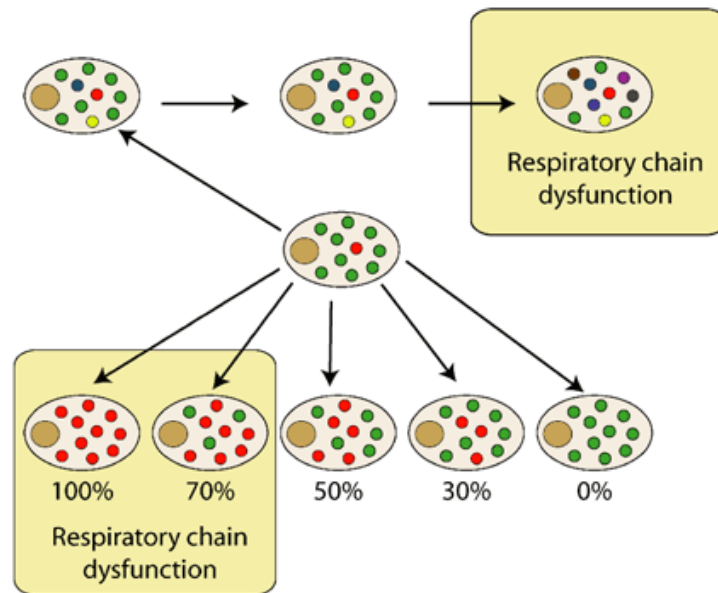


Figura 3. Expansión de mutaciones en mtDNA. La acumulación de las mutaciones en el mtDNA tiene como consecuencia la disfunción del sistema OXPHOs a causa de un bajo rendimiento de la Cadena Transportadora de Electrones. La baja proporción de mutaciones (círculos rojos) en una célula, difícilmente da lugar a un fenotipo patogénico. No obstante, si la mutación supera en número a la del mtDNA normal (círculos verdes) o pasa un umbral, se genera una disfunción mitocondrial [39].

1.2 MITOCONDRIA Y CÁNCER

Las primeras relaciones entre la importancia de la función mitocondrial y el desarrollo del proceso carcinogénico se basaron en los estudios realizados por Otto Warburg. Sus experimentos mostraron que las células tumorales sintetizaban elevadas concentraciones de lactato como producto de la vía fermentativa del piruvato. Ésta es una vía metabólica alternativa, pero las células neoplásicas la usan para la síntesis de energía en lugar de usar la vía oxidativa, incluso en condiciones normóxicas [40, 41]. Warburg, al observar una disminución de la vía oxidativa, sugirió que esta preferencia era causada por disfunciones a nivel mitocondrial y propuso que todos los tipos de cáncer dependían energéticamente de la glucólisis, siendo ésta quien aporta 50-70% de la demanda de ATP a las células tumorales (Fig. 4) [42-44].

Estudios posteriores establecieron que la disminución de la fosforilación oxidativa podía deberse a: 1) disminución en la actividad del transportador de piruvato; 2) por inhibición de la

piruvato deshidrogenasa (PDH); 3) disminución de la disponibilidad de equivalentes reductores generados en el ciclo de los ácidos tricarboxílicos debido a la baja actividad de la aconitasa e isocitrato-deshidrogenasa; así como por afectación de la 2-oxoglutarato deshidrogenasa o la citocromo *c* oxidasa a consecuencia de una disminución del pH; 4) bajo contenido mitocondrial; y 5) aumento del estrés oxidativo, por presencia de complejos respiratorios ineficientes [43].

Existe evidencia de que la reprogramación metabólica de las células cancerosas está relacionada con un estado redox alterado derivado del aumento en la concentración de ROS. Lo anterior genera daño e inestabilidad genómica que activan oncogenes o reprimen genes supresores de tumor favoreciendo la supervivencia y la proliferación de células tumorales y por ende la progresión del cáncer [45, 46].

Aunque la reprogramación metabólica en células tumorales es ahora considerada una característica del cáncer, se ha propuesto que la adquisición de alteraciones en el mtDNA impacta drásticamente en el sistema OXPHOs, convirtiéndose en un mecanismo implicado en la carcinogénesis [4, 42, 47].

La implementación de técnicas de secuenciación de nueva generación y el estudio de pacientes a gran escala han contribuido al conocimiento del panorama global de las variaciones en la secuencia del mtDNA, las cuales pueden ser germinales (heredables) o somáticas (adquiridas durante el desarrollo del individuo y no heredables). Además, se ha identificado una asociación entre estas variantes con procesos fisiológicos normales (polimorfismos) o patogénicos (mutaciones) y diferentes enfermedades, incluyendo el cáncer de mama [47-50].

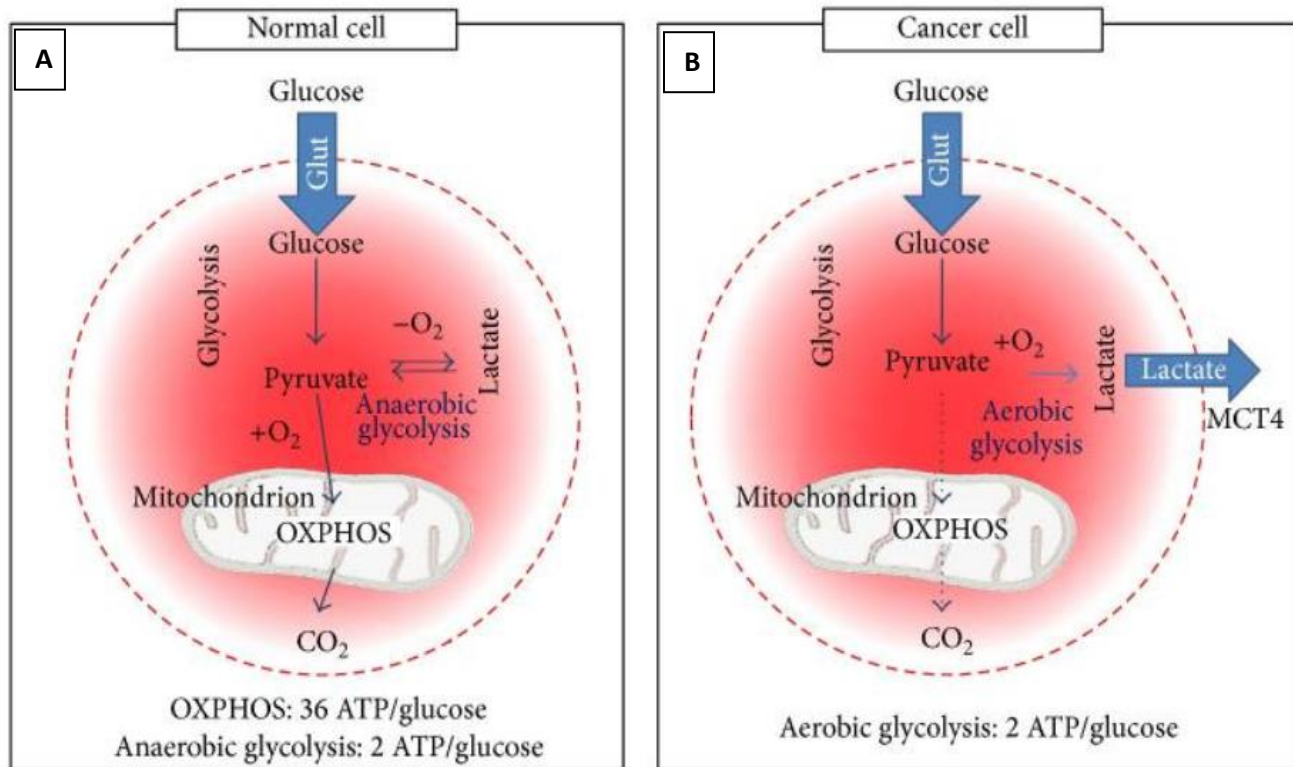


Figura 4. Efecto Warburg. Se esquematiza el metabolismo de una célula normal (A) y de una célula de cáncer (B). En condiciones normales, la glucosa entra a través del transportador Glut y es oxidado a piruvato. Cuando hay presencia de O₂, el piruvato entra a la mitocondria para generar equivalentes reductores que entran al sistema OXPHOs, generando una mayor cantidad de energía en forma de ATP. En caso de ausencia de O₂, el piruvato se oxida a lactato para restaurar los equivalentes reductores, produciendo una ganancia energética mucho menor. En una célula cancerosa (B), las deficiencias mitocondriales promueven que el piruvato sea oxidado a lactato aún en presencia de O₂, lo que generan un exceso de éste, el cual es excretado mediante el transportador de monocarboxilatos 4 (MCT4) [51].

1.3 CÁNCER DE MAMA

El cáncer de mama es el crecimiento acelerado y descontrolado de las células que conforman la glándula mamaria. Puede surgir, principalmente, en las células de los lobulillos (glándulas productoras de leche) o en las células de los conductos (que llevan la leche de los lobulillos hasta el pezón). Sin embargo, otros tipos celulares pertenecientes al tejido conjuntivo graso y fibroso no están excluidos. Además, cuando las células cancerosas están confinadas al sitio de origen (ductos o lobulillos), se trata de un cáncer no invasivo o *in situ*; sin embargo, también puede presentarse como un cáncer invasivo que eventualmente es capaz de generar metástasis e infiltrar otros órganos, cuando las células malignas se dispersan a los tejidos circundantes [52-55].

Se sabe que aproximadamente el 10% de los tumores mamarios tienen un componente hereditario, mientras que el 90% restante es el resultado de eventos esporádicos y multifactoriales. Actualmente se han identificado algunos factores de riesgo que están fuertemente asociados al padecimiento, tales como el género femenino, edad, estilo de vida (dieta, consumo de tabaco, alcohol, sedentarismo), obesidad, edad del primer embarazo, tiempo de lactancia, etnicidad, menopausia tardía, menarca temprana, una prolongada exposición a estrógenos y agentes mutagénicos, así como antecedentes familiares de cáncer. [52, 55, 56].

El cáncer de mama es un conjunto de enfermedades heterogéneas a nivel clínico, celular y molecular [53, 56]. Las causas concretas que dan origen a esta neoplasia no han sido completamente determinadas, sin embargo, se sabe que para su desarrollo, las alteraciones a nivel cromosómico, número de copias de genes, mutaciones puntuales, así como cambios en los niveles de expresión génica y fenómenos epigenéticos podrían ser un factor importante para su desarrollo [52, 54].

1.3.1 EPIDEMIOLOGÍA DEL CÁNCER DE MAMA

El cáncer de mama se ha convertido en un grave problema de salud mundial. Durante el año 2018, se estimaron 2.09 millones de casos y 627 mil muertes en el mundo a causa de esta entidad [53, 57]. Aunque en México no se tienen estudios epidemiológicos concluyentes, se estima que la incidencia y la mortalidad de esta enfermedad mantiene el mismo comportamiento que las cifras mundiales. Estudios epidemiológicos realizados en población mexicana, señalan que para el año 2016 se registraron 28.75 casos nuevos y 16 muertes por cada 100 mil mujeres [58].

Estudios realizados en pacientes pertenecientes al Instituto de Enfermedades de la Mama, FUCAM, señalaron que la edad promedio al diagnóstico fue de 51 años, un dato que varía con respecto a otras poblaciones donde ésta es de 63 años. Además, la mayor parte de los casos se diagnostican en etapas localmente avanzadas (IIB, IIA, IIB y IIIC), de acuerdo con el estadio asignado por la Unión Internacional Contra el Cáncer (UICC) y la American Joint Committee on Cancer (AJCC), lo que contribuye a que la tasa de mortalidad siga a la alza [59].

1.3.2 CLASIFICACIÓN MOLECULAR DEL CÁNCER DE MAMA

Dada la elevada heterogeneidad del cáncer de mama, muchas investigaciones se han enfocado en el estudio de diferentes moléculas que permitan clasificar a este padecimiento considerando sus características biológicas y clínicas, y que tengan impacto en el entendimiento del comportamiento biológico, diagnóstico y respuesta a tratamiento [60].

Los primeros abordajes para clasificar el cáncer de mama se basaron en pruebas inmunohistoquímicas mediante la evaluación de la expresión de proteínas como el receptor de estrógenos (RE), el receptor de progesterona (RP) y el receptor del factor de crecimiento epidérmico humano 2 (HER2). La relevancia de esta clasificación es que las pacientes pueden ser beneficiadas con algún tratamiento específico de acuerdo con la positividad de estos receptores. Por ejemplo, las pacientes con tumores positivos a receptores hormonales (RE+, RP+) son candidatas a tratamientos dirigidos contra estos receptores y tienen un mejor pronóstico a diferencia de las pacientes con tumores negativos a ellos [60, 61].

Con el advenimiento de las tecnologías de análisis masivo como el estudio del transcriptoma de tumores de mama, se ha permitido establecer una clasificación más precisa del tumor. Así, basado en los patrones de expresión génica, se han establecido cuatro principales subtipos moleculares de cáncer de mama: luminal A (LA), luminal B (LB), enriquecido en HER2 (H2) y basal (B). Esta clasificación, tiene un alto grado de correlación con el diagnóstico inmunohistoquímico, el pronóstico y la respuesta a tratamiento (Tabla 1). Cada uno de estos subtipos tiene características clínicas y biológicas que los diferencia entre ellos [59, 61].

1.3.2.1 Subtipo luminal A

El subtipo LA es el más común, representa del 30 – 70% de los tumores de mama [56]. Son tumores positivos a RE y RP, pero son HER2 negativos. Presentan una elevada expresión de genes relacionados con el RE y baja expresión de genes relacionados a proliferación celular. Presentan células bien diferenciadas y están asociados a un pronóstico favorable y mayor tasa de sobrevida. Se ha visto que este subtipo responde favorablemente a terapia antihormonal [61, 62].

1.3.2.2 Subtipo luminal B

La frecuencia del subtipo LB es del 10 – 20% y son más agresivos que los luminales A [56]. Este subtipo se caracteriza por tener células moderadamente diferenciadas, ser RE positivo, con baja (<20%) o nula expresión de RP y pueden ser HER2 positivos. Además, tienen una elevada expresión de genes relacionados a proliferación celular. Su tratamiento consiste en terapia antihormonal y quimioterapia neoadyuvante [61, 62].

1.3.2.3 Subtipo positivo a HER2

Los tumores enriquecidos en HER2 representan aproximadamente el 5 - 15% de los tumores de mama [56], se caracteriza por tener una sobreexpresión de HER2 y de genes relacionados a la vía, además de no expresan RE y RP. Otros genes que se pueden expresar en este subtipo son *GRB7* y *PGAP3*, y el 40-80% de los tumores portan mutaciones en *TP53*. Los HER2+ son tumores poco diferenciados que están asociados a un pronóstico desfavorable; sin embargo, la terapia con anticuerpos monoclonales dirigidos contra la proteína HER2 ha mejorado la sobrevida de estas pacientes [61-63].

1.3.2.4 Subtipo basal

La frecuencia de este subtipo oscila entre el 15 - 20% de los tumores [56], no expresan RE, RP ni HER2, por lo que entran dentro del grupo de los tumores triples negativos (TN) cuando son clasificados con el empleo de los marcadores inmunohistoquímicos; sin embargo, se caracterizan por expresar citoqueratinas (CK5,6,14,17), p-cadherina, el receptor del factor de crecimiento epidérmico (*EGFR*), entre otros genes del epitelio basal. Aproximadamente del 11 - 42% de estos tumores presentan mutaciones en genes *BRCA1* (9 - 100%) y *BRCA2* (2 - 12%) [64, 65]. Son tumores de poco diferenciados, tienen un curso clínico agresivo, presentan el peor pronóstico y actualmente carecen de un tratamiento específico, siendo la quimioterapia la mejor opción para tratarlos. A diferencia de los otros subtipos, este tipo de tumor muestra una distribución diferencial entre las poblaciones y su frecuencia se ve incrementada en mujeres afroamericanas y de edades más jóvenes [61-63].

Tabla 1. Clasificación inmunohistoquímica de los tumores de cáncer de mama.

Subtipo molecular	IHQ	Frecuencia (%) *	Pronóstico	Tratamiento
Luminal A	RE+ RP>20% HER2- Ki67-	65.73	Bueno	Terapia antihormonal
Luminal B	RE+ RP<20% HER2+/- Ki67+	10.91	Intermedio	Terapia antihormonal + quimioterapia neoadyuvante
HER2	RE - RP - HER2 +	8.72	Malo Probabilidad de recurrencia elevada	Anticuerpos anti-HER2
Basal	RE - RP - HER2 -	14.63	Malo Probabilidad de recurrencia alta	Quimioterapia

IHQ: marcadores inmunohistoquímicos; RE: receptor de estrógenos; RP: receptor de progesterona; HER2; receptor del factor de crecimiento epidérmico humano 2; Ki67: índice de proliferación celular.
* Frecuencias pertenecientes a población mexicana [66] [59].

A partir de la identificación de genes diferencialmente expresados entre los distintos subtipos del tumor, se han desarrollado plataformas como Mammaprint y Oncotype Dx, los cuales tratan de predecir el pronóstico y el riesgo a recaída de las pacientes. Sin embargo, muchas de estas tecnologías son de difícil acceso y por lo tanto poco aplicadas en el sector salud, por lo que la exploración de diversas moléculas que permitan clasificar, mejorar el diagnóstico y tratamiento continúan siendo de gran interés [62]. Además, el comportamiento de los tumores sigue siendo heterogéneo y se desconoce aún gran parte de los mecanismos biológicos implicados el proceso carcinogénico [62].

En estudio del cáncer de mama, la mayor parte de las líneas de investigación de hoy en día se centran en el estudio del DNA a nivel genómico, transcriptómico y epigenético. Sin embargo, un área poco estudiada es la de los productos codificados por el genoma mitocondrial, de los cuales existe evidencia que están involucrados en diferentes funciones celulares como lo son la producción de energía, apoptosis y vías de señalización relacionadas con proliferación celular.

1.3.3 VARIANTES MITOCONDRIALES EN CÁNCER DE MAMA

En cáncer de mama, el estudio de las alteraciones del mtDNA ha identificado polimorfismos donde abundan las variantes de un solo nucleótido (SNVs del inglés, single nucleotide variant), inserciones/deleciones (INDELS); mutaciones de una sola base, grandes deleciones y cambios en el número de copias del mtDNA [6, 48]. Estas alteraciones se encuentran distribuidas en toda la molécula de mtDNA; sin embargo, la región D-Loop puede acumular aún más debido a su alta tasa de mutación (Fig. 5) [6, 67, 68].

Algunos de los polimorfismos de un solo nucleótido (SNPs, del inglés single nucleotide polymorphism) localizados en la región D-Loop que incrementan el riesgo a padecer cáncer de mama son T16189C, C16207T, T16519C, T239C y A263G. Sin embargo, la variante más comúnmente asociada a esta enfermedad es el polimorfismo D310, una interrupción por una T en la posición 310 de la secuencia de citocinas que va de la base 303 a la 315 dentro de la región D-Loop. D310 se ha encontrado en el 92% de las pacientes, sugiriéndose que tiene una

participación en la oncogénesis mamaria y en la progresión tumoral [69-75]. De igual forma se ha propuesto que los SNPs de la región D-Loop contribuyen a la enfermedad a través de desregular los mecanismos de replicación y transcripción del genoma [67, 68, 73, 76, 77].

Los SNPs asociados al riesgo a padecer tumores de mama son A4727G, G9055A, C6296A, G9947A, A8860G, A10044G, A10283G, T11233C y C11503T, siendo el SNP G10398A (localizado en el gen *ND3*) el más estudiado en diversas cohortes [67, 68, 78, 79]. Se ha reportado que el alelo A de esta variante confiere susceptibilidad a cáncer de mama en población afroamericana (OR, 1.60; $p= 0.013$), china (OR, 1.77; $p= 0.05$), europea-americana (OR, 1.77; $p= 0.01$), India (OR, 1.73; $p= 0.01$) y malaya (OR, 2.29; $p = 0.007$) [67, 80-84].

Por otro lado, se ha descrito que el 73.7% de las pacientes portan mutaciones somáticas, de las cuales, una gran parte están presentes en un solo individuo y son encontradas en estado heteroplásmico que va del 5.6 al 97.4% [85-87]. El porcentaje de mutaciones en la región D-Loop en tumores de mama varía entre las poblaciones; por ejemplo, Shen et al, encontró que el 15.6% de las pacientes Taiwanesas presentan mutaciones en esta región algunas de las cuales correlacionan con metástasis [88]. Los estudios de McMahon y LaFramboise reportaron que el 55.3% de las mutaciones somáticas estaban en genes codificantes de proteína, siendo el *ND5* el que presentó una mayor carga mutacional. Se ha reportado que el 63% de las mutaciones codificantes son potencialmente causativas de disfunciones mitocondriales y que además correlacionan con una edad de inicio de cáncer de mama mayor (≥ 50 años), un elevado estadio tumoral, un pobre grado de diferenciación y contradictoriamente, una mejor sobrevida [68, 86, 89].

1.3.4 HAPLOGRUPOS EN CÁNCER DE MAMA

Como se ha mencionado, un haplogrupo es un conjunto de variantes del genoma mitocondrial que se segregan juntos. La coexistencia de diferentes variantes puede inducir cambios biológicos que dan lugar a un fenotipo normal o patogénico, como el cáncer de mama. Se ha sugerido que los cambios en la secuencia del mtDNA impactan en el sistema OXPHOs confiriendo adaptación a la célula cancerosa [47]. Estudios de asociación han observado que algunos haplogrupos modifican el riesgo de padecer cáncer de mama. Así, se ha reportado que

los haplogrupos D5 (OR, 3.11; $p=0.03$) y M (OR, 1.77; $p=0.04$) son factores de riesgo para esta neoplasia en población china, así como haplogrupo K (OR, 3.03; $p=0.0004$) en población europea-americana; además existe una sobrerepresentación del haplogrupo I en mujeres polacas y el haplogrupo X en italianas portadoras de mutaciones en *BRCA1*. En contraste, los haplogrupos H (OR, 0.636; $p=0.019$) en población polaca, el haplogrupo U (OR, 0.37; $p=0.0023$) en mujeres europeo-americanas y el haplogrupo T1a1 (OR, 0.62; $p=0.03$) en mujeres portadoras de mutaciones en *BRCA2*, son factores protectores [48, 90, 91].

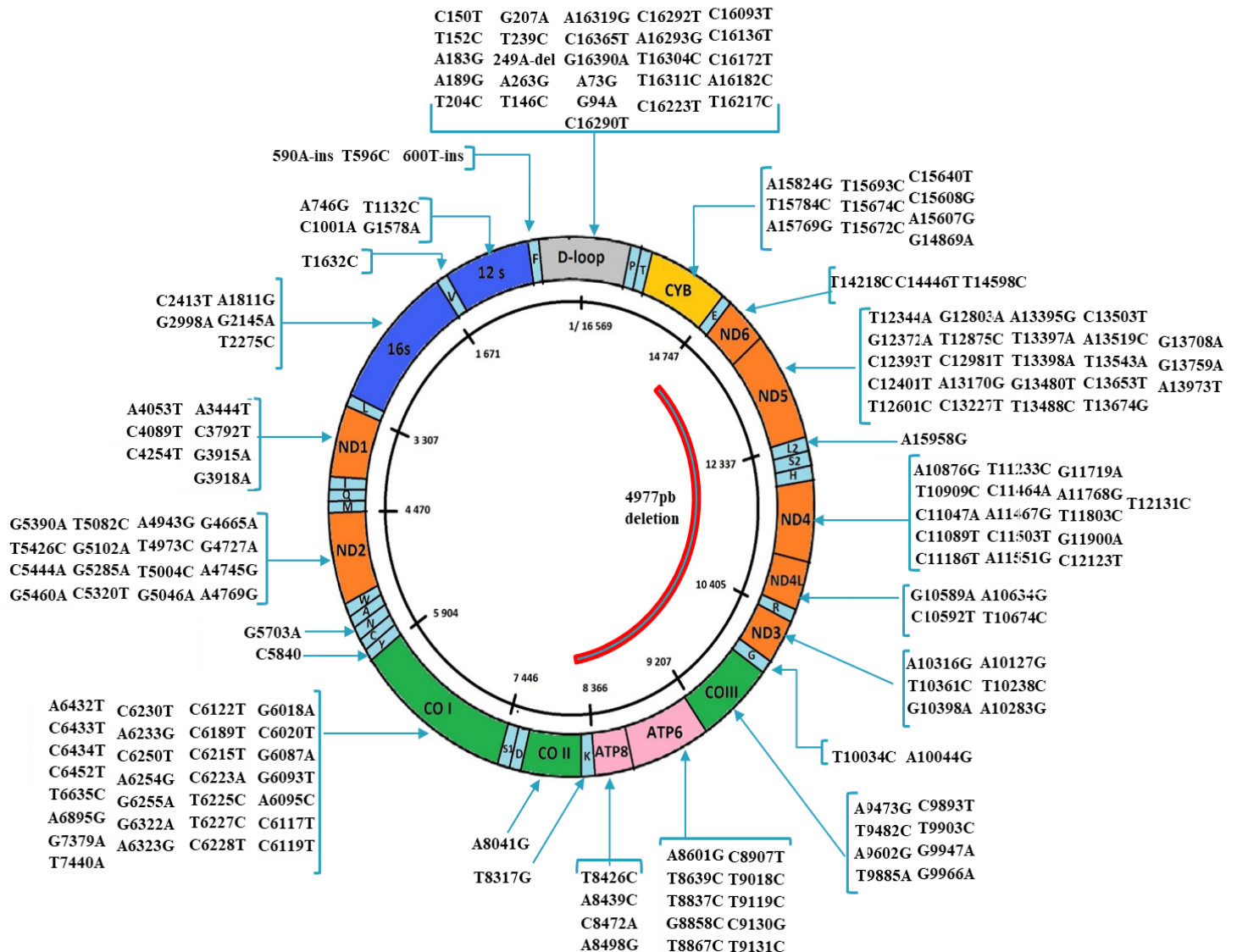


Figura 5. Alteraciones en mtDNA en cáncer de mama. El mapa del mtDNA señala los SNPs e indels más frecuentemente encontradas en tumores de mama. Datos de MITOMAP [92].

2 JUSTIFICACIÓN

El cáncer de mama es un grave problema de salud a nivel mundial y en México constituye la primera causa de muerte por cáncer en población femenina. Aunque actualmente existe una gran cantidad de estudios que se centran en dilucidar los mecanismos implicados en la biología de los tumores, la mayor parte de ellos se enfocan en el estudio de moléculas y procesos regulados por el genoma nuclear. Sin embargo, otra parte del genoma corresponde al DNA localizado en las mitocondrias, un organelo de gran importancia para el funcionamiento celular ya que se encarga de cubrir los requerimientos energéticos necesarios y tiene una participación relevante en procesos biológicos como la apoptosis y la senescencia. De hecho, se ha descrito que alteraciones en el genoma mitocondrial contribuyen al desarrollo de diversas enfermedades metabólicas, neurodegenerativas y cáncer. Las alteraciones descritas en el mtDNA son en número de copias, mutaciones puntuales y deleciones y se ha reportado que algunas de ellas se asocian al riesgo a padecer cáncer de mama, a una pobre sobrevida libre de enfermedad y a la respuesta al tratamiento. Sin embargo, aunque se han identificado un amplio número de variantes en el genoma mitocondrial de tumores de mama, su asociación con la enfermedad es aún controversial; además de que el efecto funcional y la participación que puedan tener éstas en el proceso carcinogénico es desconocido. Adicionalmente, se ha observado que el tipo de mutaciones pueden variar de una población a otra, sugiriendo la posibilidad de que existan mutaciones que sean exclusivas de cada una de ellas.

3 OBJETIVOS

3.1 OBJETIVO GENERAL

Identificar las mutaciones mitocondriales de los genes de la cadena respiratoria presentes en pacientes con cáncer de mama.

3.2 OBJETIVOS PARTICULARES

1. Secuenciar el mtDNA de muestras de sangre periférica y de tumores de pacientes con cáncer de mama.
2. Identificar las mutaciones somáticas en el mtDNA de tumores de mama.
3. Determinar si existe un perfil de mutaciones característico de cada subtipo molecular de cáncer de mama.

4 ESTRATEGIA EXPERIMENTAL

4.1 POBLACIÓN DE ESTUDIO

Se incluyeron muestras pareadas tumor-sangre periférica de 92 mujeres con diagnóstico clínico e histopatológico de cáncer de mama atendidas en el Instituto de Enfermedades de la Mama FUCAM, mayores de 18 años, que no presentaron tratamiento antineoplásico y que aceptaron participar en el estudio mediante firma de carta de consentimiento informado. Se excluyeron del estudio aquellas muestras de tumores asociados a síndromes genéticos, secundarios o de pacientes que refirieran enfermedades infecto-contagiosas.

4.2 OBTENCIÓN DE MUESTRA

Las muestras de sangre periférica fueron obtenidas por venopunción con el sistema Vacutainer al momento de la canalización de las pacientes. Los tubos de sangre se centrifugaron a 2500 rpm para separar el plasma y el paquete de células leucocitarias, las cuales fueron almacenadas a -80°C hasta el momento de la extracción de ácidos nucleicos.

Las muestras de tejido tumoral se obtuvieron al momento de la resección quirúrgica del tumor y con la ayuda del personal del área de patología se seleccionaron las zonas con alto contenido tumoral. El tumor fue dividido en dos partes, una parte fue congelada con nitrógeno líquido y almacenada a -80°C; la segunda se fijó en formol para su inclusión en parafina. A partir de la muestra incluida en parafina, se realizaron cortes de 3 µm con un microtomo. Los cortes histológicos fueron teñidos con hematoxilina-eosina para determinar el porcentaje de células tumorales. Aquellas muestras de tumor que presentaron un contenido de células tumorales mayor al 80% fueron seleccionadas para el estudio.

El subtipo molecular definido por las firmas moleculares obtenidas por microarreglos de expresión (PAM50) estuvo disponible para 55 de los tumores, por lo que se la clasificación molecular de fue asignada con los resultados de los marcadores inmunohistoquímicos (RE, RP, HER2) de acuerdo con los parámetros indicados por el Consenso Mexicano sobre diagnóstico y tratamiento del cáncer mamario [66].

4.3 ANÁLISIS MOLECULAR

4.3.1 EXTRACCIÓN DE ÁCIDOS NUCLEICOS

4.3.1.1 Obtención de DNA de sangre periférica

Para la extracción de los ácidos nucleicos a partir de sangre periférica, se usó el Maxikit de Qiagen, siguiendo las especificaciones del proveedor. Brevemente, las células leucocitarias fueron homogenizadas con 500µL de proteasa en un tubo Falcon de 50 ml y el volumen fue ajustado a 5 mL con PBS 1X (Phosphate buffered saline, pH=7.4, Gibco™). Enseguida se añadieron 6 mL de amortiguador AL, se mezcló vigorosamente y se incubó a 70°C por 10 minutos. Posteriormente se añadieron 5mL de etanol absoluto y la mezcla fue transferida a la columna QIAmp Maxi que se centrifugó a 1850 x *g* (3000 rpm) por 3 min. Se descartó el filtrado y los lavados fueron realizados añadiendo 5mL de amortiguador AW1 y centrifugando a 4500 x *g* (5000 rpm) por 1 min. Posteriormente se añadieron 5mL de amortiguador AW2 y se centrifugó a 4500 x *g* (5000 rpm) por 15 minutos. La columna fue transferida a un tubo nuevo donde se añadieron 500µL de amortiguador TE (Tris-EDTA, 1X pH 8.0, Invitrogen™) y se centrifugó a 4500 x *g* (5000 rpm) por 5 min. El DNA eluido fue almacenado y cuantificado espectrofotométricamente en NanoDrop 1000 (Thermo Fisher Scientific) a una longitud de onda de 260 nm.

4.3.1.2 Obtención de DNA de tejido tumoral

El DNA de tejido tumoral se extrajo usando el sistema AllPrep DNA/RNA MiniKit (Qiagen). En un paso inicial, las muestras de tejido fresco fueron homogenizadas usando el TissueRuptor y 600µL de amortiguador de lisis (RLT), posteriormente el homogenizado fue centrifugado por 3 minutos a $\geq 8000 \times g$ (≥ 10000 rpm). El sobrenadante se transfirió a una columna All Prep DNA, la cual se centrifugó por 30 segundos a $\geq 8000 \times g$ (≥ 10000 rpm). Se descartó el eluyente y la columna se transfirió a un nuevo tubo de recolección. Se adicionaron 500 µL de amortiguador AW1, se centrifugó durante 15 segundos a $\geq 8000 \times g$ (≥ 10000 rpm) y se descartó el filtrado. Enseguida, a la columna se le adicionaron 500µL de amortiguador AW2 y se centrifugó durante 2 minutos a $\geq 8000 \times g$ (≥ 10000 rpm). Finalmente, la columna fue transferida a otro tubo Eppendorf de recolección y se adicionaron 100 µL de amortiguador TE (Tris-EDTA, 1X pH 8.0,

Invitrogen™) sobre la membrana, se incubó a temperatura ambiente por un minuto y se centrifugó 1 minuto a $\geq 8000 \times g$ (≥ 10000 rpm). El DNA eluído se cuantificó espectrofotométricamente y se almacenó a -20°C .

4.3.2 EVALUACIÓN DE PUREZA E INTEGRIDAD DE MUESTRAS

Las muestras de DNA fueron cuantificadas mediante espectrofotometría en un NanoDrop ND-1000 a una longitud de onda de 260 nm y su integridad fue evaluada mediante electroforesis en gel de agarosa al 1% teñido con SYBR® Gold (diluido 1:500 μL de agua de grado molecular), el cual fue visualizado a una longitud de onda de 435 nm.

4.3.3 ENRIQUECIMIENTO DEL DNA MITOCONDRIAL MEDIANTE PCR LARGA

El mtDNA fue enriquecido con la generación de dos amplicones sobrelapados de aproximadamente 8.5 kb obtenidos a través de una PCR larga. La reacción fue realizada mediante el Kit SequalPrep™ Long PCR, usando 2 pares de primers (Mito1 y Mito2) diseñados con la finalidad de cubrir la totalidad del genoma mitocondrial.

Para realizar la PCR larga, las muestras de DNA se diluyeron a 5 ng/ μL . A 1 μL de DNA se le añadieron 19 μL de master mix. Los reactivos ocupados y el programa de termociclado están descritos en la Tabla 3 y Tabla 4, respectivamente.

Al finalizar la PCR, los productos fueron analizados con una electroforesis en gel de agarosa E-Gel® SizeSelect™ al 0.8 % para verificar la presencia de los amplicones de 8.5kb. Se mezclaron 5 μL de producto de PCR con 15 μL de Intercalante SYBR® Gold (diluido 1:500 μL de agua de grado molecular). El gel fue eluído a 120 Volts durante 45 minutos.

Tabla 2. Reactivos para la PCR larga

Componente	Master Mix A Volumen (μL)	Master Mix B Volumen (μL)
SequalPrep™ Buffer 10X	2	2
DMSO	0.4	0.4
SequalPrep™ Enhancer B	1	1
SequalPrep™ Polimerasa	0.36	0.36
Agua libre de DNAsas	14.24	14.24
Primer Mito1F	0.5	-
Primer Mito1R	0.5	-
Primer Mito2 F	-	0.5
Primer Mito2 R	-	0.5
DNA	1	1
Total	20	20

Tabla 3. Condiciones de termociclado para PCR larga.

Temperatura	Tiempo	Estado
94°C	2 min	Hold
94°C	10 s	30 ciclos
60°C	30 s	
68°C	8 min	
72°C	5 min	Extensión final
4°C	∞	Hold

4.3.4 PURIFICACIÓN DE PRODUCTOS DE 8.5 Kb

Para la purificación de los productos se empleó el kit Agencourt® AMPure® XP que se basa en el uso de perlas magnéticas que se unen al DNA. A los 15μL de producto de la PCR, se añadieron 27μL de perlas magnéticas (de acuerdo con las especificaciones del proveedor de añadir 1.8 μL de perlas magnéticas por 1 μL de producto) y se incubó a temperatura ambiente por 15 minutos. Los tubos que contenían el DNA y las perlas magnéticas se colocaron en la gradilla magnética por 5 minutos, posteriormente se retiró el sobrenadante. Las perlas se lavaron con 200 μL de etanol al 70% recién preparado, se retiró el sobrenadante y se dejó secar por 3 minutos. La elución del DNA se hizo con 40μL de agua de grado molecular, posteriormente se cuantificó en un NanoDrop.

4.3.5 POOL DE AMPLICONES

Los amplicones de la PCR1 y PCR2 fueron mezclados a una concentración de 100 ng/ μ L de cada una para asegurar una cobertura total del genoma mitocondrial en la secuenciación.

4.3.6 FRAGMENTACIÓN DE LOS AMPLICONES

Para realizar la fragmentación de los amplicones, los pools fueron llevados a un volumen final de 130 μ L con amortiguador Low TE y se transfirieron a un microtubo para sonicación. Los amplicones fueron fragmentados en tamaños de aproximadamente 200 pb usando el ultrasonificador Covaris®, para lo cual se corrió el programa de fragmentación “200pb mitDNA” verificando previamente las condiciones descritas en la Tabla 4. Al finalizar, los productos fueron analizados con una electroforesis en gel de agarosa E-Gel® SizeSelect™ al 2% empleándose como intercalante SYBR® Gold. El patrón adecuado de la electroforesis correspondió a un barrido localizado entre las bandas de 100 a 500 pb del marcador de pesos.

Tabla 4. Condiciones de sonicación para la fragmentación del mtDNA.

Propiedad	Especificación
<i>Número de ciclos</i>	6
<i>Temperatura de baño</i>	5°C
<i>Límite de temperatura de baño</i>	30°C
<i>Modo</i>	Frequency sweeping
<i>Función de prueba de calidad de agua</i>	OFF
<i>Ciclo de trabajo</i>	10%
<i>Intensidad</i>	5
<i>Ciclo/Horst</i>	100
<i>Tiempo</i>	60 s

4.3.7 SECUENCIACIÓN DEL DNA MITOCONDRIAL MEDIANTE PLATAFORMA MiSeq

Illumina

Los amplicones generados se emplearon para formar las bibliotecas genómicas, para esto se usó el kit Nextera DNA Flex Library Prep Kit (Illumina, San Diego, CA, USA). Este kit se basa en el uso de transposomas para la reparación de extremos y ligación de adaptadores. Se realizó una

secuenciación basada en síntesis de lecturas pareadas con longitud de 150 pb por 300 ciclos con el sistema MiSeq (Illumina, San Diego, CA, USA). Los protocolos fueron implementados en la Unidad de Secuenciación del Instituto Nacional de Medicina Genómica.

4.4 ANÁLISIS DE DATOS

Las secuencias resultantes fueron procesadas a través de la adaptación de los flujos de trabajo establecidos por el Genome Analysis Tool Kit (GATK, Broad Institute) [93] y el identificador de mutaciones somáticas Strelka2 de Illumina (Fig. 6) [94].

4.4.1 ANÁLISIS DE CALIDAD

Las lecturas fueron analizadas con FastQC (Braham Bioinformatics) para conocer su calidad [95]. Se aceptaron aquellas lecturas cuyos llamados de bases tuvieron una escala Phred mayor a 25. Aquellas lecturas que no presentaron esta condición fueron procesadas con TrimGalore (BrahamBioinformatics) para eliminar las bases de mala calidad [96]. Posteriormente, se usó el programa MultiQC para generar un reporte global de las secuencias de todas las muestras [97].

4.4.2 ALINEAMIENTO DE LAS LECTURAS

Se usó el software Burrows-Wheeler-Aligner (BWA) para alinear las lecturas al genoma mitocondrial de referencia. En particular se empleó el algoritmo BWA-MEM que es recomendado para realizar consultas de lecturas de tamaño mayor a 70 pb, con una velocidad mayor y de alta calidad [98]. La secuencia del cromosoma mitocondrial (ChrM) fue extraída de la versión hg38 del genoma humano de referencia mediante scripts diseñados en lenguaje de programación Python [99].

4.4.3 PRE-PROCESAMIENTO DE ALINEAMIENTOS

4.4.3.1 Marcado de duplicados

Para marcar los duplicados, se empleó la función MarkDuplicates del repertorio de herramientas de Picard, el cual se emplea cuando las secuencias provienen de una PCR y consiste en la identificación de lecturas que contengan errores generados durante la amplificación [100]. Las lecturas marcadas son ignoradas durante el procedimiento de descubrimiento de variantes.

4.4.3.2 Indexación

La indexación consiste en ordenar los alineamientos obtenidos, que se encuentran de forma aleatoria, los cuales son ordenados con respecto a su posición en el genoma. Este paso fue realizado con Sort, una herramienta incluida en los algoritmos de SAMtools [101].

4.4.3.3 Recalibración de calidad de bases

Durante la recalibración de bases se aplica aprendizaje automatizado para detectar y corregir patrones de errores sistemáticos en los puntajes de calidad de base emitidos por el secuenciador. En este paso se usaron las herramientas BaseRecalibrator y ApplyRecalibrator incluidas en la versión 4.0.3.0 del GATK [93].

4.4.4 IDENTIFICACIÓN DE VARIANTES EN TUMORES

El nombramiento de las variantes en tumor fue realizada con los algoritmos de Strelka2 (Illumina Inc.), que es un detector de SNVs, e INDELS a partir de secuencias pareadas de tejido normal con tejido tumoral [94]. Strelka2 emplea un enfoque Bayesiano y un modelo de haplotipo escalonado para la identificación de la variación en donde la muestra normal es representada como una mezcla de variación germinal y ruido, mientras que la muestra de tumor como una mezcla de la muestra normal con variación somática. Los pipelines documentados fueron adaptados para realizar la detección de las variantes mitocondriales.

4.4.5 DETECCIÓN DE HETEROPLASMIA

Para la evaluación de heteroplasmia, se calculó la fracción alélica de las mutaciones somáticas con las fórmulas documentadas por Strelka2. Para ello, se realizó un script con

lenguaje de programación Python [99] que conjunta todas las mutaciones diferentes contenidas en los archivos Variant Calling Format (VCF) de todas las muestras y determina la fracción alélica de cada una de ellas. El VCF construido fue procesado en lenguaje de programación R [102], donde se realizó el filtrado de aquellas mutaciones que pasaron los filtros de calidad del llamado de variante.

4.4.6 IDENTIFICACIÓN DE MUTACIONES SOMÁTICAS

Se definió como mutación somática cuando una variante identificada en tumor presentó una frecuencia alélica (FA) en sangre periférica menor a 0.001, considerando que a este punto de corte y a una profundidad de 300X, la probabilidad de que la variante exista en sangre es casi nula (no se tienen lecturas suficientes para hacer un correcto nombramiento de la variante). El filtrado de estas variantes fue realizado con la generación de scripts basados en lenguaje de programación R.

4.4.7 ANOTACIÓN Y PREDICCIÓN DEL EFECTO FUNCIONAL

Las mutaciones filtradas fueron procesadas con la versión en línea del Variant Effect Predictor (VEP), una herramienta de Ensembl que asigna el gen a la posición de las mutaciones; además de inferir el efecto funcional y su consecuencia [103]. Una vez realizada la anotación, se filtraron aquellas mutaciones presentes en los genes codificantes del mtDNA.

4.4.8 IDENTIFICACIÓN DE MUTACIONES SOMÁTICAS NUEVAS Y ASOCIADAS A CÁNCER

Las mutaciones somáticas fueron buscadas en la base de datos de MITOMAP, una base de datos que contiene los registros de la variación identificada en el mtDNA [92] para identificar aquellas que tuvieran una asociación previa con cáncer de mama o algún otro tipo de cáncer. Además, se identificaron las mutaciones no registradas por ser aquellas potencialmente nuevas.

4.4.9 IDENTIFICACIÓN DE PERFILES MUTACIONALES POR SUBTIPO MOLECULAR

Para conocer la presencia de un perfil mutacional que caracterice a cada subtipo molecular, se identificaron las mutaciones somáticas exclusivas de cada uno de ellos. La distribución de

las mutaciones fue realizada para los 90 tumores clasificados con marcadores inmunohistoquímicos y para los 55 con firmas de expresión por microarreglos (PAM50) por separado.

4.4.10 ANÁLISIS DE HAPLOGRUPOS

Se emplearon los datos obtenidos de la secuenciación del mtDNA de sangre periférica para la determinación de haplogrupos. El versión en línea de Haplogrep2 fue usado, el cual calcula un peso filogenético que corresponde a la aparición de la variante en Phylotree, una base de datos de las variantes que componen a cada haplogrupo [[104](#), [105](#)],

4.4.11 ANÁLISIS ESTADÍSTICOS

Las frecuencias mutacionales fueron determinadas por conteos simples y la significancia estadística de los datos fue estimada con la prueba de chi-cuadrada. Los conteos y los gráficos fueron elaborados empleando scripts basados en lenguaje de programación R [[102](#)]. Los scripts generados se encuentran en proceso de subirse en un repositorio online (GitHub).

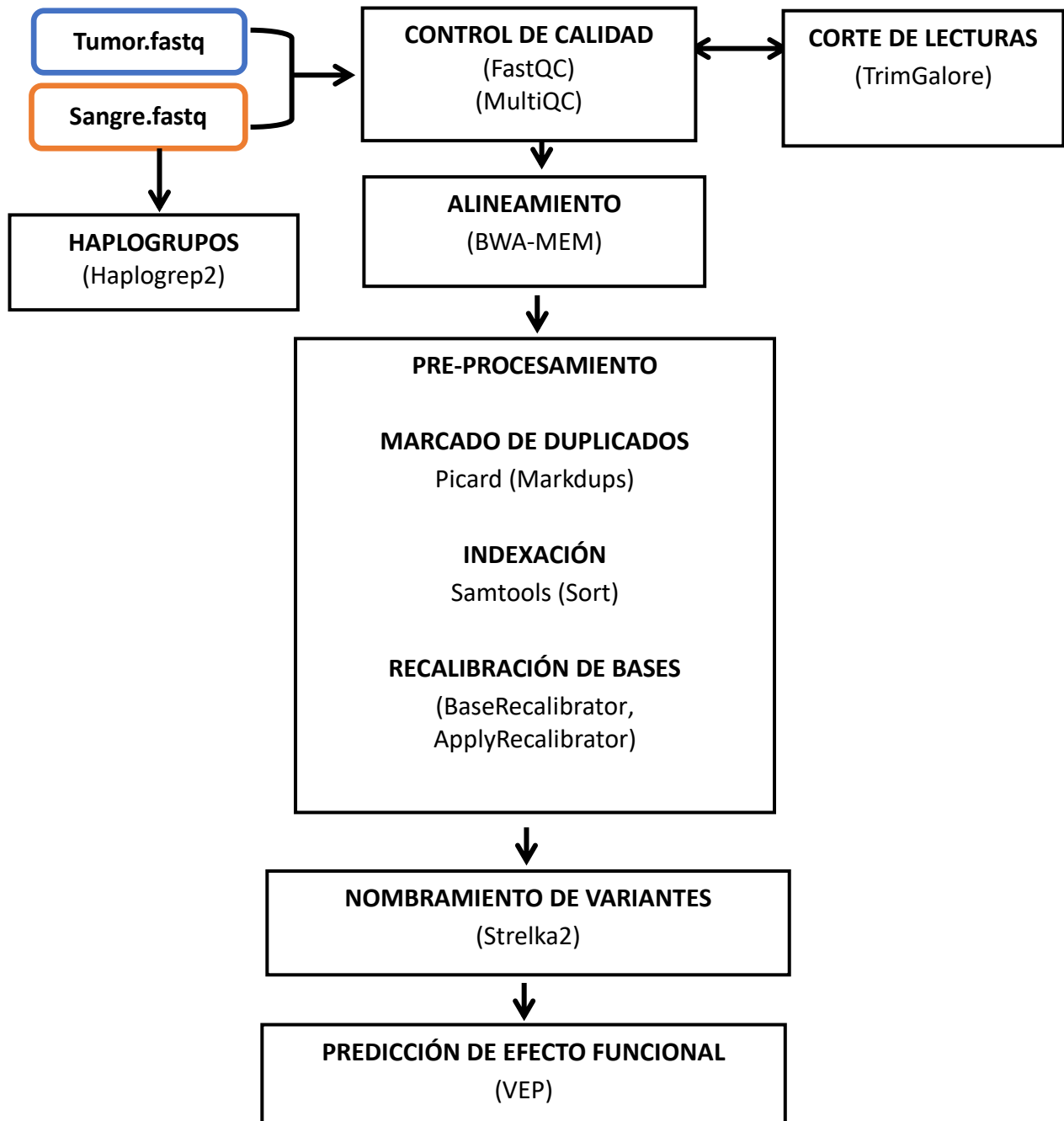


Figura 6. Análisis bioinformático de variantes mitocondriales. El flujo de trabajo consistió en evaluar la calidad de las lecturas y cortar aquellas bases con mala calidad de llamado ($Phred < 25$), realizar el alineamiento con el genoma de referencia (ChrM) y reevaluar el nombramiento de bases para la identificación de variantes tumorales a partir de la comparación entre la secuencia de sangre y tumor. Se realizó la predicción del efecto funcional de las variantes presentes solo en tumor y finalmente se asignó el haplogrupo a partir de la secuencia de sangre periférica.

5 RESULTADOS

5.1 CARACTERÍSTICAS CLÍNICAS DE LA POBLACIÓN DE ESTUDIO

Se incluyeron 92 muestras pareadas (tumor-sangre periférica) de pacientes con diagnóstico clínico e inmunohistoquímico de cáncer de mama pertenecientes al biobanco del Laboratorio de Genómica del Cáncer y captadas en el Instituto de Enfermedades de la Mama FUCAM. El intervalo de edad de las pacientes fue de 33 a 92 años con un promedio de edad de 53.83 \pm 11.5 años. El subtipo molecular fue asignado para 90 muestras, siendo el subtipo LA el más frecuente (61.9%), seguido del LB (24%), TN (6.5%) y por último del H2 (5.4%) (Tabla 5).

Tabla 5. Distribución de los subtipos moleculares en la población de estudio.

Subtipo molecular	Frecuencia n (%)
Luminal A	57 (61.9)
Luminal B	22 (24.0)
HER2	5 (5.4)
Triple Negativo	6 (6.5)
No asignado	2 (2.2)

5.2 NÚMERO Y DISTRIBUCIÓN DE LAS VARIANTES EN EL mtDNA DE MUJERES CON CÁNCER DE MAMA

Después de comparar las secuencias derivadas tanto de sangre periférica como las de tumor con el genoma de referencia el llamado de variantes realizado con el algoritmo Strelka2 identificó un total 548 variantes en las muestras de sangre y 707 variantes presentes en las muestras de tumor que pasaron los controles de calidad.

En relación con las variantes detectadas en tumor, el 97% fueron SNVs, el 2% fueron deleciones y el 1% inserciones. Estas variantes se distribuyeron a lo largo de todo el mtDNA, siendo los genes *ND5* (n=67), *16S* (n=60), *ND4* (n=58), *CO1* (n=56), *12S* (n=48), y la región D-Loop (n=53) los que presentaron mayor número de variantes en comparación con el resto de los genes. Para descartar que el total de variantes por gen estuviera influenciado por la longitud de éste, se estimó el número por kilobase. En este análisis, se observó que el D-Loop tiene una mayor tasa de variación (92.7 vars/kb); así como los genes *ATP6* (57.4 vars/kb), *ND3* (63.8 vars/kb), *12S* (50.4 vars/kb) y los tRNA-R (109.4 vars/kb), tRNA-C (76.9 vars/kb) y tRNA-Q (56.3 vars/kb) (Fig. 6). Es de hacer notar que los genes tRNA-V, -L1, -N, -S1 y -G no mostraron ninguna variante.

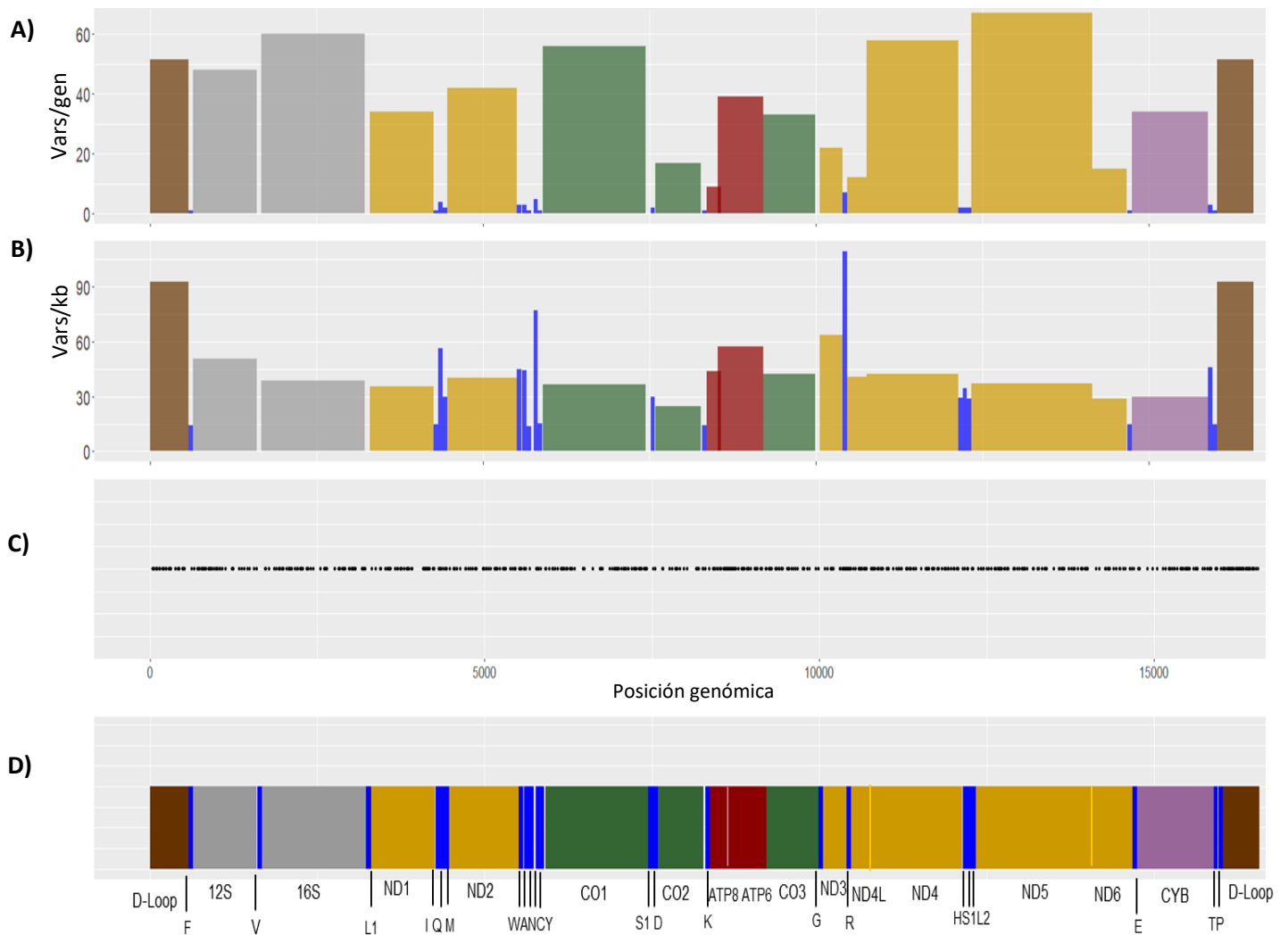


Figura 7. Distribución genómica de las variantes mitocondriales en cáncer de mama. Se muestra el número de variantes (eje Y) identificadas en cada gen (eje X) (A), el número de variantes por kilobase y por gen (B); además, se muestra la localización de las variantes, donde cada punto corresponde a una variante (C) y el mapa del genoma mitocondrial (D). Los colores representan los complejos o regiones (amarillo: complejo I; verde complejo III; morado: complejo VI; rojo: complejo V; azul: tRNA; gris: rRNA; café: región D-Loop).

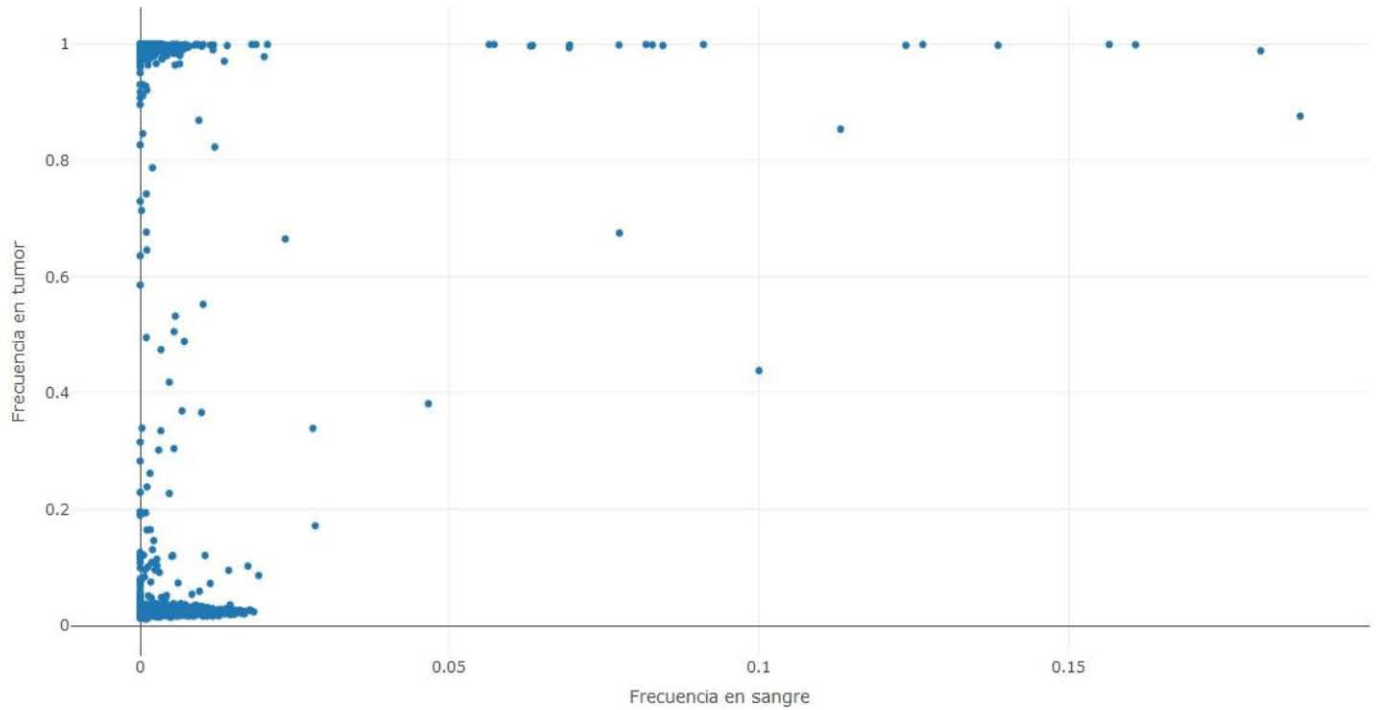
5.3 ANÁLISIS DE HETEROPLASMIA

De las 707 variantes identificadas en las muestras de tumor, 567 también fueron detectadas en sangre periférica (Fig. 7A) y únicamente 140 de ellas estuvieron presentes sólo en los tumores de mama, por lo cual, estas variantes fueron definidas como mutaciones somáticas (Fig. 7B).

Para estimar el porcentaje de heteroplasmia de las 567 variantes que estaban presentes tanto en tumor como en sangre periférica, se calculó la fracción alélica (FA) en ambos tejidos. Se observó heteroplasmia en un intervalo de 1.1 - 95.12% (Figura 7A), con variabilidad en el porcentaje de heteroplasmia entre cada paciente e incluso para la misma variante. Solo 283 de éstas estuvieron presentes en más de una muestra (frecuencia $\geq 2\%$).

Una observación importante en este análisis es que a pesar de que la FA es distinta entre individuos, todas las variantes con heteroplasmia estaban más enriquecidas en el tumor en comparación con la sangre periférica. De hecho, algunas variantes como G8027A, C11177T, G12007A, C12705T, A13105G, T16362C, entre otras, observaron bajo nivel de heteroplasmia en sangre periférica, pero en tumor alcanzaron el estado de homoplasmia (Fig. 8). Aunado a ello, variantes como T4248C y T16325C, mostraron bajo nivel de heteroplasmia en sangre periférica, pero en tumor su FA estuvo en intervalos variables para los individuos que la presentaron, siendo estos de 97.8 a 99.8% y de 1.1 a 99.8%, respectivamente (Fig. 8).

A)



B)

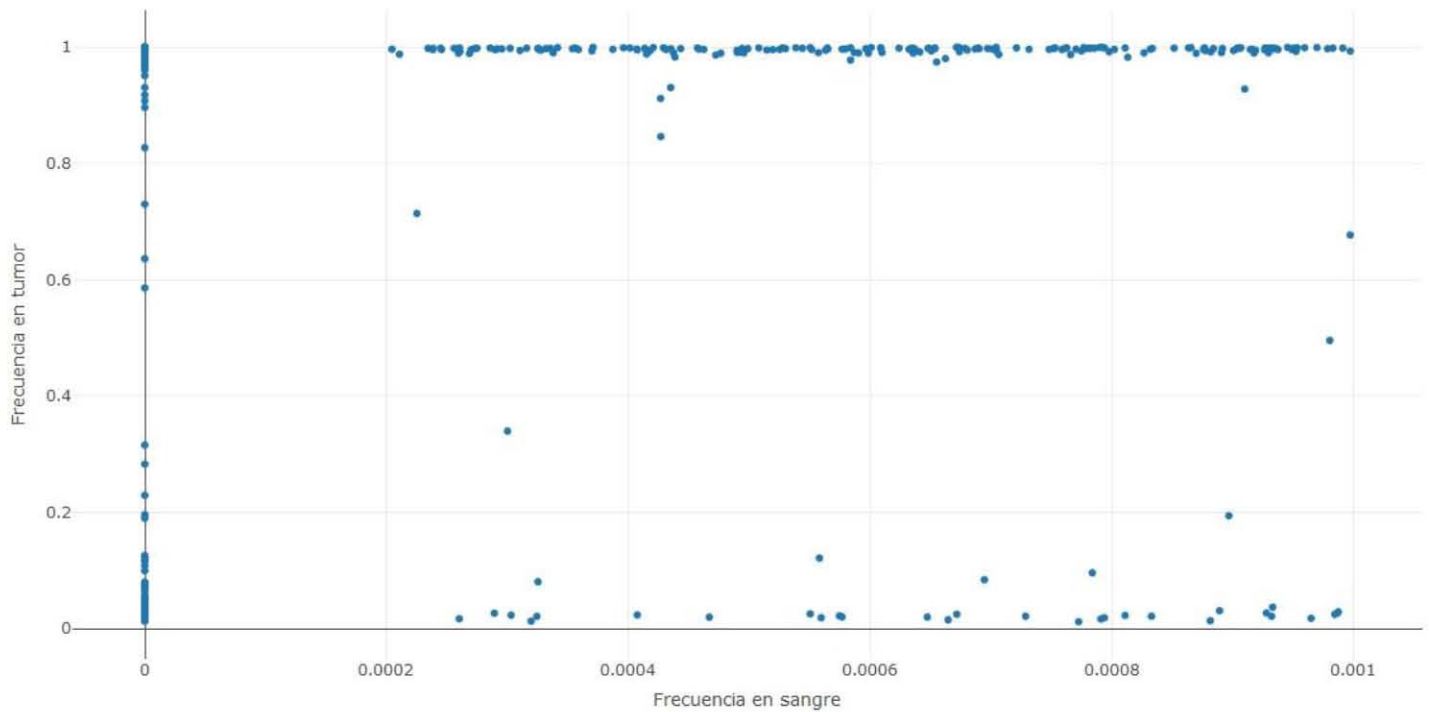


Figura 8. Frecuencia alélica de las variantes mitocondriales. Se grafica la frecuencia alélica de las variantes encontradas al comparar la sangre periférica (eje X) con el tejido tumoral (eje Y). Figura A: se representan las 707 variantes identificadas tanto en sangre periférica como en el tumor de mama; figura B: se representan las mutaciones somáticas.

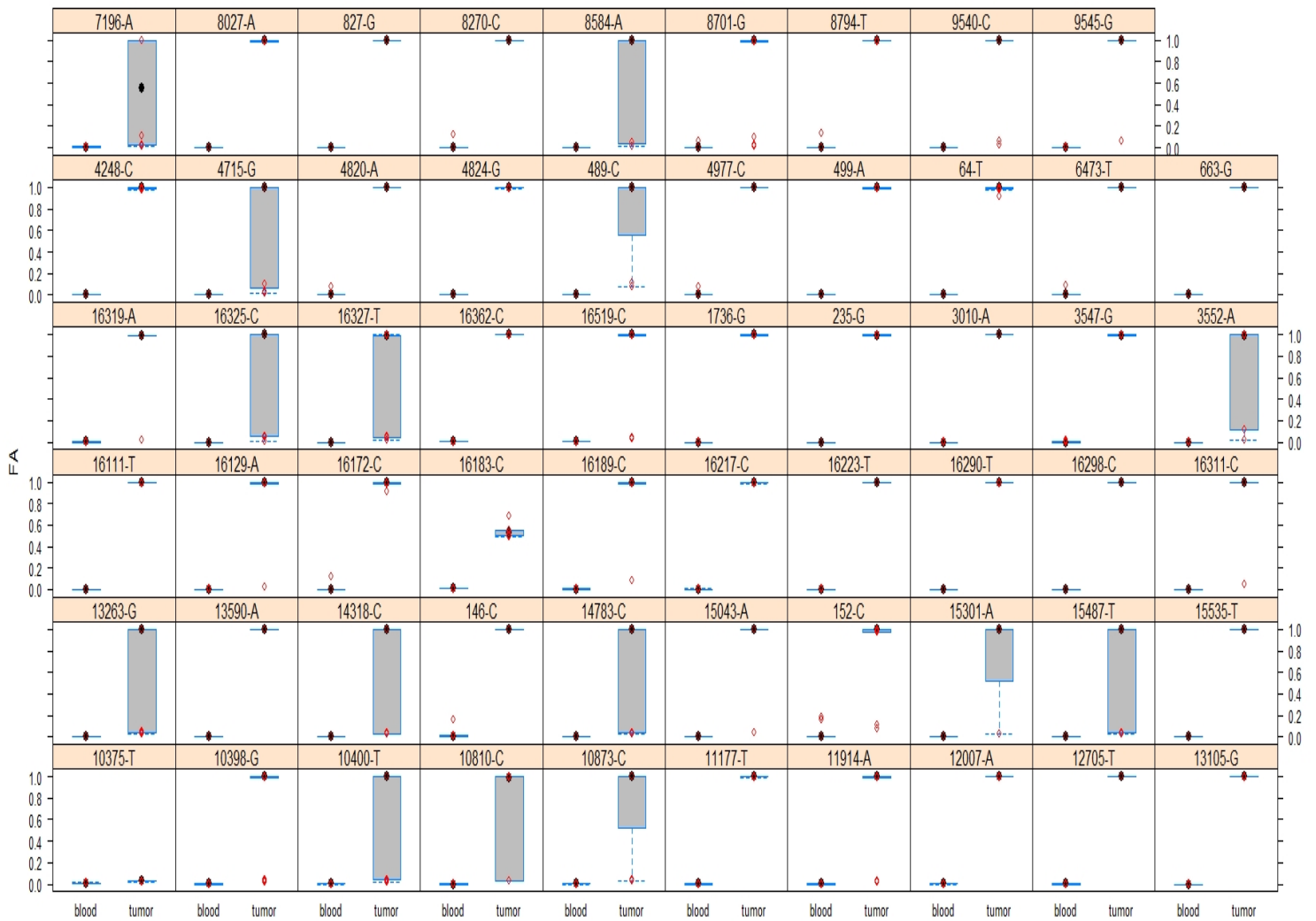


Figura 9. Distribución del grado de heteroplasmia de las variantes mitocondriales identificadas con mayor frecuencia. Cada cuadro representa una variante. En el eje y se reporta la frecuencia alélica (porcentaje de alelo mutado o heteroplasmia) y en el eje x se muestran los tejidos evaluados. Cada punto es el dato obtenido de una paciente y las barras grises representan la variación en el grado de heteroplasmia entre los individuos. Únicamente se muestran las variantes que estuvieron presentes en al menos 4 casos.

5.4 ANÁLISIS DE MUTACIONES EN TUMORES DE CÁNCER DE MAMA

En el análisis de mutaciones somáticas, se identificaron un total de 140 distribuidas en 51 (55.4%) tumores de mama, encontrándose de 1 a 18 mutaciones por muestra. Cuarenta y un (44.4%) tumores no presentaron ninguna mutación somática.

Los genes que presentaron un mayor número de mutaciones fueron el *ND5* (n=16), *16S* (n=14), el *CO1* (n=13); así como la región D-Loop (n=26). Al estimar la tasa mutacional (mutaciones/kb), la región D-Loop (23.4 muts/kb), el gen *ATP8* (14.6 muts/kb), *ND3* (14.5 muts/kb); así como los tRNA de T, H, W, A y Q fueron los que presentaron una mayor tasa mutacional (Fig. 9).

Aunque las mutaciones somáticas se presentaron con una proporción variable del alelo mutado (porcentaje de heteroplasma) entre los individuos analizados, 83 (59.3%) de ellas se presentaron con una proporción entre 95-100%, es decir, en estado homoplásmico (Fig. 10).

5.5 MUTACIONES SOMÁTICAS NUEVAS

Las mutaciones identificadas en nuestra población fueron buscadas en MITOMAP [\[92\]](#). El 72.9% (n=102) de ellas ya cuentan con registro en las bases de datos. De éstas, la T9950C, C16278T y A11002G fueron las más frecuentes en nuestras pacientes (7.6%, 5.4% y 4.3% respectivamente). El 27.1% (n=38) restante corresponden a mutaciones nuevas, sin embargo, su frecuencia en nuestra población fue menor al 2%; de hecho, una gran parte de ellas (n=33) fueron mutaciones presentes en un solo individuo (Tabla 6).

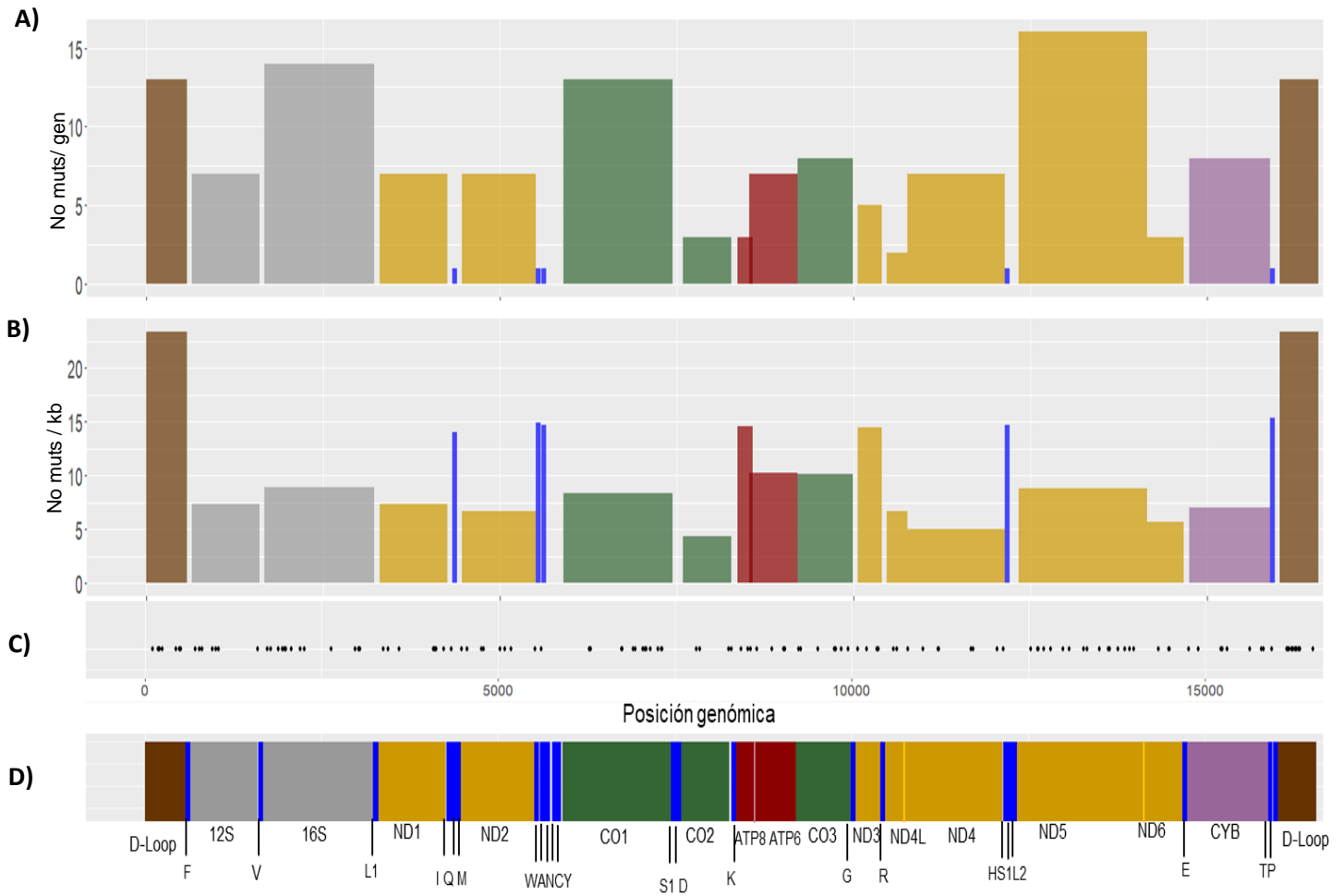


Figura 10. Distribución genómica de las mutaciones somáticas mitocondriales. Se esquematiza el número de mutaciones encontradas en cada gen (A), el número de mutaciones por kilobase y por gen (B); así como la localización de las mutaciones, donde cada punto corresponde a una mutación (C) y el mapa del genoma mitocondrial (D). Los colores representan los complejos o regiones (amarillo: complejo I; verde complejo III; morado: complejo VI; rojo: complejo V; azul: tRNA; gris: rRNA; café: región D-Loop).

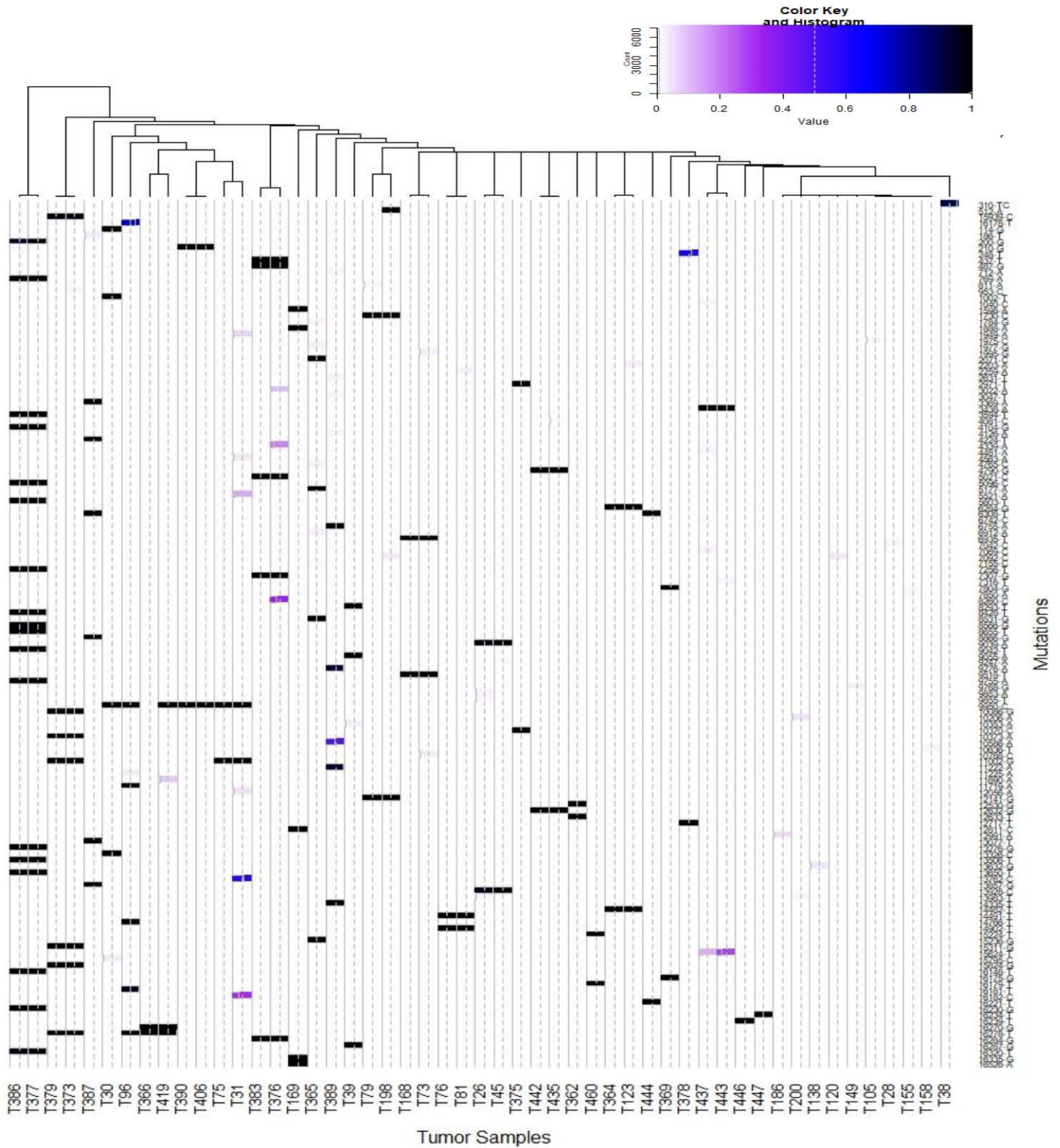


Figura 11. Niveles de heteroplasma de las mutaciones somáticas. El heatmap representa el grado de heteroplasma de cada mutación somática (filas) en cada uno de los pacientes en que se identificó (columnas). El color representa la proporción del alelo mutado; blanco: $\geq 1.1\%$, negro: 100%.

5.6 ANÁLISIS DEL EFECTO FUNCIONAL DE LAS MUTACIONES SOMÁTICAS

El 62.1% de las mutaciones somáticas se localizaron en región codificante de proteínas, mientras que el 37.9% restante se presentaron tanto en regiones regulatorias (D-Loop) como en genes que transcriben para rRNAs y tRNAs (Tabla 7).

El análisis de efecto funcional se realizó para las 87 mutaciones encontradas en regiones codificantes y las 53 no codificantes. Cabe señalar que 28 mutaciones presentaron consecuencias funcionales dobles debido a su localización en el mtDNA, ya que pueden afectar a los genes de la cadena H y de la cadena L o a dos genes cuyas secuencias se sobrelapan. Por ejemplo, por su localización, las mutaciones A8531G y A8566G afectan tanto a los genes *ATP8* y *ATP6*.

De las mutaciones codificantes, 45 (50.6%) fueron sinónimas y 44 (49.4%) no sinónimas. Sin embargo, dentro de las mutaciones no sinónimas, las de tipo missense fueron las más abundantes (97.2%).

Por otro lado, el análisis reveló que 51 (64.6%) mutaciones no codificantes podrían alterar la interacción entre esta región con proteínas que regulan el nivel de compactación de la molécula del mtDNA como el factor de unión a sitios CCCTC (CTCF).

Tabla 6. Mutaciones somáticas más frecuentes en tumores de mama analizados.

Posición	Gen	Consecuencia	Frecuencia (%)
T9950C	<i>CO3</i>	Cambio AA sinónimo: V	7.6
C16278T	D-Loop	Sitio de unión a CTCF	5.4
A11002G	<i>ND4</i>	Cambio AA sinónimo: Q	4.3

Tabla 7. Clasificación funcional de las mutaciones somáticas identificadas en cáncer de mama.

Tipo	N	Efecto n (%)		
Codificante	89*	Sinónimo	45 (50.6%)	
		No sinónimo	44 (49.4%)	
			Missense	43 (97.2%)
			Non-sense	1 (2.3%)
No codificante	79**	Regulador	51 (64.6%)	
		rRNA, tRNA	26 (32.9%)	
		Intergénicos	2 (2.5%)	
*Se consideran 2 mutaciones con función doble **Se consideran 26 mutaciones con función doble NA: no asignado				

5.7 MUTACIONES SOMÁTICAS DEL mtDNA REPORTADAS COMO ASOCIADAS A OTROS TIPOS DE CÁNCER

De todas las mutaciones que encontramos en nuestro estudio, solo 8 han sido previamente asociadas a cáncer. Así, la mutación T310TC ha sido reportada en melanoma [106]; mientras que la mutación C437T en cáncer de próstata [107]. De todas estas mutaciones, la única que ha sido descrita en tumores de mama es la A15824G (Tabla 8). Otras mutaciones como la T310TC y A16182G han sido descritas en otros tipos de tumores como polimorfismo.

Tabla 8. Mutaciones somáticas previamente relacionadas con cáncer.

Posición	Gen	Consecuencia	Función	Cambio de AA	Frecuencia (%)*	Cáncer
T310TC	D-Loop	Regulatoria	Sitio de unión a CTCF	-	1.1	Melanoma
C437T	D-Loop	Regulatoria	Sitio de unión a CTCF	-	2.1	Próstata
A487G	D-Loop	Regulatoria	Sitio de unión a CTCF	-	2.1	Próstata
C3594T	ND1	Codificante de proteína	Sinónimo	V	2.1	Tiroides
G5521A	TW	Cambio de transcrito	tRNA	-	1.1	Pulmón
C12717T	ND5	Codificante de proteína	Sinónimo	T	1.1	Páncreas
A15824G	CYB	Codificante de proteína Región regulatoria	No sinónimo Sitio de unión a CTCF	T/A	2.1	Mama
A16182C	D-Loop	Región regulatoria	Sitio de unión a CTCF	-	1.1	Próstata

*Frecuencia correspondiente a la población analizada en este estudio

5.8 DISTRIBUCIÓN DE MUTACIONES POR SUBTIPOS MOLECULARES DE CÁNCER DE MAMA

Dado que solamente 55 de los tumores contaban con clasificación molecular basada en perfiles de expresión (PAM50), la asignación de los subtipos se realizó usando los marcadores inmunohistoquímicos en 90 (97.8%) pacientes. De estas muestras, el subtipo LA presentó el mayor número de mutaciones somáticas (n=111), seguido del subtipo LB (n=44), TN (n=12) y H2 (n=4) (Tabla 9); mientras que los subtipos H2 y TN fueron los que presentaron el mayor número de casos sin mutaciones (60.0% y 50%, respectivamente), seguido del subtipo LB (45.5%) y por último del subtipo LA (42.1%) (Tabla 10).

Para determinar si existía un perfil de mutaciones asociados al subtipo molecular, se contabilizó el número de mutaciones en cada grupo. La mutación sinónima T9950C localizada en el gen *CO3* fue la única presente en los cuatro subtipos moleculares; mientras que LA y LB fueron los subtipos que compartieron más mutaciones (n=18). Así mismo, se encontraron 84 mutaciones exclusivas de LA, 23 de LB y 6 de TN; pero sus frecuencias fueron de 1% en la población analizada. El subtipo H2 no presentó ninguna mutación exclusiva (Fig. 11A). Por

otro lado, con los 55 tumores que contaron con subtipo clasificado por PAM50 también se realizó el análisis de distribución de las mutaciones, pero no se observaron diferencias significativas entre ambos clasificadores (Fig. 11B).

Tabla 9. Número de mutaciones somáticas por subtipo molecular.

Subtipo Molecular	Frecuencia n (%)	Mutaciones somáticas n (%)	Rango de mutaciones	Promedio de mutaciones
LA	57 (63.3)	111	0 – 18	2
LB	22 (24.4)	44	0 – 9	2
H2	5 (5.6)	4	0 – 2	1
TN	6 (6.7)	12	0 – 7	2

LA: Luminal A, LB: Luminal B, H2: HER2, TN: Triple negativo

Tabla 10. Número de mutaciones por individuo y subtipo molecular.

Subtipo molecular	Individuos totales* n (%)	Individuos con mutaciones n (%)	Individuos sin mutaciones n (%)	P valor
LA	57 (63.3)	33 (57.9)	24 (42.1)	0.86
LB	22 (24.4)	12 (54.5)	10 (45.5)	
H2	5 (5.6)	2 (40.0)	3 (60.0)	
TN	6 (6.7)	3 (50.0)	3 (50.0)	

*Total de muestras que contaron con clasificación inmunohistoquímica

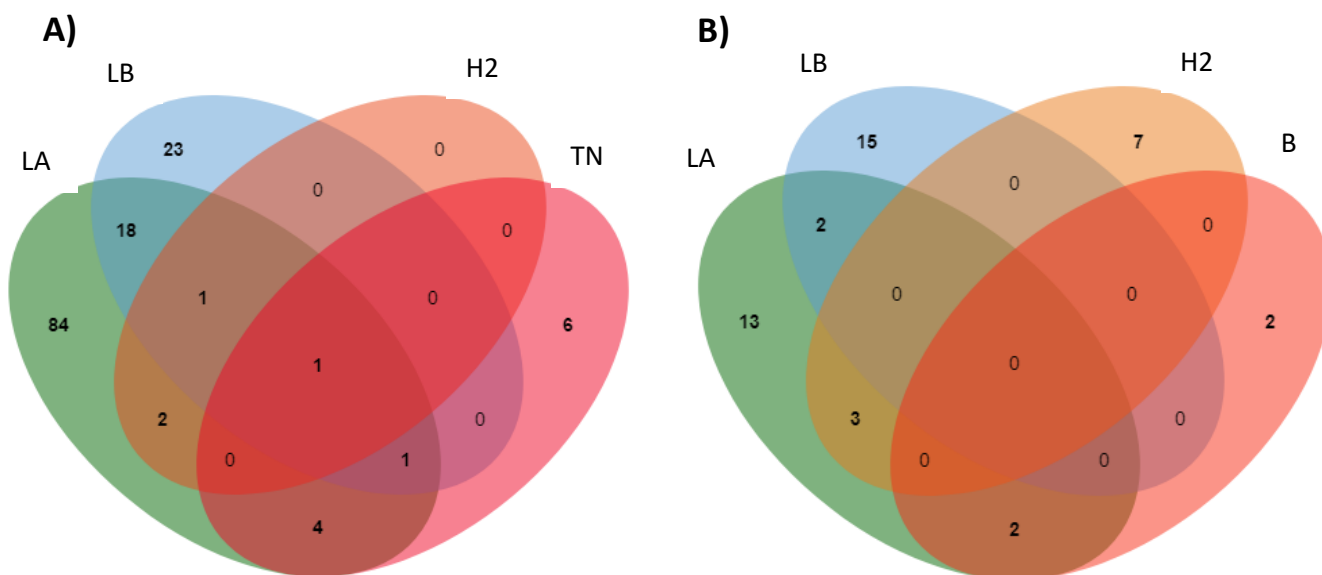


Figura 12. Distribución de mutaciones en los subtipos moleculares de cáncer de mama. Se muestra el número de mutaciones exclusivas o compartidas entre los subtipos moleculares A) tumores clasificados con marcadores inmunohistoquímicos y B) tumores con clasificación PAM50. LA: luminal A, LB: luminal B, H2: HER2, TN: Triple negativo y B: basal.

5.9 DISTRIBUCIÓN DE HAPLOGRUPOS

En el estudio se encontró que los haplogrupos A (45%), B (23%), C (12%) y D (12%) fueron los más frecuentes, a diferencia de los haplogrupos H, J y L que estuvieron en menor proporción (2%, 1% y 5%, respectivamente) (Fig. 12).

En cuanto a la distribución de los haplogrupos en los subtipos moleculares, se observó que los haplogrupos C y D solo se presentaron en los LA y LB. Así mismo, los haplogrupos A y B fueron los únicos que se detectaron en el subtipo triple negativo (Tabla 11).

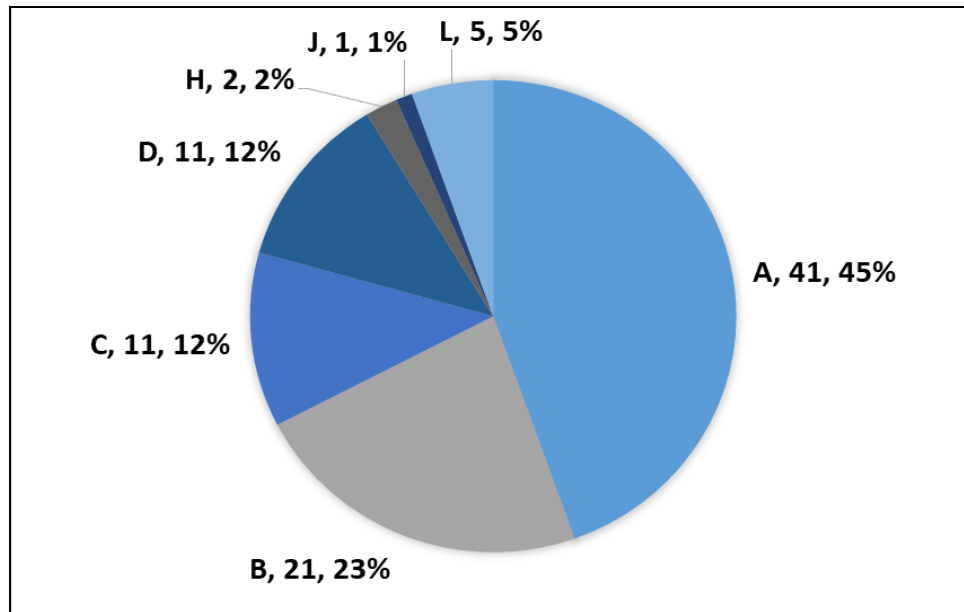


Figura 13. Distribución de haplogrupos en cáncer de mama.

Tabla 11. Distribución de los haplogrupos por subtipo molecular de cáncer de mama.

HG	Subtipo molecular n (%)				Total HG n (%)
	LA	LB	H2	TN	
A	23 (40.4)	12 (54.5)	2 (40.0)	3 (50.0)	40 (43.2)
B	13 (22.8)	3 (13.6)	1 (20.0)	3 (50.0)	20 (22.7)
C	8 (14.0)	3 (13.6)	0	0	11 (12.5)
D	9 (15.7)	2 (9.1)	0	0	11 (12.5)
H	1 (1.8)	0	1 (20.0)	0	2 (2.3)
J	0	1 (4.6)	0	0	1 (1.1)
L	3 (5.3)	1 (4.6)	1 (20.0)	0	5 (5.7)
Total	57 (100)	22 (100)	5 (100)	6 (100)	90 (100)

HG: haplogrupo, LA: luminal A, LB: luminal B, H2: HER2, TN: triple negativo

6 DISCUSIÓN

El cáncer es actualmente un grave problema de salud pública en México y en el mundo, cuya incidencia y mortalidad va en aumento a pesar de la implementación de medidas preventivas. El cáncer es un conjunto de entidades clínica y genéticamente heterogéneas, donde alteraciones tanto en genes nucleares como mitocondriales contribuyen a su desarrollo. Las primeras evidencias de la participación de las mitocondrias en la oncogénesis provienen de los estudios realizados por Otto Warburg en 1956, quien propuso que todos los tipos de cáncer se originaban de una disfunción en el proceso respiratorio de la célula [108, 109]. Posteriormente se observó que la función mitocondrial de una célula tumoral correlaciona con un estado redox particular que conllevaba a una reprogramación a nivel metabólico, hecho que actualmente es considerado una de las características principales del cáncer [42].

El empleo de las tecnologías genómicas como la secuenciación de nueva generación (NGS, del inglés, Next Generation Sequencing) ha contribuido al conocimiento de las alteraciones que ocurren en el genoma mitocondrial y que potencialmente podrían tener implicaciones relevantes en el desarrollo del cáncer, en el pronóstico de la enfermedad o en la respuesta al tratamiento [110, 111]. Para conocer la frecuencia y la naturaleza de las mutaciones en genes mitocondriales de la cadena respiratoria en pacientes con cáncer de mama y determinar si existe una correlación entre estos con el subtipo molecular del tumor, secuenciamos el genoma mitocondrial de 92 pares de sangre periférica-tumor mediante la plataforma MiSeq (Illumina). El intervalo de edad de nuestras pacientes fue de 33 a 92 (53.83 ± 11.5) años, de los cuales 45.7% de los casos fueron menores a 50 años. Así mismo, los subtipos tumorales LA (54.3%) y LB (28.3%) fueron los más representados. Estos hallazgos son similares a los resultados publicados en estudios previos en población mexicana, aunque podemos observar que la edad promedio con la que el cáncer de mama se presenta es 10 años menor comparada con la descrita para otras poblaciones (53 vs 63 años, respectivamente). Además, los subtipos H2 y TN se presentaron con una frecuencia menor (ambos 6.5%) a la ya reportada (8.7 y 14.63%, respectivamente) [59, 112].

Distribución de variantes en el genoma mitocondrial de pacientes mexicanas con cáncer de mama

El análisis de secuenciación de todo el genoma mitocondrial reveló la presencia de 707 variantes en los tumores de mama analizados, de estos, 140 fueron detectadas sólo en tumor, por lo que se consideraron como mutaciones somáticas. El 55.4% de pacientes fueron portadores de mutaciones somáticas, un porcentaje más bajo a los descrito en otros estudios. Por ejemplo, Tan y cols. estudió 19 tumores de mama y encontró que alrededor del 74% de las muestras tenían por lo menos una mutación somática [86, 87].

Heteroplasmia en los tumores de mama

En nuestro análisis, todas las variantes identificadas, y 57 de las mutaciones somáticas observaron un amplio intervalo en los niveles de heteroplasmia (1% - 95%). Dado que nuestro estudio incluyó tumores con una celularidad tumoral mayor al 80%, no podemos descartar la posibilidad de que la heteroplasmia a bajos niveles observada en el tumor, pudiera ser por contaminación de la muestra con células normales. Sin embargo, el fenómeno de heteroplasmia es característico del cáncer y de otras entidades humanas, por lo que su análisis a más detalle resulta relevante. Se sabe que durante el proceso carcinogénico ocurren dos tipos de mutaciones: las mutaciones patogénicas deletéreas que generalmente son eliminadas por comprometer la función proteica de manera irreversible, y aquellas que influyen en los mecanismos adaptativos que favorecen el crecimiento celular y por lo tanto son seleccionadas. Si las mutaciones son deletéreas, también serán pasajeras; es decir, que no son completamente fijadas y podríamos detectarlas en bajos niveles de heteroplasmia; mientras que las mutaciones que logren modificar la función del sistema OXPHOs sin comprometer la viabilidad, serían fijadas y permitirían adaptar a la célula al ambiente tumorigénico, lo que explicaría sus elevados niveles de heteroplasmia en las células tumorales, donde las mutaciones pueden alcanzar el estado homoplásmico [19, 113]. Si es así, es probable que las mutaciones homoplásmicas encontradas en las muestras analizadas, se deba a que éstas son de relevancia para el proceso carcinogénico y que quizá conceden alguna ventaja metabólica para la célula tumoral. La homoplasmia puede alcanzarse debido a

un mecanismo de selección y sobrerreplicación de la molécula de mtDNA mutada, lo cual favorecería una tasa elevada de producción de ROS por el sistema OXPHOs, y así, una proliferación celular no regulada [6, 19].

Es de hacer notar, que el grado de heteroplasmia no mostró correlación con variables clínicas como la edad, el subtipo tumoral o el grado histológico del cáncer de mama. Hasta el momento no existen datos sobre la correlación entre estas variables y el nivel de heteroplasmia, pero se ha reportado que la heteroplasmia en sangre periférica de variantes localizadas en las regiones HV1 y HV2, se asocian al riesgo a padecer cáncer de mama [114].

Las mutaciones somáticas en genes del sistema OXPHOs y RNAs no codificantes de la mitocondria

La secuenciación completa del mtDNA reveló que el 62.1% de las mutaciones somáticas se localizaron en regiones codificantes y el 37.9% se presentaron en regiones no codificantes. Además, la tasa de mutación fue más elevada para la región D-Loop y algunos tRNA, hallazgos que son consistentes con lo reportado en bases de datos públicos como MITOMAP y por otros investigadores, donde se observa que la región no codificante, incluyendo los tRNAs y el D-Loop son de 3 hasta 60 veces, respectivamente, más susceptibles de mutar en comparación con las regiones codificantes [87]. Por otro lado, los genes codificantes con más mutaciones fueron *ND5* y *CO1*. Nuestros hallazgos difieren de los descrito en otros estudios, en donde se ha reportado a *ATP6*, *ND3* y *CYB* son los genes más frecuentemente mutados. Sin embargo, nuestros resultados indicaron que el gen *ND5* presentó una mayor proporción de mutaciones somáticas, un dato similar con el reportado previamente por McMahon y LaFramboise [86].

Efecto funcional de las mutaciones somáticas

De todas las mutaciones somáticas codificantes, el 50.6% fueron no sinónimas y el 49.4% fueron sinónimas. El 97.2% de los cambios no sinónimos fueron mutaciones de pérdida de sentido, cuyo cambio de aminoácido se sugiere que pueda impactar en la estructura secundaria de la proteína, su estabilidad, alterar las modificaciones postraduccionales, su

capacidad de interacción con otras proteínas y por ende, en el funcionamiento y organización del sistema OXPHOs [115, 116]. Un ejemplo del impacto funcional de las mutaciones no sinónimas es el cambio de una leucina por una arginina en la proteína codificada por el gen *ATP6* que se genera por la mutación T8993G, la cuál ha sido identificada en tumores de próstata. Este cambio de aminoácido interrumpe un adecuado bombeo protónico por la ATP-sintasa y altera el potencial de membrana desacoplando el sistema OXPHOs; además ensayos funcionales indican su participación en el aumento de ROS y aumento de la proliferación celular en cáncer de próstata [117, 118].

Así mismo, aunque la investigación sobre el efecto funcional de las variantes sinónimas en mtDNA es escasa, es probable que al igual que lo que ocurre con algunas variantes sinónimas del DNA nuclear, éstas podrían impactar en la estabilidad del mRNA e indirectamente en los niveles de expresión, estructura y actividad de las proteínas [119]. Se ha reportado que las variantes o mutaciones sinónimas, también pueden acelerar o desacelerar la velocidad con la que el ribosoma se mueve a lo largo del mRNA, y con ello la velocidad de la traducción; así como la estructura y función de la proteína traducida. En la traducción de proteínas codificadas en el DNA nuclear, es bien conocido que el uso de codones en las familias de codones sinónimos no es aleatorio y hay una preferencia de codones determinada por la segunda o tercera base del codón [119, 120]. Así mismo, las mutaciones en cualquier parte del genoma que afectan el plegamiento y la estructura secundaria de los tres transcritos policistrónicos son potencialmente perjudiciales para la producción de las moléculas de rRNA, tRNA y mRNA maduras.

El sistema OXPHOs funciona gracias a la coordinación de la expresión de los genomas nuclear y mitocondrial, siendo este último, aquel que proporciona las subunidades proteicas más hidrofóbicas que dan lugar a componentes estructurales o subunidades catalíticas de los complejos enzimáticos. Por esta razón, alteraciones localizadas en estas regiones podrían modificar la eficiencia de la cadena respiratoria, lo que se refleja en una disminución de la producción de ATP y un aumento en la producción de ROS. Cuando las concentraciones de ROS sobrepasan el umbral de los mecanismos antioxidantes, se genera un ambiente hipóxico que activa al factor de transcripción HIF1- α , el cual conduce a la replicación del mtDNA y la expresión de genes mitocondriales, lo que lleva a la manifestación de un estado energético

alterado que da lugar a daño en estructuras celulares y resistencia a la apoptosis, promoviendo así la transformación tumoral [4, 6].

Por otro lado, el 37.9% de las mutaciones se presentaron en regiones no codificantes (D-Loop, rRNA, tRNA y regiones intergénicas), que si bien, aparentemente no comprometen la función de la proteína, sus alcances pueden llegar a la alteración de los procesos de replicación y transcripción. La predicción del efecto funcional mostró que un 64.6% de las mutaciones localizadas en zonas no codificantes tienen alta probabilidad de alterar procesos regulatorios de la expresión del mtDNA ya que las secuencias que participan en la unión de factores transcripcionales se vieron modificados, tal como el sitio de unión a CTCF, el cual es una proteína que participa en la compactación del mtDNA en nucleoides [121, 122]. La ausencia del sitio de unión para CTCF podría exponer el mtDNA a la maquinaria encargada de replicar y transcribir el mtDNA, que como se ha descrito, se ven activados dadas las condiciones hipóxicas generadas por el estrés oxidante ocurrido en la tumorigénesis.

Adicionalmente, se debe considerar que muchas de las mutaciones somáticas encontradas pueden representar mutaciones pasajeras que no desempeñan ningún papel principal en la tumorigénesis.

Mutaciones nuevas y reportadas en cáncer

Para identificar variantes nuevas, se realizó una búsqueda de las mutaciones encontradas en las bases de datos públicas, revelando que el 27.1% de ellas son potencialmente nuevas. La frecuencia del 86.8% de las mutaciones potencialmente nuevas fue del 1%, y aunque no descartamos la posibilidad de que pudieran ser originadas durante la PCR, nuestros hallazgos podrían derivarse de la variación individual, ya que en estudios realizados en diversas poblaciones también se han reportado altos porcentajes de mutaciones nuevas en los tumores de mama [79, 86].

Es de hacer notar, que la mayoría de las mutaciones que identificamos solamente en los tumores también se han registradas como variantes germinales en bases de datos y en otros estudios [33]. Por ejemplo, las mutaciones T9950C (*CO3*) y A11002G (*ND4*), mismas que están

dentro de las más frecuentes (7.6% y 5.4%, respectivamente) en nuestro estudio, han sido reportadas como polimorfismos en población asiática, y se han descrito como factor de riesgo a padecer neuropatías, cardiomiopatías, así como cáncer colorrectal y tiroideo [123-125]. Interesantemente, nuestro estudio es el primero que las reporta en cáncer de mama.

La mutación C16278T, también una de las más frecuentes (4.3%), se localiza en una región no codificante (D-Loop) cuya predicción funcional es de tipo regulatorio. Es probable que esta mutación altere una secuencia de unión del CTCF y por lo tanto promueva la expresión del mtDNA, pero aún no existen estudios funcionales. C16278T forma parte del haplogrupo X y ha sido encontrada en tumores de ovario y propuesto como marcador metastásico en DNA circulante de pacientes con cáncer de mama [126-128].

De todas las mutaciones que encontramos en nuestro estudio y que ya contaban con una asociación a cáncer, la única que ha sido descrita en tumores de mama es la A15824G, localizada en el gen *CYB* e identificada en el 10.3% de los tumores de población senegalesa. Interesantemente, su frecuencia reportada es elevada con respecto a nuestro estudio (2.1%) y aunque se desconoce su efecto funcional; el cambio que genera de una treonina por una alanina, podría impactar en el funcionamiento y estructura de la proteína [129].

Distribución de mutaciones en los subtipos moleculares de cáncer de mama

En el análisis del número y tipo de mutaciones del genoma mitocondrial por subtipo molecular de cáncer de mama, reveló que no existe un perfil de mutaciones asociados a dichos subtipos. En términos del genoma nuclear, los tumores negativos a receptores hormonales mantienen un perfil mutacional más elevado que los otros subtipos, pero no hay estudios que hayan explorado la asociación entre la carga mutacional mitocondrial y el subtipo molecular del tumor. Nuestro análisis mostró que los tumores LA y LB no sólo tienen el mayor número de mutaciones en relación con los otros subtipos moleculares; sino que, además, son los subtipos que más mutaciones comparten. La elevada carga mutacional en los tumores LA y LB podría explicarse por la formación de radicales superóxidos durante el metabolismo de los estrógenos. Así, estas moléculas altamente reactivas contribuirían a la formación de alteraciones en el mtDNA [130, 131].

Es probable que estos resultados estén sesgados por el número de tumores analizados; sin embargo, el hecho de que LA y LB sean los tipos histológicos con más mutaciones compartidas, podría ser explicado por su origen común [132]. Aunque los tumores LA y LB parecen ser muy distintos entre sí desde el punto de vista genómico, de expresión génica, y de patrones de metilación del DNA [133], también tienen similitudes biológicas, por ejemplo, baja contaminación de células del estroma y de células inmunes; así como niveles similares de la expresión de genes como *ESR1*, *PGR*, y *FOXA1* [134, 135].

Otra variable por considerar son las limitaciones de la inmunohistoquímica para establecer con precisión la clasificación molecular del tumor, en comparación con PAM50. Lo anterior, podría explicar las pequeñas diferencias en el número de mutaciones entre los subtipos moleculares cuando se usaron los dos clasificadores disponibles para 55 tumores. De hecho, se ha reportado una dificultad para distinguir adecuadamente los tumores LB de los LA y H2 usando solo marcadores inmunohistoquímicos [62].

Distribución de haplogrupos en cáncer de mama

Para conocer si existe una asociación entre los haplogrupos y el subtipo molecular de cáncer de mama, se analizaron las secuencias del mtDNA de línea germinal. Encontramos que los haplogrupos A, B, C y D son los más comunes en las mujeres que padecen esta enfermedad, observándose frecuencias similares a las reportadas para la población mestiza e indígena mexicana [35, 37, 136]. Sin embargo, se observó una mayor frecuencia de los haplogrupos D (12%) y L (5%) en nuestras muestras en comparación con la frecuencia reportada previamente para población mestiza abierta (6.2% y 1.9%, respectivamente), aunque las diferencias no son estadísticamente significativas [35].

En estudios realizados en diversas poblaciones se ha reportado que los haplogrupos K, M, T y U se asocian al riesgo a padecer cáncer de mama [67, 83, 137, 138]. Así mismo, en un estudio realizado por Blein et al 2015, se documentó un enriquecimiento del subhaplogrupo T1a1 en mujeres con cáncer de mama positivas a mutaciones en *BRCA2* [138]. Por otro lado, existen reportes de que el subhaplogrupo D5 está asociado con cáncer de mama, con una frecuencia descrita de 12.9% [83], muy similar a la encontrada en nuestro estudio (12%).

Además, el haplogrupo L es de origen africano y se ha relacionado con una elevada expresión de genes codificantes para complejos respiratorios, una menor tasa de recambio de ATP y baja producción de ROS que dan lugar a un sistema OXPHOs más eficiente; la adquisición de mutaciones somáticas podría alterar la función mitocondrial desencadenando un ambiente carcinogénico [139].

Por otro lado, al analizar la distribución de los haplogrupos en los subtipos moleculares de cáncer de mama no se observó el enriquecimiento de alguno de estos en un subtipo particular. Sin embargo, este análisis puede verse sesgado por el reducido número de muestras evaluadas ya que existen datos que sugieren que algunos haplogrupos como el D5, promueven la tumorigénesis de forma parcial activando la vía AKT mediada por un aumento de ROS que se relaciona con proliferación y supervivencia [47, 91].

Impacto del estudio del genoma mitocondrial en cáncer de mama

El cáncer es una entidad compleja y altamente heterogénea donde los mecanismos de regulación se pierden y dan lugar a una proliferación celular descontrolada. La función del genoma mitocondrial es de gran relevancia dado el grado de coordinación que tiene con la expresión del genoma nuclear. Aunque la célula tumoral pueda escapar a los mecanismos regulatorios, no pueden evadir los procesos concernientes al flujo energético. Por lo tanto, las células cancerosas deben implementar estrategias para ajustar el metabolismo energético que le proporcione adaptación al ambiente, posiblemente mediante el intercambio entre el metabolismo oxidativo y glucolítico, empleando ROS para estimular la división celular [19].

Se ha sugerido que este proceso de adaptación metabólica resulta ser más eficiente si es causada por mutaciones mitocondriales, por lo cual se ha planteado que algunas de ellas pueden estar en procesos de selección positiva durante el desarrollo del tumor, lo que explicaría la variabilidad en el grado de heteroplasmia en el tumor e incluso el estado homoplásmico, de algunas de las mutaciones encontradas en nuestro estudio. Lo anterior es sustentado con estudios que han identificado mutaciones nuevas en estados metastásicos no vistas en el estadio primario [113]. Las mutaciones nuevas no apoyan la hipótesis planteada, sin embargo, dado que un alto porcentaje de las mutaciones detectadas en este estudio sólo

se presentaron en un individuo, similar a lo que se ha descrito en otros análisis, es posible que la mayoría de las mutaciones puntuales somáticas mitocondriales no tengan una participación en el inicio de la carcinogénesis y más que tener una contribución patogénica, pueden representar simplemente las consecuencias de una inestabilidad genómica o el daño oxidativo adquirido durante los múltiples pasos de la carcinogénesis. Sin embargo, no podemos descartar que otras alteraciones como el número de copias o procesos mitoepigenéticos pudieran ser relevantes en el proceso oncogénico de la glándula mamaria, aunque son necesarios estudios a nivel funcional que permitan esclarecer los mecanismos implicados con las evidencias publicadas hasta el momento, podemos decir que algunas de ellas pueden conceder ventajas selectivas y adaptativas a la tumorigénesis, aunque son necesarios estudios a nivel funcional que permitan esclarecer los mecanismos implicados [[130](#), [140](#)].

7 CONCLUSIONES

Presentamos el primer estudio que muestra el panorama mutacional del genoma mitocondrial en pacientes mexicanas con cáncer de mama y nuestros resultados sugieren que:

- Existe una tasa de mutación similar entre los genes de la cadena OXPHOs en los tumores de mama.
- Existe un umbral en el que variantes mitocondriales como la G8027A, G13590A, A10398G, G11914A y A13105C resultan patogénicas y contribuyen al desarrollo del tumor en la glándula mamaria.
- El gen *ND5*, que codifica para una subunidad estructural de la NADH-ubiquinona oxidoreductasa, es el que presenta una mayor cantidad de mutaciones somáticas.
- No existe un perfil de mutaciones somáticas mitocondriales asociada a cáncer de mama.
- Las mutaciones más frecuentes fueron T9950C (7.6%, *CO3*), C16278T (4.3%, D-Loop) y A11002G (5.4%, *ND4*), mismas que no han sido reportadas en cáncer.
- Los tumores luminal A, fueron los que presentaron el mayor número de mutaciones en el genoma mitocondrial, pero no se observó un perfil mutacional que caracterice a cada uno de los subtipos moleculares de cáncer de mama.

8 PERSPECTIVAS

- Desarrollar estrategias experimentales que permitan la caracterización funcional de mutaciones potencialmente relevantes en cáncer de mama.
- El empleo de bases de datos públicas como TCGA (con acceso a genomas y exomas completos) podría mejorar el poder estadístico de nuestro estudio, mediante la implementación de flujos de trabajo que permitan la extracción del genoma mitocondria.
- Aumentar el número de muestras de subtipos negativos a receptores de estrógenos para descartar la inexistencia de un perfil mutacional.
- Realizar la búsqueda de estas mutaciones dentro del genoma de línea germinal con el fin de identificar potenciales marcadores que puedan contribuir en el diagnóstico y en el establecimiento del pronóstico de la enfermedad.
- Correlacionar las mutaciones somáticas y germinales con parámetros clínicos como tamaño tumoral, estadio, sobrevida, etc.

9 REFERENCIAS

1. Hou, X.S., et al., *Mitochondria: promising organelle targets for cancer diagnosis and treatment*. *Biomater Sci*, 2018. **6**(11): p. 2786-2797.
2. Akhmedov, A.T. and J. Marin-Garcia, *Mitochondrial DNA maintenance: an appraisal*. *Mol Cell Biochem*, 2015. **409**(1-2): p. 283-305.
3. Stewart, J.B. and P.F. Chinnery, *The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease*. *Nat Rev Genet*, 2015. **16**(9): p. 530-42.
4. Ksiezakowska-Lakoma, K., M. Zyla, and J.R. Wilczynski, *Mitochondrial dysfunction in cancer*. *Prz Menopauzalny*, 2014. **13**(2): p. 136-44.
5. Alston, C.L., et al., *The genetics and pathology of mitochondrial disease*. *J Pathol*, 2017. **241**(2): p. 236-250.
6. Yadav, N. and D. Chandra, *Mitochondrial DNA mutations and breast tumorigenesis*. *Biochim Biophys Acta*, 2013. **1836**(2): p. 336-44.
7. Yadava, N., et al., *Impaired mitochondrial metabolism and mammary carcinogenesis*. *J Mammary Gland Biol Neoplasia*, 2013. **18**(1): p. 75-87.
8. Lenaz, G. and M.L. Genova, *Structure and organization of mitochondrial respiratory complexes: a new understanding of an old subject*. *Antioxid Redox Signal*, 2010. **12**(8): p. 961-1008.
9. Sousa, J.S., E. D'Imprima, and J. Vonck, *Mitochondrial Respiratory Chain Complexes*. *Subcell Biochem*, 2018. **87**: p. 167-227.
10. Wallace, D.C. and D. Chalkia, *Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease*. *Cold Spring Harb Perspect Biol*, 2013. **5**(11): p. a021220.
11. Perier, C. and M. Vila, *Mitochondrial biology and Parkinson's disease*. *Cold Spring Harb Perspect Med*, 2012. **2**(2): p. a009332.
12. Urra, F.A., et al., *The Mitochondrial Complex(I)ty of Cancer*. *Front Oncol*, 2017. **7**: p. 118.
13. Wirth, C., et al., *Structure and function of mitochondrial complex I*. *Biochim Biophys Acta*, 2016. **1857**(7): p. 902-14.
14. Sazanov, L.A., et al., *A long road towards the structure of respiratory complex I, a giant molecular proton pump*. *Biochem Soc Trans*, 2013. **41**(5): p. 1265-71.
15. Kluckova, K., et al., *Mitochondrial complex II, a novel target for anti-cancer agents*. *Biochim Biophys Acta*, 2013. **1827**(5): p. 552-64.
16. Kadenbach, B. and M. Huttemann, *The subunit composition and function of mammalian cytochrome c oxidase*. *Mitochondrion*, 2015. **24**: p. 64-76.
17. Wikstrom, M. and V. Sharma, *Proton pumping by cytochrome c oxidase - A 40year anniversary*. *Biochim Biophys Acta Bioenerg*, 2018. **1859**(9): p. 692-698.
18. Habersetzer, J., et al., *ATP synthase oligomerization: from the enzyme models to the mitochondrial morphology*. *Int J Biochem Cell Biol*, 2013. **45**(1): p. 99-105.
19. Wallace, D.C., *Bioenergetic origins of complexity and disease*. *Cold Spring Harb Symp Quant Biol*, 2011. **76**: p. 1-16.
20. Hejzlarova, K., et al., *Nuclear genetic defects of mitochondrial ATP synthase*. *Physiol Res*, 2014. **63 Suppl 1**: p. S57-71.
21. Jonckheere, A.I., J.A. Smeitink, and R.J. Rodenburg, *Mitochondrial ATP synthase: architecture, function and pathology*. *J Inherit Metab Dis*, 2012. **35**(2): p. 211-25.
22. Garcia-Aguilar, A. and J.M. Cuezva, *A Review of the Inhibition of the Mitochondrial ATP Synthase by IF1 in vivo: Reprogramming Energy Metabolism and Inducing Mitohormesis*. *Front Physiol*, 2018. **9**: p. 1322.
23. Chandra, D. and K.K. Singh, *Genetic insights into OXPHOS defect and its role in cancer*. *Biochim Biophys Acta*, 2011. **1807**(6): p. 620-5.

-
24. Bogenhagen, D.F., *Mitochondrial DNA nucleoid structure*. Biochim Biophys Acta, 2012. **1819**(9-10): p. 914-20.
 25. D'Aquila, P., et al., *Mitochondrial genome and epigenome: two sides of the same coin*. Front Biosci (Landmark Ed), 2017. **22**: p. 888-908.
 26. Gorman, G.S., et al., *Mitochondrial diseases*. Nat Rev Dis Primers, 2016. **2**: p. 16080.
 27. Nicholls, T.J. and M. Minczuk, *In D-loop: 40 years of mitochondrial 7S DNA*. Exp Gerontol, 2014. **56**: p. 175-81.
 28. Penta, J.S., et al., *Mitochondrial DNA in human malignancy*. Mutat Res, 2001. **488**(2): p. 119-33.
 29. Pakendorf, B. and M. Stoneking, *Mitochondrial DNA and human evolution*. Annu Rev Genomics Hum Genet, 2005. **6**: p. 165-83.
 30. Zastawny, T.H., et al., *Comparison of oxidative base damage in mitochondrial and nuclear DNA*. Free Radic Biol Med, 1998. **24**(5): p. 722-5.
 31. Genomes Project, C., et al., *A map of human genome variation from population-scale sequencing*. Nature, 2010. **467**(7319): p. 1061-73.
 32. Schneider, S. and L. Excoffier, *Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA*. Genetics, 1999. **152**(3): p. 1079-89.
 33. Brandon, M., P. Baldi, and D.C. Wallace, *Mitochondrial mutations in cancer*. Oncogene, 2006. **25**(34): p. 4647-62.
 34. Guardado-Estrada, M., et al., *The Amerindian mtDNA haplogroup B2 enhances the risk of HPV for cervical cancer: de-regulation of mitochondrial genes may be involved*. J Hum Genet, 2012. **57**(4): p. 269-76.
 35. Martinez-Cortes, G., et al., *Maternal admixture and population structure in Mexican-Mestizos based on mtDNA haplogroups*. Am J Phys Anthropol, 2013. **151**(4): p. 526-37.
 36. Singh, K.K. and M. Kulawiec, *Mitochondrial DNA polymorphism and risk of cancer*. Methods Mol Biol, 2009. **471**: p. 291-303.
 37. Guardado-Estrada, M., et al., *A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population*. J Hum Genet, 2009. **54**(12): p. 695-705.
 38. Stefano, G.B. and R.M. Kream, *Mitochondrial DNA heteroplasmy in human health and disease*. Biomed Rep, 2016. **4**(3): p. 259-262.
 39. Trifunovic, A. and N.G. Larsson, *Mitochondrial dysfunction as a cause of ageing*. J Intern Med, 2008. **263**(2): p. 167-78.
 40. Sun, L., et al., *Metabolic reprogramming for cancer cells and their microenvironment: Beyond the Warburg Effect*. Biochim Biophys Acta Rev Cancer, 2018. **1870**(1): p. 51-66.
 41. Warburg, O., F. Wind, and E. Negelein, *The Metabolism of Tumors in the Body*. J Gen Physiol, 1927. **8**(6): p. 519-30.
 42. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
 43. Moreno-Sanchez, R., et al., *Energy metabolism in tumor cells*. FEBS J, 2007. **274**(6): p. 1393-418.
 44. Snyder, V., et al., *Cancer Stem Cell Metabolism and Potential Therapeutic Targets*. Front Oncol, 2018. **8**: p. 203.
 45. De Santis, M.C., et al., *Signaling Pathways Regulating Redox Balance in Cancer Metabolism*. Front Oncol, 2018. **8**: p. 126.
 46. Lommarini, L., et al., *Mitochondrial metabolism and energy sensing in tumor progression*. Biochim Biophys Acta Bioenerg, 2017. **1858**(8): p. 582-590.
 47. Choudhury, A.R. and K.K. Singh, *Mitochondrial determinants of cancer health disparities*. Semin Cancer Biol, 2017. **47**: p. 125-146.
 48. Jimenez-Morales, S., et al., *Overview of mitochondrial germline variants and mutations in human disease: Focus on breast cancer (Review)*. Int J Oncol, 2018. **53**(3): p. 923-936.
-

-
49. Rodic, S. and M.D. Vincent, *Reactive oxygen species (ROS) are a key determinant of cancer's metabolic phenotype*. *Int J Cancer*, 2018. **142**(3): p. 440-448.
 50. Yu, C., et al., *Deciphering the Spectrum of Mitochondrial DNA Mutations in Hepatocellular Carcinoma Using High-Throughput Sequencing*. *Gene Expr*, 2018. **18**(2): p. 125-134.
 51. Cazzaniga, M. and B. Bonanni, *Breast Cancer Metabolism and Mitochondrial Activity: The Possibility of Chemoprevention with Metformin*. *Biomed Res Int*, 2015. **2015**: p. 972193.
 52. Breastcancer.org. Enero-2019]; Available from: <https://www.breastcancer.org/>
 53. Akram, M., et al., *Awareness and current knowledge of breast cancer*. *Biol Res*, 2017. **50**(1): p. 33.
 54. Rakha, E.A. and A.R. Green, *Molecular classification of breast cancer: what the pathologist needs to know*. *Pathology*, 2017. **49**(2): p. 111-119.
 55. Society, A.C. [cited Enero-2019; Available from: <https://www.cancer.org/cancer/breast-cancer/about.html>.
 56. Rojas, K. and A. Stuckey, *Breast Cancer Epidemiology and Risk Factors*. *Clin Obstet Gynecol*, 2016. **59**(4): p. 651-672.
 57. Organization, W.H. Enero-2019]; Available from: <https://www.who.int/en/news-room/fact-sheets/detail/cancer>
 58. INEGI. Diciembre 2018]; Available from: http://www.beta.inegi.org.mx/contenidos/saladeprensa/aproposito/2018/cancer2018_Nal.pdf.
 59. Maffuz-Aziz, A., et al., *[Clinical and pathological features of breast cancer in a population of Mexico]*. *Cir Cir*, 2017. **85**(3): p. 201-207.
 60. Turashvili, G. and E. Brogi, *Tumor Heterogeneity in Breast Cancer*. *Front Med (Lausanne)*, 2017. **4**: p. 227.
 61. Provenzano, E., G.A. Ulaner, and S.F. Chin, *Molecular Classification of Breast Cancer*. *PET Clin*, 2018. **13**(3): p. 325-338.
 62. Dai, X., et al., *Breast cancer intrinsic subtype classification, clinical use and future trends*. *Am J Cancer Res*, 2015. **5**(10): p. 2929-43.
 63. Horton, J.K., et al., *Breast Cancer Biology: Clinical Implications for Breast Radiation Therapy*. *Int J Radiat Oncol Biol Phys*, 2018. **100**(1): p. 23-37.
 64. Robert, M., et al., *Emerging PARP inhibitors for treating breast cancer*. *Expert Opin Emerg Drugs*, 2018. **23**(3): p. 211-221.
 65. Peshkin, B.N., M.L. Alabek, and C. Isaacs, *BRCA1/2 mutations and triple negative breast cancers*. *Breast Dis*, 2010. **32**(1-2): p. 25-33.
 66. S.A., M.D.M., ed. *Séptima Revisión del Consenso Mexicano sobre el Diagnóstico y Tratamiento del Cáncer Mamario*. . 2017: Colima, México. 119.
 67. Bai, R.K., et al., *Mitochondrial genetic background modifies breast cancer risk*. *Cancer Res*, 2007. **67**(10): p. 4687-94.
 68. Zhu, W., et al., *Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid*. *Carcinogenesis*, 2005. **26**(1): p. 145-52.
 69. Lin, C.S., et al., *Mitochondrial DNA alterations correlate with the pathological status and the immunological ER, PR, HER-2/neu, p53 and Ki-67 expression in breast invasive ductal carcinoma*. *Oncol Rep*, 2015. **33**(6): p. 2924-34.
 70. Parrella, P., et al., *Detection of mitochondrial DNA mutations in primary breast cancer and fine-needle aspirates*. *Cancer Res*, 2001. **61**(20): p. 7623-6.
 71. Xu, C., et al., *Mitochondrial D310 mutations in the early development of breast cancer*. *Br J Cancer*, 2012. **106**(9): p. 1506-11.
 72. Yu, M., et al., *Sequence variations of mitochondrial DNA D-loop region are highly frequent events in familial breast cancer*. *J Biomed Sci*, 2008. **15**(4): p. 535-43.
 73. Alhomidi, M.A., et al., *Mitochondrial D310 instability in Asian Indian breast cancer patients*. *Tumour Biol*, 2013. **34**(4): p. 2427-32.
-

-
74. Cai, F.F., et al., *Mutations of mitochondrial DNA as potential biomarkers in breast cancer*. *Anticancer Res*, 2011. **31**(12): p. 4267-71.
 75. Kuo, S.J., et al., *Number of somatic mutations in the mitochondrial D-loop region indicates poor prognosis in breast cancer, independent of TP53 mutation*. *Cancer Genet Cytogenet*, 2010. **201**(2): p. 94-101.
 76. Aral, C., et al., *A novel approach for rapid screening of mitochondrial D310 polymorphism*. *BMC Cancer*, 2006. **6**: p. 21.
 77. Wang, Y., et al., *Microsatellite instability in mitochondrial genome of common female cancers*. *Int J Gynecol Cancer*, 2006. **16 Suppl 1**: p. 259-66.
 78. Bianchi, M.S., N.O. Bianchi, and G. Bailliet, *Mitochondrial DNA mutations in normal and tumor tissues from breast cancer patients*. *Cytogenet Cell Genet*, 1995. **71**(1): p. 99-103.
 79. Li, L., et al., *Correlational study on mitochondrial DNA mutations as potential risk factors in breast cancer*. *Oncotarget*, 2016. **7**(21): p. 31270-83.
 80. Canter, J.A., et al., *Mitochondrial DNA G10398A polymorphism and invasive breast cancer in African-American women*. *Cancer Res*, 2005. **65**(17): p. 8028-33.
 81. Covarrubias, D., et al., *Mitochondrial DNA variant interactions modify breast cancer risk*. *J Hum Genet*, 2008. **53**(10): p. 924-8.
 82. Darvishi, K., et al., *Mitochondrial DNA G10398A polymorphism imparts maternal Haplogroup N a risk for breast and esophageal cancer*. *Cancer Lett*, 2007. **249**(2): p. 249-55.
 83. Fang, H., et al., *Cancer type-specific modulation of mitochondrial haplogroups in breast, colorectal and thyroid cancer*. *BMC Cancer*, 2010. **10**: p. 421.
 84. Tengku Baharudin, N., H. Jaafar, and Z. Zainuddin, *Association of mitochondrial DNA 10398 polymorphism in invasive breast cancer in malay population of peninsular malaysia*. *Malays J Med Sci*, 2012. **19**(1): p. 36-42.
 85. Ahn, E.H., et al., *Detection of Ultra-Rare Mitochondrial Mutations in Breast Stem Cells by Duplex Sequencing*. *PLoS One*, 2015. **10**(8): p. e0136216.
 86. McMahon, S. and T. LaFramboise, *Mutational patterns in the breast cancer mitochondrial genome, with clinical correlates*. *Carcinogenesis*, 2014. **35**(5): p. 1046-54.
 87. Tan, D.J., R.K. Bai, and L.J. Wong, *Comprehensive scanning of somatic mitochondrial DNA mutations in breast cancer*. *Cancer Res*, 2002. **62**(4): p. 972-6.
 88. Shen, L., et al., *Evaluating mitochondrial DNA in patients with breast cancer and benign breast disease*. *J Cancer Res Clin Oncol*, 2011. **137**(4): p. 669-75.
 89. Tseng, L.M., et al., *Somatic mutations of the mitochondrial genome in human breast cancers*. *Genes Chromosomes Cancer*, 2011. **50**(10): p. 800-11.
 90. Dong, J., L.J. Wong, and M.P. Mims, *Mitochondrial inheritance and cancer*. *Transl Res*, 2018. **202**: p. 24-34.
 91. Ma, L., et al., *Breast cancer-associated mitochondrial DNA haplogroup promotes neoplastic growth via ROS-mediated AKT activation*. *Int J Cancer*, 2018. **142**(9): p. 1786-1796.
 92. Lott, M.T., et al., *mtDNA Variation and Analysis Using Mitomap and Mitomaster*. *Curr Protoc Bioinformatics*, 2013. **44**: p. 1 23 1-26.
 93. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data*. *Genome Res*, 2010. **20**(9): p. 1297-303.
 94. Kim, S., et al., *Strelka2: fast and accurate calling of germline and somatic variants*. *Nat Methods*, 2018. **15**(8): p. 591-594.
 95. Andrews, S. *FastQC: a quality control tool for high throughput sequence data*. 2010 [cited 2018 Enero 2018]; Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
 96. Krueger, F. *Wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files*. 2012; Available from: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
-

-
97. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report*. Bioinformatics, 2016. **32**(19): p. 3047-8.
 98. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform*. Bioinformatics, 2010. **26**(5): p. 589-95.
 99. Foundation, P.S. *Python Language Reference, version 2.7*. 2016 [cited 2018 Septiembre 2018]; Available from: <http://www.python.org>.
 100. passing, P.b. *Picard, a set of command line tools for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF*. 2015 [cited 2018 Febrero 2018]; Available from: <https://broadinstitute.github.io/picard/>.
 101. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. Bioinformatics, 2009. **25**(16): p. 2078-9.
 102. Team, R.C. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing 2014 [cited 2018 Agosto 2018]; Available from: <http://www.R-project.org/>
 103. McLaren, W., et al., *The Ensembl Variant Effect Predictor*. Genome Biol, 2016. **17**(1): p. 122.
 104. Weissensteiner, H., et al., *HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing*. Nucleic Acids Res, 2016. **44**(W1): p. W58-63.
 105. van Oven, M. and M. Kayser, *Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation*. Hum Mutat, 2009. **30**(2): p. E386-94.
 106. Ebner, S., et al., *Mitochondrial haplogroups, control region polymorphisms and malignant melanoma: a study in middle European Caucasians*. PLoS One, 2011. **6**(12): p. e27192.
 107. Chen, J.Z., et al., *Extensive somatic mitochondrial mutations in primary prostate cancer using laser capture microdissection*. Cancer Res, 2002. **62**(22): p. 6470-4.
 108. Warburg, O., [Origin of cancer cells]. Oncologia, 1956. **9**(2): p. 75-83.
 109. Seyfried, T.N., et al., *Cancer as a metabolic disease: implications for novel therapeutics*. Carcinogenesis, 2014. **35**(3): p. 515-27.
 110. Nemeth, K., et al., *Next-generation sequencing identifies novel mitochondrial variants in pituitary adenomas*. J Endocrinol Invest, 2019.
 111. Pagani, I.S., et al., *A Method for Next-Generation Sequencing of Paired Diagnostic and Remission Samples to Detect Mitochondrial DNA Mutations Associated with Leukemia*. J Mol Diagn, 2017. **19**(5): p. 711-721.
 112. Gomez-Flores-Ramos, L., et al., *Molecular Biology In Young Women With Breast Cancer: From Tumor Gene Expression To DNA Mutations*. Rev Invest Clin, 2017. **69**(4): p. 181-192.
 113. Ju, Y.S., et al., *Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer*. Elife, 2014. **3**.
 114. Shen, J., et al., *Peripheral blood mitochondrial DNA copy number, length heteroplasmy and breast cancer risk: a replication study*. Carcinogenesis, 2015. **36**(11): p. 1307-13.
 115. Letts, J.A. and L.A. Sazanov, *Clarifying the supercomplex: the higher-order organization of the mitochondrial electron transport chain*. Nat Struct Mol Biol, 2017. **24**(10): p. 800-808.
 116. Gu, J., et al., *The architecture of the mammalian respirasome*. Nature, 2016. **537**(7622): p. 639-43.
 117. Garcia, J.J., et al., *Structure, functioning, and assembly of the ATP synthase in cells from patients with the T8993G mitochondrial DNA mutation. Comparison with the enzyme in Rho(0) cells completely lacking mtdna*. J Biol Chem, 2000. **275**(15): p. 11075-81.
 118. Petros, J.A., et al., *mtDNA mutations increase tumorigenicity in prostate cancer*. Proc Natl Acad Sci U S A, 2005. **102**(3): p. 719-24.
 119. Deng, N., et al., *Single nucleotide polymorphisms and cancer susceptibility*. Oncotarget, 2017. **8**(66): p. 110635-110649.
 120. Adzhubei, A.A., et al., *Non-random usage of 'degenerate' codons is related to protein three-dimensional structure*. FEBS Lett, 1996. **399**(1-2): p. 78-82.
-

-
121. Gomez-Velazquez, M., et al., *CTCF counter-regulates cardiomyocyte development and maturation programs in the embryonic heart*. PLoS Genet, 2017. **13**(8): p. e1006985.
 122. Marinov, G.K., et al., *Evidence for site-specific occupancy of the mitochondrial genome by nuclear transcription factors*. PLoS One, 2014. **9**(1): p. e84713.
 123. Herrnstadt, C., et al., *Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups*. Am J Hum Genet, 2002. **70**(5): p. 1152-71.
 124. Yeh, J.J., et al., *Somatic mitochondrial DNA (mtDNA) mutations in papillary thyroid carcinomas and differential mtDNA sequence variants in cases with thyroid tumours*. Oncogene, 2000. **19**(16): p. 2060-6.
 125. Webb, E., et al., *Comprehensive analysis of common mitochondrial DNA variants and colorectal cancer risk*. Br J Cancer, 2008. **99**(12): p. 2088-93.
 126. Liu, V.W., et al., *High incidence of somatic mitochondrial DNA mutations in human ovarian carcinomas*. Cancer Res, 2001. **61**(16): p. 5998-6001.
 127. Weerts, M.J.A., et al., *Tumor-Specific Mitochondrial DNA Variants Are Rarely Detected in Cell-Free DNA*. Neoplasia, 2018. **20**(7): p. 687-696.
 128. Dornelles, C.L., et al., *Is haplogroup X present in extant South American Indians?* Am J Phys Anthropol, 2005. **127**(4): p. 439-48.
 129. Fatimata Mbaye, A.D., Malick Fall, Gora Diop, Babacar Mbengue, Rokhaya N. Diallo, Maguette S. Niang, Mamadou Kane, Sidy Ka, Alioune Dieye, Mbacké Sembene, *Genetic Diversity of Breast Cancer in Senegalese Women: New Insight from Somatic Mutations*. Journal of Health Science, 2014. **4**(2): p. 25-33.
 130. Rohan, T.E., et al., *Do alterations in mitochondrial DNA play a role in breast carcinogenesis?* J Oncol, 2010. **2010**: p. 604304.
 131. Roy, D., et al., *Estrogen-induced generation of reactive oxygen and nitrogen species, gene damage, and estrogen-dependent cancers*. J Toxicol Environ Health B Crit Rev, 2007. **10**(4): p. 235-57.
 132. Inic, Z., et al., *Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information*. Clin Med Insights Oncol, 2014. **8**: p. 107-11.
 133. Creighton, C.J., *The molecular profile of luminal B breast cancer*. Biologics, 2012. **6**: p. 289-97.
 134. Onuchic, V., et al., *Epigenomic Deconvolution of Breast Tumors Reveals Metabolic Coupling between Constituent Cell Types*. Cell Rep, 2016. **17**(8): p. 2075-2086.
 135. Skibinski, A. and C. Kuperwasser, *The origin of breast tumor heterogeneity*. Oncogene, 2015. **34**(42): p. 5309-16.
 136. Penalzoza-Espinosa, R.I., et al., *Characterization of mtDNA haplogroups in 14 Mexican indigenous populations*. Hum Biol, 2007. **79**(3): p. 313-20.
 137. Weigl, S., A. Paradiso, and S. Tommasi, *Mitochondria and familial predisposition to breast cancer*. Curr Genomics, 2013. **14**(3): p. 195-203.
 138. Blein, S., et al., *An original phylogenetic approach identified mitochondrial haplogroup T1a1 as inversely associated with breast cancer risk in BRCA2 mutation carriers*. Breast Cancer Res, 2015. **17**: p. 61.
 139. Kenney, M.C., et al., *Molecular and bioenergetic differences between cells with African versus European inherited mitochondrial DNA haplogroups: implications for population susceptibility to diseases*. Biochim Biophys Acta, 2014. **1842**(2): p. 208-19.
 140. Ferreira, A., et al., *Role of mtDNA-related mitopigenetic phenomena in cancer*. Eur J Clin Invest, 2015. **45 Suppl 1**: p. 44-9.
-

10 PRODUCCIÓN CIENTÍFICA

10.1 REUNIONES CIENTÍFICAS

1. **Carlos Pérez-Amado**, Hugo Tovar-Romero, Verónica Bautista-Piña, Luis Alfaro-Ruíz, Alfredo Hidalgo-Miranda, Silvia Jiménez-Morales. *Mitochondrial DNA mutation análisis in molecular subtypes of breast cancer*. European Society Human Genetics 2018, Milán, Italia. 16-19 de junio, 2018.
2. Silvia Jiménez-Morales, **Carlos Jhovani Pérez-Amado**, Hugo Tovar-Romero, Laura Gómez-Romero, Verónica Bautista-Piña, Fredy Beltrán-Anaya, Alberto Cedro-Tanda, Luis Alfaro-Ruíz, Diana Flores-Hunosta, Alfredo Hidalgo-Miranda. *Mitochondrial DNA mutations showing low heteroplasmy level in peripheral blood are enriched in tumors of breast cancer*. Human Genome Meeting 2019 Seúl, Corea del Sur, 24-26 de abril, 2019.
3. Silvia Jiménez-Morales, **Carlos Jhovani Pérez-Amado**, Hugo Tovar-Romero, Antonio Maffuz-Azziz, Verónica Bautista-Piña, Sergio Rodríguez-Cuevas, Fredy Beltrán-Anaya, Alberto Cedro-Tanda, Luis Alfaro-Ruíz, Alfredo Hidalgo-Miranda. (2018). *Mitochondrial DNA analysis in Mexican women with breast cancer*. *Human Genomics* 2018, 12(Suppl 1):9 <https://doi.org/10.1186/s40246-018-0138-6> ABSTRACT PUBLICADO

10.2 PUBLICACIONES EN CÁNCER DE MAMA

1. Jiménez-Morales, S., **Pérez-Amado, C.J.**, Langley, E., & Hidalgo-Miranda, A. (2018). *Overview of mitochondrial germline variants and mutations in human disease: Focus on breast cancer (Review)*. *International Journal of Oncology*, 53, 923-936. <https://doi.org/10.3892/ijo.2018.4468>
2. **C.J. Pérez-Amado**, H. Tovar-Romero, L.L. Gómez-Romero, FO Beltrán-Anaya, L.A. Alfaro-Ruíz, A. Cedro-Tanda, V. Bautista-Piña, A. Hidalgo-Miranda, S. Jiménez-Morales. *Landscape of mitochondrial mutations in mexican women with breast cancer*. EN REVISIÓN POR LOS COAUTORES.

10.3 PUBLICACIONES EN OTROS TEMAS

1. Rodríguez-Cortés, Y.M, Amaya-Chávez, A., García-Fabila, M., Jiménez-Morales, S., **Pérez-Amado, C.J.** & Mendieta Zerón, H. (2019). *TSH β X1 Splice Variant Expression and D2 Thr92Ala Polymorphism Analysis in Pregnant Women with Thyroid Diseases*. SN Comprehensive Clinical Medicine. pp 1–7 <https://doi.org/10.1007/s42399-019-00084-3>
2. Mendieta-Zerón, H., Jiménez-Rosales, A., **Pérez-Amado, C.J.**, & Jiménez-Morales, S. (2017). *FOXE1 Mutation Screening in a Case with Cleft Lip, Hypothyroidism, and Thyroid Carcinoma: A New Syndrome?*. Case reports in genetics, 6390545. doi:10.1155/2017/6390545

Overview of mitochondrial germline variants and mutations in human disease: Focus on breast cancer (Review)

SILVIA JIMÉNEZ-MORALES¹, CARLOS J. PÉREZ-AMADO²,
ELIZABETH LANGLEY³ and ALFREDO HIDALGO-MIRANDA¹

¹Laboratory of Cancer Genomics, National Institute of Genomic Medicine, 14610 Mexico City;

²Biochemistry Sciences Program, National Autonomous University of Mexico, 04510 Mexico City;

³Department of Basic Research, National Cancer Institute, 14080 Mexico City, Mexico

Received January 29, 2018; Accepted May 18, 2018

DOI: 10.3892/ijo.2018.4468

Abstract. High lactate production in cells during growth under oxygen-rich conditions (aerobic glycolysis) is a hallmark of tumor cells, indicating the role of mitochondrial function in tumorigenesis. In fact, enhanced mitochondrial biogenesis and impaired quality control are frequently observed in cancer cells. Mitochondrial DNA (mtDNA) encodes 13 subunits of oxidative phosphorylation (OXPHOS), is present in thousands of copies per cell, and has a very high mutation rate. Mutations in mtDNA and nuclear DNA (nDNA) genes encoding proteins that are important players in mitochondrial biogenesis and function are involved in oncogenic processes. A wide range of germline mtDNA polymorphisms, as well as tumor mtDNA somatic mutations have been identified in diverse cancer types. Approximately 72% of supposed tumor-specific somatic mtDNA mutations reported, have also been found as polymorphisms in the general population. The *ATPase 6* and *NADH* dehydrogenase subunit genes of mtDNA are the most commonly mutated genes in breast cancer (BC). Furthermore, nuclear genes playing a role in mitochondrial biogenesis and function, such as peroxisome proliferators-activated receptor gamma coactivator-1 (*PGC-1*), fumarate hydratase (*FH*) and succinate dehydrogenase (*SDH*) are frequently mutated in cancer. In this review, we provide an overview of the mitochondrial germline variants and mutations in cancer, with particular focus on those found in BC.

Contents

1. Introduction
2. Mitochondria
3. Mitochondrial DNA variants in human diseases
4. Mitochondria variants and cancer
5. Mitochondrial DNA variants in breast cancer
6. Mitochondria DNA mutations and response to therapy in breast cancer
7. Conclusions

1. Introduction

Breast cancer (BC) is the most common epidermal neoplasia worldwide, representing approximately 12% of all new cancer cases and 25% of all cancer cases in women. Belgium has the highest rate of BC, followed by Denmark and France, while the lowest incidence has been observed in Asia and Africa (1). The Breast Cancer Organization estimates that approximately 12% of women from the United States will develop invasive BC during the course of their lifetime. Over the past decade, tremendous progress has been made in identifying BC etiological factors, understanding tumor biology, developing tools for better detection and determining the most effective treatments. Nevertheless, this disease still represents an important cause of morbi-mortality due to cancer (1,2). It is widely known that the oncogenesis phenomenon is a complex biological mechanism not yet well understood. Cancer development is a multistep process, involving genome instability, which leads to sustained proliferative signaling, the evasion of growth suppressors, resistance to cell death, tumor angiogenesis, invasion and metastasis (3). Other hallmarks of the cancer cell are the capacity to evade the immune system and to generate reprogramming of the energy metabolism system (4,5). As regards the latter hallmark, the most well-known tumor cell metabolic abnormality is the Warburg effect (4).

Otto Warburg observed that cancer cells prefer aerobic glycolysis to oxidative phosphorylation; however, these cells do not uptake oxygen like normal tissue cells, even under a normal oxygen environment (4,6). Subsequently, there is ample evidence to support the hypothesis that tumor-related metabolic abnormalities involve defects in mitochondrial

Correspondence to: Dr Silvia Jiménez-Morales or Dr Alfredo Hidalgo-Miranda, Laboratory of Cancer Genomics, National Institute of Genomic Medicine, Periférico Sur No. 4809, Col. Arenal Tepepan, Delegación Tlalpan, 14610 Mexico City, Mexico
E-mail: sjimenez@inmegen.gob.mx
E-mail: ahidalgo@inmegen.gob.mx

Abbreviations: mtDNA, mitochondria DNA; BC, breast cancer; OXPHOS, oxidative phosphorylation; NEMtG, nuclear-encoded mitochondrial genes

Key words: mitochondria DNA variants, mutations, breast cancer, haplogroups

function, since many of the metabolic genes whose mutations can cause cancers are mitochondrial genes or nuclear encoded genes that are relevant to mitochondrial biogenesis and function. As an example, the metabolic enzyme, succinate dehydrogenase (SDH), is frequently mutated in cancer (7). In this review, we provide an overview of mitochondrial mutations and polymorphisms in cancer, emphasizing those in BC.

2. Mitochondria

Mitochondria are involved in fundamental cellular processes, which include the generation of adenosine triphosphate (ATP) for cellular energy, calcium storage for cell signaling, and mediating cell growth and death processes. Mitochondria, descendants of endosymbiotic α -proteobacteria that become established in a host cell, are double-membraned and semi-autonomous organelles in eukaryotic cells that can auto-replicate and are maternally inherited. This organelle has its own genome and is organized into DNA-protein complexes termed mitochondrial nucleoids, which are relatively stable genetic elements. Human cells have approximately 2,000 mitochondria per cell and an average of between 1.4 and 7.5 mitochondrial DNA (mtDNA) molecules per nucleoid (8). mtDNA is a negatively supercoiled double-stranded circular molecule of 16,569 bp in size. Based on sedimentation properties, one strand is termed light (L) and the other one heavy (H). With only <7% of the sequence considered non-coding, mtDNA contains 37 genes, 28 of which are encoded in the H strand and 9 in the L strand. Thirteen of these genes code for essential proteins for mitochondrial oxidative phosphorylation (OXPHOS). These include 7 of nearly 45 polypeptides of the electron transport chain complex I (ND1-3, ND4L and ND4-6); 1 of the 11 polypeptides of complex III (Cytb), 3 of the 13 polypeptides of complex IV (COI-III), and 2 of the 15 polypeptides of complex V (ATP6 and ATP8). mtDNA also encodes 22 transfer RNAs (tRNAs), as well as the 18S and 16S ribosomal RNAs (rRNA). tRNAs and rRNAs are necessary for the translation of the respiratory subunit mRNAs within the mitochondrial matrix (Table I) (8-10). The hypervariable region (HVR1: 16024-16383 nt and HVR2: 57-333 nt) or D-loop region is a non-coding sequence; however, it contains the origin of replication of the mtDNA (8-11). The proximity of mtDNA to the site of production of reactive oxygen species (ROS) and the absence of protective histones increase mtDNA susceptibility for acquiring mutations and oxidative damage, causing a mutation rate of 100-1,000-fold higher for mtDNA than nuclear DNA (nDNA) ($1.1-1.7 \times 10^{-8}$ per nucleotide site per generation). Moreover, mtDNA damage induced by ROS is more extensive and persists longer than ROS-nDNA injury in human cells (12,13). Mitochondrial DNA encodes only a small number of proteins needed for its molecular architecture and biological functions; hence, additional nuclear-encoded mitochondrial (NEMt) proteins (NEMtPs) are required for mitochondrial biogenesis, assembly of the respiratory chain and the maintenance of mtDNA (Fig. 1) (14). The biogenesis of these organelles, which implies variations in number, size and mass, is highly influenced by factors, such as cell division, renewal, differentiation, oxidative stress, exercise, caloric restriction, low temperature, and mutations in both nDNA and mtDNA (14,15).

Nuclear-encoded genes that participate in mitochondrial biogenesis and function. In addition to the 13 mitochondrial encoded genes, to function properly the mitochondria require >1,500 NEMtPs. These proteins are required by the mitochondria to perform DNA replication, transcription and translation processes, as well as to carry out biochemical functions and to maintain its molecular architecture (Fig. 1). Almost 1% of the products contained in the mitochondria are synthesized in the matrix and the remainder, in the cytosol. Mitochondrial proteins synthesized in the cytosol are encoded by nuclear genes derived from prokaryotic genes that were transferred to the nucleus or that originated during the evolution of eukaryotic cells. Since NEMt genes (NEMtGs or mitonuclear genes) include proteins and enzymes involved in mtDNA replication, transcription and translation, catabolytic pathways, and are components of the mitochondrial import and folding machineries, most of them are strictly essential for life (16,17). One of the most extensively studied NEMtG is the mitochondrial transcription factor A (*TFAM*) that regulates mtDNA replication/transcription and is implicated in the dynamics of mtDNA nucleoids (18-21). *TFAM* expression is promoted by the nuclear respiratory factor transcriptional regulators (NRF)-1 and NRF-2, which are activators of nuclear genes necessary for multiple mitochondrial functions. For example, the biogenesis of the mitochondria is regulated by the induction of peroxisome proliferator-activated receptor γ coactivator-1 (PGC-1) (18,19,22). Additionally, it has been reported that the absence of *TFAM* in knockout mice is responsible for a severe OXPHOS defect, marked reduction in mtDNA content and embryonic lethality (20). Furthermore, the loss of *PolgA*, encoding a subunit of mitochondrial DNA polymerase, also results in lethality in mice (23). This observation suggests that mutations in NEMtGs can affect functions, such as electron transport and oxidative phosphorylation, protein translocation and mitochondrial biogenesis. The translocase of the outer membrane (TOM) complex is a clear example of the relevant role of mitonuclear genes in the cell, since nearly all mitochondrial pre-proteins are imported via a TOM entry gate (24). From the cytosol to the TOM complex, pre-proteins are guided by molecular chaperones (heat shock protein 90 or heat shock cognate 70) (25); there, TOM20, TOM22 and TOM70 recognize the mitochondrial targeting signals of cytosolic preproteins (Fig. 1). Hence, it is clear that errors in the import and assembly of mitochondrial proteins can result in a disease state in humans (26). Additionally, nDNA genes have also been found to be involved in >30 different modified mt-tRNAs, and interestingly, Bandiera *et al* (27) found a large number of nuclear-encoded miRNAs, designated 'mitomiRs', with differential expression in the mitochondria and cytosol, demonstrating that the majority of mitochondrial miRNAs have both nuclear and mitochondrial-encoded targets.

3. Mitochondrial DNA variants in human diseases

Mitochondria are involved in the fundamental cellular process of generating ATP for cellular energy. However, energy metabolism is not the only essential role of mitochondria, since they are also crucial for stress response, cell survival and death, immune response and cell signaling. In consequence, mitochondrial impairment contributes to a wide spectrum of

Table I. Nuclear DNA and mitochondrial DNA features.

Features	Nuclear DNA	Mitochondrial DNA
Base pairs (bp)	3.2 billion	16,569
Strand	Double, lineal, forward (5'-3') and reverse (3'-5')	Double, circular, light and heavy
Organization	Chromosomes (23 pairs)	Nucleoids
Coding region	2-3%	90%
Genes		
Number	22,000 approx.	37 L (forward): 8 tRNA and 1 polypeptide H (reverse): 12 polypeptides, 14 tRNAs, 2 RNAs
Structure	Exon, intron, intergenic regions	Polycistronic
Variants	SNPs, indel, CNVs, STRs, VNTRs, translocations, inversions	SNPs, Indels, CN
Ploidy	Diploid	Polypliod
Genetic code		
Transduction codons	32	24
Initiation codon	AUG	AUG, AUA, AUU
Stop codon	UAA, UAG, UGA	AGA, AGG
Transcription	Multiple	IT _L , IT _{H1} , IT _{H2}
Genotype	Homozygote and heterozygote	Homoplasmic, heteroplasmic
Inheritance	Maternal and paternal	Maternal
Recombination	Homologous chromosomes	No evidence

SNPs, single nucleotide polymorphisms; indel, insertion/deletion; CNVs, copy number variants; STR, short tandem repeat; VNTR, variable number of tandem repeat; CV, copy number.

heterogeneous human conditions (mental disorders, cardiomyopathies, metabolic diseases, immune mediated diseases, aging and cancer); preferentially affecting those tissues with high-energy demands, such as brain, muscle and liver (12,28,29). The strongest evidence linking a mtDNA mutation to a human phenotype is the increased risk of developing blindness in subjects harboring mutations in the mitochondrial encoded complex I genes (ND1-6 and ND4L). These mutations are known to cause Leber hereditary optic neuropathy (LHON), synergistically interacting with a primary LHON mutation to cause a defect in OXPHOS complex I activity (30).

Mitochondrial disorders also cause an overlapping spectrum of diseases. In a case-control study, Hudson *et al* (29) analyzed mtDNA single nucleotide polymorphisms (SNPs) from 51,106 subjects from the Wellcome Trust Case Consortium, which included patients with ankylosing spondylitis, ischemic stroke, multiple sclerosis, Parkinson's disease, primary biliary cirrhosis, psoriasis, schizophrenia, ulcerative colitis, coronary artery disease, hypertension and type 2 diabetes. The authors observed that mtDNA polymorphisms modify the risk of developing one or more of these diseases. Of note, high-risk alleles were more frequent than protective alleles, indicating that mtDNA is not at equilibrium in the human population, and that recent mutations interact with nuclear loci to modify the risk of developing diseases (29). Notably, polymorphisms that increase the risk of developing two or more diseases were limited to mitochondrial cytochrome *b* (*CYB*: *H16R*, *T158A*) and subunit 3 of cytochrome *c* oxidase (*COIII*:

V91L, *N154S*) genes. The *CYB* variants *H16R* and *T158A* were previously associated with several diseases (31,32). The only non-synonymous polymorphism associated with a reduced risk of various entities was the G10398A (T114A) of *ND3*, which occurs twice on the human mtDNA phylogeny and has previously been associated with Parkinson's disease and cardiomyopathies (33-36). This polymorphism has been shown to reduce complex I activity, cytosolic calcium levels, and mitochondrial membrane potential, and thus may reduce the levels of ROS (33).

As regards NEMtGs, mutations in >100 genes have been implicated in mitochondrial human diseases (37-40). As an example, mutations in *elaC* ribonuclease *Z2* (*ELAC2*) have been found in individuals suffering from infantile-onset hypertrophic cardiomyopathy and complex I deficiency (39,41). Similarly, defects in any step of protein transport due to TOM complex mutations lead to oxidative stress, neurodegenerative diseases, and metabolic disorders (26). Irregularities in mitochondrial targeting signals have also been identified as a cause of pyruvate dehydrogenase (PDH) deficiency. These mutations were described in the amino-terminal targeting signal (N-MTS) of the PDH, which is a subunit of the mitochondrial matrix protein complex. X-linked mutations in the PDH E1 α subunit are the most common cause of PDH deficiency and have been associated to microcephaly and cerebral atrophy (42). Moreover, mutations in nuclear-encoded components of the mitochondrial translation machinery have been shown to be involved in other human diseases (31). To

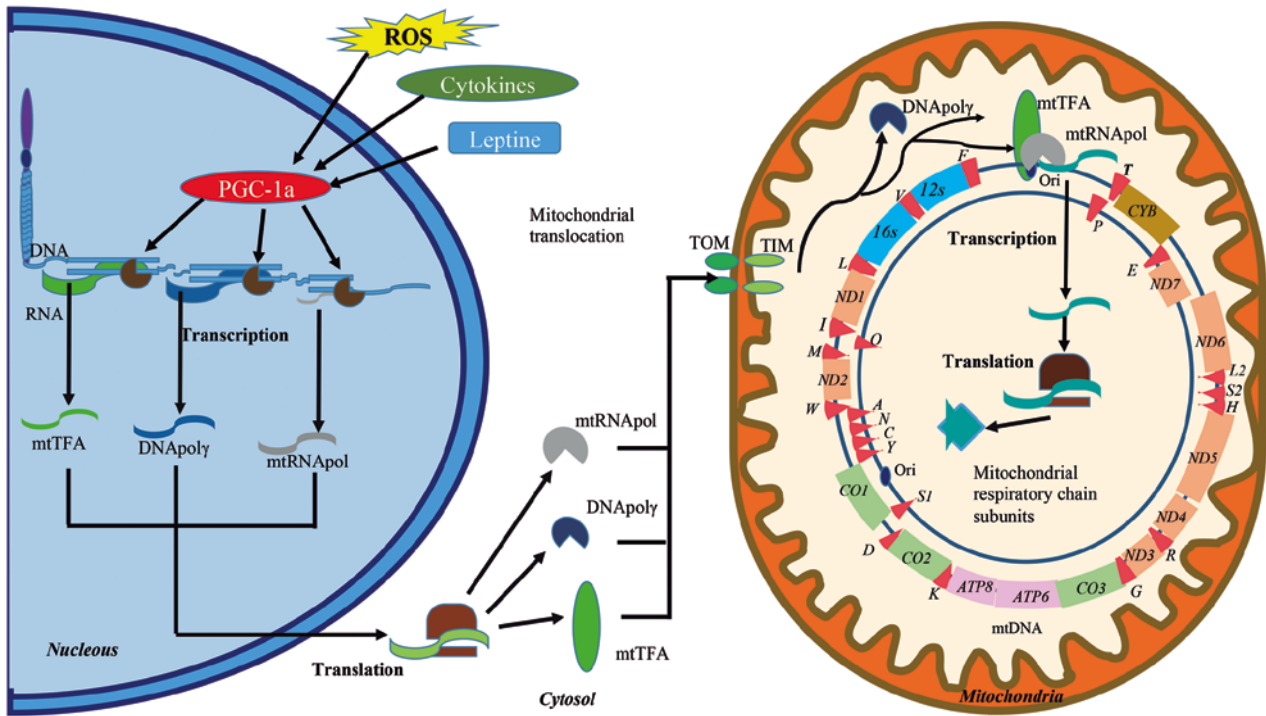


Figure 1. Signaling pathways involved in mitochondrial function and the human mitochondrial DNA (mtDNA) map. Activators, such as reactive oxygen species (ROS), cytokines and leptin activate peroxisome proliferator-activated receptor gamma coactivator-1 (PGC-1α) to promote the transcription of genes that are involved in the mitochondrial transcriptional machinery (DNAPoly, DNA polymerase γ; mtTFA, mitochondria transcription factor A; mtRNApol, mitochondrial RNA polymerase, etc.). The mRNAs are directed to the cytoplasm for translation and the resulting proteins are immediately imported into the mitochondria by the translocase of the outer and inner mitochondrial membrane (TOM and TIM, respectively) system. DNAPoly is responsible for replication, whereas mtTFA and mtRNApol are actively involved in the initiation of transcription. The formed mRNAs give rise to mitoribosomal subunits, tRNAs, or are translated into the mitoribosomes to generate protein subunits of the mitochondrial respiratory chain complex. mtDNA is a double-stranded (heavy chain: H and a light chain: L) circular DNA, containing 37 genes, of which, 28 are located in the H strand and nine in the L strand. Thirteen of these genes code for essential components of the electron transport complexes (ETC), 22 are transfer RNAs (tRNAs) and two ribosomal RNAs (rRNA). *ND1*, NADH-ubiquinone oxidoreductase chain 1; *ND2*, NADH dehydrogenase 2; *CO1*, cytochrome *c* oxidase 1; *COII*, cytochrome *c* oxidase 2; *ATP8*, ATP synthase 8; *ATP6*, ATP synthase 6; *COIII*, cytochrome *c* oxidase 3; *ND3*, NADH dehydrogenase 3; *ND4L*, NADH-ubiquinone oxidoreductase chain 4L; *ND4*, NADH dehydrogenase 4; *ND5*, NADH dehydrogenase 5; *ND6*, NADH dehydrogenase 6; *CYB*, cytochrome *b*. *ND* genes are indicated by orange boxes, *CO* genes by green boxes, the *CYB* gene by a gold box and *ATP* genes by pink boxes. rRNA is indicated by blue boxes; and tRNA by red triangles.

mention a few, tRNA modifying enzymes (*PUS1*, *TRMU*, *MTO1*), aminoacyl-tRNA synthetases (*RARS2*, *DARS2*, *YARS2*, *SARS2*, *HARS2*, *AARS2*, *MARS2*, *EARS2*, *FARS2*), ribosomal proteins (*MRPS16*, *MRPS22*, *MRPL3*, *MRPL12*), elongation factors (*GFMI*, *EFTs*, *EFTu*), translational activators (*LRPPRC*, *TACO1*, *C12orf62*, etc. (43).

4. Mitochondria variants and cancer

In pathological conditions, such as cancer, malignant transformation induces reprogramming of cell metabolism and permanent fission and fusion of mitochondria to maintain daughter cells upon cell division. During oncogenic processes, tumor cells not only acquire persistent changes in metabolism to support cell growth, but also other types of mtDNA damage (12). mtDNA mutations have been identified in all types of human tumors, both in the non-coding and coding regions of the mtDNA (Fig. 2). Larman *et al*, studying different cancer types, found that the frequency of somatic mtDNA mutations ranged from 13% in glioblastoma to 63% in rectal adenocarcinomas (44). Notably, the majority of the mtDNA mutations appeared to be homoplasmic (same sequence in all mtDNA molecules) in nature; nevertheless, heteroplasmy (the presence of >1 mtDNA sequence in a cell) could contribute to

heterogeneous mitochondrial morphologies and correlate with metabolic flexibility and cancer metastasis (28,45). According to their percentage in the cell, mutations may be homoplasmic (>95 to 100%), high heteroplasmic (>20 to <95%), low heteroplasmic (>0.5 to 20%) or rare (0.5% or less) (46). Surprisingly, 72% of the reported tumor-specific mtDNA mutations are also found in non-tumor cells (germline variants) of healthy subjects (44). As mtDNA is highly susceptible to acquiring mutations from ROS, in addition to inherited mtDNA mutations, multiple subpopulations of mtDNA could arise during the lifespan, enriching the heteroplasmic condition (mainly in cells with high division rates). Of note, although there is co-occurrence of mtDNA germline variants and somatic mutations, a high proportion of these become homoplasmic in cancer cells (47,48). However, the heteroplasmic threshold effect and the mechanisms underlying their selection towards a homoplasmic state during tumor development are not yet well understood (48).

Warburg's observation was the cornerstone for the study of mitochondrial function in tumor cells. Later, it was observed that the mRNA levels for defined mtDNA encoded genes were upregulated in many human cancer types (44,49-51). In this regard, the first significant finding was the observation that hepatoma cells express hexokinase II, while normal hepatic

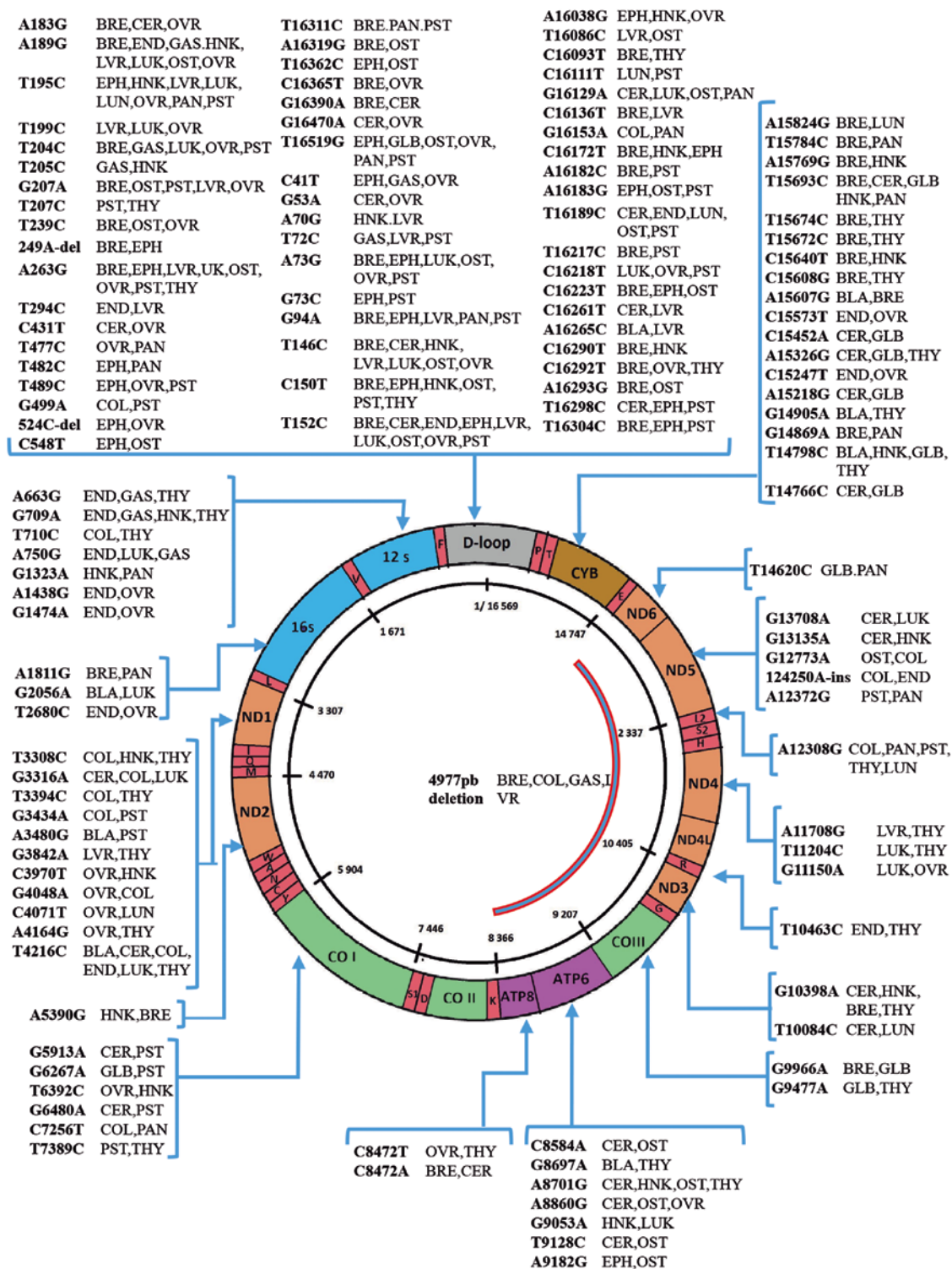


Figure 2. Mitochondrial DNA mutations reported in different types of cancer. The figure shows the mitochondrial (mtDNA) genome diagram and mutations reported in human cancer. mtDNA is a double-stranded circular DNA comprised by 16,569 bases and contain 37 genes. Of these, 13 encode essential components of the electron transport complexes (ETC), 22 are tRNAs and 2 rRNAs. Mutations are named by the locations of the mutated base (C: cytosine, G: guanine, T: thymine, A: adenine). The non-coding D-loop region (gray color) exhibits the higher number of mutations. *ND1*, NADH-ubiquinone oxidoreductase chain 1; *ND2*, NADH dehydrogenase 2; *COI*, cytochrome *c* oxidase I; *COII*, cytochrome *c* oxidase 2; *ATP8*, ATP synthase 8; *ATP6*, ATP synthase 6; *COIII*, cytochrome *c* oxidase 3; *ND3*, NADH dehydrogenase 3; *ND4L*, NADH-ubiquinone oxidoreductase chain 4L; *ND4*, NADH dehydrogenase 4; *ND5*, NADH dehydrogenase 5; *ND6*, NADH dehydrogenase 6; *CYB*, cytochrome *b*; BLA, bladder cancer; BRE, breast cancer; CER, cervical cancer; COL, colon cancer; END, endometrial cancer; EPH, esophageal cancer; GAS, gastric cancer; GLB, glioblastoma; HNK, head and neck cancer; LUK, leukemia; LVR, liver cancer; LUN, lung cancer; OTS, osteosarcoma; OVR, ovarian cancer; PNC, pancreatic cancer; PST, prostate cancer; THY, thyroid cancer.

cells express hexokinase IV (glucokinase) (49). Metabolic deregulation and its associated consequences induce a strong selective pressure on tumor cells. Hence, acquiring somatic

mtDNA mutations that impact oxidative phosphorylation seems to be an alternate mechanism for enhancing tumor growth (44). Another important observation is that mutations

in OXPHOS genes and accumulation of ROS can lead to the oxidation of proteins and their missfolding. To balance this, metastatic cells show high expression of the unfolded protein response (UPR^m), which allows them to survive in this highly stressed environment (45,52).

The mtDNA of cancer cells was first studied by Polyak *et al* in 1998 (47). They sequenced the mtDNA of 10 human colorectal cancer (CC) cell lines. The authors identified mutations in 70% of all studied cell lines; many of these occurred in rRNA genes or were missense mutations, with only one frame-shift and one chain-termination mutation (47). Another interesting finding was that tumor cells frequently contained a relatively small number of mitochondria with few mtDNA molecules per mitochondrion. In addition, most of the mtDNA mutations were homoplasmic, derived from a mtDNA somatic mutation initially represented in a heteroplasmic state (44). These data suggested that at least one of the observed tumoral mtDNA mutations could confer a selective advantage to the cancer cell, further supporting the hypothesis that mtDNA somatic mutations play an important role in promoting the proliferation of several types of tumor cells (47). Moreover, Habano *et al* (53) reported that changes in the length of the polycytosine tract of the HVR2 in the mitochondrial D-loop are the most frequent (44%) mutations in colorectal tumors. Currently, it is well known that this region is highly mutated in a variety of human cancers (ovarian, thyroid, kidney, liver, lung, colon, gastric, brain, bladder, head and neck, prostate, breast and leukemia), although most of these mutations do not have pathogenic implications, as they are not generating stop-codons, frame-shifts or altering tRNA structures (44). Furthermore, Larman *et al* (44), analyzing whole-genome data generated by The Cancer Genome Atlas Research Network (CGARN), reported that the majority of coding mutations occurred in the subunits of the electron transport chain complex (NADH dehydrogenase, cytochrome *b* and cytochrome *c* oxidase). The atlas includes paired tumor and normal tissue samples from 226 individuals with 5 types of cancer: Colon adenocarcinoma (COAD), rectal adenocarcinoma (READ), acute myeloid leukemia (AML), glioblastoma (GLB) and ovarian serous cystadenocarcinoma (OVR). Highly variable proportions of deleterious tumor-specific mutations across tumor types were found, ranging from 13% in GLB to 63% in READ. Another interesting finding was the differential distribution of nonsynonymous (missense and truncating) and synonymous mutations between inherited variants and somatic mutations in all types of cancer (COAD, READ, AML, GLB, OVR). The nonsynonymous mutations were in higher proportion in the tumor mtDNA (86%) in comparison with the germline mtDNA (31%). The majority of tumors (94%) that harbored somatic mtDNA mutations carried at least one nonsynonymous mutation and only 37% bearer somatic nonsynonymous mutations. Tumors that carried synonymous mutations, but no nonsynonymous mutations accounted for only 3%. OVR (93%) exhibited the highest proportion of nonsynonymous mtDNA somatic mutations compared to the rest of the tumor types (COAD, 89%; READ, 80%; AML, 78%; and GLB, 80%).

The first direct causal connection between a mtDNA mutation and cancer was reported by Petros *et al* (54) who identified the T8993G (L156R) germline mutation in the *ATP6* gene as heteroplasmic in a patient with prostate cancer and

demonstrated its tumorigenic capacity in a nude mouse prostate cancer model. The T8993G mutation in *ATP6* is known to induce enhanced ROS production, which may be the cause of enhanced tumorigenicity, since ROS are known to affect transcription factors, signal transduction kinases, regulatory phosphatases and angiogenic factors (44). Recently developed techniques, such as next generation sequencing (NGS) have improved our knowledge of the nature of mtDNA mutations thanks to the comparison of tumor and adjacent normal tissue or blood in a wide variety of cancer types. mtDNA variants commonly observed either as germline polymorphisms (or germline mutations) or somatic mutations (cancer cells) include point mutations, mono- or dinucleotide repeats, insertions, deletions and mtDNA copy number variations (50,51,54). The most common mtDNA polymorphisms which have been associated with an increased risk of cancer are T16189C (D-Loop), G10398A (A114T at *ND3* gene) and the deletion Δ mtDNA4977 (55-59). Δ mtDNA4977 is the most common deletion in the mitochondria, which occurs between nucleotides 8,470 and 13,477 of the human mtDNA. This deletion includes 5 tRNA genes, 4 genes encoding subunits of NADH dehydrogenases, *COIII* and ATPase genes, creating a smaller mtDNA molecule that leads to a decrease in ATP production and to abnormal ROS generation (Fig. 2) (60).

In addition to the structural abnormalities in mtDNA, fluctuations in mtDNA copy number are frequently described in cancer. For instance, an elevated mtDNA content has been found in primary head and neck squamous cell carcinoma, papillary thyroid carcinoma and endometrial cancer, in contrast to gastric cancer, which exhibits mtDNA depletion (61-63). The mtDNA copy number in cancer may depend upon the specific site of mutation associated with that cancer. Thus, an increase in mtDNA copy number can occur as a compensatory response to mitochondrial dysfunction or to mutations in nuclear genes indirectly involved in controlling mtDNA copy number. Conversely, mutations in the D-loop region, which control mtDNA replication, would be expected to result in a decrease in copy number. In many cases, fewer copies of mtDNA are accompanied by a decrease in the expression of mitochondrial genes, suggesting a suppressed mitochondrial activity in these tumor types. mtDNA impairment may result, not only in mtDNA copy number alterations, but also in the acquisition of new point mutations and deletions in mtDNA, damaged mitochondrial function and changes in cell and tissue viability. Defects in the components of the mtDNA repair machinery, which mediate mtDNA replication and repair, may increase the accumulation of mtDNA mutations, even though mitochondria have nDNA repair pathways, such as base excision repair, miss-match repair and recombinational repair (12,13). Indeed, Reznik *et al* (64) found that the mtDNA copy number is related to the incidence of key driver mutations that cause cells to become cancerous (13,14). Of note, although the mtDNA copy number can influence the level of transcription of mtDNA genes, not all cancer types exhibit a correlation between the expression of respiratory genes and mtDNA copy number (64). Exogenous and environmental factors, such as industrial by-products, ultraviolet and ionizing radiation, tobacco smoke, chemicals, environmental toxins and therapeutic drugs may also affect mtDNA expression (13,14).

Haplogroups and cancer risk. Ancient mtDNA SNPs are grouped into haplogroups, which are groups of haplotypes that share several polymorphisms acquired during the expansion and migration of the human population. Haplogroups reflect specific ancestral populations and are the result of sequential accumulation of mitochondrial mutations through maternal lineages; hence, they present continent-specific distributions. The molecular mechanisms underlying the association between mtDNA haplogroups and cancer development remains unclear and controversial results between populations have been reported (65-73). However, specific mtDNA haplogroups have been associated with the risk of prostate cancer, BC, CC, nasopharyngeal cancer, gastric cancer, myelodysplastic syndromes (MS), acute lymphoblastic leukemia, etc. (65-72,74). For example, an over-representation of haplogroup I in Polish cancer patients has been reported (75), haplogroup U has been shown to be associated with an increased risk of developing renal and prostate cancers (65), and haplogroup JT has been shown to be associated with MS susceptibility in North American caucasian subjects (70). Additionally, haplogroups H, N and L have been shown to be associated with the risk of developing pancreatic cancer, whereas haplogroup K has been identified as protective for this type of cancer in participants from the San Francisco Bay Area in California (76). In China, haplogroups D4a and D5a have been found to be associated with an increased the risk of esophageal cancer, while the N and M haplogroups have been reported as biomarkers for the good and poor survival of patients with gastric cancer, respectively (67,69). Another study, including 1,503 autopsy cases from Japan (76), reported an association between the haplogroup M7b2, previously associated with leukemia (77) and an increased risk of developing hematologic cancers (76). Nevertheless, some of the most robust association data have been reported for haplogroup T (M184) and the risk of CC in European-American descendants, following the analysis of 2,453 CC cases and 11,930 multi-ethnic controls (68).

Nuclear-encoded genes, mtDNA biogenesis and function in cancer. While a number of studies on cancer have focused on nDNA mutations, less research has been directed at determining the influence of germline variants and somatic mutations in NEMtGs and mitochondrial genes on organelle dysfunction and cancer development. It is well known that mutations in nuclear-encoded components of the mitochondrial translation machinery accumulate non-repaired nDNA and mtDNA that may promote cellular transformation and tumor development. Mutations in nuclear genes responsible for mtDNA biosynthesis and integrity maintenance, such as *POLG*, *mtTFA*, *PGC-1*, *NRF-1*, *NRF-2*, isocitrate dehydrogenase (*IDH*), *SDH* family members (*SDHA*, *SDHB*, *SDHC* and *SDHD*), fumarate hydratase (*FH*), *SIRT1-3*, *POLG*, *TWINKLE* and *OPA1* are frequently implicated in carcinogenic processes (74,78,79). In fact, heterozygous germline mutations in *SDHB*, *SDHC*, *SDHD* and *SDHAF2* cause hereditary paragangliomas and pheochromocytomas (80). *POLG* is known as the only DNA polymerase that functions in human mitochondria. Due to its ability to bypass lesions by translation synthesis, *POLG* is vital for maintaining the genetic integrity of mtDNA. Mutations in this gene can promote mtDNA mutations, the accumulation of unrepaired lesions or can even block DNA replication in

cancer (4,6-13). Furthermore, *IDH1* and *IDH2*, mitochondrial and TCA cycle-related proteins, are frequently mutated in tumors, such as glioblastoma. In fact, studies on GLB and AML have identified recurrent missense mutations that impact specific residues of the nuclear genes *IDH1* and *IDH2* (78,79). Notably, it has been observed that tumors carrying *IDH* mutations have few or no somatic mtDNA mutations (79).

The abnormal expression of several NEMtGs is a common finding in all steps of cancer development and progression. For example, TNF receptor-associated protein 1 (TRAP1) is a protein that plays a role in the regulation of proteostasis, mitochondrial apoptosis and suppression of ROS production in the mitochondria of tumor cells, through regulation of folding and stability of selective proteins. In fact, it has been reported that TRAP1 overexpression favors resistance to standard chemotherapeutic agents (81). The overexpression of genes involved in the activation of the UPR^{mt} has also been observed in metastatic cells. The UPR^{mt} system protects from protein misfolding and oxidation; however, it might also maintain deleterious mtDNA mutations conferring advantages that facilitate cancer cell metastasis (52,82).

5. Mitochondrial DNA variants in breast cancer

The recognition of the importance of mtDNA mutations as a relevant feature in BC development, metastasis and resistance to treatment, has occurred over the past decade (48,83,84). Both mtDNA germline variants and tumor mutations have been involved in BC (Table II). As with other types of cancer, the D-loop region (16024-576 nt), mainly the HVR1 and HVR2 regions, is an important mitochondrial mutation hotspot (48,83,85). This region has been reported to be ~60-fold more susceptible to mutations than other regions of the mitochondrial genome (55,86-88). The T16189C, C16207T, T16519C, T239C and A263G polymorphisms, located in the D-loop region, are frequently associated with the risk of BC (85-87). By contrast, the T16183C, C16223T, T16362C, A73G and C150T polymorphisms, have been found to be enriched in controls compared to BC patients, suggesting that they could protect for this disease (86). Of these, T16189C is the most common polymorphism that is associated with the risk of BC (75,89,90). The T to C change in the wild-type (C5TC4) generates a long run of C residues considered to be a source of enhanced instability (55,85,87,88,91). This abnormality often leads to heteroplasmic length variation of the poly-C tract (>10 cytosines) and a correlation has been observed between the poly-C variants and the mtDNA copy numbers (59), as subjects who have an interrupted poly-C carry higher mtDNA copy number, in contrast to those with an uninterrupted poly-C, suggesting that the T16189C variant may affect mtDNA replication (92).

Coding mitochondrial variants have also been implicated in the risk of BC. On one hand, the A4727G, G9055A, C6296A, G9947A, A8860G, A10044G, A10283G, T11233C and C11503T variants increase the risk, and on the other hand, the T3197C and G13708A SNPs decrease the risk of developing this neoplasia (85,86,88,93-95). In terms of deletions, the 2463delA, 6298delT and the Δ mtDNA4977 polymorphisms have been identified as susceptibility factors for BC; however, Δ mtDNA4977 is the most commonly implicated in the breast

carcinogenesis process (88,96). The mechanisms that promote tolerance to these deletions or contribute to their propagation in the tumor cells are unknown (82). Additionally, differential contribution and interaction between germline variants have been reported to increase BC susceptibility. For example, G10398A has been shown to be associated with an increased risk of BC in European-American, Polish and Malay populations, but not in Hindu and African-American women (57,75,84,97-100). Furthermore, T4216C confers a risk of BC only in combination with the G10398A germline variant (97).

It is relevant to state that in addition to the G6267A somatic mutation, the germline variants T14110C, G14207A and D310 (from nucleotide positions 303 to 315 and interrupted by a T in position 310) are very common in BC cases (96,101,102). The poly-cytidine stretch D310 is the most common, reaching 92% of the tumor samples; thus, it has been suggested as a potential starting point for the clonal expansion of malignant cells and may be a clinical marker for breast tumorigenesis (85,88,101,103-108). Alhomidi *et al*, in 2013, supported this hypothesis, reporting that 32.5% of patients from India with BC were D310 carriers (108).

Apart from somatic mtDNA mutations, studies have identified mtDNA mutation (varying degrees, different sites, and different types) in nearly 60% of BCs tumors (87,88,95,101); however, the question of whether the mutations help drive the tumor or are bystander events remains unanswered (76,83). McMahon and LaFramboise (83), in a study that included 99 normal-tumor paired BC samples, identified 142 somatic mutations; however, many of these have been reported as germline variants. G567A, G766A, G1333A, A2085G, G2621A, G5043 and G11346A were identified in only one patient, as examples of somatic mutations (described exclusively in breast tumor tissue) (83,109). Other matched blood-tumor mtDNA sample studies have also shown that most somatic mutations are singletons arising in a single patient and interestingly, are present in a heteroplasmic state in the tumor, but not in mtDNA from normal tissues. It has been suggested that many putative somatic mutations are in undetected germline heteroplasmies that have undergone clonal expansion in the tumor (46,83,110).

With regard to the impact of mtDNA content in BC, a lower mtDNA content has been observed in BC tumors when compared to their surrounding normal epithelium; however, these data are inconsistent (111).

Haplogroups and breast cancer. The important role of haplogroups in different human conditions such as aging, neurodegenerative diseases, metabolic diseases, as well as BC (Table III) have also been explored. Currently, it is considered that metabolism and sensitivity to oxidative stress differs among mitochondrial haplogroups (112).

Studies performed on the Polish population have demonstrated the overrepresentation and under-representation of haplogroups I and H, respectively, in patients with BC, in comparison with healthy groups. These data suggest that the presence of different polymorphisms given in the haplogroup may be fundamental for the carcinogenic process (75). Furthermore, an association between haplogroups and *BRCA1* and *BRCA2* mutations has been reported. Tommasi *et al* (113) found that haplogroups X and H were significantly more frequent in subjects bearing mutations in the *BRCA1* and

BRCA2 genes, respectively. Additionally, it has recently been observed that haplogroup T1a1 may modify the individual risk of acquiring the disease in *BRCA2* mutation carriers (114). Likewise, Rao *et al* (115), studying 92 triple-negative BC (TNBC) patients of African-American (n=31), caucasian (n=31) and Hispanic (n=30) origin, found that Nigerian (29%), Cameroon (26%), Sierra Leone (16%) and Guinea-Bissau (13%) haplogroups were highly prevalent in African-American patients; while the H (29%), U (23%), K (13%), V (7%), and J (7%) haplogroups accounted for 79% of all haplogroups identified in caucasian-derived populations. As has been previously reported, haplogroups A (34%), B (17%) and C (13%) are the most frequent in Hispanic patients. Notably, the authors found that non-caucasian haplogroups represent almost 64% of TNBC patients, which is in accordance with data showing that TNBC is more common in African- and Hispanic-derived populations (115). Although some of these studies have provided controversial findings and statistically underpowered results, the European haplogroup N, defined by the G10398A germline variant, has been reported as a risk factor for BC, particularly in women of African-American and Indian origin (96,116-118). Additionally, Bai *et al* (85), studying a caucasian population, found that haplogroup K (G10398A and T16189C) increased the risk of BC, whereas haplogroup U conferred protection against this disease. Moreover, Darvishi *et al* (99) re-analyzed approximately 1,000 mitochondrial DNA sequences and found an association between the distribution of haplogroup N and the incidence of sporadic BC. This was corroborated through a case control study, where they also observed association of the A10398C germline variant with sporadic BC.

Mutations in nuclear-encoded genes that participate in mtDNA biogenesis and BC. It is well known that mutations in nuclear encoded genes, such as *BRCA1*, *BRCA2*, *TP53*, *PTEN*, *CHEK2* and *CHEK1*, are commonly associated with a high risk of developing BC. Moreover, as the majority of mitochondrial proteins are nuclear encoded and post-translationally imported in the mitochondria, the nuclear/mitochondrial connection in BC is also altered. However, the most notable association between the mitochondrial function and nuclear encoded genes relates to the OXPHOS system since 79 of its components are encoded by nDNA. In fact, none of the subunits of complex II are encoded by mtDNA (74,78,119). Hossein and Houshmand (120) used algorithms to determine that the mitochondrial deletion Δ mtDNA4977 in association with alterations in nuclear genes (such as *BRCA*, *ER* and *TP53* genes) led to a phenotype of premature aging and BC (95). Moreover, Jandova *et al* (121) reported that the copy number of mtDNA can produce differential expression levels of matrix metalloproteinase 9 (*MMP9*) and collagen, type I, $\alpha 1$ (*Colla*), which are capable of triggering the malignancy. Notably, *MMP9* has been shown to be involved in several types of cancer and metastasis. Additionally, expression analyses have also revealed an association between the expression of NEMtGs and BC subtypes. Furthermore, mitochondrial membrane respiratory chain complex I and IV and ATP synthesis were reduced, in both epithelial and stromal cancer cells compared to normal breast cells. Additionally, *MRPL12*, *POLG* and *RNASEH1* were upregulated in cancerous stromal cells (122).

Table II. Common germline variants associated with breast cancer.

Population	Mutation	Gene	OR (CI 95%)	P-value	Refs.
African-American	G10398A	<i>ND3</i>	1.6 (1.10-2.31)	0.013	(57)
Canadian	310C-ins	D-loop	-	0.0001	(106)
Chinese	2463A-del	<i>RNR2</i>	8.05 (1.023-63.34)	0.022	(66,87,96)
	C6296T	<i>COI</i>	4.464 (0.55-35.90)	0.038	
	6298T-del	<i>COI</i>	4.464 (0.55-35.90)	0.038	
	8460-13477del	<i>ATP8, ATP6, CO3, TG, ND3, TR, ND4L, ND4, TH, TS2, TL2, ND5</i>	-	-	
	A8860	<i>ATP6</i>	5.254 (0.65-41.92)	0.021	
	A10397G	<i>ND3</i>	3.11 (1.07-9.06)	0.03	
	G10398A	<i>ND3</i>	1.77 (1.00-3.14)	0.05	
	13237A-del	<i>ND5</i>	4.85 (0.60-38.86)	0.037	
European-American	T16189C	D-loop	2.36 (1.14-4.88)	0.019	
	T3197C	<i>RNR2</i>	0.31 (0.13-0.75)	0.004	(85,98)
	G9055A	<i>ATP6</i>	3.03 (1.63-5.63)	0.0004	
	A10397G	<i>ND3</i>	-	0.03	
	G10398A	<i>ND3</i>	1.79 (1.14-2.81)	0.01	
	G13708A	<i>ND5</i>	0.47 (0.24-0.92)	0.002	
	T16519C	D-loop	1.98 (1.25-3.12)	0.003	
	T16519C	D-loop	1.98 (1.25-3.12)	0.003	
Indian	G10398A	<i>ND3</i>	1.73 (1.13-2.66)	0.01	(99)
Italian	A153G	D-loop	19 (1.8-201.9)	0.009	(107,113)
	T195C	D-loop	6 (1.12-31.99)	0.04	
	G225A	D-loop	12.7 (1.18-136.28)	0.03	
	T226C	D-loop	12.7 (1.18-136.28)	0.03	
	G3918A	<i>ND1</i>	-	-	
	G11719A	<i>ND4</i>	13.2 (2.13-82.13)	0.005	
	G11900A	<i>ND4</i>	-	-	
	T12344A	<i>ND5</i>	-	-	
	G13708A	<i>ND5</i>	-	-	
	G14869A	<i>CYB</i>	-	-	
	C16093T	D-loop	-	-	
	T16183C	D-loop	12.7 (1.18-136.28)	0.03	
	C16278T	D-loop	7.3 (1.14-4.88)	0.03	
	T16519C	D-loop	21 (2.15-204.6)	0.003	
Malay	G10398A	<i>ND3</i>	2.29 (1.25-4.20)	0.007	(100)
Polish	A73G	D-loop	-	0.001	(75,94)
	C150T	D-loop	-	0.001	
	T152C	D-loop	-	0.059	
	T239C	D-loop	-	0.001	
	A263G	D-loop	-	0.001	
	A4727G	<i>ND2</i>	-	-	
	G9947A	<i>COIII</i>	-	-	
	A10044G	<i>TG</i>	-	-	
	A10283G	<i>ND3</i>	-	-	
	T11233C	<i>ND4</i>	-	-	
	C11503T	<i>ND4</i>	-	-	
	T16183C	D-loop	-	0.036	
	T16189C	D-loop	-	0.004	
	C16192T	D-loop	-	0.023	
	T16207C	D-loop	-	0.023	
	C16223T	D-loop	-	0.001	
	T16362C	D-loop	-	0.001	
	T16519C	D-loop	-	0.003	
Spanish	G3010A	<i>RNR2</i>	0.73 (0.44-1.00)	0.047	(116)
	T3197C	<i>RNR2</i>	2.72 (1.14-7.18)	0.015	
	A12308G	<i>TL2</i>	0.66 (0.46-0.94)	0.019	
	T16519C	D-loop	0.74 (0.55-0.99)	0.039	
Taiwanese	mtDNA depletion	All	-	0.0008	(125)

OR, odds ratio; CI, confidence interval; *RNR2*, rRNA subunit 16s; *ND1*, NADH dehydrogenase subunit 1; *ND2*, NADH dehydrogenase subunit 2; *ND3*, NADH dehydrogenase subunit 3; *ND4*, NADH dehydrogenase subunit 4; *ND4L*, NADH dehydrogenase subunit 4L; *ND5*, NADH dehydrogenase subunit 2; *COI*, cytochrome oxidase subunit 1; *COII*, cytochrome oxidase subunit 2; *CYB*, cytochrome b; *ATP6*, ATP synthase subunit 6; *ATP8*, ATP synthase subunit 8; *TG*, tRNA glycine; *TR*, tRNA arginine; *TH*, tRNA histidine; *TS2*, tRNA serine 2; *TL2*, tRNA leucine.

Table III. Frequency of haplogroups associated with breast cancer.

Haplo-group	SNPs	Population	N Controls/cases	Frequency (%)	Effect	OR (CI 95%)	P-value	Refs.
D5	C150T, T1107, A5301G, A10397G	Chinese	104/114	12.9	Increased risk	3.11 (1.07-9.06)	0.030	(66)
K	A10550G, T11299C, T14798C, T16224C, T16311C	European-American	156/260	18.6	Increased risk of hereditary cancer	3.03 (1.63-5.63)	0.0004	(85)
H	G2706A, T7028C	Polish Italian	44/100 10/10	36 40	Protection Increased risk of <i>BRCA2</i> mutation carriers	0.636 (0.51-0.74) -	0.019 0.05	(75) (113)
I	T10034C, G16129A	Polish	44/100	14	Increased risk	-	0.017	(75)
M	T489C, C10400T, T14783C, G15043A	Chinese	104/114	66.3	Increased risk. Protection against metastasis	1.77 (1.03-3.07)	0.040	(62)
N	G8701A, C9540T, G10398A, C10873T, A15301G	Chinese	34/70	47	Risk of metastasis	0.39 (0.17-0.94)	0.036	(66)
U	A11467G, A12308G, G12372A	European-American	156/260	7.7	Decreased risk of hereditary cancer	0.37 (0.19-0.73)	0.0023	(85)
X	T6221C, C6371T, A13966, T14470C, T16189C, C16278T	Italian	10/10	60	Increased risk of <i>BRCA1</i> mutation carriers	-	0.04	(113)
T1a1	T9899C, A16163G, C16186T, T152C, T195C	Multi-national	3989/3698	9.3	Decreased risk of <i>BRCA2</i> mutations carriers	0.62 (0.40-0.95)	0.03	(114)

SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

mtDNA variants and BC subtypes and survival. Few studies have been conducted to clarify whether there is a correlation between mtDNA germline variants or somatic mutations and BC subtypes. Thus far, there has been no direct association of mutations with the intrinsic subtypes of this neoplasia; however, Lin *et al* (123) observed that the D310 germline variant was associated with an advanced stage of the tumor and with negative Her2/neu expression. Tommasi *et al* (113) found that the T16126C germline variant was more frequent, but not exclusive, in estrogen- and progesterone receptor-negative tumors. Nevertheless, the authors detected that C150T was exclusive to patients with hereditary BC, as was T16519C, which was found in 90% of individuals carrying mutations in *BRCA1*. Blein *et al* (124) sequenced blood mtDNA from 436 women, with a positive familial BC history, but negative for *BRCA1/2* pathogenic mutations, diagnosed with BC, and found an enrichment of mutations in *MT-ATP6* (86.5 variants/Mb) and *MT-CYB* (84.1 variants/Mb), but no association between the rate or type of mutations. Moreover, McMahan and LaFramboise (83) observed that alterations in heteroplasmy levels were more significant in *HER2*-positive tumors than in *HER2*-negative tumors.

The impact of the mtDNA mutation rate and mtDNA content and their correlation with prognosis and survival have also been investigated. A higher mutational burden and a low copy number of mtDNA have been associated with a worse overall survival, in contrast to patients carrying a low mutation

rate or high mtDNA content; however, these results remain inconclusive (111,125).

6. Mitochondria DNA mutations and response to therapy in breast cancer

Widespread attention has been paid to mitochondria mutations. This is a not a recent phenomenon, and it fits into the broader historical interest in studying the size, evolution and function of genomes. The first complete human mtDNA sequenced was in 1981; currently, >3,000 complete mitochondrial genomes have been sequenced (109). The unique structure and function of mitochondria make them an excellent model system to identify and answer biological and medical issues where other approaches exhibit technical difficulties. As the mitochondria control both energy metabolism and susceptibility towards apoptotic cell death, mtDNA mutations have been suggested to play important roles in breast carcinogenesis and drug-resistance. Thus, the landscape of driving tumor mtDNA mutations could be used for the design of clinical trials and development of new anti-neoplastic therapies. In this regard, the role of mitochondrial dysfunction and cancer therapy has been widely investigated. Recently, Farnie *et al* (126) demonstrated that tumors with high mitochondrial mass (mito-high) were specifically enriched in the number of cancer stem-like cells (CSCs), which are involved in tumor recurrence and metastasis. Furthermore, these mito-high breast CSCs exhibited resistance

to DNA damage following antineoplastic treatment. Based on the 'Endo-symbiotic Theory of Mitochondrial Evolution', it is suggested that antibiotics against mitochondria could be used to modify mitochondrial mass within CSCs. In fact, the FDA has already approved drugs/antibiotics that target mitochondria to eradicate CSC activity (metformin and pyruvate, salinomycin) as well as inhibitors of mitochondrial biogenesis and mitochondrial translation (erythromycins, tetracyclines and glycolcyclines) (126,127). Notably, doxycycline and azithromycin have already exhibited significant efficacy in treatment-resistant cancer patients with lymphoma and non-small cell lung tumors, respectively (128,129).

The steep rise in interest in the study of mitochondrial mutations is also due to the fact that mtDNA mutations play a role in many other human diseases, opening a paradigm shift that suggests that mitochondrial mutations will contribute to our understanding of tumor biology in the mammary gland. In fact, Teixeira *et al* (130) found that the mitochondria were also involved in stem cell differentiation, showing that blocking any of the 13 key proteins linked to ATP synthase disrupted or stalled egg development in fruit flies. Certainly, the recognition of the widespread applicability of studying the role of mitochondria and the biological implications of mtDNA may will strongly improve the development of new means with which to treat human common diseases, including BC.

7. Conclusions

As mtDNA mutations and mtDNA copy number have been found to play important roles in breast carcinogenesis, the landscape of driving tumor mtDNA mutations may be used for the design of clinical trials and the development of novel anti-neoplastic therapies. Moreover, the suggestion that mtDNA copy number is associated with the severity of the disease could identify cases with better or worse survival. It is clear that NGS technologies have resulted in an unprecedented ability to understand mtDNA structure; however, the challenge is to determine whether and which mutations or germline variants carry biological significance. In addition, meriting further investigation are other aspects of mitochondrial genomes which undoubtedly are involved in human diseases, these aspects include chromosome structure, transcriptional and translational architecture, and modes of repair and replication.

Acknowledgements

Not applicable.

Funding

This study was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) through the program Fondo Sectorial de Investigación en Salud y Seguridad Social (S0008-2016-1-272618) and by the Internal Research Program of the National Institute of Genomic Medicine, Mexico (C2_19/2009).

Availability of data and materials

Not applicable.

Authors' contributions

SJM and CJP reviewed the literature and wrote the first draft of the manuscript. SJM, EL and AHM were involved in the conception of the review and critically reviewed the final version. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. World Cancer Research Fund International: BC Statistics. <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics> Accessed January 19, 2018.
2. Breastcancer.org: U.S. BC statistics. http://www.breastcancer.org/symptoms/understand_bc/statistics. Accessed January 19, 2018.
3. Hanahan D and Weinberg RA: Hallmarks of cancer: The next generation. *Cell* 144: 646-674, 2011.
4. Warburg O: On the origin of cancer cells. *Science* 123: 309-314, 1956.
5. Havas KM, Milchevskaya V, Radic K, Alladin A, Kafkia E, Garcia M, Stolte J, Klaus B, Rotmensz N, Gibson TJ, *et al*: Metabolic shifts in residual breast cancer drive tumor recurrence. *J Clin Invest* 127: 2091-2105, 2017.
6. Wu W and Zhao S: Metabolic changes in cancer: Beyond the Warburg effect. *Acta Biochim Biophys Sin (Shanghai)* 45: 18-26, 2013.
7. Bardella C, Pollard PJ and Tomlinson I: SDH mutations in cancer. *Biochim Biophys Acta* 1807: 1432-1443, 2011.
8. Bogenhagen DF: Mitochondrial DNA nucleoid structure. *Biochim Biophys Acta* 1819: 914-920, 2012.
9. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM and Howell N: Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23: 147, 1999.
10. Wallace DC and Chalkia D: Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harb Perspect Biol* 5: a021220, 2013.
11. Stewart JB and Chinnery PF: The dynamics of mitochondrial DNA heteroplasmy: Implications for human health and disease. *Nat Rev Genet* 16: 530-542, 2015.
12. Akhmedov AT and Marín-García J: Mitochondrial DNA maintenance: An appraisal. *Mol Cell Biochem* 409: 283-305, 2015.
13. Yakes FM and Van Houten B: Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci USA* 94: 514-519, 1997.
14. Egea PF, Stroud RM and Walter P: Targeting proteins to membranes: Structure of the signal recognition particle. *Curr Opin Struct Biol* 15: 213-220, 2005.
15. Baker MJ, Frazier AE, Gulbis JM and Ryan MT: Mitochondrial protein-import machinery: Correlating structure with function. *Trends Cell Biol* 17: 456-464, 2007.
16. Neupert W and Herrmann JM: Translocation of proteins into mitochondria. *Annu Rev Biochem* 76: 723-749, 2007.
17. Lill R and Mühlenhoff U: Maturation of iron-sulfur proteins in eukaryotes: Mechanisms, connected processes, and diseases. *Annu Rev Biochem* 77: 669-700, 2008.
18. Virbasius JV and Scarpulla RC: Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: A potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. *Proc Natl Acad Sci USA* 91: 1309-1313, 1994.

19. Hendrickson SL, Lautenberger JA, Chinn LW, Malasky M, Sezgin E, Kingsley LA, Goedert JJ, Kirk GD, Gomperts ED, Buchbinder SP, *et al*: Genetic variants in nuclear-encoded mitochondrial genes influence AIDS progression. *PLoS One* 5: e12862, 2010.
20. Scarpulla RC: Nucleus-encoded regulators of mitochondrial function: Integration of respiratory chain expression, nutrient sensing and metabolic stress. *Biochim Biophys Acta* 1819: 1088-1097, 2012.
21. Kasashima K and Endo H: Interaction of human mitochondrial transcription factor A in mitochondria: Its involvement in the dynamics of mitochondrial DNA nucleoids. *Genes Cells* 20: 1017-1027, 2015.
22. Lee HC and Wei YH: Mitochondrial biogenesis and mitochondrial DNA maintenance of mammalian cells under oxidative stress. *Int J Biochem Cell Biol* 37: 822-834, 2005.
23. Hance N, Ekstrand MI and Trifunovic A: Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum Mol Genet* 14: 1775-1783, 2005.
24. Mokranjac D and Neupert W: Cell biology: Architecture of a protein entry gate. *Nature* 528: 201-202, 2015.
25. Fan AC, Bhangoo MK and Young JC: Hsp90 functions in the targeting and outer membrane translocation steps of Tom70-mediated mitochondrial import. *J Biol Chem* 281: 33313-33324, 2006.
26. MacKenzie JA and Payne RM: Mitochondrial protein import and human health and disease. *Bi-ochim Biophys Acta* 1772: 509-523, 2007.
27. Bandiera S, Rübberg S, Girard M, Cagnard N, Hanein S, Chrétien D, Munnich A, Lyonnet S and Henrion-Caude A: Nuclear outsourcing of RNA interference components to human mitochondria. *PLoS One* 6: e20746, 2011.
28. Stefano GB and Kream RM: Mitochondrial DNA heteroplasmy in human health and disease. *Biomed Rep* 4: 259-262, 2016.
29. Hudson G, Gomez-Duran A, Wilson IJ and Chinnery PF: Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet* 10: e1004369, 2014.
30. Hudson G, Carelli V, Spruijt L, Gerards M, Mowbray C, Achilli A, Pyle A, Elson J, Howell N, La Morgia C, *et al*: Clinical expression of Leber hereditary optic neuropathy is affected by the mitochondrial DNA-haplogroup background. *Am J Hum Genet* 81: 228-233, 2007.
31. Soini HK, Moilanen JS, Vilmi-Kerälä T, Finnilä S and Majamaa K: Mitochondrial DNA variant m.15218A>G in Finnish epilepsy patients who have maternal relatives with epilepsy, sensorineural hearing impairment or diabetes mellitus. *BMC Med Genet* 14: 73, 2013.
32. Fragaki K, Procaccio V, Bannwarth S, Serre V, O'Hearn S, Potluri P, Augé G, Casagrande F, Caruba C, Lambert JC, *et al*: A neonatal polyvisceral failure linked to a de novo homoplasmic mutation in the mitochondrially encoded cytochrome *b* gene. *Mitochondrion* 9: 346-352, 2009.
33. Ghezzi D, Marelli C, Achilli A, Goldwurm S, Pezzoli G, Barone P, Pellecchia MT, Stanzione P, Brusa L, Bentivoglio AR, *et al*: Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur J Hum Genet* 13: 748-752, 2005.
34. Huerta C, Castro MG, Coto E, Blázquez M, Ribacoba R, Guisasola LM, Salvador C, Martínez C, Lahoz CH and Alvarez V: Mitochondrial DNA polymorphisms and risk of Parkinson's disease in Spanish population. *J Neurol Sci* 236: 49-54, 2005.
35. Galliche L, Serre V, Beinat M, Assouline Z, Lebre AS, Chretien D, Nietschke P, Benes V, Boddaert N, Sidi D, *et al*: Exome sequencing identifies MRPL3 mutation in mitochondrial cardio-myopathy. *Hum Mutat* 32: 1225-1231, 2011.
36. Götz A, Tyynismaa H, Euro L, Ellonen P, Hyötyläinen T, Ojala T, Hämäläinen RH, Tommiska J, Raivio T, Oresic M, *et al*: Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *Am J Hum Genet* 88: 635-642, 2011.
37. Bayat V, Thiffault I, Jaiswal M, Tétreault M, Donti T, Sasarman F, Bernard G, Demers-Lamarque J, Dicaire MJ, Mathieu J, *et al*: Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. *PLoS Biol* 10: e1001288, 2012.
38. Elo JM, Yadavalli SS, Euro L, Isohanni P, Götz A, Carroll CJ, Valanne L, Alkuraya FS, Uusimaa J, Paetau A, *et al*: Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. *Hum Mol Genet* 21: 4521-4529, 2012.
39. Haack TB, Kopajtich R, Freisinger P, Wieland T, Rorbach J, Nicholls TJ, Baruffini E, Walther A, Danhauser K, Zimmermann FA, *et al*: ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. *Am J Hum Genet* 93: 211-223, 2013.
40. Soiferman D, Ayalon O, Weissman S and Saada A: The effect of small molecules on nuclear-encoded translation diseases. *Biochimie* 100: 184-191, 2014.
41. Powell CA, Nicholls TJ and Minczuk M: Nuclear-encoded factors involved in post-transcriptional processing and modification of mitochondrial tRNAs in human disease. *Front Genet* 6: 79, 2015.
42. Takakubo F, Cartwright P, Hoogenraad N, Thorburn DR, Collins F, Lithgow T and Dahl HH: An amino acid substitution in the pyruvate dehydrogenase E1 alpha gene, affecting mitochondrial import of the precursor protein. *Am J Hum Genet* 57: 772-780, 1995.
43. Weraarpachai W, Sasarman F, Nishimura T, Antonicka H, Auré K, Rötig A, Lombès A and Shoubridge EA: Mutations in C12orf62, a factor that couples COX I synthesis with cytochrome *c* oxidase assembly, cause fatal neonatal lactic acidosis. *Am J Hum Genet* 90: 142-151, 2012.
44. Larman TC, DePalma SR, Hadjipanayis AG, Protopopov A, Zhang J, Gabriel SB, Chin L, Seidman CE, Kucherlapati R and Seidman JG: Cancer Genome Atlas Research Network: Spectrum of somatic mitochondrial mutations in five cancers. *Proc Natl Acad Sci USA* 109: 14087-14091, 2012.
45. Kenny TC, Hart P, Ragazzi M, Sersinghe M, Chipuk J, Sagar MA, Eliceiri KW, LaFramboise T, Grandhi S, Santos J, *et al*: Selected mitochondrial DNA landscapes activate the SIRT3 axis of the UPRmt to promote metastasis. *Oncogene* 36: 4393-4404, 2017.
46. Ahn EH, Hirohata K, Kohn BF, Fox EJ, Chang CC and Loeb LA: Detection of ultrarare mitochondrial mutations in breast stem cells by duplex sequencing. *PLoS One* 10: e0136216, 2015.
47. Polyak K, Li Y, Zhu H, Lengauer C, Willson JK, Markowitz SD, Trush MA, Kinzler KW and Vogelstein B: Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat Genet* 20: 291-293, 1998.
48. Yadav N and Chandra D: Mitochondrial DNA mutations and breast tumorigenesis. *Biochim Biophys Acta* 1836: 336-344, 2013.
49. Pedersen PL, Mathupala S, Rempel A, Geschwind JF and Ko YH: Mitochondrial bound type II hexokinase: A key player in the growth and survival of many cancers and an ideal prospect for therapeutic intervention. *Biochim Biophys Acta* 1555: 14-20, 2002.
50. Brandon M, Baldi P and Wallace DC: Mitochondrial mutations in cancer. *Oncogene* 25: 4647-4662, 2006.
51. Kirches E: Mitochondrial and nuclear genes of mitochondrial components in cancer. *Curr Genomics* 10: 281-293, 2009.
52. Kenny TC and Germain D: mtDNA, metastasis, and the mitochondrial unfolded protein response (UPRmt). *Front Cell Dev Biol* 5: 37, 2017.
53. Habano W, Sugai T, Yoshida T and Nakamura S: Mitochondrial gene mutation, but not large-scale deletion, is a feature of colorectal carcinomas with mitochondrial microsatellite instability. *Int J Cancer* 83: 625-629, 1999.
54. Petros JA, Baumann AK, Ruiz-Pesini E, Amin MB, Sun CQ, Hall J, Lim S, Issa MM, Flanders WD, Hosseini SH, *et al*: mtDNA mutations increase tumorigenicity in prostate cancer. *Proc Natl Acad Sci USA* 102: 719-724, 2005.
55. Yu M, Shi Y, Zhang F, Zhou Y, Yang Y, Wei X, Zhang L and Niu R: Sequence variations of mitochondrial DNA D-loop region are highly frequent events in familial breast cancer. *J Biomed Sci* 15: 535-543, 2008.
56. Liu VW, Wang Y, Yang HJ, Tsang PCK, Ng TY, Wong LC, Nagley P and Ngan HY: Mitochondrial DNA variant 16189T>C is associated with susceptibility to endometrial cancer. *Hum Mutat* 22: 173-174, 2003.
57. Canter JA, Kallianpur AR, Parl FF and Millikan RC: Mitochondrial DNA G10398A polymorphism and invasive breast cancer in African-American women. *Cancer Res* 65: 8028-8033, 2005.
58. Zhang J, Asin-Cayuela J, Fish J, Michikawa Y, Bonafe M, Olivieri F, Passarino G, De Benedictis G, Franceschi C and Attardi G: Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes. *Proc Natl Acad Sci USA* 100: 1116-1121, 2003.
59. Liou CW, Lin TK, Chen JB, Tiao MM, Weng SW, Chen SD, Chuang YC, Chuang JH and Wang PW: Association between a common mitochondrial DNA D-loop polycytosine variant and alteration of mitochondrial copy number in human peripheral blood cells. *J Med Genet* 47: 723-728, 2010.

60. Peng TI, Yu PR, Chen JY, Wang HL, Wu HY, Wei YH and Jou MJ: Visualizing common deletion of mitochondrial DNA-augmented mitochondrial reactive oxygen species generation and apoptosis upon oxidative stress. *Biochim Biophys Acta* 1762: 241-255, 2006.
61. Su X, Wang W, Ruan G, Liang M, Zheng J, Chen Y, Wu H, Fahey TJ III, Guan M and Teng L: A comprehensive characterization of mitochondrial genome in papillary thyroid cancer. *Int J Mol Sci* 17: 1594, 2016.
62. Wang Y, Liu VW, Xue WC, Tsang PC, Cheung AN and Ngan HY: The increase of mitochondrial DNA content in endometrial adenocarcinoma cells: A quantitative study using laser-captured microdissected tissues. *Gynecol Oncol* 98: 104-110, 2005.
63. Singh B, Owens KM, Bajpai P, Desouki MM, Srinivasasainagendra V, Tiwari HK and Singh KK: Mitochondrial DNA polymerase POLG1 disease mutations and germline variants promote tumorigenic properties. *PLoS One* 10: e0139846, 2015.
64. Reznik E, Miller ML, Şenbabaoğlu Y, Riaz N, Sarungbam J, Tickoo SK, Al-Ahmadie HA, Lee W, Seshan VE, Hakimi AA, *et al*: Mitochondrial DNA copy number variation across human cancers. *eLife* 5: e10769, 2016.
65. Booker LM, Habermacher GM, Jessie BC, Sun QC, Baumann AK, Amin M, Lim SD, Fernan-dez-Golarz C, Lyles RH, Brown MD, *et al*: North American white mitochondrial haplogroups in prostate and renal cancer. *J Urol* 175: 468-473, 2006.
66. Fang H, Shen L, Chen T, He J, Ding Z, Wei J, Qu J, Chen G, Lu J and Bai Y: Cancer type-specific modulation of mitochondrial haplogroups in breast, colorectal and thyroid cancer. *BMC Cancer* 10: 421, 2010.
67. Hu SP, Du JP, Li DR and Yao YG: Mitochondrial DNA haplogroup confers genetic susceptibility to nasopharyngeal carcinoma in Chaoshanese from Guangdong, China. *PLoS One* 9: e87795, 2014.
68. Li Y, Beckman KB, Caberto C, Kazma R, Lum-Jones A, Haiman CA, Le Marchand L, Stram DO, Saxena R and Cheng I: Association of genes, pathways, and haplogroups of the mitochondrial genome with the risk of colorectal cancer: The multiethnic Cohort. *PLoS One* 10: e0136796, 2015.
69. Jiang J, Zhao JH, Wang XL, Di JI, Liu ZB, Li GY, Wang MZ, Li Y, Chen R and Ge RL: Analysis of mitochondrial DNA in Tibetan gastric cancer patients at high altitude. *Mol Clin Oncol* 3: 875-879, 2015.
70. Poynter JN, Richardson M, Langer E, Hooten AJ, Roesler M, Hirsch B, Nguyen PL, Cioc A, Warlick E and Ross JA: Association between mitochondrial DNA haplogroup and myelodysplastic syndromes. *Genes Chromosomes Cancer* 55: 688-693, 2016.
71. Singh KK and Kulawiec M: Mitochondrial DNA polymorphism and risk of cancer. *Methods Mol Biol* 471: 291-303, 2009.
72. Yacoub HA, Mahmoud WM, El-Baz HA, Eid OM, El-Fayoumi RI, Mahmoud MM, Harakeh S and Abuzinadah OH: New haplotypes of the ATP synthase subunit 6 gene of mitochondrial DNA are associated with acute lymphoblastic leukemia in Saudi Arabia. *Asian Pac J Cancer Prev* 15: 10433-10438, 2014.
73. Cano D, Gomez CF, Ospina N, Cajigas JA, Groot H, Andrade RE and Torres MM: Mitochondrial DNA haplogroups and susceptibility to prostate cancer in a colombian population. *ISRN Oncol* 2014: 530675, 2014.
74. Yu FY, Xu Q, Wu DD, Lau AT and Xu YM: The Prognostic and Clinicopathological roles of Sirtuin-3 in various cancers. *PLoS One* 11: e0159801, 2016.
75. Czarnecka AM, Krawczyk T, Plak K, Klemba A, Zdrozny M, Arnold RS, Kofler B, Golik P, Szybinska A, Lubinski J, *et al*: Mitochondrial genotype and breast cancer predisposition. *Oncol Rep* 24: 1521-1534, 2010.
76. Verma M, Naviaux RK, Tanaka M, Kumar D, Franceschi C and Singh KK: Meeting report: Mitochondrial DNA and cancer epidemiology. *Cancer Res* 67: 437-439, 2007.
77. van Gisbergen MW, Voets AM, Starmans MH, de Coo IF, Yadak R, Hoffmann RF, Boutros PC, Smeets HJ, Dubois L and Lambin P: How do changes in the mtDNA and mitochondrial dysfunction influence cancer and cancer therapy? Challenges, opportunities and models. *Mutat Res Rev Mutat Res* 764: 16-30, 2015.
78. Parker SJ and Metallo CM: Metabolic consequences of oncogenic IDH mutations. *Pharmacol Ther* 152: 54-62, 2015.
79. Reitman ZJ, Jin G, Karoly ED, Spasojevic I, Yang J, Kinzler KW, He Y, Bigner DD, Vogelstein B and Yan H: Profiling the effects of isocitrate dehydrogenase 1 and 2 mutations on the cellular metabolome. *Proc Natl Acad Sci USA* 108: 3270-3275, 2011.
80. Burnichon N, Brière JJ, Libé R, Vescovo L, Rivière J, Tissier F, Jouanno E, Jeunemaitre X, Bénit P, Tzagoloff A, *et al*: SDHA is a tumor suppressor gene causing paraganglioma. *Hum Mol Genet* 19: 3011-3020, 2010.
81. Zhang B, Wang J, Huang Z, Wei P, Liu Y, Hao J, Zhao L, Zhang F, Tu Y and Wei T: Aberrantly upregulated TRAP1 is required for tumorigenesis of breast cancer. *Oncotarget* 6: 44495-44508, 2015.
82. Lin YF, Schulz AM, Pellegrino MW, Lu Y, Shaham S and Haynes CM: Maintenance and propagation of a deleterious mitochondrial genome by the mitochondrial unfolded protein response. *Nature* 533: 416-419, 2016.
83. McMahon S and LaFramboise T: Mutational patterns in the breast cancer mitochondrial genome, with clinical correlates. *Carcinogenesis* 35: 1046-1054, 2014.
84. Yu Y, Lv F, Lin H, Qian G, Jiang YS, Pang LX, Wang YP, Wang XF, Kang YM, Li CB, *et al*: Mitochondrial ND3 G10398A mutation: A biomarker for breast cancer. *Genet Mol Res* 14: 17426-17431, 2015.
85. Bai RK, Leal SM, Covarrubias D, Liu A and Wong LJ: Mitochondrial genetic background modifies breast cancer risk. *Cancer Res* 67: 4687-4694, 2007.
86. Zhu W, Qin W, Bradley P, Wessel A, Puckett CL and Sauter ER: Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid. *Carcinogenesis* 26: 145-152, 2005.
87. Wang Y, Liu VW, Tsang PC, Chiu PM, Cheung AN, Khoo US, Nagley P and Ngan HY: Microsatellite instability in mitochondrial genome of common female cancers. *Int J Gynecol Cancer* 16 (Suppl 1): 259-266, 2006.
88. Bianchi MS, Bianchi NO and Bailliet G: Mitochondrial DNA mutations in normal and tumor tissues from breast cancer patients. *Cytogenet Cell Genet* 71: 99-103, 1995.
89. Shakhssalim N, Houshmand M, Kamalidehghan B, Faraji A, Sarhangnejad R, Dadgar S, Mobaraki M, Rosli R and Sanati MH: The mitochondrial C16069T polymorphism, not mitochondrial D310 (D-loop) mononucleotide sequence variations, is associated with bladder cancer. *Cancer Cell Int* 13: 120, 2013.
90. Tipiriseti NR, Govatati S, Pullari P, Malempati S, Thupurani MK, Perugu S, Guruvaiah P, Rao KL, Digumarti RR, Nallanchakravarthula V, *et al*: Mitochondrial control region alterations and breast cancer risk: A study in South Indian population. *PLoS One* 9: e85363, 2014.
91. Fendt L, Niederstätter H, Huber G, Zelger B, Dünser M, Seifarth C, Röck A, Schäfer G, Klock-er H and Parson W: Accumulation of mutations over the entire mitochondrial genome of breast cancer cells obtained by tissue microdissection. *Breast Cancer Res Treat* 128: 327-336, 2011.
92. Ebner S, Lang R, Mueller EE, Eder W, Oeller M, Moser A, Koller J, Paulweber B, Mayr JA, Sperl W, *et al*: Mitochondrial haplogroups, control region polymorphisms and malignant melanoma: A study in middle European Caucasians. *PLoS One* 6: e27192, 2011.
93. Wang CY, Wang HW, Yao YG, Kong QP and Zhang YP: Somatic mutations of mitochondrial genome in early stage breast cancer. *Int J Cancer* 121: 1253-1256, 2007.
94. Czarnecka AM, Klemba A, Krawczyk T, Zdrozny M, Arnold RS, Bartnik E and Petros JA: Mitochondrial NADH-dehydrogenase polymorphisms as sporadic breast cancer risk factor. *Oncol Rep* 23: 531-535, 2010.
95. Weigl S, Paradiso A and Tommasi S: Mitochondria and familial predisposition to breast cancer. *Curr Genomics* 14: 195-203, 2013.
96. Li L, Chen L, Li J, Zhang W, Liao Y, Chen J and Sun Z: Correlational study on mitochondrial DNA mutations as potential risk factors in breast cancer. *Oncotarget* 7: 31270-31283, 2016.
97. Rohan TE, Wong LJ, Wang T, Haines J and Kabat GC: Do alterations in mitochondrial DNA play a role in breast carcinogenesis? *J Oncol* 2010: 604304, 2010.
98. Covarrubias D, Bai RK, Wong LJ and Leal SM: Mitochondrial DNA variant interactions modify breast cancer risk. *J Hum Genet* 53: 924-928, 2008.
99. Darvishi K, Sharma S, Bhat AK, Rai E and Bamezai RN: Mitochondrial DNA G10398A polymorphism imparts maternal Haplogroup N a risk for breast and esophageal cancer. *Cancer Lett* 249: 249-255, 2007.
100. Tengku Baharudin N, Jaafar H and Zainuddin Z: Association of mitochondrial DNA 10398 polymorphism in invasive breast cancer in Malay population of peninsular Malaysia. *Malays J Med Sci* 19: 36-42, 2012.
101. Tan DJ, Bai RK and Wong LJ: Comprehensive scanning of somatic mitochondrial DNA mutations in breast cancer. *Cancer Res* 62: 972-976, 2002.

102. Gallardo ME, Moreno-Loshuertos R, López C, Casqueiro M, Silva J, Bonilla F, Rodríguez de Córdoba S and Enríquez JA: m.6267G>A: a recurrent mutation in the human mitochondrial DNA that reduces cytochrome *c* oxidase activity and is associated with tumors. *Hum Mutat* 27: 575-582, 2006.
103. Aral C, Kaya H, Ataizi-Celikel C, Akkiprik M, Sonmez O, Gulluoglu BM and Ozer A: A novel approach for rapid screening of mitochondrial D310 polymorphism. *BMC Cancer* 6: 21, 2006.
104. Ma Y, Bai RK, Trieu R and Wong LJ: Mitochondrial dysfunction in human breast cancer cells and their trans-mitochondrial cybrids. *Biochim Biophys Acta* 1797: 29-37, 2010.
105. Santos GC Jr, Góes AC, Vitto H, Moreira CC, Avvad E, Rumjanek FD and Moura Gallo CV: Genomic instability at the 13q31 locus and somatic mtDNA mutation in the D-loop site correlate with tumor aggressiveness in sporadic Brazilian breast cancer cases. *Clinics (São Paulo)* 67: 1181-1190, 2012.
106. Xu C, Tran-Thanh D, Ma C, May K, Jung J, Vecchiarelli J and Done SJ: Mitochondrial D310 mutations in the early development of breast cancer. *Br J Cancer* 106: 1506-1511, 2012.
107. Parrella P, Xiao Y, Fliss M, Sanchez-Cespedes M, Mazzarelli P, Rinaldi M, Nicol T, Gabriel-son E, Cuomo C, Cohen D, *et al*: Detection of mitochondrial DNA mutations in primary breast cancer and fine-needle aspirates. *Cancer Res* 61: 7623-7626, 2001.
108. Alhomidi MA, Vedicherla B, Movva S, Rao PK, Ahuja YR and Hasan Q: Mitochondrial D310 instability in Asian Indian breast cancer patients. *Tumour Biol* 34: 2427-2432, 2013.
109. Lott MT, Leipzig JN, Derbeneva O, Xie HM, Chalkia D, Sarmady M, Procaccio V and Wallace DC: mtDNA variation and analysis using mitomap and mitomaster. *Curr Protoc Bioinformatics* 44: 1.23.1-1.23.26, 2013.
110. Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, Taylor RW, Samuels DC, Santibanez-Koref M and Chinnery PF: Universal heteroplasmy of human mitochondrial DNA. *Hum Mol Genet* 22: 384-390, 2013.
111. Weerts MJ, Sieuwerts AM, Smid M, Look MP, Foekens JA, Sleijfer S and Martens JW: Mitochondrial DNA content in breast cancer: Impact on *in vitro* and *in vivo* phenotype and patient prognosis. *Oncotarget* 7: 29166-29176, 2016.
112. Mueller EE, Brunner SM, Mayr JA, Stanger O, Sperl W and Kofler B: Functional differences between mitochondrial haplogroup T and haplogroup H in HEK293 cybrid cells. *PLoS One* 7: e52367, 2012.
113. Tommasi S, Favia P, Weigl S, Bianco A, Pilato B, Russo L, Paradiso A and Petruzzella V: Mitochondrial DNA variants and risk of familial breast cancer: An exploratory study. *Int J Oncol* 44: 1691-1698, 2014.
114. Blein S, Bardel C, Danjean V, McGuffog L, Healey S, Barrowdale D, Lee A, Dennis J, Kuchenbaecker KB, Soucy P, *et al*: Breast Cancer Family Registry; EMBRACE; GEMO Study Col-laborators; HEBON: An original phylogenetic approach identified mitochondrial haplogroup T1a1 as inversely associated with breast cancer risk in BRCA2 mutation carriers. *Breast Cancer Res* 17: 61, 2015.
115. Rao R, Rivers A, Rahimi A, Wooldridge R, Rao M, Leitch M, Euhus D and Haley BB: Genetic ancestry using mitochondrial DNA in patients with triple-negative breast cancer (GAMiT study). *Cancer* 123: 107-113, 2017.
116. Mosquera-Miguel A, Alvarez-Iglesias V, Carracedo A, Salas A, Vega A, Carracedo A, Milne R, de León AC, Benitez J, Carracedo A, *et al*: Is mitochondrial DNA variation associated with sporadic breast cancer risk? *Cancer Res* 68: 623-625, author reply 624, 2008.
117. van Oven M and Kayser M: Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30: E386-E394, 2009.
118. Salas A, García-Magariños M, Logan I and Bandelt HJ: The saga of the many studies wrongly associating mitochondrial DNA with breast cancer. *BMC Cancer* 14: 659, 2014.
119. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, Roach KC, Mandell J, Lee MK, Ciernikova S, *et al*: Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. *JAMA* 295: 1379-1388, 2006.
120. Hossein R and Houshmand M: Diagnostic algorithm for identification of individuals with hereditary predisposition to breast cancer. *Lik Sprava* 1-2: 103-108, 2008.
121. Jandova J, Janda J and Sligh JE: Changes in mitochondrial DNA alter expression of nuclear encoded genes associated with tumorigenesis. *Exp Cell Res* 318: 2215-2225, 2012.
122. Zhang Q, Liang Z, Gao Y, Teng M and Niu L: Differentially expressed mitochondrial genes in breast cancer cells: Potential new targets for anti-cancer therapies. *Gene* 596: 45-52, 2017.
123. Lin CS, Chang SC, Ou LH, Chen CM, Hsieh SS, Chung YP, King KL, Lin SL and Wei YH: Mitochondrial DNA alterations correlate with the pathological status and the immunological ER, PR, HER-2/neu, p53 and Ki-67 expression in breast invasive ductal carcinoma. *Oncol Rep* 33: 2924-2934, 2015.
124. Blein S, Barjhoux L, Damiola F, Dondon MG, Eon-Marchais S, Marcou M, Caron O, Lortholary A, Buecher B, Vennin P, *et al*; GENESIS investigators: Targeted sequencing of the mitochondrial genome of women at high risk of breast cancer without detectable mutations in BRCA1/2. *PLoS One* 10: e0136192, 2015.
125. Hsu CW, Yin PH, Lee HC, Chi CW and Tseng LM: Mitochondrial DNA content as a potential marker to predict response to anthracycline in breast cancer patients. *Breast J* 16: 264-270, 2010.
126. Farnie G, Sotgia F and Lisanti MP: High mitochondrial mass identifies a sub-population of stem-like cancer cells that are chemo-resistant. *Oncotarget* 6: 30472-30486, 2015.
127. Ferreri AJ, Ponzoni M, Guidoboni M, Resti AG, Politi LS, Cortelazzo S, Demeter J, Zallio F, Palmas A, Muti G, *et al*: Bacteria-eradicating therapy with doxycycline in ocular adnexal MALT lymphoma: A multicenter prospective trial. *J Natl Cancer Inst* 98: 1375-1382, 2006.
128. Lamb R, Harrison H, Hulit J, Smith DL, Lisanti MP and Sotgia F: Mitochondria as new therapeutic targets for eradicating cancer stem cells: Quantitative proteomics and functional validation via MCT1/2 inhibition. *Oncotarget* 5: 11029-11037, 2014.
129. Ferreri AJ, Ponzoni M, Guidoboni M, De Conciliis C, Resti AG, Mazzi B, Lettini AA, De-meter J, Dell'Oro S, Doglioni C, *et al*: Regression of ocular adnexal lymphoma after *Chlamydia psittaci*-eradicating antibiotic therapy. *J Clin Oncol* 23: 5067-5073, 2005.
130. Teixeira FK, Sanchez CG, Hurd TR, Seifert JRK, Czech B, Preall JB, Hannon GJ and Lehmann R: ATP synthase promotes germ cell differentiation independent of oxidative phosphorylation. *Nat Cell Biol* 17: 689-696, 2015.



TSH β X1 Splice Variant Expression and D2 Thr92Ala Polymorphism Analysis in Pregnant Women with Thyroid Diseases

Yesica María Rodríguez Cortés¹ · Araceli Amaya Chávez¹ · Magdalena García Fabila¹ · Silvia Jiménez Morales² · Carlos Jhovani Pérez Amado³ · Hugo Mendieta Zerón⁴

Accepted: 29 May 2019
© Springer Nature Switzerland AG 2019

Abstract

Thyroid diseases are pathologies that frequently affect pregnant women. The objective was to find out whether the expression of *TSH β X1* splice variant and D2 Thr92Ala polymorphism in the *DIO2* gene are associated with thyroid disease in pregnant women. It was a comparative, prospective clinical, and cross-sectional study. Ninety-two pregnant patients, of which 38 were normothyroid and 54 had thyroid disease, were included. The expression of the *TSH β X1* splice variant was quantified by real-time PCR and the D2 Thr92Ala (rs225014 CT) polymorphism was genotyped using TaqMan probes. The *TSH β X1* splice variant was detected in both groups without significant differences in the relative expression. The T allele was associated with the thyroid dysfunction (OR = 2.117, $p = 0.0339$). We confirmed the presence of the predicted *TSH β X1*, and our results suggest that 92Th (rs225017T) allele is a risk factor to thyroid dysfunction in pregnant women.

Keywords Hyperthyroidism · Hypothyroidism · Polymorphism · Pregnancy · TSH β X1 · D2 THR92ALA polymorphism

Introduction

Thyroid diseases are reported in 5 to 10% of all pregnancies and are associated with serious complications that can contribute to perinatal death and adverse obstetric events [1, 2]. In pregnancy, there is excessive thyroid stimulation [3] due to physiological changes such as increased thyroid-binding globulin (TBG), increased degradation of T4 to T3 by the inner ring of D3 abundantly expressed in the placenta, amnion, and chorion that generates reverse T3 and the effect of human chorionic gonadotropin (hCG). By sharing alpha chain with TSH, hCG generates a false stimulation of receptors

increasing the levels of T3 and T4, doing pregnancy a favorable scenario for thyroid dysfunctions [4].

On the other hand, deiodinase 2 (DIO2) is a selenoenzyme responsible for the deiodination of T4 to T3, which makes it crucial for the proper functioning of thyroid hormones. D2 is produced by the *DIO2* gene (NC_000014.9), which is located on chromosome 14 in the long arm position 14q24.3 [5]. A change of the T allele by C generates an amino acid coding alanine instead of threonine in position 92, which makes the protein differ from the canonical structure. In 2017, Wouters et al. proposed that the presence of the Thr92Ala (rs225014 CT) polymorphism of *DIO2* influences enzymatic function

This article is part of the Topical Collection on *Medicine*

✉ Hugo Mendieta Zerón
drmendietaz@yahoo.com

Yesica María Rodríguez Cortés
jmariarodriguez@unicolmayor.edu.co

Araceli Amaya Chávez
amayacha8789@gmail.com

Magdalena García Fabila
mmgafa@yahoo.com.mx

Silvia Jiménez Morales
sjimenez@inmegen.gob.mx

Carlos Jhovani Pérez Amado
jhamado24@gmail.com

- ¹ Autonomous University of the State of Mexico, Toluca, Mexico
- ² National Institute of Genomic Medicine, Mexico, Mexico
- ³ Biochemical Science Program, National Autonomous University of Mexico, Mexico, Mexico
- ⁴ Faculty of Medicine, Autonomous University of the State of Mexico and Ciprés Grupo Médico S.C. (CGM), Mexico, Mexico

and its presence is thought to predict a favorable response in a combination therapy of levothyroxine (LT4) and T3 [6]. Furthermore, the Ala92Ala genotype was reported to be related with a reduction in the placental activity of D2, which could worsen gestational complications [7].

In addition to the above, *TSH* variants have been proven to be related with thyroid dysfunction, osteoprotection in patients with hyperthyroidism (*TSH-βv*), or with immune processes (R55G variant) [8–10]. In this line of knowledge, it has been predicted by automated computational analysis a *Homo sapiens* *TSH* beta transcript variant X1 (*TSHβX1*) (XM_011542065.2) derived from a genomic sequence (NC_000001.11) using the Gnomon gene prediction method [11]. The predicted variant belongs to the BioProject: PRJNA168, related to the GenBank project: PRJNA31257.

The aim of this study was to describe the allele frequency of the Thr92Ala in pregnant women with or without thyroid disease and to confirm the *TSHβX1* splice variant expression in this population.

Material and Methods

Setting

This was a clinical, comparative, prospective, and transversal study. Pregnant women aged 18–44 years old, attended at the “Mónica Pretelini Sáenz” Maternal-Perinatal Hospital (HMPMPS), Health Institute of the State of Mexico (ISEM), Toluca, Mexico, were invited to participate.

Two groups were conformed: (A) Normothyroid pregnant women (NT) and (B) patients with a thyroid disease named hypothyroidism or hyperthyroidism (TD). Patients with other chronic diseases were excluded and those whose clinical follow-up was lost or that in the puerperium had to be treated in the obstetric intensive care unit were eliminated from the study, besides the samples that were not suitable for genetic analysis were eliminated too.

Sample Calculation

The sample was calculated with Eq. 1.

$$n_0 = \frac{2(Z_\alpha + Z_\beta)^2 S^2}{d^2} \quad (1)$$

where n_0 : necessary size of the sample, S : standard deviation, d : difference to find, $Z_\alpha = 1.96$, $Z_\beta = 0.482$. The relative units (UR) of gene expression were taken into account as measures, accepting an alpha risk of 0.05 and a beta risk of 0.2 in a bilateral contrast, obtaining a sample size of 48 subjects per group, to detect an equal or greater difference at 4 RU with a standard deviation of 7.

Sociodemographic Information

All the participants answered a questionnaire that included demographic information such as name, age and address, obstetric data (number of pregnancies, deliveries, cesareans, and abortions), and familiar thyroid disease background. The anthropometric and laboratory data (weight, height, body mass index (BMI), glucose, creatinine, triglycerides, total cholesterol, and TSH) from all the patients were obtained from the clinical files.

Blood Samples

Peripheral blood samples (2 tubes of 5 mL, a total blood tube with EDTA anticoagulant for the genetic analysis, and a dry tube for the determination of the hormone TSH) were collected by venipuncture in the arm. All samples were stored at -80°C until the analysis.

TSH Quantification

TSH measurements were performed using the ELISA technique (diagnostics A) at the Research Laboratory of Ciprés Grupo Médico S.C. (CGM).

Identification of the D2 Thr92Ala Polymorphism

DNA Extraction

The DNA was extracted in the Laboratory of Genetics, Faculty of Medicine, UAEMex, using the kit QIAmp DNA Mini and Blood kit from Qiagen DNA, (Cat no./ID: 51106). The genetic material purity and integrity were measured with the Nanodrop 1000 and was stored at -20°C until genotyping. DNA integrity was verified by E-Gel gels in the E-Gel iBase and E-Gel Safe Imager (Invitrogen).

Genotyping

The genotyping was carried out with a total concentration of 30 ng/ μl in each reaction well, using 0.06 μl of the rs225014 (Thr92Ala) primers and probes contained in Human Custom TaqMan Genotyping Assay C__15819951_10, 2.5 μl of Master Mix, 2.44 μl of water and 5 μl of DNA (all from Thermo Fisher Scientific, Waltham, MA, USA). Then, a conventional PCR was performed in a GeneAmp System 9700 thermal cycler using the following cycling program: denaturation of: 50°C 2 min, and 95°C 10 min, followed by 40 cycles of 94°C 15 s, and 40 cycles 60°C 1 min; to finish at 4°C 7 min. At the end of the PCR program, the amplified plate was taken to the QuantStudio 3 Real-Time PCR System

genotyping equipment using TaqMan probes to process each sample of the patients. Genotyping data were validated by sequencing three samples from each genotype as we previously reported [12] and using the following primers: 5' CTGGCTCGTGAAAGGAGGTCAA 3' and 3'CCAATTCCAGTGTGGTGCATGT 5' [13].

TSH β X1 Expression Analysis

RNA Isolation

The RNA was extracted in the Laboratory of Genetics, Faculty of Medicine, UAEMex, using the Norgen's Total RNA Purification Kit (Norgen Biotek Corp.). The purity and integrity of total RNA were assessed using a NanoPhotometer (Implen GmbH, Germany).

Real-Time PCR

In the Research Laboratory of Ciprés Grupo Médico S.C. (CGM), the *TSH β X1* expression was determined by the real-time PCR technique in a PrimeQ (Techne, UK) equipment. The primers used were synthesized at the Synthesis and DNA Sequencing Unit of the National Autonomous University of Mexico (UNAM), Institute of Biotechnology (Cuernavaca, Morelos, Mexico). The SYBR Green RNA to Ct 1 step Kit (Cat. number 4389986) was used with the following oligonucleotides: *GADPH* (House keeping gene): 5'CTTGGTATCGTGGAAGGACTC 3', 3'GTAGAGGCAGGGATGATGTTCT 5' and *TSH β X1*: 5'TGTGGGCAAGCGATGTCTTTT 3', 3'GATGGTTAGGCAATAAGCACACT 5'. The reaction was done with a final volume of 50 μ l, the primers' concentration were of 200 nM and the samples of 100 ng. Cycling was programmed in the following way: retrotranscription step (48 °C 30 min \times 1 cycle), enzyme activation (95 °C 10 min \times 1 cycle), denaturation (95 °C 15 s \times 40 cycles), alignment (59 °C 30 s \times 40 cycles), extension (60 °C 30 s \times 40 cycles), and melting curve (72 °C 1 min \times 61 cycles). The fold change in *TSH β X1* splice variant was

normalized against *GADPH* and then compared with the controls through the $2^{-\Delta\Delta CT}$ method.

Statistical Analysis

Hardy-Weinberg equilibrium (HWE) was evaluated to the Thr92Ala genotypes by using Finetti program (<https://ihg.gsf.de/cgi-bin/hw/hwa2.pl>). Continuous data were expressed as means \pm standard deviation (SD). Statistical analysis was performed using the Mann-Whitney *U* test after performing the Levene test. The normality hypothesis was tested using the Kolmogorov-Smirnov test. Odds ratio (OR) test was used to compare the *TSH β X1* splice variant positive identification, D2 Thr92Ala genotypes, and alleles with the presence/absence of thyroid disease. Statistical significance was tested at the $p \leq 0.05$ level using SPSS ver. 23.0 software (IBM Corp., Armonk, NY, USA).

Ethics

All participants gave their informed consent and both the Research Committee and the Ethics on Research Committee of the HMPMPS approved the protocol (code: 2016-09-481). All procedures were performed in accordance with the Declaration of Helsinki and the General Health Law in Mexico.

Results

Sociodemographic and Anthropometric Data

Of all the 96 samples collected, four were excluded because their clinical follow-up was lost; obtaining an $N = 92$, of which 54 had thyroid dysfunction (cases) and 38 were NT (controls). The initial proportion of both groups changed because in the analysis of TSH serum levels, it was found that more patients had altered hormone through the follow-up period. The description of the study population is depicted in Table 1.

There were no differences in the number of abortions among groups, which led to infer that in the patients analyzed,

Table 1 General description of study population

Group	Age (years) mean \pm SD	Gestational antecedent		Previous abortions <i>N</i> (%)	Family history of thyroid dysfunction <i>N</i> (%)
		Primigravid <i>N</i> (%)	> 1 pregnancies <i>N</i> (%)		
Control group (<i>N</i> = 38)	27 \pm 8	15 (39.47)	23 (60.52)	11 (28.94)	5 (13.15)
Thyroid dysfunction (<i>N</i> = 54)	28 \pm 5	22 (40.7)	32 (59.25)	14 (20.37)	19 (35.18)
Total (<i>N</i> = 92)	28 \pm 7	37 (40.21)	55 (59.78)	25 (27.17)	24 (26.08)

the thyroid pathology was not influencing the gestational losses. Regarding the relationship of the health status of the patients with the presence or absence of a family history of thyroid disease, it was observed that while 86% ($n = 33$) of the patients in the NT group had no family history, by contrast, 35.18% ($n = 19$) of the patients with a thyroid dysfunction had a family member with a thyroid-related disease. A significant $p < 0.05$ was obtained when looking for this relationship, evidencing that there is a connection between family history and thyroid dysfunction. The descriptive statistics of the groups are observed in Table 2.

In terms of BMI, most of the patients belong to the category of overweight and obesity. However, there was no significant difference between the two groups. Regarding blood pressure, only one patient in each study group presented hypertension. In addition, 13 different dosage types of levothyroxine or thiamazol were registered and only one patient was untreated.

TSH Screening

Taking into account the reference values for TSH per trimester [14], seven patients were reassigned to the group of thyroid diseases because their serum TSH levels exceeded the limits allowed in the specific gestational trimester.

D2 Thr92Ala Allele Frequency

The genotypes were in HWE, showing the next distribution: 40 (48.2%) patients presented the CC genotype, of which 22 (55%) had no thyroid disease, while 18 (45%) were from the group with thyroid dysfunction. For the CT genotype, which was found in 33 (39.8%) patients, the opposite happened that the majority of the ($n = 23$, 69.7%) patients with thyroid dysfunction had it and 10 (30.3%) were from the NT group. The

TT genotype was present in 10 (12%) patients who, in their majority ($n = 7$, 70%), were patients with thyroid dysfunction. In summary, the T allele was the most common in cases with thyroid dysfunction in comparison with NT (OR = 2.117, $p = 0.0339$) (Table 3). Comparisons between TSH levels and each genotype were performed without finding significant differences.

TSH β X1 Expression

This *TSHB*X1 splice variant was present in 14 (25.92%) of the cases and in 13 (34.21%) of the controls. Eleven cDNA samples were not adequate for expression analysis. When performing the OR analysis between the state of health and the presence/absence of the gene transcript, there was not confirmation of a significant statistical difference (OR 0.6731, 95% CI 0.2722–1.6643, $p = 0.3914$). Otherwise, the fold change was of 3.33 in the TD group vs ND group for the relative expression of this splice variant.

Discussion

The thyroid diseases do not have a unique origin; genetic, endogenous, and environmental factors are involved [15, 16]. It has been previously published that a relationship between obesity and hypothyroidism in which leptin levels intervene and in which TSH is directly responsible for weight gain [17]. Perhaps, a population without these basic problems would have seen a significant difference between the two study groups.

Of particular concern is that not suffering a thyroid disease in pregnancy does not exonerate a healthy patient from having complications in postpartum due to the thyroid gland. In fact,

Table 2 Anthropometric, clinical characteristics, and biochemical data of the population

Variable	Total study population	Thyroid dysfunction group ($n = 54$)	Normothyroid group ($n = 38$)
Weight (kg)	70.0 (± 14.3)	70.1 (± 14.7)	69.9 (± 14.1)
Height (cm)	155 (± 0.0)	155 (± 0.06)	155 (± 0.05)
SBP (mmHg)	108.4 (± 11.8)	109.3 (± 13.4)	107.3 (± 9.5)
DBP (mmHg)	68.3 (± 10.3)	69.2 (± 11.9)	67.2 (± 7.9)
BMI (kg/m^2)	28.9 (± 5.0)	29.01 (± 5.3)	28.8 (± 4.6)
Glucose (mg/dL)	80.0 (± 16.0)	79 (± 18)	81.3 (± 13.3)
TC (mg/dL)	205 (± 54.7)	216.3 (± 62.0)	188.5 (± 37.9)
TG (mg/dL)	213.4 (± 94.9)	225.3 (± 105.1)	198.4 (± 80.2)
Hb (g/dL)	13.2 (± 1.5)	12.9 (± 1.7)	13.5 (± 1.1)
Creatinine (mg/dL)	0.6 (± 0.1)	0.60 (± 0.1)	0.58 (± 0.1)

BMI body mass index, DBP diastolic blood pressure, Hb hemoglobin, SBP systolic blood pressure, TC total cholesterol, TG triglycerides

Table 3 Genotype and allele distribution per group

Thr92Ala rs225014 CT	No thyroid disease <i>n</i> (%)	Thyroid disease <i>n</i> (%)	OR IC (95%)	<i>p</i> value
Genotype	<i>N</i> = 35	<i>N</i> = 48		
Ala92Ala (CC)	22 (62.8)	18 (37.5)		0.07
Thr92Ala (TC)	10 (28.6)	23 (47.9)		
Thr92AThr (TT)	3 (8.6)	7 (14.6)		
CC vs TC			2.811 [1.067–7.409]	0.0343
CC vs TT			2.852 [0.643–12.642]	0.1573
Allele				
C	54 (73.3)	59 (61.5)		
T	16 (26.7)	37 (38.5)	2.1165 [1.0585–4.2322]	0.0339

99% of genotyping (low DNA quality in one sample)

5 to 18% of healthy patients present in the first year a postpartum thyroiditis [18, 19]. What is clear from this study is the relationship between the family history of thyroid disease and the health status of our patients. The previous thing verifies the great influence of the genetic background in the presentation of the disease. In this case, there are many genes involved in the development of these diseases, which can follow a pattern in terms of anomalies over several generations. In addition, the most frequent pathology at the thyroid level was clinical hypothyroidism, which is consistent with the epidemiological data regarding the incidence of the disease [20].

For instance, *DIO2* and *TSH* have been implicated in the thyroid disease development. Polymorphisms in *DIO2* could interfere in the expression level or the activity of this enzyme, potentially altering the metabolism of thyroid hormones [7, 21]. On this field, comparing the Thr92Thr homozygous and Thr92Ala heterozygous genotypes, the D2 activity has been found to decline among Ala92Ala homozygous [22]. In fact, the D2 Ala92Ala genotype is related with a reduction in the placental activity of D2, which could worsen gestational complications [7]. Studying pregnant women with or without thyroid disease, we identified the Thr92 (T) allele as a risk factor to thyroid disease in pregnant women. So far, most of the studies attribute the C allele as the diseased genotype since it has been associated with drug resistance, lower T3 values, and complications with diabetes [23–25]. However, other group reported the C allele as more common in healthy patients suggesting that it is a low-risk factor for the disease [26] which is in accordance with our data. A differential distribution of these alleles between ethnic groups have been reported, being the C allele the less frequent in populations with Asian and European ancestry than subjects with Amerindian background [27]. These contradictory results have to be elucidated; however, we cannot discard that sample sizes and the genetic background could bias our results.

An important issue to take into account is the identification of possible risk factors in order to improve the prevention methods in pregnancies' complications. According to the 2017 Guidelines of the American Thyroid Association (ATA) for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum, if any of the following risk factors are identified, testing for serum TSH is recommended: (1) a history of hypothyroidism/hyperthyroidism or current symptoms/signs of thyroid dysfunction, (2) known thyroid antibody positivity or presence of a goiter, (3) history of head or neck radiation or prior thyroid surgery, (4) age > 30 years, (5) type 1 diabetes or other autoimmune disorders, (6) history of pregnancy loss, preterm delivery, or infertility, (7) multiple prior pregnancies (≥ 2), (8) family history of autoimmune thyroid disease or thyroid dysfunction, (9) morbid obesity (BMI ≥ 40 kg/m²), (10) use of amiodarone or lithium, or recent administration of iodinated radiologic contrast, and (11) residing in an area of known moderate to severe iodine insufficiency [14].

From the above mentioned risk factors, autoimmunity in our population seems to be one of the most important factors, since we must take into account that the percentage of mothers with AITD and with an affected daughter is 13% [28], but the challenge we face is to detect this autoimmunity process before pregnancy. Even more, the results of this survey support the usefulness of the identification of polymorphisms and genetic variants that increase the risk of developing thyroid disease during pregnancy.

Our study constitutes a characterization of the Mexican population and also confirms a role for a strict thyroid function monitoring during pregnancy [4]. In case of institutional budget limitations, TSH must be measured in the first trimester to all pregnant women with a strong family history of thyroid disease, over 30 years of age, with autoimmune diseases, morbid obesity, or who have received radiation in head and neck [29].

Finally, in this initial approach, the higher fold change in the expression of the *TSH β X1* splice variant in the TD group makes suppose an implication of this variant to develop a thyroid dysfunction. Notwithstanding, the information is not conclusive and needs further corroboration.

At this stage, the presence of the *TSH β X1* splice variant in the analyzed population opens the possibility of a new study, in which the role of the interactome in thyroid dysfunctions can be evidenced, to see if the association of the resulting protein with other molecules is involved with the disease. In other words, the studies have to establish if the protein-protein interactions are affecting the cellular response [30].

Limitations of our study include the relatively low number of patients, even for our institution, attributed to the rate of acceptance of the patients to participate into any study but at the same time, it is valuable to see ethical recommendations avoiding coercion in such an important decision. Secondly, the possibility of stratification bias cannot be ruled out in our sample due to the treatment of the patients. It is also a limitation the scarce information of the *TSH β X1* splice variant thus restricting the conclusions of its positive expression. Further, the TSH serum increments in the NT group during the following period made us lose the initial equal distribution of the groups, forcing us to use nonparametric tests in the analysis.

Conclusions

Our data suggests that the allele T of Ala92Ala is associated with thyroid disease. To the best of our knowledge, this is the first study showing expression of the predicted *TSH β X1* splice variant in human pregnant women, being an example of the certainty of the informatics sequence genome prediction.

Acknowledgments The authors thank MSc. Cristian F. Layton Tovar for his help in the TSH quantification and to all the clinical staff of the Maternal-Fetal Medicine Service of the HMPMPS for their support to develop this project.

Funding This project was partially funded by Ciprés Grupo Médico, Grant 2017-07.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval The Ethics on Research Committee of the HMPMPS approved the protocol (code: 2016-09-481).

Informed Consent All participants gave their informed consent.

References

- González-Velázquez A, Ávalos-Guerrero Á, Ramírez-Montiel ML, et al. Incidencia de patología tiroidea durante el embarazo. *Méd Sur*. 2003;20:11–6.
- Nijkamp JW, Korteweg FJ, Groen H, Timmer A, van den Berg G, Bossuyt PM, et al. Thyroid function testing in women who had a stillbirth. *Clin Endocrinol*. 2016;85:291–928. <https://doi.org/10.1111/cen.13002>.
- Nazarpour S, Tehrani FR, Simbar M, et al. Thyroid and pregnancy in Tehran, Iran: objectives and study protocol. *Int J Endocrinol Metab*. 2016;14:e33477. <https://doi.org/10.5812/ijem.33477>.
- Leung AM. Thyroid function in pregnancy. *J Trace Elem Med Biol*. 2012;26:137–40. <https://doi.org/10.1016/j.jtemb.2012.03.004>.
- Araki O, Murakami M, Morimura T, et al. Assignment of type II iodothyronine deiodinase gene (DIO2) to human chromosome band 14q24.2–q24.3 by in situ hybridization. *Cytogenet Cell Genet*. 1999;84:73–4. <https://doi.org/10.1159/000015218>.
- Wouters HJCM, van Loon HCM, van der Klauw MM, Elderson MF, Slagter SN, Kobold AM, et al. No effect of the Thr92Ala polymorphism of deiodinase-2 on thyroid hormone parameters, health-related quality of life, and cognitive functioning in a large population-based cohort study. *Thyroid*. 2017;27:147–55. <https://doi.org/10.1089/thy.2016.0199>.
- Dora JM, Wajner SM, Costa JD, Pinto Ribeiro RV, Leiria LB, Lopes MG, et al. Type 2 deiodinase Thr92Ala polymorphism is associated with disrupted placental activity but not with dysglycemia or adverse gestational outcomes: a genetic association study. *Fertil Steril*. 2014;101:833–40. <https://doi.org/10.1016/j.fertnstert.2013.11.018>.
- Moreno-Navarrete JM, Moreno M, Ortega F, Xifra G, Hong S, Asara JM, et al. TSHB mRNA is linked to cholesterol metabolism in adipose tissue. *FASEB J*. 2017;31:4482–91. <https://doi.org/10.1096/fj.201700161R>.
- Baliram R, Latif R, Zaidi M, Davies TF. Expanding the role of thyroid-stimulating hormone in skeletal physiology. *Front Endocrinol (Lausanne)*. 2017;8:252. <https://doi.org/10.3389/fendo.2017.00252>.
- Pappa T, Johannesen J, Scherberg N, Torrent M, Dumitrescu A, Refetoff S. A TSH β variant with impaired immunoreactivity but intact biological activity and its clinical implications. *Thyroid*. 2015;25:869–76. <https://doi.org/10.1089/thy.2015.0096>.
- PREDICTED: *Homo sapiens* thyroid stimulating hormone beta (TSHB), transcript variant X1, mRNA. Available from: https://www.ncbi.nlm.nih.gov/nuccore/XM_011542065.2. Accessed 30 March 2019.
- Mendieta Zerón H, Jiménez Rosales A, Pérez Amado CJ, et al. FOXE1 mutation screening in a case with cleft lip, hypothyroidism and thyroid carcinoma: a new syndrome? *Case. Rep Genet*. 2017;2017:6390545. <https://doi.org/10.1155/2017/6390545>.
- Luo M, Zhou XH, Zou T, Keyim K, Dong LM. Type II deiodinase polymorphisms and serum thyroid hormone levels in patients with mild cognitive impairment. *Genet Mol Res*. 2015;14:5407–16. <https://doi.org/10.4238/2015.May.22.10>.
- Alexander EK, Pearce EN, Brent GA, Brown RS, Chen H, Dosiou C, et al. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease during pregnancy and the postpartum. *Thyroid*. 2017;27:315–89. <https://doi.org/10.1089/thy.2016.0457>.
- Janssen BG, Saenen ND, Roels HA, Madhloum N, Gyselaers W, Lefebvre W, et al. Fetal thyroid function, birth weight, and in utero exposure to fine particle air pollution: a birth cohort study. *Environ Health Perspect*. 2017;125:699–705. <https://doi.org/10.1289/EHP508>.
- Kahr MK, Antony KM, DelBeccaro M, Hu M, Aagaard KM, Suter MA. Increasing maternal obesity is associated with alterations in

- both maternal and neonatal thyroid hormone levels. *Clin Endocrinol*. 2016;84:551–7. <https://doi.org/10.1111/cen.12974>.
17. Sanyal D, Raychaudhuri M. Review article hypothyroidism and obesity: an intriguing link. *Indian J Endocrinol Metab*. 2016;20:554–7. <https://doi.org/10.4103/2230-8210.183454>.
 18. Alemu A, Terefe B, Abebe M, Biadgo B. Thyroid hormone dysfunction during pregnancy: a review. *Int J Reprod Biomed (Yazd)*. 2016;14:677–86.
 19. Pedersen C, Leserman J, Garcia N, Stansbury M, Meltzer-Brody S, Johnson J. Late pregnancy thyroid-binding globulin predicts perinatal depression. *Psychoneuroendocrinology*. 2016;65:84–93. <https://doi.org/10.1016/j.psyneuen.2015.12.010>.
 20. Panicker V. Genetics of thyroid function and disease. *Clin Biochem Rev*. 2011;32:165–75.
 21. Leiria LB, Dora JM, Wajner SM, Estivalet AAF, Crispim D, Maia AL. The rs225017 polymorphism in the 3'UTR of the human DIO2 gene is associated with increased insulin resistance. *PLoS One*. 2014;9(8):e103960. <https://doi.org/10.1371/journal.pone.0103960>.
 22. Canani LH, Capp C, Dora JM, Meyer ELS, Wagner MS, Harney JW, et al. The type 2 deiodinase A/G (Thr92Ala) polymorphism is associated with decreased enzyme velocity and increased insulin resistance in patients with type 2 diabetes mellitus. *J Clin Endocrinol Metab*. 2005;90:3472–8. <https://doi.org/10.1210/jc.2004-1977>.
 23. Carlé A, Laurberg P. Hypothyroid patients encoding combined MCT10 and DIO2 gene polymorphisms may prefer L-T3 + L-T4 combination treatment – data using a blind, randomized, clinical study. *Eur Thyroid J*. 2017;6:143–51. <https://doi.org/10.1159/000469709>.
 24. Castagna MG, Dentice M, Cantara S, Ambrosio R, Maino F, Porcelli T, et al. DIO2 Thr92Ala reduces deiodinase-2 activity and serum-T3 levels in thyroid-deficient patients. *J Clin Endocrinol Metab*. 2017;102:1623–30. <https://doi.org/10.1210/jc.2016-2587>.
 25. Yalakanti D, Dolia PB. Association of Type II 5' monodeiodinase Thr92Ala single nucleotide gene polymorphisms and circulating thyroid hormones among type 2 diabetes mellitus patients. *Indian J Clin Biochem*. 2016;31:152–61. <https://doi.org/10.1007/s12291-015-0518-9>.
 26. Gałecka E, Talarowska M, Orzechowska A, Górski P, Bieńkiewicz M, Szemraj J. Association of the DIO2 gene single nucleotide polymorphisms with recurrent depressive disorder. *Acta Biochim Pol*. 2015;62:297–302. https://doi.org/10.18388/abp.2015_1002.
 27. Rajeevan H, Osier MV, Cheung KH, Deng H, Druskin L, Heinzen R, et al. ALFRED: the Allele FREquency database. Update. *Nucleic Acids Res*. 2003;31:270–1.
 28. Conzuelo Rodríguez G, Mendieta Zerón H. Familial autoimmune thyroid disease and PTPN-22. *Med Glas (Zenica)*. 2015;12:151–6. <https://doi.org/10.17392/803-15>.
 29. Stagnaro-Green A, Abalovich M, Alexander E, et al. American Thyroid Association taskforce on thyroid disease during pregnancy and postpartum. American Thyroid Association taskforce on thyroid disease during pregnancy and postpartum. Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and postpartum. *Thyroid*. 2011;21:1081–125. <https://doi.org/10.1089/thy.2011.0087>.
 30. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, et al. The BioPlex network: a systematic exploration of the human Interactome. *Cell*. 2015;162:425–40. <https://doi.org/10.1016/j.cell.2015.06.043>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Case Report

FOXE1 Mutation Screening in a Case with Cleft Lip, Hypothyroidism, and Thyroid Carcinoma: A New Syndrome?

Hugo Mendieta-Zerón,^{1,2} Angélica Jiménez-Rosales,¹
Carlos Jhovani Pérez-Amado,³ and Silvia Jiménez-Morales⁴

¹Ciprés Grupo Médico S.C. (CGM), Felipe Villanueva 700, Col. Morelos, 50120 Toluca, MEX, Mexico

²Universidad Autónoma del Estado de México (UAEMex), Felipe Villanueva Sur 1209, Col. Rancho Dolores, 50170 Toluca, MEX, Mexico

³Programa de Maestría en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico

⁴Instituto Nacional de Medicina Genómica (INMEGEN), Mexico City, Mexico

Correspondence should be addressed to Hugo Mendieta-Zerón; mezh_74@yahoo.com

Received 3 May 2017; Revised 19 July 2017; Accepted 30 July 2017; Published 27 August 2017

Academic Editor: Balraj Mittal

Copyright © 2017 Hugo Mendieta-Zerón et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A 26-year-old woman is referred to the Internal Medicine consultation due to increases in laboratory studies associated with Papillary Thyroid Carcinoma (PTC) that was confirmed by histopathological studies. Her clinical history revealed that, at 3 months of age, she was successfully treated with surgery for cleft lip (CL) and at the age of 24 years was diagnosed with hypothyroidism. Single nucleotide polymorphisms (SNPs) in *FOXE1* and its promoter regions have been associated with various etiologies related to the thyroid, including orofacial clefting, specially cleft palate (CP) and CL, hypothyroidism (HT), and thyroid cancer. The association of CL, HT, and PTC might be component of a new syndrome; however *FOXE1* coding region, which has been involved with these entities, has not exhibited mutations or SNPs. Further study of other genes may help in better characterization of the possible syndrome.

1. Introduction

Transcription factors NKX2-1, *FOXE1*, *HHEX*, and *PAX8* are involved in cellular differentiation during embryogenesis. They play a critical role in the morphogenesis, differentiation, and maintenance of the thyroid gland. The *FOXE1* gene encodes for a transcription factor protein that is expressed from the embryonic stage in the thyroid primordium until the development of the thyroid gland by the regulation of thyroid promoters [1]. Single nucleotide polymorphisms (SNPs) in this gene and its promoter regions have been associated with various etiologies related to the thyroid, including orofacial clefting, especially cleft palate (CP) and cleft lip (CL), hypothyroidism (HT), and thyroid cancer (TC).

CP and CL are common birth defects, particularly in Asian and Native American population, which have the highest rates of prevalence, in contrast to African population, which has the lowest. Both have a complex etiology, in

that they possess multigenetic and multifactorial causes that have not been elucidated. However, research and mapping of the 9q22-q33 guided the associations with several SNPs located in the *FOXE1* locus [2]. Actually, the genotypes involving the commonest alleles of rs3758249 (GG) and rs4460498 (CC) were the most associated with CP in Caucasian and Asian derived populations [3–5]. To note, both SNPs are located in a high linkage disequilibrium (LD) region, which also harbors the rs1867277 (–283G>A *FOXE1*) SNPs that have been reported as functional in thyroid cancer [6].

HT is the most common thyroid disorder. It is characterized by low production of thyroid hormones T3 and T4. Through Genome-Wide Association Studies (GWAS), various polymorphisms, including SNP located in the *FOXE1* gene, have been identified with problems in the formation and differentiation of the thyroid, producing a predisposition to the disease [5].

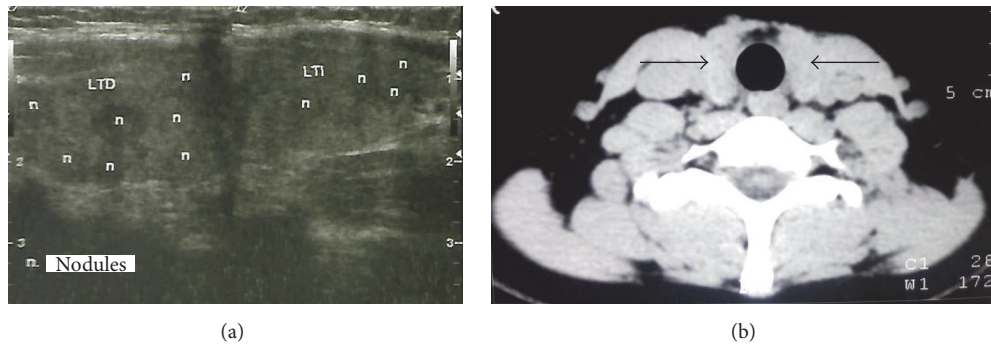


FIGURE 1: The imaging studies showed multinodularity (US) (a) and thyroid growth (CAT) (b). Arrows in (b) show an enlarged thyroid gland.

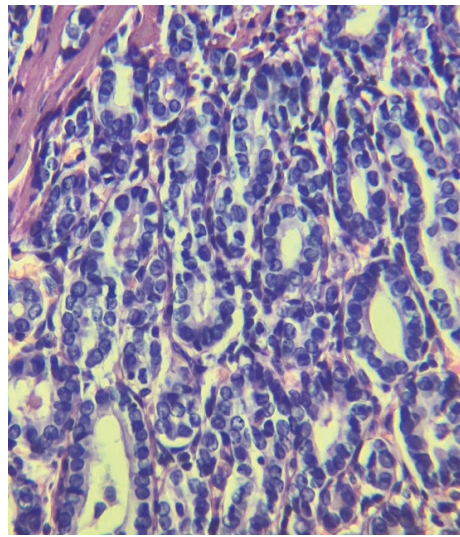


FIGURE 2: Thyroid fine needle aspiration (FNA) biopsy.

TC is the most common endocrine malignancy, with a strong genetic component that has been shown to extend beyond the nuclear family. Histologically, it is classified as Papillary (PTC), Follicular (FTC), and Medullary (MTC) Carcinomas, and undifferentiated anaplastic thyroid carcinomas, and studies have suggested that gender, age, tumor size, histologic type, tumor infiltration, and vascular/lymphatic invasion affect clinical outcome and treatment options [13]. Evidence in multiple ethnicities has associated SNP with the same LD region as *FOXE1* with PTC and MTC [1, 6]. The case described here represents unexpected combination of CL, HT, and PTC.

2. Case Report

A 26-year-old woman from Jalisco (Mexico) is referred to the Internal Medicine consultation due to increases in laboratory studies associated with PTC. Her clinical history revealed that, at 3 months of age, she was successfully treated with surgery for CL, at 14 years of age underwent a cholecystectomy, and at the age of 24 years was diagnosed with HT. In April 2016, her clinical examinations revealed that she had anti-thyroglobulin antibodies ($>4,000 \mu\text{g/L}$) and

anti-peroxidase antibodies $> 42 \text{ IU/mL}$ (normal reference value, 34). Imaging studies demonstrated multinodularity (Ultrasound [US]) (Figure 1(a)), discrete thyroid growth (Figure 1(b)) and the thyroid fine needle aspiration (FNA) biopsy corroborated PTC Bethesda Class V (Figure 2).

Based on the contribution of *FOXE1* mutations and SNPs variants (Figure 3) in the development of the CL, HT, and PC, we sequenced the exon 1 of this gene (coding region and 3'UTR) in order to identify mutations or SNPs potentially associated with the patient's diseases. At the Research Laboratory of Ciprés Grupo Médico S.C. (CGM), Genomic DNA was isolated from peripheral white blood cells using the DNA purification kit (Quick-DNA Miniprep Plus Kit, Zymo Research Corp) according to the manufacturer's instructions. At the INMEGEN Exon 1 of *FOXE1* gene was analysed using the primers described in Table 1. PCR was performed in a final volume of $50 \mu\text{l}$, using 50 ng of genomic DNA, 2.5 pmoles of each primer, 2 mM of each dNTP, 1 U of AmpliTaq Gold DNA Polymerase, 4 mM MgCl_2 , and 10x PCR Gold buffer (Applied Biosystems provided polymerase, MgCl_2 , and buffer). Cycling conditions for PCR consisted of a first denaturation step of 95°C for 10 min, followed by 35 cycles of denaturation at 95°C for 20 s, annealing at

TABLE 1: Sequence of the primers used.

ID	Sequence	Length	Tm	Amplicon
1F	GTCACTCCCGAGCCTCTGT	19	60.41	395 pb
1R	GTAGCTGTAGGGCGGCTTC	19	59.99	
2F	AGCCGCCCTACAGCTACAT	19	59.87	488 pb
2R	CGCGGGGTAGTAGACTGG	18	59.26	
3F	GCCGTCTATGCAGGCTACG	19	61.8	639 pb
3R	GGGTCCCAGTTGAGTCCTCT	20	60.5	

F: forward; R: reverse.

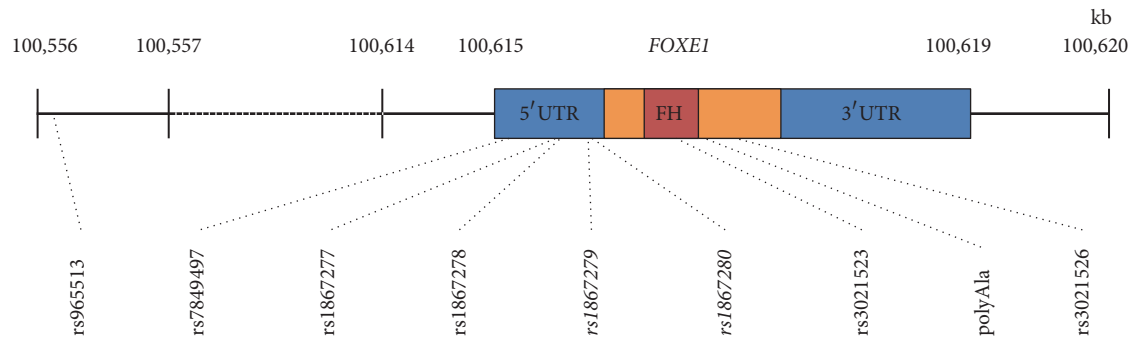


FIGURE 3: Diagram of the *FOXE1* region marking some single nucleotide polymorphisms (SNPs) reported. The 5' and 3' untranslated regions (UTR) are marked in a blue box, the *FOXE1* gene exon is marked in an orange box, the forkhead region is marked in a red box, and the SNPs present in our patient are marked in italic.

60°C for 20 s, and extension at 72°C for 30 s, followed by a final extension at 72°C for 7 min. The runs were carried out on a GeneAmp PCR system 9700 (Applied Biosystems). PCR products were sequenced directly with a DNA Sequencing Kit with Big Dye Terminator on an automated ABI 3730XL sequencer (Applied Biosystems, Foster City, CA, USA) and data analyses were performed by Lasergene software (DNASTAR Inc.). Neither mutations nor SNPs were found in *FOXE1* coding region. Two SNPs (rs1867279 and rs1867280) that have already been associated with CP were found in a homozygote wild type genotype. Additionally, we found the rs71369530 (Ala-14) in homozygote fashion. Functional prediction analysis using the *SNPinfo* Web Server (<https://snpinf.niehs.nih.gov/snpinf/snpfunc.html>) showed that only the ancestral alleles of rs1867279 and rs1867280 might recruit the E2F and NFκB transcription factors, respectively.

3. Discussion

3.1. Generalities. Several studies in genetics have shown that HT, TC, and isolated CL and CP are related to the *FOXE1* locus. *FOXE1* (Forkhead box E1; UniGene accession number Hs.159234) is an intron-less single exon gene that encodes transcription factor FOXE1 (or TTF-2). FOXE1 is a 373 amino acid protein (38 kDa) that contains a forkhead winged helix DNA binding domain and a polyalanine (polyAla) tract of variable length (11–19 Ala, but 14-Ala is most abundant) and negatively regulates thyroglobulin (TG) and Thyroid Peroxidase (TPO) expression, which regulates, in conjunction with

NKX2-1 (Homeobox protein or TTF-1), PAX8, and HHEX, thyroid differentiation and function [9, 17].

RNA in situ hybridization studies in zebrafish in the fish ortholog of *FOXE1*, *foxel*, indicated that during the postfertilization development, the gene is early expressed in the nervous system, afterward in the oral epithelium and thyroid gland, later in the developing heart, and, in the final stages, in the pharyngeal arch epithelium [5]. Moreover, genome-wide association studies of genotypes of electronic medical records in DNA biobanks have associated the *FOXE1* locus with thyroiditis, nodular and multinodular goiters, and thyrotoxicosis, among others [18], indicating that mutations in this gene and its promoters exert an important effect on normal thyroid function.

3.2. Orofacial Clefts. Orofacial clefts (OC) comprise clefts, disruption, or gaps in structures such as lips, palate, eyes, and nose and arise from failure of normal craniofacial development, in which cells grow, migrate, and differentiate in a coordinated fashion to fuse together the lateral, medial, and maxillary processes. Moreover, given the different developmental origins of lip and palate, these are classified in CL with or without CP [19].

Isolated CL and CP are common genetic defects with a strong genetic component that has not been completely elucidated, and approximately 70–80% of CL/CP that are nonsyndromic (absence of additional structural or cognitive abnormalities) represent complex human disorders with the interaction of genetic risk factors and environmental exposures that increase susceptibility [19]. Evidence suggests

that rare DNA variations in several genes and SNPs (many of these are present in noncoding DNA) are linked to CL and CP etiology [3, 8].

3.3. Relation of *FOXE1* with Orofacial Clefts. Genome-wide association studies (GWAS) have related *FOXE1* with CL and CP in different population, within the region 9q22-33 [3, 5, 8]; however the SNP and haplotype frequencies found within each population indicate multiple risk alleles for such a complex disease, due in part to maternal or paternal overtransmission [3]. Moreover, as its role with the expression of TG and TPO suggests, mutations at the *FOXE1* locus in 9q22 have been associated with thyroid disorders such as congenital HT due to thyroid dysgenesis, primary HT, goiters, nonmedullary thyroid cancer, PTC, and TC [5, 8].

FOXE1 is responsible for palatogenesis, and it appears that casual mutations in noncoding regions that regulate *FOXE1* expression are related to CL and CP [3], such as SNPs rs7850258, rs12342417, and rs10984103 within the *FOXE1* locus that have been associated with high risk for orofacial clefting [8].

Few studies have been carried out on Mexican population or in persons with American-Indian ethnicity; however, research has shown that SNP rs4460498 and rs375829, located in the *FOXE1* locus, are associated with nonsyndromic CP and CL for European and Mayan-Mesoamerican descendants [7]. This study obtained important significance in high risk of CL/CP for rs4460498 $p_{\text{Europe}} = 6.50 \times 10^{-6}$ and $p_{\text{Mayan}} = 0.051$, and for rs375829 $p_{\text{Europe}} = 2.41 \times 10^{-5}$ and $p_{\text{Mayan}} = 0.0299$.

3.4. Relation of *FOXE1* with Hypothyroidism. HT is the most common thyroid disorder characterized by deficiencies of thyroid hormones T3 and T4 (triiodothyronine and thyroxine, resp.), which are regulators of metabolism and development. A common marker of HT is the high levels of Thyroid-Stimulating Hormone (TSH), which indicate impaired thyroid function; however, in other cases, reduced levels of TSH cause low levels of T3 and T4. Iodine deficiency is the most common cause of HT, but congenital HT and autoimmune HT are other varieties with different genetic causes [20] that cause central HT, primary HT, alterations in thyroid transcription factors (with which the *FOXE1* gene is associated), dyshormonogenesis, and other genes such as the *TG* and *TSH* genes [19].

A mutation in the coding sequence of *FOXE1* that shifts Arginine-102 to a Cysteine (R102C), a highly conserved residue within the forkhead domain of the protein, impairs *FOXE1* from binding DNA; thus, it is inactive toward transcription. This inactivity leads to thyroid dysgenesis in the patient, which causes congenital HT, CP, spikey hair, and bilateral choanal atresia. Interestingly, because athyreosis is not present, human thyroid formation and development do not depend entirely on *FOXE1* activity [21].

SNPs that are present in other autoimmune diseases, such as rs6679677 near *PTPN22*, rs3184504 in *SH2B3*, and rs2517532 in HLA class I region, are linked to HT, while

rs4915077 near *VAV3* [20] and other SNPs such as rs7850258, rs965513, rs925489, and rs10759944, all located near *FOXE1* in chromosome 9, exhibit genome-wide significance with primary HT [18]. Other SNPs in chromosome 9 that appear to be associated with primary HT are rs4979402, rs4979397, rs1408528, and rs1535971, located in *DFNB31* that is 16.6 Mb from *FOXE1* and not in LD with other SNP of *FOXE1*; however, these later regions do not correlate when TSH is not expressed [18].

3.5. Relation of *FOXE1* with Thyroid Cancer. TC is a complex polygenic disorder and the networking of the molecular mechanisms involved in its development, the genes involved in the process, and the risk environmental factors remain to be understood. It has been proposed that TC is the result of multiple low-to moderate penetrance genes interacting with each other, acting together with other transcriptional regulators that produce loss- or gain-of-function and that, together with environmental and other genetic risk factors, can lead to the development of cancer [6, 14, 21]. However, this explains only a small portion of the incidence, as there are other genetic factors involved. TC types, such as FTC, Differentiated Thyroid Cancer (DTC, which includes PTC and MTC), and undifferentiated anaplastic thyroid carcinomas, have different biomarkers and cellular conformations and thus different molecular pathways involved in their development.

rs7850258, located in the *FOXE1* promoter region near craniofacial enhancer hscCNE-67, has direct involvement in the development of thyroid and heart, and is also associated with HT and TC, as it was found to alter the enhancer function in both oral epithelial and thyroid cell lines. In this SNP, allele G is associated with HT and CL, whereas allele A is associated with TC [5].

Studies in several countries [6, 9–11, 14–16] have revealed that *FOXE1* locus polymorphisms in chromosome 9q22.33, such as rs965513 and polymorphisms (rs7849497, rs1867277, rs1867278, rs1867279, and rs1867280) in the promoter regions for this factor are associated with an increased risk of non-medullary thyroid cancer, differentiated thyroid cancer, and the Bamforth–Lazarus syndrome, in which the mutation of a serine-57 to asparagine (S57N) in the forkhead-DNA binding domain of *FOXE1* leads to CP, choanal atresia, bifid epiglottis, thyroid agenesis or dysgenesis, hypothyroidism, and spikey hair [22]. Although there is no clear evidence of the role of the rs1867279 and rs1867280 in CL, HT, and PTC pathogenesis, functional prediction analysis showed that both SNPs could modify the transcription factor binding site and thus modulate the expression of relevant genes in the thyroid morphogenesis. Otherwise, the rs1867279 and rs1867280 might be in LD with the true casual variant, such as rs1867277, because a high LD has been described in the *FOXE1* region. On this regard, it has been reported that the rs1867277A allele recruits the USF1/USF2 transcription factors and has been shown to be involved in an allele-specific transcriptional regulation of *FOXE1* in thyroid cancer [6]. Actually, the rs1867277 have been found related to TC in several populations [21, 23, 24].

The information about the association of the *FOXE1* SNP rs1867277 with TC appears to vary with ethnicity and cancer type [16, 23]. For example Bychkov et al. [12] found overexpression of rs1867277 and rs965513 by immunohistochemical staining in PTC in Japanese persons. On the contrary, neither Maillard et al. [14] in French Polynesians nor Bonora et al. [25] in European descendants found an association of this SNP with nonmedullary thyroid cancer.

SNP rs1867277, within the $-283G>A$ promoter region, was assessed functionally by Landa et al. [6] and their results suggest that this SNP increases the transcriptional activity of the *FOXE1* gene promoter by recruitment of leucine zipper USF 1 and 2. Thus, it is likely that recruiting USF1 and USF2 proteins produces a deregulation of *FOXE1* that leads to malignant behavior of thyroid cells [9].

It has been demonstrated that, through a number of signalling pathways, *FOXE1*, along with the remainder of the family of FOX transcription factors, can induce *WNT5A*, a gene (located in 3p14.3) expressed in a variety of human tumors [26]. Thus, by deregulation of *FOXE1*, it is likely that this protein can become a tumor suppressor or an oncogene through the Wnt pathway, either dependently or independently of β -catenin [17]. Also, there is evidence that the *RET/PTC3* protooncogene that causes TC decreases *FOXE1* expression. Similarly, in squamous cell carcinoma, the 9q22 region of the chromosome (where *FOXE1* is localized) is lost and, in other cancers, hypermethylation of promoter is also observed [5].

Levels of *FOXE1* and *PAX8* messenger RNA (mRNA) in thyroid nodules have been found to be decreased in malignant thyroid lesions and higher in benign lesions, and even the gene expression patterns of *FOXE1* and *PAX8* can help in differentiating FTC from MTC and the Follicular Variant of Papillary Thyroid Carcinoma (FVPTC) [27].

Because we did not sequence the entire gene, our analysis did not exclude the hypothesis that *FOXE1* mutations contribute to the development of these diseases. Numerous evidences have been published suggesting that this gene is involved in these three diseases. In the same haplotype analysis of the members of a Portuguese family with non-medullary thyroid carcinoma, the selected individuals with the disease not only shared the *FOXE1* p.A248G variant, but also the same haplotype between markers rs965512 and D99S180 (located in chromosome 9), including rs1867279 and rs1867280 (SNPs present in our patient). However, these persons did not exhibit an association with any of the other polymorphic loci with the cancer because, as according to the authors and as reported in a previous paper, 15% of Portuguese control population have these SNP [17]. However, we propose that ethnicity and environmental factors are very important in triggering a disruptive response in rs1867279 and rs1867280, as in the case of our patient, in whom these SNPs are acting together, along with other risk factors (ethnic and environmental), to produce CL, HT, and TC.

Given the puzzling findings reported with rs1867277 [12, 16, 22] in *FOXE1* or rs944289 in *NKX2* [15], it is obvious that ethnicity and environmental factors play a key role for the development of TC and other malignancies. Additionally, the

involvement of *FOXE1* in thyroid formation and function, as well as the formation of craniofacial structures, relates mutations in the noncoding region of *FOXE1* to the phenotype of our patient.

Evidence such as the Bamforth–Lazarus syndrome [22], or mutations in *FOXE1* that explain an association of hypothyroidism and cleft palate [28], in addition to observational and comparative data of human and animal models null for *Foxe1* [5, 29] that are consistent with CP, CL, and other thyroid abnormalities, suggests that mutations or SNP in *FOXE1* can be a common cause for a possible new syndrome characterized by CL, HT, and PTC. In this line, Tables 2 and 3 illustrate some of the most studied SNPs and genes probably associated with these diseases.

Seeking an explanation for the possible identification of the syndrome that has been discussed, the *FOXE1* mutation has been tested, being negative as far as could be analysed. A possible candidate gene in this situation is E-cadherin *CDH1* on 16q22.1, which is also expressed in critical stages of lip and palate development in embryos [30, 31]. Hence, there are many more genetic alterations that could explain this possible association.

4. Conclusions

The association of CL, HT, and PTC might be component of a new syndrome; however *FOXE1* coding region, which has been involved with these entities, has not exhibited mutations or SNPs. Based on this case study and a review of the literature on the role that the *FOXE1* gene plays in these pathologies, a more extensive sequence analysis (including other genes) is needed to rule out its role in the above-mentioned diseases.

Ethical Approval

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Consent

Informed consent was obtained from all patients for being included in the study.

Conflicts of Interest

All authors declare that there are no conflicts of interest.

Authors' Contributions

All authors made substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work, drafting the work, or revising it critically for important intellectual content, gave final approval of the version to be published, and agreed to be accountable for all aspects of the work in ensuring that

TABLE 2: SNPs found in *FOXE1* related to orofacial clefts: cleft lip (CL) and cleft palate (CP), papillary thyroid cancer, and hypothyroidism.

Pathology	SNP in <i>FOXE1</i> or in its promoter regions	Sample/population
Orofacial clefts CL/P	rs4460498	Americans, Colombians, Danish, Filipinos, Norwegian [3].
	rs894673	Americans, Colombians, Filipinos [3].
	rs3758249	African-Brazilians, Americans, Central-Europeans, Colombians, Danish, Filipinos, Mayan-Mesoamerican, Norwegian [3, 4, 7].
	rs1867278	Americans, Filipinos, Norwegians, Danish [3, 4].
	rs1867280	Americans, Colombians, Danish, Filipinos, Norwegian [3].
	rs7850258	Human fetal oral epithelial thyroid cell line, Caucasian Europeans, Hondurans [3, 5].
	rs12342417	Europeans [5, 7].
	rs10984103	Europeans, Filipinos [8].
	rs4460498	Central-Europeans, Mayan-Mesoamerican [8].
	rs1867277	Caucasian Australians, Italians, Japanese, Portuguese, Spanish, Turkish [6, 9–12].
rs7850258	Rat FRTL epithelial thyroid cell line, zebra fish, mouse [5].	
rs965513	Colombians, European descents, Germans, Icelandics, Japanese, Polynesians, Portuguese [12–16].	
rs894673	Turkish [10].	
rs3758249	Caucasian Australians, Turkish [10, 11].	
rs907577	Caucasian Australians [11].	
rs3021526	Caucasian Australians [11].	
rs1443434	Caucasian Australians [11].	
rs907580	Caucasian Australians [11].	
Rs7849497	Portuguese [13].	
Rs1867278	Portuguese [13].	
Rs1867279	Portuguese [13].	
Rs1867280	Portuguese [13].	
rs7850258	Rat FRTL epithelial thyroid cell line, zebra fish, mouse, European [5, 17].	
rs965513	Europeans [17].	
rs925489	Europeans [17].	
rs10759944	Europeans [17].	
Papillary thyroid cancer		
Hypothyroidism		

TABLE 3: Genes related to orofacial clefts: cleft lip (CL) and cleft palate (CP), papillary thyroid cancer, or hypothyroidism.

Pathology	Genes
Orofacial clefts CL/P	<i>BMP4</i>
	<i>CRISPLD2</i>
	<i>FGF8</i>
	<i>FGFR1</i>
	<i>FGFR2</i>
	<i>FOXE1</i>
	<i>GLI2</i>
	<i>GSTM1</i>
	<i>GSTT1</i>
	<i>JAG2</i>
	<i>LHX8</i>
	<i>MSX1</i>
	<i>MSX1</i>
	<i>MSX2</i>
	<i>NAT1</i>
	<i>NAT2</i>
	<i>PTCH</i>
	<i>PVRL1</i>
	<i>RYK</i>
	<i>SATB2</i>
Papillary thyroid cancer	<i>SKI</i>
	<i>SPRY2</i>
	<i>TBX10</i>
	<i>TGFB3</i>
	<i>ATM</i>
Hypothyroidism	<i>NKX2</i>
	<i>RET</i> (RET/PTC rearrangement)
	<i>DFNB31</i>
	<i>PTPN22</i>
	<i>SH2B3</i>
	<i>VAV3</i>

questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgments

Funding was provided by Ciprés Grupo Médico S.C. (CGM) and by the Instituto Nacional de Medicina Genómica (INMEGEN).

References

- [1] L. P. Fernández, A. López-Márquez, Á. M. Martínez, G. Gómez-López, and P. Santisteban, “New Insights into FoxE1 Functions: Identification of Direct FoxE1 Targets in Thyroid Cells,” *PLoS ONE*, vol. 8, no. 5, Article ID e62849, 2013.
- [2] M. L. Marazita, J. C. Murray, A. C. Lidral et al., “Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35,” *American Journal of Human Genetics*, vol. 75, no. 2, pp. 161–173, 2004.
- [3] L. M. Moreno, M. A. Mansilla, S. A. Bullard et al., “FOXE1 association with both isolated cleft lip with or without cleft palate, and isolated cleft palate,” *Human Molecular Genetics*, vol. 18, no. 24, pp. 4879–4896, 2009.
- [4] A. do Rego Borges, J. Sá, R. Hoshi et al., “Genetic risk factors for nonsyndromic cleft lip with or without cleft palate in a Brazilian population with high African ancestry,” *American Journal of Medical Genetics, Part A*, vol. 167, no. 10, pp. 2344–2349, 2015.
- [5] A. C. Lidral, H. Liu, S. A. Bullard et al., “A single nucleotide polymorphism associated with isolated cleft lip and palate, thyroid cancer and hypothyroidism alters the activity of an oral epithelium and thyroid enhancer near *FOXE1*,” *Human Molecular Genetics*, vol. 24, no. 14, Article ID ddv047, pp. 3895–3907, 2015.
- [6] I. Landa, S. Ruiz-Llorente, C. Montero-Conde et al., “The variant rs1867277 in *FOXE1* gene confers thyroid cancer susceptibility through the recruitment of USF1/USF2 transcription factors,” *PLoS Genetics*, vol. 5, no. 9, Article ID e1000637, p. e1000637, 2009.
- [7] K. U. Ludwig, A. C. Böhmer, M. Rubini et al., “Strong association of variants around *FOXE1* and orofacial clefting,” *Journal of Dental Research*, vol. 93, no. 4, pp. 376–381, 2014.
- [8] K. U. Ludwig, A. C. Böhmer, J. Bowes et al., “Imputation of orofacial clefting data identifies novel risk loci and sheds light on the genetic background of cleft lip ± cleft palate and cleft palate only,” *Human Molecular Genetics*, vol. 26, no. 4, pp. 829–842, 2017.
- [9] R. A. Tomaz, I. Sousa, J. G. Silva et al., “*FOXE1* polymorphisms are associated with familial and sporadic nonmedullary thyroid cancer susceptibility,” *Clinical Endocrinology*, vol. 77, no. 6, pp. 926–933, 2012.
- [10] E. Somuncu, A. Karatas, S. Ferahman et al., “The investigation of foxe1 variations in papillary thyroid carcinoma,” *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 10, pp. 13458–13464, 2015.
- [11] M. Bullock, E. L. Duncan, C. O’Neill et al., “Association of *FOXE1* polyalanine repeat region with papillary thyroid cancer,” *Journal of Clinical Endocrinology and Metabolism*, vol. 97, no. 9, pp. E1814–E1819, 2012.
- [12] A. Bychkov, V. Saenko, M. Nakashima et al., “Patterns of *FOXE1* expression in papillary thyroid carcinoma by immunohistochemistry,” *Thyroid*, vol. 23, no. 7, pp. 817–828, 2013.
- [13] A. Chrisoulidou, M. Boudina, A. Tzemailas et al., “Histological subtype is the most important determinant of survival in metastatic papillary thyroid cancer,” *Thyroid Research*, vol. 4, no. 1, article no. 12, 2011.
- [14] S. Maillard, F. Damiola, E. Clero et al., “Common variants at 9q22.33, 14q13.3, and *ATM* loci, and risk of differentiated thyroid cancer in the French Polynesian population,” *PLoS ONE*, vol. 10, no. 4, Article ID e0123700, 2015.
- [15] M. Penna-Martinez, F. Epp, H. Kahles et al., “*FOXE1* association with differentiated thyroid cancer and its progression,” *Thyroid*, vol. 24, no. 5, pp. 845–851, 2014.
- [16] A. P. Estrada-Florez, M. E. Bohórquez, R. Sahasrabudhe et al., “Clinical features of Hispanic thyroid cancer cases and the role of known genetic variants on disease risk,” *Medicine (United States)*, vol. 95, no. 32, Article ID e4148, 2016.
- [17] J. S. Pereira, J. G. da Silva, R. A. Tomaz et al., “Identification of a novel germline *FOXE1* variant in patients with familial non-medullary thyroid carcinoma (FNMTc),” *Endocrine*, vol. 49, no. 1, pp. 204–214, 2015.
- [18] J. C. Denny, D. C. Crawford, M. D. Ritchie et al., “Variants near *FOXE1* are associated with hypothyroidism and other thyroid conditions: Using electronic medical records for genome- and phenome-wide studies,” *American Journal of Human Genetics*, vol. 89, no. 4, pp. 529–542, 2011.

- [19] S. Lesage, V. Drouet, E. Majounie et al., "Loss of VPS13C function in autosomal-recessive parkinsonism causes mitochondrial dysfunction and increases pink1/parkin-dependent mitophagy," *The American Journal of Human Genetics*, vol. 98, no. 3, pp. 500–513, 2016.
- [20] N. Eriksson, J. Y. Tung, A. K. Kiefer et al., "Novel associations for hypothyroidism include known autoimmune risk loci," *PLoS ONE*, vol. 7, no. 4, Article ID e34442, 2012.
- [21] I. Bariş, A. E. Arisoy, A. Smith et al., "A novel missense mutation in human TTF-2 (FKHL15) gene associated with congenital hypothyroidism but not athyreosis," *Journal of Clinical Endocrinology and Metabolism*, vol. 91, no. 10, pp. 4183–4187, 2006.
- [22] A. Carré, R. T. Hamza, D. Kariyawasam et al., "A novel FOXE1 mutation (R73S) in bamforth-lazarus syndrome causing increased thyroidal gene expression," *Thyroid*, vol. 24, no. 4, pp. 649–654, 2014.
- [23] A. Carré, M. Castanet, S. Sura-Trueba et al., "Polymorphic length of FOXE1 alanine stretch: Evidence for genetic susceptibility to thyroid dysgenesis," *Human Genetics*, vol. 122, no. 5, pp. 467–476, 2007.
- [24] C. M. Pereda, F. Lesueur, M. Pertesi et al., "Common variants at the 9q22.33, 14q13.3 and ATM loci, and risk of differentiated thyroid cancer in the Cuban population," *BMC Genetics*, vol. 16, no. 1, 2015.
- [25] E. Bonora, C. Rizzato, C. Diquigiovanni et al., "The FOXE1 locus is a major genetic determinant for familial nonmedullary thyroid carcinoma," *International Journal of Cancer*, vol. 134, no. 9, pp. 2098–2107, 2014.
- [26] M. Katoh and M. Katoh, "Transcriptional mechanisms of WNT5A based on NF- κ B, Hedgehog, TGF β , and Notch signaling cascades," *International Journal of Molecular Medicine*, vol. 23, no. 6, pp. 763–769, 2009.
- [27] A. E. Walts, J. M. Mirocha, and S. Bose, "Follicular variant of papillary thyroid carcinoma (FVPTC): histological features, BRAF V600E mutation, and lymph node status," *Journal of Cancer Research and Clinical Oncology*, 2015.
- [28] M. Castanet, S.-M. Park, A. Smith et al., "A novel loss-of-function mutation in TTF-2 is associated with congenital hypothyroidism, thyroid agenesis and cleft palate," *Human Molecular Genetics*, vol. 11, no. 17, pp. 2051–2059, 2002.
- [29] M. De Felice, C. Ovitt, E. Biffali et al., "A mouse model for hereditary thyroid dysgenesis and cleft palate," *Nature Genetics*, vol. 19, no. 4, pp. 395–398, 1998.
- [30] T. Frebourg, C. Oliveira, P. Hochain et al., "Cleft lip/palate and CDH1/E-cadherin mutations in families with hereditary diffuse gastric cancer," *Journal of Medical Genetics*, vol. 43, no. 2, pp. 138–142, 2006.
- [31] P. R. Benusiglio, O. Caron, E. Consolino et al., "Cleft lip, cleft palate, hereditary diffuse gastric cancer and germline mutations in CDH1," *International Journal of Cancer*, vol. 132, no. 10, p. 2470, 2013.