UNIVERSITY OF TARTU
FACULTY OF BIOLOGY AND GEOGRAPHY, INSTITUTE OF MOLECULAR AND CELL BIOLOGY, DEPARTMENT OF EVOLUTIONARY BIOLOGY

Table of contents

Abbreviations

(Y)BP	(years) before present
CR	Control Region
CRS	Cambridge Reference Sequence
D-loop	displacement loop (= control region) of mtDNA
HVSI	the first hypervariable segment of the control region of mitochondrial genome
HVSII	the second hypervariable segment of the control region of mitochondrial genome
LHON	Leber hereditary optic neuropathy

Introduction

Literature overview

West–Eurasian mtDNA tree

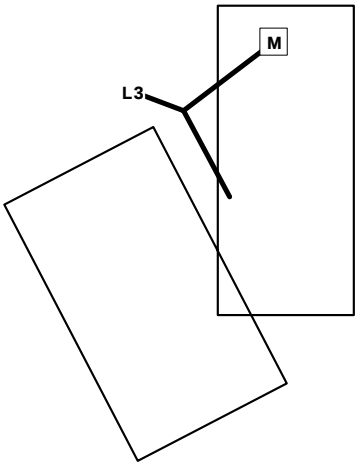
One of the great virtues of mitochondrial DNA (mtDNA) is the opportunity it provides for detailed estimation of the human maternal genealogy. Usefulness of mtDNA molecule in studying humanity's demographic history arises from its large copy number, strictly maternal inheritance (Giles et al. 1980) and general homoplasmy. MtDNAs rarely, if ever mix and recombine (Elson et al. 2001; Ingman et al. 2000). Human mtDNA studies could be broadly categorized on the basis of how much of the molecule was assayed. Human mtDNA was fully sequenced already more than 20 years ago (Anderson et al. 1981), making it possible to design strategies to study its variation in human populations. Historically "low-restriction analysis" with 5 or 6 restriction enzymes, was used to reveal variation in mtDNA pool whereas later on "high-resolution analysis" included already 12 or 14. Sequencing studies have mostly been focused on the control region (CR), or alternatively called displacement loop (D-loop), of the molecule. CR includes first hypervariable segment (HVSI), the region between nucleotide positions (np) 16024–16383 numbered

letters. Four European-specific haplogroups were determined (H, I, J, K) and were defined as follows. Haplogroup J was characterised by the *Bst*NI 13704 restriction site (G to A transition at np 13708) and the *Hinf*I 16065 site (C to T transition at np 16069) losses. Haplogroup H has lost an *Alu*I restrition site at np 7028. Haplogroup I was defined by the *Dde*I np 1715 and the *Hae*II restriction site losses at np 4529 and the *Alu*I

several sub-variants of this theory, largely evolving towards allowing extensive long-distance gene flows (“trellis” models). S

mutations (reversions) that cause difficulties in the estimation of genetic distances and make phylogenetic inferences difficult (Maddison 1991). Thus, one of the problems confusing phylogenetic reconstructions is raised by particularly fast mutation rates -

members of any of the clusters branching from that node, an asterisk (*) to the list of clusters was appended (Macaulay et al. 1999).



particularly variable and were thus called as hypervariable regions (Vigilant et al. 1991). Transitions at sites like 16093, 16129, 16209, 16311 and 16362 in HVSI and 146, 150 and 152 in HVSII occur frequently in many different phylogenetic contexts and these sites can be considered as “mutational hotspots” (Hasegawa et al. 1993; Ingman et al. 2000; Stoneking 2000; Wakeley 1993). The variation of mutation rate is higher in HVSII, where one finds a few sites where substitutions occur very often (observed frequently in different lineages), while most of HVSII shows rather little sequence variation (Aris-Brosou and Excoffier 1996). That is why HVSII segment is usually considered to be less informative than HVSI and has gained little attention in mtDNA studies resolving humanity’s past. Nevertheless, HVSII has provided molecular markers, useful to resolve haplogroup X’ phylogeny (Reidla et al. 2003)**.** In addition to “mutational hotspots” there are some polymorphisms in HVSII that can not be used in phylogenetic analyses. Firstly, transition (G>A) at np 263 represents rare polymorphisms in the CRS mtDNA (Andrews et al. 1999). Secondly, transition (A>G) at np 73 has shared anchestry in haplogroups U, K, T, J, I, W, X , and Z (e.g. [Torroni, 1996; Macaulay et al. 1999), thus can not be regarded as a defining polymorphism. Thirdly, there is a tandem repetitive polycytosine tract, from np 302 to 309 in HVSII know4 Tc 199tes lu6 Tw(rek ahh22.74 -156 ad12 analyses. Fir1,)T0.0742

Coalescent theory tells us what gene genealogies are expected to look like if populations have different demographic histories – that is, how genealogies are affected by changes in population size and structure. Populations may be reduced dramatically in size and subsequently recover (a “bottleneck effect”).

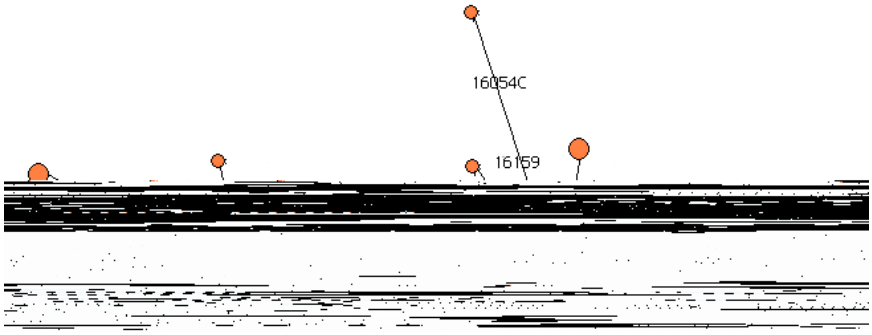
Topology of mtDNA haplogroup J

The first studies with high-resolution restriction mapping divided global mtDNA variation into a number of major ancient clades, called haplogroups (for a review see (Wallace 1995). Haplogroup J was first described in 1994 by Torroni and colleagues (Torroni et al. 1994), it was characterized by the 13704 *Bst*NI site and the 16065 *Hinf*I site losses.

The first study that revealed some inner features of haplogroup J structure was performed by Torroni and et al. in 1997. 37 Italian subjects affected by Leber hereditary optic neuropathy (LHON) were screened for most of the mutations that were known to be associated with LHON at that time. LHON is a maternally transmitted disease in which the primary clinical manifestation is acute or subacute bilateral loss of central vision, leading to blindness. Approximately 90% of LHON cases are associated with and likely directly caused by mtDNA mutations at nps 3460, 11778 or 14484. These are designated as "primary" mutations because they impart a high risk for LHON expression [Wallace, 1988 #957; Howell et al. 1991; Johns et al. 1992; 1993). Mutation at np 3460 is distributed randomly along the phylogenetic tree, without any preferential association with the nine haplogroups (H, I, J, K, T, U, V, W, and X) that characterize European populations, whereas the mutations at nps 11778 and 14484 show a strong preferential association with haplogroup J. Based on RFLP data of LHON affected individuals belonging to haplogroup J, a phylogenetic network was constructed. Haplogroup J was divided into two subclusters, defined the *Alu*I np 7474 (transition C>T at np 7476) and the *Acc*I np 15257 (transition G>A at np 15257) restriction site losses (Torroni et al. 1997).

More detailed phylogenetic networks for European mtDNA were constructed by use of sequence data from HVSI (Richards et al. 1998). Skeleton network for European mitochondrial phylogenetic structure was constructed (Figure 2), based on HVSI data, but also including informative coding region variation, established earlier (Torroni et al. 1996). The major founder subcluster, J*, harboring HVSI mutations 16069 (C>T) and 16126 (T>C) was defined. Several HVSI polymorphisms were found to be haplogroup J specific, i.e. an array of transitions comprising nps 16145 (G>A), 16193 (C>T), 16222 (C>T), 16231 (T>C) and, 16261 (C>T) (Richards et al. 1998).

Figure 2. Schematic tree for European mtDNA variation. Clusters of sequences comprising named clades are outlined and labelled. The node marked CRS corresponds



proportional in lenght to the number of substitutions. Transitions are indicated on the lines by np number; a single transversion present is indicated by np and base abbreviation for a derived state. A reticulation in the network, in which it has been impossible to resolve a recurrent mutation at one or more sites, is represented by a parallelogram-like square where the parallel lines offer alternative evolutionary trajectories in the phylogeny. CRS stands for Cambridge reference sequence [figure from (Helgason et al. 2000)].

A new and an "ultimate" phylogenetic tree-building era came when data about complete mtDNA sequences started to accumulate. For example, phylogenetic networks for mtDNA haplogroup JT were constructed based on 28 sequence data of the complete mtDNA coding region and HVSI motif (Finnilä and Majamaa 2001). According to this study, haplogroup J network found on sequence variation in the coding region, has been divided into subclusters that conformed to the topology proposed before (Torroni et al. 1997). Accordingly, np 7474 *Alu*I and np 15257 *Acc*I site losses were found to define subcluster J2, whereas subcluster J1, which has been defined by the lack of np 15257 *Acc*I site loss (Torroni et al. 1997), was found to be characterized by the G>A transition in the np 3010 that created a restriction site for *Bsh*12306I restriction enzyme at np 3008. A major part of this subcluster was further determined by the transition in the np 14798 (T>C) in the cytochrome b gene. Median-joining networks were constructed separately for complete coding region sequence data (Figure 4) and for the HVSI sequence data (Figure 5). Haplogroup JT network included parallel mutation at np 16172, 16182, 16183, 16186, 16189, 16304 and 16363 and one reticulation. In case of

haplogroup T, the HVSI network correlates well with the network based on the coding region sequence, with the exception of one sample (number 23), which was discrepant between the two networks. In case of haplogroup J, parallel mutations at np 16145,

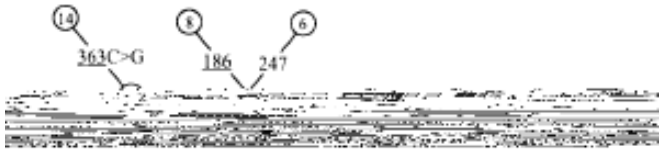


Figure 5. Phylogenetic network of haplogroup JT based on variation in the HVSI sequence. The numbers inside the nodes denote samples and numbers on the lines connecting the nodes denote polymorphic nucleotides, with the first two digits, 16, omitted. Underlined digits indicate parallel mutations. CRS, Cambridge Reference Sequence [figure from (Finnilä and Majamaa 2001)].

In the same year, article by Nicole Maca-Meyer et al. was published (Maca-Meyer et al. 2001). 42 complete human mtDNA lineages were sequenced. Phylogenetic network based on complete mtDNA genome sequences was presented (Figure 6). Due to small sample size, (the tree topology is based on two haplogroup J and two haplogroup T complete sequences), the resolution becomes poor and the tree topology bifurcating, and no confounding conclusions can be drawn. Several parallel mutations were detected, for example, 3010 G>A, has occurred three times in European populations (in the haplogroups J, U and H) and was also found in African haplogroup L2, suggesting either parallel mutations at these sites or very old mtDNA alleles that arose in African mtDNA sequences. Those subsequently evolved from a common ancestor to “newer” European mtDNA sequences in which they underwent “reversion“ on multiple occasions. As for haplogroup J and subhaplogroup H1, the 3010 transition has arisen in the root of the haplogroup, in other cases, as for haplogroups U and L2, the 3010 polymorphism has arisen on the tips of the already formed branches of the phylogenetic tree, suggesting parallel mutations at this site.

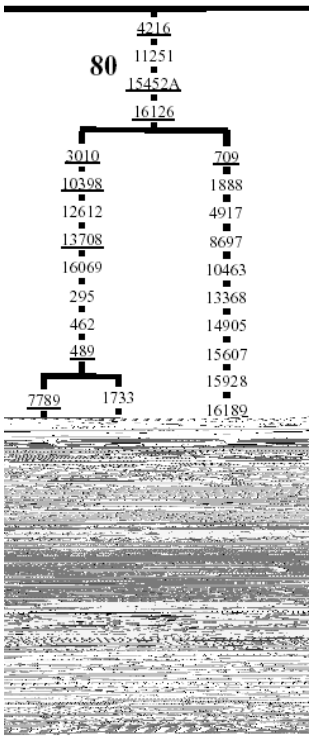


Figure 6. Phylogenetic network for haplogroups J and T based on complete mtDNA genome sequences. Numbers along the links refer to nucleotide positions, suffixes are transversions, underlining indicates recurrent mutations; the order of the mutations on a path not interrupted by any branching or distinguished nodes is arbitrary [figure from [(Maca-Meyer et al. 2001)].

A year later, in 2002, a study based on the nucleotide variation in the coding region of human mtDNA was performed by Corinna Herrnstadt and colleagues (Herrnstadt et al. 2002). 560 complete European, Asian, and African mtDNA coding region sequences from unrelated individuals were analyzed. Phylogenetic networks of African, Asian and European mtDNA sequences based on coding region variations were developed. Here, on Figure 7, haplogroup JT (based on 33 haplogroup J and 46 haplogroup T variants) of the European phylogenetic network from the previous study is presented. The transition in the position 3010 was, in addition to haplogroups H, J, U and L2, also discovered from the Asian haplogroup D, confirming its fast-evolving nature. Several polymorphisms proved to be unique to J haplogroup: variations at nps 5633 (C>T), 7476 (C>T), 10172 (G>A), 10499 (A>G), 12612 (A>G), and others. Several new coding region polymorphisms were found to be unique to haplogroup J:

Hypervariable regions from D-loop were not analysed hence as for J haplogroup, valuable information became lost.

Geographic spread of mtDNA haplogroup J

Average frequency of J haplogroup is the highest in the Near – East (12%) [Richards, 2000 #961; our data) reaching to the highest value, 25%, in the Arabian Peninsula among the Bedouins (Di Rienzo and Wilson 1991). Average haplogroup J frequency starts to decline towards Europe (11%)(Richards et al. 2000), Caucasus 8% (Macaulay et al. 1999; Richards et al. 2000), North – Africa (6%) (Corte-Real et al. 1996; Krings et al. 1999; Pinto et al. 1996; Rando et al. 1998) and becomes practically missing in the East – Asia 1% (Kolman et al. 1996; Yao et al. 2002) and Australia 0% (Huoponen et al. 2001; Redd and Stoneking 1999). In the European populations, J haplogroup is evenly spread (on average 10%), giving occasional peaks i.e. 12% among the Spaniards (Corte-Real et al. 1996; Crespillo et al. 2000; Larruga et al. 2001; Pinto et al. 1996) and 18% among the Italians [Richards, 2000 #961; our data). Frequency pattern can, at least partly, be expounded by the colonization of Europe, North – Africa, and Asia during the Upper Palaeolithic and the Neolithic times. There has been a considerable debate on the respective roles of Palaeolithic (50 000 – 15 000) and Neolithic (7000 – 12 000 YBP) population expansions to the contemporary Europeans' mtDNA gene pool. Studies based on mtDNA haplogroup and haplotype frequencies (Richards et al. 1996; Richards et al. 1998) and the large scale analysis of Y-chromosome markers (Underhill et al. 2001) indicate that European gene pool has mostly Upper Paleolithic origin and the genetic contribution of geographically

growth along an expansion front, of the type defined by (Ammerman and Cavalli-Sforza 1984) for Neolithic Europe as a wave of advance fueled by “demic diffusion”. At another extreme, one might have a progression of saltatory jumps from one suitable

The aim of the present study

Haplogroup J is a variety of human mtDNA pool of a special interest for several important reasons, general for the demographic history of western Eurasian populations, as well as to explore deeper some aspects

Subjects and methods

Samples

712 samples belonging to haplogroup J according to their RFLP and HVSI data were collected from our mtDNA bank. MtDNA genomes from the following populations were represented. Samples were divided into groups, based on geographic of ethnic affinities.

European samples (417 samples).

Scandinavia: 37 Swedes (from different regions in Sweden, including Gotland). Eastern Europe: 38 Estonians, 9 Latvians and 5 Lithuanians, 28 Ukrainians, 12 Russians from Krasnodar, 6 Hungarians, and 31 Russians. Caucasia: 8 Armenians, 23 Ossetians, 2 Adyges, 6 Russians from Adygea, 5 individuals from Karachaev, 6 Nogays, 2 Kumyks, 6 Lezgines, 2 Abazines, and 4 Kabardins. Volga-Uralic region: 3 Permiac Komis, 6 Zyrian Komis, 2 Chuvashes, 24 Khants, 2 individuals of Mordva-Eryza and 4 individuals of Mordva-Moksha ancestry. Central Europe: 18 Slovaks, and 7 Czechs. Southern Europe: 16 Albanians, 27 Turks, 15 Bosnians, 13 Italians, 10 Greeks, 9 Moldovans, 10 Moldaevian Russians, 9 Gagauzes, 11 Cypriots, and 1 Cretan.

Near – Eastern samples (207 samples).

60 Iranians, 36 Kuwaitis, 44 Saudi-Arabians, 8 Lebanese, 24 Syrians, 12 Jordanians, 7 Yemenitess, 11 Omanis, and 5 Jews (Ashkenazim).

Central and North – Asian samples (45 samples).

9 Kazakh-Shoris, 9 persons from Tuva, 1 Yakutian, 1 Ket, 2 Uighurs, 3 Kalmyks, 5 Arabs from Uzbeckistan, 2 Uzbecks, and 13 Kazakhs.

North – African samples (41 samples).

28 Moroccans and 13 Egyptians.

RFLP analyses were performed to the whole haplogroup J sample (712 individuals). HVSII region was sequenced for 306 samples chosen on random from the European samples data set. Although the whole European J haplogroup was planned to be screened for HVSII variation, due to technical problems some of the European nations (i.e. from Volga - Uralic region and Caucasia) that belong to J haplogroup and are

represented in our mtDNA bank, remained unsequenced. In the current study HVSII variation was screened for Armenian, Ossetian, Sweden, Slovakian, Czech, Bosnian, Ukrainian, Latvian, Lithuanian, Estonian, Hungarian, Turkish, Albanian, Greek, Italian, Cypriot, Kumyk, Jewish, and Moroccan populations (altogether 306 samples). Obtained results are shown in Table 2 (see appendixes). According to the results obtained from HVSII variation and combinig that with the coding region and HVSI polymorphisms, three different median networks were constructed. In Figure 9, European J haplotypes are presented in a median joining network. In addition to the European samples that were screened for HVSII and RFLP polymorphisms in the current study, 17 Finns whose CR and coding region polymorphisms were described previously in literature

DNA amplification and sequencing

DNA amplification was performed on 10 – 20 ng of template DNA and was carried out with the thermocycler "Biometra UNO II" usually in total volume of 15-20µl.

Component	Concentration	Concentration in PCR reaction
Tartrazine Buffer or "8,3" buffer or "B" buffer	750 mM Tris-HCl, pH 8.8, 200 mM $(NH_4)_2SO_4$, 0.1% Tween 20; 5% Ficoll 400. 400mM Tri-HCl, pH 8,3, 110µg/ml BSA Courtesy of Solis Biodyne	1/10
$MgCl_2$ dNTP mix (dATP, dCTP, dGTP, dTTP) Taq DNA polymerase (FIREPol) from Solis Biodyne and/or from Juhan Sedman	25 mM $MgCl_2$ 10 mM 2 U/µl	2,5 mM 1 mM 0.125 –0.2 U
F primer R primer Deionized water DNA sample	10 pmol/µl 10 pmol/µl different	~0,2 pM ~0,2 pM 1-3 µl

Oligonucleotide sequences used for PCR and sequencing reactions are listed in Table 1. The cycle profile started with 94^{o}C for 1 min, followed by 35 – 45 cycles of 94^{o}C for 20 s, 52 - 59^{o}C for 15 s and 72^{o}C for 1 min. Number of cycles and annealing temperature depended on primer specifity and mtDNA quality. Incubation with restrictases (0.3-0.5 units per reaction) was done overnight. Sequencing reactions were made using the DYEnamic ET Terminator Cycle Sequencing Kit from Amersham Pharmacia Biotech on a MegaBase 1000 DNA sequencer. PCR primers were destroyed using exonuclease I and free deoxynucleotides were eliminated by shrimp alkaline phosphatase (both from Amersham Pharmacia Biotech). Following reactions were

carried out in 10-µl volume, with using 5µl of purified PCR product and 2µl DYE premix, 1µl of one. As for HVSII region, sequences were obtained between sites 70 and 500. The cycle sequencing profile was 33 to 35 cycles of 94^{o}C for 20s, 50^{o}C for 15 s and 60^{o}C for 1 min. The sequences were aligned and manually checked in SeqLab (GCG Wisconsin Package 10, Genetics Computer Group).

Table 1. Detection of polymorphisms. Amplified sequence or endonuclease restriction

Detection of polymorphisms in haplogroup J samples and data analysis

All the samples belonging to J haplogroup were collected from our mtDNA bank according to their HVSI and RFLP motifs. Information about polymorphic sites in mtDNA coding region was obtained from published complete sequences and RFLP studies (Finnilä and Majamaa 2001; Helgason et al. 2000; Herrnstadt et al. 2002; Maca-Meyer et al. 2001; Macaulay et al. 1999; Richards et al. 1996; Richards et al. 1998; Torroni et al. 1994; Torroni et al. 1997). The whole haplogroup J sample (712 individuals) was screened for the polymorphism at np 13708 and 10398 that are two of the many polymorphisms specific to J haplogroup (Finnilä and Majamaa 2001), in order to exclude the possible earlier misidentifications or mixing sample numbers. All the samples were additionally tested for the polymorphisms at nps 3010 and 7476, in order to divide the samples between two major clades of J haplogroup (Torroni et al. 1997; Finnilä et al. 2001). The samples whose 3010 and 7476 RFLP data or/and 13708 and 10398 RFLP data was controversial were additionally screened for the polymorphisms at np 4216 and 12612, the former defines the common root of JT haplogroup (Macaulay et al. 1999; Richards et al. 1998) and the latter is known to be specific to haplogroup J (i.e. (Finnilä and Majamaa 2001)). The samples harbouring polymorphism at np 16193 according to their HVSI data and loss of *Alu* I restriction site at np 7474, were tested for the polymorphism at np 5633 (Finnilä and Majamaa 2001). The samples harbouring *Alu* I restriction site loss at np 7474, were additionally screened for *Xmi* I (an isoscisomer of *Acc* I) restriction site loss at np 15257. The samples harbouring polymorphism at np 16231 in their HVSI region and *Alu* I restriction site loss at np 7474 (thus belonging to J2 subhaplogroup), were screened by sequencing for polymorphisms at np 7789 and 10499, in order to increase the phylogenetic resolution of the J2 branch.

The second hypervariable region was]TJT*0.aln1

(see Figure 10). In order to obtain true pictures of J haplogroup spread in Europe and in

Results and discussion

In the current study HVSII variation was screened for 306 mtDNA samples of mostly European origin. Polymorphisms at np 7789, 10499 and 14798 were screened by mtDNA sequencing of 306 mtDNA samples. RFLP analysis was performed to coding region polymorphisms at np 3010, 7476, 5633, 10398, 4216, 12612 and 15257 where necessary to mtDNA samples belonging to J haplogroup (712 individuals). RFLP analysis provided us information about the subdivision of J haplogroup samples between two major branches J1 and J2, and in addition to that, J2b. Skeleton network for haplogroup J was constructed according to RFLP data. Further analysis of haplogroup J topology was conducted using the information obtained from the control region variation – by sequencing HVSI and HVSII. 306 HVSI and HVSII haplotypes as well as RFLP data is shown in Table 3 (see appendixes). Nucleotide positions that according to literature were stated as fast evolving (i.e. HVSII positions 146, 150 and 152) were excluded from data analysis. Transition (G>A) at np 263 represents a set of rare polymorphisms in the CRS mtDNA and was excluded from the study. Transition (A>G) at np 73 is not haplogroup J specific but has shared anchestry in haplogroups U, K, T, J, I, W, X and was not included in the present study. Three individuals - Turk 271, Turk 289 and Greek 88 - were excluded from the analysis. The Turks showed controversial HVSI, HVSII and coding region polymorphisms while the Greek was

12306I restriction site at np 3008 and *Alu* I restriction site at np 7474 losses. Loss of *Bsh* 12306I restriction site at np 3008 defines J1 subhaplogroup whereas loss of *Alu* I restriction site at np 7474 defines J2 subhaplogroup. In addition to the loss of *Bsh* 12306I restriction site at np 3008 is subhaplogroup J1 characterized by HVSII polymorphism at np 462. Subhaplogroup J2 is additionally defined by loss of *Acc* I restriction site at np 15257. According to the current study out of 712 haplogroup J samples 86% (613 individuals) belonged into J1 subhaplogroup, 14% (99 individuals) belonged to subhaplogroup J2. Subhaplogroup J1 is divided into 3 subclusters J1a, J1b and J1c. Polymorphism at np 16222 is one of the polymorphisms defining J1a subcluster. Polymorphism at np 14798 is one of the polymorphisms defining J1b subcluster, polymorphism at np 16193 is one of the polymorphisms defining subcluster J1c. J haplogroup samples can be divided between these three J1 subclusters only when their HVSI and coding region polymorphisms are combined and compared. As subhaplogroup J1 is comparatively well represented in all the populations included in this study, then subhaplogroup J2 is on the contrary, less abundant.

Especially confusing in constructing median network for haplogroup J were HVSI polymorphisms at nps 16145 and 16261. An interesting phenomenon was observed. Namely, HVSI polymorphisms 16145 and 16261 seem to be coupled in J haplogroup. Both of them occur as a couple on the branches leading to J1a and J2a subclusters. That was clear already from earlier studies, in particular from complete mtDNA sequence data (Finnilä and Majamaa 2001). In addition, as Figure 9 suggests that the combination 16145 and 16261 has also formed inside J1b subcluster. In case we regard HVSII polymorphisms the defining ones (as it is shown in Figure 11) it is also possible to allocate the subset of J1b linegaes carrying HVSI 16261 into a separate, the fourth subcluster according their HVSII polymorphism at np 188. It is absolutely clear that assuming neutral evolution, probablility of arising this tandem of transitions independently, four times in one single haplogroup, is extremely low ($p<0.01$). In order to study the evolutionary background of these two polymorphisms, the following haplogroup J HVSI lineages 16069-16126-16145 and 16069-16126-16261 and 16069-16126-16145-16261 were selected from the data set and analysed. The task was easier for J1 subhaplogroup since the sample size for this subcluster is much larger. The results for J2a subcluster are statistically less representative because of smaller sample

size. Firstly, hypothesis was considered that 16145-16261 motifs have not formed as parallel mutations on both J1 and J2 branches. As it was mentioned in the literature overview, transition (G>A) at np 3010 has occurred several times in mitochondrial phylogeny and can be excluded from the analyses for being recurrent. However, one may notice that inside haplogroup J, the transition (G>A) at np 3010 has not reverted nor shown any parallel mutation. Furthermore, an even stronger argument in favour of independent origins of 16145, 16261 tandem motifs in J1 and J2 is provided by the fact that the two sub-clades under consideration, differ, in addition, by three mutations that are stable according to mtDNA sequence databases – namely at nps 462 in HVSII and 7476 plus 15257 in the coding region of mtDNA genome (see Figure 9). Even more convincingly, the 16145, 16261 motif in J2a clade is always seen at the background of coding region mutations at nps. 7789 and 10499, not reported to occur within J1 clade. Taken together, irrespective of a low probability of an independent origin of this tandem

in one particular haplogroup of the human mtDNA pool. Neither has this assumed selection led to the fixation of the 16145, 16261 motif in haplogroup J, suggesting that even if true, this selective advantage cannot be high. Alternatively, other coding or/and control region mutations in the remaining branches of haplogroup J that do not encompass this tandem HVSI motif, may have pleiotropic effects that diminish a putative selective effect of the motif under discussion. Summing up, our current knowledge is insufficient to provide an answer to the question whether an apparently independent occurrence of 16145, 16261 tandem motif in two limbs of haplogroup J topology – and only in haplogroup J – is a stocha

glance, some polymorphisms in HVSII seemed to defining, i.e. 3 Turkes formed a lineage acoording to HVSII polymorphism at

at np 16172 seemed to be specific to several populations belonging to subcluster J1a. Unfortunately, certain lineages like the Ossetians and the Armenians lack both the HVSI polymorphism at np 16172 as well as HVSII polymorphism at np 242. Besides, HVSII polymorphism at np 242 has underwent reversion in J1a subcluster and so might have done the HVSI polymorphism at np 16222. Therefore markers additional to 16222 and 242 that would define also the Caucasian share of J1a subcluster would be benefitial to have at hand. All in all additional polymorphic markers besides polymorphisms at np 242, 16172 and 16222 are required in order to specify the topology of J1a subcluster. According to literature, several coding region polymorphisms have been reported to define J1a subcluster (Finnilä and Majamaa 2001). Still it is not known yet wether these coding region polymorphisms coincide with the formation of transition at np 16222.

present. Subclade J2a seems to be the clade that can very well be described by HVSI polymorphisms like 16231 and coding region polymorphisms like 7789 and 10499 and HVSII polymorphism 319. Among the J subclusters all of these polymorphisms are specific to J2a subclade only. J2a subclade is homogenously spread in Europe, the Caucasian nations lack J2a subclade. Regrettably there is no information available so far about Near Eastern J2a subclade haplotypes.

Subcluster J2b1 is characterized by the formation of *Alu* I restriction site at np 5633. Subcluster J2b1 haplotypes are virtually absent in Europe (only 8 Moroccans are allocated into this cluster). In Near East however, subcluster J2b1 is more abundant and diverse. Most of the lineages allocated into J2b subcluster comprise in addition to coding region polymorphism at np 5633 HVSI polymorphism at np 16193 and form J2b2 subclade. J2b2 subclades are present both

Geographic spread of J haplotypes

132 different haplotypes were detected among 417 European samples according to their HVSI and RFLP polymorphisms. All of the haplotypes were homogenously distributed across Europe, showing no clustering to specific geographical regions. Only the Ossetians and the Albanians seem to have experienced population bottleneck or

populations. Additionally, J2 subclade is represented at high frequency (55%) among the Greeks, that of course, might indicate sampling effect.

Estimation of coalescence times

Coalescence times (state the beginning of population expansion in the certain region) for all the major nodes within J haplogroup that expressed more or less star-like topology were calculated (see Table 2). In some cases, when the branching pattern was being far from star-like for both Near Eastern as well as European haplogroup J graphs (as for J1c subcluster), the coalescence time was not estimated. When comparing two median joining networks, constructed of the European and Near Eastern J haplotypes (Figure 9 and Figure 10, respectively), we can see that J1b subclade harbors the most star-like topology in the European median joining network. Thus, the time estimate dating the subclade J1b population expansion will be more precise than that calculated i.e. for J2b2 subclade in Europe. As for J1 branch, J1b subclade is much younger in the European populations than in the Near Eastern ones. According to the present study, subclusters J1b as well as J1a2 date to the Neolithic times in Europe. More difficulties arise with the "Caucasian" share of J1 subcluster, J1a1. In the European median joining network subcluster J1a1 looks star-like, well, in a way. Coalescence time estimate obtained for the European J1a1 subcluster becomes, for some reason, unexpectedly high. Whereas the tip of J1a subclade, J1a2 in Europe, remains much younger. So, J1a1 is older in Caucasus than in Near East. There are some long branches protruding from the central node of Caucasian J1a1 subcluster (i.e. 16193-16362, 16189-16235 and 16189-16187) that may account for the old age of the Caucasian J1a1. When evaluating the coalescence time for this subcluster, the Ossetians, who have experienced population bottleneck or founder effect during their demographic history were excluded from the calculations for statistical reasons. Still, the time of population expansion becomes about 27000 years. Phenomenon alike the one described here, has been reported earlier, in case of subhaplogroup T1, whose expansion time was also evaluted to be about of the same magnitude in Caucasus (Metspalu et al. 1999). According to the coalescence times obtained for J1a1 subcluster we may assume that the population expansion in Caucasus occurred during the Upper Palaeolithic. Human dispersals might have started from Caucasus region , towards Near East and then, later on, during the

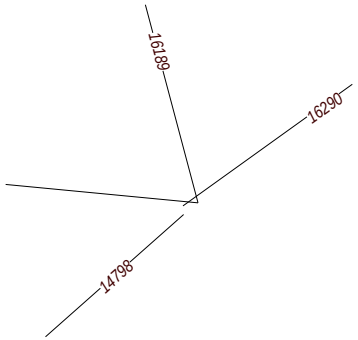
Table 2. Coalescence times calculated for the major nodes within J haplogroup in Europe and in Near East.

Subclade of J haplogroup	Coalescence time in Europe	Coalescence time in Near East
J1a1	27300 (standard error +/- 8000 years)	17700 (standard error +/- 2500 years)
J1a2	7700 (standard error +/- 3500 years)	-
J1b	5000 (standard error +/- 2200 years)	23300 (standard error +/- 4300 years)
J2a	19200 (standard error +/- 6900 years)	-
J2b1	-	15000 (standard error +/- 5000)
J2b2	161600 (standard error +/- 8100 years)	16000 (standard error +/- 5700 years)
J2b3	5800 (standard error +/- 2900 years)	-

Neolihtic, towards Europe. Nevertheless, subcluster J1a1 haplotype diversities in Caucasus are lower than those in Near East, increasing credibility to the fact that J1a1 subcluster still may descend from Near East. However, haplotype diversity is tightly connected with the number of samples under investigation, therefore, we can attribute the low haplotype diversity in Caucasus to small sample size. However, as discussed before, when dating population expansions by coalescence time estimates one has to be cautious, due to probable errors in time estimates that may arise from non-starlikeness of an evolutionary tree.

In case of J2 subhaplogroup, the time estimates evaluating the population expansions, become older compared to those of J1 subhaplogroup. That will apparently account for Europe as well as for Near East although in the current study there were no Near Eastern haplotypes belonging to J2a subclade represented. Therefore, J2 subhaplogroup in Europe can be considered as “old”, probably descending from the population expansions during the Upper Palaeolithic. The expansion of the major clades in J1 subhaplogroup (J1a and J1b) has started around Last Glacial Maximum (about 20 000 YBP) in Near East, J2b subcluster expanded about at the same time or later, during the early Natufian, when extreme aridity in Anatolia was over and the Caucasus glaciers started to melt. In conclusion, the picture emerges showing that J haplogroup (as a whole) is older in Near East than in Europe. There has been a remarkable gene flow

from the Near East into Europe during European prehistory. All the European major haplogroup J clusters date back to the Neolithic, the only exception is the Caucasian J1a1, that might rather originate from Caucasus than from Near East. Due to its location between the Near East, Europe and Central Asia, genetics of the Caucasus region populations may have an important role in the reconstruction of the dispersal routes of modern humans, including from the very beginning of the colonization of Eurasia. On the other hand, a possibility exists that some lineages that are by now widely spread in Near East, might have emerged from there.


16189
16290
14798


16086
16290
16189
16185
16373
16309
3010
16169
4216 16069

Summary

The current study is in concordance with the previous statements referring to the Near Eastern ancestry of some European J subclusters dating back to the Neolithic times. This probably accounts for the vast majority of European J subclusters, especially for subcluster J1b that is being the most frequent and diverse among the European J lineages. Still, J1a1 subcluster in Near East may originally descend from the Caucasus that according to time estimates dating the population dispersal, is about 27 000 years old in the Caucasus region.

According to the present study, in order to refine mtDNA J haplogroup's genealogy, information about polymorphisms in the mtDNA coding region as well as from HVSI and HVSII was combined. . The most parsimonious evolutionary tree for J haplogroup was presented according to HVSI, HVSII and coding region polymorphisms.According the results of the current thesis, conclusion can be drawn that HVSII region contains little information in resolving J haplogroup's topology. Although several HVSII polymorphisms are specific to J haplotypes on the whole, they can nthoug2057f Europholefine -6 -118 Twn3/579u0.099

Kokkuvõte (Summary in Estonian)

Inimese mitokondri genoom on 16569 aluspaari pikkune DNA rõngasmolekul. Enamuses inimese rakkudest on kuni mitu tuhat mitokondriaalse DNA (mtDNA) koopiat. MtDNA-d iseloomustab unilineaarne päritavus, ta pärandub järglastele ainult emapoolselt vanemalt ega rekombineeru. Mitokondriaalse DNA võib jagada kaheks fuktsionaalselt erinevaks osaks: kodeerivaks ja mittekodeerivaks piirkonnaks. Mittekodeeriv ala omakorda jagatakse kaheks hüpervarieeruvaks piirkonnaks. Mittekodeeriv osa evolutsioneerub umbes kümme korda kiiremini kui kodeeriv osa, mille evolutsiooni aeglustavad funktsionaalsed piirangud. Mitterekombineeruvus, kiire ja aeglase evolutsiooniga piirkondade olemasolu ning suur koopiaarv teevad mitokondriaalsest DNA-st sobiva tööriista inimpopulatsioonide populatsioonide päritolu ja ajaloo uurimisel. MtDNA nö. võidukäik inimkonna demograafilise ajaloo uurimisel algas 1980-ntatel aastatel, kui sekveneeriti esimene inimese mtDNA täisjärjestus.

Käesoleva töö eesmärgiks oli suurendada mtDNA haplogupp J fülogeneesi lahutusvõimet, uurida J haplogrupi topoloogiat ja vanust nii Lähis-Idas kui ka Euroopas. Selleks analüüsiti nii eelnevalt määratud J haplogrupi täisjärjestusi kui ka I hüpervarieeruva piirkonna polümorfisme ja vastavalt kirjanusest saadud andmetele, otsustati täpsemalt uurida II hüpervarieeruva ala polümorfset mustrit. Käesolevas töös uuriti täpsemalt 306 Eurooplase ja 207 Lähis-Idast pärineva J haplogruppi kuuluva indiviidi mtDNA II hüpervarieeruva regiooni polümorfismide mustrit ja kombineerides seda nii I hüpervarieeuva piirkonna kui ka kodeeriva ala polümorfismidega, esitati vastavalt saadud andmetele kõige “säästlikum” J haplogrupi evolutsiooniline puu. Selgus, et II hüpervarieeruv piirkond üldiselt ei ole piisavalt informatiivne selleks, et seda edspidi J haplogrupi topoloogia uurimisel kasutada. Sellest hoolimata leiti II hüpervarieeruvast piirkonnast kolm polümorfismi (295, 319, 462), mida saab J haplogrupi puhul lugeda defineerivateks. Vastavalt analüüsitud I hüpervarieeruva piirkonna polümorfismidele jagati J haplogrupp J1 ja J2 alamhaplogruppideks, need alamhaplogrupid omakorda alamklastriteks.

J haplogrupi vanus Lähis-Idas on tunduvalt suurem kui Euroopas. Enamus Euroopa J haplogrupi alamklastreid pärineb ilmselt neoliitilisest ekspansioonist Lähis-Idast. Seda seostatakse nö. “neoliitilise revolutsiooniga”, maaviljeluse ja karjakasvatamise algusega Lähis-Idas, mis tõi kaas rahvaarvu suurenemise vastavas piirkonnas ning populatsiooni ekspansiooni Euroopa suunas. Samas J2 alamhaplogrupp Euroopas võib pärineda

neoliitikumi eelsest paleoliitilisest ekspansioonist. Vastavalt J haplogrupile koonduvad Kaukaasia rahvad põhiliselt J1a1 alamklastrisse, mille vanuseks käesolevas töös hinnati ligikaudu 27 000 aastat. Sama alamklastri vanuseks Lähis-Idas aga ligikaudu 18 000 aastat. Seega hoopis Kaukaasia võib olla mõnede J hapogrupi liinide algkoduks.

Kokkuvõttes võib öelda, et töö tulemusena saavutati J haplogrupi senisest märksa sügavam fülogeneetiline resolutsioon ja saadud tulemused on nüüdsest peale kasutatavad fülogeograafilises analüüsis. Siiski nõuab J haplogrupi evolutsioonilise puu topoloogia detailiseerimine suurt edasist tööd: kodeeriva ala täielikku polümorfismide andmestikku ja selle käigus leitud uute mutatsioonide uurimist nende fülogeograafilise informatiivsuse selgitamiseks.

Acknowledgements

It has been a privilege to have Professor Richard Villems and Dr. Ene Metspalu as supervisors. I thank them both for the professional guidance in this interesting field of science. Many thanks to Ille Ilpus, Hela Help, Jaan Lind and Jüri Parik for technical assistance. Special thanks to Professor Elsa Khusnutdinova for access of large set of samples, Marina Bermisheva, Oleg Balanovsky, Larissa Tamba, Marina Gubina, Vladislava Gusar, Lovorka Barać, Marijana Perečić, Svetlana Cvietan, Erwan Pennarun, and Doron Behar. I would like to express my gratitude to all the co-workers in our laboratory whose previous work has supplied me with J haplogroup HVSI sequences: Katrin Kaldma, Ene Metspalu, Toomas Kivisild, Kristiina Tambets, Jüri Parik, Sirle Laos and Karmen Mae.

Publications

Tambets K, Rootsi S, Kivisild T, Help H, **Serk P**, Loogvali EL, Tolk HV, Reidla M, Metspalu E, Pliss L, Balanovsky O, Pshenichnov A, Balanovska E, Gubina M, Zhadanov S, Osipova L, Damba L, Voevoda M, Kutuev I, Bermisheva M, Khusnutdinova E, Gusar V, Grechanina E, Parik J, Pennarun E, Richard C, Chaventre A, Moisan JP, Barac L, Pericic M, Rudan P, Terzic R, Mikerezi I, Krumina A, Baumanis V, Koziel S, Rickards O, De Stefano GF, Anagnou N, Pappa KI, Michalodimitrakis E, Ferak V, Furedi S, Komel R, Beckman L, Villems R.

The western and eastern roots of the Saami--the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes.
Am J Hum Genet. 2004 Apr;74(4):661-82

References

Alves-Silva J, da Silva Santos M, Guimarães PE, Ferreira AC, Bandelt H-J, Pena SD, Prado VF (2000) The ancestry of Brazilian mtDNA lineages. Am J Hum Genet 67:444-61

Ammerman AJ, Cavalli-Sforza LL (1984) The Neolithic transition and the genetics of

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

Corte-Real HB, Macaulay VA, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, Papiha S, et al (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. Ann Hum Genet 60:331-50

Crespillo M, Luque JA, Paredes M, Fernandez R, Ramirez E, Valverde JL (2000) Mitochondrial DNA sequences for 118 individuals from northeastern Spain. Int J Legal Med 114:130-2

Di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human

Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, et al (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.

Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, Taylor L (1991) Leber hereditary optic neuropathy: indentification of the same mitochondrial ND1 mutation in six pedigrees. Am J Hum Genet 49:939-950

Huoponen K, Schurr TG, Chen Y, Wallace DC (2001) Mitochondrial DNA variation in an aboriginal Australian population: evidence for genetic isolation and regional differentiation. Hum Immunol 62:954-69.

Ingman M, Gyllensten U (2001) Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. J Hered 92:454-61.

Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. Nature 408:708-713

Johns DR, Neufeld MJ (1992) Cytochrome b mutations Leber hereditary optic neuropathy. Biochem Biophys Res Commun 181:1358-1364

Johns DR, Heher KL, Miller NR, Smith KH (1993) Leber's hereditary optic neuropathy: clinical manifestations of the 14484 mutation. Arch Ophthalmol 11:495-498

Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, et al (1999) Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. Curr Biol 9:1331-1334

Kivisild T, Tolk H-V, Parik J, Wang Y, Papiha SS, Bandelt H-J, Villems R (2002) The emerging limbs and twigs of the East Asian mtDNA tree. Mol Biol Evol 19:1737-1751

Kolman C, Sambuughin N, Bermingham E (1996) Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142:1321-34

Krings M, Salem AE, Bauer K, Geisert H, Malek AK, Chaix L, Simon C, et al (1999) mtDNA analysis of Nile River Valley populations: A genetic corridor or a barrier to migration? Am J Hum Genet 64:1166-76.

Larruga JM, Diez F, Pinto FM, Flores C, Gonzalez AM (2001) Mitochondrial DNA characterisation of European isolates: the Maragatos from Spain. Eur J Hum Genet 9:708-16.

Maca-Meyer N, Gonzalez AM, Larruga JM, Flores C, Cabrera VM (2001) Major genomic

mitochondrial lineages delineate early human expansions. BMC Genet 2:13

Macaulay VA, Richards MB, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, et al (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64:232-249

Maddison D (1991) African origin of human mitochondrial DNA reexamined. Syst Zool 40:355-363

Metspalu E, Kivisild T, Kaldma K, Parik J, Reidla M, Tambets K, Villems R (1999) The Trans-Caucasus and the Expansion of the

pool. Am J Hum Genet 59:185-203

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, et al (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67:1251-1276

Richards MB, Macaulay VA, Bandelt H-J, Sykes BC (1998) Phylogeography of mitochondrial DNA in western Europe. Ann Hum Genet 62:241-60

Rose G, Passarino G, Carrieri G, Altomare K, Greco V, Bertolini S, Bonafè M, Franceschi C, De Benedictis G (2001) Eur Jour of Hum Genet 9:701-707

Saillard J, Forster P, Lynnerup N, Bandelt H-J, Nørby S (2000) mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. Am J Hum Genet 67:718-726

Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R Martin W (2002) Genetics and geography of wild cereal domestication in Near East Nat Rev Genet 3:429-41

Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613-23

Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS (1996) A view of the neolithic demic diffusion in Europe through two Y chromosome-specific markers. Am J Hum Genet 59:964-8

Sigurđardóttir, Helgason A, Gulcher JR, Stefansson K, Donnelly P (2000) Am J Hum Genet 66:1599-1609

Stoneking M (2000) Hypervariable sites in the mtDNA control region are mutational hotspots. Am J Hum Genet 67:1029-32.

Stringer CB, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. Science 239:1263-8

Templeton AR (1992) Human origins and analysis of mitochondrial DNA sequences. Science 255:737

Torroni A, Bandelt H-J, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, et al (1998) mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137-52

Torroni A, Chen YS, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, et al (1994) mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico [see comments]. Am J Hum Genet 54:303-18

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, et al (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835-50

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 55:760-76

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994) Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci U S A 91:1158-62

Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, Carducci C, et al (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 60:1107-21

Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley R, Oefner PJ, et al (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet. 65:43-62

Van Andel T, Runnels C (1995) The earliest farmers in Europe. Antiquity 69:481-500

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science 253:1503-7

Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. J Mol Evol 37:613-23

Wallace D (1995) Mitochondrial DNA variation in human evolution, degenerative disease, and aging. Am J Hum Genet 57:201-23

Wolpoff M, Wu XZ, Thorne AG (1984) Modern Homo sapiens origins: A general theory of hominid evolutioninvolving the fossil evidence from East Asia. In: Smith FH, Spencer F (eds) The origin of modern humans: A world survey of the fossil evidence. Alan L. Riss, New York, pp 411-483

Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635-51.

Appendixes

[redacted]

										-				T		J1a
										-				T		J1a
Ossetia	86	69-126-145-222-261-274	73-199-263-295-ins310C-ins316C-462	-	+	-			-				T		J1a	
Ossetia	88	69-126-145-222-261-274	73-199-263-295-ins310C-ins317C-462	-	+	-			-				T		J1a	

Slovakkia	46	69-126-189	73-151-185-228-263-295-ins310C-ins316C-462	-	+	-			+				C		J1b
Slovakkia	87	69-126-290	73-185-188-228-263-295-ins310C-ins316C-462	-	+	-			-				C		J1b
Slovakkia	79	69-126-223	73-146-185-228-263-295-ins316C-462	-	+	-			-				C		J1b
Slovakkia	33	69-126-160-222	73-185-228-263-295-ins310C-ins316C-462	-	+	-			-				C		J1b
Slovakkia	86	69-126-293	73-228-263-295-ins316C-462	-	+	-			-				C		J1b
Slovakkia	27	69-126-301	73-185-188-228-263-295-ins310C-ins316C-462	-	+	-			-				C		J1b

Ukraine	93-01	69-126-222	73-195-228-263-295-ins316C-462-489	-	+	-			-				C		J1b
Ukraine	102-02	69-126-172-189	73-228-263-295-ins316C-462	-	+	-			+				C		J1b
Ukraine	78-01	69-126-189	73-151-185-228-263-295-ins310C-ins316C-462	-	+	-			+				C		J1b
Ukraine	25-02	69-126-287	73-185-263-295-ins316C-368-462	-	+	-			-				C		J1b
Ukraine	131-02	69-126-145-231-261-355	73-150-152-215-195-263-295-ins311T-ins316C-319-430	+	-	-	-	A	-			G	T	-	J2a
Ukraine	161-02	69-126-145-172-222-261-362	73-242-263-295-ins316C-462	-	+	-			-				T		J1a
Ukraine	52-01	69-126-193-278-311	73-150-152-263-295-ins310C-ins316C	+	-	-	+		-				T	-	J2b
Ukraine	104-01	69-126	73-185-228-263-295-ins316C-462	-	+	-			-				C		J1b
Ukraine	111-01	69-126	73-146-185-188-228-263-295-ins316C-462	-	+	-			-				C		J1b
Ukraine	126-01	69-126	73-185-263-265-ins316C-462	-	+	-			-				C		J1b

Estonia	29	69-261	73-146-185-228-263-295-ins316C-462	-	+	-	-	C	J1b

Greece	38	69-126-145-187-189-261	73-199-242-263-265-295-ins310C-ins316C-462	-	+	-			+				T		J1a
Greece	58	69-126-145-231-284	73-150-152-195-215-263-295-ins310C-ins316C-319	+	-	-	-	A	-			G	T	-	J2a
Greece	83	69-126-145-231-284	73-152-195-215-263-295-ins310C-ins311C-ins316C-319-441	+	-	-	-	A	-			G	T	-	J2a
Greece	64	69-126-145-231-261-311	73-150-152-195-203-215-263-295-ins316C-319	+	-	-	-	A	-			G	T	-	J2a
Greece	74	69-126-145-231-261-311	73-152-195-203-215-228-263-295-ins316C-319	+	-	-	-	A	-			G	T	-	J2a
Greece	100	69-126-145-231-261-311	73-150-152-195-203-215-263-295-ins316C-319	+	-	-	-	A	-			G	T	-	J2a

Marocco	91E	69-126-169-193-278	73-150-152-185-263-295-ins316C-462	+	-	-	+		-			T	-	J2b
Marocco	1F													