



Mitochondrial DNA Mutation Analysis in Breast Cancer: Shifting From Germline Heteroplasmy Toward Homoplasmy in Tumors

Carlos Jhovani Pérez-Amado^{1,2}, Hugo Tovar³, Laura Gómez-Romero³, Fredy Omar Beltrán-Anaya⁴, Verónica Bautista-Piña⁵, Carlos Domínguez-Reyes⁵, Felipe Villegas-Carlos⁵, Alberto Tenorio-Torres⁵, Luis Alberto Alfaro-Ruiz¹, Alfredo Hidalgo-Miranda^{1*} and Silvia Jiménez-Morales^{1*}

¹ Laboratorio de Genómica del Cáncer, Instituto Nacional de Medicina Genómica, Mexico City, Mexico, ² Programa de Doctorado, Posgrado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México, Mexico City, Mexico, ³ Genómica Computacional, Instituto Nacional de Medicina Genómica, Mexico City, Mexico, ⁴ Laboratorio de Investigación en Epidemiología Clínica y Molecular, Facultad de Ciencias Químico-Biológicas, Universidad Autónoma de Guerrero, Chilpancingo, Mexico, ⁵ Instituto de Enfermedades de la Mama, FUCAM, Mexico City, Mexico

OPEN ACCESS

Edited by:

Obul Reddy Bandapalli,
Hopp Children's Cancer Center
Heidelberg (KITZ), Germany

Reviewed by:

Pawel Buczkowicz,
PhenoTips, Canada
Kiran Gona,
Yale Medicine, United States

*Correspondence:

Alfredo Hidalgo-Miranda
ahidalgo@inmegen.gob.mx
Silvia Jiménez-Morales
sjimenez@inmegen.gob.mx

Specialty section:

This article was submitted to
Cancer Genetics,
a section of the journal
Frontiers in Oncology

Received: 15 June 2020

Accepted: 18 September 2020

Published: 27 October 2020

Citation:

Pérez-Amado CJ, Tovar H, Gómez-Romero L, Beltrán-Anaya FO, Bautista-Piña V, Domínguez-Reyes C, Villegas-Carlos F, Tenorio-Torres A, Alfaro-Ruiz LA, Hidalgo-Miranda A and Jiménez-Morales S (2020) Mitochondrial DNA Mutation Analysis in Breast Cancer: Shifting From Germline Heteroplasmy Toward Homoplasmy in Tumors. *Front. Oncol.* 10:572954. doi: 10.3389/fonc.2020.572954

Studies have suggested a potential role of somatic mitochondrial mutations in cancer development. To analyze the landscape of somatic mitochondrial mutation in breast cancer and to determine whether mitochondrial DNA (mtDNA) mutational burden is correlated with overall survival (OS), we sequenced whole mtDNA from 92 matched-paired primary breast tumors and peripheral blood. A total of 324 germline variants and 173 somatic mutations were found in the tumors. The most common germline allele was 663G (12S), showing lower heteroplasmy levels in peripheral blood lymphocytes than in their matched tumors, even reaching homoplasmic status in several cases. The heteroplasmy load was higher in tumors than in their paired normal tissues. Somatic mtDNA mutations were found in 73.9% of breast tumors; 59% of these mutations were located in the coding region (66.7% non-synonymous and 33.3% synonymous). Although the *CO1* gene presented the highest number of mutations, tRNA genes (T, C, and W), rRNA 12S, and *CO1* and *ATP6* exhibited the highest mutation rates. No specific mtDNA mutational profile was associated with molecular subtypes of breast cancer, and we found no correlation between mtDNA mutational burden and OS. Future investigations will provide insight into the molecular mechanisms through which mtDNA mutations and heteroplasmy shifting contribute to breast cancer development.

Keywords: mitochondrial DNA, breast cancer, mutations, heteroplasmy, haplogroups, molecular subtypes

Abbreviations: AKT, AKT serine/threonine kinase; ATP6, ATP synthase membrane subunit 6; ATP8, ATP synthase membrane subunit 8; BRCA, BRCA DNA repair associated; BWA, Burrows-Wheeler aligner; CN, copy number; CO1, cytochrome C oxidase subunit 1; CO2, cytochrome C oxidase subunit 2; CO3, cytochrome C oxidase subunit 3; CYB, cytochrome B; ER, estrogen receptor; GATK, genome analysis tool kit; H2, HER2 positive breast cancer subtype; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; IHC, immunohistochemistry; LA, luminal A breast cancer subtype; LB, luminal B breast cancer subtype; MAF, mutant allele fraction; mtDNA, mitochondrial DNA; mtDNA-CN, mitochondrial DNA copy number; ND (1–6), NADH dehydrogenase subunits (1–6); ND4L, NADH dehydrogenase subunits 4L; OXPHOS, oxidative phosphorylation; PR, progesterone receptor; ROS, reactive oxygen species; rRNA, ribosomal RNA; SNV, single-nucleotide variant; TN, triple negative; tRNA, transfer RNA; VEP, variant effect predictor.

INTRODUCTION

Breast cancer is the most common cancer type in women worldwide (1); in 2018, 2.09 million cases and 627,000 deaths were estimated around the world (2). In Mexico, this malignancy is currently the leading cause of cancer-related deaths in women (3–5). Breast cancer is a heterogeneous disease in terms of the molecular features and clinical outcome; knowledge of the genetic causes is still incomplete. To comprehensively address this issue, the analysis of nuclear genes has been frequently done (4, 6–10). However, mutations in mitochondrial genes have also been reported to play important roles in human cancer development (11). Experimental evidence in mouse models demonstrated that G13997A and 13885insC mutations in the reduced form of the nicotinamide adenine dinucleotide dehydrogenase subunit 6 (*ND6*) gene are associated with the transition of poor metastatic tumor cells into a highly metastatic phenotype (12).

Mitochondria are indispensable for numerous cellular processes, such as respiratory energy metabolism, cell proliferation, apoptosis, senescence, and immune response (13). Mitochondrial DNA (mtDNA) is a circular DNA molecule of 16,569 bp containing 2 rRNA, 22 tRNA, and 13 genes encoding protein subunits for the mitochondrial complexes of the oxidative phosphorylation system (OXPHOS). The mitochondrial genome is characterized by many DNA molecules per cell and a high mutation rate that creates a state known as heteroplasmy (a mixture of mutant and normal mtDNA in cells) (14, 15). The relative abundance of certain *de novo* mutations might be pathogenic, although the heteroplasmic threshold effect and the molecular mechanisms underlying the mutation selection during disease development have not yet been established (15). However, the association among several mtDNA germline variants and a broad spectrum of human malignancies has been widely documented (16–18). For instance, the T16189C (D-Loop) and G10398A (A114T at *ND3* gene) germline variants have been commonly associated with an increased risk of cancer (17, 19–21). The important role of germline mtDNA variants in tumors has been noted in groups of variants inherited without any recombination, called haplogroups, conferring risk for cancer and as modifiers for potential metastasis and for the response to treatment in affected patients (22).

Genomic evidence suggests that mtDNA germline variants and tumor mutations are also involved in breast cancer development. mtDNA mutations have been reported in almost 60% of breast tumors, varying in nature (mainly single base substitutions and indels) and heteroplasmy levels (23–26); however, whether these mutations are drivers or passengers remains largely uncharacterized (27, 28). mtDNA sequencing from matched normal–tumor breast tissues has shown that a high proportion of the somatic mutations are singletons arising in a single patient and that many of those were under clonal expansion in the primary or metastatic tumors (17, 22, 28–33). To explore the landscape of mtDNA mutations in breast cancer from Mexican women and to know if mtDNA mutations and haplogroups are associated with the molecular subtypes of breast tumors or with overall survival (OS), we performed whole

mtDNA sequencing in matched-paired primary breast tumors and peripheral blood.

MATERIALS AND METHODS

Patient Data

We included matched tumor samples and peripheral blood from 92 women with breast cancer. Patients were recruited from the Instituto de Enfermedades de la Mama FUCAM (Mexico City, Mexico), and all patients provided written informed consent. Patients with chemotherapy, radiotherapy, and neoadjuvant treatment before the surgery were excluded. Based on the current international criteria and using immunohistochemical (IHC) markers, breast tumors were classified as luminal A (LA), luminal B (LB), HER2-positive (H2), and triple negative (TN) (34). Patients were followed up for at least 30 months (30–124) after the clinical diagnosis.

Sample Processing and DNA Extraction

Peripheral blood was obtained before surgical treatment using EDTA vacutainer tubes, and the buffy coat (leukocytes) was separated out by centrifugation at $1850 \times g$. Fresh tumor tissues were obtained during surgical treatment and stored at -80°C until further analysis. Breast tumor samples with $>80\%$ of tumor cells were included. Total DNA from leukocytes and tumor samples were extracted using a Maxi Kit (Qiagen, Hilden, Germany) and AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany), respectively, according to the manufacturer's instructions. The nucleic acids were quantified by spectrophotometry (NanoDrop, Thermo Fisher Scientific, Waltham, MA, United States) and stored at -20°C until library preparation.

Whole mtDNA Sequencing

To avoid contamination with nuclear mitochondrial segments, the whole mtDNA was enriched using long PCR with two overlapping primer sets (**Supplementary Table 1**) and subsequently fragmented by sonication. Indexed paired-end mtDNA libraries were generated with the Nextera DNA Flex Library Prep Kit (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. Libraries were sequenced on an Illumina MiSeq using 150-bp paired-end read chemistry and 300 cycles (MiSeq Reagent Kit v2, Illumina, San Diego, CA, United States).

Bioinformatic Analysis

We included mitochondrial genome data with depth coverage with at least 300X *per* base. A quality control (QC) evaluation was performed on the raw sequence data using the FastQC algorithm (35), and reads with Phred quality >25 were accepted for further analysis. The Burrows–Wheeler alignment (BWA-MEM) algorithm was used to align all sequences against the revised Cambridge Reference Sequence (GRCh38, GCF_000001405.38) (36). Duplicated reads on mtDNA were removed. Then, reads were realigned and quality scores were recalibrated using Genome Analysis Toolkit (GATK) (37). Haplogroups were

assigned on the mtDNA variants detected in blood samples and using Haplogrep software (v2.1.19) and Phylotree databases (38, 39). The Strelka2 algorithm with default parameters (40) was used to identify somatic mutations by comparing germline and matched mtDNA variants. Only variants labeled PASS were considered high-quality variants (40). To reduce false positives, a somatic mutation was considered when the mutant allele fraction (MAF) was <0.001 in normal tissue and >0.01 in tumor tissue in all samples where it occurred. Germline variants were defined if $MAF > 0.01$ in the blood sample regardless of the proportion in tumors and if this occurred in at least one patient. Variants not fulfilling these criteria were considered as probable mutations. These thresholds were estimated based on the depth coverage. All variants displaying mutant allele abundance from 1% to $<95\%$ were deemed heteroplasmic. Homoplasmic was defined as a mutant allele $>95\%$ (32). Mutations were reviewed in the MITOMAP database to determine if they had been reported in breast cancer or in other malignancies (41). To predict their role as a driver or passenger, we used the Variant Effect Predictor (VEP) software with the default parameters (42). To assess whether an mtDNA profile was associated with the tumor subtype, we performed an unsupervised hierarchical clustering analysis accounting for all mtDNA mutations using R (version 3.6.0). The distribution of mtDNA somatic mutations among breast cancer subtypes was visualized using Venn diagrams (43).

Statistical Analysis

Mutations were counted according to gene distribution, and frequencies were obtained for all analyzed patients. To assess the association between mtDNA mutations and haplogroups and clinical and molecular characteristics among groups, chi-square and Fisher's exact tests were calculated when appropriate. Mann-Whitney U tests were used to evaluate differences in heteroplasmy levels between the blood and tumor tissues. An OS analysis was performed, including patients who were followed up to 124 months after diagnosis and initial treatment, defining OS as the time between initial diagnosis and death. mtDNA mutational burden was stratified into two groups: patients without any mutation and patients with mutations, as well as according to the mutational burden mean (patients with a low or high mutational burden). To evaluate whether somatic mtDNA mutational burden is associated with patient outcome, Kaplan-Meier survival curves were obtained and the hazard risk was estimated using a Cox proportional hazard model adjusting for variables that have been reported to influence the OS (age, tumor stage, clinical stage, and hormone receptor status). P -values < 0.05 were considered statistically significant. Statistical analysis was conducted using SPSS (Version 25, IBM Corp., Armonk, NY, United States). Plots were generated using R (version 3.6.0).

RESULTS

Study Population

We included 92 unrelated patients with a mean age of 53.83 ± 11.5 years (range 33 to 92 years; **Supplementary Table 2**).

The tumor subtype was assigned by IHC markers in 90 patients. Luminal A (63.3%) was the most frequent subtype, followed by LB (24.4%), TN (6.7%), and H2 (5.6%) subtypes.

Distribution of Haplogroups

From 564 germline variants identified in the normal tissues, 296 were used by Haplogrep for haplogroup typing. After haplotyping the mtDNA of all patients, haplogroup A was the most frequent (44.6%), followed by haplogroups B (22.8%), C (11.9%), D (12%), and L (5.4%). Haplogroups H and J were found in 2.2 and 1.1% of patients, respectively. Stratification analysis by IHC subtypes revealed that haplogroups C and D were present only in LA and LB subtypes, and TN tumors carried haplogroups A and B. No statistically significant differences in the haplogroup distribution was found among the tumor subtypes ($P > 0.05$).

Mitochondrial DNA Variant Distribution

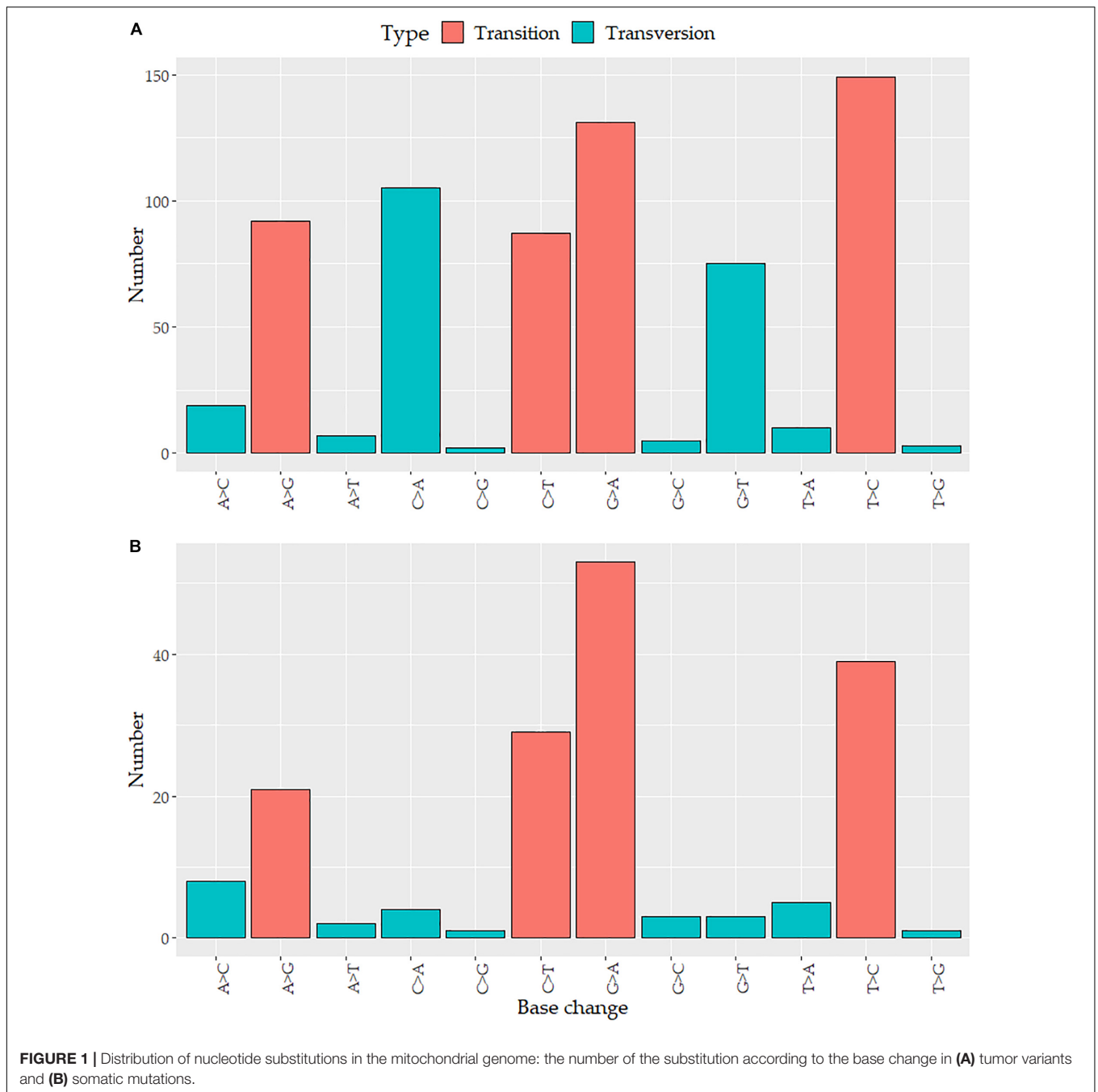
The depth coverage of mtDNA for the examined samples was from $300\times$ to $6000\times$ per base (**Supplementary Figure 1**). 3404 variants carried by normal tissue (located in 564 positions across the mtDNA) and 3876 variants identified in tumor samples passed all quality controls. Blood samples ranged from 16 to 75 variants ($\bar{X} = 37 \pm 9$), whereas tumor samples carried from 24 to 89 variants ($\bar{X} = 42 \pm 11$).

Variants detected in breast tumor tissues were located in 709 positions across the mtDNA, 685 (96.6%) of these variants were single-nucleotide variants (SNVs), 12 (1.7%) were small deletions, and 12 (1.7%) were small insertions. The variants were distributed around all mtDNA: 438 (61.8%) were located in coding genes (*ND5*: 69, *CO1*: 57, *ND4*: 50, *ND1*: 47, *CYB*: 43, *ATP6*: 40, *ND2*: 39, *CO3*: 25, *CO2*: 21, *ND3*: 17, *ND6*: 17, *ATP8*: 9, and *ND4L*: 7) and 271 (38.2%) in non-coding regions (D-Loop region: 114, rRNA: 101, tRNA: 47, and intergenic regions: 9). Given the nature of the mtDNA molecules, three variants were overlapped in the *ATP8* and *ATP6* genes. The mutation rate of mtDNA per kilobase (kb) in tumors was 42.79 variants/kb (vars/kb). Genes showing the highest number of vars/kb were *tRNA-C* (92.3), *tRNA-T* (92.3), *tRNA-L1* (67.6), *tRNA-W* (59.7), *ATP6* (58.8), *12S* (51.4), *ND1* (49.2), and the D-Loop region (95.4 vars/kb) (**Supplementary Figure 2**). None of the 709 variants were detected in *tRNA-A*, *tRNA-D*, or *tRNA-S1*. Transitions and transversions accounted for 67% (459) and 33% (226) of the SNVs, respectively (**Figure 1A**).

Of the total variants detected in tumor tissues, 324 were defined as germline variants and 173 as somatic mutations. Two hundred eleven variants were considered as probable mutations.

Germline Variants Are Highly Heterogeneous in the Heteroplasmic Levels in Breast Tumor Tissue

One hundred sixty-three (50.3%) germline variants were detected in more than two cases, and 161 (49.7%) were singletons, arising in only one patient. Only 15 out of 324 (4.6%) were identified in more than 10 ($>11\%$) patients: T146C, A153G, A235G, A663G, A1736G, T4248C, A4824G, G8027A, C8794T, C16111T,



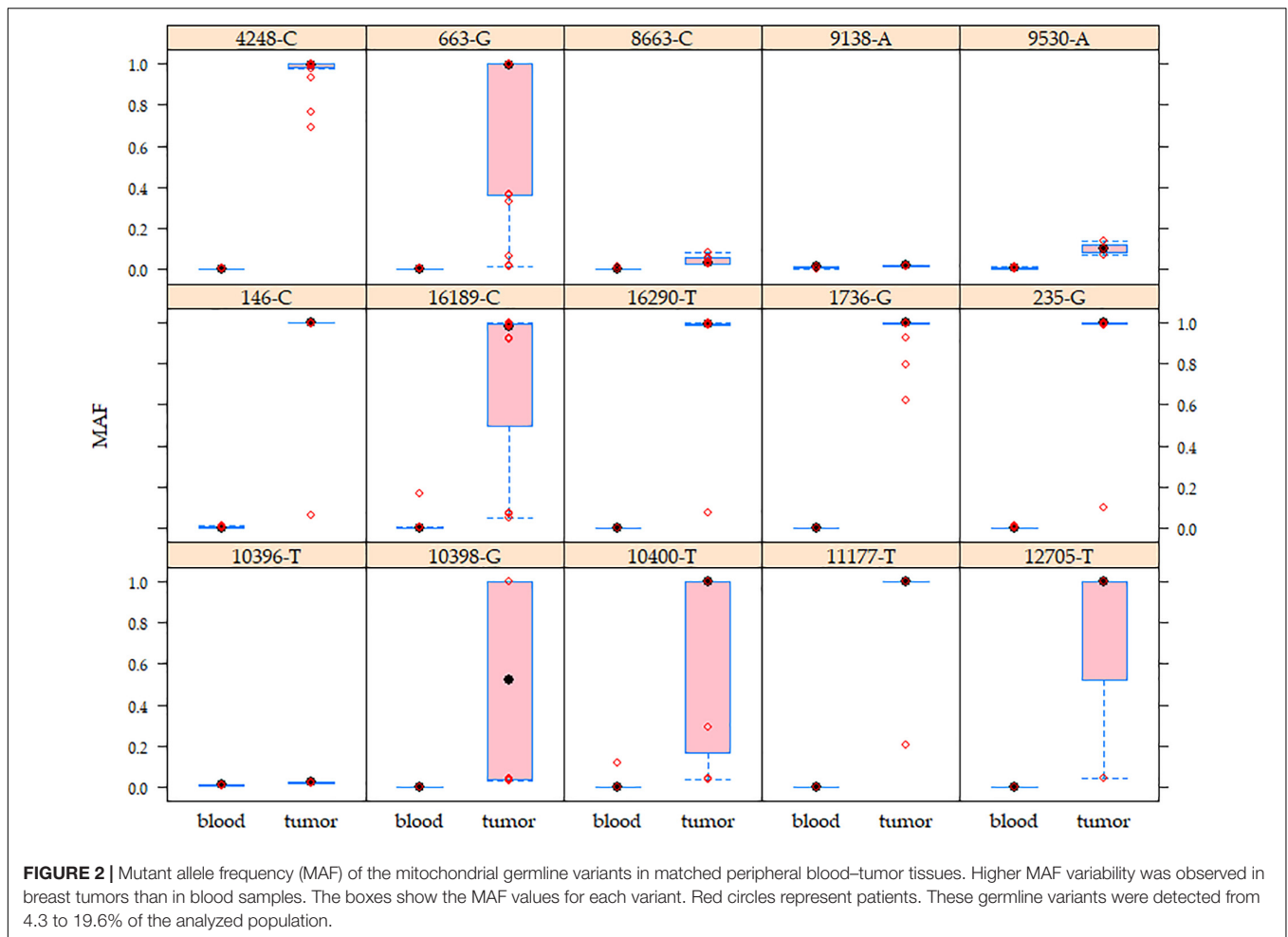
T16189C, C16290T, G16319A, T16362C, and T16519C, with A663G (19.6%) being the most frequent germline variant.

Heteroplasmy levels differed among tissues (blood <20%, tumors >1% and <95%) (**Supplementary Figure 3**). 127/324 (39.2%) germline variants exhibited a very low heteroplasmy level (1–2%) in both tissues, of which 42 (33.1%) were identified in more than two patients and the remaining 85 (66.9%) were singleton germline variants. To note, 25 (29.4%) of the singleton germline variants are reported in the MITOMAP database. Regarding the non-singleton germline variants, 66 (40.5%) displayed a shift in heteroplasmy since they presented

higher heteroplasmy levels in tumors than normal tissues in all positive patients, as exemplified by A1736G, T4248C, and C16290T (**Figure 2** and **Supplementary Figure 3**). Variants such as A663G, A10398G, C10400T, C12705T, and T16189C showed a wide range of heteroplasmy levels among tumors (**Figure 2**).

Distribution of Somatic Mutations

We detected 384 variants displaying MAF < 0.01 in blood samples, but >0.01 in tumor samples. The 211 variants showing MAFs between 0.001 and 0.01 in normal tissues were named as probable mutations, and the remaining 173 were identified as



somatic mutations (<0.001 MAF in normal tissue). Although 14 (6.6%) of the probable mutations presenting very low MAF in normal tissue (<0.008) were homoplasmic in tumors and 47 (22.3%) of all probable mutations were detected in more than two breast tumors (being T8430A: 9.8% and A8439C: 6.5% the most frequent), toward reducing biases derived from the technology and bioinformatic tools used, our further analysis was focused only on those variants classified as somatic mutations.

All somatic mutations were found in 68 (73.9%) tumors, ranging from 1 to 17 mutations per sample ($\bar{X} = 2 \pm 3$). Single-nucleotide variants comprised the major type of somatic mutations (98%), followed by deletions (1%), and insertions (1%). Transversions were less common than transitions (16 and 84%, respectively) (Figure 1B). Of the 173 somatic mutations, 102 (59%) were distributed along protein-coding regions, whereas 71 (41%) were found in non-coding regions (D-Loop, rRNA, tRNA, intergenic regions). The D-Loop region (12.7%) and *CO1* (10.9%), *16S* (9.2%), *12S* (8.1%), and *ND5* (8.1%) genes harbored more mutations than the other genes. Our data showed that 147 (85%) somatic mutations were singletons while the remaining 26 (15%) were in more than two cases, with G3219T (*16S*, 3.3%), T3631A (*ND1*, 3.3%), and A7124G (*CO1*, 3.3%) being the most frequent mutations.

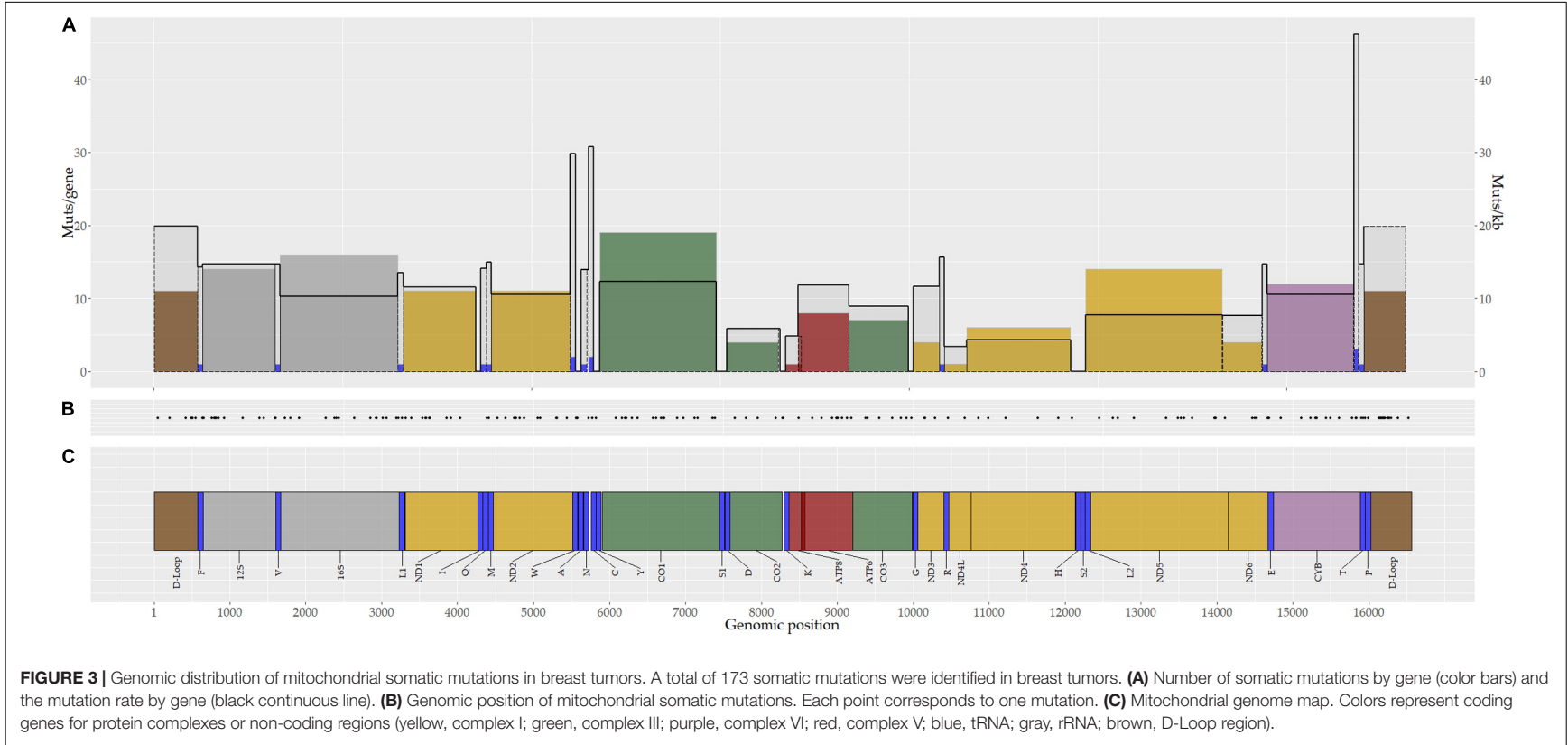
Based on the mutation rate per kilobase (mut/kb), the *tRNA-T* (46.2 mut/kb), *tRNA-C* (30.8 mut/kb), and *tRNA-W* (29.8 mut/kb) genes, as well as D-Loop region (19.8 mut/kb), showed the highest mutation rates (Figure 3). Notably, no somatic mutations were identified in 10 tRNAs genes (I, A, Y, S1, D, K, G, H, S2, and L2) (Figure 3).

Regarding the heteroplasmy status of the somatic mutations, 37 (21.4%) were homoplasmic, whereas the remaining 136 (78.6%) showed a high range of heteroplasmy levels. The T15115C, C16270A, and C16527T mutations, which were detected in two patients, exhibited a homoplasmic status in one patient but heteroplasmic in the other one.

From all somatic mutations, 56 (32.4%) are not recorded in the MITOMAP database (41), including two of the most frequent mutations detected in our cohort (G3219T and T3631A).

Functional Effect Prediction Analysis of Somatic Mutation

To determine whether the somatic mutations are either driver or passenger mutations, we annotated them using VEP software (42). Of 102 coding mutations, 34 (33.3%) were synonymous



and 68 (66.7%) were non-synonymous (missense: 91.2%; non-sense: 5.9%; frameshift: 2.9%). Regarding non-coding mutations, 8 (8.4%) were located along regulatory regions, 46 (64.8%) within the rRNA and tRNA genes, and 19 (26.8%) fell into D-Loop and intergenic regions (Table 1). Given the nature of the mtDNA molecules, the T5726C and G5777A mutations (located in *tRNA-N* and *tRNA-C* genes, respectively) overlap in L-strand replication origin 1 (MT-OLR1) sequences and have a double function (Supplementary Table 3). Functional prediction analysis suggested that non-coding mutations (regulatory regions) might modify sequences associated with mtDNA replication and transcription processes (Table 2).

Somatic Mutations Associated With Breast Cancer and Other Cancer Types

To determine if the mtDNA somatic mutations identified in this study have previously been associated with cancer, all mutations were examined in the MITOMAP database. We found that 25 (14.6%) of these mutations have been identified in several types of human cancer. Only G13333A and A15836C mutations, which were detected in one patient (1.1%), have been recorded in breast tumors (Table 2).

Distribution of Somatic Mutations by Breast Tumor Subtype

We assessed the association between mitochondrial mutations and clinical characteristics of breast cancer. No significant differences were found in terms of age, status, or tumor subtype among patients carrying any number of mutations and with no mutations ($P > 0.05$).

Unsupervised hierarchical clustering analysis with all 173 somatic mutations did not show any association between the mtDNA mutational profile and breast tumor subtypes. The distribution of all mutations and their heteroplasmy levels differed between patients (Supplementary Figure 4). The LA subtype displayed the highest mutation rate (5.97 mut/kb), followed by LB, TN, and H2 subtypes (1.32 mut/kb, 1.08 mut/kb, and 0.72 mut/kb, respectively). The LA subtype also showed the highest number of mutations (99/173), followed by LB (53/173), TN (18/173), and H2 (12/173) tumor subtypes. Most

mutations were exclusively found in one tumor subtype (LA subtype: 88, LB: 44, TN: 17, and H2:7) (Supplementary Table 4). Eleven (A2641C, T3535A, G5777A, C6371T, G8292A, T10463C, A11221G, G15928A, C16150T, C16294T, and C16527T), and 2 (A15496C and A644C) mutations were detected in 2% of the LA and LB tumors subtype, respectively. Seven mutations were shared by LA and LB tumors subtypes, but no one mutation was present in all tumor subtypes (Supplementary Figure 5). No statistical significance was found in terms of the distribution of somatic mutations and breast cancer subtypes ($P = 0.86$; Supplementary Table 4).

Tumor Mutational Burden Association and Overall Survival

To determine the use of mtDNA tumor mutational burden as a potential predictor of survival in breast cancer, a Kaplan–Meier analysis was performed (Supplementary Figure 6). Regarding OS, our analysis revealed no statistically significant differences among positive *versus* negative mtDNA somatic mutation groups (HR = 1.637; 95% CI = 0.401–6.683), or when comparing a low (≤ 2 mutations) or high (> 2 mutations) mutational burden (HR = 1.385; 95% CI = 0.329–5.833). The threshold was established according to the mutational burden mean ($n = 2$) of somatic mutations.

DISCUSSION

Breast cancer, the leading cancer type in women around the world, is characterized by its complexity and its highly heterogeneous genetic background, resulting from the interaction between known environmental factors and mutations in the nuclear genome and mtDNA. The mitochondrial genome function is relevant for the transformation processes; tumor cells are able to escape from different regulatory mechanisms, but they cannot evade energy flow mechanisms. Thus, cancer cells could gain a functional mtDNA mutation as a strategy to adjust energy metabolism during adaptation to oncogenic conditions (44–46). To evaluate the landscape of somatic mutation in breast tumors and to determine whether the mtDNA mutational burden is correlated with the OS of breast cancer, we sequenced 92 whole mtDNA paired peripheral blood–tumor samples from Mexican women.

Native American Haplogroups in Mexican Women With Breast Tumors

Seven different haplogroups were identified in our samples. The Native American haplogroups (A, B, C, and D) were more common than the major European (H and J) and African (L) haplogroups. Although the present study is limited by the selected population (women with breast cancer), we found similar frequencies of the A, B, and C haplogroups with those reported in the Mexican-Mestizo (mixed) populations and Mexican-Americans living in the United States (47, 48). Considering our findings and those previously published, the Native American haplogroups' frequencies agree with the demographic history of

TABLE 1 | Functional classification of the somatic mutations identified in breast cancer.

Effect type	Number of mutations, <i>n</i> (%)
Non-synonym	68 (66.7)
Missense	62 (91.2)
Non-sense	4 (5.9)
Frameshift	2 (2.9)
Synonym	34 (33.3)
Regulatory ^a	8 (8.4)
tRNA, rRNA	46 (64.8)
Intergenic	19 (26.8)

^aTwo mutations in tRNA genes were overlapped with L-strand replication origin 1 (OLR1) and had a double function.

TABLE 2 | Somatic mutation previously associated with cancer.

Position	Gene	Consequence	Function	AA change	Frequency (%)*	Tumor type
G207A	D-Loop	Regulatory	LPS	–	1.1	Prostate, thyroid
T3394C	ND1	Coding	Missense	Y/H	1.1	Leukemia, colorectal, lung
T4823C	ND2	Coding	Synonym	V	1.1	Bladder, lung, head, and neck
T5567C	TW	Transcript change	tRNA	–	1.1	Endometrial, ovary
A5581G	Intergenic	Non-coding	Non-coding	–	2.2	Thyroid
C6371T	CO1	Coding	Synonym	S	2.2	Colorectal
T7389C	CO1	Coding	Missense	Y/H	1.1	Thyroid
G10290A	ND3	Coding	Missense	A/T	1.1	Endometrial
T10463C	TR	Transcript change	tRNA	–	2.2	Endometrial, thyroid, colorectal
G12630A	ND5	Coding	Synonym	W	1.1	Colorectal
G13333A	ND5	Coding	Missense	A/T	1.1	Breast
A13966G	ND5	Coding	Missense	T/A	1.1	Prostate
T14470C	ND6	Coding	Synonym	G	1.1	Colorectal
T15831C	CYB	Coding	Missense	I/T	1.1	Endometrial
A15836C	CYB	Coding	Missense	I/L	1.1	Breast
G15928A	TT	Transcript change	tRNA	–	2.2	Colorectal
G15995A	TP	Transcript change	tRNA	–	1.1	Endometrial
C16134T	D-Loop	Non-coding	Non-coding	–	1.1	Glioblastoma
A16162G	D-Loop	Regulatory	TAS	–	1.1	Ovary
A16163G	D-Loop	Regulatory	TAS	–	1.1	Ovary
C16248T	D-Loop	Non-coding	Non-coding	–	1.1	Ovary
C16261T	D-Loop	Non-coding	Non-coding	–	1.1	Nasopharyngeal, ovary
C16294T	D-Loop	Non-coding	Non-coding	–	2.2	Ovary
C16296T	D-Loop	Non-coding	Non-coding	–	1.1	Ovary
C16527T	D-Loop	Non-coding	Non-coding	–	1.1	Ovary, pancreatic

AA, amino acid; LPS, L-strand promoter sequence; TAS, termination associated sequence. *Frequency corresponding to population analyzed in this study.

the Mexican-Mestizo population that is composed of Native-American (53%), European (42%), and African (5%) ancestry (48, 49).

We found a higher frequency of the D haplogroup (12%) than in previously published works about the Mexican-Mestizo population (6.2%) (50, 51). The D haplogroup has been associated with predisposition to breast cancer in the Chinese population and experimental data revealed that the D5 branch haplogroup promotes tumorigenesis through AKT activation, mediated by a high concentration of reactive oxygen species (ROS) (52, 53). However, we need to include more patients and to perform a case-control analysis to determine whether the haplogroup D has clinical significance in Mexican women with breast cancer.

High Heteroplasmy Levels of Germline Variants in Tumors and Heteroplasmy Shifting

In this study, we were able to distinguish acquired mutations from germline variants and to detect heteroplasmy shifting by analyzing normal-tumor matched tissues. On the one hand, <33% of the germline variants exhibiting low heteroplasmy levels (1–2%) were detected in more than two subjects, and 29.4% of singletons germline variants have been annotated in the public databases, suggesting that these could be rare variants. On the other hand, for several germline variants (as A1736C, T4248C, and C16290T), the mutant allele was enriched in tumors in

comparison with normal tissues (Figure 2). A663G, a marker of haplogroup A, was the most common germline variant. No evidence of an association among haplogroup A, or this variant with breast cancer risk, has been documented. However, the 663G allele has been reported as associated with various mitochondrial diseases (54–57). Additionally, A663G has been found as a rare variant in papillary thyroid carcinoma and nasopharyngeal carcinoma (frequency <1%) (58, 59). *In silico* predictions suggest that the change of A > G could destabilize the structure of the rRNA by interrupting the hydrogen bond between U and A, then altering its function (54, 60). Nevertheless, its role in cancer development needs to be addressed.

The remaining common variants, T146C, A153G, A235G, A663G, A1736G, T4248C, A4824G, G8027A, C8794T, C16111T, T16189C, C16290T, G16319A, T16362C, and T16519C, are part of well-known mtDNA haplogroups and have been detected in some types of cancer such as prostate, ovarian, nasopharyngeal, thyroid, colorectal, gastric, and glioblastoma (58, 59, 61–67). Notwithstanding, to the best of our knowledge, only C16290T, which was identified in 14.1% of our cases, has been reported in breast tumors with a frequency of <1% (68).

Differences of heteroplasmy levels among tissues (blood <20%, tumors 1–95%) observed in this study could be biased by contamination by a different proportion of germline mtDNA during sequencing (69) or that the unregistered singletons germline variants showing <2% of heteroplasmy load arose during the genomic sequencing process (70). Our current strategy

does not allow us to identify whether the heteroplasmy comes from the enrichment of mutated mitochondria or from mtDNA copy numbers (mtDNA-CN) variations. Diverse analyses have highlighted the great heterogeneity of the mtDNA-CN in different tissues and cancer types (71–74), but studies comparing mtDNA-CN among blood and normal tissue adjacent to the tumor are scarce and have found inconclusive results (28, 74–77).

mtDNA heteroplasmy is a normal condition of randomized processes and regulated homeostatic mechanisms, such as the mitochondrial segregation during the cell division, and mitochondrial fusion and fission process (78, 79). The changes in allelic frequencies have been suggested to be the response to selective pressures generated under physiological and pathological environmental conditions (17). However, it could be also as a result of changes in mtDNA-CN that occurred during the mtDNA replication (which is cell cycle independent) (80–82), or under oxidative stress induced by endogenous and exogenous factors (hormones, age, dietary, etc.) (83). The correlation among heteroplasmy and mtDNA-CN has scarcely been explored, and higher levels of heteroplasmy were correlated with lower mtDNA-CN in the central nervous system, but not in other tissues (73). In breast cancer, controversial results are found (83, 84), and it has been suggested that both processes are not interacting with each other and play different roles in the development of breast tumors (83). Although we have to consider that the shift toward homoplasmy observed in the present work could be influenced by mtDNA-CN rather than an enrichment of mutated mitochondria, heteroplasmy shifting has been reported in kidney and thyroid carcinomas (85), as well as breast cancer (28). Studies indicate that tumor cells show a neutral evolution of their mtDNA and even have the ability to tolerate pathogenic mutations; however, some mutant alleles reaching a critical threshold can contribute to cancer development and progression (85). The molecular mechanisms by which these alleles (whose heteroplasmy levels shift in the tumor) contribute to carcinogenesis is still unclear.

High Frequency of Heteroplasmic mtDNA Mutations in Breast Tumors

In our study, 24 (26.1%) breast tumors were negative for somatic mutations and 68 (73.9%) were positive; over 78% of these mutations were heteroplasmic. Studies of breast cancer and diverse types of tumors have reported different proportions of tumors carrying mtDNA mutations (73.7 and 45.7%, respectively) (26, 28, 69, 86). The discrepancies among these findings could be partially explained by limitations derived from the sequencing methods, or due to the different thresholds used to define a somatic mtDNA mutation (28, 86). We found a high frequency of heteroplasmic mtDNA mutations and coding mutations (mainly non-synonym mutations), as well as a high mutation rate in the D-loop region and tRNA genes, as previously reported in several types of tumors (86, 87).

Coding mtDNA mutations could alter the protein function and OXPHOS system, and non-coding mutations might affect fundamental biological processes such as mtDNA replication, transcription, and structural mtDNA organization, thus conferring advantages for tumor cell proliferation (88, 89).

It has been suggested that deleterious pathogenic mutations, arising during the carcinogenic process are eliminated when protein function and viability cells are compromised, but selected if they enable cell proliferation (44, 90–93). In fact, it has been proposed that due to the *ND5* proton channel function, mutations in this gene might contribute to a deficient yield of the OXPHO system, leading to a high rate of oncogenic mutations in nuclear DNA and mtDNA (94, 95). Studies describing mutations in metastatic lesions, which were not found in primary tumors, support the hypothesis of the relevance of mtDNA mutations in cancer progression (90, 96). Li et al., by analyzing 26 types of cancer, proposed that regions and functional units are under negative and positive selection processes, suggesting the important role of mtDNA mutations in tumors (86).

Various authors stated that missense mutations of mtDNA are non-tumorigenic since they do not determine the impairment of the respiratory chain (28, 97). Likewise, since tRNA mtDNA pathogenic mutations are rarely fixed, the deleterious effect of these mutations might be functionally compensated for by gaining new mutations or altering mtDNA-CN (98). Other authors have proposed random processes defining the landscape of mtDNA mutations in cancer, suggesting that the heteroplasmic status arises in tumor progenitor cells by chance, without conferring any physiological or tumorigenic advantages (87). Primary tumors may have multiple clonal subpopulations that emerged during the numerous selective events required to complete the tumorigenesis process (99); thus, we have to consider that in addition to mtDNA-CN changes, clonal selection occurring during breast cancer progression could influence the levels of heteroplasmy observed in the present study and in other types of tumor (80, 100). More experiments are needed to understand the mechanisms underlying the mtDNA germline variants selection toward a homoplasmic state in breast cancer development.

As several of the recurrent mutations identified in this analysis (carried by >3% of our patients) have been related to other human diseases (58, 62, 101, 102), these mutations could also play a role in breast tumor development. However, it is necessary to validate our findings in an external cohort, including a population with a different haplogroup background. As well, we have to take into consideration that driver and passenger mutations may acquire oncogenic potential when, next to them, mtDNA gains more variants. In animal models, it was observed that the mtDNA containing two mutations in the *ND6* gene (G13997A and 13588insC) enhances the metastatic potential of tumor cells (12). Functional experiments could improve our understanding of the oncogenic biological processes mediated by these mutations (103, 104). The use of transmitochondrial hybrids might be helpful to address these issues (12).

The differentiation between a somatic mutation and a germline variant is a very complex task during the study of the mitochondrial genome. Technical and bioinformatics limitations might derive in misclassification of the variants. An interesting observation was the detection of the two coding variants T8430A (L22H) and A8439C (Q25P) as part of the probable mutation group. 8430A allele has been proposed to be associated with elevated ROS concentrations and reduced ATP synthesis (67, 105, 106). The 8439C allele has been described

as a mutation in breast cancer and *in silico* analysis suggested that this mutation is highly deleterious and affects the *ATP8* function (107). Greater depth coverage of the mtDNA analysis could precisely determine the nature of the probable mutations (germline variants or somatic mutation) found in this work. As well, additional studies on T8430A and A8439C are needed to decipher their role in breast cancer development and progression.

mtDNA Mutations and Clinical Relevance

In addition to the description of the landscape of the mtDNA mutations in breast tumors, we also evaluated their clinical relevance in tumor classification and prognosis. Over 85% of the somatic mutations were found with low frequency (<1%), probably due to individual variations (28). The LA and LB tumor subtypes displayed the highest number of mutations, sharing 4.1% of the mutations. However, we found neither an mtDNA mutation profile nor haplogroups associated with tumor subtypes. Few studies have explored the association between mtDNA mutational burden and haplogroups with breast tumor molecular subtypes, and their results are still inconclusive (51). Based on nuclear DNA mutation data showing that hormone-receptor-negative tumors have a higher mutational burden than hormone-receptor-positive tumors (108), we expected to observe the same behavior regarding mtDNA mutations. We suggest that the elevated mutational burden of hormone-receptor-positive tumors (LA and LB) could be explained by superoxide radical formation during estrogen metabolism, which promotes mtDNA alterations (103, 109). A recent study revealed a differential expression profile of mitochondria-related genes among molecular subtypes of breast cancer, perhaps derived from the partial contribution of the mtDNA mutations in breast tumor biology (110). Increasing sample size and balancing tumor subtypes could support our observations.

Overall survival analysis showed no differences among patients carrying a high *versus* low mtDNA mutational burden in tumors, which is in contrast with the data reported in breast cancer (28). A high mtDNA mutational burden has been associated with a worse outcome in pulmonary adenocarcinoma, acute myeloid leukemia, and pancreatic ductal adenocarcinoma (18, 111, 112); the opposite was observed in oral squamous cell carcinoma, and even acute myeloid leukemia, and breast cancer (28, 113, 114). We cannot exclude potential bias due to the nature of the studied mutations, the clinical characteristics of the patients included, or the ethnic/genetic background of the populations studied. Our small sample size and potential population stratification, since all patients were recruited in Mexico City, might also bias our results. As well, we cannot ignore the fact that other alterations, such as the mtDNA-CN (suggested as a potential prognostic biomarker for breast cancer), mito epigenetic processes, and nuclear genes that are involved in the mitochondrial biogenesis, may be clinically relevant (18, 71, 115, 116). Thus, studies are necessary for larger cohorts of patients to determine the significance of the mtDNA mutational burden as a biomarker in breast cancer.

This study describes the distribution of mtDNA germline variants and mutations in breast tumors in a population of Mexican women. We found a high mutation rate in the

D-loop region and tRNA genes. Heteroplasmy analysis suggested that negative and positive selection processes are shaping the landscape of mtDNA mutations in breast cancer, but functional experiments are needed to further understand the oncogenic biological processes.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: European Nucleotide Archive (<https://www.ebi.ac.uk/ena>) (Accession: PRJEB40354).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Instituto Nacional de Medicina Genómica. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SJ-M, CP-A, and AH-M: conceptualization and design. CP-A, VB-P, FB-A, LA-R, CD-R, FV-C, and AT-T: sample identification and clinical follow-up. CP-A, HT, LG-R, and FB-A: data analysis. SJ-M and AH-M: supervision. CP-A and SJ-M: writing of the original draft. All authors: manuscript revision and approval of the submitted version.

FUNDING

This study was funded by grants from the Consejo Nacional de Ciencia y Tecnología (CONACyT) through the program Fondo Sectorial de Investigación en Salud y Seguridad Social (S0008-2016-1-272618) and the Instituto Nacional de Medicina Genómica (C2_19/2009 and 24/2016/I). CP-A was supported by CONACyT (CVU 821714). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank the Unit of Sequencing (USeq) at INMEGEN for technical support and the Department of Computational Genomics at INMEGEN for the bioinformatic support. Finally, we are extremely grateful to all the patients who participated in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.572954/full#supplementary-material>

REFERENCES

- Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and Mortality and Epidemiology of Breast Cancer in the World. *Asian Pac J Cancer Prev.* (2016) 17:43–6. doi: 10.7314/apjcp.2016.17.3.43
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
- Maffuz-Aziz A, Labastida-Almendares S, Espejo-Fonseca A, Rodriguez-Cuevas S. [Clinical and pathological features of breast cancer in a population of Mexico]. *Cir Cir.* (2017) 85:201–7. doi: 10.1016/j.circir.2016.08.004
- Provenzano E, Ulaner GA, Chin SF. Molecular classification of breast cancer. *PET Clin.* (2018) 13:325–38. doi: 10.1016/j.cpet.2018.02.004
- Reynosio-Noveron N, Villarreal-Garza C, Soto-Perez-de-Celis E, Arce-Salinas C, Matus-Santos J, Ramirez-Ugalde MT, et al. Clinical and epidemiological profile of breast cancer in Mexico: results of the seguro popular. *J Glob Oncol.* (2017) 3:757–64. doi: 10.1200/JGO.2016.007377
- Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res.* (2015) 5:2929–43.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature.* (2000) 406:747–52. doi: 10.1038/35021093
- Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res.* (2010) 16:5222–32. doi: 10.1158/1078-0432.CCR-10-1282
- Rheinbay E, Parasuraman P, Grimsby J, Tiao G, Engreitz JM, Kim J, et al. Recurrent and functional regulatory mutations in breast cancer. *Nature.* (2017) 547:55–60. doi: 10.1038/nature22992
- Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature.* (2012) 486:405–9. doi: 10.1038/nature11154
- Scheid AD, Beadnell TC, Welch DR. The second genome: effects of the mitochondrial genome on cancer progression. *Adv Cancer Res.* (2019) 142:63–105. doi: 10.1016/bs.acr.2019.01.001
- Ishikawa K, Takenaga K, Akimoto M, Koshikawa N, Yamaguchi A, Imanishi H, et al. ROS-generating mitochondrial DNA mutations can regulate tumor cell metastasis. *Science.* (2008) 320:661–4. doi: 10.1126/science.1156906
- Hou XS, Wang HS, Mugaka BP, Yang GJ, Ding Y. Mitochondria: promising organelle targets for cancer diagnosis and treatment. *Biomater Sci.* (2018) 6:2786–97. doi: 10.1039/c8bm00673c
- Lee SR, Han J. Mitochondrial nucleoid: shield and switch of the mitochondrial genome. *Oxid Med Cell Longev.* (2017) 2017:8060949. doi: 10.1155/2017/8060949
- Stefano GB, Bjenning C, Wang F, Wang N, Kream RM. Mitochondrial Heteroplasmy. *Adv Exp Med Biol.* (2017) 982:577–94. doi: 10.1007/978-3-319-55330-6_30
- Larman TC, DePalma SR, Hadjipanayis AG, Cancer Genome Atlas Research Network, Protopopov A, Zhang J, et al. Spectrum of somatic mitochondrial mutations in five cancers. *Proc Natl Acad Sci USA.* (2012) 109:14087–91. doi: 10.1073/pnas.1211502109
- Jimenez-Morales S, Perez-Amado CJ, Langley E, Hidalgo-Miranda A. Overview of mitochondrial germline variants and mutations in human disease: focus on breast cancer (Review). *Int J Oncol.* (2018) 53:923–36. doi: 10.3892/ijo.2018.4468
- Wu S, Akhtari M, Alachkar H. Characterization of mutations in the mitochondrial encoded electron transport chain complexes in acute myeloid leukemia. *Sci Rep.* (2018) 8:13301. doi: 10.1038/s41598-018-31489-0
- Kumari T, Vachher M, Bansal S, Bamezai RNK, Kumar B. Meta-analysis of mitochondrial T16189C polymorphism for cancer and Type 2 diabetes risk. *Clin Chim Acta.* (2018) 482:136–43. doi: 10.1016/j.cca.2018.03.041
- Tipiriseti NR, Govatati S, Pullari P, Malempati S, Thupurani MK, Perugu S, et al. Mitochondrial control region alterations and breast cancer risk: a study in South Indian population. *PLoS One.* (2014) 9:e85363. doi: 10.1371/journal.pone.0085363
- Czarnecka AM, Krawczyk T, Plak K, Klemba A, Zdrozny M, Arnold RS, et al. Mitochondrial genotype and breast cancer predisposition. *Oncol Rep.* (2010) 24:1521–34. doi: 10.3892/or.00001014
- Weerts MJA, Sleijfer S, Martens JWM. The role of mitochondrial DNA in breast tumors. *Drug Discov Today.* (2019) 24:1202–8. doi: 10.1016/j.drudis.2019.03.019
- Wang Y, Liu VW, Tsang PC, Chiu PM, Cheung AN, Khoo US, et al. Microsatellite instability in mitochondrial genome of common female cancers. *Int J Gynecol Cancer.* (2006) 16(Suppl. 1):259–66. doi: 10.1111/j.1525-1438.2006.00412.x
- Bianchi MS, Bianchi NO, Bailliet G. Mitochondrial DNA mutations in normal and tumor tissues from breast cancer patients. *Cytogenet Cell Genet.* (1995) 71:99–103. doi: 10.1159/000134072
- Weigl S, Paradiso A, Tommasi S. Mitochondria and familial predisposition to breast cancer. *Curr Genomics.* (2013) 14:195–203. doi: 10.2174/1389202911314030005
- Tan DJ, Bai RK, Wong LJ. Comprehensive scanning of somatic mitochondrial DNA mutations in breast cancer. *Cancer Res.* (2002) 62:972–6.
- Verma M, Naviaux RK, Tanaka M, Kumar D, Franceschi C, Singh KK. Meeting report: mitochondrial DNA and cancer epidemiology. *Cancer Res.* (2007) 67:437–9. doi: 10.1158/0008-5472.CAN-06-4119
- McMahon S, LaFramboise T. Mutational patterns in the breast cancer mitochondrial genome, with clinical correlates. *Carcinogenesis.* (2014) 35:1046–54. doi: 10.1093/carcin/bgu012
- Bai RK, Leal SM, Covarrubias D, Liu A, Wong LJ. Mitochondrial genetic background modifies breast cancer risk. *Cancer Res.* (2007) 67:4687–94. doi: 10.1158/0008-5472.CAN-06-3554
- Zhu W, Qin W, Bradley P, Wessel A, Puckett CL, Sauter ER. Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid. *Carcinogenesis.* (2005) 26:145–52. doi: 10.1093/carcin/bgh282
- Lin CS, Chang SC, Ou LH, Chen CM, Hsieh SS, Chung YP, 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) 33:2924–34. doi: 10.3892/or.2015.3887
- Ahn EH, Hirohata K, Kohn BF, Fox EJ, Chang CC, Loeb LA. Detection of Ultra-rare mitochondrial mutations in breast stem cells by duplex sequencing. *PLoS One.* (2015) 10:e0136216. doi: 10.1371/journal.pone.0136216
- Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, et al. Universal heteroplasmy of human mitochondrial DNA. *Hum Mol Genet.* (2013) 22:384–90. doi: 10.1093/hmg/dd5435
- Senkus E, Kyriakides S, Penault-Llorca F, Poortmans P, Thompson A, Zackrisson S, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* (2013) 6:v17–23. doi: 10.1093/annonc/mdt284
- Andrews S. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* Babraham Bioinformatics: Cambridge (2010).
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* (2010) 26:589–95. doi: 10.1093/bioinformatics/btp698
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* (2010) 20:1297–303. doi: 10.1101/gr.107524.110
- Weissensteiner H, Pacher D, Kloss-Brandstatter A, Forer L, Specht G, Bandelt HJ, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* (2016) 44:W58–63. doi: 10.1093/nar/gkw233
- van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* (2009) 30:E386–94. doi: 10.1002/humu.20921
- Kim S, Scheffler K, Halpern AL, Bekrity MA, Noh E, Kallberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods.* (2018) 15:591–4. doi: 10.1038/s41592-018-0051-x
- Lott MT, Leipzig JN, Derbeneva O, Xie HM, Chalkia D, Sarmady M, et al. mtDNA variation and analysis using mitomap and mitomaster. *Curr Protoc Bioinformatics.* (2013) 44:1–6. doi: 10.1002/0471250953.bi0123s44

42. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The ensembl variant effect predictor. *Genome Biol.* (2016) 17:122. doi: 10.1186/s13059-016-0974-4
43. Bardou P, Mariette F, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics.* (2014) 15:293. doi: 10.1186/1471-2105-15-293
44. Wallace DC. Bioenergetic origins of complexity and disease. *Cold Spring Harb Symp Quant Biol.* (2011) 76:1–16. doi: 10.1101/sqb.2011.76.010462
45. Chandra D, Singh KK. Genetic insights into OXPHOS defect and its role in cancer. *Biochim Biophys Acta.* (2011) 1807:620–5. doi: 10.1016/j.bbabo.2010.10.023
46. Wallace DC. A mitochondrial bioenergetic etiology of disease. *J Clin Invest.* (2013) 123:1405–12. doi: 10.1172/JCI61398
47. Baptista RRC, Hernandez OL, Vega AJ, Mercado SA, Arreola CAA. Reevaluación de polimorfismos asociados a riesgo de desarrollo de cáncer mamario en población mexicana a partir de información disponible en bases de datos públicas. *Genét Méd Genóm.* (2018) 3:359.
48. Baptista Rosas RC, Mercado Sesma A, Hernandez Ortega L, Hernandez Gonzalez L, Vega Avalos J, Arreola Cruz AA. The utility of genomic public databases to mitochondrial haplotyping in contemporary Mestizo population of Mexican origin. *Mitochondrial DNA A DNA Mapp Seq Anal.* (2019) 30:567–72. doi: 10.1080/24701394.2019.1580271
49. Moreno A, Sandoval K. Diversidad genómica en México: pasado indígena y mestizaje. *J Cuicuilco.* (2013) 20:249–75.
50. Martinez-Cortes G, Salazar-Flores J, Haro-Guerrero J, Rubi-Castellanos R, Velarde-Felix JS, Munoz-Valle JF, et al. Maternal admixture and population structure in Mexican-Mestizos based on mtDNA haplogroups. *Am J Phys Anthropol.* (2013) 151:526–37. doi: 10.1002/ajpa.22293
51. Tommasi S, Favia P, Weigl S, Bianco A, Pilato B, Russo L, et al. Mitochondrial DNA variants and risk of familial breast cancer: an exploratory study. *Int J Oncol.* (2014) 44:1691–8. doi: 10.3892/ijo.2014.2324
52. Choudhury AR, Singh KK. Mitochondrial determinants of cancer health disparities. *Semin Cancer Biol.* (2017) 47:125–46. doi: 10.1016/j.semcancer.2017.05.001
53. Ma L, Fu Q, Xu B, Zhou H, Gao J, Shao X, et al. Breast cancer-associated mitochondrial DNA haplogroup promotes neoplastic growth via ROS-mediated AKT activation. *Int J Cancer.* (2018) 142:1786–96. doi: 10.1002/ijc.31207
54. Lu J, Li Z, Zhu Y, Yang A, Li R, Zheng J, et al. Mitochondrial 12S rRNA variants in 1642 Han Chinese pediatric subjects with aminoglycoside-induced and nonsyndromic hearing loss. *Mitochondrion.* (2010) 10:380–90. doi: 10.1016/j.mito.2010.01.007
55. Ryzdzanicz M, Wrobel M, Cywinska K, Froehlich D, Gawrecki W, Szyfter W, et al. Screening of the general Polish population for deafness-associated mutations in mitochondrial 12S rRNA and tRNA Ser(UCN) genes. *Genet Test Mol Biomarkers.* (2009) 13:167–72. doi: 10.1089/gtmb.2008.0098
56. Wang Q, Li QZ, Han D, Zhao Y, Zhao L, Qian Y, et al. Clinical and molecular analysis of a four-generation Chinese family with aminoglycoside-induced and nonsyndromic hearing loss associated with the mitochondrial 12S rRNA C1494T mutation. *Biochem Biophys Res Commun.* (2006) 340:583–8. doi: 10.1016/j.bbrc.2005.12.045
57. Nishigaki Y, Yamada Y, Fuku N, Matsuo H, Segawa T, Watanabe S, et al. Mitochondrial haplogroup A is a genetic risk factor for atherothrombotic cerebral infarction in Japanese females. *Mitochondrion.* (2007) 7:72–9. doi: 10.1016/j.mito.2006.11.002
58. Yeh JJ, Lunetta KL, van Orsouw NJ, Moore FD Jr., Mutter GL, Vijg J, et al. Somatic mitochondrial DNA (mtDNA) mutations in papillary thyroid carcinomas and differential mtDNA sequence variants in cases with thyroid tumours. *Oncogene.* (2000) 19:2060–6. doi: 10.1038/sj.onc.1203537
59. Pang LJ, Shao JY, Liang XM, Xia YF, Zeng YX. Mitochondrial DNA somatic mutations are frequent in nasopharyngeal carcinoma. *Cancer Biol Ther.* (2008) 7:198–207. doi: 10.4161/cbt.7.2.5256
60. Sawabe M, Tanaka M, Chida K, Arai T, Nishigaki Y, Fuku N, et al. Mitochondrial haplogroups A and M7a confer a genetic risk for coronary atherosclerosis in the Japanese elderly: an autopsy study of 1,536 patients. *J Atheroscler Thromb.* (2011) 18:166–75. doi: 10.5551/jat.6742
61. Jeronimo C, Nomoto S, Caballero OL, Usadel H, Henrique R, Varzim G, et al. Mitochondrial mutations in early stage prostate cancer and bodily fluids. *Oncogene.* (2001) 20:5195–8. doi: 10.1038/sj.onc.1204646
62. Webb E, Broderick P, Chandler I, Lubbe S, Penegar S, Tomlinson IP, et al. Comprehensive analysis of common mitochondrial DNA variants and colorectal cancer risk. *Br J Cancer.* (2008) 99:2088–93. doi: 10.1038/sj.bjc.6604805
63. Chen JZ, Gokden N, Greene GF, Mukunyadzi P, Kadlubar FF. Extensive somatic mitochondrial mutations in primary prostate cancer using laser capture microdissection. *Cancer Res.* (2002) 62:6470–4.
64. Brandon M, Baldi P, Wallace DC. Mitochondrial mutations in cancer. *Oncogene.* (2006) 25:4647–62. doi: 10.1038/sj.onc.1209607
65. Bragoszewski P, Kupryjanczyk J, Bartnik E, Rachinger A, Ostrowski J. Limited clinical relevance of mitochondrial DNA mutation and gene expression analyses in ovarian cancer. *BMC Cancer.* (2008) 8:292. doi: 10.1186/1471-2407-8-292
66. Kirches E, Krause G, Warich-Kirches M, Weis S, Schneider T, Meyer-Puttitz B, et al. High frequency of mitochondrial DNA mutations in glioblastoma multiforme identified by direct sequence comparison to blood samples. *Int J Cancer.* (2001) 93:534–8. doi: 10.1002/ijc.1375
67. Wu CW, Yin PH, Hung WY, Li AF, Li SH, Chi CW, et al. Mitochondrial DNA mutations and mitochondrial DNA depletion in gastric cancer. *Genes Chromosomes Cancer.* (2005) 44:19–28. doi: 10.1002/gcc.20213
68. Tseng LM, Yin PH, Chi CW, Hsu CY, Wu CW, Lee LM, et al. Mitochondrial DNA mutations and mitochondrial DNA depletion in breast cancer. *Genes Chromosomes Cancer.* (2006) 45:629–38. doi: 10.1002/gcc.20326
69. Stafford P, Chen-Quin EB. The pattern of natural selection in somatic cancer mutations of human mtDNA. *J Hum Genet.* (2010) 55:605–12. doi: 10.1038/jhg.2010.76
70. Castellani CA, Longchamps RJ, Sun J, Guallar E, Arking DE. Thinking outside the nucleus: mitochondrial DNA copy number in health and disease. *Mitochondrion.* (2020) 53:214–23. doi: 10.1016/j.mito.2020.06.004
71. Shen J, Platek M, Mahasneh A, Ambrosone CB, Zhao H. Mitochondrial copy number and risk of breast cancer: a pilot study. *Mitochondrion.* (2010) 10:62–8. doi: 10.1016/j.mito.2009.09.004
72. Shen H, Yu M, Tsoli M, Chang C, Joshi S, Liu J, et al. Targeting reduced mitochondrial DNA quantity as a therapeutic approach in pediatric high-grade gliomas. *Neuro Oncol.* (2020) 22:139–51. doi: 10.1093/neuonc/noz140
73. Scholle LM, Zierz S, Mawrin C, Wickenhauser C, Urban DL. Heteroplasmy and copy number in the common m.3243A>G mutation—a post-mortem genotype-phenotype analysis. *Genes (Basel).* (2020) 11:212. doi: 10.3390/genes11020212
74. Yuan Y, Ju YS, Kim Y, Li J, Wang Y, Yoon CJ, et al. Comprehensive molecular characterization of mitochondrial genomes in human cancers. *Nat Genet.* (2020) 52:342–52. doi: 10.1038/s41588-019-0557-x
75. Xia P, Wang HJ, Geng TT, Xun XJ, Zhou WJ, Jin TB, et al. Mitochondrial DNA levels in blood and tissue samples from breast cancer patients of different stages. *Asian Pac J Cancer Prev.* (2014) 15:1339–44. doi: 10.7314/apjcp.2014.15.3.1339
76. Park SY, Shin MG, Kim HR, Oh JY, Kim SH, Shin JH, et al. Alteration of mitochondrial DNA sequence and copy number in nasal polyp tissue. *Mitochondrion.* (2009) 9:318–25. doi: 10.1016/j.mito.2009.04.006
77. Reznik E, Miller ML, Senbabaoglu Y, Riaz N, Sarungbam J, Tickoo SK, et al. Mitochondrial DNA copy number variation across human cancers. *Elife.* (2016) 5:10769. doi: 10.7554/eLife.10769
78. Hahn A, Zuryn S. The cellular mitochondrial genome landscape in disease. *Trends Cell Biol.* (2019) 29:227–40. doi: 10.1016/j.tcb.2018.11.004
79. Naeem MM, Sondheimer N. Heteroplasmy shifting as therapy for mitochondrial disorders. *Adv Exp Med Biol.* (2019) 1158:257–67. doi: 10.1007/978-981-13-8367-0_14
80. Grady JP, Pickett SJ, Ng YS, Alston CL, Blakely EL, Hardy SA, et al. mtDNA heteroplasmy level and copy number indicate disease burden in m.3243A>G mitochondrial disease. *EMBO Mol Med.* (2018) 10:262. doi: 10.15252/emmm.201708262
81. Guo W, Yang D, Xu H, Zhang Y, Huang J, Yang Z, et al. Mutations in the D-loop region and increased copy number of mitochondrial DNA in human laryngeal squamous cell carcinoma. *Mol Biol Rep.* (2013) 40:13–20. doi: 10.1007/s11033-012-1939-7

82. Zhang H, Burr SP, Chinnery PF. The mitochondrial DNA genetic bottleneck: inheritance and beyond. *Essays Biochem.* (2018) 62:225–34. doi: 10.1042/EBC20170096
83. Shen J, Wan J, Song R, Zhao H. Peripheral blood mitochondrial DNA copy number, length heteroplasmy and breast cancer risk: a replication study. *Carcinogenesis.* (2015) 36:1307–13. doi: 10.1093/carcin/bgv130
84. Zhao H, Shen J, Medico L, Platek M, Ambrosone CB. Length heteroplasms in human mitochondrial DNA control regions and breast cancer risk. *Int J Mol Epidemiol Genet.* (2010) 1:184–92.
85. Grandhi S, Bosworth C, Maddox W, Sensiba C, Akhavanfard S, Ni Y, et al. Heteroplasmic shifts in tumor mitochondrial genomes reveal tissue-specific signals of relaxed and positive selection. *Hum Mol Genet.* (2017) 26:2912–22. doi: 10.1093/hmg/ddx172
86. Li D, Du X, Guo X, Zhan L, Li X, Yin C, et al. Site-specific selection reveals selective constraints and functionality of tumor somatic mtDNA mutations. *J Exp Clin Cancer Res.* (2017) 36:168. doi: 10.1186/s13046-017-0638-6
87. Collier HA, Khrapko K, Bodyak ND, Nekhaeva E, Herrero-Jimenez P, Thilly WG. High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. *Nat Genet.* (2001) 28:147–50. doi: 10.1038/88859
88. Gomez-Velazquez M, Badia-Careaga C, Lechuga-Vieco AV, Nieto-Arellano R, Tena JJ, Rollan I, et al. CTCF counter-regulates cardiomyocyte development and maturation programs in the embryonic heart. *PLoS Genet.* (2017) 13:e1006985. doi: 10.1371/journal.pgen.1006985
89. Marinov GK, Wang YE, Chan D, Wold BJ. Evidence for site-specific occupancy of the mitochondrial genome by nuclear transcription factors. *PLoS One.* (2014) 9:e84713. doi: 10.1371/journal.pone.0084713
90. Ju YS, Alexandrov LB, Gerstung M, Martincorena I, Nik-Zainal S, Ramakrishna M, et al. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife.* (2014) 3:2935. doi: 10.7554/eLife.02935
91. Errichiello E, Venesio T. Mitochondrial DNA variants in colorectal carcinogenesis: drivers or passengers? *J Cancer Res Clin Oncol.* (2017) 143:1905–14. doi: 10.1007/s00432-017-2418-2
92. Deng N, Zhou H, Fan H, Yuan Y. Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget.* (2017) 8:110635–49. doi: 10.18632/oncotarget.22372
93. Adzhubei AA, Adzhubei IA, Krashennnikov IA, Neidle S. Non-random usage of 'degenerate' codons is related to protein three-dimensional structure. *FEBS Lett.* (1996) 399:78–82. doi: 10.1016/s0014-5793(96)01287-2
94. Lets JA, Sazanov LA. Clarifying the supercomplex: the higher-order organization of the mitochondrial electron transport chain. *Nat Struct Mol Biol.* (2017) 24:800–8. doi: 10.1038/nsmb.3460
95. Gu J, Wu M, Guo R, Yan K, Lei J, Gao N, et al. The architecture of the mammalian respirasome. *Nature.* (2016) 537:639–43. doi: 10.1038/nature19359
96. Yadav N, Chandra D. Mitochondrial DNA mutations and breast tumorigenesis. *Biochim Biophys Acta.* (2013) 1836:336–44. doi: 10.1016/j.bbcan.2013.10.002
97. Kondrashov FA. Prediction of pathogenic mutations in mitochondrially encoded human tRNAs. *Hum Mol Genet.* (2005) 14:2415–9. doi: 10.1093/hmg/ddi243
98. Kern AD, Kondrashov FA. Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. *Nat Genet.* (2004) 36:1207–12. doi: 10.1038/ng1451
99. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell.* (2017) 168:613–28. doi: 10.1016/j.cell.2017.01.018
100. Greaves M, Maley CC. Clonal evolution in cancer. *Nature.* (2012) 481:306–13. doi: 10.1038/nature10762
101. Fatimata MAD, Malick F, Diop G, Mbengue B, Diallo RN, Niang MS, et al. Genetic diversity of breast cancer in senegalese women: new insight from somatic mutations. *J Health Sci.* (2014) 4:25–33. doi: 10.5923/j.health.20140402.01
102. Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson 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) 70:1152–71. doi: 10.1086/339933
103. Rohan TE, Wong LJ, Wang T, Haines J, Kabat GC. Do alterations in mitochondrial DNA play a role in breast carcinogenesis? *J Oncol.* (2010) 2010:604304. doi: 10.1155/2010/604304
104. Ferreira A, Serafim TL, Sardao VA, Cunha-Oliveira T. Role of mtDNA-related mitopeigenetic phenomena in cancer. *Eur J Clin Invest.* (2015) 45(Suppl. 1):44–9. doi: 10.1111/eci.12359
105. Weiss H, Wester-Rosenloef L, Koch C, Koch F, Baltrusch S, Tiedge M, et al. The mitochondrial Atp8 mutation induces mitochondrial ROS generation, secretory dysfunction, and beta-cell mass adaptation in conplastic B6-mtFVB mice. *Endocrinology.* (2012) 153:4666–76. doi: 10.1210/en.2012-1296
106. Lee J, Ding S, Walpole TB, Holding AN, Montgomery MG, Fearnley IM, et al. Organization of subunits in the membrane domain of the bovine F-ATPase revealed by covalent cross-linking. *J Biol Chem.* (2015) 290:13308–20. doi: 10.1074/jbc.M115.645283
107. Grzybowska-Szatowska L, Slaska B, Rzymowska J, Brzozowska A, Florianczyk B. Novel mitochondrial mutations in the ATP6 and ATP8 genes in patients with breast cancer. *Mol Med Rep.* (2014) 10:1772–8. doi: 10.3892/mmr.2014.2471
108. Budczies J, Bockmayr M, Denkert C, Klauschen F, Lennner JK, Gyorffy B, et al. Classical pathology and mutational load of breast cancer - integration of two worlds. *J Pathol Clin Res.* (2015) 1:225–38. doi: 10.1002/cjp.225
109. Roy D, Cai Q, Felty Q, Narayan S. Estrogen-induced generation of reactive oxygen and nitrogen species, gene damage, and estrogen-dependent cancers. *J Toxicol Environ Health B Crit Rev.* (2007) 10:235–57. doi: 10.1080/15287390600974924
110. Sarmiento-Salinas FL, Delgado-Magallon A, Montes-Alvarado JB, Ramirez-Ramirez D, Flores-Alonso JC, Cortes-Hernandez P, et al. breast cancer subtypes present a differential production of reactive oxygen species (ROS) and susceptibility to antioxidant treatment. *Front Oncol.* (2019) 9:480. doi: 10.3389/fonc.2019.00480
111. Kazdal D, Harms A, Endris V, Penzel R, Oliveira C, Kriegsmann M, et al. Subclonal evolution of pulmonary adenocarcinomas delineated by spatially distributed somatic mitochondrial mutations. *Lung Cancer.* (2018) 126:80–8. doi: 10.1016/j.lungcan.2018.10.024
112. Hopkins JF, Denroche RE, Aguiar JA, Notta F, Connor AA, Wilson JM, et al. Mutations in mitochondrial DNA from pancreatic ductal adenocarcinomas associate with survival times of patients and accumulate as tumors progress. *Gastroenterology.* (2018) 154:1620–4. doi: 10.1053/j.gastro.2018.01.029
113. Lin JC, Wang CC, Jiang RS, Wang WY, Liu SA. Impact of somatic mutations in the D-loop of mitochondrial DNA on the survival of oral squamous cell carcinoma patients. *PLoS One.* (2015) 10:e0124322. doi: 10.1371/journal.pone.0124322
114. Damm F, Bunke T, Thol F, Markus B, Wagner K, Gohring G, et al. Prognostic implications and molecular associations of NADH dehydrogenase subunit 4 (ND4) mutations in acute myeloid leukemia. *Leukemia.* (2012) 26:289–95. doi: 10.1038/leu.2011.200
115. Thyagarajan B, Wang R, Nelson H, Barcelo H, Koh WP, Yuan JM. Mitochondrial DNA copy number is associated with breast cancer risk. *PLoS One.* (2013) 8:e65968. doi: 10.1371/journal.pone.0065968
116. Campa D, Barrdahl M, Santoro A, Severi G, Baglietto L, Omichessan H, et al. Mitochondrial DNA copy number variation, leukocyte telomere length, and breast cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Breast Cancer Res.* (2018) 20:29. doi: 10.1186/s13058-018-0955-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Pérez-Amado, Tovar, Gómez-Romero, Beltrán-Anaya, Bautista-Piña, Domínguez-Reyes, Villegas-Carlos, Tenorio-Torres, Alfaro-Ruiz, Hidalgo-Miranda and Jiménez-Morales. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.