Molecular diversity of mitochondrial DNA in Iranian Azeri ethnicities vis-à-vis other Azeris in Asia

Mohammad Asgharzadeh[1*], Hossein Samadi Kafil[2], Ahmad Ranjbar[3], Ali Rahimipour[4], Kazem Najati[1], Mohammad Rahbani Nobar[1]

[1]Biotechnology Research Center, Tabriz University of Medical Sciences, P.O. Box 51666-14766, Tabriz, I.R. Iran [2]Epidemiology Research Center, Aja University of Medical Sciences, P.O. Box 1411718541, Tehran, I.R. Iran [3]Hematology and Oncology Research Center, Tabriz Univeristy of Medical Sciences, P.O. Box 51666-14766, Tabriz, I.R. Iran [4]Department of Clinical Biochemistry, Shahid Beheshti University of Medical Sciences, P.O.Box 1985717443, Tehran, I.R. Iran

Abstract
In order to investigate the molecular diversity of mtDNA in Azeri population, 133 Azeri subjects inhabiting different regions of Azerbaijan (Iran) were selected. Blood samples were taken from these subjects for mtDNA extraction. The extracted mtDNA samples were then studied by the PCR-RFLP method. Fourteen haplogroups were characterized from which 82% were identified as European specific haplogroups. The H haplogroup was the most frequent and 79 haplotypes were specified. In this study, the Iranian Azeri population was found to be a heterogenic population where all the specific haplogroups of Asians, Europeans and Africans were present in the studied population. Comparing the haplogroups of the present investigation with other populations indicated a very close similarity with other Iranian populations, but was different from haplogroups of other Asian populations who also speak the Azeri language.
Keywords**:** Azeri; Population; Iran; mtDNA; Diversity; Haplogroup

INTRODUCTION

Human mitochondrial DNA (mtDNA) is a circular molecule of 16569 bp, present in thousands of copies per cell (Tagliabracci *et al*., 2001; Pääbo *et al*., 1988). It possesses maternal inheritance, lacks recombination and has a relatively high rate of mutation (Elson *et al*., 2001). The mtDNA mutations lead to changes in nucleotides, which are inherited maternally and spread with migration and colonization of human populations (Torroni *et al*., 1996). Therefore, the study of mtDNA provides a potential explanation for genetic polymorphisms and evolution of populations, and the pattern of migration in different parts of the world.

Polymerase chain reaction-Restriction fragment length polymorphism (PCR-RFLP) is an easy and fast technique for analysis of genetic diversity and variations of mtDNA in human populations (Baudouin *et al*., 2005; Torroni *et al*., 1996). Consequently, it determines population specific haplogroups and haplotypes (Torroni *et al*., 1994; Wallace, 1994). Most of the haplogroups are continent-specific, such as the L haplogroup prevalent in the African population (Chen *et al*., 1995), In fact Africans represent the most ancient human mtDNA group (going back to 200000 years ago), and all modern humans have a common and recent African origin (Vigilant *et al*., 1991). In Siberia, East and South Asia the A, B, C, D, E, G and M haplogroups are predominant, and in native Americans the four Asian haplogroups of A,B,C and D are widespread. In contrast, the H, I, J, K, T, U, V, W and X haplogroups constitute the most common haplogroups in Europe (Herrnstadt *et al*., 2002).

Azeris are one of the biggest Iranian populations who reside mostly in East and West Azerbaijan, Ardabil and the Zanjan provinces; however, they have also settled in other provinces of Iran including Qazvin, Hamadan, Markazi and Tehran (Fig. 1). Historians believe that their historical origin goes back

*Correspondence to: **Mohammad Asgharzadeh, Ph.D.***
Tel: +98 411 3357126; Fax: +98 411 3364666
E-mail: asgharzadehmo@yahoo.com

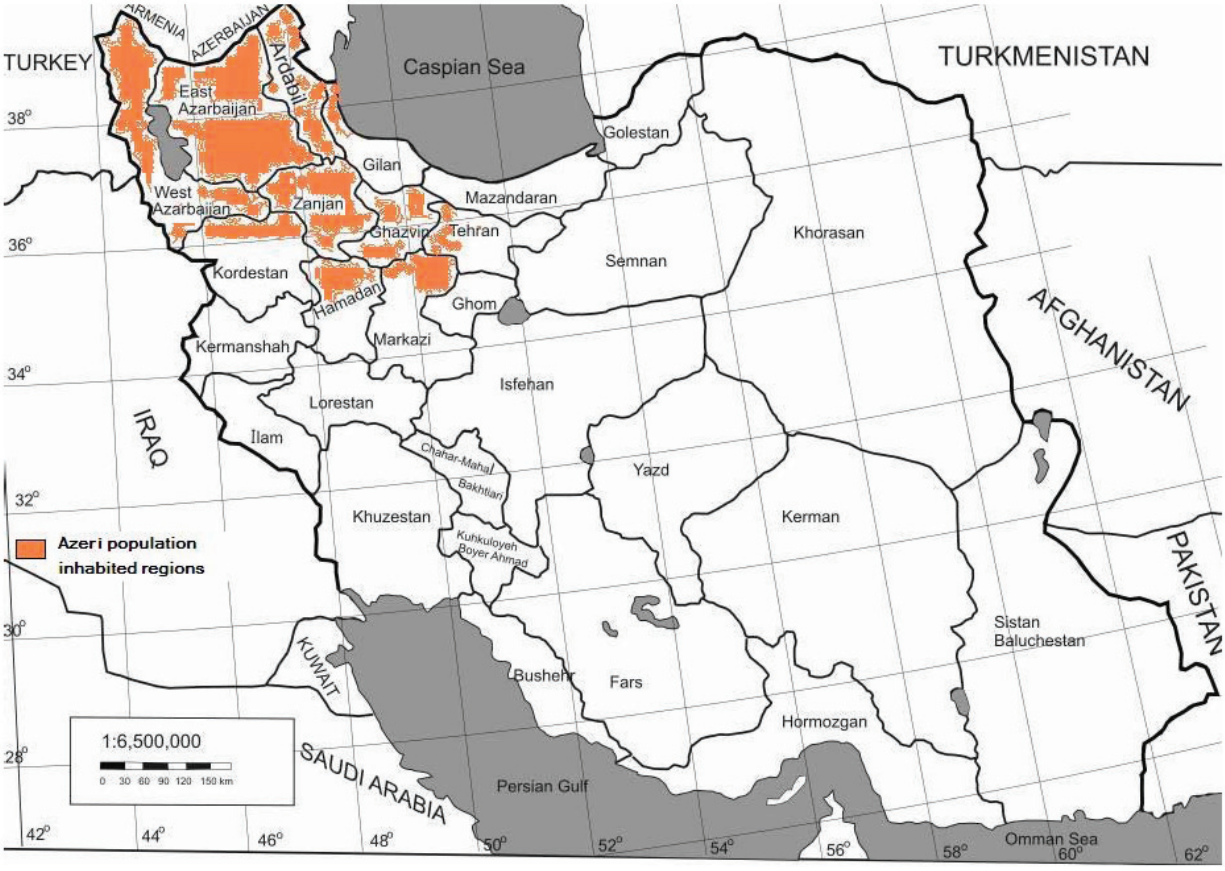

Figure 1. Azeri-inhabited places in Iran.

to the "Maad" ethnicity. The Mads are members of the Indo-Iranian family (Aryans) who migrated to Iran around 1700 BC and settled in the Northwest of Iran (Gharagheshlaghi., 2007). The aim of this study is to investigate the molecular diversity of mtDNA in the Azeri populatiom and to determine the relationship of the present studied population with other Iranian populations, Asian Azeris, Middle Easterners, Central Asians and European populations.

MATERIALS AND METHODS

Subjects: Blood samples were collected from 133 subjects who are currently living in different parts of Azerbaijan. The subjects were not familial relatives, who along with their mothers and grandmothers, were born in Azerbaijan and spoke the Azeri language.

mtDNA molecular analysis: mtDNA was separated from blood cells using proteinase K, sodium dodecyl sulfate (SDS), Cetyltrimethyl ammonium bromide (CTAB) and chloroform (Asgharzadeh *et al*., 2008) and all mtDNAs were amplified in nine different PCR reactions. Primers that were used are presented in Table 1. The properties of the primer pairs were based on previous published work (Anderson *et al*., 1981; Andrews *et al*., 1999).

PCRs were performed in 100 µl volumes which contained 0.5 µM each of the specific primer, 1.5 mM $MgCl_2$, 200 µM each of the dNTPs, 500 mM KCl (pH 8.5), 200 mM Tris-HCl, 2U of recombinant *Taq* DNA polymerase (Cinnagen, Iran) and the extracted mtDNA as template. Finally, the PCR products were electrophoresed on an 0.8% agarose gel and stained with 0.5 µg/ml of ethidium bromide for eventual visualization. The size of the resulting fragments was determined in accordance with a DNA ladder (100 bp DNA ladder plus, Fermentas, Lithuania).

PCR products were purified using high pure PCR product purification kit (Roche, Germany) and digested with *AluI, AvaII, DdeI, HaeII, HaeIII, HhaI, HincII, HInfI, HPAI, RsaI* restriction enzymes. The digested fragments were then electrophoresed on a 2% agarose

Table 1. Primers used for amplification of mtDNA by PCR.

Number of primer	fragment length (bp)	Annealing temperature	Characteristics of 5'→3' primers Forward, reverse*
1	2155	54	1562-1581, 3717-3701
2	2085	54	3108-3127, 5193- 5173
3	2282	62	5151-5170, 7433-7414
4	1796	57	7367-7384, 9172-9154
5	1692	56	8416-8436, 10107-10088
6	2428	61	9911-9932, 12338-12309
7	2196	59	12190-12209, 14385-14366
8	2284	55	14260-14279, 16543-16526
9	1812	65	16453-16472, 1696-1677

*(5'→ 3' before comma refers to forward primers and after comma refers to reverse primer).

gel, and stained as mentioned above. Different mtDNA haplogroups were determined based on the effect of the restriction enzymes on the polymorphic regions with respect to published criteria (Barnabas *et al*., 2006; Al-Zahery *et al*., 2003; Finnilä *et al*., 2001; Macaulay *et al*., 1999; Torroni *et al*., 1996).

RESULTS

In this study, 133 mtDNA samples from the Azeri population were analysed using the PCR-RFLP method and 14 haplogroups were found, where more than 90% of the Azeri population belonged to the I, U, K, T, J, H, HV, W and X haplogroups. The most prevalent of these was the H haplogroup (25.57%) followed by the U haplogroup (20.3%). From the European haplogroups, 8 haplogroups H, I, J, K, T, U, W, X were found in 109 Azeris (81.96%) the West Asian haplogroup, HV, was found to be present in 13 Azeris (9.77%). Three Azeris (2.26%) with the N haplogroup and 4 others with unknown haplogroups (3%) were also observed. From the 8 Asian haplogroups (A-G, M), the A, M and D haplogroups were found in 3 Azeris (2.26%) and the African haplogroup (L) was observed in one Azeri subject (0.75%) (Table 2).

Seventy nine haplotypes were found in 133 Azeri subjects, from which 55 haplotypes were exclusive and 24 haplotypes were prevalent in more than one subject. Seventeen haplotypes were observed in 2 subjects, 2 haplotypes were seen in 3 subjects (one haplotype from the V haplogroup in 4 subjects and one haplotype from the K haplogroup in 5 subjects), one haplotype from the T haplogroup was found in 8 subjects, one haplotype from the U haplogroup was observed in 8 subjects, and one haplotype from the H haplogroup was found to be frequent in 13 subjects.

DISCUSSION

The Azeri population represents one of the Iranian shia-Muslim populations. In this population, specific European (H, I, J, K, T, U, V, W, X), Asians (M, A, D) and African (L) haplogroups of the mtDNA have been observed, but most of this population (82%) possess the European specific haplogroups. In a way, the H haplogroup was shown to be the most frequent hap-

Table 2. Prevalence (frequency) of mtDNA haplogroups in the Azeri population.

Haplogroups	fragment length (bp) Prevalence (%)
HV	13(9.77)
H	34(25.57)
J	15(11.28)
T	15(11.28)
K	8(6.02)
U	27(20.3)
I	6(4.5)
X	1(0.75)
W	3(2.26)
N	3(2.26)
M	1(0.75)
A	1(0.75)
D	1(0.75)
L	1(0.75)
Others	4(3.01)
Total	133(100)

Table 3. Comparing prevalence of Azeris mtDNA haplogroups with other populations.

Population	No. of samples	Haplogroup's prevalence (%)																	
		HV	H	V	J	T	K	U	I	X	W	N	M	A	C	D	G	L	others
Azeris (present study)	133	9.77	25.57	-	11.28	11.28	6.02	20.3	4.5	0.75	2.26	2.26	0.75	1	-	1	-	0.75	3.01
Iranians (Quintana-Murci *et al*., 2004)	42	19.1	14.3	-	16.7	9.5	7.1	16.8	2.4	-	2.4	4.8	4.8	-	-	-	-	-	2.4
Iraqis (Al-Zahery *et al*., 2003)	216	10.6	19.9	0.5	9.3	8.8	3.2	19	1.9	2.8	1.9		1.4					4.2	16.6
Anatolian Turks (Quintana-Murci *et al*., 2004)	50	6	26	-	8	8	8	24	2	6	-	2	-	4	-	2	2	-	4
Turkmen (Quintana-Murci *et al*., 2004)	41	4.8	22	-	9.8	7.3	-	4.8	-	2.4	-	2.4	4.9	2.4	7.3	22	-	-	9.9
Uzbeks (Quintana-Murci *et al*., 2004)	42	7.2	21.4	-	7.1	4.8	-	12	-	-	2.4	-	11.9	7.1	2.4	9.5	2.4	-	11.8
South Siberians (Derenko *et al*., 2003)	480	-	3.5	-	3.5	1.5	0.2	6.7	0.6	0.6	-	0.6	2.7	2.7	36.3	17.9	6.9	-	16.3
Uygurs (Liuqi *et al*., 2008)	50	12	6	-	2	24	-	14	-	-	6	-	14	6	-	10	-	-	6
Pakistanis (Quintana-Murci *et al*., 2004)	100	4	12	-	1	1	-	17	-	1	1	3	47	-	-	-	1	1	11
Italians (Torroni *et al*., 1997)	99		33.3	5.1	7.1	9.1	8.1	22.2	4	3	2		-	-	-	-	-	-	6.1
Norwegians (Passarino *et al*., 2002)	74	-	39.19	5.4	10.81	12.16	4.06	17.5	4.06	1.35	2.7	1.35	-	-	-	-	-	1.35	-
Bosnians (Malyarchuk *et al*., 2003)	144	0.69	47.92	-	6.94	4.86	4.17	19.45	2.78	1.39	1.39	0.69	1.39	-	-	-	-	0.69	7.64
Slovenian (alyarchuk *et al*., 2003)	104	-	47.12	-	9.62	5.77	3.85	19.22	1.92	0.96	4.81	-	-	-	-	-	-	-	6.73
Finishes (Finnilä *et al*., 2001)	480		39.17	5.63	5.4	2.5	2.5	27.92	3.1	1.46	9.6								2.72
Swedish (Torroni *et al*., 1996)	37	-	40.5	5.4	2.7	21.6	13.5	16.2	-	-	-	-	-	-	-	-	-	-	-
Hungarians (Tömöry *et al*., 2007)	101	3	39.6	4.8	7.9	9	7.9	15.9	2	-	7.9	-	-		-				2

logroup (25.57%) in the present population; it is also the most frequent in Europe with a 40-60% frequency in Western Europe (Torroni *et al*., 1998). Moving from West to the East, the frequency of this haplogroup diminishes gradually, i.e. in the Italian population, it has a 33.3% frequency (Torroni *et al*., 1997), with 26% in observed in the Turkish Anatolian population (Quintana-Murci *et al*., 2004), 25.57% in the subjects of the present study, 3.5 % in the South Siberian population (Derenko *et al*., 2003) and the lowest frequency observed in the Indian population (Torroni *et al*., 1998). As shown previously, in one of the studied sub-

jects, the L haplogroup was observed, which is specific to Africa (Chen *et al*., 1995), thus reflecting gene flow from Africa to Azerbaijan. In three subjects the M, A and D haplogroups were observed at a frequency of (2.26%), thus indicating the transfer of gene from Central Asia to Azerbaijan, and as with other Iranians this rate is low (Quintana-Murci *et al*., 2004). The M haplogroup is specific to the Indian peninsula with a 65% frequency (Metspalu *et al*., 2004), but in the Iranian population, it is 2.34% (Houshmand *et al*., 2004). However, in the Azeri population just one subject (0.75%) had this haplogroup, thus indicating that gene flow India to Iran, especially Azerbaijan, was very limited.

Considering the fact that specific haplogroups from the three continents including Europe, Asia and Africa exist in the Azeri population, it can be claimed that the Iranian Northwestern Azeri population is heterogeneous. In the Azeri population, the U haplogroup is the most frequent followed by the H haplogroup (Table 3). This is similar to the finns (Finnilä *et al*., 2001), Italians (Torroni *et al*., 1997) and Bosnians (Liuqi *et al*., 2008). Furthermore, I haplogroup is more frequent in the Azeri population than other populations (4.5%), followed by Norwegians (4.06%) (Passarino *et al*., 2002) and Italians (4%) (Torroni *et al*., 1997).

The Azeri population speaks Azeri, which is one of the Ural-Altaic languages; however, subjects in the present study were genetically different from other Central and East Asian populations that speak the Azeri language (Table 3). Almost 82% of the Azeri subjects in the Northwest of Iran had European haplogroups, and were similar to other Iranians; however, the other Azeri speaking people who live in Asia had mostly the Asian haplogroups, with the European haplogroups being of low frequency among them. For example, the European haplogroup in 7 Azeri speaking populations in Southern Siberia was 16.7% (Derenko *et al*., 2003), in Turkmens 46.3%, and in Uzbeks 47.7% (Quintana-Murci *et al*., 2004).

Considering the similarity between the Azeri and Iranian haplogroups (Table 3), it can be concluded that the Iranian population have been living in Azerbaijan since ancient times, and a small ethnic group who spoke the Altaic language subsequently invaded this region, as confirmed by historical documents. In the 11th century this region was invaded by "Seljuq" Turks (Gharagheshlaghi *et al*., 2007) and "Oghuz" nomadic riders (Johanson *et al*., 1998). A similar situation can also be seen in Iranian Arabs whose mtDNA is closely related to Iranian groups rather than other populations who speak Arabic (a Semitic language) (Nasidze *et al*., 2008). In recent study which compares the mtDNA (by the sequencing method) of Turks and Indo-Iranian Tajik ethnicities in Central Asia, Heyer *et al* (2009) differences have been found between these two ethnic groups. However, the results of this study indicate that the Azeri population and other Iranians have the same genetic properties notwithstanding the fact that they live in different geographical regions and speak different languages.

The Azeri population has similar cultural and religious ties with the rest of the Iranian population, and this research further confirms this by showing that their mtDNA is highly similar to other Iranian populations, but is different from other Sunni-Muslim Turkish populations (Quintana-Murci *et al*., 2004).

Acknowledgments

This study was supported by a grant from Tabriz university of Medical Sciences. We thank Dr. Hossein Navidinia and Dr. Mohammad Momenian for their comments on prepareation of this paper.

References

Al-Zahery N, Semino O, Benuzzi G, Magri C, Passarino G, Torroni A, Santachiara-Benerecetti AS (2003). Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. *Mol Phylogenet Evol*. 28: 458-472.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.

Andrews RM, Kubacka I, Chinnery PF Lightowlers RN, Turnbull DM, Howell N (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 23: 147.

Asgharzadeh M, Kafil HS, Khakpour M (2008). Comparison of mycobacterial interspersed repetitive unit-variable number tandem repeat and IS6110-RFLP methods in identifying epidemiological links in patients with tuberculosis in Northwest of Iran. *Ann Microbiol*. 58: 333-339.

Barnabas S, Shouche Y, Suresh CG (2006). High-resolution mtDNA studies of the Indian population: implications for palaeolithic settlement of the Indian subcontinent. *Ann Hum Genet*. 70: 42-58.

Baudouin S, Saunders D, Tiangyou W, Elson JL, Poynter J, Pyle A, Keers S, Turnbull DM, Howell N, Chinnery PF (2005). Mitochondrial DNA and survival after sepsis: A prospective study. *Lancet* 366: 2118-2121.

Chen YS, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC (1995). Analysis of mtDNA variation in African

populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet*. 57: 133-149.

Derenko MV, Grzybowski T, Malyarchuk BA, Dambueva IK, Denisova GA, Czarny J, Dorzhu CM, Kakpakov VT, Miścicka-Sliwka D, Woźniak M, Zakharov IA (2003). Diversity of mitochondrial DNA lineages in South Siberia. *Ann Hum Genet*. 67: 391-411.

Elson JL, Andrews RM, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (2001). Analysis of European mtDNAs for recombination. *Am J Hum Genet*. 68: 145-53.

Finnilä S, Lehtonen MS, Majamaa K (2001). Phylogenetic network for European mtDNA. *Am J Hum Genet*. 68: 1475-1484.

Gharagheshlaghi AT (2007). Influence of Turish migration in evolution of azerian peoples language. *Faslname olome ensanie Daneshgahe Akzahra*. 17: 93-123.

Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N (2002). Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet*. 70:1152-1171.

Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet*. 70: 1152-1171.

Heyer E, Balaresque P, Jobling MA, Quintana-Murci L, Chaix R, Segurel L, Aldashev A and Hegay T (2009). Genetic diversity and the emergence of ethnic groups in Central Asia. *BMC Genetic*. 10: 49.

Houshmand M, Sanati MH, Vakilian M (2004). Investigation of the mitochondrial haplogroups M, BM, N, J, K and their frequencies in five regions in Iran. *Iranian J Biotech*. 2: 44-48.

Johanson L (1998). The history of Turkic. In the Turkic languages (ed. L Jonson & E Csato), London: Rout Ledge. 81-83.

Liuqi Y, Sijie T, Haijing Y, Zheng B, Qiao E, Dong Y, Zan R and Xiao C (2008). Gene admixture in ethnic populations in upper parts of silk road revealed by mtDNA polymorphism. *Sci Cina ser C-Life Sci*. 51: 435-444.

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonné-Tamir B, Sykes B, Torroni A (1999). The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet*. 64: 232-249.

Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Miścicka-Sliwka D (2003). Mitochondrial DNA variability in Bosnians and Slovenians. *Ann Hum Genet*. 67: 412-425.

Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, Serk P, Karmin M, Behar DM, Gilbert MT, Endicott P, Mastana S, Papiha SS, Skorecki K, Torroni A, Villems R (2004). Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet*. 5: 26.

Nasidze I, Quinque D, Rahmani M, Alemohamad SA, Stoneking M (2008). Close genetic relationship between Semitic-speaking and Indo-European-speaking groups in Iran. *Ann Hum Genet*. 72: 241-252.

Pääbo S, Gifford JA, Wilson AC (1988). Mitochondrial DNA sequences from a 7000-year-old brain. *Nucleic Acids Res*. 16: 9775-9778.

Passarino G, Cavalleri GL, Lin AA, Cavalli-Sforza LL, Børresen-Dale AL, Underhill PA (2002). Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. *Eur J Hum Genet*. 10: 521-529.

Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS, Coppa A, Ayub Q, Mohyuddin A, Tyler-Smith C, Qasim Mehdi S, Torroni A, McElreavey K (2004). Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet*. 74: 827-845.

Tagliabracci A, Turchi C, Buscemi L, Sassaroli C (2001). Polymorphism of the mitochondrial DNA control region in Italians. *Int J Legal Med*. 114: 224-228.

Tömöry G, Csányi B, Bogácsi-Szabó E, Kalmár T, Czibula A, Csosz A, Priskin K, Mende B, Langó P, Downes CS, Raskó I (2007). Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. *Am J Phys Anthropol*. 134: 354-368.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonné-Tamir B, Scozzari R (1998). mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet*. 62: 1137-1152.

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144: 1835-1850.

Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC (1994). mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am J Hum Genet*. 55: 760-776.

Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, Carducci C, Leuzzi V, Carelli V, Barboni P, De Negri A, Scozzari R (1997). Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *Am J Hum Genet*. 60: 1107-1121.

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991). African populations and the evolution of human mitochondrial DNA. *Science* 253: 1503-1507.

Wallace DC (1994). Mitochondrial DNA sequence variation in human evolution and disease. *Proc Natl Acad Sci*. 91: 8739-8746.