

Article

Mitochondrial Control Region Variants Related to Breast Cancer

Luis Enrique Hernández ¹, Jorge Hermilo Vega Avalos ¹, Laura Yareni Zuñiga ², María Guadalupe Sánchez-Parada ^{3,4}, Ana Elizabeth González Santiago ^{3,4}, Luis Miguel Román Pintos ^{2,4}, Rolando Castañeda Arellano ^{3,4}, Luis Daniel Hernández-Ortega ^{3,4}, Arie Roldán Mercado-Sesma ^{2,4}, Felipe de Jesús Orozco-Luna ⁵ and Raúl C. Baptista-Rosas ^{2,4,6*}.

¹ Programa de Licenciatura de Médico Cirujano y Patero, Centro Universitario de Tonalá, Universidad de Guadalajara, México

² departamento de Ciencias de la Salud-Enfermedad como Proceso Individual. Centro Universitario de Tonalá, Universidad de Guadalajara, México

³ Departamento de Ciencias Biomédicas, Centro Universitario de Tonalá, Universidad de Guadalajara, México

⁴ Centro de Investigación Multidisciplinaria en Salud, Centro Universitario de Tonalá, Universidad de Guadalajara, México.

⁵ Centro de Análisis de Datos y Supercómputo, Universidad de Guadalajara, México

⁶ Hospital General de Occidente, Secretaría de Salud Jalisco, México

* Correspondence: raul.baptista@academicos.udg.mx

Abstract: Breast cancer has an important incidence in the worldwide female population. Although alterations in the mitochondrial genome probably play an important role in carcinogenesis, the actual evidence is ambiguous and inconclusive. Our purpose was to explore mitochondrial sequences of cases with breast cancer from different origins and determine the associated polymorphisms. We identified 124 mtDNA sequences associated to breast cancer cases of which 86 were complete and 38 were partial sequences. Of these 86 complete sequences, 52 belonged to patients with a confirmed diagnosis of breast cancer, and 34 sequences were obtained from healthy mammary tissue of the same patients used as controls. From the mtDNA analysis, two polymorphisms with significant statistical differences were found: m.310del (rs869289246) in 34.6% (27/78) of breast cancer cases and 61.7% (21/34) in the controls; m.315dup (rs369786048) in 60.2% (47/78) of breast cancer cases and 38.2% (13/34) in the controls. In addition, the variant, m.16519T>C (rs3937033), was found in 59% of the control sequences and 52% of the breast cancer sequences with a significant statistical difference. Polymorphic changes are evolutionarily related to the haplogroup H of Indo-European and Euro-Asiatic origins; however, they were found in all non-European breast cancer.

Keywords: breast cancer; polymorphism; mitochondrial genomic; D310

1. Introduction

Breast cancer has ranked first among malignant diseases for more than a decade, and it is one of the leading causes of death in the female population, mainly impacting age groups between the fourth and sixth decade of life and affecting all socioeconomic levels [1]. Mortality rates for this neoplastic disease have shown a significant increase over the last five decades. Between 1955 and 1960, the rate was approximately 2 to 4 deaths per 100,000 women, rising steadily in adult women of all age groups but with a greater impact on women 30 years of age or older, until reaching a figure close to 9 per 100,000 inhabitants in the mid-1990s, remaining stable since. Short-term estimates suggest that this same trend will continue, with an increase in the number of cases [2].

The actual evidence indicates that breast cancer is a multifactorial disease with modifiable and nonmodifiable risk factors. Obesity, postmenopausal state, smoking, and physical inactivity are the most important modifiable factors. Inherited genetic autosomal mutations have been associated in ~10% of total cases of breast cancer. On the other hand,

pregnancy at an early age and breastfeeding for at least six months have a protective effect [3]. In Latin America, there are important regional differences regarding the incidence and prevalence, with a higher frequency of breast cancer in areas with the highest socioeconomic incomes compared with indigenous areas, where the socioeconomic level is lower [4,5]. This evidence is compatible with the epidemiology of breast cancer worldwide, where it affects developed countries with greater importance and has lower incidence and prevalence rates in third world populations.

In addition to socioeconomic and environmental components, several authors have mentioned its relationship with genetic variants in human populations, with special attention paid to the different mitochondrial haplogroups. However, there is a lack of evidence of mitochondrial haplogroups with sporadic [6] or familial hereditary breast cancer [7].

Many authors have suggested that alterations in mitochondrial DNA (mtDNA) play an important role in carcinogenesis [8–16] involving the control region, due to the fact that it contains sequences essential for transcription and replication. Previous studies have suggested that polymorphisms in these noncoding sequences of the control region could play an important role in breast cancer pathogenesis [17–21]. These changes in the sequence can be associated with a specific phenotype and serve as markers for the development of various malignancies [22]. However, some studies suggest that different mitochondrial polymorphisms vary from one population to another, and the available evidence suggests that some polymorphisms are more related to the presence of breast cancer than others [23–25].

Previous works have focused on variants associated with coding regions related to mitochondrial metabolism. Among the most cited in literature is the polymorphism m.10398A>G, related to the synthesis of the protein NADH-ubiquinone oxidoreductase 3 (ND3), which has been identified as a biomarker in different populations such as in Polish [26], Indian [22,27], Chinese [24,28], and, mainly, Afro-descendant groups [22,29–31]. This polymorphism has also been associated with metabolic syndrome and mental disorders in populations of Asiatic origin [32]. However, in a meta-analysis, it was determined that there was no association when analyzing this polymorphism individually, without correlating it with other mitochondrial polymorphisms in women affected by this malignant disease [33]. Likewise, the frequency of two novel polymorphisms in the *D-loop* of the control region, m.16290insT and m.16293delA, was statistically more prevalent in breast cancer patients than in control subjects (position 16290: OR = 6.011, 95% CI = 1.2482–28.8411, $p = 0.002$; position 16293: OR = 5.6028, 95% CI = 1.4357–21.8925, $p = 0.010$). In addition, previous observations found one novel mutation in the ND3-coding region at position 10316 (A > G) in 69% of breast cancer patients but not in the control subjects. The study suggests that two novel polymorphisms in the *D-loop* may be candidate biomarkers for breast cancer diagnosis in Bangladeshi women [23].

Although in Caucasian American groups of European descent, the variant m.10398A>G is related with an increased risk of breast cancer, other polymorphisms with more statistical significance have been identified such as m.16519T>C located in the control region [33–38].

Other studies have focused on m.16519T>C [18,29,30,32,36], particularly mutations in the *D-loop* due to the fact of the increased risk of breast cancer, occurring either singularly or in association with other mitochondrial protein-coding genes alterations such as m.10398A>G, m.13368G>A, or m.14766C>T [32,36]. Moreover, the association of several variants resulted in a significant predictive breast cancer factor. Indeed, m.10398A>G, together with other mutations, such as m.4216T>C, m.9055G>A, m.12308A>G, or m.16519T>C, is considered to increase the risk of developing breast cancer in women [32,37].

Despite these promising results, as well as the fact that most of these mtDNA polymorphisms have functional consequences, associations between specific polymorphisms and the risk of cancer have been the subject of intense debate. Several studies involving associations of specific polymorphisms with cancer risk have been thoroughly analyzed

due to the fact that are characterized by erroneous experimental designs, misinterpretations, and low-quality data [39,40]. Previous studies, carried out through the analysis of data published in case-control studies of breast cancer with a phylogenetic-based approach and using complete genome sequences and partial sequences that corresponded mainly to segments from the region of control, found inconsistencies and contradictions in the nonsynonymous polymorphism, m.10398A>G, in breast cancer, mainly in the selection of control groups as well as in the use of inadequate statistics and errors in the nomenclature [6,40,41]. However, due to the fact of its potential usefulness as a diagnostic tool, the study of mitochondrial DNA and its relation to cancer must remain an important focus of oncological biomarker research using adequate study designs, population stratification, and independent replication of the results.

The purpose of the present work was to explore mitochondrial sequences of clinical cases of breast cancer diagnosis, by reusing the available information in the public, free-access database, *GenBank*, and determine the prevalence of polymorphisms associated with this neoplastic disease.

2. Materials and Methods

The search for complete and partial mtDNA sequences was performed using the National Center for Biotechnology Information's (NCBI) *GenBank*, focusing on those stored in the repository *Nucleotide* database (<https://www.ncbi.nlm.nih.gov/nucleotide>). In searching for complete chromosomes, fragments smaller than 15,400 base pairs were not considered. Sequences of a small size were obtained, considering fragments no greater than 1500 bp.

The search strategy was carried out using the keywords "*Homo sapiens*", "*mitochondrion*", and "*Breast Cancer*", using *Booleans* and filters to select the results and excluding the term "*ancient human remains*" to facilitate the search in the database. As control sequences, using the same approach but adding the keywords "*Eskimo*" and "*Inuit*", which is considered to be a low-prevalence population of breast cancer. Moreover, we considered that in some identified breast cancer mtDNA sequences, it was possible to identify control individuals used in those previous works and even healthy breast mtDNA sequences obtained from the contralateral breast of the same patients.

Once the sequences were identified in the database, the accompanying metadata were analyzed to validate the diagnosis of breast cancer and eliminate those that did not correspond or did not specify the presence of the neoplastic disease. From the selected sequences, the reference of the work appointment was obtained to evaluate the end date of the sampling and its objectives, which were contained in the experimental design of each particular study.

Using the sequence identification numbers of *GenBank*, the complete sequences were obtained in *FASTA* format files. Subsequently, the *FASTA* files were used as inputs fed into the University of California Santa Cruz Genome Browser (<http://genome.ucsc.edu/> [42]) and the sequences were aligned against the *revised Cambridge Reference Sequence (rCRS)* with the latest available version (*GRCh38/hg38 Assembly*). The *FASTA* files were aligned using *BLAT*, a *BLAST*-like alignment tool [43], and custom sequence tracks were created in the browser with the mtDNA obtained from breast cancer cases, allowing us to observe the variants in a general way and from a panoramic perspective of the mitochondrial genomic landscape.

The haplotification and accounting of polymorphisms were carried out using the MITOMASTER application (<https://www.mitomap.org/foswiki/bin/view/MITOMASTER>), building a database using a *.csv* file for tabulation and statistical analysis of frequencies. The criteria for the classification of the different haplogroups can be found in the *Phylotree* database (<http://www.phylotree.org/>) [44]. Once the haplotype was performed and the variants in the sequences identified, a database was constructed to quantify the haplogroups, haplotypes, and main subclades in the population analyzed with breast cancer.

To evaluate the presence of phantom polymorphisms and discordance among the different haplogroups, the *Haplogrep* package was used (available at <http://haplogrep.uibk.ac.at/>). Furthermore, using it as a reference, the sequence NC_012920, deposited in *GenBank* as an *rCRS* [27], and with the tool *Haplogrep 2* (v.2.1.19) accessible in the same platform, we constructed a graphical phylogenetic tree to explore the population structure and identify possible phantom polymorphisms not assigned to the resulting haplogroup. All variants with a score of <3 and which were found in at least two different sequences according to the criteria of Soares et al. [45] were considered.

Only complete mtDNA sequences were analyzed, and a phylogenetic tree was generated, graphically representing the presence of polymorphisms associated to the various haplogroups, global and local private mutations, and regressive mutations as well as the expected loss of polymorphisms for each haplogroup.

For the analysis of the structure of the populations, a single file with all of the sequences in the FASTA format was used as an input for the multiple sequence alignment application with *CLUSTAL W*, using the *Molecular Evolutionary Genetics Analysis (MEGA)* package (<https://www.megasoftware.net/>). Once aligned, the tools, including the mtDNA nomenclature, mtSNP profile, and mtSNP frequency, in the package *mtDNAprofler* were used (<http://mtprofler.yonsei.ac.kr/>). From the use of these tools, changes in the sequences were identified as well as insertion and deletion sites. In addition, all of the sequence alignments were reviewed to detect polymorphic heteroplasmic sites.

The number of total polymorphic loci (mtSNPs), the number of fixed differences, the polymorphic and monomorphic mutations among the populations, the shared polymorphisms, and the average number of different nucleotides between populations were also estimated. In this way, the identification of the most common polymorphisms in the control region and in the coding region of the mtDNA was facilitated.

Specific variants were searched for in the regions associated with MDP mitochondria-derived proteins in the region of the MT-RNR1 gene that transcribes for 12s RNA located at the genomic coordinates chrM:648-1,601 (the MOTS-c sequence of 51 bp with localization at chrM:1,343-1,393) as well as the variants in the MT-RNR2 gene, which transcribes for 16s RNA at the coordinates chrM:1,669-3,231 (*Humanin* of 75 bp with localization at chrM:2,633-2,707; *SHLP1* of 75 bp with localization at chrM:2,485-2,559; *SHLP2* with 81 bp with localization at chrM:2,088-2,168; *SHL3* with 117 bp with localization at chrM:1,703-1,819; *SHLP4* with 81 bp with localization at chrM:2,442-2,522; *SHLP5* with 75 bp with localization at chrM:2,485-2,559; *SHLP6* with 63 bp with localization at chrM:2,990-3,052).

To evaluate its correlation, the breast cancer group's frequency was compared with selected polymorphisms in breast cancer mtDNA sequences and controls; the X^2 test was performed using *R v.4.2.0* (<https://cran.r-project.org/>) and *RStudio v.2022.02.2+485* (<https://www.rstudio.com/>).

3. Results

From the search of the *NCBI GenBank* database using the previously defined criteria, 124 mtDNA sequences were identified, where 90 complete mtDNA (16.5 kb) sequences were identified, 52 belonged to patients with a confirmed diagnosis of breast cancer, and 38 sequences were used as controls (Table S1). These control sequences were distributed in three categories: 9 distant normal tissue sequences, 10 paracancerous normal tissue sequences obtained from Wang et al. [46], and 15 sequences obtained from cells identified as normal under laser capture microdissection from biopsies of patients diagnosed with invasive breast carcinoma reported in the work of Fendt et al. [47]. Another five complete sequences were obtained from Hernandez de la Cruz et al. and two sequences from Imanishi et al. (2011) by direct submission to the *NCBI GenBank* (Table S1).

There were 38 partial sequences (i.e., smaller than 0.5 kb) of breast cancer cases, of which 12 sequences covered the HVR1 region, and 24 sequences were from the HVR2

region of the mitochondrial chromosome. It is important to note that 48.3% (i.e., 60 sequences, where 45 were obtained from breast cancer tissue and 15 were controls) of the sequences were reported in Europe by Gasparre et al. [48] and Fendt et al. [47] (Table S1). Only 23.3% (i.e., 29 sequences, where 10 were obtained from breast cancer tissue, and 19 were controls) were of Chinese origin, reported by Wang et al. [46], which explains why the haplotyping of the sequences was related with Caucasians via the haplogroup H (Table S2).

Regarding the partial sequences, 14 sequences from the HVR1 region were published by Ghatak et al. [49], and the 24 remaining sequences (12 were from the HVR1 region, and the other 12 were from the HVR2 region) were associated with direct referrals without related publications and provided by Darvishi et al. [50] (Table S1).

The number of total polymorphic loci, without considering potential phantom variants, was analyzed, and near half of the variations were in the control region with an average of 10.7. The results of haplotyping in the population with breast cancer showed different clades and subclades distributed in 15 haplogroups (Supplementary Materials Table S2 and Figure S1). Seven possible phantom polymorphisms were identified in the forty-seven complete mtDNA sequences, and they were excluded from the subsequent analyses [51] (Table S3).

After haplotyping and navigation in the *UCSC genome browser*, we found 214 different polymorphisms that shared sequences, which we used as controls, and 241 unique polymorphisms were found only in breast cancer, while only 1 polymorphism, A3480G, was found exclusively in a control sample (Supplementary Materials Figure S2).

The cumulative analysis showed that most of the variations in the breast cancer sequences were unique polymorphisms (Figure 1). Only 13 mutations were related in almost 50% of the sequences studied. In contrast, in the control sequences, there were up to 22 different polymorphisms. Most of the variations were found in unique polymorphisms distributed along the mitochondrial genome (Supplementary Materials Figure S5); however, the most frequent repetition was concentrated in the control region, between positions 576 and 16024 of the mtDNA (Supplementary Materials Figure S4).

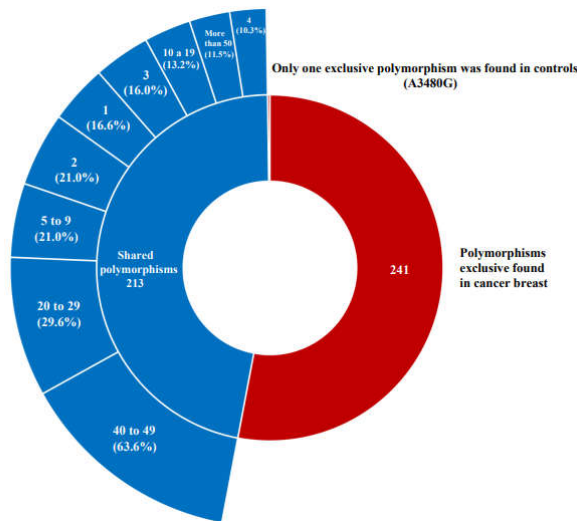


Figure 1. The frequency of the variant repetitions among the different complete sequences obtained from individuals diagnosed with breast cancer and from the control groups.

The most frequently shared polymorphisms were m.310del (rs869289246), found in 61.7% (21/34), and m.16519T>C (rs3937033), found in 58.8% (20/34) of the controls, while m.315dup (rs369786048) was found in 60.2% (47/78) of the breast cancer cases, with a statistically significant difference ($p < 0.05$) (Table 1).

Table 1. Frequent polymorphisms in different mtDNA regions of the breast cancer cases and the controls.

mtDNA Region (Positions)	Polymorphism	NCBI dbSNP ID	Breast Cancer, %	Controls, %	X ²
HRV-II (57-372)	m.310del	rs869289246	34.6 (27/78)	61.7 (21/34)	0.007 *
	m.310T>C	rs1556422421	5.1 (4/78)	0 (0/34)	0.178
	m.310delinsCCTC	rS46492218	1.2 (1/78)	2.9 (1/34)	0.542
	m.315dup	rs369786048	60.2 (47/78)	38.2 (13/34)	0.030 *
MT-ND3 (10059-10404)	m.10398A>G	rs2853826	18.0 (9/50)	29.4 (10/34)	0.219
HRV-I (16024-16383)	m.16183A>C	rs28671493	12.8 (10/78)	14.7 (5/34)	0.787
	m.16189T>C	rs28693675	18.8 (17/90)	26.5 (9/34)	0.354
	m.16311T>C	rs34799580	12.2 (11/90)	23.5 (8/34)	0.118
	m.16519T>C	rs3937033	51.5 (34/66)	58.8 (20/34)	0.007 *

* The difference was statistically significant between the breast cancer cases and the controls at $p < 0.05$

Moreover, 32 polymorphisms were present in two sequences and 60 sequences with a single polymorphism were present in each sequence. In m.16147C>T, there were two different polymorphisms: a transition C/T in three sequences and C/A in one.

When we compared the relative frequencies of each sequence, we observed that there was a significant proportion of exclusive polymorphisms in the breast cancer cases that were practically absent in the control samples (Supplementary Materials Figure S5).

The most frequent variants in the breast cancer cases, with a lower incidence in the controls, were m.3010G>A, m.16311T>C, m.16189T>C, and m.16519T>C; however, m.16519T>C, the most important polymorphism, represented at least half of the total sequences shared. It is an important point of focus that m.16189T>C and m.16311T>C were positioned in HVR1 and that m.16519T>C was placed in a noncoding position in the control region. Of note, at position 16189, the sequences HG825995.1, HG826004.1, and GU592040.1 had a deletion that was not registered in the dbSNP database, in addition to an insertion in the sequence AB626609.1 (variant m.16189delinsCC with the dbSNP ID rs369574569).

From the total 86 mtDNA complete sequences obtained from patients diagnosed with breast cancer, six common polymorphisms were identified in the sequences m.263A>G, m.750A>G, m.1438A>G, m.4769A>G, m.8860A>G, and m.15326A>G, representing more than 98% of the analyzed sequences obtained from malignant tissue. However, they were also found in most of the control samples without statistical differences.

The polymorphism, m.10398A>G (rs2853826), usually associated with breast cancer in other studies, was found only in 18.0% (9/50) of the breast cancer mtDNA sequences analyzed as well as in 29.4% (10/34) of the control samples, without statistical significance (Table 4). The m.10398A>G variant was found in the coding region of the MT-ND3 gene that codes for subunit ND3 of complex I (NADH dehydrogenase), although no significant differences were identified between the cases with malignant neoplasia and their respective controls (Table 4). The frequencies of m.315dup (rs369786048) and m.16519T>C were the most recurrent polymorphisms found in our analysis, either in isolation or both in the same sequence, with statistical significance (Table 4 and Supplementary Materials Figure S5).

In terms of heteroplasmic changes in mtDNA sequences, our search showed 22 heteroplasmic changes distributed in only 17 complete sequences (Tables S4 and S5). On the other hand, partial sequences presented only seven heteroplasmic changes in only four

sequences, all obtained by laser capture microdissection (LCM). These changes were not identified in the rest of the sequences with heteroplasmic somatic transitions described previously by Wang et al. (2007). The polymorphism, m.2275T>C, was found in the primary cancerous tissue identified as EF114285.1 but not in other normal tissues. The polymorphism, m.8601A>G, was related to the primary cancerous sequence EF114276.1 and the paracancerous normal sample EF114277.1, obtained from cells without dysmorphic or malignant changes under macro- and microscopic analysis but with polymorphic changes related to the neoplastic genotype.

Other common findings were polymorphisms in the control region that clearly showed a characteristic pattern related to the presence of one polymorphism and the absence of others (Supplementary Materials Figure S5). We were able to identify that in 47 mitochondrial sequences analyzed, all had alterations in positions 309 to 315, characterized by only three situations: the first, manifested by a deletion (m.310del with the dSNP ID rs869289246); the second, a variation characterized by the insertion of three bases in m.310delinsCCTC (rs46492218); the third, characterized by the translocation of a thymine for cytosine m.310T>C (rs1556422421). An interesting observation was that the changes in m.310T>C (34.6% of cases; 27 of the 78 sequences) and m.315dup (60.2% of the cases; 47 of the 78 sequences) (Table 2) were observed in almost half of the sequences of non-European origin.

When the variant m.315dup was absent, we identified the presence of the polymorphism m.310del. Even though the most frequent polymorphism in position 310 was m.310del, our analysis found m.310T>C in 4 of the 78 sequences (5.1%) and m.310delinsCCTC in 1 of the 78 sequences (1.2%). After the statistical analysis using the χ^2 test, m.310del and m.315dup were significantly different when comparing cases and controls; we also ruled out a potential artifact associated with ghost mutations (Tables S3 and S4).

Variations in the repeat sequences between nucleotides 303 and 316-318 were defined by D310, although they were related with oxidative stress and resulted in a compensatory increase in the mtDNA copy number of homopolymeric C repeats in the noncoding mitochondrial control region. These variants in D310 were present in 98% of the analyzed sequences. Although most of these polymorphisms were commonly associated with specific haplogroups, it was revealed that the m.315dup polymorphism was evolutionarily associated with haplogroup H of Caucasian origin and infrequent in Indo-European haplogroups, with a frequency lower than 1.5%. In our analysis, 99% of sequences of non-European origin with breast cancer (Figures S4 and S5) were present.

Regarding the variants found in MPD, we identified 11 polymorphisms of which 5 were not reported in the NCBI dbSNP database, but none of them were significantly different between the group with breast cancer and the control sequences (Table 2).

Table 2. Polymorphisms in mitochondrial-derived protein sequences related to breast cancer cases and the controls.

mtDNA Region Genes (positions)	Polymorphism	NCBI <i>dbSNP</i> ID	Breast Cancer, %	Controls, %	X ²
MT-RNR1					
MOTS (1343-1393)	m.1346A>G	rs879104061	1.9 (1/52)	0 (0/34)	0.416
MT-RNR2					
Humanin (2633-2707)	m.2706A>G	rs2854128	61.5 (32/52)	76.4 (26/34)	0.418
SHLP1 (10059-10404)	-	-	-	-	-
SHLP2 (2088-2168)	m.2124A>G	-	1.9 (1/52)	0 (0/34)	0.416
	m.2145del	-	1.9 (1/52)	0 (0/34)	0.416
	m.1709A>G	rs200251800	3.8 (2/52)	11.8 (4/34)	0.158
	m.1716A>G	rs1556422559	1.9 (1/52)	5.9 (2/34)	0.327
SHLP3 (1703-1819)	m.1719A>G	rs3928305	13.4 (7/52)	5.9 (2/34)	0.261
	m.1736A>G	rs193303006	3.8 (2/52)	11.8 (4/34)	0.158
	m.1811A>G	rs28358576	1.9 (1/52)	2.9 (1/34)	0.759
SHLP4 (2442-2522)	m.2445A>G	-	1.9 (1/52)	0 (0/34)	0.416
SHLP5 (2780-2854)	-	-	-	-	-
SHLP6 (2990-3052)	m.3010A>G	rs3928306	21.1 (11/52)	17.6 (6/34)	0.689

4. Discussion

In humans, the mutation rate of the mitochondrial genome is at least ten times higher than that of the nuclear genome. Most mutations in mtDNA accumulate in the displacement loop, called the *D-loop*, in the noncoding control region. The *D-loop* functions as a promoter for both heavy and light strands of mtDNA and, similar to the control region, does not encode any functional peptides.

D310, a term that defines these changes, previously assigned in [52], was described in 2001 by Sanchez-Cespedes et al. [53] in several types of tumors and, apparently, it is the most frequent hotspot of mitochondrial genome somatic mutations in various cancers located in the *D-loop* region. D310 refers to a mononucleotide repeat sequence between positions 303–309 and 316–318 of mtDNA with variable lengths of cytosine repetitions, described as a polycytosine stretch, poly-C repeat, or C-tract. This poly-C repeat, located between the 303 and 315 nucleotides, is usually interrupted by a T base at position 310 and has been reported as a hotspot for deletion or insertion mutations in primary tumors [53–56].

This length of cytosine bases is highly polymorphic, ranging from seven to nine nucleotides; however, usually seven cytosines are the most common finding. The D310 repeat is a part of conserved sequence block II of the hypervariable region II, located 92 bp away from the heavy-strand replication region, and mitochondrial chromosome replication is initiated when this loci forms a persistent RNA–DNA hybrid that results in the initiation of mtDNA heavy-strand replication with the conserved sequence block I and conserved sequence block III regions [56–58]. Previous genomic analysis revealed that only a homopolymeric sequence in conserved sequence block II, ranging from six to twelve nucleotides, showed high variability. Transcriptional analyses revealed that most of the common polymers can support accurate transcriptional initiation [59].

Alterations in the mtDNA replication rate can result from a deletion or insertion in the number of C residues in the poly-C repeat, which may interfere with the binding of

mtDNA polymerase γ and other trans-acting elements. It is believed that D310 plays a vital role in preserving the number of mtDNA contents. Because the control region is a susceptible region to oxidative stress, many researchers have hypothesized that the D310 polymorphisms are related to an increased stress status.

Apparently, there was no relationship between the different haplogroups associated with risk of or protection against breast cancer. However, when we analyzed the sequence information, it was possible to observe polymorphisms commonly related to cancer. It is important to note that there is a statistical issue related to the fact that most of the samples were of European origin, as they comprised close to half of the population with breast malignant disease (44.9%), they were more similar to the sequences used as the *rCRS* reference genome in the database and genomic browsers.

Regarding the presence of phantom polymorphisms, defined as systematic artifacts generated during the sequencing process, and despite the common belief that these artificial changes in the sequence are almost ubiquitous in the results, their frequency and impact on the interpretation of the results can vary dramatically [49]. In mtDNA, they generate an access point pattern quite different from that of natural polymorphisms, which were found in the analyzed sequences [45]; therefore, it is important to identify them before making diagnostic inferences or related hypotheses.

Rejecting false polymorphisms as artifacts, taking the phylogenetic knowledge into account, the procedure incorporated all variants with frequency scores based on 2,196 complete mitochondrial genomes as references [26]. All of the remaining variants in a sample were annotated, and if a variant occurred with a frequency of less than three and at least two samples shared this change, then it was listed as a rare polymorphism in the report. The reason for this threshold is that a known phantom variant with score of two is, in fact, also present within the rare polymorphism list. Therefore, this procedure allows for the identification of potential phantom changes in the sequences [26].

Otherwise, of the 241 polymorphisms exclusive to the neoplastic tissue, only 9 were repeated in more than 3 different sequences in our analysis: m.16183A>C repeated in 11 different sequences; m.16217T>C in 5; m.16298T>C in 4. The rest of the polymorphisms were repeated three times in 6 sequences, twice in 22 sequences, and only once in 193 sequences. Some of these changes in the mtDNA sequences were related in the biomedical literature to other types of malignant neoplastic diseases [46].

In a previous study, we identified that the polymorphism, m.315dup, was not associated, in the literature, with breast cancer. However, it had as a reference two samples of medium prevalence, which is very low for neoplastic disease [60]. There was a change in the sequence m.315dup, where cytosine was inserted between positions 311 and 315, compared to the *rCRS*; thus, it was translated into these sequences, and the *rCRS* was larger than many other sequences in the mitochondrial genomes (five or six nucleotides at this position).

The polymorphism, m.315dup, is considered to be one of the most recent mutations, occurring in the last 60,000 years. This is a type of mutation, now well accepted and detailed in many of the phylogenetic trees, which is related to haplogroup H. Although this haplogroup is related to populations of ethnic origin, it has been found to have a high prevalence in Latin American mestizos and in other populations with different haplogroups [60–65].

The m.315dup and m.310T>C polymorphisms, which were found in the noncoding HVR2 fragment of the *D-loop* in the mtDNA control region, are considered as access points due to the high frequency of changes. Furthermore, the modification of these were allowed and may have implications for the appropriate transcription and regulation of mitochondrial genome expression. This poly-C hotspot area in D310 is considered very polymorphic and can be different between direct relatives by the maternal line [45,66–68], and due to its high prevalence in Caucasian groups, m.315dup is usually not considered during bioinformatic analysis, and it is not used in the construction of phylogenetic trees, because it is usually identified. However, there are reports in the literature where the m.315insC polymorphism is associated with various forms of cancer and other chronic

degenerative diseases [50,67]. This poly-C tract in the mitochondrial *D-loop*, located commonly between the 303 and 315 nucleotides, was identified as a frequent *hotspot* mutation region in human neoplasia, including breast cancer [61,62], suggesting that mtDNA instability at this site may be a common characteristic in this malignant disease.

The alteration hallmarks of D310, found in all of the analyzed malignant sequences and identified in only two related options, were m.315dup and m.310del, but in neither case were both in the same sequence, and both polymorphisms are evolutionarily related to haplogroup H of Caucasian, Indo-European, and Euro-Asiatic origins. However, that they were found in all of the non-European origin samples with breast cancer could be a clue for future research to explore. Our analysis of D310 identified more frequent variants, including m.315dup and other variants, such as m.309dup (rs878871521), related to insertions of C in positions 303 to 309 (a change of seven to eight cytosines); m.309del (rs1556422420), related to a change of CTC by only one cytosine, with a deletion of CT (left-shifted) or deletion of TC (right-shifted); m.310T>C (rs1556422421) and m.310del (rs869289246), with the loss of a nucleotide.

The other important finding in our analysis was the presence of m.16519T>C, found in nearly two-thirds of the control sequences and half of the breast cancer sequences with significant statistical differences. Other studies have focused on this polymorphism, apparently, without relation with other variants previously described in the available literature [21,30–33,37]. Moreover, the association of several variants resulted in a significant predictive breast cancer factor. Indeed, m.10398A>G, usually associated with breast cancer in other studies, [24,27,29–32,49], was found in only 24% of the mtDNA sequences analyzed without statistical significance, and in association with other mutations, such as m.4216T>C, m.9055G>A, m.12308A>G, or m.16519T>C, it was found to increase the risk of a woman developing breast cancer [36].

m.16519T>C was associated in the *Phylotree* database between the branches of haplogroup L5 and haplogroup L2 with 11 other mutations (i.e., m.16311T>C, m.16189T>C, m.16187C>T, m.15301A>G, m.13506C>T, m.13105A>G, m.10810T>C, m.10688G>A, m.8655C>T, m.825T>A, and m.247G>A) back to its CRS value and usually found in Euro-Asiatic mitogenomes. These ancient polymorphisms arose 140,000 years ago and are considered part of the suggested *Mitochondrial Eve genome*, supporting the widely accepted theory related to first human migration outside of the African continent. Once again, the high number of mutations suggests there was a significant bottleneck in human evolution at the time, perhaps around 120,000 years ago, which might have lasted for many thousands of years.

Finally, the molecular mechanisms underlying the increased risk of cancer due to the presence of these specific mtDNA polymorphisms are still unclear. The control region is important for the regulation of mitochondrial genome replication and expression. The polymorphisms in this region might affect mtDNA replication and lead to electron transport chain alterations, resulting in compensatory increases in glycolytic ATP production. However, one of the inevitable products of these alterations is an increased release of highly reactive oxygen species, which may lead to mitochondrial abnormalities. These abnormalities invoke a mitochondria-to-nucleus retrograde response and, finally, result in nuclear genome damage, which contributes to initial events related to carcinogenesis [69–71]. The regulation of mitochondrial genome replication from the control region might also lead to mtDNA damage [72–74] and, with a critical number of mitochondrial genome changes, to cellular apoptosis [75], which finally could induce cancer development.

5. Conclusions

In conclusion, the changes in the mtDNA breast cancer were polymorphic and not associated with a few unique mutations. Most of the polymorphisms related to breast cancer were unique to each sequence. The combination of the position of the polymorphism provides some clues regarding mitochondrial alterations in breast cancer.

The polymorphism, m.10398A>G, usually associated with breast cancer in other studies, was found only in one-quarter of the mtDNA sequences analyzed, without statistical significance.

Some polymorphisms placed in the control region of the mitochondrial chromosome were present in the breast cancer sequences, but only m.16519T>C was related to the control sequences with statistical significance. Alterations in the poly-cytosine area, defined as D310, were found in all the analyzed malignant sequences. Only two variants related to changes in D310 were found, m.315dup and m.310del, but in neither case were both in the same sequence. These D310 polymorphisms are evolutionarily related to haplogroup H of Caucasian, Indo-European, and Euro-Asiatic origins. However, both were found in all non-European origins with breast cancer. Mitochondrial variants in D310 related with deletions/insertions and base changes mainly involved purine transitions, which were assumed to have occurred after the action of reactive oxygen species. Many frequent variants accumulated in the D-loop, a specialized sequence in the control region that functions as a promoter for both heavy and light strands of mitochondrial chromosomes. A mononucleotide repeat between nucleotides 303 and 316-318 was identified as a frequent hotspot for mutations in primary tumors and in our research.

Further analysis is required with the objective to evaluate more sequences and to calculate the correlation with the risk in the development of malignancy. This approach will provide us with a general perspective on the importance of consolidating the evidence in specialized repositories and that have not been crossed yet. Finally, it is important to include populations of different origins than Caucasian, with the purpose of showing the genetic differences related to breast cancer, regardless of high and low prevalence. With this focus, our understanding of these malignant diseases will improve through the interpretation of the complete mitogenome.

Supplementary Materials: Table S1: Origin and references of mtDNA sequences associated with breast cancer and controls used for the analysis; Table S2: Comparative summary of macrohaplogroups' distribution obtained from complete and partial mitochondrial sequences from patients diagnosed with breast cancer; Table S3: Comparative summary of the distribution of possible phantom polymorphisms in complete mtDNA sequences with breast cancer diagnosis. Table 1: Frequent polymorphisms in different mtDNA regions of breast cancer cases and controls; Table S5: Polymorphisms in mitochondrial-derived proteins sequences related to breast cancer cases and controls; Table S4: Summary of the heteroplasmic in mtDNA sequences by position in cases diagnosed with breast cancer; Table S5: Summary of the heteroplasmic polymorphisms detected by sequence and not related with haplotyping in the mtDNA sequences of cases diagnosed with breast cancer and controls; Figure S1: Phylogenetic tree of mtDNA complete sequences obtained from breast cancer tissue; Figure S2: Mitochondrial genome landscape of mtDNA sequences obtained from breast cancer tissue; Figure S3: Comparison of breast cancer and control polymorphisms frequency; Figure S4: Pareto chart of breast cancer polymorphisms; Figure S5: Proportional relative frequency by the position of breast cancer and control polymorphisms; Figure S4: Positions 315 and 309-310 in mtDNA shows common hallmark features found in breast cancer tissue and controls.

Author Contributions: Conceptualization, A.R.M.S. and R.C.B.R.; methodology, R.C.B.R. and A.R.M.S., software, F.J.O.L.; validation, R.C.B.R.; formal analysis and investigation, L.E.H., J. H.V.A. and R.C.B.R.; resources, L.M.R.P., R.C.A, L.D.H.O. and A.E.G.S.; data curation, R.C.B.R.; writing—original draft preparation, L.E.H. and J.H.V.A.; writing—review and editing, L.Y.Z., A.R.M.S. and R.C.B.R.; visualization, F.J.O.L. and R.C.B.R.; supervision, R.C.B.R.; project administration, A.R.M.S.; references, L.E.H., J.H.V.A., A.E.G.S. and L.Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: No funding was obtained.

Institutional Review Board Statement: No ethical approval was required.

Informed Consent Statement: Not required, as it did not involve humans, only public and open-source database mining.

Data Availability Statement: Sequence data are available at <https://www.ncbi.nlm.nih.gov/nucleotide/>. All the information used in this article is available in public databases at the links provided in Section 2.

Acknowledgments: We thank the Data Analysis and Supercomputing Center of the University of Guadalajara for the facilities for the research program of L.E.H. and J.H.V.A. and the space for coordination and teamwork. We appreciate the comments and recommendations of Robert Kuhn to improve the analysis using UCSC Genome Browser and Ian Capstick for manuscript style correction.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians* **2018**, *68*, 394–424, doi:10.3322/caac.21492.
- Carioli, G.; Malvezzi, M.; Rodriguez, T.; Bertuccio, P.; Negri, E.; La Vecchia, C. Trends and Predictions to 2020 in Breast Cancer Mortality: Americas and Australasia. *Breast* **2018**, *37*, 163–169, doi:10.1016/j.breast.2017.12.004.
- Rojas, K.; Stuckey, A. Breast Cancer Epidemiology and Risk Factors. *Clin Obstet Gynecol* **2016**, *59*, 651–672, doi:10.1097/GRF.0000000000000239.
- Cuevas, S.A.R.; García, M.C. Epidemiología del cáncer de mama. *Ginecol Obstet Mex* **2006**, *74*, 585–593.
- Sifuentes-Álvarez, A.; Castañeda-Martínez, L.Y.; Lugo-Nevarés, M.; Reyes-Romero, M.A. Factores de riesgo asociados con el cáncer de mama en mujeres del estado de Durango, México. *Ginecol Obstet Mex* **2015**, *83*, 662–669.
- Riley, V.; Erzurumluoglu, A.M.; Rodriguez, S.; Bonilla, C. Mitochondrial DNA Haplogroups and Breast Cancer Risk Factors in the Avon Longitudinal Study of Parents and Children (ALSPAC). *Genes* **2018**, *9*, 395, doi:10.3390/genes9080395.
- Gutiérrez Povedano, C.; Salgado, J.; Gil, C.; Robles, M.; Patiño-García, A.; García-Foncillas, J. Analysis of BRCA1 and MtDNA Haplotypes and MtDNA Polymorphism in Familial Breast Cancer. *Mitochondrial DNA* **2015**, *26*, 227–231, doi:10.3109/19401736.2013.825773.
- Mambo, E.; Chatterjee, A.; Xing, M.; Tallini, G.; Haugen, B.R.; Yeung, S.-C.J.; Sukumar, S.; Sidransky, D. Tumor-Specific Changes in MtDNA Content in Human Cancer. *Int. J. Cancer* **2005**, *116*, 920–924, doi:10.1002/ijc.21110.
- Birch-Machin, M.A. The Role of Mitochondria in Ageing and Carcinogenesis. *Clin Exp Dermatol* **2006**, *31*, 548–552, doi:10.1111/j.1365-2230.2006.02161.x.
- Brandon, M.; Baldi, P.; Wallace, D.C. Mitochondrial Mutations in Cancer. *Oncogene* **2006**, *25*, 4647–4662, doi:10.1038/sj.onc.1209607.
- Putignani, L.; Raffa, S.; Pescosolido, R.; Aimati, L.; Signore, F.; Torrisi, M.R.; Grammatico, P. Alteration of Expression Levels of the Oxidative Phosphorylation System (OXPHOS) in Breast Cancer Cell Mitochondria. *Breast Cancer Res Treat* **2008**, *110*, 439–452, doi:10.1007/s10549-007-9738-x.
- Rohan, T.E.; Wong, L.-J.; Wang, T.; Haines, J.; Kabat, G.C. Do Alterations in Mitochondrial DNA Play a Role in Breast Carcinogenesis? *Journal of Oncology* **2010**, *2010*, 1–11, doi:10.1155/2010/604304.
- Chandra, D.; Singh, K.K. Genetic Insights into OXPHOS Defect and Its Role in Cancer. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2011**, *1807*, 620–625, doi:10.1016/j.bbabi.2010.10.023.
- Yadava, N.; Schneider, S.S.; Jerry, D.J.; Kim, C. Impaired Mitochondrial Metabolism and Mammary Carcinogenesis. *J Mammary Gland Biol Neoplasia* **2013**, *18*, 75–87, doi:10.1007/s10911-012-9271-3.
- Jezierska-Drutel, A.; Rosenzweig, S.A.; Neumann, C.A. Role of Oxidative Stress and the Microenvironment in Breast Cancer Development and Progression. In *Advances in Cancer Research*; Elsevier, 2013; Vol. 119, pp. 107–125 ISBN 978-0-12-407190-2.
- Singh, R.K.; Srivastava, A.; Kalaiarasan, P.; Manvati, S.; Chopra, R.; Bamezai, R.N.K. MtDNA Germ Line Variation Mediated ROS Generates Retrograde Signaling and Induces Pro-Cancerous Metabolic Features. *Sci Rep* **2015**, *4*, 6571, doi:10.1038/srep06571.
- Parrella, P.; Xiao, Y.; Fliss, M.; Sanchez-Céspedes, M.; Mazzarelli, P.; Rinaldi, M.; Nicol, T.; Gabrielson, E.; Cuomo, C.; Cohen, D.; et al. Detection of Mitochondrial DNA Mutations in Primary Breast Cancer and Fine-Needle Aspirates. *Cancer Res* **2001**, *61*, 7623–7626.
- Tan, D.-J.; Bai, R.-K.; Wong, L.-J.C. Comprehensive Scanning of Somatic Mitochondrial DNA Mutations in Breast Cancer. *Cancer Res* **2002**, *62*, 972–976.
- Yu, M.; Shi, Y.; Zhang, F.; Zhou, Y.; Yang, Y.; Wei, X.; Zhang, L.; Niu, R. Sequence Variations of Mitochondrial DNA D-Loop Region Are Highly Frequent Events in Familial Breast Cancer. *J Biomed Sci* **2008**, *15*, 535–543, doi:10.1007/s11373-007-9229-4.
- Ye, C.; Shu, X.O.; Pierce, L.; Wen, W.; Courtney, R.; Gao, Y.-T.; Zheng, W.; Cai, Q. Mutations in the Mitochondrial DNA D-Loop Region and Breast Cancer Risk. *Breast Cancer Res Treat* **2010**, *119*, 431–436, doi:10.1007/s10549-009-0397-y.

21. Tommasi, S.; Favia, P.; Weigl, S.; Bianco, A.; Pilato, B.; Russo, L.; Paradiso, A.; Petruzzella, V. Mitochondrial DNA Variants and Risk of Familial Breast Cancer: An Exploratory Study. *International Journal of Oncology* **2014**, *44*, 1691–1698, doi:10.3892/ijo.2014.2324.
22. Singh, K.K.; Kulawiec, M. Mitochondrial DNA Polymorphism and Risk of Cancer. In *Cancer Epidemiology*; Verma, M., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 471, pp. 291–303 ISBN 978-1-58829-987-1.
23. Sultana, G.N.N.; Rahman, A.; Shahinuzzaman, A.D.A.; Begum, R.A.; Hossain, C.F. Mitochondrial DNA Mutations--Candidate Biomarkers for Breast Cancer Diagnosis in Bangladesh. *Chin J Cancer* **2012**, *31*, 449–454, doi:10.5732/cjc.012.10024.
24. Jiang, H.; Zhao, H.; Xu, H.; Hu, L.; Wang, W.; Wei, Y.; Wang, Y.; Peng, X.; Zhou, F. Peripheral Blood Mitochondrial DNA Content, A10398G Polymorphism, and Risk of Breast Cancer in a Han Chinese Population. *Cancer Sci* **2014**, *105*, 639–645, doi:10.1111/cas.12412.
25. Grzybowska-Szatowska, L.; Ślaska, B.; Rzymowska, J.; Brzozowska, A.; Floriańczyk, B. Novel Mitochondrial Mutations in the ATP6 and ATP8 Genes in Patients with Breast Cancer. *Molecular Medicine Reports* **2014**, *10*, 1772–1778, doi:10.3892/mmr.2014.2471.
26. Weissensteiner, H.; Pacher, D.; Kloss-Brandstätter, A.; Forer, L.; Specht, G.; Bandelt, H.-J.; Kronenberg, F.; Salas, A.; Schönherr, S. HaploGrep 2: Mitochondrial Haplogroup Classification in the Era of High-Throughput Sequencing. *Nucleic Acids Res* **2016**, *44*, W58–W63, doi:10.1093/nar/gkw233.
27. Francis, A.; Pooja, S.; Rajender, S.; Govindaraj, P.; Tipiriseti, N.R.; Surekha, D.; Rao, D.R.; Rao, L.; Ramachandra, L.; Vishnupriya, S.; et al. A Mitochondrial DNA Variant 10398G>A in Breast Cancer among South Indians: An Original Study with Meta-Analysis. *Mitochondrion* **2013**, *13*, 559–565, doi:10.1016/j.mito.2013.08.004.
28. Yu, Y.; Lv, F.; Lin, H.; Qian, G.; Jiang, Y.S.; Pang, L.X.; Wang, Y.P.; Wang, X.F.; Kang, Y.M.; Li, C.B.; et al. Mitochondrial ND3 G10398A Mutation: A Biomarker for Breast Cancer. *Genet. Mol. Res.* **2015**, *14*, 17426–17431, doi:10.4238/2015.December.21.12.
29. Canter, J.A.; Kallianpur, A.R.; Parl, F.F.; Millikan, R.C. Mitochondrial DNA G10398A Polymorphism and Invasive Breast Cancer in African-American Women. *Cancer Research* **2005**, *65*, 8028–8033, doi:10.1158/0008-5472.CAN-05-1428.
30. Mims, M.P.; Hayes, T.G.; Zheng, S.; Leal, S.M.; Frolov, A.; Ittmann, M.M.; Wheeler, T.M.; Prchal, J.T. Mitochondrial DNA G10398A Polymorphism and Invasive Breast Cancer in African-American Women. *Cancer Research* **2006**, *66*, 1880–1881, doi:10.1158/0008-5472.CAN-05-3774.
31. Setiawan, V.W.; Chu, L.-H.; John, E.M.; Ding, Y.C.; Ingles, S.A.; Bernstein, L.; Press, M.F.; Ursin, G.; Haiman, C.A.; Neuhausen, S.L. Mitochondrial DNA G10398A Variant Is Not Associated with Breast Cancer in African-American Women. *Cancer Genetics and Cytogenetics* **2008**, *181*, 16–19, doi:10.1016/j.cancergencyto.2007.10.019.
32. Kishida, K.; Tominaga, M.; Matsubara, K.; Taguchi, M.; Noguchi, M.; Tsunawake, N.; Shidoji, Y. An Association Analysis between Mitochondrial DNA A10398G Polymorphism and Temperament in Japanese Young Adults. *PLoS ONE* **2009**, *4*, e7763, doi:10.1371/journal.pone.0007763.
33. Czarnecka, A.M.; Krawczyk, T.; Zdrożny, M.; Lubiński, J.; Arnold, R.S.; Kukwa, W.; Ścińska, A.; Golik, P.; Bartnik, E.; Petros, J.A. Mitochondrial NADH-Dehydrogenase Subunit 3 (ND3) Polymorphism (A10398G) and Sporadic Breast Cancer in Poland. *Breast Cancer Res Treat* **2010**, *121*, 511–518, doi:10.1007/s10549-009-0358-5.
34. Verma, M.; Naviaux, R.K.; Tanaka, M.; Kumar, D.; Franceschi, C.; Singh, K.K. Meeting Report: Mitochondrial DNA and Cancer Epidemiology. *Cancer Research* **2007**, *67*, 437–439, doi:10.1158/0008-5472.CAN-06-4119.
35. Bai, R.-K.; Leal, S.M.; Covarrubias, D.; Liu, A.; Wong, L.-J.C. Mitochondrial Genetic Background Modifies Breast Cancer Risk. *Cancer Research* **2007**, *67*, 4687–4694, doi:10.1158/0008-5472.CAN-06-3554.
36. Covarrubias, D.; Bai, R.-K.; Wong, L.-J.C.; Leal, S.M. Mitochondrial DNA Variant Interactions Modify Breast Cancer Risk. *J Hum Genet* **2008**, *53*, 924–928, doi:10.1007/s10038-008-0331-x.
37. Tipiriseti, N.R.; Lakshmi, R.K.; Govatati, S.; Govatati, S.; Vuree, S.; Singh, L.; Raghunadha Rao, D.; Bhanoori, M.; Vishnupriya, S. Mitochondrial Genome Variations in Advanced Stage Breast Cancer: A Case–Control Study. *Mitochondrion* **2013**, *13*, 372–378, doi:10.1016/j.mito.2013.04.010.
38. Weigl, S.; Paradiso, A.; Tommasi, S. Mitochondria and Familial Predisposition to Breast Cancer. *CG* **2013**, *14*, 195–203, doi:10.2174/1389202911314030005.
39. Chatterjee, A.; Dasgupta, S.; Sidransky, D. Mitochondrial Subversion in Cancer. *Cancer Prevention Research* **2011**, *4*, 638–654, doi:10.1158/1940-6207.CAPR-10-0326.
40. Salas, A.; García-Magariños, M.; Logan, I.; Bandelt, H.-J. The Saga of the Many Studies Wrongly Associating Mitochondrial DNA with Breast Cancer. *BMC Cancer* **2014**, *14*, 659, doi:10.1186/1471-2407-14-659.
41. Mao, Q.; Gao, L.; Liu, Q.; Zou, Y.; Yang, X.; Wang, H.; Wang, Q.; Yu, H. The A10389G Polymorphism of ND3 Gene and Breast Cancer: A Meta-Analysis. *Biomedical Reports* **2013**, *1*, 259–264, doi:10.3892/br.2013.58.
42. Kent, W.J.; Sugnet, C.W.; Furey, T.S.; Roskin, K.M.; Pringle, T.H.; Zahler, A.M.; Haussler, and D. The Human Genome Browser at UCSC. *Genome Res.* **2002**, *12*, 996–1006, doi:10.1101/gr.229102.
43. Kent, W.J. BLAT —The BLAST -Like Alignment Tool. *Genome Res.* **2002**, *12*, 656–664, doi:10.1101/gr.229202.
44. van Oven, M.; Kayser, M. Updated Comprehensive Phylogenetic Tree of Global Human Mitochondrial DNA Variation. *Hum. Mutat.* **2009**, *30*, E386–E394, doi:10.1002/humu.20921.
45. Soares, P.; Ermini, L.; Thomson, N.; Mormina, M.; Rito, T.; Röhl, A.; Salas, A.; Oppenheimer, S.; Macaulay, V.; Richards, M.B. Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *The American Journal of Human Genetics* **2009**, *84*, 740–759, doi:10.1016/j.ajhg.2009.05.001.

46. Wang, C.-Y.; Wang, H.-W.; Yao, Y.-G.; Kong, Q.-P.; Zhang, Y.-P. Somatic Mutations of Mitochondrial Genome in Early Stage Breast Cancer. *Int. J. Cancer* **2007**, *121*, 1253–1256, doi:10.1002/ijc.22822.
47. Fendt, L.; Niederstätter, H.; Huber, G.; Zelger, B.; Dünser, M.; Seifarth, C.; Röck, A.; Schäfer, G.; Klocker, H.; Parson, W. Accumulation of Mutations over the Entire Mitochondrial Genome of Breast Cancer Cells Obtained by Tissue Microdissection. *Breast Cancer Res Treat* **2011**, *128*, 327–336, doi:10.1007/s10549-010-1092-8.
48. Gasparre, G.; Porcelli, A.M.; Bonora, E.; Pennisi, L.F.; Toller, M.; Iommarini, L.; Ghelli, A.; Moretti, M.; Betts, C.M.; Martinelli, G.N.; et al. Disruptive Mitochondrial DNA Mutations in Complex I Subunits Are Markers of Oncocytic Phenotype in Thyroid Tumors. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9001–9006, doi:10.1073/pnas.0703056104.
49. Ghatak, S.; Lallawmzuali, D.; Lalmawia; Sapkota, R.; Zothanpuia; Pautu, J.L.; Muthukumaran, R.B.; Senthil Kumar, N. Mitochondrial D-Loop and Cytochrome Oxidase C Subunit I Polymorphisms among the Breast Cancer Patients of Mizoram, North-east India. *Curr Genet* **2014**, *60*, 201–212, doi:10.1007/s00294-014-0425-2.
50. Darvishi, K.; Sharma, S.; Bhat, A.K.; Rai, E.; Bamezai, R.N.K. Mitochondrial DNA G10398A Polymorphism Imparts Maternal Haplogroup N a Risk for Breast and Esophageal Cancer. *Cancer Letters* **2007**, *249*, 249–255, doi:10.1016/j.canlet.2006.09.005.
51. Brandstätter, A.; Sängler, T.; Lutz-Bonengel, S.; Parson, W.; Béraud-Colomb, E.; Wen, B.; Kong, Q.-P.; Bravi, C.M.; Bandelt, H.-J. Phantom Mutation Hotspots in Human Mitochondrial DNA. *Electrophoresis* **2005**, *26*, 3414–3429, doi:10.1002/elps.200500307.
52. Marchington, D.R.; Hartshorne, G.M.; Barlow, D.; Poulton, J. Homopolymeric Tract Heteroplasmy in MtDNA from Tissues and Single Oocytes: Support for a Genetic Bottleneck. *Am J Hum Genet* **1997**, *60*, 408–416.
53. Sanchez-Cespedes, M.; Parrella, P.; Nomoto, S.; Cohen, D.; Xiao, Y.; Esteller, M.; Jeronimo, C.; Jordan, R.C.; Nicol, T.; Koch, W.M. Identification of a Mononucleotide Repeat as a Major Target for Mitochondrial DNA Alterations in Human Tumors. *Cancer research* **2001**, *61*, 7015–7019.
54. Parrella, P.; Seripa, D.; Matera, M.G.; Rabitti, C.; Rinaldi, M.; Mazzarelli, P.; Gravina, C.; Gallucci, M.; Altomare, V.; Flammia, G.; et al. Mutations of the D310 Mitochondrial Mononucleotide Repeat in Primary Tumors and Cytological Specimens. *Cancer Letters* **2003**, *190*, 73–77, doi:10.1016/S0304-3835(02)00578-5.
55. Lee, H.-C.; Wei, Y.-H. Mitochondrial DNA Instability and Metabolic Shift in Human Cancers. *IJMS* **2009**, *10*, 674–701, doi:10.3390/ijms10020674.
56. Yusoff, A.; Radzak, S.; Mohd Khair, S. Research Highlights on Contributions of Mitochondrial DNA Microsatellite Instability in Solid Cancers – an Overview. *wo* **2022**, *26*, 8–26, doi:10.5114/wo.2022.115674.
57. Xu, B.; Clayton, D.A. A Persistent RNA-DNA Hybrid Is Formed during Transcription at a Phylogenetically Conserved Mitochondrial DNA Sequence. *Mol Cell Biol* **1995**, *15*, 580–589, doi:10.1128/MCB.15.1.580.
58. Lee, D.Y.; Clayton, D.A. Initiation of Mitochondrial DNA Replication by Transcription and R-Loop Processing. *Journal of Biological Chemistry* **1998**, *273*, 30614–30621, doi:10.1074/jbc.273.46.30614.
59. Hauswirth, W.W.; Clayton, D.A. Length Heterogeneity of a Conserved Displacement-Loop Sequence in Human Mitochondrial DNA. *Nucl Acids Res* **1985**, *13*, 8093–8104, doi:10.1093/nar/13.22.8093.
60. Baptista-Rosas, R.C.; Mercado-Sesma, A.R.; Hernández, L.E.; Avalos, J.H.V.; Miranda, J.A.O.; Cruz, A.A.A.; Ramos, J.A.C.; Ortega, L.D.H.; Muñoz, A.D.R.; Luna, F. de J.O.; et al. *Insights Breast Cancer through Mitochondrial Genomics*; In Review, 2019;
61. Picornell, A.; Giménez, P.; Castro, J.A.; Ramon, M.M. Mitochondrial DNA Sequence Variation in Jewish Populations. *Int J Legal Med* **2006**, *120*, 271–281, doi:10.1007/s00414-006-0083-0.
62. Fridman, C.; Cardena, M.M.S.G.; Kanto, E.A.; Godinho, M.B.C.; Gonçalves, F.T. SNPs in Mitochondrial DNA Coding Region Used to Discriminate Common Sequences in HV1–HV2–HV3 Region. *Forensic Science International: Genetics Supplement Series* **2011**, *3*, e75–e76, doi:10.1016/j.fsigs.2011.08.037.
63. Röck, A.W.; Dür, A.; van Oven, M.; Parson, W. Concept for Estimating Mitochondrial DNA Haplogroups Using a Maximum Likelihood Approach (EMMA). *Forensic Science International: Genetics* **2013**, *7*, 601–609, doi:10.1016/j.fsigen.2013.07.005.
64. Verma, K.; Sharma, S.; Sharma, A.; Dalal, J.; Bhardwaj, T. Sequence Polymorphism Data of the Hypervariable Regions of Mitochondrial DNA in the Yadav Population of Haryana. *Data in Brief* **2018**, *18*, 164–171, doi:10.1016/j.dib.2018.03.004.
65. Bhatti, S.; Abbas, S.; Aslamkhan, M.; Attimonelli, M.; Trinidad, M.S.; Aydin, H.H.; de Souza, E.M.S.; Gonzalez, G.R. Genetic Perspective of Uniparental Mitochondrial DNA Landscape on the Punjabi Population, Pakistan. *Mitochondrial DNA Part A* **2018**, *29*, 714–726, doi:10.1080/24701394.2017.1350951.
66. Bandelt, H.-J.; Quintana-Murci, L.; Salas, A.; Macaulay, V. The Fingerprint of Phantom Mutations in Mitochondrial DNA Data. *The American Journal of Human Genetics* **2002**, *71*, 1150–1160, doi:10.1086/344397.
67. Aral, C.; Kaya, H.; Ataizi-Çelikel, Ç.; Akkiprik, M.; Sönmez, Ö.; Güllüoğlu, B.M.; Özer, A. A Novel Approach for Rapid Screening of Mitochondrial D310 Polymorphism. *BMC Cancer* **2006**, *6*, 21, doi:10.1186/1471-2407-6-21.
68. Alhomidi, M.A.; Vedicherla, B.; Movva, S.; Rao, P.K.; Ahuja, Y.R.; Hasan, Q. Mitochondrial D310 Instability in Asian Indian Breast Cancer Patients. *Tumor Biol.* **2013**, *34*, 2427–2432, doi:10.1007/s13277-013-0793-0.
69. Bandy, B.; Davison, A.J. Mitochondrial Mutations May Increase Oxidative Stress: Implications for Carcinogenesis and Aging? *Free Radical Biology and Medicine* **1990**, *8*, 523–539, doi:10.1016/0891-5849(90)90152-9.
70. Gille, J.J.P.; Joenje, H. Cell Culture Models for Oxidative Stress: Superoxide and Hydrogen Peroxide versus Normobaric Hyperoxia. *Mutation Research/DNAging* **1992**, *275*, 405–414, doi:10.1016/0921-8734(92)90043-O.
71. Scarpulla, R.C. Nuclear Control of Respiratory Gene Expression in Mammalian Cells. *J. Cell. Biochem.* **2006**, *97*, 673–683, doi:10.1002/jcb.20743.

-
72. Lee, H. Somatic Mutations in the D-Loop and Decrease in the Copy Number of Mitochondrial DNA in Human Hepatocellular Carcinoma. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2004**, 547, 71–78, doi:10.1016/j.mrfmmm.2003.12.011.
 73. Xie, H.; Lev, D.; Gong, Y.; Wang, S.; Pollock, R.E.; Wu, X.; Gu, J. Reduced Mitochondrial DNA Copy Number in Peripheral Blood Leukocytes Increases the Risk of Soft Tissue Sarcoma. *Carcinogenesis* **2013**, 34, 1039–1043, doi:10.1093/carcin/bgt023.
 74. Carew, J.S.; Huang, P. Mitochondrial Defects in Cancer. *Molecular Cancer* **2002**, 1, 9, doi:10.1186/1476-4598-1-9.
 75. Wang, X. The Expanding Role of Mitochondria in Apoptosis. *Genes Dev* **2001**, 15, 2922–2933.