Potential Association of Mitochondrial Haplogroups and A8860G Mutation with Breast Cancer Risk

Han N. Mohammed Fadhl,^{1¶} and Farhad M. Abdulkarim^{2,3 *¶}

¹ Basic science, College of dentistry, University of Sulaimani, Sulaimaniyah, Iraq

² Microgene Diagnostic Center, Harem hospital, Sulaimaniyah, Iraq

³ Kurdistan Institute for Strategic Studies and Scientific Research (KISSR), Sulaimaniyah, Iraq

*Corresponding author: abdulkarimf@microgene.org (FMA)

Contributionship: [¶](HNM) and (FMA)

Abstract:

The last decade has witnessed great progresses regarding the molecular basis of breast cancer with discovery of different nuclear susceptibility genes; in addition investigations and researches regarding mitochondrial DNA (mtDNA) mutations in breast cancer have been started. Mitochondrial haplogroup determinants (single nucleotide polymorphism SNP) and somatic mitochondrial mutations have recently been studied as possible risk factors for carcinogenic processes in different tissues, hence in order to identify breast cancer related SNPs and haplogroups among the population of Sulaimaniyah city/Iraq, the entire mitochondrial genome of 20-breast cancer samples and comparable controls were sequenced. Haplogrep 2.0 was used for haplogroup identification; Chi-square and Fishers exact test were applied to assess relational significance. HV haplogroup in the cancer samples appeared to be a risk factor for breast cancer compared to the most common H haplogroup in control samples with a p-values of 0.002 and 0.006 respectively and an Odd Ratio (OR) = 28.00. Besides, SNP (A8860G) was also identified as a risk factor for breast cancer as compared to other randomly selected SNPs (A750G, A1438G and C7028T) with p values ≤ 0.05 and OR > 1 .

Key-words: Breast Cancer, Mitochondrial DNA, Haplogroup, SNP, Sulaimaniyah City

28 **1.Introduction:**

29 Breast cancer is the most frequently diagnosed cancer in women [1]; representing a
 30 heterogeneous, debilitating disease of a multistep carcinogenic background [2]. The incidence
 31 of breast cancer in Kurdish women in north of Iraq is the same as in western countries and even
 32 higher especially in the young age group [3, 4]. Although, major progresses have been made
 33 over the past decade in exploring the molecular basis of breast cancer and several susceptibility
 34 nuclear-genes of high, moderate and low penetrance have been identified as predisposing
 35 factors for breast cancer, still these genes can explain only 10-15% of breast cancer cases [5, 6,
 36 and 7]. Recently, many attempts have been made to study possible relation of mitochondrial
 37 genome mutations with incidence of breast cancer and other types of cancer, including
 38 mitochondrial haplogroups that are determinants of a number of single nucleotide
 39 polymorphisms (SNP)'s gained by geographical and environmental impacts during the process
 40 of evolution throughout history [8]. To date several studies in different populations have been
 41 performed, figuring out significant relation between specific haplogroups and incidence of
 42 different cancers, including breast cancer [9, 10, 11, and 12] as well as identification of specific
 43 mutational variants in association with cancer cases in general as T16189C, G10398A and the
 44 deletion of mtDNA 4977 [13]. Mitochondrial DNA (mtDNA) is a circular molecule of 16569
 45 bp, coding for two 2 ribosomal RNAs (12S and 16S), 22 transfer RNAs and 13 essential
 46 protein subunits of the oxidative phosphorylation system (OXPHOS) [14]. It is well known that
 47 (mtDNA) has a mutation rate several times higher than nuclear DNA [15], due to its limited
 48 repair mechanisms, lack of protective histones and its close proximity to the electron transport
 49 chain, which continuously generates free radicals [16]; in addition, mtDNA is organized in an
 50 economic pattern with its genes lacking introns [17]. It was in 1956, when Otto Warburg
 51 observed that cancers ferment glucose in the presence of oxygen, proposing that abnormalities

in mitochondrial respiration may be responsible for cancer production [16, 18, and 19]. Now it's well-established that changes in the biochemical processes that accompany the process of carcinogenesis as the aerobic glycolysis do not impair mitochondrial function but are rather essential for cancer cell viability [15].

The current study was performed to identify possible relation of specific haplogroups with incidence of breast cancer in the population of Sulaimaniyah city (a Kurd ethnic occupant city in north of Iraq), as well as recognition of significantly common SNPs among the breast cancer cases in the studied population, considering no such previous studies have been performed neither in the city nor in the locality

2. Materials and methods:

-Sample selection:

The study was approved by the Ethical Committee of Collage of Medicine/University of Sulaimani (Number 44, on January 30, 2017) and written consents were obtained from all participants. A total of 40 subjects (20 breast cancer tissue and 20 control samples) were recruited for the study. Control samples were taken from benign breast tissue specimens (fibroadenoma and non-proliferative fibrocystic breast disease). Breast cancer tissue samples were taken from mastectomy specimens of subjects already diagnosed as having invasive ductal carcinoma (Grade II and III) by core biopsy and with no family history of breast cancer. All samples were from unrelated individuals of the Kurdish ethnic population from the center of Sulaymaniyah city.

74 -DNA extraction:

75 Total genomic DNA was extracted using DNA extraction kit (GeNet bio/South Korea). The
76 extraction was performed according to the manufacturer's instruction. Purity and concentration
77 of the extracted DNA were obtained using a Biophotometer (Eppendorf/Germany).

79 -PCR amplification and sequencing:

80 The entire mitochondrial genome was amplified in the form of four overlapping PCR
81 fragments using long Taq kit (Dongsheng Biotech/China) and the primers listed in (Table 1).

83 **Table 1: Sequence of the amplification Primers of four overlapping mitochondrial DNA**
84 **fragments**

	Primer name	Sequence
1	1F	5`-AGG TCT ATC ACC CTA TTA ACC ACT CA-3`
2	2F	5`-CAA GAG CCT TCA AAG CCC TCA GTA-3`
3	3F	5`-ACG CCA CTT ATC CAG TGA ACC ACT-3`
4	4F	5`-CCT AGC AAT AAT CCC CAT CCT CCA-3`
5	1R	5`-TGA GCA AGA GGT GGT GAG GTT GAT-3`
6	2R	5`-GGG CAC CGA TTA TTA GGG GAA CTA-3`
7	3R	5`-TAT GAG AAT GAC TGC GCC GGT GAA-3`
8	4R	5`-CGT GAT GTC TTA TTT AAG GGG AAC GT-3`

The amplified PCR products were purified using PCR purification kit (NORGEN biotek/Canada) and the primers listed in (Table 2) were used for sequencing of the amplified mtDNA fragments.

90

Table 2: The Sequencing primers of the four overlapping fragments of mitochondrial

DNA

Primer name	Sequence	Position
5MT	5'-TGA ACT CAC TGG AAC GGG GAT GCT-3'	723-700
6MT	5'- GCA GAA GGT ATA GGG GTT AGT CCT-3'	1852-1829
7MT	5'- ATG CCT GTG TTG TGA GAG TGA-3'	2439-2416
8MT	5'-TCT TGT CCT TTC GTA CAG GGA GGA-3'	3138-3115
9MT	5'-CTG AGA CTA GTT CGG ACT CCC CTT-3'	3934-3911
10MT	5'- CGG TTG CTT GCG TGA GGA AAT ACT-3'	4665-4642
11MT	5'- GGA GTA GTG TGA TTG AGG TGG AGT-3'	5385-5362
12MT	5'-GGA GTG TGG CGA GTC AGC TAA ATA-3'	6885-6862
13MT	5'- AAG GGC ATA GAG GAC TAG GAA GCA-3'	7711-7688
14MT	5'- AGG GAG GTA GGT GGT AGT TTG TGT-3'	8477-8454
15MT	5'- GGG GTC ATG GGC TGG GTT TTA CTA-3'	9258-9235
16MT	5'- TAT AGG GTCGAA GCC GCA CTC GTA-3'	10190-10167
17MT	5'- GTG AGG GGT AGG AGT CAG GTA GTT-3'	10986-10963
18MT	5'- TAG GGA AGT CAG GGT TAG GGT GGT-3'	12381-12358
19MT	5'- AGT GCT TGA GTG GAG TAG GGC TGA-3'	13089-13066
20MT	5'- AAT CCT GCG AAT AGG CTT CCG GCT-3'	13733-13710
21MT	5'- GCT ATT GAG GAG TAT CCT GAG GCA-3'	14454-14431

22MT	5'- TGC AAG CAG GAG GAT AAT GCC GAT-3'	15112-15089
23MT	5'- GGT AGC TTA CTG GTT GTC CTC CGA-3'	15782-15759

93

94 -Data analysis:

95 The algorithm implemented in the HaploGrep 2.0 was used for identification of haplogroups
 96 [20]. Chi-square and Fishers exact test were used to determine the significance of relations of
 97 breast cancer with haplogroups and SNPs. The <https://www.mitomap.org/MITOMAP> website
 98 which provides a comprehensive database for human mitochondrial DNA was used for
 99 allocation of mutations, identifying types of the mutations and determining amino acid
 100 substitutions.

101 3. Results:

102 A total of 344 mutations in the cancer samples and 203 mutations in the control samples were
 103 identified in the current study. The majority of the mutations were point mutations with only
 104 four insertion mutation regions. Based on the MITOMAP databases, single nucleotide
 105 polymorphisms (SNP) of the Kurdish ethnicity were identified and accounted for 74% of all
 106 mutations in the breast cancer samples, of which 61% were distributed in the coding region and
 107 39% were in the non-coding region (Table 3). Whereas, in the control samples SNPs
 108 constituted 90% and the pattern of distributions were 58% in the coding region and 42% in the
 109 non-coding region (Table 4).

110

111

112

113

114 **Table3: Total single nucleotide polymorphism (SNP) mutations in breast cancer samples**

SNP	Gene/region	Nucleotide substitution	Amino acid Substitution	SNP	Gene/ region	Nucleotide substitution
3394	ND1	T>C	Tyr-His	16086	HV1	T>C
3834	ND1	G>A	Sync	16172	HV1	T>C
3741	ND1	C>T	Sync	16186	HV1	C>T
4011	ND1	C>T	Sync	16187	HV1	C>T
4216	ND1	T>C	Tyr-His	16189	HV1	T>C
4769	ND2	A>G	Sync	16192	HV1	C>T
4917	ND2	A>G	Asn-Asp	16193	HV1	C>T
7028	Cox1	C>T	Sync	16217	HV1	T>C
8137	Cox2	C>T	Sync	16223	HV1	C>T
8684	ATPase 6	C>T	Thr-Ile	16274	HV1	G>A
8697	ATPase 6	G>A	Sync	16209	HV1	T>C
8860	ATPase 6	A>G	Thr-Ala	16234	HV1	C>T
9755	Cox3	G>A	Sync	16249	HV1	T>C
9899	Cox3	T>C	Sync	16291	HV1	C>T
10142	ND3	C>T	Sync	16294	HV1	C>T
10586	ND4L	G>A	Sync	16309	HV1	A>G
11251	ND4	A>G	Sync	16318	HV1	A>T
11467	ND4	A>G	Sync	16362	HV1	T>C
11719	ND4	G>A	sync	16519	HV1	T>C
12372	ND5	G>A	Sync	146	HV2	T>C
12612	ND5	A>G	Sync	151	HV2	C>T
12618	ND5	G>A	Sync	152	HV2	T>C
12705	ND5	C>T	Sync	195	HV2	T>C
12879	ND5	T>C	Sync	Poly C insertion (309, 310 C	HV2	

				and 311)		
13104	ND5	C>T	Sync	263	HV2	A>G
13188	ND5	C>T	Sync	417	HV2	G.>A
13368	ND5	G>A	Sync	499	HV2	G>A
13500	ND5	T>C	Sync	462	HV3	C>T
13708	ND5	G>A	Ala-Thr	489	HV3	T>C
14139	ND5	A>G	Sync	10463	tRNA-Arg	T>C
14364	ND6	G>A	Sync			
14569	ND6	G>A	Ala-Thr			
14905	Cyt-B	G>A	Sync			
14766	Cyt-B	C>T	Thr-Ile			
15148	Cyt-B	G>A	Synch			
15326	Cyt-B	A>G	Thr-Ala			
15452	Cyt-B	C>A	Leu-Ile			
15607	Cyt-B	T>C	Sync			
12308	tRNA	A>G				
10463	tRNA	A>G				
709	rRNA	G>A				
750	rRNA	A>G				
980	rRNA	T>C				
1438	rRNA	A>G				
1811	rRNA	A>G				
1888	rRNA	G>A				
2259	rRNA	C>T				
3010	rRNA	G>A				
2706	rRNA	A>G				

115 Red coloured residues were identified in more than one case

116

Table 4: Total single nucleotide polymorphism (SNP) mutations in control group samples

SNP	Gene/region	Nucleotide substitution	a.a substitution	SNP	Gene/region	Nucleotide substitution	a.a substitution
73	HV2	A>G	-	1811	rRNA	A>G	
146	HV2	T>C	-	1438	rRNA	A>G	
152	HV2	T>C	-	2706	rRNA	A>G	
189	HV2	A>G	-	3010	rRNA	G>A	
195	HV2	T>C	-	4216	ND1	T>C	Tyr-His
204	HV2	T>C	-	4769	ND2	A>G	Synch, Met
263	HV2	A>G	-	5046	ND2	G>A	Val-Ile
295	HV2	C>T	-	5460	ND2	G>A	Ala-Thr
Poly C insertion (309, 310 C and 311)			-	7028	Cox1	C>T	Synch
			-	7521	tRNA	G>A	
462	HV3	C>T	-	7789	Cox2	G>A	Synch
489	HV3	T>C	-	8860	ATPase6	A>G	Thr-Ala
709	rRNA	G>A	-	8994	ATP6	G>A	Synch
750	rRNA	A>G	-	11674	ND4	C>T	Synch
1243	rRNA	T<C		11719	ND4	G>A	Sync
13145	ND5	G>A	Ser-Asn	12705	ND5	C>T	Synch
14766	Cyt-B	C>T	Thr-Ile				
15326	Cyt-B	A>G	Thr-Arg				
16189	HV1	T>C					
16218	HV1	C>T					
16223		C>T					
16519	HV1	T>C					

Red coloured residues were identified in more than one case

119

Nine Western Eurasian haplogroups and their subclasses were identified in both cancer and control samples using the Haplogrep 2.0 program. Haplogroups: HV, N, R, U, J, T and H were identified in breast cancer samples, on the other hand H, HV, N, R0, J, X and W haplogroups were identified in the control samples. The most common haplogroup in the control samples was the H-haplogroup (60%), while in breast cancer samples the frequency was less common (5%). In contrast, commonly occurring haplogroup in breast cancer samples was HV (35%) followed by N (25%) (Table 5 A & B).

Table 5: The identified Haplogroups in breast cancer (A) and control (B) subjects

Breast cancer samples			Control samples		
Haplogroups	Frequency of occurrence	%	Haplogroups	Frequency of occurrence	%
HV	7	35%	H	12	60%
N	5	25%	HV	3	20%
U7	2	10%	N	1	5%
R0	2	10%	R0	1	5%
J	1	5%	J1	1	5%
U1	1	5%	X	1	5%
T	1	5%	W	1	5%
H	1	5%	-	-	-
Total	20	100%	Total	20	100%

A statistically significant association between haplogroup HV and breast cancer was identified using Chi-square and Fisher's exact tests, where the p-values were 0.002 and 0.006, respectively, and the odd ratio (OR) = 28, indicating that the HV haplogroup is a high risk factor for the incidence of breast cancer.

Furthermore, the homoplasmic mutation, SNP (A8860G) (Figure 1) was identified in all 20 breast cancer samples (100%), while in control samples it was less frequent and identified in only 4 samples (20%).

Figure 1: Electropherogram and sequence of the A8860G region

Point mutation site is indicated by an arrow

To identify the significance of this mutation, Chi-square, Fishers exact test were used and OR were calculated and compared with three other randomly selected SNPs (A750G, A1438G and C7028T) (Figure 2a, 2b and 2c).

Figure 2a: Electropherogram and sequence of the A750G region

Point mutation site is indicated by an arrow

Figure 2b: Electropherogram and sequence of the A1438G region

Point mutation site is indicated by an arrow

Figure 2c: Electropherogram and sequence of the C7028T region

Point mutation site is indicated by an arrow

As indicated in (Table 6), the OR values were greater than 1 and the p-values were less than 1, indicating SNP (A8860G) as a risk factor for developing breast cancer.

Table 6: Calculated OR and p values for SNP A8860G with three randomly selected SNPs

SNPs	Odd ratio	P value (chi-square)	P value (Fishers exact)
8860/750	4.722	0.0012	0.0014
8860/1438	5	0.009	0.013
8860/7028	5	0.011	0.021

4. Discussion:

Incidence of cancer in Sulaymaniyah city (latitude 33.314690 and longitude 44.376759) in the northeast part of Iraq [21], has lately shown a great increase, as is the case in the rest of Iraq; and breast cancer was the commonest among women [22, 23]. Mitochondrial DNA mutations and polymorphisms have been increasingly reported in a wide variety of cancers, including breast, prostate, and colorectal cancers [24], indicating that mutations and haplogroup determining polymorphisms may influence mitochondrial protein changes that may affect the OXPHOS process and promote the production of reactive oxidative species [25]. For breast cancer it is well established that reactive oxygen species play an important role in the process of carcinogenesis [26, 27]; hence, the current study was conducted to determine common SNPs and haplogroups among the breast cancer cases

According to the results most of the SNPs were in the coding region, both in the cancer cases and the controls (61% and 58%, respectively). This shows the importance of whole genomic

170 sequencing for precise haplogroup determination and the lesser value of the polymorphisms
171 traditionally used in the hypervariable regions for forensic purposes [28, 29].

172 Several previous studies have been conducted and showed a significant relation between specific
173 haplogroups and cancer incidence in general [13]. In breast, Chinese women of haplogroups M
174 and subhaplogroup D5 had shown a higher incidence for cancer [30, 31], while no such a
175 remarkable relation was identified between cancer and specific haplogroups in European and
176 Caucasian women [32], still haplogroup K showed a significant association with breast cancer in
177 European-American women [33]. Nevertheless in the current study a significant relation between
178 haplogroup HV and breast cancer was identified with p value = 0.002 and 0.006 for Chi square
179 and Fisher's exact test respectively and OR of 28.

180 In addition to the haplogroups, several distinct SNPs have been previously discovered to be
181 associated with cancers in general as T16189C, G10398A and the deletion of mtDNA 4977 [13];
182 in regard to breast, certain SNPs were identified as well to be associated with increased cancer
183 incidence. A10398G is one of the well-known SNPs in breast cancer detected in European-
184 American, Malaysian and African-American women [33, 34, 35, 36 and 37]; in addition SNPs
185 G9055A and T16519C were also identified as risk factors for breast cancer in European-
186 American females [33]. Furthermore several other germ line mutations as 2463 A-deletion,
187 C6296A, 6298 T-deletion, A8860G, and 8460-13327deletion, were detected in chines women
188 with breast cancer [15].

189 Although many SNPs were identified in breast cancer samples in the current study shown in
190 (Table 3), but the only mutation showed a significantly high incidence among breast cancer
191 samples compared to the control samples was homoplasmic SNP (A8860G). It is a non-
192 synchronous mutation in the Mt-ATP 6 gene that was detected in all 20 breast cancer samples
193 while only in 4 of the control samples and this result was compatible with Li et al [15]. This gene
194 encodes ATP synthase 6 (681 amino acids), a subunit of complex V, whose mutation results in

195 substitution of a polar uncharged amino acid (threonine) with a non-polar aliphatic amino acid
196 (alanine); this may affect hydrophobic interactions and hence the structure of the protein.
197 However, such a prediction of protein structure is not absolute as these mutations may be
198 followed by other compensatory mutations (suppressor mutations) in order to minimize the initial
199 mutation's effect [38], these compensatory and suppresser mutations may explain the presence of
200 the mutation A8860G in 20% of phenotypically healthy control samples

201

202 **5. Conclusion:**

203 As the results show, both haplogroup HV and SNP A8860G are risky factors for developing
204 breast cancer in the studied population; however these results are not compatible with the
205 previously identified risky SNPs and haplogroups in breast cancer studies performed in other
206 populations, except for SNP (A8860G) which was compatible with that of Li et al [15]; this
207 could be explained by the effect of other parameters on the mitochondrial genome, such as
208 individual physiology and influences of geographical location, suggesting a population specific
209 effect of these haplogroups and SNPs in the carcinogenic processes. Hence in an attempt to
210 fortify the current results and identify possible associations with other tissue cancers, further
211 studies are required with larger sample size and other cancer tissues to be included.

212

213 **Acknowledgement:**

214 We are grateful for the sample donors for their cooperation. Efforts of Mr. Pola Abdalla Othman
215 are gratefully acknowledged. This work was supported by University of Sulaimani and KISSR

216

217 **Data availability:**

218 All data are available upon request from the corresponding author via the following e-
219 mail: abdulkarimf@microgene.org

220

221 **References:**

- 222 1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018:
223 GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185
224 countries. CA: A Cancer Journal for Clinicians. 2018;68 (6):394-424..
- 225 2. Ingvarsson S. Breast cancer: introduction. Seminars in Cancer Biology. 2001;11(5):323-326
- 226 3. Majid R, Mohammed H, Saeed H, Safar B, Rashid R, Hughson M. Breast cancer in kurdish
227 women of northern Iraq: incidence, clinical stage, and case control analysis of parity and
228 family risk. BMC Women's Health. 2009;9(1).
- 229 4. Molah Karim S, Ali Ghalib H, Mohammed S, Fattah F. The incidence, age at diagnosis of
230 breast cancer in the Iraqi Kurdish population and comparison to some other countries of
231 Middle-East and West. International Journal of Surgery. 2015;13:71-75.
- 232 5. Frey J, Salibian A, Schnabel F, Choi M, Karp N. Non-BRCA1/2 Breast Cancer
233 Susceptibility Genes. Plastic and Reconstructive Surgery - Global Open. 2017;5(11):e1564.
- 234 6. Russnes H, Lingjærde O, Børresen-Dale A, Caldas C. Breast Cancer Molecular
235 Stratification. The American Journal of Pathology. 2017; 187(10):2152-2162.
- 236 7. Hamdi Y, Boujemaa M, Ben Rekaya M, Ben Hamda C, Mighri N, El Benna H et al. Family
237 specific genetic predisposition to breast cancer: results from Tunisian whole exome
238 sequenced breast cancer cases. Journal of Translational Medicine. 2018;16(1).

- 239 8. Wallace D. Bioenergetics in human evolution and disease: implications for the origins of
240 biological complexity and the missing genetic variation of common diseases. Philosophical
241 Transactions of the Royal Society B: Biological Sciences. 2013;368(1622):20120267.
- 242 9. Canter J, Kallianpur A, Fowke J. Re: North American White Mitochondrial Haplogroups in
243 Prostate and Renal Cancer. Journal of Urology. 2006;176(5):2308-2309.
- 244 10. Verma M, Kumar D. Application of mitochondrial genome information in cancer
245 epidemiology. Clinica Chimica Acta. 2007;383(1-2):41-50.
- 246 11. Riley V, Erzurumluoglu A, Rodriguez S, Bonilla C. Mitochondrial DNA Haplogroups and
247 Breast Cancer Risk Factors in the Avon Longitudinal Study of Parents and Children
248 (ALSPAC). Genes. 2018;9(8):395.
- 249 12. Lan Q, Xie T, Jin X, Fang Y, Mei S, Yang G et al. MtDNA polymorphism analyses in the
250 Chinese Mongolian group: Efficiency evaluation and further matrilineal genetic structure
251 exploration. Molecular Genetics & Genomic Medicine. 2019;7(10).
- 252 13. Jimenez-Morales S, Perez-Amado C, Langley E, Hidalgo-Miranda A. Overview of
253 mitochondrial germline variants and mutations in human disease: Focus on breast cancer
254 (Review). International Journal of Oncology. 2018.
- 255 14. Errichiello E, Venesio T. Mitochondrial DNA Variations in Tumors: Drivers or Passengers?.
256 Mitochondrial DNA - New Insights. 2018;. DOI: 10.5772/intechopen.75188
- 257 15. Li L, Chen L, Li J, Zhang W, Liao Y, Chen J et al. Correlational study on mitochondrial
258 DNA mutations as potential risk factors in breast cancer. Oncotarget. 2016;7(21):31270-
259 31283.
- 260 16. Mohamed Yusoff A, Wan Abdullah W, Mohd Khair S, Abd Radzak S. A comprehensive
261 overview of mitochondrial DNA 4977-bp deletion in cancer studies. Oncology Reviews.
262 2019;13(1).

- 263 17. Alexeyev M, Shokolenko I, Wilson G, LeDoux S. The Maintenance of Mitochondrial DNA
264 Integrity--Critical Analysis and Update. Cold Spring Harbor Perspectives in Biology.
265 2013;5(5):a012641-a012641..
- 266 18. Zong W, Rabinowitz J, White E. Mitochondria and Cancer. Molecular Cell. 2016;61(5):667-
267 676.
- 268 19. Germain D, Papa L, Kenny T, Takabatake Y, Riar A. Mitochondrial dysfunction in breast
269 cancer. Research and Reports in Biology. 2015;:137.
- 270 20. Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G et al.
271 HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA
272 haplogroups. Human Mutation. 2010;32(1):25-32.
- 273 21. Al-Zahery N, Saunier J, Ellingson K, Parson W, Parsons T, Irwin J. Characterization of
274 mitochondrial DNA control region lineages in Iraq. International Journal of Legal Medicine.
275 2012;127(2):373-375.
- 276 22. Khoshnaw N, Mohammed H, Abdullah D. Patterns of Cancer in Kurdistan - Results of Eight
277 Years Cancer Registration in Sulaymaniyah Province-Kurdistan-Iraq. Asian Pacific Journal
278 of Cancer Prevention. 2016;16(18):8525-8531.
- 279 23. Iraqi cancer data for 2017–2018, available at [http://www.emro.who.int/irq/iraq-](http://www.emro.who.int/irq/iraq-news/cancer-data-for-20172018-announced-in-iraq.html)
280 [news/cancer-data-for-20172018-announced-in-iraq.html](http://www.emro.who.int/irq/iraq-news/cancer-data-for-20172018-announced-in-iraq.html)
- 281 24. Parr R, Dakubo G, Thayer R, McKenney K, Birch-Machin M. Mitochondrial DNA as a
282 potential tool for early cancer detection. Human Genomics. 2005;2(4):252
- 283 25. Li H, Slone J, Fei L, Huang T. Mitochondrial DNA Variants and Common Diseases: A
284 Mathematical Model for the Diversity of Age-Related mtDNA Mutations. Cells.
285 2019;8(6):608.

286 26. Okoh V, Deoraj A, Roy D. Estrogen-induced reactive oxygen species-mediated signalings
287 contribute to breast cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*.
288 2011;1815(1):115-133.

289 27. Gurer-Orhan H, Ince E, Konyar D, Saso L, Suzen S. The Role of Oxidative Stress
290 Modulators in Breast Cancer. *Current Medicinal Chemistry*. 2018;25(33):4084-4101.

291 28. Fridman C, Cardena M, Kanto E, Godinho M, Gonçalves F. SNPs in mitochondrial DNA
292 coding region used to discriminate common sequences in HV1–HV2–HV3 region. *Forensic*
293 *Science International: Genetics Supplement Series*. 2011;3(1):e75-e76.

294 29. Weng S, Lin T, Wang P, Chen S, Chuang Y, Liou C. Single nucleotide polymorphisms in the
295 mitochondrial control region are associated with metabolic phenotypes and oxidative stress.
296 *Gene*. 2013;531(2):370-376.

297 30. Fang H, Shen L, Chen T, He J, Ding Z, Wei J et al. Cancer type-specific modulation of
298 mitochondrial haplogroups in breast, colorectal and thyroid cancer. *BMC Cancer*. 2010;10(1).

299 31. Ma L, Fu Q, Xu B, Zhou H, Gao J, Shao X et al. Breast cancer-associated mitochondrial
300 DNA haplogroup promotes neoplastic growth via ROS-mediated AKT activation.
301 *International Journal of Cancer*. 2017; 142 (9):1786-1796.

302 32. Gutiérrez Povedano C, Salgado J, Gil C, Robles M, Patiño-García A, García-Foncillas J.
303 Analysis of BRCA1 and mtDNA haplotypes and mtDNA polymorphism in familial breast
304 cancer. *Mitochondrial DNA*. 2013; 26(2):227-231.

305 33. Bai R, Leal S, Covarrubias D, Liu A, Wong L. Mitochondrial Genetic Background Modifies
306 Breast Cancer Risk. *Cancer Research*. 2007;67(10):4687-4694.

307 34. Tengku Baharudin N, Jaafar H, Zainuddin Z. Association of Mitochondrial DNA 10398
308 Polymorphism in Invasive Breast Cancer in Malay Population of Peninsular Malaysia.
309 *Malays J Med Sci*. 2012; 19(1): pp.36–42.

- 310 35. Mims M, Hayes T, Zheng S, Leal S, Frolov A, Ittmann M et al. Mitochondrial DNA
311 G10398A Polymorphism and Invasive Breast Cancer in African-American Women. *Cancer*
312 *Research*. 2006;66(3):1880-1881.
- 313 36. Darvishi K, Sharma S, Bhat A, Rai E, Bamezai R. Mitochondrial DNA G10398A
314 polymorphism imparts maternal Haplogroup N a risk for breast and esophageal cancer.
315 *Cancer Letters*. 2007;249(2):249-255.
- 316 37. Jahani M, Azimi Meibody A, Karimi T, Banoei M, Houshmand M. An A10398G
317 mitochondrial DNA alteration is related to increased risk of breast cancer, and associates with
318 Her2 positive receptor. *Mitochondrial DNA Part A*. 2019;31(1):11-16.
- 319 38. Schaefer C and Rost B. Predict impact of single amino acid change upon protein
320 structure. *BMC Genomics*. 2012;13(Suppl 4), p.S4

321 **Supporting information**

322 **S1aTable. Frequency difference of haplogroups H and HV among breast cancer samples**
323 **and control samples**

324 **S1bTable. Odd ratio for haplogroup HV/H (Risk estimation)**

325 **S2aTable. Frequency difference of mutations (8860) and (750) among breast cancer**
326 **samples and control samples**

327 **S2bTable. Odd ratio for 8860/750 (Risk estimation)**

328 **S3aTable. Frequency difference of mutations (8860) and (1438) among breast cancer**
329 **samples and control samples**

330 **S3bTable. Odd ratio for 8860/1438 (Risk estimation)**

- 331 **S4aTable. Frequency difference of mutations (8860) and (7028) among breast cancer**
332 **samples and control samples**
- 333 **S4bTable. Odd ratio for 8860/7028 (Risk estimation)**