

Mitochondrial DNA (Hypervariable region II) diversity from Basrah population

Bassim Muften Ohied¹, Adnan Issa Al-Badran^{2*}

^{1,2}Biology Department, College of Science, University of Basrah, Iraq

Email: Adnan.albadran@uobasrah.edu.iq

DOI: 10.47750/pnr.2022.13.S06.431

Abstract

Mitochondrial DNA hypervariable regions II of control region were sequenced from 139 random healthy unrelated individuals ethnic of Iraqi population in Basrah city. The aim of this study was to detect and classify the mtDNA haplotypes into mtDNA haplogroups will be better in applications of the forensic genetics and will determine the history of the Iraqi Population The variation in sequence within the D-loop control region was analyzed for the haplogroups composition which displayed a high haplogroups frequency. H, J, U, B, N and R (25%, 12%, 12%, 7%, 5% and 5%, respectively, Moderate Haplogroups Frequency L, T and W was (3%) HV, I, X and M (1%), and low haplogroups frequency in pre-HV (0.7%). This study Reported absence of haplogroups V, P, Y, O, Z, Q, G, E and C.

Keywords: mitochondrial DNA, haplogroups, DNA Sequencing, Basrah Iraqi population, mtDNA diversity.

INTRODUCTION

In Human, mitochondrial DNA (mtDNA) is a circular double-stranded molecule with a length of 16,569 base pairs (bp) encoding 13 subunits of the oxidative phosphorylation scheme, 2 ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs) (Anderson et al., 1981). In each cell found in hundreds to thousands of copies not in the nucleus but the mitochondria, the energy-generating organelles of the cell. MtDNA primarily consists of coding DNA, Except for a long fragment of a maximum of 1100-bp, which mostly has regulatory functions and is therefore called the control region. 25 years ago, since the first in-depth study of human mtDNA variability (Brown, 1980), It has been widely used in studies of human evolution, migration and the history of population. This widespread use is due to the unique features of mtDNA, which make it particularly suited for evolutionary studies. Such characteristics include a high number of copies, maternal inheritance, a lack of recombination and a generally mutation rate than those seen in nuclear DNA.

In 1981 Anderson et al. published on the human mitochondrial DNA (mtDNA) nucleotides sequences (Anderson et al, 1981). In the control region, including the displacement loop system (D-loop region), hypervariable region 1 (HV1), hypervariable region 2 (HV2) and hypervariable region 3 (HV3) rich in polymorphism were discovered (Horai and Hayasaka 1990; Lutz et al 2000).

Each cell contains several thousand copies of mtDNA, which is useful because minute specimens or old specimens have been contaminated or degraded (Lutz et al, 1996; Koyama et al ,2002). So the value of mtDNA in the forensic medicine field is therefore extremely high. In addition, MtDNA is transmitted through cytoplasmic inheritance or maternal inheritance, or put it another way. To this purpose, it is theoretically useful in defining maternal relationships, but since the father's mtDNA (sperm) is eradicated after fertilization, could not use mtDNA to identified a parental relationship (Giles et al, 1980; Gill et al ,1994).

Material and Methods

Population sample

A total of 139 blood samples from unrelated females were gathered, representing informed consent was obtained from all donors and ethnic background details. The sample included only the native people.

DNA extraction

Genomic DNA extracted used gSYNC™ DNA Extraction Kit Quick protocol by Gene aid company.

PCR amplification

According to (Tsutsumi et al. 2004) report; nt 29 to 408 of HV2 have been amplified by synthesizing forward primer: 5'-GGTCTATCACCTATTAACCAC'-3' and reverse primer: 5'-CTGTTAAAAGTGCATACCGCCA-3'. The mixture of PCR reactions used was PCR mixture 3.0 µl, 2.0 mM dNTP mix 3.0 µl, 20 pmol of each primer, Gold Taq DNA polymerase 2U and 15 ng template DNA applied to a total volume of 25 µl of sterilized water. The step of PCR was set initial denaturation at 95°C for 9 min, followed by thirty cycles (denaturation at 95°C for 60 sec, annealing at 55°C for 60 sec, and extension at 72°C for 120 sec). The samples were incubated at 72°C for an additional 7 min after the last cycle. The PCR product was sent to a microgen company in South Korea to electrophoresis it by using genetic analyzer.

Phylogeny analysis

Clustal X ver. 2.0 used to align the DNA sequences (Larkin et al. 2007). The MEGA software (Kumar et al. 2008) was used to identify each haplotype using the complete deletion method.

The genetic structure of the population includes statistics of Genetic diversity, pairwise F_{ST} values (with 10,000 stages), and analyses of molecular variance (AMOVA) in view of various criteria were determined utilizing the Arlequin3.5 (Excoffier and Lischer 2010). The F_{ST} estimates were utilized to make a Neighbor-Joining (NJ) tree of population. Authentic demography was inspected following two unique approaches: firstly, neutrality tests including Tajima's D (Tajima 1989), Fu's F_s (Fu 1997) and R_2 (Ramos-Onsins and Rozas 2002); secondly, mismatch distribution (MMD; Harpending 1994). The statistical significance of the observed data to the normal dissemination displayed for unexpected for sudden expansion growth determined dependent on sum of squared deviations (SSD) and Harpending's raggedness index (Harpending 1994).

Result and Dissection

PCR products were separated by the agarose gel electrophoresis (Figure .1). HVII amplification, HV2 F and HV2 R primers were used producing a 379bp product

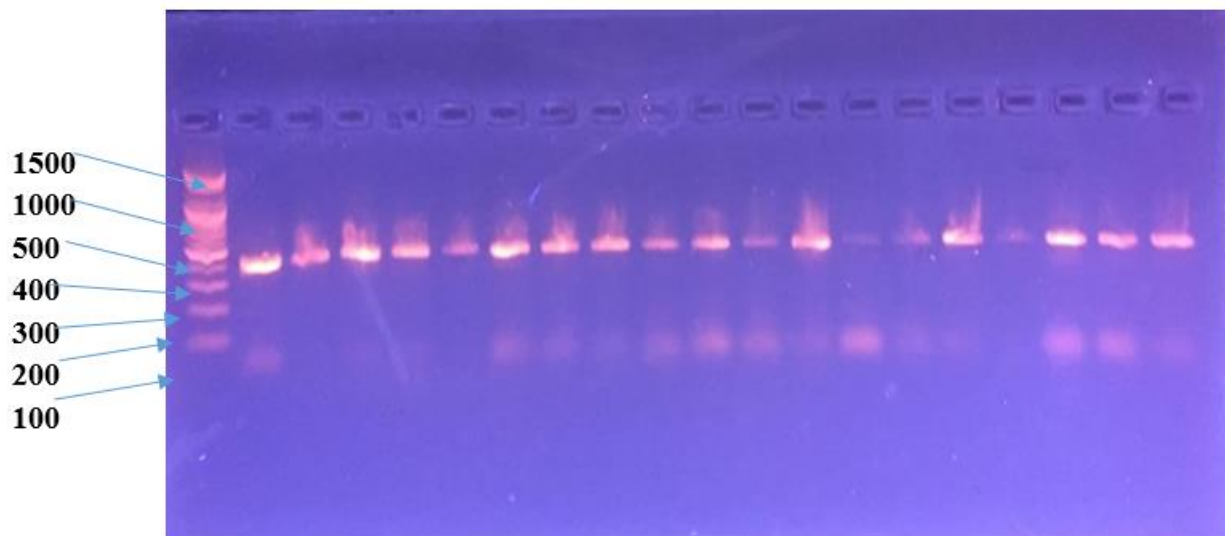


Figure 1: Conventional PCR products of the mitochondrial hypervariable region amplification. PCR products were separated by electrophoresis on 2% agarose gel. 100 bpDNA ladder was used as a molecular marker; Lane M. HV2 PCR products (379 bp).

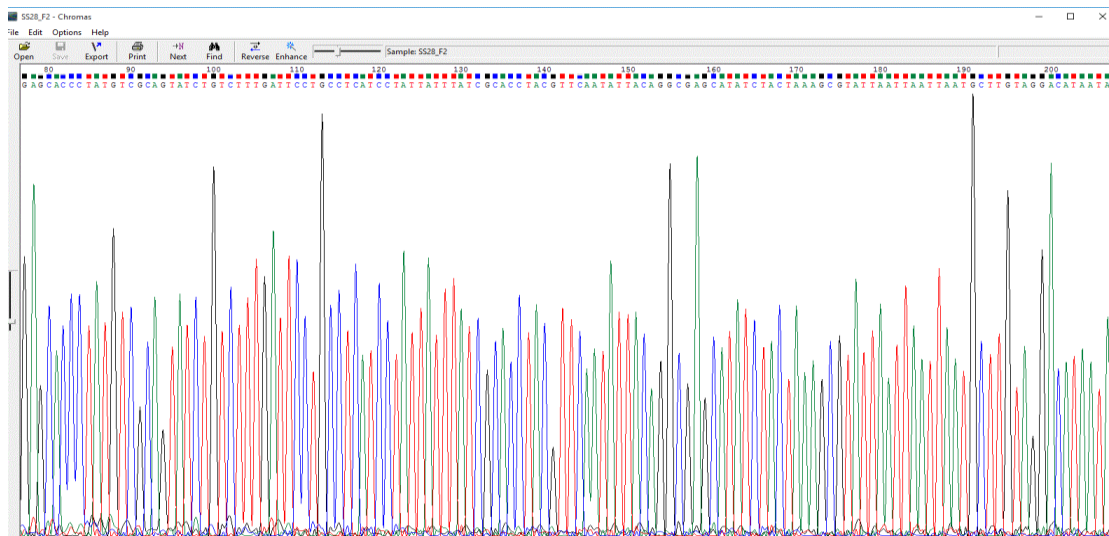


Figure 2: The sequence of HV2 for one sample

The diversity levels and neutrality test scores for the HVS-II data of Basrah and comparative populations show in Table (1). The diversity value in Basrah population was 0.9284, while Iranians have 0.999 genetic diversity, the highest genetic diversity was observed in Armenians, Georgians, Azeris (1.000). The measure of genetic diversity for Turks and Tatars (0.993), (0.998). For the neutrality test statistics, Basrah population shows a significant negative Tajima's D (-0.052) accompanied with a significant Fu's Fs (-2.69062). These results indicate that Basrah population may be undergoing an expansion in a relatively short period of time.

Table1: The studies diversity indices and neutrality test comparing with other populations based on HV2 mtDNA sequences.

Population	No. of samples	No. of haplotypes, h	Haplotype diversity, Hd (S.D.)	Nucleotide diversity, Pi (S.D.)	Tajima's D	reference
Basrah	139	128	0.9284	(0,0012) (0,00005)	-0.0528687(P<0.05	This study
Iranians (Total)	352	315	0,999 (0,000)	0,00208 (0,0004)	-2,58 (P < 0,001)	Derenko <i>et al.</i> 2013
Iranians	30	30	1,000 (0,009)	0,00200 (0,00014)	-2,14 (P < 0,05) *	*Schönberg <i>et al.</i> 2011
Turks	29	27	0,993 (0,013)	0,00173 (0,00015)	-2,41 (P < 0,01) *	*Schönberg <i>et al.</i> 2011
Persians	181	164	0,999 (0,001)	0,00213 (0,00006)	-2,54 (P < 0,001) °	° Derenko <i>et al.</i> 2013
Qashqais	112	94	0,996 (0,002)	0,00204 (0,00007)	-2,4 (P < 0,01) °	° Derenko <i>et al.</i> 2013
Azeris	22	22	1,000 (0,014)	0,00212 (0,00014)	-1,81 (P < 0,05) °	° Derenko <i>et al.</i> 2013
Georgians	28	28	1,000 (0,010)	0,00165 (0,00014)	-2,22 (P < 0,01)*	Fraumene <i>et al.</i> 2006
Azeris	30	29	0,998 (0,009)	0,00208 (0,00012)	-2,09 (P < 0,05) *	Fraumene <i>et al.</i> 2006
Armenians	30	30	1,000 (0,009)	0,00172 (0,00011)	-2,15 (P < 0,05) *	Fraumene <i>et al.</i> 2006
Sardinians	63	50	0,992 (0,004)	0,00147 (0,00011)	-1,79 (P < 0,05) *	Fraumene <i>et al.</i> 2006
Tatars	73	68	0,998 (0,003)	0,00213 (0,00008)	-2,33 (P < 0,01) #	#Malyarchuk <i>et al.</i> 2010

HV2 Haplogroups predicated and variants

An aggregate of 128 haplotypes,(Table: 2 , 3) haplogroups have been recognized in this study. In fact, a large portion of the remainder is spoken to by individualizations. As indicated by their defined or assumed geographic/ethnic starting point (Torroni *et al.*,2006, Roostalu *et al.*,2007, Soares,2010) notwithstanding a solid West Eurasian component (77.8 in the Basrah it is conceivable to perceive commitments from North/East and Sub-Saharan Africa and from East and South Asia. West Eurasian mtDNAs saw in this study are around similarly appropriated into macro Hgs R0, K, U, and J, T, haplogroups and sub-haplogroups In the Basrah contemplate Hg H wins (25%) trailed by Hgs J (12.9%), U (12.3%) and T (3%). On the other hand, both the less spoken to N and W haplogroups display low frequencies (marginally significant) frequencies in this study. The most continuous large macro Hg R0 incorporates atoms R0a (preHV),(I) increasingly spoke to among the Basrah (2%)HV, watched fundamentally as HV , however particularly H mtDNAs. In spite of the fact that most of the H mtDNAs (4%) has not dropped into any of tried sub-haplogroups, a set number of H subsets (H32, H57) . Practically all the primary U sub-haplogroups and the settled branch K has been identified in the Basrah study, however just a sub-set of them (K1, U1, U3, U4, U5) notwithstanding the South West Asian (U7) were seen in the Basrah. The settled Hg K, fundamentally K1, was watched The circumstance of full macro Hg J, T is progressively intricate. Significant differences (P < 0.05) developed in the circulation of J1 and J2 sub-clades, with the last significantly more continuous in the Basrah (10% and 1%). On the other hand, Hg T showed a lower recurrence in the Basrah (3%) because of a huge lower incidence of its T1 sub-clade and T2 sub-clade (1.4% 2.1%, P < 0.05). On the other hand, Hgs N1 (2.8%) and W (3.5%), they were both in Basrah. Haplogroups X was recognized as X2 with

a frequency lower than 1.4%. African haplogroups are of North/East and sub-Saharan African source and minor components recorded at Basrah. The North/East African commitment is principally spoken to by Hg M18 which records for 0.7% of Basrah, the last showing likewise 0.7% of Hg U5. The sub-Saharan African part included Hgs L0, L1, L2 and L3 and represented 3.5% in the Basrah. The Asian commitment was fundamentally higher ($P < 0.01$) in the Basrah, It incorporates mtDNAs having a place with the Southern Asian Hgs M (M18, M70) and R1, R2 and R8a and U1a, U3b. Haplogroups U7a, frequent in Southwest Asia, was watched. The East Asian haplogroups B2, B4 was identified at an exceptionally low frequency in Basrah.

Table 2: summary of HV2 haplo group predicated and variants

No	Sequence	Predicted Haplogroup	Total Variants	Variants
1	HV2b10_F2	H1a(H1a)	4	"A73G, C140d,
2	HV2b13_F2	J1b(J1b7)	4	C166d, G171d"
3	HV2b15_F2	H1a(H1a)	4	"A28C, A73G,
4	HV2b17_F2	H2a(H2a2a1)	2	C141d, C150T"
5	HV2b18_F2	U7a(U7a4a1)	7	"A73G, C140d,
6	HV2b21_F2	B2a(B2a5)	4	T146d, T152d"
7	HV2b22_F2	T1a(T1a1+@152)	4	T346G, T392G
8	HV2b23_F2	J1b(J1b7)	4	"A73G, C141d, T146C, C151T, T152C, T195C,

Table 3: Haplogroups frequency with other population

haplogroup	Estimated mt-DNA haplogroups frequency (%) in population			
	Basrah	Norway	Gurna Egept	Scandinavia
Sample size	139	838		2203
Pre-HV	0.7	Nr	Nr	4.13
HV	1.0	3.6	Nr	
H	25	37.5	14.7	43.26
V	0	0.5	Nr	nr
J	12	15	5.9	10.12
T	3.0	8.5	5.9	1.13
K	0.7	5.5	Nr	5.17
U	12	17.9	8.8	16
I	1.0	1.7	5.9	3.31
X	1.0	nr	nr	nr
W	3.0	1.8	nr	nr
B	7.0	nr	nr	nr
M	1.0	nr	17.6	nr
N or R	5.0	1	8.8	nr
L0-L3#	3.0	Nr	11.7	nr
Other	25.6	7.1	2.9	
	This study	Krzewińska <i>et al</i> , 2015	Stevanovitch <i>et al</i> , 2003	Kivisild <i>et al</i> 2007

HV2 Multiple alignment

The mtDNA control region nucleotide sequences for 139 right whale samples received a consensus region of 379 -bp . Nucleotide diversity and heterogeneity among the Basrah population was evaluated by means of analysis of the sequencing after effects of HV2 region of the mitochondrial genome. The got nucleotide sequence were adjusted against the revised Cambridge Reference Sequence (rCRS) (accession number NC_012920) utilizing CLUSTAL W analysis tool of Biology Workbench at (www5). Occurrences of point heteroplasmy were assessed on the extent of minor versus major component. Varieties from the rCRS spoke to mutation, nucleotide positions of these mutations were resolved and in this manner mitochondrial haplotypes were finished up and connected to the distributed databases.

Table 4: The type, position, and frequency of mutation for haplogroup depend of rCRS

Query	rCRS Position	Query Position	rCRS NT	Query NT	Mut type	GB Freq *	Freq % in haplo	Refs
HV2b17_F2	346	280	T	G	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/746)	G at 346>1 refs
HV2b17_F2	392	326	T	G	transversion	FL: (0.00%)CR: 3 (0.00%)	0.00 (0/746)	G at 392>1 refs
HV2b27_F2	228	181	G	A	transition	FL: 1207 (2.55%)CR: 612 (0.86%)	79.87 (960/1202)	A at 228>52 refs
HV2b27_F2	254	207	T	C	transition	FL: 1 (0.00%)CR: 4 (0.01%)	0.00 (0/1202)	C at 254>1 refs
HV2b27_F2	255	208	G	:	deletion	FL: 0 (0.00%)CR: 0 (0.00%)	0.00 (0/1202)	
HV2b27_F2	257	209	A	G	transition	FL: 118 (0.25%)CR: 77 (0.11%)	0.00 (0/1202)	G at 257>16 refs
HV2b27_F2	259	211	A	G	transition	FL: 22 (0.05%)CR: 27 (0.04%)	0.00 (0/1202)	G at 259>4 refsG at 259>1 unpub
HV2b27_F2	263	215	A	T	transversion	FL: (0.00%)CR: 3 (0.00%)	0.00 (0/1202)	T at 263>1 refs
HV2b27_F2	295	247	C	T	transition	FL: 2224 (4.69%)CR: 1231 (1.73%)	95.76 (1151/1202)	T at 295>55 refs
HV2b27_F2	302	254	A	AA	insertion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/1202)	AA at 302>1 refs
HV2b27_F2	315	268	C	CC	insertion	FL: 14322 (30.21%)CR: 13908 (19.56%)	42.35 (509/1202)	CC at 315>69 refsCC at 315>1 unpub
HV2b27_F2	324	278	C	G	transversion	FL: 4 (0.01%)CR: 47 (0.07%)	0.00 (0/1202)	G at 324>4 refs
HV2b27_F2	325	279	C	G	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/1202)	G at 325>1 refs
HV2b27_F2	328	282	A	G	transition	FL: 6 (0.01%)CR: 38 (0.05%)	0.00 (0/1202)	G at 328>1 refs
HV2b27_F2	330	284	C	G	transversion	FL: 2 (0.00%)CR: 10 (0.01%)	0.00 (0/1202)	G at 330>2 refs
HV2b31_F2	73	185	A	T	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	T at 73>1 refs

HV2b31_ F2	80	192	C	T	transition	FL: 0 (0.00%)CR: 0 (0.00%)	0.00 (0/656)	T at 80>1 refs
HV2b31_ F2	87	199	A	T	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	T at 87>1 refs
HV2b31_ F2	95	207	A	G	transition	FL: 1 (0.00%)CR: 8 (0.01%)	0.00 (0/656)	G at 95>1 refs
HV2b31_ F2	96	208	C	A	transversion	FL: (0.00%)CR: 2 (0.00%)	0.00 (0/656)	A at 96>1 refsA at 96>1 unpub
HV2b31_ F2	98	210	C	A	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	A at 98>1 refs
HV2b31_ F2	134	246	T	G	transversion	FL: 0 (0.00%)CR: 0 (0.00%)	0.00 (0/656)	
HV2b31_ F2	135	247	T	:	deletion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	: at 135>1 refs
HV2b31_ F2	142	253	T	:	deletion	FL: (0.00%)CR: 2 (0.00%)	0.00 (0/656)	: at 142>1 refs
HV2b31_ F2	149	259	T	C	transition	FL: (0.00%)CR: 5 (0.01%)	0.00 (0/656)	C at 149>1 refs
HV2b31_ F2	151	261	C	T	transition	FL: 1627 (3.43%) CR: 1519 (2.14%)	1.52 (10/656)	T at 151>49 refsT at 151>1 unpub
HV2b31_ F2	158	268	T	A	transversion	FL: 1 (0.00%)CR: 1 (0.00%)	0.00 (0/656)	A at 158>1 refs
HV2b31_ F2	165	275	A	C	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	C at 165>1 refsC at 165>1 unpub
HV2b31_ F2	167	277	C	G	transversion	FL: (0.00%)CR: 1 (0.00%)	0.00 (0/656)	G at 167>1 refs

rCRS: revised Cambridge Reference Sequence, GB Freq * :GenBank frequency, ‡ :High Frequency Haplogroups, Refs: MITOMAP References, Unpub: Unpublished Variant.

Neighbor joining tree

(Figure 3)show the comparing study samples with rCRS and drawing the neighborhood tree for samples with rCRS we observe the genetic distance among the samples that is due to the genetic mutations that occurred in haplotype compared with rCRS

12. Kivisild T, Tambets K, Viivi Tolk, Helle, Metspalu E, An mtDNA perspective of French genetic variation, 2007, *Annals of Human Biology* · January 2007, DOI: 10.1080/03014460601076098 ·
13. Koyama H, Iwasa M, Ohtani S, Ohira H, Tsuchimochi T, Maeno Y, Isobe I, Matsumoto T, Yamada Y, Nagao M (2002) Personal identification from human remains by mitochondrial DNA sequencing. *Am J Forensic Med Pathol* 23, 272-276.
14. Krzewińska M, Bjørnstad G, Skoglund P, Olason PI, Bill J, Götherström A, Hagelberg E. *Philos Trans R Soc Lond B Biol Sci.* 2015 Jan 19;370(1660):20130384. doi: 10.1098/rstb.2013.0384. PMID:25487335.
15. Kumar S., Nei M., Dudley J. and Tamura K. 2008 MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* 9, 299–306.
16. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H, Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins D.G. Clustal W and Clustal X version 2.0. (2007) *Bioinformatics*, Volume 23, Issue 21, 1 November 2007, Pages 2947–2948.
17. Lutz S, Weisser HJ, Heizmann J, Pollak S (1996) mtDNA as a tool for identification of human remains. Identification using mtDNA. *Int J Legal Med* 109, 205-209.
18. Lutz S, Wittig H, Weisser HJ, Heizmann J, Junge A, Dimo-Simonin N, Parson W, Edelmann J, Anslinger K, Jung S, Augustin C (2000) Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? *Forensic Sci Int* 113, 97-101.
19. Malyarchuk B., Derenko M., Denisova G. and Kravtsova O. 2010 Mitogenomic diversity in Tatars from the Volga-Ural region of Russia. *Mol. Biol. Evol.* 27, 2220–2226.
20. Ramos-Onsins, S., Rozas, J. Statistical Properties of New Neutrality Tests Against Population Growth, 2003. *Molecular Biology and Evolution* 19(12):2092-100 ·
21. Roostalu U, Kutuev I, Loogväli E-L, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. (2007). Origin and Expansion of Haplogroup H, the Dominant Human Mitochondrial DNA Lineage in West Eurasia: The Near Eastern and Caucasian Perspective; *Molecular Biology and Evolution*, Volume 24, Issue 2, February 2007, Pages 436–448, <https://doi.org/10.1093/molbev/msl173>.
22. Schneider, S., and L. Excoffier. 1999. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics* 152:1079-1089.
23. Schönberg A., Theunert C., Li M., Stoneking M. and Nasidze I. 2011 High-throughput sequencing of complete human mtDNA genomes from the Caucasus and West Asia: high diversity and demographic inferences. *Eur. J. Hum. Genet.* 19, 988–994.
24. Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt H-J, Torroni A, Richard MB (2010). The archaeogenetic of Europe. *Curr Biol* 20:R174-R183.
25. Stevanovitch, A., Gilles, A., Bouzaid, E., Kefi, R., Paris, F., Gayraud, R. P., Spadoni, J. L., El-Chenawi, F., eraud-Colomb, E. B. 2003. Mitochondrial DNA Sequence Diversity in a Sedentary Population from Egypt. *Annals of Human Genetics* (2003) 68,23– doi: 10.1046/j.1529-8817.2003.00057.x 39.
26. Tajima F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
27. Torroni A, Achilli A, Macaulay V, Richards M, Bandelt H-J. Harvesting the fruit of the human mtDNA tree. *Trends in Genetics.* 2006;22:339–45.
28. Tsutsumi H, Mukoyama R, Komuro T (2004) Mitochondrial DNA polymorphism from dental pulp. In *DNA Polymorphism*, Vol 12, Nippon DNA Takei Gakkai, Takatori T eds, Toyo Shoten, Tokyo, 249-251 (in Japanese).
29. Weir BS, Cockerham CC. Estimating F-Statistics for the Analysis of Population-Structure. *Evolution* 38: 1358-1370.