DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
92

TOWARDS THE UNDERSTANDING OF POST-GLACIAL SPREAD OF HUMAN MITOCHONDRIAL DNA HAPLOGROUPS IN EUROPE AND BEYOND: A PHYLOGEOGRAPHIC APPROACH

KRISTIINA TAMBETS

TARTU UNIVERSITY
PRESS

Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy (in molecular biology) on May 18, 2004 by the Council of the Institute of Molecular and Cell Biology, University of Tartu.

Opponent: Prof. Kari Majamaa, MD, PhD, Department of Neurology, University of Oulu, Finland

Commencement: Room No 217, Riia 23, Tartu on June 18, at 9.15

The publication of this dissertation is granted by the University of Tartu

Tartu Ülikooli Kirjastus
www.tyk.ut.ee
Tellimus nr. 258

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The current dissertation is based on the following publications referred to in the text by their Roman numerals:

I. **Tambets, K.,** Rootsi, S., Kivisild, T., Help, H., Serk, P., Loogväli, E.-L., Tolk H.-V., 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, J.-P., Barać, L., Peričić, M., Rudan, P., Terzić, R., Mikerezi, I., Krumina, A., Baumanis, V., Koziel, S., Rickards, O., De Stefano, GF., Anagnou, N., Pappa, K.I., Michalodimitrakis, E., Ferák, V., Füredi, S., Komel, R., Beckman, L., Villems, R. (2004) The western and eastern roots of the Saami — the story of genetic "outliers" told by mtDNA and Y-chromosome. *American Journal of Human Genetics* 74, 661–682.

II. Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H.-V., **Tambets, K.**, Forster, P., Karger, B., Francalacci, P., Janicijevic, B., Rudan, P., Rickards, O., Savontaus, M.-L., Huoponen, K., Laitinen, V., Koivumäki, S., Sykes, B., Novelleto, A., Moral, P., Sellitto, D., Santachiara-Benerecetti, A.S., Semino, O., Scozzari, R. (2001) A Signal, from human mtDNA, of postglacial recolonization in Europe. *American Journal of Human Genetics* 69, 844–852.

III. Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., **Tambets, K.,** Tolk, H.-V., Parik, J., Loogväli, E.-L., Derenko, M., Malyarchuk, B., Bermisheva, M., Zhadanov, S., Pennarun, E., Gubina, M., Golubenko, M., Damba, L., Feodorova, S., Gusar, V., Grechanina, E., Mikerezi, I., Moisan, J.-P., Chaventre A., Khusnutdinova, E., Osipova, L., Stepanov, V., Voevoda, M., Achilli, A., Rengo C., Rickards, O., De Stefano, G. F., Papiha, S., Beckman, L., Janicijevic, B., Rudan P., Anagnou N., Michalodimitrakis, E., Koziel, S., Usanga, E., Geberhiwot, T., Herrnstadt, C., Howell, N., Torroni, A., Villems, R. (2003). Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics* 73, 1178–1190.

IV. **Tambets, K.**, Tolk, H.-V., Kivisild, T., Metspalu, E., Parik, J., Reidla, M., Voevoda, M., Damba, L., Bermisheva, M., Khusnutdinova, E., Golubenko, M., Stepanov, V., Puzyrev, V., Usanga, E., Rudan, P., Beckman, L., Villems, R. (2003) Complex signals for population expansions in Europe and beyond. — In: *Examining the Farming/Language Dispersal Hypothesis* (Cambridge University Press), 449–457.

V. **Tambets, K.**, Rootsi, S., Kivisild, T., Villems, R. (2001) The concepts of Richard Indreko about the origin of the Finno-Ugric speakers and the population genetics of the extant North-East European populations. *TRAMES*, 5 (55/50), 1, 59–74.

My contribution to the articles referred in the current thesis is as follows:

Ref. **I** — a) conceived and designed the mtDNA experiments; b) performed the mtDNA experiments of 160 U5 genomes, analyzed the mtDNA variation of the populations listed in *Subjects and Methods*; c) performed the statistical and phylogeographical analysis of mtDNA; d) wrote the paper;

Ref. **II** — a) participated in performing the experiments: analyzed the mtDNA variation (HVS-I sequencing and RFLP analysis) of 148 Estonians and 199 Albanians; b) assisted in the analysis of the data and; c) in the preparation of the manuscript;

Ref. **III** — a) participated in performing the experiments: analyzed the mtDNA variation (HVS-I sequencing and RFLP analysis) of 1622 DNA samples from 409 Estonians, 192 Latvians, 318 Swedes, 199 Albanians, 116 Hungarians and 388 Turks; b) assisted in the analysis of the data and; c) in the preparation of the manuscript;

Ref. **IV** — a) conceived and designed the experiments; b) analyzed the data; c) performed the phylogeographic analysis of mtDNA; d) wrote the paper;

Ref. **V** — a) conceived and designed the experiments; b) analyzed the mtDNA data; d) wrote the paper;

ABBREVIATIONS

AMH	anatomically modern human
bp/kbp	base pair/thousand (kilo) base pairs
CRS	Cambridge Reference Sequence
D-loop	displacement loop/control region of mtDNA
DNA	deoxyribonucleic acid
hg(s)	haplogroup(s)
HVS-I/HVS-II	the first/second hypervariable segment
LD	linkage disequilibrium
LGM	the Last Glacial Maximum
MRCA	the most recent common ancestor
mtDNA	mitochondrial DNA
np(s)	nucleotide position(s)
OXPHOS	oxidative phosphorylation
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
tRNA	transfer ribonucleic acid
YBP	years before present

Definitions of basic terms, used in current dissertation

Haplotype (= lineage)	mtDNA sequence with characteristic polymorphisms, encompasses all identical sequences;
Haplogroup	monophyletic clade of haplotypes sharing characteristic defining sequence polymorphisms;
Founder haplotype	common ancestral haplotype to which all haplotypes under concern coalesce to;
Coalescence time	time to MRCA;
Phylogeography	the study of the spatial distribution of genealogical lineages;
Star-like phylogeny	phylogeny of a set of sequences that mostly (or all) coalesce to the same haplotype

1. INTRODUCTION

For the past two decades the maternally inherited mitochondrial DNA (mtDNA) and, for a slightly shorter time, its paternally inherited counterpart — the Y chromosome — have been in the limelight of genetic studies to elucidate the demographic history of humankind. MtDNA was the first polymorphic DNA system examined in humans for evolutionary purposes. The study of maternal lineages in extant populations and efforts to also add ancient DNA into this analysis has provided a great deal of information about the evolution and dispersals of anatomically modern humans (AMHs).

The first clear evidence about the existence of genetic differences between individuals was obtained more than a century ago when Karl Landsteiner in 1901 described different blood groups of the AB0-system. More extensive studies in this area started in the 1950s–1960s when systematic analysis of the variation of proteins in world populations became possible. The transition from the analysis of protein polymorphisms to the studies of diversity of genes started in 1980s. While the main aim of the first studies was to get information about how genetic variation is associated with diseases, they also widened the understanding of how this variation may reflect demographic history of humans.

The analysis of uniparentally inherited marker systems allows population geneticists to add a truly novel dimension compared to that of classical genetic marker studies. Above scoring the differences in allele frequencies and seeking similarities and differences between populations, the value of which is sometimes questionable, one may construct phylogenetic trees of individuals that rely on well-understood genealogical relationships.

The focus of this work is the analysis of genetic diversity of maternal lineages in different Eurasian populations, as well as the environmental and cultural processes that might have been involved in the shaping of this variety. Here, under my particular interest is a population whose place in the European genetic landscape has been obscure — the Saami. Novel experimental mtDNA data offers an opportunity to use the phylogeographic approach in the study of different branches of the mtDNA tree to analyze the time depth of their expansion and, less directly, of their split from sister branches of the tree and to follow the directions of their spread. In this study, three different phylogenetic surveys, all of them related to the post-Last Glacial Maximum (LGM) recolonization of Eurasia by AMHs, are discussed.

Knowledge about the general topology of the global mtDNA tree provides the basis for investigating many interesting details of mtDNA variation in different regions. Thus, the first section of this study gives an overview about the special properties of mtDNA and its worldwide variation, with an emphasis on the European mtDNA variety. It also includes a short description of methods used in phylogenetic analyses, and an overview of current hypotheses on the origin of AMHs.

2. LITERATURE OVERVIEW

2.1. Structure and organization of mtDNA

MtDNA is an extranuclear genome of eukaryotic cells. Each mitochondrion contains usually 5–10 mtDNAs in its matrix. The number of mtDNAs in the somatic cell is about 1,000–10,000 (Lightowlers *et al.* 1997). Human mtDNA is a circular double-stranded molecule with a size of approximately 16.6 kbp (Anderson *et al.* 1981; Andrews *et al.* 1999, see fig. 1). It codes for 22 distinct transfer RNAs, two ribosomal RNAs and for 13 protein genes, which are mostly involved in the electron transport and oxidative phosphorylation (OXPHOS) pathway — the energy producing system of the cell located within the mitochondrial inner membrane. This system is made up and controlled both by the products of genes, encoded in nuclear DNA and mtDNA. The majority of proteins of OXPHOS enzyme complexes are encoded in the cell nucleus and transported to mitochondria from cytosol. Most of the genes are transcribed from guanine-rich heavy (H) strand of mtDNA, only the sixth subunit of complex I and eight tRNAs are transcribed from cytosine-rich light (L) strand (for a review, see Wallace *et al.* 1999).

The organization of the mtDNA genome is conserved in most of the metazoan organisms (Saccone *et al.* 1999), which denotes the essential role of the same minimal set of proteins for respiration and oxidative phosphorylation. Mammalian mtDNA has been built up extremely economically — it almost lacks noncoding regions and all of the coding sequences are contiguous (Anderson *et al.* 1981). The largest noncoding region is the D-loop or control region, which is a 1.1 kbp segment that contains the control elements for replication and transcription of mtDNA (Lightowlers *et al.* 1997, see fig. 1).

According to a generally accepted view, mitochondrion is a remnant of a prokaryotic organism, which became endosymbiotic with a eukaryotic cell soon after the appearance of eukaryots *ca* 1.5×10^9 years ago (Margulis 1975). Phylogenetic analyses based on mtDNA genes indicate that they are most closely related to those found in genomes of alpha-proteobacteriae (Andersson *et al.* 2003). Mitochondria have co-evolved with their "hosts" and during the evolution many of their genes have been exported into the nucleus of the cell. A wide variety of fragments, very likely deriving from the mitochondrial genome, are also found from the human genome. Some of them code for essential mitochondrial proteins and are transported, via specific mechanisms, back to the mitochondrion. There is also a wide variety of pseudogenes or just fragments of DNA of a likely mtDNA origin (Zischler *et al.* 1995; 1998; Bensasson *et al.* 2003). However, a multitude of mitochondrial proteins encoded in the nucleus display no homology to bacterial proteins, indicating that they have originated within the eukaryotic cell subsequent to the acquisition of the endosymbiont (Andersson *et al.* 2003).

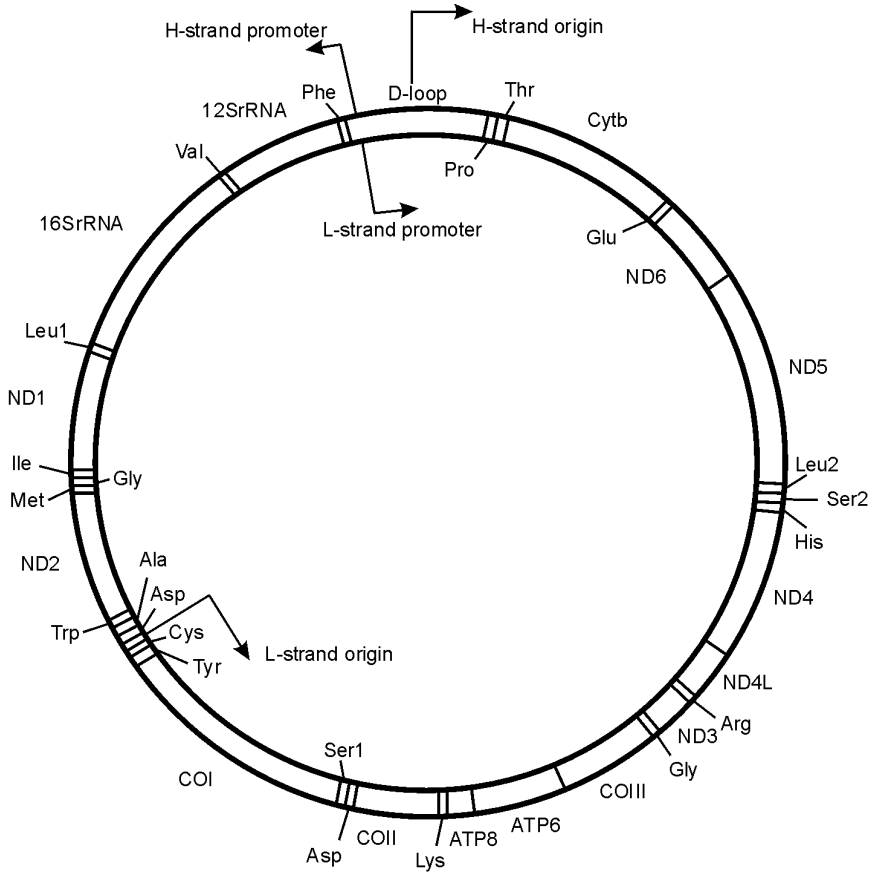

Figure 1 Human mtDNA (16.6 kb) genomic map. Location of the genes of seven subunits of OXPHOS complex I (ND1, 2, 3, 4, 4L, 5 and 6), one subunit of complex III (Cytb), three subunits of complex IV (COI, COII and COIII), two ribosomal RNAs (12SrRNA and 16SrRNA), 22 tRNAs and D-loop region are shown. Genes that are transcribed from the H-strand and L-strand are shown outside or inside the circle, respectively. Arrows indicate the location of promoters for transcription and replication origins for both strands.

2.2. Special features of mtDNA

MtDNA, compared to nuclear genes and nuclear DNA in general, has a number of special characteristics that make it a useful marker for phylogenetic studies. These are maternal inheritance, lack of recombination, homoplasmy and a relatively high mutation rate. A special aspect is the neutrality of its evolution — a problem that has been discussed extensively again only very recently.

2.2.1. Maternal inheritance and the lack of recombination. MtDNA as a single locus

Mammalian mtDNA is maternally inherited (Giles *et al.* 1980). In fertilization, the sperm mitochondria are imported into the oocyte, but selectively eliminated in early stages of embryogenesis (Manfredi *et al.* 1997; Sutovsky *et al.* 2004). Ubiquination of the mid-piece of the sperm has been proposed to act as one of the signals for destruction of paternal mtDNA in a fertilized cell (Sutovsky *et al.* 1999). The penetrance of paternal mtDNA in a developing organism is further hindered by a much larger amount of maternal mtDNA, whose copy-number in mature oocytes exceeds 100,000, while the sperm cell contains only 50–1,200 mtDNA genomes (Michaels *et al.* 1982; Diez-Sanchez *et al.* 2003). The leakage of paternal mtDNA has been reported in human polyploid as well as in some normal embryos, generated by *in vitro* fertilization techniques (St John *et al.* 2004). This might suggest that the transmission of paternal mtDNA is more frequent in the case of the poor-quality oocytes. One case of paternal leakage of mtDNA in with a patient on the background of severe metabolic disease has been reported (Schwartz and Vissing 2002) and discussed (Bromham *et al.* 2002), but there is no evidence for mtDNA paternal transfer in normal conditions.

Maternal mode of inheritance and the lack of recombination (Olivo *et al.* 1983; Merriwether *et al.* 1991; Elson *et al.* 2001; Piganeau and Eyre-Walker 2004) offer a possibility to track individual genealogies and their evolution through the genetic history of human populations. From time to time, several groups have questioned the lack of recombination in humans. Because only non-recombining loci have a single unique genealogical history, a presumed recombination, if true, would have a profound effect to the current interpretation of human mtDNA variation. In 1999, three studies challenging the view of non-recombining mode of human mtDNA heritage were published. Awadalla *et al.* (1999) tested statistically the level of pairwise linkage disequilibrium (LD) as a function of the distance between sites and found that LD declines with increasing distance, hinting to recombination. Eyre-Walker *et al.* (1999) used the phylogenetic approach for the same purpose. They analyzed the amount of homoplasies among human mtDNA coding region sequences and concluded that the frequency of parallel mutations at the same nucleotide position (np) is much higher than expected on the basis of single rate of synonymous mutations and suggested that this finding can be explained by recombination. Hagelberg *et al.* (1999) discovered a putative recombinant control region haplotype among individuals from one Pacific island population, but later the authors corrected their interpretation after finding out the sequence alignment errors (Hagelberg *et al.* 2000). However, claims about the presence of recombination in mtDNA have received strong criticism both on methodological and on data-quality grounds (*e.g.* Macaulay *et al.* 1999*a*; Kivisild and Villems 2000; Kumar *et al.* 2000). Other analyses (Ingman *et al.* 2000; Jorde and Bamshad 2000; Elson *et*

al. 2001) and reanalyses (Piganeau and Eyre-Walker 2004) have not found evidences of occurrence of recombination in mtDNA.

Due to the lack of recombination, mtDNA acts as a single locus. The effective population size of the mitochondrial genome is only one fourth that of the autosomal loci. The influence of genetic drift thus makes the mitochondrial genome more sensitive to random fluctuations of allele frequencies than that for the autosomal loci.

2.2.2. Homoplasmy

Various tissues of the same individual usually share only one type of mtDNA — this condition is referred to as homoplasmy. When a mutation arises, there may be complete switching to the new mtDNA variant within a single generation (Poulton *et al.* 1998 and references therein). If the switching is incomplete then two or more types of mtDNA can be observed in a cell. This state is called heteroplasmy.

Homoplasmy is thought to be important for the maintenance of the normal mitochondrial function through the coordinated expression of the mitochondrial and nuclear genes (*e.g.* Hirata *et al.* 2002) and is believed to be preserved by the genetic bottleneck mechanism in the oogenesis, most likely in the stage of primordial germ cells. There is no strict definition for the mtDNA bottleneck — it has been described as an event or series of events at one or several stages of oogenesis, which leads to a reduction in the number of mtDNAs, followed by a rapid increase in mature oocytes (Thorburn and Dahl 2001). Different numbers (from 1 to 200) of segregating mtDNA units have been suggested for mammals (Koehler *et al.* 1991; Jenuth *et al.* 1996; Marchington *et al.* 1998; Poulton *et al.* 1998).

2.2.3. Mutation rate in human mtDNA

On average, mtDNA accumulates mutations more than 10 times faster than does the nuclear genome (Brown *et al.* 1979; Ingman and Gyllensten 2001). This phenomenon leads to high level of within-population polymorphisms. A high mutation rate of mtDNA is associated with several properties of mtDNA, different from those of nuclear genome, and with the peculiarities of processes taking place in the mitochondrial OXPHOS pathway. Firstly, mtDNA lacks protective proteins like histons. Secondly, mtDNA is exposed to oxidative damage by reactive oxygen radicals — by-products of OXPHOS. Thirdly, it has been suggested that the reparation system of mtDNA is not as effective as that in the nucleus (Bogenhagen 1999), although many elements of the mtDNA repair system have been shown to exist in animal mitochondria (for recent review see Mason and Lightowlers 2003).

The fast evolving mtDNA provides more information about recent events in evolution than does a strech of DNA of an equal length in the nuclear genome. At the same time, the high mutation rate creates possibilities for homoplasy — the same mutation could arise in different branches of the mtDNA phylogenetic tree, thus blurring the possibility to establish an unambiguous (more precisely — the most parsimonious) order of evolutionary events.

The mutation rate is different both for mtDNA regions and nps within a region. Pesole *et al.* (1999) showed that: 1) nonsynonymous nps, D-loop central domain, tRNA and rRNA genes evolve at about a five to ten times lower rate than synonymous nps and two peripheral domains of D-loop; 2) the rate of synonymous sites is quite uniform over the genome, whereas the rate of non-synonymous sites differs considerably between genes; 3) nonsynonymous sites show comparable rates in mtDNA and nuclear genome, synonymous sites and rRNA evolve *ca* 20 times and tRNAs *ca* 100 times more rapidly in mammalian mitochondria. Additionally, transitions have been shown to occur *ca* 12 to 37 times more often than transversions (see Meyer *et al.* 1999 and references therein).

The considerable rate variation has also been observed between different D-loop nps (Hasegawa *et al.* 1993; Wakeley 1994; Macaulay *et al.* 1997; Finnilä *et al.* 2001). For example, the transitions at nps 16093, 16129, 16189, 16311 and 16362 in HVS-I and 73, 146, 150, 152, 195 in HVS-II are considered as mutational "hotspots" and are often observed in different phylogenetic branches of mtDNA. It also raises a possibility that, in "hotspots", potential recurrent mutations may be missed or overlooked in phylogenetic reconstructions.

In phylogenetic tree-building, differences in mutation rates can be normalized by assigning different weights to the nps with known rate variation (*e.g.* Richards *et al.* 1998). However, "hidden" and/or parallel mutations do not likely harm the outcome in any profound way, provided the level of resolution between the branches of an mtDNA phylogenetic tree is sufficient, i.e. the information of basal nodes of the tree is available. Then, the fast evolving positions may, theoretically, blur only the terminal tips of the tree, not distorting its basic topology. The combined usage of information both from fast evolving control region sequences and diagnostic coding region sites has justified itself in many mtDNA population genetic studies (*e.g.* Torroni *et al.* 1996; Richards *et al.* 1998; 2000; Macaulay *et al.* 1999*b*; Schurr *et al.* 1999; Kivisild *et al.* 2002).

2.2.4. The role of natural selection in the evolution of human mtDNA

The assumption of the neutrality of a genetic marker is one of the cornerstones of many methods that are used in phylogenetic analysis. The theory of neutral evolution was first introduced by Kimura (1968). According to this theory, mutations occur stochastically and their fixation is the result of random drift rather than of natural selection. Occurring deleterious mutations are removed by

purifying selection; positive selection does not play any significant role. This means, theoretically, that the rate of evolution solely depends on the mutation rate. Correspondingly, the simplest model for explaining the present mtDNA variation is the following: mutations have accumulated sequentially along radiating female lineages and have reached to polymorphic frequencies only because of random genetic drift in its various manifestations, whereas the influence of (positive) natural selection has been negligible. The demographic history of a population, or more precisely — of carriers of particular variants — has likely played a decisive role.

According to the "near-to-neutral" theory of evolution there may also be, among the mutations of recent origin in the evolutionary time scale, slightly deleterious ones that are not yet removed by purifying selection (Hasegawa *et al.* 1998; Nachman 1998; Gerber *et al.* 2001). The lack of recombination in human mtDNA makes it a subject of "Muller's rachet" — of a genetic mechanism that predicts an accumulation of slightly deleterious mutations. Many authors have tested the neutrality of mtDNA evolution, primarily by estimating the differences between the fixation of nonsynonymous and synonymous substitutions in the mtDNA of different species. They have observed that there is an excess of nonsynonymous mtDNA polymorphisms relative to fixed sequence change (*e.g.* Graven *et al.* 1995; Nachman 1998; Excoffier and Yang 1999). It has been also shown that the "older" branches of the human mtDNA tree contain relatively less nonsynonymous substitutions than the "younger" ones, which can be interpreted by inferring the action of purifying selection on the mtDNA (Moilanen *et al.* 2003; Moilanen and Majamaa 2003; Elson *et al.* 2004; Ruiz-Pesini *et al.* 2004). Recent studies of human populations have compared the ratios of nonsynonymous and synonymous substitutions in different lineages of complete mtDNA sequences from Africa, Asia and Europe (Mishmar *et al.* 2003; Elson *et al.* 2004; Ruiz-Pesini *et al.* 2004). Mishmar *et al.* (2003) and Ruiz-Pesini *et al.* (2004) interpreted the differences between the results of different haplogroups spread in tropical, temperate and arctic zones as evidence for climatic adaptations, and noted that the present continental distribution of mtDNA haplogroups might be non-random, shaped by positive selection. The analysis, however, did not consider the possibility that analogous differences might also exist between lineages that have not been selected according to their geographical origin, as it has been demonstrated (Excoffier 1990; Torroni *et al.* 2001; Moilanen and Majamaa 2003; Elson *et al.* 2004). It has to be noted that the neighbour-joining tree used by Ruiz-Pezini and co-authors was incorrect since the branching order of its clusters did not correspond to the established phylogenetic relationships between the clades of mtDNA.

Due to these special properties, mtDNA offers an opportunity to reconstruct the maternal genealogies, unhindered by the genetic fog of recombination and to detect the differences even among closely related groups that have diverged

within a relatively short time scale. Thus, the genealogy of maternal lineages provides a link between observable sequence variation and evolutionary events that have shaped this diversity. However, one should not forget that in order to draw conclusions about the demographic history of a population the data of different genetic systems (Y chromosome, autosomes) must be analysed and combined with those obtained from maternally inherited mtDNA.

2.3. Phylogenetic trees and networks

The most common way to express the phylogenetic relationships of different genes or organisms is to present those in the tree-like form. Different phylogenetic methods exist that can be used for reconstructing phylogenetic trees from molecular data (*e.g.* Saitou and Nei 1987; Fitch 1977; Felsenstein 1981).

While useful for reconstructing phylogenies from interspecific data, the traditional tree-building methods are often unsatisfactory when applied to human mtDNA data. The reasons for this are short genetic distances between individuals; large sample sizes; homoplasy or parallel mutation events; and reversals of character changes. The resulting mass of equally plausible trees can be best presented by a network, which expresses the alternative evolutionary pathways in a form of unresolved reticulations (fig. 2). The network is generated under the assumption that the evolutionary process has proceeded by employing the smallest number of character changes to create the present sequence variability. For smaller datasets (for example, less than 100 HVS-I sequences) the reduced median network (Bandelt *et al.* 1995), which consists of almost all possible maximum parsimony trees, can be used. Obvious recurrent characters in a network, presented first as reticulations, are reconstructed by splitting characters into new characters that account for the hypothetical multiple hits. Each reduction step employs the parsimony and frequency (of sampled sequence types) criteria as well as the knowledge of different mutation rates of nps. For larger datasets of several hundred sequences, the median joining algorithm (Bandelt *et al.* 1999), where inner branches between shortly connected nodes are sequentially introduced, is most commonly used.

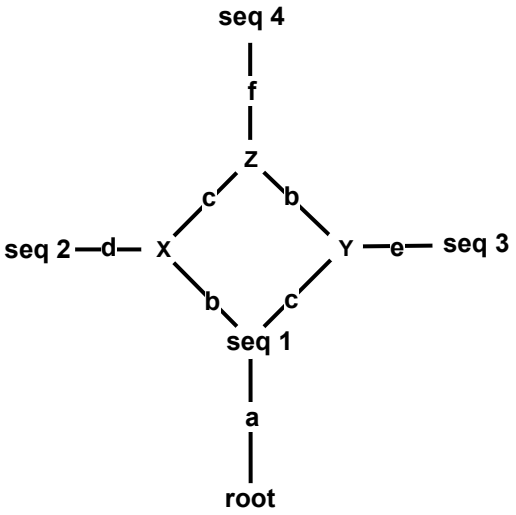

Figure 2 Phylogenetic network of four sequences (seq 1 — seq 4). Alternative possibilities for tree-construction are shown as a reticulation. a, b, c, d, e and f are different characters; X, Y, Z are median vectors.

2.4. Rooting of the phylogenetic tree

To show the temporal stratification of branching events, a phylogenetic tree must have a root. For that, the outgroup that already earlier in evolution has separated from the most recent common ancestor (MRCA) of the studied group will be chosen. In human mtDNA analysis, the corresponding sequences of chimpanzee (Vigilant *et al.* 1991; Ingman *et al.* 2000; Maca-Meyer *et al.* 2001), Neanderthal (Krings *et al.* 1997; Ovchinnikov *et al.* 2000), as well as of the recent insertion of D-loop segment into the nuclear genome (Watson *et al.* 1997), have been used as an outgroup. If the outgroup cannot be determined, the midpoint rooting will be employed. The midpoint root is specified so that the distance from that to all terminal nodes of the tree would be minimal. Midpoint rooting was used in the first studies of human mtDNA (Cann *et al.* 1987), when the data of suitable outgroups was lacking.

2.5. Calibration of the mtDNA molecular clock

The accurate estimation of substitution rates and divergence times is one of the central questions to be answered when discussing the origin and demographic history of modern humans. The main assumptions while linking together the observable sequence variation and the time that has passed to produce it have been as follows: 1) the mutation rate is constant in different lineages; and 2) the particular loci that have been used for calculations are selectively neutral. Important information for calibration of the molecular clock comes from historical records and fossils.

One of the approaches is to estimate the mean rate of mtDNA divergence in geographical regions for which the more or less exact time of colonization is available from historical sources. The extent of differentiation within clusters specific to New Guinea, Australia and the Americas has been used and the divergence rate (twice the substitution rate) estimate between 2–4% per million years has been calculated for whole human mtDNA molecule (Wilson *et al.* 1985; Cann *et al.* 1987; Torroni *et al.* 1994*a*). The same value for transitions of D-loop HVS-I region (between nps 16090 to 16365) was found to be 36% per million years (Forster *et al.* 1996).

The other possibility is to use an outgroup method, which compares the average amount of sequence variation between two species considering their distance from the MRCA. The time of their split is taken from paleontological evidences. For human and chimpanzee, assuming that orangutan and African apes diverged *ca* 13 million YBP (Andrews 1992), the split has been inferred to occur around 5 million years ago (Horai 1995). The substitution rate at synonymous sites and in the control region has been calculated as 3.9×10^{-8} and 7×10^{-8} per site/per year, respectively (Horai 1995) and in coding region as 1.7×10^{-8} per site/per year by Ingman *et al.* (2000). Close estimates were obtained also by Mishmar *et al.* (2003), who found substitution rate 1.3×10^{-8} for the coding region part of mtDNA assuming 7 millions years for the split between humans and chimpanzees. In the same way, the divergence rate of both hyper-variable segments of D-loop was calculated to be *ca* 15% per million years by Vigilant *et al.* (1991) and 23% by Stoneking *et al.* (1992); for HVS-I alone the estimate was 33% per million years (Ward *et al.* 1991). These estimates suggest that the MRCA of the human mtDNA phylogeny occurred around 200,000 years ago.

One approach that does not require additional information from fossils or historical records considers the pedigree data for calibration and estimates the mutation rate directly from the samples with known genealogy. In those studies (*e.g.* Howell *et al.* 1996; Parsons *et al.* 1997; Howell *et al.* 2003) the obtained mutation rate estimates for control region have been approximately ten times higher than those calculated with other methods. This phenomenon pointed to a possibility that the conventional phylogenetic rates might be underestimated and

together with that also the dates of past divergences might be wrong (Pääbo 1996). However, it is also possible that large proportions of mutations that can be observed in pedigrees are in fact mildly deleterious and will not be fixed — i.e. they will be removed from the mtDNA gene pool. Although the improved knowledge of the differences between the mutation and fixation rates and differences in rates between sites in control region suggest that there is no need to thoroughly revise the phylogenetic estimates of mutation rates (Macaulay *et al.* 1997), the possible role of selection, undetected homoplasy *etc*. must be kept in mind.

2.6. Early studies of phylogeographic diversity of human mtDNA

The first application of mtDNA data to elucidate the origin of modern humans took place in the late 1970s and early 1980s, when Brown *et al.* (1980) discovered that restriction fragment pattern of mtDNAs among individuals from diverse geographic and ethnic origin differed substantially. The time of coalescence to the global mtDNA variation was estimated to be 180–360,000 YBP. The complete sequence of mitochondrial genome was published in 1981 (Anderson *et al.* 1981) and soon after that many studies of different populations were undertaken. While Brown *et al.* (1980) had used many restriction enzymes (18) in few samples (21), Denaro and colleagues (1981) treated a large number of samples (235) with only a single enzyme. They found that the presence of *Hpa*I restriction site that corresponds to the transition at np 3954 separates most of the Africans from Eurasian individuals. Notably, the Africans had a derived state of this np when compared to other primates, whereas the rest of studied populations shared an ancestral state. Based on this, Asia as a possible starting point for mtDNA radiation was proposed. This study was supported by the investigation of Blanc *et al.* (1983) who discovered the frequent presence of ancestral state of one *Hinc*II polymorphism (which corresponds to transition in np 12,406) among Asians and was upheld also by studies of Nepalese (Brega *et al.* 1986) and Chinese (Yu *et al.* 1988) populations. The mtDNA variation was shown to be high also among the Japanese population (Horai *et al.* 1984). Cann *et al.* (1982) demonstrated that mtDNA diversity among aboriginal Australians is as diverse as in any other population tested in the Old World. Johnson *et al.* (1983) showed that the mtDNA diversity is greatest in Africa, and that all mtDNA variants present today can be seen as deriving from a single phylogenetic tree. Although they argued that the highest diversity in Africa can be caused by the longer age of African variants as well as by different mutation rates in different mtDNA lineages, the mid-point root of their mtDNA tree indicated an African origin. Excoffier and Langaney (1989) suggested that the

high diversity of Africans could be explained equally well by assuming that selection has played a significant role in creating their present mtDNA variation. Relethford and Jorde (1999) added that the high genetic diversity in Africa could be the consequence of a larger effective population size there.

The possibility to trace the origin and to study the demographic history of humankind by the use of mtDNA as a population genetic marker became highly visible in 1987 with the well-known *Nature* paper of Allan Wilson's group (Cann *et al.* 1987). They used high-resolution restriction analysis of 147 mtDNAs from a worldwide sample and concluded that the root of human mtDNA phylogenetic tree is in Africa. Compared to earlier studies (*e.g.* Johnson *et al.* 1983) that allowed the screening of 2–4% of the total mtDNA sequence variation, in this analysis a ten times higher resolution level was obtained and the clustering of mtDNA types favoured again African origin of the mtDNA tree. This particular paper, published more than 15 years ago, was instrumental in coining the term "The African Eve" (Ayala 1995).

2.6.1. Multiregional model of human evolution

The results of early mtDNA studies led to the construction of a starlike mtDNA tree with a central node shared by a high number of individuals from all over the world (Excoffier and Langaney 1989). Other lineages, some of which where population specific, radiated from that central haplotype. The root of this tree was disputed and it was sometimes interpreted as support for the multiregional model of the origin of modern humans (Templeton 1992). This model has a long history and is based largely on archaeological evidence and the interpretation of morphological markers of fossil findings. The multiregional model was first proposed by Franz Weidenreich (1943), but was further developed and promoted by others, most visibly by Milford Wolpoff and colleagues (Wolpoff *et al.* 1984; Wolpoff and Caspari 1997). According to the theory, all modern humans evolved from their common archaic ancestor, *Homo erectus*, who left Africa probably as early as *ca* 1.8 million years ago. Each continental group of populations developed to modern humans, or *Homo sapiens*, already *in situ*, participating in a permanently ongoing gene flow.

2.6.2. Recent African origin of "Mitochondrial Eve" and of anatomically modern humans

An opposite model to "multiregionalism" is the "recent out-of-Africa" model, known also as the *Garden of Eden* model, according to which AMHs evolved only in Africa and replaced all other preexisting archaic hominids. This theory was proposed in its extreme form — assuming total replacement of all archaic ancestors — by archaeologists Stringer and Andrews (1988) at about the same

time when the genetic evidences of recent African origin of human mtDNA appeared (Cann *et al.* 1987).

The most important results of the study of Wilson's group were: 1) the midpoint-rooted maximum parsimony tree obtained from the results of the high-resolution RFLP analysis was not starlike — it revealed a deep split between the two main branches, one of which consisted only of African lineages and the other encompassed both Africans and non-Africans; 2) the mtDNA variation was highest in Africa. These findings were interpreted as evidence of the African origin of mtDNA in extant populations — all mtDNAs stem from one woman who lived ca 200,000 years ago. Cann *et al.* (1987) was criticized mainly on five grounds: 1) for using the indirect RFLP-method of comparing DNAs instead of sequencing; 2) for poor sampling, where native African mtDNAs were represented by a small number of African Americans; 3) for using less powerful midpoint-rooting of the obtained phylogenetic tree instead of the outgroup-method; 4) that the authors did not provide statistical support for inferring an African origin of human mtDNA variation and; 5) that they inadequately calibrated the rate of human mtDNA evolution (Darlu and Tassy 1987; Saitou and Omoto 1987; Excoffier and Langaney 1989; Kruger and Vogel 1989; Maddison 1991).

Keeping in mind the criticism of their first analysis, Wilson's group (Vigilant *et al.* 1991) analyzed 189 sequences of two hypervariable segments of mtDNA control region, including those from 121 native Africans. The branches of the maximum parsimony tree, rooted by the more powerful outgroup–method, which used chimpanzee sequence, started again exclusively from Africa. The calibration of this tree with the mutation rate calculated from comparisons of average amount of sequence differences between human and using an outgroup, resulted in approximately the same age of "mitochondrial Eve" that was obtained in the previous publication (Cann *et al.* 1987). Again, criticisms of those studies stayed unmoved because equally parsimonious trees from the same dataset showing other results were obtained (Hedges *et al.* 1992; Templeton 1992). The conclusion that could be drawn from those early papers was that the available sequence data and the genealogical resolution level were insufficient to solve statistically the place of origin of human mtDNA.

As for today, the recent out-of-Africa model has found preponderant support from the mtDNA analysis of large datasets analyzed with different methods (*e.g.* Ruvolo *et al.* 1993; Chen *et al.* 1995; 2000; Horai *et al.* 1995; Penny *et al.* 1995; Watson *et al.* 1997; Ingman *et al.* 2000) as well as from the studies of Y-chromosome (*e.g.* Hammer 1995; Underhill *et al.* 1997; 2000), autosomal loci (*e.g.* Armour *et al.* 1996; Nei and Takezaki 1996; Tishkoff *et al.* 1996; 2000; Jin *et al.* 1999; Kaessmann *et al.* 1999; Jorde *et al.* 2000) and from paleo-anthropological findings (Stringer and Andrews 1988; Foley 1998; Stringer 2000; 2003). However, this model is certainly over-simplified, as it does not account for continental subdivisions, environmental changes affecting the

demographic history of populations, potential sex-linked differences in generation time *etc.*, and lacks a mechanism to explain the present diversity of AMH (for a review, see Excoffier 2002). Phylogenetic studies on mtDNA in Europeans (Torroni *et al.* 1996; Macaulay *et al.* 1999*b*), Asians (Torroni *et al.* 1993*a*; 1994*b*; Schurr *et al.* 1999; Kivisild *et al.* 2002; Yao *et al.* 2002*a*; 2002*b*; Kong *et al.* 2003*a*; 2003*b*), Papuans (Stoneking *et al.* 1990; Forster *et al.* 2001) and Native Americans (Torroni *et al.* 1993*b*; 1994*a*; 1994*c*) have confirmed that each continent has distinct sets of phylogenetically deep mtDNA branches that speak in favour of the so-called *weak Garden of Eden* model (Harpending *et al.* 1993). According to this model, present genetic variation has not been the result of uninterrupted demographic expansion of the out-of-Africa founders but rather by the different maturation and range expansion phases of regional gene pools that followed the initial expansion (Forster *et al.* 2001; 2004). Various alternative modes and pathways of dispersal of AMHs, following the northern (over Sinai) and/or southern (over southern Arabia) routes have been put forward, based on paleoanthropological (Lahr and Foley 1994; 1998; Stringer 2003) and genetic data (Cavalli-Sforza *et al.* 1994; Hammer *et al.* 1997; 1998; Jin *et al.* 1999; Kivisild *et al.* 1999*a*; 2000; 2003*a*; Quintana-Murci *et al.* 1999; Underhill *et al.* 2000; Cann 2001; Templeton 2002). Thus, modern humans originated from a recent single evolutionary event, whereas modern human diversity is the result of multiple evolutionary events brought about by multiple geographic dispersals.

2.7. Global mtDNA variation. Nomenclature of mtDNA haplogroups

The first studies of mtDNA variation in human populations used either the RFLP analysis or sequencing of the hypervariable segments (usually HVS-I) of the control region. A more refined picture about the spread of different mtDNA variants started to emerge during early 1990s with the application of the high-resolution restriction fragment analysis to study the mtDNA variation from one continent at a time (*e.g.* Ballinger *et al.* 1992; Torroni *et al.* 1992; 1993*a*; 1993*b*; 1994*a*; 1994*b*; 1994*c*; Chen *et al.* 1995). This series of investigations established "the backbone" of the human mtDNA phylogenetic tree that has largely been in use since then. Meanwhile, an independent classification was proposed, based on the phylogenetic analysis of the variation of HVS-I part of mtDNA (Richards *et al.* 1996). Thereafter, a new series of studies allowed the combination of the control region and coding region (RFLP) data (Torroni *et al.* 1996; Richards *et al.* 1998; Macaulay *et al.* 1999*b*; Schurr *et al.* 1999).

or mtDNA haplogroups in different world populations

	H	Pre V + V	(Pre-HV)1 + HV	U	U*	U1	U2	U3	U4	U5	U6	U7	U8	K	T	J	R*	R9	B	N*	W	N1	X	A	N9	M*	Mi[b]	C	D	G	Z	L0	L1	L2	L3
5	4	42		50						50																			3		1				
3	40	5		31						27			1	3	2	4					9	3	1								2				
4	38	5		23	1			1		16			1	4	12	11					3	5	1										1		
1	42	5	2	21		1	1	1	3	11				4	11	8					2	2	3												
4	48	6	2	24		1		1	6	12				4	5	7					1	3	1			1									
0	48	4	2	23	2		2		1	8	1		1	8	11	6					2	3	1												1
7	25		6	26	1	4	1	5	1	5		2	1	6	11	11	2			1	4	5	5	1	1	1		1	1	1	1				
3	15			20			3		14	3					4				1					4	2			30	14	5	5				
1	3		1	2					1	1					1	1		6	1		2			2	2	3		44	30	4					
0	6			16			5	2	5	4					1	4		9	4			5	3		5	7		19	15	2	5				
3																	1	16	16					7	6	23		3	22	3	2				
2				4			4										5	45	13					3	1	7			17	5					
5	3		2	15		1	10		1	1		2			2	1	15	3			1					40	19	1							
0																			25				2	33				20	20						
3	18	2		16				1		4	8			3	4	3										1						13	25	17
6																															25	12	36	27
2				5						3	2															1					5	16	43	30
3																															4	70	8	18

cified derivatives of a particular haplogroup. Frequencies of hg U sub-hgs are given separately. [a]Data are from: [1]Sajantila 96; Delghandi *et al.* 1998; Tambets *et al.* 2004; [2]Meinilä *et al.* 2001; [3]Passarino *et al.* 2002; [4]Malyarchuk *et al.* 2002; 2001; Dubut *et al.* 2004; [7]Tambets *et al.* 2000; [8]Saillard *et al.* 2000*b*; Derbeneva *et al.* 2002*a*; 2002*b*, our unpublished ko *et al.* 2003; [11]Yao *et al.* 2002*b*; [12] Fucharoen *et al.* 2001; Oota *et al.* 2001; Yao *et al.* 2002*a*; [13]Bamshad *et al.* 2001, . 1993*b*; [15]Rando *et al.* 1998; [16]Pereira *et al.* 2001; Salas *et al.* 2002; [17]Rosa *et al.* in press; [18]Vigilant *et al.* 1991; Chen *et* up M sub-clades spread mostly only in India.

As an outcome, it was established that all mtDNA variants, or *haplotypes*, with characteristic polymorphisms can be divided into a number of monophyletic clades, or *haplogroups* (hgs), defined by single or several mtDNA coding region polymorphisms (often defined by a gain or a loss of a restriction site), associated with polymorphisms (if any) of the mtDNA control region. The main haplogroups are denoted by capital letters and consist of different sub-haplogroups defined by characteristic polymorphisms (see figures 2–4). Most of the haplogroups exhibit restricted geographical spread (table 1), thus distinguishing populations from different continents/geographical regions. The refinement of the mtDNA tree and the nomenclature of mtDNA is an ongoing process. In particular, data emerging from complete mtDNA sequences (Ingman *et al.* 2000; Finnilä *et al.* 2001; Finnilä and Majamaa 2001; Maca-Meyer *et al.* 2001; Herrnstadt *et al.* 2002; Kivisild *et al.* 2002; Yao *et al.* 2002*a*; Kong *et al.* 2003*b*; Reidla *et al.* 2003) offer better phylogenetic resolution of different variants of maternal lineages. In the next chapters the global mtDNA tree will be discussed in greater detail.

2.7.1. The skeleton of the human mtDNA phylogenetic tree

The general scheme of the basal branches of the global mtDNA tree in humans is presented in figure 3. The likely root of the human mtDNA tree is between hgs L0 and L1, dividing the phylogenetic tree into two basic clades: L0 and all the rest. An approximate place of the root has been suggested using different approaches: comparisons with chimpanzee mtDNA complete sequences (Ingman *et al.* 2000; Maca-Meyer *et al.* 2001) as well as by use of Neanderthal mtDNA control region sequences (Krings *et al.* 1997; 1999; Ovchinnikov *et al.* 2000).

The oldest lineage cluster L was initially defined by the presence of the *Hpa*I restriction site at np 3592 (Chen *et al.* 1995). This study, however, was based on a phylogenetic analysis where Asian and European mtDNA sequences were used as the outgroups for the African tree. Therefore, L is not a "real" clade and includes several distinctive paraphyletic clusters of African mtDNA lineages. Using chimpanzee mtDNA sequence as an outgroup for human mtDNA, four major African clades within L can now be recognized — L0, L1, L2 and L3 (Salas *et al.* 2004). Each of these clades can be further divided into several subclusters (fig. 3). Different estimates of the age of MRCA for African lineages have given very similar results with coalescence values of 100,000–170,000 YBP (Chen *et al.* 1995; Graven *et al.* 1995; Horai *et al.* 1995; Watson *et al.* 1997; Chen *et al.* 2000). The MRCA of the oldest clade containing both African and all non-African individuals lies in hg L3 and dates to *ca* 50,000–80,000 YBP (Watson *et al.* 1997; Ingman *et al.* 2000).

Two branches of L3 – M and N (fig. 3) — cover mtDNA variation outside Africa. Most western Eurasians are characterized by clades within hg N

(Torroni *et al.* 1996; Richards *et al.* 1998; Macaulay *et al.* 1999*b*), whereas M and N contributed both to the current eastern Eurasian mtDNA pool (*e.g.* Torroni *et al.* 1993*a*; Kivisild *et al.* 2002; Yao *et al.* 2002*a*; Kong *et al.* 2003*a*; Comas *et al.* 2004). Because of its great time depth and virtual absence in western Eurasians, it is not excluded that hg M was brought to Asia from East Africa, along the southern route, by the earliest migration wave of AMHs (Kivisild *et al.* 1999*a*; 2000; Quintana-Murci *et al.* 1999). Recently, the same scenario was proposed also for the spread of hg N (Kivisild *et al.* 2003, but see also Kivisild *et al.* 1999*a*). The lack of L3 lineages other than M and N among non-Africans (Ingman *et al.* 2000; Herrnstadt *et al.* 2002; Kivisild *et al.* 2002) and, particularly, in South Asia, suggests that the earliest migration(s) of modern humans already carried these two mtDNA ancestors (Kivisild *et al.* 2003*b*). This scenario is also consistent with the fact that the founder ages of M

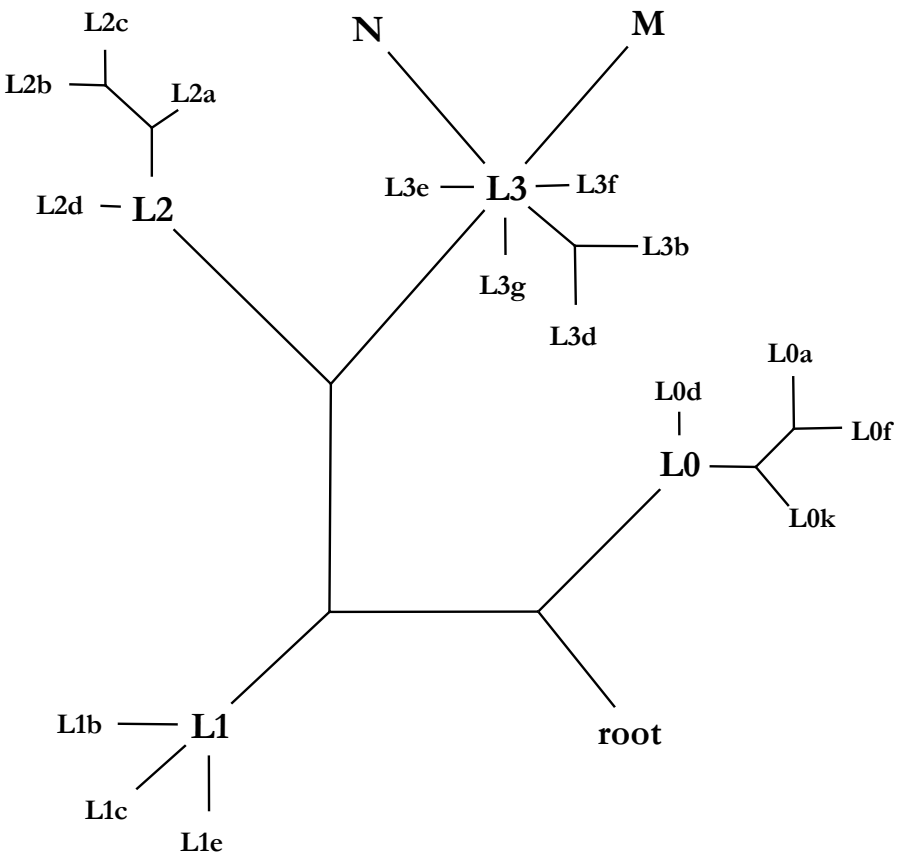

Figure 3 Scheme of global human mtDNA tree (based on Chen *et al.* 1995; Watson *et al.* 1997; Chen *et al.* 2000; Torroni *et al.* 2001; Herrnstadt *et al.* 2002; Salas *et al.* 2002; 2004). Tree is rooted with chimpanzee mtDNA sequence. For further details, see figures 4, 5.

and N have been shown to be very similar — 54,200 ± 11,400 and 53,400 ± 11,700 YBP, respectively (Forster *et al.* 2001). Also, it has been proposed by Kivisild *et al.* (2003*b*) that, considering the mtDNA variation in South Asia, the N branch had relatively early given rise to its large daughter clade R, which later, among eastern Eurasians, differentiated into clusters B and R9 (see fig. 4) and gave rise to hgs HV, TJ, and U (see fig. 5) among western Eurasians.

2.7.2. MtDNA variation in Asia

A simplified scheme of the topology of the branches of the mtDNA tree in eastern Eurasia is given in figure 4. The phylogeographic patterns of hgs M and N are very different, signifying their distinct expansions into Eurasia, which yielded the geographical structuring of external branches of these two haplogroups in western and eastern Eurasia.

All hg M subclades show marked frequency variations throughout Asia. In Europe, apart from a few exceptions in the easternmost regions (Bermisheva *et al.* 2002), lineages of this clade have been found only occasionally (table 1). Hg N includes a large lineage cluster R (see fig. 4 and 5), which is widely spread both in eastern and western parts of Eurasia. Almost all eastern Eurasian R lineages belong to the two major hgs B and R9 (fig. 4). Sub-clades of R that encompass the majority of mtDNA variants spread in western Eurasia (HV, TJ, U, see fig. 5) are rare or absent in most eastern Eurasian populations. Eastern and western Eurasian-specific mtDNA packages meet in Central Asia, which is a contact zone between those regionally differentiated groups. There, the contribution of eastern and western Eurasian mtDNAs to the total mtDNA pool is more or less equal (Comas *et al.* 2004; Quintana-Murci *et al.* 2004, see also table 1). The contribution of western Eurasian components (U4, H, JT) account for one third also in some western and southern Siberian populations (Derbeneva *et al.* 2002*a*; 2002*b*; Derenko *et al.* 2003). In northern Siberia, among Mansis, a novel branch, a sister group to hg W (see fig. 5), was recently characterized (Derbeneva *et al.* 2002). In Central Asia, less than 5% of the haplotypes belong to South Asian-specific sub-branches of hg U2 and hg M (Comas *et al.* 2004). These hgs have probably arisen through *in situ* diversification in early Upper Palaeolithic (Kivisild *et al.* 1999; Bamshad *et al.* 2001).

The derived lineage groups of hg M — hgs D (D4 and D5), G (G1 and G2), C, Z —, and those of hg N — hgs A and Y (fig. 4) form the majority of the mtDNA pool in northern and northeastern Asia (Torroni *et al.* 1993*a*; Starikovskaya *et al.* 1998; Schurr *et al.* 1999; Kivisild *et al.* 2002; Derenko *et al.* 2003; Fedorova *et al.* 2003; Puzyrev *et al.* 2003) and are common in East Asia (Kivisild *et al.* 2002; Yao *et al.* 2002*a*; 2002*b*). These hgs are predominant eastern Eurasian-specific hgs also in the Central Asian mtDNA gene pool

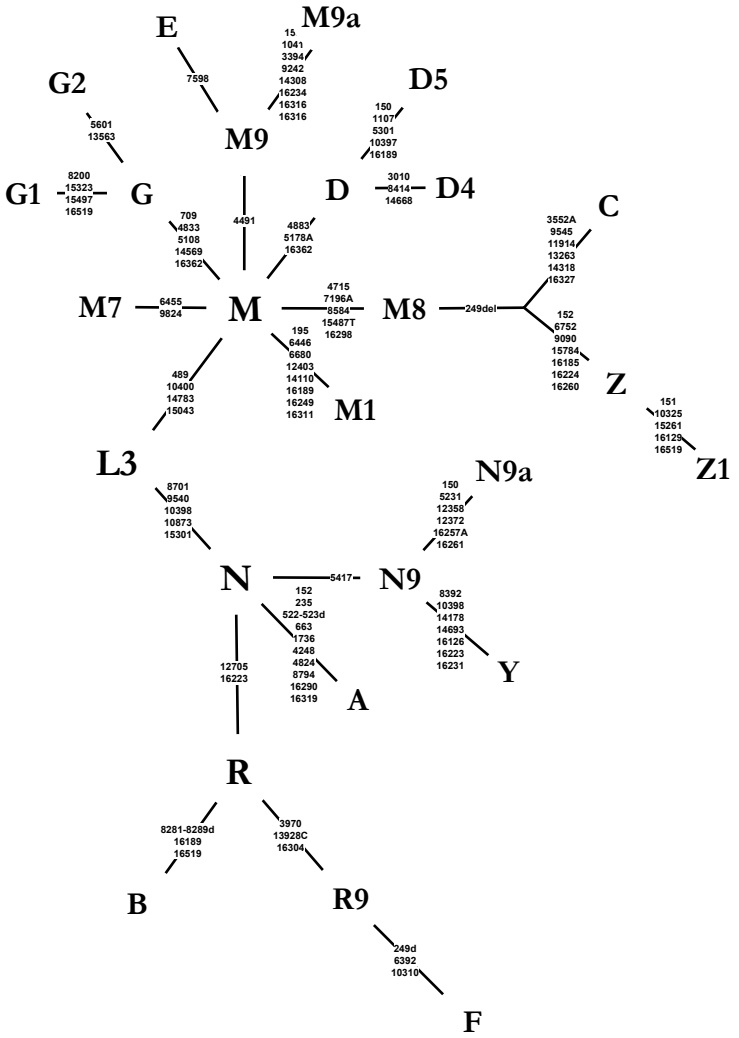

Figure 4 Schematic tree of human mtDNA haplogroups spread in eastern Eurasian populations (based on Quintana-Murci *et al.* 1999; Kivisild *et al.* 2002; Yao *et al.* 2002*b*; Kong *et al.* 2003*b*). Haplogroup-defining polymorphisms, relative to CRS, are shown on the links. Only transversion are further specified.

(Kivisild *et al.* 2002; Comas *et al.* 2004; Quintana-Murci *et al.* 2004). The spread of hgs C, D and G shows a decreasing gradient towards South and Southeast Asia. These hgs are rare in South Asia (Kivisild *et al.* 1999*a*; 2003*b*; Bamshad *et al.* 2001) and in most Southeast Asian populations (groups L, R and K, respectively, in Ballinger *et al.* 1992; Schurr and Wallace 2002). However, hg D, which is common in East Asia and Siberia, is rather unevenly spread in

Southeast Asia — it is rare or absent among Vietnamese, Malay and Sabah Aborigines (group L in Ballinger *et al.* 1992), but found at moderate frequencies (17%), for example, in Thailand (Fucharoen *et al.* 2001; Oota *et al.* 2001; Yao *et al.* 2002). Some sub-branches of hgs D4 and G2 might have had their earliest diversification in Central Asia (Comas *et al.* 2003). Hgs A and Y have phylogeographic patterns, similar to hgs C and D — these hgs are more frequent in northern Asian populations (Starikovskaya *et al.* 1998; Schurr *et al.* 1999; Saillard *et al.* 2000*a*) and are, likewise, virtually absent in South Asia (Kivisild *et al.* 1999*a*; 2003*b*; Bamshad *et al.* 2001) and rare in Southeast Asia (Fucharoen *et al.* 2001; Oota *et al.* 2001; Schurr and Wallace 2002; Yao *et al.* 2002*a*).

Hg Z, the sister clade of C, stemming from the common node M8 (Kivisild *et al.* 2002), which was first described in north Siberians (Schurr *et al.* 1999), is, interestingly, present also in North European populations, among Finns (Meinilä *et al.* 2001) and Saami (Sajantila and Pääbo 1995; Dupuy and Olaisen 1996; Lahermo *et al.* 1996; Delghandi *et al.* 1998), where the hg Z1 haplotypes spread in northern Eurasian populations (with HVS-I motif 16129-16185-16223-16224-16260-16298) can be found. In those populations hg D5b haplotypes with HVS-I transitions at nps 16126-16136-16189-16223-16390, otherwise found from southern Siberian populations (Derenko *et al.* 2003), can also be observed (Meinilä *et al.* 2001).

A large and rather superficially characterized variety of sub-branches of hg M including M2 – M6 (not shown in fig. 4), are spread at high frequencies only in South Asia (Kivisild *et al.* 1999*a*; 2003*b*; Bamshad *et al.* 2001). Hgs M7, M8 (other than C and Z), M9 and N9a are common in East and Southeast Asian populations (Fucharoen *et al.* 2001; Oota *et al.* 2001; Kivisild *et al.* 2002; Yao *et al.* 2002*a*; 2002*b*; Kong *et al.* 2003*a*; 2003*b*), and some of those (M7 and M9) have also been found occasionally from some Siberian populations (Derbeneva *et al.* 2002*b*; Derenko *et al.* 2003).

The Southeast Asian mtDNA pool consists mostly of hg B and hg R9, the latter includes also hg F (fig. 4). These haplogroups might have had their earliest diversification in Southern China and/or Southeast Asia (Yao *et al.* 2002*b*). In some populations, like in Polynesians, hg B is almost the only variant of mtDNA (haplotypes with 9bp deletion in Ballinger *et al.* 1992; Redd *et al.* 1995; Sykes *et al.* 1995; Lum and Cann 1998). In most of the Siberian populations its frequency is low (Derenko *et al.* 1999; 2003; Fedorova *et al.* 2003).

2.7.3. MtDNA variation in Europe

The analysis of classical genetic markers has shown that Europe as a whole is quite homogeneous — the genetic distances between different populations are relatively short and the genetic landscape, if to compare it with other Old World continents, is rather uniform (Di Rienzo and Wilson 1991; Piazza 1993; Cavalli-Sforza *et al.* 1994). Only some clear outliers, like the Saami, Sardinians,

Basques and some others, have been shown to emerge from this homogenuous entity (Cavalli-Sforza *et al.* 1994).

More than 90% of maternal lineages present in European populations can be classified into 8 major haplogroups (H, V, T, J, N1, U, X, W; see fig. 5), characteristic to western Eurasians in general (*e.g.* Torroni *et al.* 1994*d*; 1996; Richards *et al.* 1998; 2000; Macaulay *et al.* 1999*b*, see also table 1). All of them coalesce to the common node N in mtDNA phylogeny (fig. 5). As it was already indicated above, only a minor part of maternal lineages that have been found in Europe belong to East or South Asian-specific haplogroups. Yet it is worthwhile to note here that the uniformity of the genetic landcape of the distribution of mtDNA haplogroups in Europe is a term that should not be used in its absolute meaning. It has already been shown (Richards *et al.* 2002; Richards 2003) that a deeper phylogenetic analysis allows the revealing of significant differences in the spread of mtDNA sub-haplogroups in Europe — and not only among genetic outliers like Saami.

The phylogenetic classification of European mtDNA relies mostly on the combined usage of diagnostic coding region polymorphisms and sequence information of both hypervariable segments of mtDNA. Further refinement of the present nomenclature, due to higher phylogenetic resolution, can be achieved by the complete sequencing of mtDNA genomes (*e.g.* Finnilä *et al.* 2001; Herrnstadt *et al.* 2002).

The majority of European-specific haplogroups (HV, TJ and U) stem from a large nested lineage cluster R of the macrohaplogroup N (for defining polymorphisms of western Eurasian hgs, see fig. 5).

Hg H is by far the most frequent maternal lineage cluster in Europe. Its frequency is the highest (40–60%) in western and northern European populations, but it is also common in the populations of the Near East (Torroni *et al.* 1998; Richards *et al.* 2000) and is well visible in North Africans (Corte-Real *et al.* 1996; Rando *et al.* 1998; Stevanovitch *et al.* 2004) and Central Asians (Metspalu *et al.* 1999; Comas *et al.* 2004; Quintana-Murci *et al.* 2004), less so in South Asians (Passarino *et al.* 1996; Kivisild *et al.* 1999*a*; 1999*b*; 2003*b*; Bamshad *et al.* 2001) and in Native Siberians (Torroni *et al.* 1998; Derbeneva *et al.* 2002*a*; 2002*b*; Derenko *et al.* 2003; Fedorova *et al.* 2003). Phylogenetically closely related clades to hg H — (preHV)1, HV1 and HV2 (fig. 5) — are present predominantly in the Near East, Middle East and in the Caucasus (Metspalu *et al.* 1999; Richards *et al.* 2002; 2003). It is believed that the initial expansion of hg H took place most probably in the Near East about 25,000 YBP (Richards *et al.* 2000). Hg H includes the Cambridge Reference Sequence (CRS, Anderson *et al.* 1981), which is one of the most frequent HVS-I haplotypes found in Europe and occupies the central node in the hg H topology, provided additional coding region mutations are not considered. There have been several attempts to refine the inner structure of H (*e.g.* Helgason *et al.* 2001; Malyarchuk and Derenko 2001) to be able to trace the phylogeographical spread of different branches of this major haplogroup. Based on mtDNA

complete sequences, seven sub-branches of hg H — H1 and H2 (Finnilä *et al.* 2001) along with H3 and H4 (Herrnstadt *et al.* 2002) and H5-H7 (Quintans *et al.* 2004) have been defined, but a broad phylogeographic analysis on hg H sub-clades is still needed to understand its spread pattern in western Eurasia.

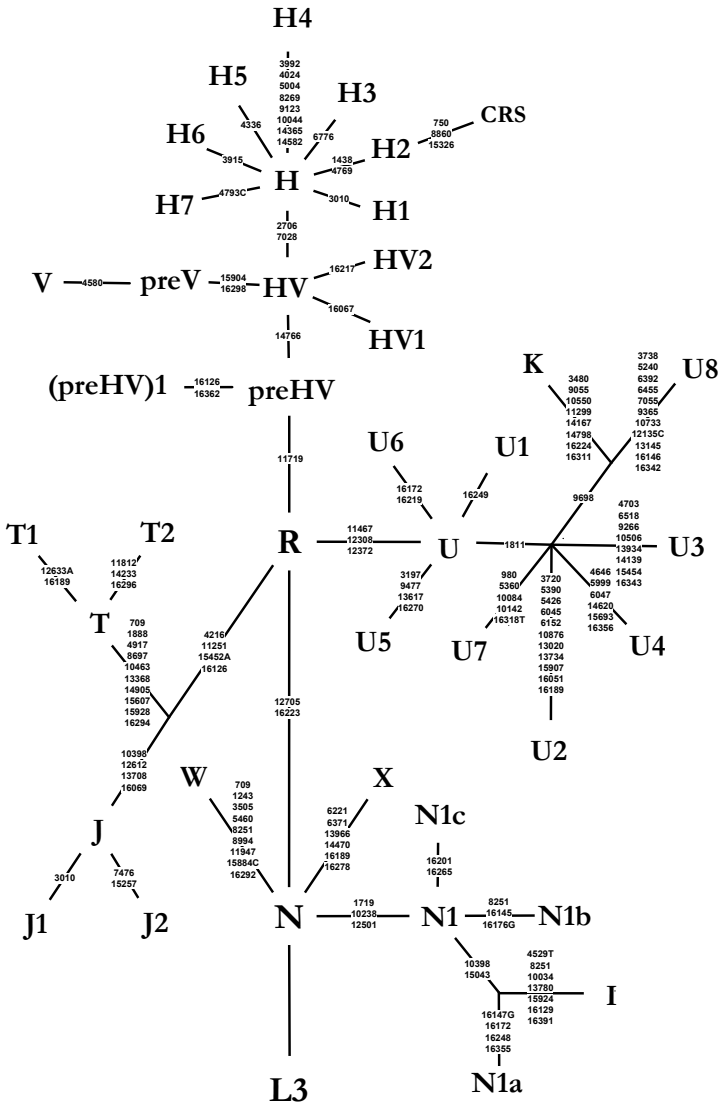

Figure 5 Schematic tree of human mtDNA haplogroups spread in western Eurasian populations (based on Macaulay *et al.* 1999; Richards *et al.* 2000; Finnilä *et al.* 2001; Herrnstadt *et al.* 2002; Quintana-Murci *et al.* 2004; Quintans *et al.* 2004; Reidla *et al.* 2004). Only defining HVS-I and coding region nps have been used. See also the legend for figure 4.

Hg V is the sister-clade of hg H (Torroni *et al.* 1998). The frequency of hg V does not exceed the level of 1–6% in the mtDNA gene pool of most European populations. As exceptions, the Basques and Catalans in the Iberian Peninsula can be named, among whom the frequency of hg V is more than 20%. A particular exception are also the Saami in the Scandinavian Peninsula. Among the latter, the frequency of HVS-I sequences carrying characteristic to hg V transition in np 16298 has been shown to be especially high, ranging over 40% (Sajantila *et al.* 1995). The diversity of hg V is, however, very low among the Saami (Torroni *et al.* 1998). It has been proposed that hg V is a mtDNA marker for the population expansion after LGM that started from the Iberian Peninsula 10,000–15,000 years ago (Torroni *et al.* 1998).

Hg U is the most ancient haplogroup in Europe and embraces numerous phylogeographically different sub-clades (fig. 5), some of which can be found from Africa (*e.g.* Rando *et al.* 1998; Rosa *et al.* in press), Siberia (*e.g.* Derbeneva *et al.* 2002*b*; Derenko *et al.* 2003; Fedorova *et al.* 2003), the Near-East and the Caucasus-area (Metspalu *et al.* 1999; Richards *et al.* 2000), as well as in South Asia (Kivisild *et al.* 1999*a*; 1999*b*; 2003*b*; Bamshad *et al.* 2001). The expansion of hg U has likely started more than 50,000 YBP (Torroni *et al.* 1996; Richards *et al.* 1998; 2000).

Hg U5 is the largest and most diverse branch of hg U. It can be further divided into a number of subclades (Richards *et al.* 1998; 2000, see fig. 5), but because many of them are thus far defined only by fast evolving nps in HVS-I, more extensive knowledge about its coding region variation is needed before the in-depth phylogenetic analysis of different subclades can be vigorously applied. However, some of the sub-clades can be recognized also by their specific HVS-I motifs. One of these is the so-called "Saami-motif", distinguished from the root of hg U and the CRS by three transitions in nps 16189, 16144 and 16270 (Sajantila *et al.* 1995).

Hg U4 is the second largest sub-clade of U in Europe. According to Richards *et al.* (1998), hg U4 dates back to more than 25,000 years and, similarly to hg U5, started to expand already before the LGM. Its frequency seems to be higher in eastern and southern than in western European populations and is the highest (16%), surprisingly, among western Siberian Ugric-speaking Mansis (Derbeneva *et al.* 2002*b*). Hgs U1, U2, U3, U6, U7 and U8 are, particularly in Europe, less frequent subclades of hg U. From those, U1, U3, U6 and U7 are present at relatively higher frequencies in southern European populations and in the Near East, North Africa (Macaulay *et al.* 1999*b*; Richards *et al.* 2000) and the Caucasus area (Macaulay *et al.* 1999*b*; Metspalu *et al.* 1999; Tambets *et al.* 2000), whereas U2 can be found throughout the Near East and Europe (Macaulay *et al.* 1999*b*). As it was indicated above, distinct sub-branches of U2 are specific for South Asia (Kivisild *et al.* 1999*a*; 2003*b*; Bamshad *et al.* 2001). Hg K, one of the first described haplogroups in the European population (Torroni *et al.* 1994*d*), was shown to belong to hg U (Hofmann *et al.* 1997). Hg K seems to be more common in western than in eastern

European populations (table 1), and is well present in the Near East and Anatolia (Kivisild *et al.* 2003*a*). The highest frequency of hg K has been observed among Ashkenazi Jews (30%, Ritte *et al.* 1993; Behar *et al.* 2004).

Hgs T and J (fig. 5) are sister clades, present in European populations at similar frequencies (6%–14%, see also table 1). Both of them have a complex inner topology with several sub-founders (*e.g.* Finnilä and Majamaa 2001; Herrnstadt *et al.* 2002), some of which might have arrived to Europe with the Neolithic immigration from the Near East (Richards *et al.* 2000).

Hgs that stem directly from the common node N (I, N1a, N1b, N1c, W, X, see fig. 5) are relatively rare in Europe and do not usually exceed the level of 5% there (Richards *et al.* 1998). Hg I, together with N1a, N1b and N1c, belongs to the clade N1 (see fig. 5) and is spread mostly in northern and western Europe (Richards *et al.* 2000). Hg X, despite its minor contribution to the package of maternal lineages of most of the populations, has a surprisingly wide geographical distribution. Described first as a European specific mtDNA variant (Torroni *et al.* 1996) it has by now been found from North and East Africa, Near East, Siberia as well as among Native Americans (Forster *et al.* 1996; Brown *et al.* 1998; Smith *et al.* 1999; Derenko *et al.* 2001). Hg W has higher diversity in southern than in northern European populations (Richards *et al.* 1998) and it has been found also, for example, in South Asia (Kivisild *et al.* 1999*b*). The frequency of hg W is particularly high among Finns (table 1), its diversity, however, is very low there (Meinilä *et al.* 2001).

Most European haplogroups show greater haplotype diversities as well as deeper coalescence ages in the Near East than in Europe (Torroni *et al.* 1998; Richards *et al.* 2000), supporting their Near Eastern origin. The coalescence ages of the primary founder haplotypes have shown that the majority of them have migrated to Europe in the Late Upper Palaeolithic-Mesolithic period (Richards *et al.* 2000). One exception is hg V, which seems to have originated and expanded within Europe 10,000–15,000 YBP (Torroni *et al.* 1998). Another is hg U5. Although it occurs at about 2% in the Near East, its phylogeographical spread suggests that it evolved mainly within Europe during the past 50,000 years and its presence in the Near East can be the result of back-migration from Europe (Richards *et al.* 2000).

2.7.3.1. The Saami — genetic "outliers" of Europe

The Saami — aboriginal inhabitants of Fennoscandia — are regarded as extreme genetic outliers among European populations (Cavalli-Sforza *et al.* 1994). At present, *ca* 60,000 Saami live in the northern regions of Norway, Sweden and Finland (Haetta 1996). Saami languages belong to the Finno-Ugric branch of the Uralic language family. The closest linguistic neigbours of the Saami are the Finns, the Karelians, and the Estonians.

Analyses of classical genetic markers have demonstrated that the genetic distances between the Saami and other Europeans are significantly larger than between any other pair of European populations (Cavalli-Sforza *et al.* 1994). It has been suggested that the genetic composition of the Saami arose from extensive admixture between Caucasoid and Mongoloid populations (Guglielmino *et al.* 1990; Cavalli-Sforza *et al.* 1994, but see also *e.g.* Beckman et al. 1988, 1993). Similarly, studies of mtDNA have identified large genetic distances between the Saami and other Europeans, including the Finns (Sajantila and Pääbo 1995; Sajantila *et al.* 1995). Likewise, Lahermo et al. (1996) found no overlap between Saami and the remaining European mtDNA patterns and concluded that the Saami and the Finns must have different genetic histories. One hypothesis to explain the presence of genetic differences and language similarities in the Finns and the Saami involves a language shift by the Finns from Indo-European to Finno-Ugric (Sajantila and Pääbo 1995). However, further studies of mtDNA variation revealed that the majority of mtDNA lineages of the Saami are clustered in a subset of the European mtDNA pool (Torroni *et al.* 1998; Villems *et al.* 1998; Tambets *et al.* 2000).

Nearly half of the Y chromosomes of the Saami share a TatC allele or haplogroup N3 with most Finno-Ugric and Siberian populations. This variant is found at high frequencies among Siberian populations, such as the Yakuts and the Buryats, but is virtually absent in western Europe. These findings have been interpreted according to the classic view that a substantial element of the Saami genetic lineages originated in a recent migration from Asia (Zerjal et al. 1997, 2001).

2.8. Pre-historical, archaeological and linguistic context for present European mtDNA variation

The beginning of the Upper Palaeolithic, 40,000–50,000 YBP, marks the first appeareance of AMHs in Europe. Before that, Neanderthals had occupied Europe for about 250,000 years. It appears now that the coexistance of Neanderthals and AMHs in Europe lasted for about 10,000 years or even longer (Stringer and Davies 2001).

The mtDNA studies of the extant human populations are in accordance with the theory of recent total replacement of earlier forms of humans by *Homo sapiens sapiens*. Already before the first DNA sequences for Neanderthals became available, it was predicted, based on palaeontological material, the estimations of the divergence times of Neanderthals and AMHs, and the observed mtDNA sequence variations in extant European populations, that the genetic continuity between Neanderthals and AMHs is highly unlikely (Torroni *et al.* 1994*d*; Richards *et al.* 1996). Soon, these predictions got direct support

from genetic data, gathered from partial sequencing of the control region of fossile DNA, extracted from Neanderthals. It has been unequivocally shown that the Neanderthal mtDNA falls outside the variation of mtDNA of modern humans (Krings *et al.* 1997; 1999; Ovchinnikov *et al.* 2000; Serre *et al.* 2004). Phylogenetic analysis and the age of the MRCA for the mtDNAs of the Neanderthal and modern humans was estimated to be about four times older than the age of the MRCA of modern human mtDNAs. These results indicate that there is no genetic contribution of Neanderthals to the mtDNA gene pool of modern European populations. This discontinuity is supported as well by the mtDNA analyses of the remains of two ca 24,000 old Cro-Magnon type individuals (Caramelli *et al.* 2003). The obtained results have revealed that the mtDNA types of these early anatomically modern pre-LGM European individuals are well within the range of the variation of present-day humans.

During the peak of the Last Ice Age, about 24,000–20,000 YBP, a large part of Europe was unsuitable for human occupation. In the cold period the population of Europe was concentrated to the so-called refuge areas or refugia — regions that maintained more favorable environmental conditions at the LGM-period. After the LGM, northern regions of Europe were gradually re-colonized by Late Upper Palaeolithic populations spreading out from different refugia — the Franco-Cantabrian, Ukrainian and possibly some others, situated, *e.g.* in the southern Alps, in so far less well-identified locations south of the lower Danube and Balkans (Housley *et al.* 1997; Dolukhanov 2000).

One still hotly debated issue on the possible origin of modern Europeans is associated with the onset of farming, its spread in Europe and its genetic influence to the European Mesolithic populations. Two competing hypotheses — the demic and the cultural diffusion models — have been presented. According to the demic diffusion model, food production led to a slow expansion of Neolithic source population from the Near East due to the population growth, resulting from the advantages of agriculture ("wave of advance", Ammerman and Cavalli-Sforza 1984; Cavalli-Sforza *et al.* 1994). According to the most extreme views, it led to the total replacement of less numerous Mesolithic hunter-gatherers, who had lived in Europe before (Barbujani *et al.* 1998; Chikhi *et al.* 1998; 2002). Yet the classical presentation assumes that the demic diffusion involved a substantial minority of Neolithic newcomers (*ca* 27%), arriving from the Near East — either from Levant or Anatolia (Cavalli-Sforza *et al.* 1994). The model of demic diffusion was embraced also by some linguists (Renfrew 1987) and it was proposed that: 1) the genes of Anatolian and Near Eastern populations; 2) the technology of farming; and 3) Indo-European languages, spoken today in most regions of Europe, were together brought to Europe in the course of the same migration. However, most linguists did not agree with this model and preferred the theory of Marija Gimbutas (1970), according to which Proto-Indo-Europeans had spread west with the Bronze-Age "Kurgan culture" of the eastern European steppe (Mallory 1989; Hines 1991). More recently, an extensive study on 87

languages yielded a support for the Anatolian theory of Indo-European origin (Gray and Atkinson 2003). The study also showed evidence of a period of rapid divergence giving rise to the Italic, Celtic, Balto-Slavic and Indo-Iranian families in a close time frame to that suggested for a possible Kurgan expansion. Thus, these two linguistic hypotheses need not be mutually exclusive. Several theories have been postulated also on the formation of the Finno-Ugric branch of the Uralic language family, spoken by a minority of inhabitants of Europe, like, for example, by the Finns, Karelians, Estonians and the Saami. The earliest hypotheses associated the arrival of Finno-Ugric speakers from the Volga-Oka River region, which was assumed to be the Uralic homeland, in Europe *ca* 6000 YBP with a Neolithic Combed Ware culture. According to the latest hypotheses, the migration of Finno-Ugric speakers occurred from the south, almost simultaneously with the reatreat of the ice sheets (for review, see *e.g.* Poikalainen 2001; Wiik 2000).

The opposite model of the origin of modern Europeans assumes, on the contrary, that the onset of farming took place not because of any significant migration of people from the Near East, but because of a cultural transition in the form of diffusion of ideas (Barker 1985; Whittle 1996). One intermediate model, the pioneer colonization, assumes the selective migration of only small groups from western Asia (Zvelebil 1986; 2000; van Andel 2000) and the continuity of indigenous Mesolithic populations. This last model, with some modifications (Wilson *et al.* 2001), explaining the presence of the Neolithic component in comparable frequencies to this of Central and South Europe, also in North and northeastern Europe, has also gained support from mtDNA studies (Richards *et al.* 1996; 2000; Torroni *et al.* 1998; Richards 2003), where it was suggested that most of the contemporary maternal lineages in Europe have their ancestry in the Late Glacial expansions within Europe, associated with the climatic improvements following the last Ice Age, whereas only a minor part is dated either to the initial early Upper Palaeolithic settlement of the continent by AMHs (hg U), or brought to Europe by later immigrants during the early Neolithic (hgs J, T1, U3 and few subclusters of H and W, Richards *et al.* 2000). The scenario of limited contribution (20–25%) of the Neolithic component to the gene pool of modern Europeans has been supported also by studies of Y-chromosomal markers (Rosser *et al.* 2000; Semino *et al.* 2000; Scozzari *et al.* 2001).

3. AIMS OF THE PRESENT STUDY

The main goal of the present study was to improve the knowledge about mtDNA variation in Europe and in its surrounding regions. Inasmuch as there is evidence that the vast majority of the present-day European genes descend from indigenous Palaeolithic ancestors (Richards *et al.* 2000; Semino *et al.* 2000; Richards 2003), a better understanding of genetic variation requires the reconstruction and distinction of the gene pools in different glacial refugia. In order to draw conclusions about the processes that might have been involved in the shaping of present mtDNA diversity, the aim of this study was to apply, where possible, the phylogeographic approach on the mtDNA lineage clusters, which are common and widespread in Europe. Hgs U, V and X were taken into specific focus to analyze their possible expansion times and distribution in different populations.

One particularly fascinating example in the European genetic landscape is the Saami population, shown to be a genetic "outlier" in Europe. Here, my aim was to shed further light on the questions of their genetic history.

The following questions were posed:

1. What might have been the main processes that have shaped present mtDNA variation in Eurasia? Which are the "molecular signals" that can be associated with the major prehistoric population movements?

2. What is the phylogenetic affiliation of mtDNA lineages spread among the Saami in the Eurasian context?

3. From where did these mtDNA lineages possibly arise? How did they reach northernmost Fennoscandia?

4. SUBJECTS AND METHODS

The experimental basis of the current thesis relies on the analysis of mtDNA variation in the following populations, analysed by the author: 73 Swedish Saami (from Norrbotten county, Sweden), 557 Estonians (from Tartumaa, Järvamaa, Jõgevamaa, Ida- and Lääne-Virumaa, Läänemaa, Põlvamaa, Valgamaa, Viljandimaa, Võrumaa, Harjumaa, Pärnumaa and Saaremaa county, Estonia), 398 Swedes (from Skåne, Södermanland, Uppland, Norrbotten county and from island Gotland, Sweden), 299 Latvians (from North and South Curonia, Lettigallia and Semigallia, Latvia), 45 Lithuanians (from Lithuania), 199 Albanians (from Tirana, Albania), 102 Mordvin (from Mordovia, Russia), 136 Komis (from Komi Republic and Komi-Permyak Autonomous district, Russia), 101 Udmurts and 40 Bashkirs (from Bashkortostan, Russia), 55 Chuvashes (from Chuvashia, Russia), 40 Tatars (from Tatarstan, Russia), 116 Hungarians (from Pest county, Hungary), 40 Khants and 38 Mansis (from Khanty-Mansiisk Autonomous District, Russia) and 388 Turks (from Cappadocia, Turkey). DNA samples were collected in collaboration with the Institute of Biochemistry and Genetics, Ufa, Russia (Dr. Elza Khusnutdinova), Gotland University, Visby, Sweden (Dr. Lars Beckman), Institute of Forensic Sciences, Budapest, Hungary (Dr. S. Füredi), Faculty of Natural Sciences of Tirana University, Tirana, Albania (Dr. I. Mikerezi), Medical Academy of Latvia, Riga, Latvia (Dr. A. Krumina) and University of Kiel, Kiel, Germany (Dr. M. Gölge). DNA samples were collected from healthy unrelated individuals after obtaining informed consent. In additon, all U5 samples (in ref. I) were analyzed by the author for markers in the mtDNA coding region.

DNA samples were amplified by PCR reaction and sequenced for 360 bp in HVS-I region (nps 16024–16383), using the primers from Vigilant *et al.* (1989). Descriptions of methods are given in references I–III. The sequences were aligned and polymorphic nps were determined using the Genetics Computer Group Wisconsin Package. The HVS-I sequences were reported in terms of mutations of CRS (Anderson *et al.* 1981). The haplogroup affiliations were affirmed by RFLP analysis of informative coding region nps, according to the present mtDNA nomenclature (see ref. I). Haplogroup and haplotype frequencies as well as diversities (calculated as in Nei 1987) of particular lineage clusters and sub-clusters were calculated and compared, including the available published data. Population data were compared by the use of principal component analysis (program POPSTR), based on haplogroup frequencies. For further analysis, the phylogenetic networks relating different haplotypes (as described in Bandelt *et al.* 1995; 1999) were constructed. The time estimates for founding haplotype clusters were calculated by use of statistic ρ, the average transitional distance from the putative founder sequence, and calibrated using a transition rate of 1 in 20,180 years for the HVS-I region between nps 16090–16365 (Forster *et al.* 1996). The standard error (σ) was calculated as in Saillard *et al.* (2000*a*).

5. RESULTS AND DISCUSSION

There are several possibilities to analyze the data of uniparentally inherited marker systems. One may compare the mtDNA variation (*e.g.* frequencies of haplogroups) between different populations or population groups in order to find out the similarities and differences in the composition of their maternal gene pools (reference **V**). Alternatively, one may apply a phylogeographic approach in order to study the global spread of particular lineages (references **II, III** and **IV**). The two approaches can be combined (references **I**), so that a wide phylogeographical analysis of the major maternal lineages present in a population (here, in the Fennoscandian Saami) can provide information, which helps to shed light on the pre-history of the population under study. This approach can be particularly informative in combination with a parallel and phylogeographically comparable study of Y-chromosomal heritage. In this dissertation, the main stress will be on the results associated with mtDNA analysis.

5.1. The phylogenetic affiliation of the maternal lineages of the Saami on the Eurasian mtDNA tree (Ref. I, II, V)

5.1.1. MtDNA analysis of the Saami and some notes about their Y chromosomal variation

Our analysis of Swedish Saami, combined with the reanalysis of previously published sequences of other Saami populations from Finland and Norway (Sajantila *et al.* 1995; Dupuy and Olaisen 1996; Delghandi *et al.* 1998) showed that the "outlying" status of the Saami is caused by relative haplogroup proportions in their mtDNA pool, not by distinctive phylogeographical affiliations of their maternal lineages (fig. 2A in ref. **I**). As a specific novelty, we have shown that there is little, if any, historic gene flow from Samoyedic- and Ugric-speaking populations from Siberia to Fennoscandia alongside the Arctic zone. The Saami mtDNA pool consists predominantly of two haplogroups that are widely spread in Europe: V and U5 (table 1 and 2 in ref. **I**; fig. 2 in ref. **V**; table 1 in ref. **II**). These haplogroups together cover more than 90% of the Saami mtDNA lineages. The rest of the variation is shared by European-specific hgs H, W and T (fig. 1 in ref. **I**). Only a small fraction of mtDNA variants shows the contribution from eastern Eurasia (hgs D5 and Z1, table 1 in ref. **I**). The most plausible explanation for the unconventional frequency pattern of mtDNA haplogroups among the Saami is that genetic drift (bottleneck and founder-effect) has had a major role in shaping the mtDNA pool of this small European population. One can bring many examples from literature showing that small

isolated populations may drastically differ, in haplogroup frequencies, from their "parent population". For example, the Maori mtDNA pool in New Zealand is composed almost exclusively of hg B (Murray-McIntosh *et al.* 1998), though that of the likely ancestors of the Maori in Southeast Asia is diverse.

Although the Saami Y-chromosomal heritage is not the topic of this PhD dissertation, it seems appropriate to discuss it here briefly. In contrast to mtDNA, the Saami Y-chromosomal pool does not differ profoundly from that characteristic to their geographical neighbours, in particular as far as the Finnic-speaking populations are concerned (fig. 2B in ref. **I**). Compared to Germanic-speaking populations of Fennoscandia — Swedes and Norwegians — the Saami Y-chromosomal repertoire is rich in hg N3 (*alias* hg 16 or the TatC allele, in older literature, incl. ref. **V**). N3 encompasses most of the Y-chromosomal lineages among the Saami and in many other North European as well as northern Asian populations, whereas its frequency is less than 10% in the Norwegians and Swedes (table 3 in ref. **I**). The highest frequency of N3 (TatC) has been observed among Yakuts and Buryats (fig. 6 in ref. **V**). Here, it is important to note that N3 is by far more diverse in eastern Europe than in Siberia (fig. 5 in ref. **V**). Moreover, sister-branch of hg N3, N2 (previously hg 12, which was earlier and in ref. **V** considered as the ancestral group of N3) is also well present in eastern European, and, in particular, in several Siberian populations (table 3 in ref. **I**). Therefore, although the Saami share their predominant hg N3 with many Siberian populations, we do not need to assume recent directional gene flow from Siberians to the Saami. Yet a common Y-chromosomal ancestry of many northern Eurasian populations and the Saami (as well as the Finns, Estonians, Latvians *etc*.) is obvious, probably going back to deep prehistoric, likely Palaeolithic times.

5.1.2. Phylogeography of haplogroup U5b1b

Most of the U5b lineages spread among the Saami possess the so-called "Saami-motif" — a combination of three transitions, 16144 (T to C), 16189 (C to T) and 16270 (C to T) in their HVS-I sequences, which was previously believed to be restricted only to the Saami — occasional findings from neighbouring populations, in the Karelians and Finns, were explained by their admixture with the Saami (Sajantila *et al.* 1995; Meinilä *et al.* 2001).

We analysed the informative coding region nps (table 4 in ref. **I**) in U5b topology and showed that "the Saami variant" of U5b (U5b1b1), predominant among the Saami, is widely spread among different eastern European populations, extending, at very low frequencies, also to western Europe and to the Caucasus (fig. 3B and table 1 in ref. **I**).

On the other hand, U5b1b1 forms a sub-branch of a lineage cluster, defined here as U5b1b (5656G, 7385G, 10927C). Phylogeographic analysis of U5b1b

revealed that this haplogroup, similarly to V, another major haplogroup in Saami (see also ref. **II**), has a greater diversity in western than in eastern Europe, supporting a scenario, according to which they both started their expansion after the LGM from western parts of Europe, possibly from Franco-Cantabrian refugium. Meanwhile, the analysis of different European populations suggests that these haplogroups have likely reached northern Fenno-scandia not along the Atlantic coast of the Scandinavian Peninsula, but rather via the "eastern route" (fig. 4 in ref. **I**), as the haplotypes of hgs U5 and V of the Saami are well present in eastern European populations. The phylogeographic pattern of U5b1b1 suggests that this particular sub-clade might have arisen in East Europe.

The solitary location of the Saami in the European genetic landscape has been interpreted both as a significant Mongoloid component in their gene pool and their possible Siberian ancestry (Cavalli-Sforza *et al.* 1994, and references therein) or by an early split and further genetic isolation of the Saami from other Europeans (Sajantila *et al.* 1995). The main reason why earlier studies of uni-parentally inherited genetic markers could not shed clear enough light on the question about the place of maternal lineages of the Saami in the Eurasian mtDNA tree was the lack of appropriate phylogenetic resolution and phylo-geographic width. In light of new data, the phylogenetic analysis of mtDNA (table 2) as well as Y-chromosomal markers clearly revealed that both maternal and paternal lineages of the Saami originate from the same sources as those of other northern European populations.

Table 2 Frequencies (%) of predominant mtDNA haplogroups of the Saami in some northern Eurasian populations (data from ref. I)

Population	Region	n	V	U5b1b1	H1	D5	Z
Saami	Europe	445	41.6	47.6	2.5	3.1	1.3
Swedes	Europe	503	3.2	0.8	3.4	0	0.4
Finns	Europe	641	6.4	6.7	3.6	0.2	1.5
Estonians	Europe	545	3.3	0.7	4.8	0.2	0
Russians	Europe	761	3.8	1.2	1.2	0	0.3
Mordvin	Europe	111	3.6	2.7	0	0	0
Khants	Siberia	255	0	0	0	0	0
Mansis	Siberia	138	0.7	0	0	0.7	0
Nganasans	Siberia	131	0	0	0.8	0	2.3
Nenets	Europe/Siberia	137	0	0	0	0	0
Selkups	Siberia	120	0	0	0	0	0

5.2. Post-LGM re-colonization of Europe and beyond as revealed from phylogeographical pattern of the maternal lineages (Ref. II, III, IV)

In order to find traces of putative post-LGM large-scale demographic events, including migrations linked to climatic and environmental changes, from the genetic heritage of European populations, we carried out several phylogenetic and phylogeographic case studies involving mtDNA haplogroups that are spread over Europe (and beyond).

5.2.1. MtDNA haplogroup V (ref. II)

The analysis of hg V, and its presumable phylogenetic precursors — different variants of pre*V, in several populations from Europe and surrounding regions showed that while pre*V tends to be scattered throughout Europe as well as in northern Africa, hg V has two peaks of frequency: one in southwestern Europe (at highest frequency in the Basques — 12%) and the other in Fennoscandia, among the Saami (52%, table 1 and fig. 3 in ref. **II**). This frequency pattern suggests that one of those regions might be the place from where the expansion of hg V started. However, relatively recent extreme manifestations of random genetic drift (bottlenecks and subsequent founder-effects) might have strongly affected the observed haplogroup frequency patterns. To address this issue, the diversity values for populations contributing more than 5% to the total V dataset were calculated (table 2 in ref. **II**). The exclusion of populations with low hg V diversity values erased the frequency peak in Scandinavia but did not affect the one in southwestern Europe (fig. 4 in ref. **II**). The age estimates for different subsets of hg pre*V showed that the expansion of pre*V probably took place already before the LGM (table 3 in ref. **II**). The same calculations for hg V showed that the expansion of hg V was more recent. At the same time, the coalescence age of hg V in the west (16,300 ± 4,800 YBP) is almost twice of that in the east (8,500 ± 2,300 YBP, see table 3 in ref. **II**). Geographical distributions and time estimates can be interpreted in such a way that while hg pre*V originated in Europe before the LGM, then hg V seems to have arisen in a southwestern European refugium after the LGM and reached eastern and southern Europe even much later.

5.2.2. MtDNA haplogroup X (ref. III)

While different mtDNA haplogroups, as a rule, have relatively restricted geographical spread, hg X is an exception. It is found at low frequencies both in western Eurasia (Richards *et al.* 2000) and in northern populations of Native Americans (Ward *et al.* 1991; Forster *et al.* 1996; Brown *et al.* 1998; Malhi *et al.* 2001; Malhi and Smith 2002). Most importantly in this context, hg X has not been found among modern northern Siberian populations or from East Asian populations — the possible source populations for Native Americans. Hg X, however, has been detected in Altaian population and it was suggested that this is a trace that speaks in favour of the Beringian flow of hg X chromosomes to North America (Derenko *et al.* 2001).

In this study, five mtDNAs from hg X were completely sequenced. A maximum parsimony tree, which additionally included 16 published complete sequences of hg X, was constructed in order to dissect the hg X variation into sub-clades and to analyze their spatial distribution. Our analysis showed that hg X lineages can be grouped into two basic branches, X1 and X2. The former is largely restricted to North and East Africa, whereas X2 is spread widely throughout western Eurasia (fig. 1 in ref. **III**).

X2, which started its spread most probably from the Near East, can be further dissected into six sub-clades (fig. 1 and 2 in ref. **III**). One of those, X2a, is found only among some northern Native American tribes. This American-specific sub-clade has derived from X2 by a unique combination of five mutations (fig. 1 in ref. **III**) and is not related to those few hg X2 members that are so far found from Siberia (Derenko *et al.* 2001), speaking strongly against a cited above suggestion that brought together the Altaian and Amerind hg X variants.

One third of the European hg X sequences belong to sister-groups X2b and X2c. Populations from southern Europe possess hg X2 at higher frequencies than those from northern and northeastern populations ($p<0.05$). Hg X2 is very rare in eastern European populations (table 1 in ref. **III**). Expansion time estimates of X2, based on HVS-I and coding region variation, are 17,900 ± 2,900 YBP and 21,600 ± 4,000 YBP, respectively (fig. 1 and 2 in ref. **III**). These estimates are consistent with the expansion of X2 around and, possibly, just after the LGM, when the climate started to improve. For X1, the HVS-I based time depth is significantly older (42,900 ± 18,100 YBP), though associated with large standard errors, likely due to smaller sample size and less star-like tree topology. This peculiar topology of X1 suggests that possibly only a restricted variation of this cluster has survived from the early Upper Palaeolithic period; it also suggests an early split between X1 and X2.

5.2.3. MtDNA haplogroups U4 and U5 (ref. IV)

Phylogeographic analysis of the two largest subclusters of common and widespread western Eurasian hg U – U5 and U4 – was performed.

U5 is a "major prototype" western Eurasian mtDNA clade. Its coalescence age is 45,000–53,000 YBP and the phylogeny of U5 is shown to be non-star-like (Richards *et al.* 2000). The essential question for our work was to understand whether this topologically complex clade can be "resolved" into subclusters of a more ordered topology, thus offering, compared with hg U5 as a whole, better insights into the recent patterns of the spread and expansion of western Eurasians.

Our analysis of a large dataset of U5 mtDNAs (n = 526) based mostly on HVS-I sequences and essential diagnostic site for hg U (+12,308 *Hinf*I) revealed that U5 ramifies into many potential sub-founders (fig. 35.1 in ref **IV**; fig. 6). The important aspect of this analysis was that almost all of them show a star-like topology and exhibit post-LGM coalescence ages around the end of the Pleistocene period. These findings suggest that hg U5 underwent a post-LGM expansion phase. Furthermore, this temporal pattern is similar to that of hgs V and X (ref. **II** and **III**, see also table 3).

Hg U5 is relatively frequent all over western Eurasia (Richards *et al.* 1998; Richards *et al.* 2000), whereas U4 seems to be more frequent in eastern Europe (Bermisheva *et al.* 2002) and is either absent or very rare in the Near East (Al-Zahery *et al.* 2003; Kivisild *et al.* 2003*a*) and elsewhere. Its highest frequencies can be actually observed in Ob-Ugric Khants and Mansis, who live nowadays on the eastern side of the Ural Mountains (Derbeneva *et al.* 2002*b*, our unpublished data). The coalescence age of U4 was suggested to lay around 16,000–25,000 YBP (Richards *et al.* 2000). The phylogenetic tree of U4 (fig. 35.2 in ref. **IV**) that we constructed from hg U4 mtDNAs (n = 386) shows a limited number of subfounders. Sub-clade U4a (defined by HVS-I motif 16134–16356) and U4b (defined by HVS-I motif 16179–16356) likely encompass monophyletic lineages. Lineage cluster U4c (defined by HVS-I motif 16356–16362) is more likely a paraphyletic group because np 16362 is one of mutational "hotspots" (see section 2.2.3.). We calculated the coalescence ages for different subfounders in different geographical subsets of U4 and found that in eastern Europe, the coalescence ages are more or less the same for all main founder-haplotypes (except for U4b, that is very rare in eastern regions of Europe), being around the LGM. Taking into account the data from other disciplines (*e.g.* Dolukhanov 2000), this might suggest that the expansion of U4 in eastern Europe started during the formation of the Periglacial (northern Ukrainian) refugium around the peak of the LGM. The corresponding values in western Europe are, again, around the end of Pleistocene, close to those for hg V (ref. **II**) and subclusters of hg U5.

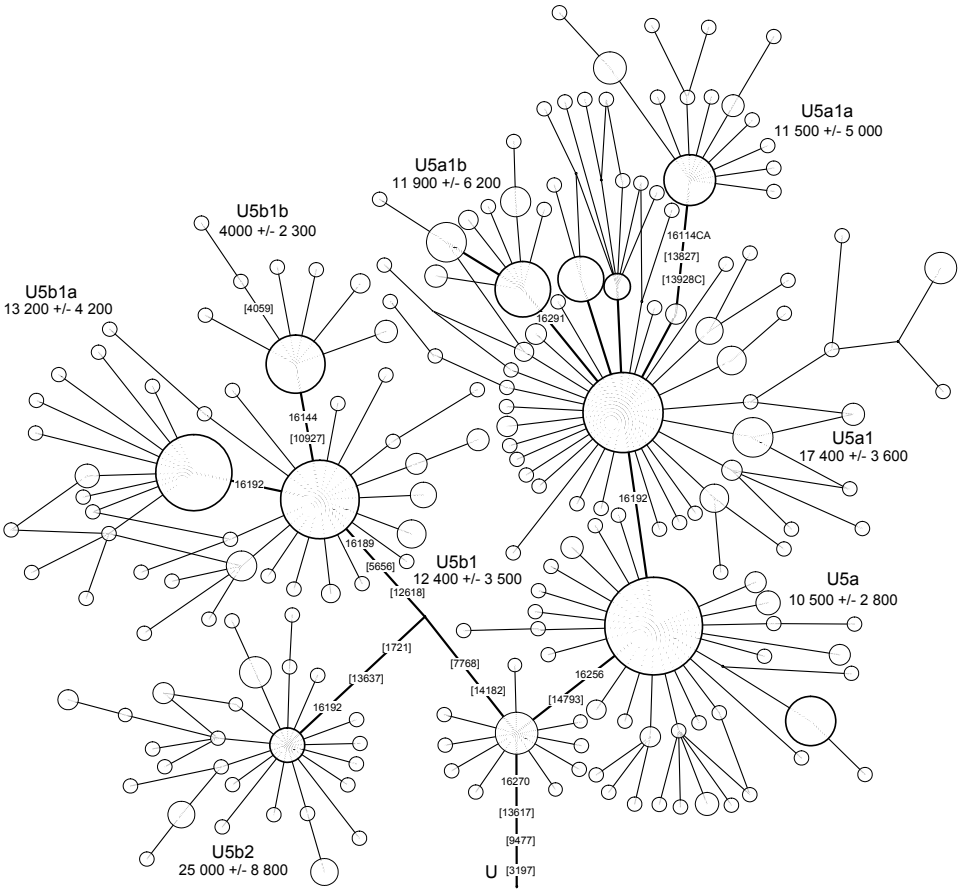

Figure 6 Phylogenetic network of haplogroup U5 HVS-I sequences (n = 526). Coalescence ages for larger subhaplogroups are indicated. Coding region information is from Finnilä *et al.* 2001 and Herrnstadt *et al.* 2002 (note that the definition of subhgs is in some cases different here). Populations used in coalescence age calculations are given in the footnote of table 3.

Table 3 Coalescence ages of different subsets of haplogroups U4, U5, V and X

Hg	Region	n	ρ	T	ΔT	reference
U4a	Europe	82	0.829	16,700	4,300	IV
U4a	Eastern Europe	50	1.000	20,200	5,900	IV
U4a	Western Europe	27	0.519	10,500	3,500	IV
U4b	Europe	31	0.355	7,200	2,700	IV
U4b	Western Europe	17	0.529	10,700	4,600	IV
U4c	Europe	31	1.000	20,200	7,500	IV
U5a	Europe	94	0.521	10,500	2,800	IV
U5a1	Europe	87	0.862	17,400	3,600	IV
U5b1	Europe	67	0.612	12,400	3,500	IV
U5b1b1	Europe	118	0.212	4,300	1,400	I
U5b2	Europe	25	1.240	25,000	8,800	IV
V	Europe	205	0.556	11,200	2,700	II
V	Western Europe	98	0.806	16,300	4,800	II
V	Eastern Europe	81	0.420	8,500	2,300	II
X2	Eurasia, Africa	159	0.887	17,900	2,900	III
X2	Europe	28	1.143	23,100	2,100	III

NOTE: Hg — haplogroup; n — sample size; ρ — the average mutational distance to the assumed founder haplotype of the cluster; T — time to the most recent common haplotype of sequences under survey (calculated as in Forster *et al.* 1996); ΔT — standard error (as in Saillard *et al.* 2000*a*). Coalescence age calculations are based on the phylogenetic networks in fig. 3B (ref. I); fig. 2 (ref. II); fig. 2 (ref. III); fig. 35.1 and 35.2 (ref. IV); fig. 6. In case of U4 and U5, the HVS-I sequences of the following populations were used (given in alphabetical order):

Albanians (Belledi *et al.* 2000; our unpublished data); Austrians (Parson *et al.* 1998); Bashkirs (Bermisheva *et al.* 2002), Basques (Bertranpetit *et al.* 1995; Corte-Real *et al.* 1996); Bulgarians (Calafell *et al.* 1996; Richards *et al.* 2000); Chuvashes (Richards *et al.* 2000; Bermisheva *et al.* 2002); Crete (our unpublished data); Czechs (Richards *et al.* 2000; our unpublished data); Croats (our unpublished data); Danes (Richards *et al.* 1996); English (Piercy *et al.* 1993; Richards *et al.* 1996; 2000; Helgason *et al.* 2000); Estonians (Sajantila *et al.* 1996; our unpublished data); Finns (Sajantila *et al.* 1995; Richards *et al.* 1996; Kittles *et al.* 1999; Finnilä *et al.* 2001); French (Rousselet and Mangin 1998; Danan *et al.* 1999; Richards *et al.* 2000; our unpublished data); Germans (Richards *et al.* 1996; Hofmann *et al.* 1997; Baasner *et al.* 1998; Lutz *et al.* 1998; Pfeiffer *et al.* 1999); Greeks (Richards *et al.* 2000); Hungarians (our unpublished data); Icelanders (Helgason *et al.* 2000); Irish (Richards *et al.* 2000); Italians (Francalacci *et al.* 1996; Richards *et al.* 2000; our unpublished data); Karelians (Sajantila *et al.* 1995); Khants (our unpublished data); Komis (Bermisheva *et al.* 2002); Lithuanians (our unpublished data); Mansi (Derbeneva *et al.* 2002*b*; our unpublished data); Maris (Bermisheva *et al.* 2002); Mordvin (Bermisheva *et al.* 2002); Nenets (Saillard *et al.* 2000*b*; Norwegians (Opdal *et al.* 1998; Helgason *et al.* 2000; Richards *et al.* 2000); Poles (Richards *et al.* 2000; our unpublished data); Portuguese(Corte-Real *et al.* 1996); Romanians (Richards *et al.* 2000); Russians (Orekhov *et al.* 1999; our unpublished data); Saami (Sajantila *et al.* 1995; Dupuy and Olaisen 1996; Delghandi *et al.* 1998); Sardinians (Di Rienzo and Wilson 1991; Richards *et al.* 2000); Scots (Helgason *et al.* 2000); Sicilians (Richards *et al.* 2000; Forster *et al.* 2002; our unpublished data); Slovaks (our unpublished data); Spaniards (Corte-Real *et al.* 1996; Salas *et al.* 1998; Crespillo *et al.* 2000); Swedes (Sajantila *et al.* 1996); Swiss (Pult *et al.* 1994; Dimo-Simonin *et al.* 2000); Tatars (Bermisheva *et al.* 2002); Udmurts (Bermisheva *et al.* 2002); Ukrainians (our unpublished data); Welsh (Richards *et al.* 1996; 2000).

Whereas Richards *et al.* (1996) re-opened discussion about the role of neolithization in the extant mtDNA gene pool of European populations and while Torroni *et al.* (1998) gave the first hint that the phylogeography of hg V may be interpreted as a sign of post-glacial population expansion from refugia, it became also clear that by far more comprehensive data sets are needed to reveal whether such kind of molecular signals are indeed general or do they reflect demographic events of isolated nature. Why such a question was relevant to ask arose also from the fact that hg V is only a minor haplogroup in the European mtDNA pool. Secondly, hg V was also well present in Northwest Africa (Torroni *et al.* 1998), which made the place of its initial expansion unclear. Thirdly, at that time, next to nothing was known about the phylogeography of hg V in eastern Europe — except of its very high frequency among the Saami population.

Furthermore, while a paper by Richards and collaborators (Richards et al. 2000) suggested ages for major (and some minor) western Eurasian mtDNA haplogroups, it left unanswered an equally important question about their "demographic history" after their arrival in Europe. The task of my work was to see whether there are systematic signals for expansion, present in many haplogroups, and if there is, then what is their possible time frame.

Presented here results confirm and extend a concept (Richards *et al.* 1996; 2000) that the current basic branches of the European mtDNA tree that were present in Europe/western Eurasia already before the LGM, were involved in the post-LGM recolonization processes. This conclusion is based not only on the coalescence age calculations of various subsets of mtDNA haplogroups, but equally so on phylogeographic evidence. In this respect, one may suggest that the sum of demographic events that consists of the LGM-generated contraction of pre-LGM European populations into a limited number of largely isolated refugia for many millennia, followed by recolonization of northern Europe, has had a profound effect on shaping the present phylogeographic pattern of mtDNA lineage clusters. As was first suggested based on an example of hg V (Torroni *et al.* 1998), "molecular signals" from such events can be traced from a detailed phylogeographic study. The results presented here demonstrate that the "post-LGM expansion signal" is seen equally well in hgs V, X, U4 and U5 (table 3), allowing to generalize the concept.

6. CONCLUSIONS

The results of the present study can be summarized as follows:

1. We have shown that the dissection of mtDNA variation into precisely — defined monophyletic units. *i.e.* sub-branches of the more widely spread and therefore less informative "classical" haplogroups like U5, U4 and X, is essential in the identification of phylogeographically informative spatial frequency patterns, in order to find their possible association with demographic events. The estimations of expansion times of mtDNA hgs V, U and X let us suggest that their spread began from different European refugia soon after the end of the LGM and the melting of the Fennoscandian Ice Sheet.

2. We have shown — with the example of the Saami — that calculations of genetic distances between populations, without knowledge about the phylogenetic relationships between compared characters, can lead to interpretations of a dubious value concerning questions like the genetic "origin" of a population. Seemingly quantitative, such an analysis tends to expose aspects that do not allow one to distinguish "outliers" — in the strict meaning of this term, from phenomena generated by genetic drift. The mtDNA and Y-chromosomal analysis of the Saami showed that the "outlying" position of the Saami in the genetic landscape of Europe is generated by the effects of genetic drift and not by the different origin of their maternal and paternal lineages from those of other Europeans/European Finnic-speakers.

3. Our results confirm that all sub-populations of the Saami are characterized by a restricted variation of maternal lineages predominantly European in origin (hg V and U5). Less than 5% of mtDNA variants, spread among the Saami, belong to eastern Eurasian-specific haplogroups. A deep phylogenetic analysis of the "Saami-specific" hg U variants in a comprehensive European-Siberian context allowed us to demonstrate that:

    3.1. the "Saami U5" — U5b1b1 — is actually present, though at low frequencies (up to 7%), but highly diversely, in a wide area all over Europe, in particular in East Europe;
    3.2. U5b1b1 is virtually absent in the mtDNA pool of Siberian populations;
    3.3. the phylogeographic pattern of the phylogenetic precursor of U5b1b1, clade U5b1b, suggests that it, like hg V, has started to spread after the LGM from the western, rather than from eastern Europe. Both of them most probably reached northern Fennoscandia via eastern Europe.

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

Ammerman AJ, Cavalli-Sforza LL (1984) The Neolithic transition and the genetics of populations in Europe. Princeton University Press, Princeton

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, *et al* (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Andersson SG, Karlberg O, Canback B, Kurland CG (2003) On the origin of mitochondria: a genomics perspective. Philos Trans R Soc Lond B Biol Sci 358:165–77; discussion 177–179

Andrews P (1992) Evolution and environment in the Hominoidea. Nature 360:641–646

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

Armour JA, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, Kidd KK, *et al* (1996) Minisatellite diversity supports a recent African origin for modern humans. Nat Genet 13:154–160

Awadalla P, Eyre-Walker A, Smith JM (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. Science 286:2524–2525

Ayala FJ (1995) The myth of Eve: molecular biology and human origins. Science 270:1930–1936

Baasner A, Schafer C, Junge A, Madea B (1998) Polymorphic sites in human mitochondrial DNA control region sequences: population data and maternal inheritance. Forensic Science International 98:169–178

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen KH, *et al* (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152

Bamshad M, Kivisild T, Watkins WS, Dixon ME, Ricker CE, Rao BB, Naidu JM, *et al* (2001) Genetic evidence on the origins of Indian caste populations. Genome Research 11:994–1004

Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Barbujani G, Bertorelle G, Chikhi L (1998) Evidence for Paleolithic and Neolithic gene flow in Europe. Am J Hum Genet 62:488–492

Barker G (1985) Prehistoric Farming in Europe. Cambridge University Press, Cambridge

Beckman L, Beckman G, Nylander PO (1988) Gc subtypes in Finns, Swedes and Swedish Lapps. Hum Hered 38:18–21

Beckman G, Beckman L, Sikstrom C (1993) Serum complement (C3, BF, C4) types in Swedish Saamis. Hum Hered 43:362–365

Behar DM, Hammer MF, Garrigan D, Villems R, Bonne-Tamir B, Richards M, Gurwitz D, *et al* (2004) MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. Eur J Hum Genet 12:355–364

Belledi M, Poloni ES, Casalotti R, Conterio F, Mikerezi I, Tagliavini J, Excoffier L (2000) Maternal and paternal lineages in Albania and the genetic structure of Indo-European populations. Eur J Hum Genet 8:480–486

Bensasson D, Feldman MW, Petrov DA (2003) Rates of DNA duplication and mitochondrial DNA insertion in the human genome. J Mol Evol 57:343–354

Bermisheva M, Tambets K, Villems R, Khusnutdinova E (2002) Diversity of mitochondrial DNA haplotypes in ethnic populations of the Volga-Ural region of Russia. Mol Biol 36:990–1001

Bertranpetit J, Sala J, Calafell F, Underhill PA, Moral P, Comas D (1995) Human mitochondrial DNA variation and the origin of Basques. Ann Hum Genet 59:63–81

Blanc H, Chen KH, D'Amore MA, Wallace DC (1983) Amino acid change associated with the major polymorphic Hinc II site of Oriental and Caucasian mitochondrial DNAs. Am J Hum Genet 35:167–176

Bogenhagen DF (1999) Repair of mtDNA in vertebrates. Am J Hum Genet 64:1276–81.

Brega A, Gardella R, Semino O, Morpurgo G, Astaldi Ricotti GB, Wallace DC, Santachiara Benerecetti AS (1986) Genetic studies on the Tharu population of Nepal: restriction endonuclease polymorphisms of mitochondrial DNA. Am J Hum Genet 39:502–512

Bromham L, Eyre-Walker A, Smith NH, Smith JM (2002) Mitochondrial Steve: paternal inheritance of mitochondria in humans. TREE 18:2–4

Brown MD, Hosseini SH, Torroni A, Bandelt H-J, Allen JC, Schurr TG, Scozzari R, *et al* (1998) mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? Am J Hum Genet 63:1852–1861

Brown WM, George M, Jr., Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76:1967–1971

Brown WM (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc Natl Acad Sci U S A 77:3605–3609

Calafell F, Underhill P, Tolun A, Angelicheva D, Kalaydjieva L (1996) From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. Ann Hum Genet 60:35–49

Cali F, Le Roux MG, D'Anna R, Flugy A, De Leo G, Chiavetta V, Ayala GF, *et al* (2001) MtDNA control region and RFLP data for Sicily and France. Int J Legal Med 114:229–231

Cann RL, Brown WM, Wilson AC (1982) Evolution of human mitochondrial DNA: a preliminary report. Prog Clin Biol Res 103:157–165

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

Cann RL (2001) Genetic clues to dispersal in human populations: retracing the past from the present. Science 291:1742–1748

Caramelli D, Lalueza-Fox C, Vernesi C, Lari M, Casoli A, Mallegni F, Chiarelli B, *et al* (2003) Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. Proc Natl Acad Sci U S A 100:6593–6597

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The History and Geography of Human Genes. Princeton University Press, Princeton

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

Chen YS, Olckers A, Schurr TG, Kogelnik AM, Huoponen K, Wallace DC (2000) mtDNA variation in the South African Kung and Khwe-and their genetic relationships to other African populations. Am J Hum Genet 66:1362–1383

Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G (1998) Clines of nuclear DNA markers suggest a largely neolithic ancestry of the European gene pool. Proc Natl Acad Sci U S A 95:9053–9058

Chikhi L, Nichols RA, Barbujani G, Beaumont MA (2002) Y genetic data support the Neolithic demic diffusion model. Proc Natl Acad Sci U S A 99:11008–11013

Comas D, Plaza S, Wells RS, Yuldaseva N, Lao O, Calafell F, Bertranpetit J (2004) Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. Eur J Hum Genet 11:1–10

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 60:331–350

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 114:130–132

Danan C, Sternberg D, Van Steirteghem A, Cazeneuve C, Duquesnoy P, Besmond C, Goossens M, et al (1999) Evaluation of parental mitochondrial inheritance in neonates born after intracytoplasmic sperm injection. Am J Hum Genet 65:463–473

Darlu P, Tassy P (1987) Disputed African origin of human populations. Nature 329:111–112

Delghandi M, Utsi E, Krauss S (1998) Saami mitochondrial DNA reveals deep maternal lineage clusters. Hum Hered 48:108–114

Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen E, Cavalli-Sforza LL, Wallace DC (1981) Ethnic variation in Hpa 1 endonuclease cleavage patterns of human mitochondrial DNA. Proc Natl Acad Sci U S A 78:5768–5772

Derbeneva OA, Starikovskaia EB, Volod'ko NV, Wallace DC, Sukernik RI (2002*a*) [Mitochondrial DNA variation in Kets and Nganasans and the early peoples of Northern Eurasia]. Genetika 38:1554–1560

Derbeneva OA, Starikovskaya EB, Wallace DC, Sukernik RI (2002*b*) Traces of early Eurasians in the Mansi of northwest Siberia revealed by mitochondrial DNA analysis. Am J Hum Genet 70:1009–1014

Derenko MV, Dambueva IK, Malyarchuk BA, Dorzhu CM, Zakharov IA (1999) Structure and diversity of the mitochondrial gene pool in indigenous populations of Tuva and Buryatia based on restriction polymorphism data. Genetika 35:1706–1712

Derenko MV, Grzybowski T, Malyarchuk BA, Czarny J, Miscicka-Sliwka D, Zakharov IA (2001) The presence of mitochondrial haplogroup X in Altaians from South Siberia. Am J Hum Genet 69:237–241

Derenko MV, Grzybowski T, Malyarchuk BA, Dambueva IK, Denisova GA, Czarny J, Dorzhu CM, *et al* (2003) Diversity of Mitochondrial DNA Lineages in South Siberia. Ann Hum Genet 67:391–411

Di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci U S A 88:1597–1601

Diez-Sanchez C, Ruiz-Pesini E, Lapena AC, Montoya J, Perez-Martos A, Enriquez JA, Lopez-Perez MJ (2003) Mitochondrial DNA content of human spermatozoa. Biol Reprod 68:180–185

Dimo-Simonin N, Grange F, Taroni F, Brandt-Casadevall C, Mangin P (2000) Forensic evaluation of mtDNA in a population from south west Switzerland. Int J Legal Med 113:89–97

Dolukhanov PM (2000) "Prehistoric revolutions" and languages in Europe. In: Künnap A (ed) The roots of peoples and languages of Northern Eurasia: II and III. University of Tartu. Division of uralic Languages, Societas Historiae Fenno-Ugricae, Tartu, pp 71–84

Dubut V, Chollet L, Murail P, Cartault F, Beraud-Colomb E, Serre M, Mogentale-Profizi N (2004) mtDNA polymorphisms in five French groups: importance of regional sampling. Eur J Hum Genet 12:293–300

Dupuy BM, Olaisen B (1996) MtDNA sequences in the Norwegian Saami and main population. In: Carracedo A, Brinkmann B, Bär W (eds) Advances in forensic haemogenetics. 6. Springer-Verlag, Berlin, Heidelberg, New York, pp 23–25

Elson JL, Samuels DC, Turnbull DM, Chinnery PF (2001) Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. Am J Hum Genet 68:802–806

Elson JL, Turnbull DM, Howell N (2004) Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. Am J Hum Genet 74:229–238

Excoffier L, Langaney A (1989) Origin and differentiation of human mitochondrial DNA. Am J Hum Genet 44:73–85

Excoffier L (1990) Evolution of human mitochondrial DNA: evidence for departure from a pure neutral model of populations at equilibrium. J Mol Evol 30:125–39

Excoffier L, Yang Z (1999) Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. Mol Biol Evol 16:1357–68

Excoffier L (2002) Human demographic history: refining the recent African origin model. Curr Opin Genet Dev 12:675–682

Eyre-Walker A, Smith NH, Smith JM (1999) How clonal are human mitochondria? Proc R Soc Lond B Biol Sci 266:477–483

Fedorova SA, Bermisheva MA, Villems R, Maksimova NR, Khusnutdinova EK (2003) Analysis of mitochondrial DNA lineages in Yakuts. Mol Biol (Mosk) 37:643–653

Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. Journal of Mol Evol 17:368–376

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

Finnilä S, Majamaa K (2001) Phylogenetic analysis of mtDNA haplogroup TJ in a Finnish population. J Hum Genet 46:64–69

Fitch W (1977) On the problem of discovering the most parsimonious tree. Am. Nat. 111:1169–1175

Foley R (1998) The context of human genetic evolution. Genome Res 8:339–347

Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945

Forster P, Torroni A, Renfrew C, Rohl A (2001) Phylogenetic Star Contraction Applied to Asian and Papuan mtDNA Evolution. Mol Biol Evol 18:1864–1881

Forster P, Cali F, Rohl A, Metspalu E, D'Anna R, Mirisola M, De Leo G, et al (2002) Continental and subcontinental distributions of mtDNA control region types. Int J Legal Med 116:99–108
Forster P (2004) Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. Phil Trans R Soc Lond 359, 255–264
Francalacci P, Bertranpetit J, Calafell F, Underhill PA (1996) Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe. Am J Phys Anthropol 100:443–460
Fucharoen G, Fucharoen S, Horai S (2001) Mitochondrial DNA polymorphisms in Thailand. J Hum Genet 46:115–125
Gerber AS, Loggins R, Kumar S, Dowling TE (2001) Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? Annu Rev Genet 35:539–566
Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A 77:6715–6719
Gimbutas M (1970) Proto-Indo-European Culture: The Kurgan Culture during the fifth, fourth, and third millenia B.C. In: Cardona G, Hoenigswald HM, Senn A (eds) Indo-European and Indo-Europeans. University of Pennsylvania Press, Philadelphia, pp 155–195
Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L (1995) Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. Mol Biol Evol 12:334–345
Gray RD, Atkinson QD (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. Nature 426:435–439
Guglielmino CR, Piazza A, Menozzi P, Cavalli-Sforza LL (1990) Uralic genes in Europe. Am J Phys Anthropol 100:443–460
Haetta OM (1996) The Sami: an indigenous people of the Arctic. Davvi Girji, Karasjohka/Karasjoki, Vaasa
Hagelberg E (1997) Ancient and modern mitochondrial DNA sequences and the colonization of the Pacific. Electrophoresis 18:1529–1533
Hagelberg E, Goldman N, Lio P, Whelan S, Schiefenhovel W, Clegg JB, Bowden DK (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. Proc R Soc Lond B Biol Sci 266:485–492
Hagelberg E, Goldman N, Lio P, Whelan S, Schiefenhovel W, Clegg JB, Bowden DK (2000) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. Proc R Soc Lond B Biol Sci 267:1595–1596
Hammer MF (1995) A recent common ancestry for human Y chromosomes. Nature 378:376–378
Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, *et al* (1997) The geographic distribution of human Y chromosome variation. Genetics 145:787–805
Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, *et al* (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441
Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Curr Anthropol 34:483–496

Hasegawa M, Di Rienzo A, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. J Mol Evol 37:347–354
Hasegawa M, Cao Y, Yang Z (1998) Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. Mol Biol Evol 15:1499–1505
Hedges SB, Kumar S, Tamura K, Stoneking M (1992) Human origins and analysis of mitochondrial DNA sequences. Science 255:737–739
Helgason A, Sigurdadottir S, Gulcher J, Ward R, Stefanson K (2000) mtDNA and the origins of the Icelanders: deciphering signals of recent population history. Am J Hum Genet 66:999–1016
Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, Sykes B (2001) mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am J Hum Genet 68:723–737
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
Hines J (1991) Reviews. Nor Archael Rev 24:49–54
Hirata S, Hoshi K, Shoda T, Mabuchi T (2002) Spematozoon and mitochondrial DNA. Reprod Med Biol 1:41–47
Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, Paprotta A, Gerbitz KD (1997) Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. Hum Mol Genet 6:1835–1846
Horai S, Gojobori T, Matsunaga E (1984) Mitochondrial DNA polymorphism in Japanese. I. Analysis with restriction enzymes of six base pair recognition. Hum Genet 68:324–332
Horai S (1995) Evolution and the origins of man: clues from complete sequences of hominoid mitochondrial DNA. Southeast Asian J Trop Med Public Health 26:146–154
Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci U S A 92:532–536
Housley RA, Gamble CS, Street M, Pettitt P (1997) Radiocarbon evidence for the late glacial human recolonisation of northern Europe. Proc Prehist Soc 63:25–54
Howell N, Kubacka I, Mackey DA (1996) How rapidly does the human mitochondrial genome evolve? Am J Hum Genet 59:501–509
Howell N, Oostra RJ, Bolhuis PA, Spruijt L, Clarke LA, Mackey DA, Preston G, *et al* (2003) Sequence analysis of the mitochondrial genomes from dutch pedigrees with leber hereditary optic neuropathy. Am J Hum Genet 72:1460–1469
Ingman M, Kaessmann H, Pääbo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. Nature 408:708–713
Ingman M, Gyllensten U (2001) Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. J Hered 92:454–461
Jenuth JP, Peterson AC, Fu K, Shoubridge EA (1996) Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. Nat Genet 14:146–151

Jin L, Underhill PA, Doctor V, Davis RW, Shen P, Cavalli-Sforza LL, Oefner PJ (1999) Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations. Proc Natl Acad Sci U S A 96:3796–800

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL (1983) Radiation of human mitochondria DNA types analyzed by restriction endonuclease cleavage patterns. J Mol Evol 19:255–271

Jorde LB, Bamshad M (2000) Questioning evidence for recombination in human mitochondrial DNA. Science 288:1931

Jorde LB, Watkins WS, Bamshad M, Dixon ME, Ricker CE, Seielstad MT, Batzer MA (2000) The distribution of human genetic diversity: A comparison of mitochondrial, autosomal, and Y-chromosomal data. Am J Hum Genet 66:979–988

Kaessmann H, Heissig F, von Haeseler A, Pääbo S (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. Nat Genet 22:78–81

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217:624–626

Kittles RA, Bergen AW, Urbanek M, Virkkunen M, Linnoila M, Goldman D, Long JC (1999) Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. Am J Phys Anthropol 108:381–399

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

Kivisild T, Kaldma K, Metspalu M, Parik J, Papiha SS, Villems R (1999*b*) The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World. In: Papiha SS, Deka R, Chakraborty R (eds) Genomic Diversity. Kluwer Academic/Plenum Publishers, pp 135–152

Kivisild T, Papiha SS, Rootsi S, Parik J, Kaldma K, Reidla M, Laos S, *et al* (2000) An Indian Ancestry: a key for understanding human diversity in Europe and beyond. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and the population prehistory of Europe. McDonald Institute for Archaeological Research University of Cambridge, Cambridge, pp 267–279

Kivisild T, Villems R (2000) Questioning evidence for recombination in human mitochondrial DNA. Science 288:1931

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

Kivisild T, Rootsi S, Metspalu M, Metspalu E, Parik J, Kaldma K, Usanga E, *et al* (2003*a*) The Genetics of Language and Farming Spread in India. In: Bellwood P, Renfrew C (eds) Examining the farming/language dispersal hypothesis. The McDonald Institute for Archaeological Research, Cambridge, pp 215–222

Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, Metspalu E, *et al* (2003*b*) The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. Am J Hum Genet 72:313–332

Koehler CM, Lindberg GL, Brown DR, Beitz DC, Freeman AE, Mayfield JE, Myers AM (1991) Replacement of bovine mitochondrial DNA by a sequence variant within one generation. Genetics 129:247–255

Kong QP, Yao YG, Liu M, Shen SP, Chen C, Zhu CL, Palanichamy MG, *et al* (2003*a*) Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. Hum Genet 21:391–405

Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, Zhang YP (2003*b*) Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from Complete Sequences. Am J Hum Genet 73:671–676

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA sequences and the origin of modern humans. Cell 90:19–30

Krings M, Geisert H, Schmitz RW, Krainitzki H, Pääbo S (1999) DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. Proc Natl Acad Sci U S A 96:5581–5585

Kruger J, Vogel F (1989) The problem of our common mitochondrial mother. Hum Genet 82:308–312

Kumar S, Hedrick P, Dowling T, Stoneking M (2000) Questioning evidence for recombination in human mitochondrial DNA. Science 288:1931

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen L, Savontaus ML (1996) The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 58:1309–1322

Lahr M, Foley R (1994) Multiple dispersals and modern human origins. Evolutionary Anthropology 3:48–60

Lahr MM, Foley RA (1998) Towards a theory of modern human origins: geography, demography, and diversity in recent human evolution. Am J Phys Anthropol Suppl:137–176

Lightowlers RN, Chinnery PF, Turnbull DM, Howell N (1997) Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. Trends Genet 13:450–455

Lum JK, Cann RL (1998) mtDNA and language support a common origin of Micronesians and Polynesians in Island Southeast Asia. Am J Phys Anthropol 105:109–119

Lutz S, Weisser HJ, Heizmann J, Pollak S (1998) Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med 111:67–77

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, Forster P, Bendall KE, Watson E, Sykes B, Bandelt HJ (1997) mtDNA mutation rates--no need to panic. Am J Hum Genet 61:983–90

Macaulay V, Richards M, Sykes B (1999*a*) Mitochondrial DNA recombination — no need to panic. Proc R Soc Lond B Biol Sci 266:2037–2039

Macaulay VA, Richards MB, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, *et al* (1999*b*) 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

Malhi RS, Schultz BA, Smith DG (2001) Distribution of mitochondrial DNA lineages among Native American tribes of Northeastern North America. Hum Biol 73:17–55

Malhi RS, Smith DG (2002) Brief communication: Haplogroup X confirmed in prehistoric North America. Am J Phys Anthropol 119:84–86

Mallory JP (1989) In Search of the Indo-Europeans. Thames and Hudson, London

Malyarchuk BA, Derenko MV (2001) Mitochondrial DNA variability in Russians and Ukrainians: Implications to the origin of the Eastern Slavs. Ann Hum Genet 65:63–78

Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Wozniak M, Miscicka-Sliwka D (2002) Mitochondrial DNA variability in Poles and Russians. Ann Hum Genet 66:261–283

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

Manfredi G, Thyagarajan D, Papadopoulou LC, Pallotti F, Schon EA (1997) The fate of human sperm-derived mtDNA in somatic cells. Am J Hum Genet 61:953–60.

Marchington DR, Macaulay V, Hartshorne GM, Barlow D, Poulton J (1998) Evidence from human oocytes for a genetic bottleneck in an mtDNA disease. Am J Hum Genet 63:769–775

Margulis L (1975) Symbiotic theory of the origin of eukaryotic organelles; criteria for proof. Symp Soc Exp Biol 29:21–38

Mason PA, Lightowlers RN (2003) Why do mammalian mitochondria possess a mismatch repair activity? FEBS Lett 554:6–9

Meinilä M, Finnilä S, Majamaa K (2001) Evidence for mtDNA admixture between the Finns and the Saami. Hum Hered 52:160–170

Merriwether DA, Clark AG, Ballinger SW, Schurr TG, Soodyall H, Jenkins T, Sherry ST, *et al* (1991) The structure of human mitochondrial DNA variation. J Mol Evol 33:543–555

Metspalu E, Kivisild T, Kaldma K, Parik J, Reidla M, Tambets K, Villems R (1999) The Trans-Caucasus and the Expansion of the Caucasoid-Specific Human Mitochondrial DNA. In: Papiha S, Deka R, Chakraborty R (eds) Genomic Diversity: Application in Human Population Genetics. Kluwer Academic / Plenum Publishers, New York, pp 121–134

Meyer S, Weiss G, von Haeseler A (1999) Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. Genetics 152:1103–1110

Michaels GS, Hauswirth WW, Laipis PJ (1982) Mitochondrial DNA copy number in bovine oocytes and somatic cells. Dev Biol 94:246–251

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, *et al* (2003) Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci U S A 100:171–176

Moilanen JS, Finnilä S, Majamaa K (2003) Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J. Mol Biol Evol 20:2132–2142

Moilanen JS, Majamaa K (2003) Phylogenetic network and physicochemical properties of nonsynonymous mutations in the protein-coding genes of human mitochondrial DNA. Mol Biol Evol 20:1195–1210

Murray-McIntosh RP, Scrimshaw BJ, Hatfield PJ, Penny D (1998) Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequences. Proc Natl Acad Sci U S A 95:9047–9052

Nachman MW (1998) Deleterious mutations in animal mitochondrial DNA. Genetica 103:61–69

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, pp 145–163
Nei M, Takezaki N (1996) The root of the phylogenetic tree of human populations. Mol Biol Evol 13:170–177
Olivo PD, Van de Walle MJ, Laipis PJ, Hauswirth WW (1983) Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. Nature 306:400–402
Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M (2001) Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29:20–21
Opdal SH, Rognum TO, Vege A, Stave AK, Dupuy BM, Egeland T (1998) Increased number of substitutions in the D-loop of mitochondrial DNA in the sudden infant death syndrome. Acta Paediatr 87:1039–1044
Orekhov V, Poltoraus A, Zhivotovsky LA, Spitsyn V, Ivanov P, Yankovsky N (1999) Mitochondrial DNA sequence diversity in Russians. FEBS Letters 445:197–201
Ovchinnikov IV, Gotherstrom A, Romanova GP, Kharitonov VM, Liden K, Goodwin W (2000) Molecular analysis of Neanderthal DNA from the northern Caucasus. Nature 404:490–493
Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, *et al* (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nat Genet 15:363–368
Parson W, Parsons TJ, Scheithauer R, Holland MM (1998) Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: application of mtDNA sequence analysis to a forensic case. Int J Legal Med 111:124–132
Passarino G, Semino O, Bernini LF, Santachiara-Benerecetti AS (1996) Pre-Caucasoid and Caucasoid genetic features of the Indian population, revealed by mtDNA polymorphisms. Am J Hum Genet 59:927–934
Passarino G, Cavalleri GL, Lin AA, Cavalli-Sforza LL, Borresen-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
Penny D, Steel M, Waddell PJ, Hendy MD (1995) Improved analyses of human mtDNA sequences support a recent African origin for Homo sapiens. Mol Biol Evol 12:863–882
Pereira L, Macaulay V, Torroni A, Scozzari R, Prata MJ, Amorim A (2001) Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. Ann Hum Genet 65:439–458
Pesole G, Gissi C, De Chirico A, Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol 48:427–434
Pfeiffer H, Brinkmann B, Huhne J, Rolf B, Morris AA, Steighner R, Holland MM, et al (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112:291–298
Piazza A (1993) Who are the Europeans? Science 260:1767–1769
Piercy R, Sullivan KM, Benson N, Gill P (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Legal Med 106:85–90

Piganeau G, Eyre-Walker A (2004) A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. Heredity 92:282–288
Poikalainen V (2001) Palaeolithic art from the Danube to Lake Baikal. Folklore 18/19:7–16
Poulton J, Macaulay V, Marchington DR (1998) Mitochondrial genetics '98 is the bottleneck cracked? Am J Hum Genet 62:752–757
Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, Paabo S (1994) Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. Biol Chem Hoppe Seyler 375:837–840
Puzyrev VP, Stepanov VA, Golubenko MV, Puzyrev KV, Maximova NR, Kharkov VN, Spiridonova MG, *et al* (2003) MtDNA and Y-Chromosome Lineages in the Yakut Population. Genetika 39:975–981
Pääbo S (1996) Mutational hot spots in the mitochondrial microcosm. Am J Hum Genet 59:493–496
Quintana-Murci L, Semino O, Bandelt H-J, Passarino G, McElreavey K, Santachiara-Benerecetti AS (1999) Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. Nat Genet 23:437–441
Quintana-Murci L, Chaix R, Wells S, Behar D, Sayar H, Scozzari R, Rengo C, *et al* (2004) Where West meets East: The complex mtDNA landscape of the Southwest and Central Asian corridor. Am J Hum Genet 74:827–845
Quintans B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A (2004) Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. Forensic Sci Int 140:251–257
Rando JC, Pinto F, Gonzalez AM, Hernandez M, Larruga JM, Cabrera VM, Bandelt HJ (1998) Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, near-eastern, and sub-Saharan populations. Ann Hum Genet 62:531–550
Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro AS, Stoneking M (1995) Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol 12:604–615
Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk H, Parik J, *et al* (2003) Origin and Diffusion of mtDNA Haplogroup X. Am J Hum Genet. 73:1178–90
Relethford JH, Jorde LB (1999) Genetic evidence for larger African population size during recent human evolution. Am J Phys Anthropol 108:251–260
Renfrew C (1987) Archaeology and Language: the Puzzle of Indo-European Origins. Jonathan Cape, London
Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, *et al* (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203
Richards MB, Macaulay VA, Bandelt H-J, Sykes BC (1998) Phylogeography of mitochondrial DNA in western Europe. Ann Hum Genet 62:241–260
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 M, Macaulay V, Torroni A, Bandelt HJ (2002) In search of geographical patterns in European mitochondrial DNA. Am J Hum Genet 71:1168–1174

Richards M, Rengo C, Cruciani F, Gratrix F, Wilson J, Scozzari R, Macaulay V, *et al* (2003) Extensive female-mediated gene flow from sub-Saharan Africa into near eastern Arab populations. Am J Hum Genet. 72:1058–1064

Richards M (2003) The Neolithic invasion of Europe. Annu Rev Anthropol 32:135–162

Ritte U, Neufeld E, Broit M, Shavit D, Motro U (1993) The differences among Jewish communities--maternal and paternal contributions. J Mol Evol 37:435–440

Rosa A, Brehm A, Kivisild T, Metspalu E, Villems R (in press) MtDNA profile of West Africa Guineans: towards a better understanding of the Senegambia region. Ann Hum Genet

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, *et al* (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543

Rousselet F, Mangin P (1998) Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. Int J Legal Med 111:292–298

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303:223–226

Ruvolo M, Zehr S, von Dornum M, Pan D, Chang B, Lin J (1993) Mitochondrial COII sequences and modern human origins [published erratum appears in Mol Biol Evol 1994 May;11(3):552]. Mol Biol Evol 10:1115–1135

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A (1999) Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238:195–209

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

Saillard J, Evseva I, Tranebjaerg L, Norby S (2000*b*) Mitochondrial DNA diversity among Nenets. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and and the population prehistory of Europe. McDonald Institute for Archaeological Research Monograph Series, Cambridge University, Cambridge, pp 255–258

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Saitou N, Omoto K (1987) Time and place of human origins from mt DNA data. Nature 327:288

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus ML, Aula P, *et al* (1995) Genes and languages in Europe: an analysis of mitochondrial lineages. Genome Res 5:42–52

Sajantila A, Pääbo S (1995) Language replacement in Scandinavia. Nat genet 11:359–360

Sajantila A, Salem AH, Savolainen P, Bauer K, Gierig C, Pääbo S (1996) Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci U S A 93:12035–9

Salas A, Comas D, Lareu MV, Bertranpetit J, Carracedo A (1998) mtDNA analysis of the Galician population: a genetic edge of European variation. Eur J Hum Genet 6:365–375

Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sanchez-Diz P, Macaulay V, *et al* (2002) The making of the African mtDNA landscape. Am J Hum Genet 71:1082–1111

Salas A, Richards M, Lareu MV, Scozzari R, Coppa A, Torroni A, Macaulay V, *et al* (2004) The African diaspora: mitochondrial DNA and the Atlantic slave trade. Am J Hum Genet 74:454–465

Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC (1999) Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. Am J Phys Anthropol 108:1–39

Schurr TG, Wallace DC (2002) Mitochondrial DNA diversity in Southeast Asian populations. Hum Biol 74:431–452

Schwartz M, Vissing J (2002) Paternal inheritance of mitochondrial DNA. N Engl J Med 347:576–580

Scozzari R, Cruciani F, Pangrazio A, Santolamazza P, Vona G, Moral P, Latini V, *et al* (2001) Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. Hum Immunol 62:871–884

Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, *et al* (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290:1155–1159

Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, Mennecier P, Hofreiter M, *et al* (2004) No Evidence of Neandertal mtDNA Contribution to Early Modern Humans. PLoS Biol 2:E57

Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA (1999) Distribution of mtDNA haplogroup X among Native North Americans. Am J Phys Anthropol 110:271–284

St John JC, Lloyd R, El Shourbagy S (2004) The potential risks of abnormal transmission of mtDNA through assisted reproductive technologies. Reprod Biomed Online 8:34–44

Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC (1998) mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of Ancient Beringia and the peopling of the New World. Am J Hum Genet 63:1473–1491

Stevanovitch A, Gilles A, Bouzaid E, Kefi R, Paris F, Gayraud RP, Spadoni JL, *et al* (2004) Mitochondrial DNA Sequence Diversity in a Sedentary Population from Egypt. Ann Hum Genet 68:23–39

Stoneking M, Jorde LB, Bhatia K, Wilson AC (1990) Geographic variation in human mitochondrial DNA from Papua New Guinea. Genetics 124:717–733

Stoneking M, Sherry ST, Redd AJ, Vigilant L (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. Phil. Trans. Royal Soc. 337:167–175

Stringer CB, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. Science 239:1263–1268

Stringer C (2000) Coasting out of Africa. Nature 405:24–25, 27

Stringer C, Davies W (2001) Archaeology. Those elusive Neanderthals. Nature 413:791–792

Stringer C (2003) Human evolution: Out of Ethiopia. Nature 423:692–3, 695.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G (1999) Ubiquitin tag for sperm mitochondria. Nature 402:371–372

Sutovsky P, Van Leyen K, McCauley T, Day BN, Sutovsky M (2004) Degradation of paternal mitochondria after fertilization: implications for heteroplasmy, assisted reproductive technologies and mtDNA inheritance. Reprod Biomed Online 8:24–33

Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. Am J Hum Genet 57:1463–1475

Tambets K, Kivisild T, Metspalu E, Parik J, Kaldma K, Laos S, Tolk H-V, *et al* (2000) The topology of the maternal lineages of the Anatolian and Trans-Caucasus populations and the peopling of the Europe: some preliminary considerations. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and the population prehistory of Europe: McDonald Institute for Archaeological Research Monograph Series. Cambridge University Press, Cambridge, pp 219–235

Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogvali EL, Tolk HV, *et al* (2004) The Western and Eastern Roots of the Saami--the Story of Genetic "Outliers" Told by Mitochondrial DNA and Y Chromosomes. Am J Hum Genet 74:661–682

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

Templeton A (2002) Out of Africa again and again. Nature 416:45–51

Thorburn DR, Dahl HH (2001) Mitochondrial disorders: genetics, counseling, prenatal diagnosis and reproductive options. Am J Med Genet 106:102–114

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonne-Tamir B, *et al* (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271:1380–1387

Tishkoff SA, Pakstis AJ, Stoneking M, Kidd JR, Destro-Bisol G, Sanjantila A, Lu RB, *et al* (2000) Short tandem-repeat polymorphism/alu haplotype variation at the PLAT locus: implications for modern human origins. Am J Hum Genet 67:901–925

Torroni A, Schurr TG, Yang CC, Szathmary EJ, Williams RC, Schanfield MS, Troup GA, *et al* (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 130:153–162

Torroni A, Sukernik RI, Schurr TG, Starikorskaya YB, Cabell MF, Crawford MH, Comuzzie AG, *et al* (1993*a*) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. Am J Hum Genet 53:591–608

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, *et al* (1993*b*) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994*a*) Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci U S A 91:1158-1162Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC (1994*b*) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93:189–199

Torroni A, Chen YS, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, *et al* (1994*c*) mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54:303–318

Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC (1994*d*) 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, 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 144:1835–1850

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 62:1137–1152

Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A, Calderon FL, *et al* (2001) Do the four clades of the mtDNA haplogroup L2 evolve at different rates? Am J Hum Genet 69:1348–1356

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, *et al* (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res 7:996–1005

Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, *et al* (2000) Y chromosome sequence variation and the history of human populations. Nat Genet 26:358–361

van Andel TH (2000) Where received wisdom fails: the Mid-Palaeolithic and early Neolithic climates. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and the population prehistory of Europe. McDonald Institute for Archaeological Research Monograph Series, Cambridge University, Cambridge, pp 31–39

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci U S A 86:9350–9354

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

Villems R, Adojaan M, Kivisild T, Metspalu E, Parik J, Pielberg G, Rootsi S *et al.* (1998) Reconstruction of maternal lineages of Finno-Ugric speaking people and some remarks on their paternal inheritance. In: Wiik K, Julku K (eds) The roots of people and languages of Northern Eurasia I. Societas Historiae fenno-Ugricae, Turku, pp 180–200

Wakeley J (1994) Substitution-rate variation among sites and the estimation of transition bias. Mol Biol Evol 11:436–442

Wallace DC, Brown MD, Lott MT (1999) Mitochondrial DNA variation in human evolution and disease. Gene 238:211–230

Ward RH, Frazier B, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci U S A 88:8720–8724

Watson E, Forster P, Richards M, Bandelt HJ (1997) Mitochondrial footprints of human expansions in Africa. Am J Hum Genet 61:691–704

Weidenreich F (1943) The ‘‘Neanderthal man’’ and the ancestors of ‘‘Homo sapiens’’. Am Anthropologist 45:39–48

Whittle A (1996) Europe in the Neolithic. Cambridge University Press, Cambridge

Wiik K (2000) European *Lingua Francas*. In: Künnap A (ed) The Roots of Peoples and Languages of Northern Eurasia II and III. University of Tartu and Societas Historiae Fenno-Ugricae, Tartu, pp 202–236

Wilson A, Cann R, Carr S, George M, Gyllensten U, Helm-Bychowski K, Higuchi R, *et al* (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biological Journal of the Linnean Society 26:375–400

Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB (2001) Genetic evidence for different male and female roles during cultural transitions in the British Isles. Proc Natl Acad Sci U S A 98:5078–5083

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

Wolpoff M, Caspari R (1997) Race and Human Evolution. Simon & Schuster, New York

Yao Y, Nie L, Harpending H, Fu Y, Yuan Z, Zhang Y (2002*a*) Genetic relationship of Chinese ethnic populations revealed by mtDNA sequence diversity. Am J Phys Anthropol. 118:63–76

Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002*b*) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. Am J Hum Genet 70:635–651

Yu MS, Qiu XF, Xue JG, Liu ZD, Tan JZ, Li HJ, Liu DX, *et al* (1988) Mitochondrial DNA polymorphism in Chinese. Sci Sin 31:860–872

Zerjal T, Beckman L, Beckman G, Mikelsaar AV, Krumina A, Kucinskas V, Hurles ME, et al (2001) Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in Northern European populations. Mol Biol Evol 18:1077–1087

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, et al (1997) Genetic relationships of Asians and Northern Europeans, revealed by Y- chromosomal DNA analysis. Am J Hum Genet 60:1174–1183

Zischler H, Geisert H, von Haeseler A, Pääbo S (1995) A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. Nature 378:489–492

Zischler H, Geisert H, Castresana J (1998) A hominoid-specific nuclear insertion of the mitochondrial D-loop: implications for reconstructing ancestral mitochondrial sequences. Mol Biol Evol 15:463–469

Zvelebil M (1986) Mesolithic prelude and Neolithic revolution. In: Zvelebil M (ed) Hunters in transition: Mesolithic societies of temperate Eurasia and their transition to farming. Cambridge University Press, Cambridge, pp 5–15

Zvelebil M (2000) The social context of the agricultural transition in Europe. In: Renfrew C, Boyle K (eds) Archaeogenetics: DNA and and the population prehistory of Europe. McDonald Institute for Archaeological Research Monograph Series, Cambridge University, Cambridge, pp 57–79

KOKKUVÕTE

Eukarüootsetes rakkudes leidub pärilikkusainet lisaks rakutuumale veel ka organellide — mitokondrite ja kloroplastide — maatriksis. Imetajate mitokondri umbes 16 600 aluspaariline rõngasjas genoom pärandub põlvkonnast põlvkonda vaid emaliinis ja muteerub tuuma omast umbes suurusjärgu võrra kiiremini — see lihtsustab evolutsioonilises ajaskaalas suhteliselt hiljuti lahknenud geeniliinide eristamist ja sellise "noore" liigi nagu *Homo sapiens* geneetilise mitmekesisuse uurimist. Eeldatakse, et mitokondriaalse DNA (mtDNA) varieeruvuse ja erinevate geeniliinide geograafilise leviku kujundamisel on põhiroll olnud juhuslikul geneetilisel triivil ja erinevaid mtDNA variante kandvate inimrühmade demograafilisel ajalool. Tänu sellele on mtDNA, eriti viimase kümnekonna aasta jooksul, leidnud laialdast rakendust kaasaegse inimese kujunemise, tema eelajalooliste migratsioonide ja geneetilise varieeruvuse uurimisel.

Töö kirjanduse ülevaate osas kirjeldatakse mtDNA kui geneetilise markersüsteemi omadusi ja põhilisi uurimismeetodeid ning käsitletakse mtDNA globaalsest varieeruvust. Lühidalt puudutatakse ka erinevaid teooriaid kaasaegse inimese päritolu ja leviku kohta ning käsitletakse tähtsamaid mtDNA-ga seotud uurimistöid, mis vaekausi hiljutise Aafrikast väljarände hüpoteesi kasuks on kallutanud. Selle hüpoteesi kohaselt pärineb kõikide kaasaegsete inimeste mtDNA umbes 200 000 aastat tagasi Aafrikas elanud eellase mtDNAst. Mõned selle nn. "mitokondriaalse Eeva" järeltulijad rändasid umbes 60 000 aastat tagasi Aafrikast välja ja koloniseerisid ajapikku kõik kontinendid peale Antarktika. Sellega kaasnes teistel kontinentidel elanud perekond *Homo* vormide väljasuremine, kuid nende kahe protsessi vaheline võimalik vahetu põhjuslik-tagajärgne seos nõuab veel kinnitust. Praegu kasutusel olevas mtDNA tüüpide nomenklatuuris jagatakse mtDNA erinevad variandid kindlate polümorfismide alusel haplogruppideks, mis tänapäevases, mtDNA varieeruvust käsitlevas kirjanduses on ühtlasi iseloomustatavad kui monofüleetilisted klaadid. Haplogrupp hõlmab erinevaid haplotüüpe, mis kõik omavad haplogruppi defineerivat ühte või enamat polümorfismi, kuid erinevad haplogrupisiseselt tekkinud mutatsioonide osas. MtDNA analüüs erinevates maailma populatsioonides on näidanud, et haplogruppide geograafiline levik on suuremal või vähemal määral kontinendispetsiifiline. Käesolevas doktoritöös on põhirõhk asetatud Euroopa populatsioonide mtDNA varieeruvuse ja selle kujunemise põhjuste uurimisele. MtDNA fülogeneesipuu detailide uurimisel on oluline tunda puu üldist topoloogiat — erinevate harude ajalise ja ruumilise hargnemise seaduspärasusi. Seetõttu ei saa puu tipmiste harude, mille hulka kuuluvad ka Euroopa mtDNA variandid, kirjeldamisel mööda basaalsete harude tundmisest — lisaks Euroopa mtDNA mitmekesisuse tutvustamisele antakse kirjanduse ülevaates lühikokkuvõte mtDNA globaalsest varieeruvusest.

Ehkki Euroopas on emaliinide varieeruvust teiste Vana ja Uue Maailma kontinentidega võrreldes oluliselt rohkem uuritud, on siinselgi geneetilisel

kaardil "valgeid laike", mille puhul olemasolevaid andmeid interpreteerida ja üldisesse konteksti sobitada ei osata. Käesolevas töös on püütud saamide ema- (ja isa-) liinide uurimisega valgustada üht sellist erandlikku juhtu (artiklid I ja V). Töö teiseks lähenemisnurgaks on Euroopas ja kaugemalgi levinud emaliinide, mis kuuluvad haplogruppidesse U, V ja X, fülogeograafiliste mustrite ja nende kujunemise võimalike põhjuste analüüs (artiklid I–IV).

Töö autori poolt läbi viidud eksperimentaalne osa hõlmab 73 Rootsi saami, 557 eestlase, 398 rootslase, 299 lätlase, 45 leedulase, 199 albaanlase, 102 mordvalase, 136 komi, 101 udmurdi, 55 tšuvaši, 40 baškiiri, 40 tatarlase, 116 ungarlase, 40 handi, 38 mansi ja 388 türklase mtDNA analüüsi. MtDNA varieeruvuse väljaselgitamiseks amplifitseeriti iga proovi mtDNA esimene hüpervarieeruv regioon (HVS-I) ja sekveneeriti selle 360 aluspaari vahemikus 16 023–16 383. Uuritavate mtDNAde haplogrupiline kuuluvus tehti kindlaks restriktsioonanalüüsi või polümorfsete piirkondade sekveneerimisega. Populatsioonide mtDNA profiilide võrdlemiseks kasutati haplogruppide sagedustel rajanevat põhikomponentanalüüsi. Erinevate haplotüüpide fülogeneetiliste suhete analüüsiks kasutati mediaanvõrgustiku meetodit. Haplogruppide ekspansiooniaja alguse hindamine toimus järjestuste varieeruvuse alusel: selleks arvutati uuritavasse haplogruppi kuuluvate haplotüüpide keskmine transitsiooniline kaugus (ρ) arvatavast fülogeneetilisest eellashaplotüübist. Saadud väärtus teisendati ajaliseks hinnanguks HVS-I piirkonna (vahemikus 16 090–16 365) transitsioonide keskmise fikseerumiskiiruse (1 transitsioon 20 180 aasta kohta) alusel.

Töö põhilised tulemused on:

1. Euroopas laialt levinud mtDNA haplogruppide U4, U5 ja X fülogeograafiline analüüs näitas et "klassikaliste" emaliiniklastrite klassifitseerimine hästidefineeritud monofüleetilisteks alamhulkadeks on nende erinevate harude fülogeograafilise levikumustri uurimisel esmatähtis ja võimaldab ühtlasi teha järeldusi emaliinide olemasoleva ruumilise jaotuse kujunemise aja ja põhjuste kohta. MtDNA haplogruppide V, U, X alamklastrite ekspansiooniaja arvutused lubavad oletada, et nende laialdane levik sai alguse Põhja-Euroopa rekoloniseerimisega kaasnenud populatsioonide ekspansiooniga erinevatest Euroopa refuugiumidest;

2. Euroopa geneetilisel maastikul üheks ekstreemseks erandiks peetava populatsiooni — saamide — ema- (ja isa)liinide analüüs näitas, et populatsioonide geneetiliste distantside arvutamine fülogeneetilist aspekti arvesse võtmata võib viia väära interpretatsioonini populatsiooni võimaliku päritolu kohta ja ei võimalda näilisele kvantitatiivsele lähenemisele vaatamata teha vahet geneetiliste "autsaiderite" ja väikeses populatsioonis aset leidnud juhuslikust geenitriivist tingitud eripärade vahel. Meie poolt saadud tulemused näitasid selgesti, et saamide mtDNA ja Y-kromosoomi

markerite osas on saamide geneetilise eripära põhjuseks mitte teistest eurooplastest erinev päritolu, vaid geenitriivist põhjustatud iseäralik alleelisageduste muster;

3. Valdav enamus (umbes 95%) saamide emaliinidest kuulub Euroopas laialt levinud mtDNA variantide, eelkõige haplogruppide U5 ja V hulka. Euraasia idaosale iseloomulike emaliinide (hg M derivaadid) osakaal ei ületa viit protsenti saamidel leitud mtDNA variantide koguhulgast. Siiani üksnes saamidele iseloomulikuks peetud mtDNA haplogrupi U5 variantide fülogeograafiline analüüs näitas, et:

   3.1. saamide seas väga sage haplogrupi U5 alamharu (50% emaliinide koguhulgast), mis siin defineeriti U5 alamhulgaks U5b1b1, on, ehkki madala sagedusega (kuni 7%), esindatud paljudes erinevates Euroopa populatsioonides;

   3.2. U5b1b1 puudub enamiku Siberi populatsioonide geenitiigist;

   3.3. U5b1b1 eellasklastri U5b1b geograafiline levikumuster lubab oletada, et haplogrupi U5b1b, nagu ka haplogrupi V, ekspansioon sai alguse viimase jääaja maksimumi ajal Euroopa edelaosas asunud refuugiumist. Samas on tõenäone, et haplogruppide V ja U5 tee Skandinaavia põhjaossa on kulgenud Ida-Euroopa kaudu, osana protsessist, mis algas Euroopa põhjapiirkondade taasasustamisega viimase jääaja järgselt.

ACKNOWLEDGEMENTS

This study was carried out at the Department of Evolutionary Biology, in the Institute of Molecular and Cell Biology of the University of Tartu and in the Estonian Biocentre.

My deepest gratitude goes to Prof. Richard Villems, who has been the main driving force formulating the most interesting ideas and questions to investigate. His help has been invaluable for the completion of this work.

I am also grateful to all our co-authors and collaborators who have collected the priceless DNA samples from very different parts of the world. I would especially like to thank Prof. Lars Beckman, for the DNA samples of the Swedish Saami, Dr. Elza Khusnutdinova, for the samples of Volga-Uralic region populations and Dr. Ludmila Osipova and Dr. Mikhail Voevoda, for the samples of Siberian populations.

I would like to express my sincere thanks to all the staff and fellow students from the Department of Evolutionary Biology, for creating an inspiring and pleasant atmosphere in the laboratory, for interesting discussions and valuable criticism during the writing of this dissertation.

I place great value in the friendly environment and people of our Institute and Estonian Biocentre.

I appreciate the excellent technical assistance of Ille Hilpus and Jaan Lind.

And then, I owe a lot to my dearest ones, some of who — both elder and younger — share the identical variant of mtDNA with me. They have encouraged and supported me constantly during my studies and behaved extraordinarily well during all those years when we have together studied the wonders of life.

PUBLICATIONS

Tambets, K., Rootsi, S., Kivisild, T., Help, H., Serk, P., Loogväli, E.-L., Tolk H.-V., 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, J.-P., Barać, L., Peričić, M., Rudan, P., Terzić, R., Mikerezi, I., Krumina, A., Baumanis, V., Koziel, S., Rickards, O., De Stefano, GF., Anagnou, N., Pappa, K.I., Michalodimitrakis, E., Ferák, V., Füredi, S., Komel, R., Beckman, L., Villems, R. (2004) The western and eastern roots of the Saami — the story of genetic "outliers" told by mtDNA and Y-chromosome. *American Journal of Human Genetics* 74, 661–682.

Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H.-V., **Tambets, K.**, Forster, P., Karger, B., Francalacci, P., Janicijevic, B., Rudan, P., Rickards, O., Savontaus, M.-L., Huoponen, K., Laitinen, V., Koivumäki, S., Sykes, B., Novelleto, A., Moral, P., Sellitto, D., Santachiara-Benerecetti, A.S., Semino, O., Scozzari, R. (2001) A Signal, from human mtDNA, of postglacial recolonization in Europe. *American Journal of Human Genetics* 69, 844–852.

Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., **Tambets, K.,** Tolk, H.-V., Parik, J., Loogväli, E.-L., Derenko, M., Malyarchuk, B., Bermisheva, M., Zhadanov, S., Pennarun, E., Gubina, M., Golubenko, M., Damba, L., Feodorova, S., Gusar, V., Grechanina, E., Mikerezi, I., Moisan, J.-P., Chaventre A., Khusnutdinova, E., Osipova, L., Stepanov, V., Voevoda, M., Achilli, A., Rengo C., Rickards, O., De Stefano, G. F., Papiha, S., Beckman, L., Janicijevic, B., Rudan P., Anagnou N., Michalodimitrakis, E., Koziel, S., Usanga, E., Geberhiwot, T., Herrnstadt, C., Howell, N., Torroni, A., Villems, R. (2003). Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics* 73, 1178–1190.

Tambets, K., Tolk, H.-V., Kivisild, T., Metspalu, E., Parik, J., Reidla, M., Voevoda, M., Damba, L., Bermisheva, M., Khusnutdinova, E., Golubenko, M., Stepanov, V., Puzyrev, V., Usanga, E., Rudan, P., Beckman, L., Villems, R. (2003) Complex signals for population expansions in Europe and beyond. — In: *Examining the Farming/Language Dispersal Hypothesis* (Cambridge University Press), 449–457.

Tambets, K., Rootsi, S., Kivisild, T., Villems, R. (2001) The concepts of Richard Indreko about the origin of the Finno-Ugric speakers and the population genetics of the extant North-East European populations. *TRAMES*, 5 (55/50), 1, 59–74.

CURRICULUM VITAE

Kristiina Tambets

Citizenship:	Estonian
Date and place of birth:	April 20th, 1970, Tallinn, Estonia
Family status:	married, two children
Address:	Institute of Molecular and Cell Biology, University of Tartu and Estonian Biocentre, 23 Riia Street, 51010, Tartu, Estonia
Telephone:	+372 7 375 053
Fax:	+372 7 420 286
E-mail:	ktambets@ebc.ee

Education and professional employment

1977–1985	Tallinn Secondary School no. 10
1985–1988	Tallinn Secondary School no. 20
1988–1992	University of Tartu, Faculty of Biology and Geography, raduated as geneticist
1992–1996	University of Tartu, Institute of Molecular and Cell Biology, Department of Evolutionary Biology, MSc student in evolutionary biology, MSc thesis "MtDNA control region polymorphism in Estonian population"
1996–2003	PhD student in the Department of Evolutionary Biology, research subject: human mtDNA variation
From 2003	Estonian Biocentre, scientist

Scientific work

Starting from 1992 I have been studying and working in the Department of Evolutionary Biology of the University of Tartu under the supervision of Prof. Richard Villems. I have been involved in projects, associated with the analysis of human mtDNA variation in different Eurasian populations. The results of studies are published in the following articles:

1. **Tambets, K.,** Rootsi, S., Kivisild, T., Help, H., Serk, P., Loogväli, E.-L., Tolk H.-V., 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, J.-P., Barać, L., Peričić, M., Rudan, P., Terzić, R., Mikerezi, I., Krumina, A., Baumanis, V., Koziel, S., Rickards, O., De Stefano, GF., Anagnou, N., Pappa, K.I., Michalodimitrakis, E., Ferák, V., Füredi, S., Komel, R., Beckman, L., Villems, R. (2004) The western and eastern roots of the Saami — the story of genetic "outliers" told by mtDNA and Y-chromosome. *American Journal of Human Genetics* 74, 661–682.

2. Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., **Tambets, K.,** Tolk, H.-V., Parik, J., Loogväli, E.-L., Derenko, M., Malyarchuk, B., Bermisheva, M., Zhadanov, S., Pennarun, E., Gubina, M., Golubenko, M., Damba, L., Feodorova, S., Gusar, V., Grechanina, E., Mikerezi, I., Moisan, J.-P., Chaventre A., Khusnutdinova, E., Osipova, L., Stepanov, V., Voevoda, M., Achilli, A., Rengo C., Rickards, O., De Stefano, G. F., Papiha, S., Beckman, L., Janicijevic, B., Rudan P., Anagnou N., Michalodimitrakis, E., Koziel, S., Usanga, E., Geberhiwot, T., Herrnstadt, C., Howell, N., Torroni, A., Villems, R. (2003). Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics* 73, 1178–1190.

3. **Tambets, K.**, Tolk, H.-V., Kivisild, T., Metspalu, E., Parik, J., Reidla, M., Vevoda, M., Damba, L., Bermisheva, M., Khusnutdinova, E., Golubenko, M., Stepanov, V., Puzyrev, V., Usanga, E., Rudan, P., Beckman, L., Villems, R. (2003) Complex signals for population expansions in Europe and beyond. — In: *Examining the Farming/Language Dispersal Hypothesis* (Cambridge University Press), 449–458.

4. Villems, R., Rootsi, S., **Tambets, K.**, Adojaan, M., Orekhov, V., Khusnutdinova, E., Yankovsky, N. (2002) Archaeogenetics of Finno-Ugric-speaking populations. In *The Roots of Peoples and Languages of the Northern Eurasia*. (Societas Historiae Fenno-Ugricae, Gummerus Kirjapaino Oy), 271–284.

5. Bermisheva, M., **Tambets, K.**, Villems, R., Khusnutdinova, E. (2002) Diversity of mitochondrial DNA haplogroups in ethnic populations of the Volga-Ural region. *Moleculyarnaya Biologiya* 36, 990–1001.

6. Torroni, A., Bandelt, H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villems, R., Kivisild, T., Metspalu, E., Parik, J., Tolk, H.-V., **Tambets, K.**, Forster, P., Karger, B., Francalacci, P., Janicijevic, B., Rudan, P., Rickards, O., Savontaus, M.-L., Huoponen, K., Laitinen, V., Koivumäki, S., Sykes, B., Novelleto, A., Moral, P., Sellitto, D., Santachiara-Benerecetti, A.S., Semino, O., Scozzari, R. (2001) A Signal, from human mtDNA, of postglacial recolonization in Europe. *American Journal of Human Genetics* 69, 844–852.

7. **Tambets, K.**, Rootsi, S., Kivisild, T., Villems, R. (2001) The concepts of Richard Indreko about the origin of the Finno-Ugric speakers and the population genetics of the extant North-East European populations. *TRAMES*, 5 (55/50), 1, 59–74.

8. **Tambets, K.**, Kivisild, T., Metspalu, E., Parik, J., Kaldma, K., Laos, S., Tolk, H.-V., Gölge, M., Demirtas, H., Geberhiwot, T., Papiha, S.S., De Stefano, G.F., Villems, R. (2000) The Topology of the Maternal Lineages of the Anatolian and Trans-Caucasus Populations and the Peopling of Europe: Some Preliminary Considerations. In *Archaeogenetics: DNA and the population prehistory in Europe*. (Cambridge University Press), 219–235.

9. Rootsi, S., Kivisild, T., **Tambets, K.**, Adojaan, M., Parik, J., Reidla, M., Metspalu, E., Laos, S., Tolk, H.-V., Villems, R. (2000) On the Phylo-geographic Context of Sex-Specific Genetic Markers of Finno-Ugric Populations. In *The Roots of Peoples and languages of Northern Eurasia II and III.* (University of Tartu, Division of Uralic Languages/Societas Historiae Fenno-Ugricae), 148–164.

10. Metspalu E., Kivisild T., Kaldma K., Parik J., Reidla M., **Tambets K.**, Villems R. (1999) The Trans-Caucasus and the expansion of the Caucasoid-specific human mitochondrial DNA. In *Genome Diversity: Applications to Human Population Genetics.* (Kluwer Academic/Plenum Publishers), 241–261.

11. Villems, R., Adojaan, M., Kivisild, T., Parik, J., Pielberg, G., Rootsi, S., **Tambets, K.**, Tolk, H.-V. (1998) Reconstruction of maternal lineages of Finno-Ugric speaking people and some remarks on their paternal inheritance. In *The Roots of Peoples and languages of Northern Eurasia I.* (Societas Historiae Fenno-Ugricae, Gummerus Kirjapaino Oy), 180–200.

CURRICULUM VITAE

Kristiina Tambets

Kodakondsus:	Eesti
Sünniaeg ja -koht:	20 aprill, 1970, Tallinn, Eesti
Perekonnaseis:	abielus, peres kaks last
Aadress:	Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituut, Eesti Biokeskus, Riia mnt.23, 51010, Tartu, Eesti
Telefon:	+372 7 375 053
Faks:	+372 7 420 286
E-post:	ktambets@ebc.ee

Haridus ja erialane teenistuskäik

1977–1985	Tallinna 10. Keskkool
1985–1988	Tallinna 20. Keskkool
1988–1992	Tartu Ülikool, Bioloogia- ja geograafiateaduskond, Diplomeeritud bioloog geneetika erialal
1992–1996	Tartu Ülikool, Molekulaar- ja Rakubioloogia Instituut, molekulaarbioloogia, seejärel evolutsioonilise bioloogia õppetool, Magistritöö teemal "MtDNA kontrollregiooni polümorfism eestlaste populatsioonis"
1996–2003	Evolutsioonilise bioloogia õppetooli doktorant Uurimistöö teema: inimese mtDNA polümorfism
Alates nov. 2003	Eesti Biokeskuse teadur

Teadustegevus

Alates 1992 a. olen õppinud ja töötanud molekulaarbioloogia ning seejärel evolutsioonilise bioloogia õppetooli juures. Olen olnud seotud mitmesuguste teaduslike projektidega, mille eesmärgiks on inimese mitokondriaalse DNA varieeruvuse uurimine erinevates Euroopa ja Aasia populatsioonides.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.

13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004, 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004, 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004, 109 p.