Chapter 25

The Topology of the Maternal Lineages of the Anatolian and Trans-Caucasus Populations and the Peopling of Europe: Some Preliminary Considerations

Kristiina Tambets, Toomas Kivisild, Ene Metspalu, Jüri Parik, Katrin Kaldma, Sirle Laos, Helle-Viivi Tolk, Mukaddes Golge, Halil Demirtas, Tarekegn Geberhiwot, Surinder S. Papiha, Gian Franco de Stefano & Richard Villems

Here we discuss how our understanding of the peopling of Europe by modern humans may be improved by results which can be obtained in the investigation of genetic lineages of populations living in Anatolia and the Trans-Caucasus: Turks, Armenians, Georgians and Ossetes (Fig. 25.1). These four populations occupy a geographic area commonly believed to have great importance for the peopling of Europe. The present paper is directly related to our other paper in this volume (Kivisild et al.) which primarily addresses the genetics of Indian populations and our understanding of the first waves of migration of modern humans out of Africa and the peopling of Eurasia in general.

The historical background

Anatolia and the Trans-Caucasus are famous for their extraordinarily deep and rich history. They are close to and partially overlap with the Fertile Crescent; they include the upper parts of the Tigris and Euphrates basins and Lake Urmia. With the oldest signs of food production anywhere and with the oldest urban settlements which even seem to predate these, the region continued to play an important role for many millennia. Somewhat surprisingly, Anatolia itself is so far very poor in Palaeolithic finds. In contrast, the western Caucasus, in particular the Black Sea coastal area, Colchis and the Upper Kura and Arax River basins have yielded findings that cover a long time span starting from the Upper Palaeolithic and extending well into the Lower Palaeolithic (Rybakov *et al.* 1984). There are numerous (uncorrected) radiocarbon dates for the region, starting from about 35,000 before present that presumably belong to modern humans. The same area is equally rich in Mesolithic findings (Rybakov et al. 1989). However, this surprisingly different picture between Anatolia and Georgia may have a trivial explanation: lack of Lower Palaeolithic finds in Anatolia may be the result of the limited fieldwork. (P. Dolukhanov, pers. comm.).

The Neolithic and Chalcolithic periods of western Asia, including Anatolia and the Trans-Caucasus area, are much better documented, with numerous findings throughout southern Anatolia and Upper Mesopotamia as well as the surroundings of Lakes Van, Urmia and Sevan. The Bronze and Iron Ages of Anatolia and the Trans-Caucasus are again extraordinarily rich in important and often culturally overlapping civilizations (Hattian, Hittite, Hurrian and others).

The Hittite language was one of the earliest written Indo-European languages. It is believed that migration starting from Thrace around 1300–1200 BC, bringing to Anatolia yet another wave of Indo-European-speaking populations. This invasion has

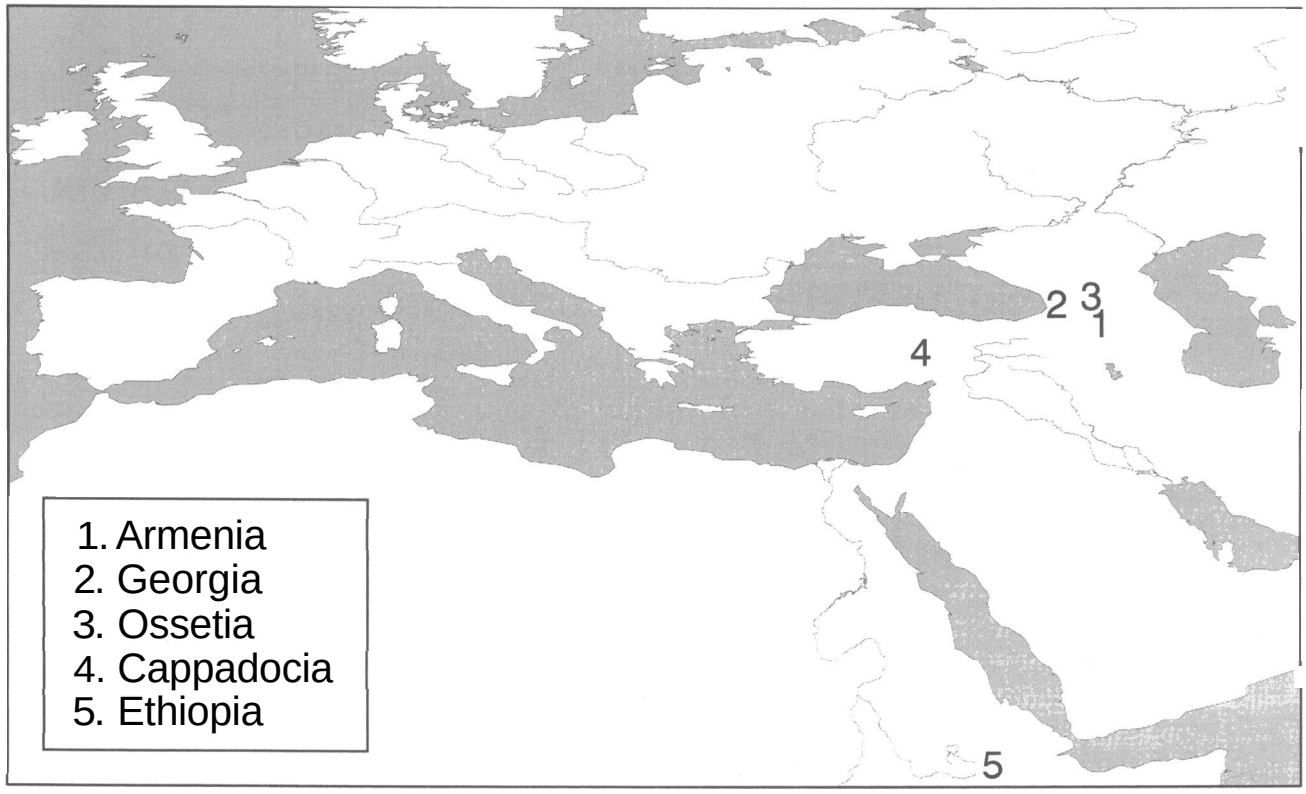


Figure 25.1. *Map of the area. Ossetes live in both sides of the High Caucasus; our sample was collected in Ossetian villages in Georgia.*

been correlated with the destruction of Troy VIIb and the end of Hattusas (the latter dated around 1180 BC because of an abrupt end of written sources), but sophisticated neo-Hittite and Hurrian-Urartian cultures remained and developed in southeastern and eastern Anatolia. Meanwhile, Ionian Greeks started to build their civilization in western Anatolia and these, together with new Indo-European-speaking states further east — Lydians and Medians — became, for a while, part of Achemenid Persia and were taken over by Alexander's empire and its successor states. The following centuries saw the presence of Rome and Byzantium in western Anatolia as far as Armenia and successively Seleucid and Parthian empires in the east. Slow westward advancement of various Turkic nomads, their conversion to Islam in Central Asia by AD 970 and, finally, the fall of Constantinople in AD 1453 brought this period to an end.

The populations which we have studied are linguistically diverse. Turkish, a novelty in Anatolia, belongs to the Altaic group of languages, while Ossetian and Armenian are branches of the Indo-European phylum of languages. Kartvelian, spoken by Georgians (self-name Kartvels), belongs to the southern branch of the Caucasian group of languages, considered to be as distant from the Indo-European phylum as, for example, the Afro-Asiatic (Hamito-Semitic) group. Furthermore, Georgians are often considered as an autochthonous population of the region. The Nostratic school of linguists suggests that the Caucasian languages may have had, until the Neolithic, a much wider geographic distribution, including southern Europe. Apart from this, they also argue that the homeland of the Indo-European languages is in eastern Anatolia — Armenian Highlands (e.g. Gamkrelidze & Ivanov 1984).

Despite their language, maternal lineages of present-day Turks have little in common genetically with the Siberian Altaic populations; instead Anatolian genetic continuity is likely. However, this continuity is itself diverse and complex as well as deep. Linguistically, Hurrians and Urartians did not speak Indo-European languages and the extent of genetic change during the replacement of Urartian (itself perhaps a branch of Hurrian) by Indo-European Armenian, which occurred at about 700–600 BC, is unclear. The nature of the Hurrian language is obscure, since being of the agglutinative type, it was neither Semitic nor Indo-European. By analogy with the Turkish invasion of at that time largely Greek-speaking Anatolia, the much earlier linguistic change

from Urartian to Armenian might have had only a minor effect on the genes of the population living in this area, particularly so for maternally inherited mtDNA. There seems to be an agreement now that out of the languages extant today, Greek is the closest to Armenian. The subsequent colourful history of Armenians is much better documented: here it is worthwhile noting that Greater Armenia covered a much wider area than the present-day Republic of Armenia and that at some point in Armenian history, about 800 years ago, 'Lesser Armenia' functioned even as a state largely within the borders of historic Cilicia, situated south and east of Cappadocia.

Finally, the Indo-Iranian speaking Ossetes are considered to be direct descendants of the historic Alans. Relative to Sarmatians, it is believed (depending on the interpretation of names in Chinese chronicles), that at some time they lived much further east, being a threat to China's northern borders. Defeated by the Huns in AD 375, most Alans fled the northern Pontic-Caspian steppes and took refuge in the northern Caucasus, pushed further south by subsequent attacks by Mongols and other invaders. At present, they live both north and south of the High Caucasus, in the Russian Federation and in Georgia, respectively.

Samples

Our sample of Ossetes was collected in Ossetian settlements in Georgia, south of the High Caucasus, Armenian samples in Armenia and Georgian samples from Tbilisi and the Megrelian part of Georgia. The Turkish samples are from the historic Cappadocian area. Ethiopian samples were collected from four different localities in Ethiopia. Population and sub-population (tribal, where appropriate) affiliation and birthplaces of the four grandparents1two parents of a donor were established by questioning. It turned out that a substantial fraction of the grandparents of Armenians in our sample were from the Lake Van area, from the neighbourhood of Erzurum and even from the shores of Lake Urmia. Samples were analyzed as explained in the accompanying paper (Kivisild *et al.* Chapter 31).

Anatolia and particularly the Caucasus area, are ethnically and linguistically most diverse and we are well aware that our analysis covers only a fraction of it. On the other hand, all populations under the study here are numerous; except for the Ossetes, their numbers are in millions and they constitute an absolute majority of the people living in the area.

Formulation of the questions

Out of a potentially long list of questions, we concentrate below on only a few. The first is, in the most general terms, the topology of the phylogenetic tree of the maternal lineages of the Anatolian and the Trans-Caucasus populations. The second is the comparison of these results with data available for other major geographic areas / groups of populations such as Europeans, Indians and Mongoloids. However, our main thrust is an attempt to contribute to the understanding of the place of these populations in the formation of genetic diversity we see among the extant Europeans. To do so, we use two approaches. Firstly, we compare European and Anatolian-Trans-Caucasian populations with respect to the topology of the mtDNA tree. Secondly, we seek to identify the sub-founders within the reconstructed lineage clusters and to compare the corresponding inferred coalescence ages for the populations, living in Europe and in the Anatolia–Trans-Caucasus. Here, we limit our analysis to some case studies. Finally, we bring in the 'sub-Saharan African angle' in order to discuss a few specific aspects of the mtDNA tree, first of all related to the varieties of haplogroup M, found in Anatolia.

Previous results

Although studies using classical genetic markers to compare the four populations investigated here are limited (Cavalli-Sforza *et al.* 1994), they do provide a clear general picture. Among western Asian populations, Turks, Georgians and Armenians all belong to separate branches of one of the two major genetic clusters (e.g. fig. 4.15.1 in Cavalli-Sforza *et al.* 1994). In a more general tree, covering 39 Asian populations (Cavalli-Sforza *et al.* 1994, fig. 4.10.1), they group together with Indian populations and Arabs in a cluster that is well separated from Mongoloids and other Asians. Synthetic genetic maps of the world, based on the first, second and third principal components, all show that the corresponding gene frequencies in Anatolia and the Trans-Caucasus area map together with those characteristic for western Eurasia and northern Africa (Cavalli-Sforza *et a].* 1994).

Three previously published papers deal with mtDNA hypervariable region (HVR) I (or I and II) sequences of Turks and cover, together, about 100 individuals (Richards *et al.* 1996; Calafell *et al.* 1996; Comas *et al.* 1996). All of them lack RFLP identification of mtDNA haplogroups: that was one of the

reasons why we investigated a further 400 Turks, using a combined HVR I and RFLP analysis. The Adygeis, a small population living in the northern Caucasus, are well studied by extensive RFLP and HVR I analysis (Macaulay *et al.* 1999). Partial RFLP analysis of the Ethiopian mtDNA haplogroups revealed the presence of haplogroups U and M among Ethiopians (Passarino *et al.* 1998) and a recent paper (Quintana-Munci *et al.* 1999) links Ethiopia and Eurasia via 'the South Route'. Our three previous papers (Metspalu *et al.* 1999; Kivisild *et al.* 1999a,b) cover selected general and specific aspects of the distribution of mtDNA haplogroups in western Eurasia and India.

Table 25.1. *MtDNA haplogroups in Anatolian–Trans-Caucasus populations (%).*

Haplogroup	Armenians 192	Georgians 139	Ossetes 187	Turks 388
A	0	0	0	0.5
Ä	2.6	2.2	11.8	0.8
B	0.5	0	0	0
F	0	0	0	0.3
H	30.9	17.3	18.7	25.0
I	1.6	2.2	4.3	2.3
J	8.9	3.6	18.7	10.9
K	7.9	10.1	1.1	5.9
L	0.5	0	1.1	0.3
M	0	2.2	2.1	4.1
M1	0	0.7	0	0.3
Õ	1.0	0	0	1.3
O*	7.3	7.2	13.4	3.6
P*	0	1.4	0	0.5
V+pV	0	0.7	0	0.3
R*	1.0	5.0	0	1.7
pJT/pHV	0.5	0.7	1.1	2.3
T	11.5	12.9	6.9	11.9
U	20.4	21.6	17.9	19.1
W	1.0	1.4	2.1	3.9
X	2.1	10.1	0.5	4.4
Ü	2.1	0	0	0.3

Notes:
1. A is defined by transitions at nps 16,223,16,145 and a transversion at np 16,176G or A and + *Ava*II 8252; +*Hph*I 10,237; -*Nla*III 12,501; -*Dde*I 1719.
2. Õ is defined by transitions at nps 15,043, 16,223, 16,172, 16,248, a transversion at np 16,147A and +*Hph*I 10,237; -*Nla*III 12,501; -*Dde*I 1719.
3. pJT/pHV is defined by transitions at 16,126, but without+*Nla*III 4126 and 00073A (see also text).
4. O* is defined as an internal node with 00073A -*Mse*I 14,766 and +*Alu*I 7025 (see Fig. 25.4).
5. R* is defined as an internal node with -*Mbo*II 12,704; +*Mse*I 14,766 and 16,223C; here are included haplotypes deriving from R* which do not belong to any of the described so far haplogroups (e.g. U, T, J etc.).
6. Ml is defined as M with an array of transitions at nps 16,223, 16,129, 16,189 and 16,249 (see Fig. 25.2).
7. P* is defined by +*Mse*I 14,766; 00073G and 16,217C (see Fig. 25.3).
8. Ü is defined by transitions at nps 16,201, 16,223 and 16,265 and -*Dde*I 10,398; +*Hph*I 10,238;-*Nla*III 12,501; -*Dde*I 1719; *Hph*I 10,237.
9. Haplogroup K is a sub-cluster of haplogroup U.

Empirical data

Frequencies of mtDNA haplogroups in Anatolian-Trans-Caucasus populations

Table 25.1 compares mtDNA haplogroup frequencies of Armenians, Georgians, Ossetes and Turks among each other and with Europeans and Indians as 'outgroups'. The overall result is clear: the distribution of haplogroups in all of the four populations studied is much closer to that typical of European rather than to Indian populations. Yet the differences between mtDNA haplogroups in Europeans and Anatolian-Trans-Caucasians are both quantitative and qualitative. Unfortunately, this table lacks data on other Turkish language group populations. However, putative assignment of approximately 200 mtDNA HVR-I sequences of different Central Asian Turkish-speaking populations (Comas *et al.* 1998) into a median network tree (see reconstruction in Kivisild *et al.* 1999a) shows that significantly more than a half of their mtDNA variants belong to haplogroups M, A, F and B, infrequent in all four Anatolian - - Trans-Caucasus populations (Table 25.1), but typical for Mongoloids (Torroni *et al.* 1993). The major European haplogroup H in Anatolians-Trans-Caucasians is somewhat less frequent than in European Caucasoids (Torroni *et al.* 1994), specifically so among Georgians and Ossetes (Table 25.1). The other major haplogroup, U, is nearly equally frequent in Europeans (Torroni *et al.* 1996; Kivisild *et al.* 1999b), Anatolians-Trans-Caucasians and in Indians (Kivisild *et al.* 1999b). However, the three geographic areas differ substantially when sub-clusters of U are compared (Table 25.2). Some of these differences will be discussed below; here we indicate one aspect of these differences which seems to be interesting. Namely, excluding a few exceptions such as the lack of sub-cluster U4 mtDNAs in our Ossetian sample, all basic western Eurasian varieties of haplogroup U are much more evenly represented among Anatolians-Trans-Caucasians than in Europeans. This finding can be understood in the context

Table 25.2. *Distribution of haplogroup U varieties in Anatolians-Trans-Caucasians and in some other populations*

Population	*N*	*n*	%	Per cent from haplogroup U							
				U°	U1	U2	U3	U4	U5	U6	U7	K
Armenians	192	54	28.1	1.8	16.7	7.4	16.7	13.0	13.0	0	3.7	27.8
Georgians	139	44	31.7	0	11.4	4.5	13.6	25.0	11.4	0	2.3	31.7
Ossetes	187	36	19.3	0	11.1	25.0	16.7	0	16.7	0	25.0	18.9
Turks	388	92	23.7	5.4	15.2	4.3	22.8	4.3	21.5	0	6.5	24.7
A+G+T	719	190	26.6	3.1	14.1	5.2	18.3	11.5	22.0	0	4.7	27.2
Europe-1	509	118	23.2	5.1	2.5	10.2	6.8	15.3	41.5	0	0	18.6
Europe-2	484	109	22.5	2.8	8.3	4.6	3.7	10.1	33.0	1.8	0.9	35.8
Europe-3	382	100	26.2	3.0	1.0	2.0	1.0	20.0	62.0	0	0	11.0
Indians	550	72	13.1	0	2.3	77.9	0	4.7	1.2	0	12.7	1.2
Ethiop.	270	15	5.6	6.7	6.7	13.3	13.3	0	0	40.0	0	20.0
Nile Valley	255	12	4.7	0	8.3	0	25.0	0	16.7	16.7	8.3	25.0

Notes:
1. U° is defined as +12,308 *Hinf*I and -11,465 *Tru1*I mtDNAs that do not belong to any of the indicated above sub-clusters of haplogroup U.
2. A+G+T: Armenians, Georgians and Turks: this paper.
3. Europe-1: Russians, Poles, Czechs, Slovaks: our unpublished data and Orekhov *et al.* 1999.
4. Europe-2: Sienans, Tuscans, Sardinians, French, Albanians: our unpublished data and Di Rienzo & Wilson 1991; Torroni *et al.* 1996.
5. Europe-3: Estonians, Finns, Karelians: our unpublished data and Sajantila *et al.* 1995; Richards *et al.* 1996; Villems *et al.* 1998.
6. Indians: Kivisild *et al.* 1999b.
7. Nile Valley populations: deduced from HVR I sequence information (Krings *et al.* 1999).

of early and subsequent migrations of modern humans to Europe with respect to the coalescence ages of individual sub-clusters of this complex mtDNA haplogroup in Europe and in western and southern Eurasia.

Comments on haplogroup M

Haplogroup M is, somewhat surprisingly, infrequent in Anatolian-Trans-Caucasus populations, not exceeding a few per cent in Turks, Ossetes and Georgians and was absent among 200 Armenians studied by us (Table 25.1). Although the overall frequency of haplogroup M lineages is low among Turks, our large sample size makes their total number sufficient for phylogenetic analysis (Fig. 25.2). Most of the lineages are typically Mongoloid-specific varieties of this cluster, found equally frequently among Central Asian Turkish-speaking peoples but not in Indians. It is likely that their presence indicates a trace of the Altaic Turkish maternal lineages among contemporary Turks. Interestingly, some of the haplogroup M lineages found among Turks and Georgians belong to the northeastern African specific sub-cluster of haplogroup M (M1), characterized by an HVR I motif of four transitions at nps 16,189, 16,223, 16,249 and 16,311 (Rando *et al.* 1998). We have also found the same sub-cluster of haplogroup M among southern Sicilians and the characteristic motif of four transitions (without haplogroup identification) is present in the published Iberian and Sardinian mtDNA HVR-I sequences (Corte-Real *et al.* 1996; Di Rienzo & Wilson 1991). This variety of haplogroup M is found in northern and northwestern Africans (Rando *et al.* 1998), but at considerably lower frequency and — what is more important — at lower diversity, than can be observed among Ethiopians (Fig. 25.3), where its frequency in our sample was 15.5 per cent (*N* = 270). Further search of the literature revealed that the same motif is also frequent in populations living in the Nile Valley from Mediterranean to southern Sudan (Krings *et al.* 1999). However, diversity of haplogroup M among Africans and in the European Mediterranean islands is practically restricted to a few founder lineages, and, in contrast to Indian and Mongoloid populations, we have not found in Ethiopian populations haplotypes, corresponding to M* node, frequent among southern, eastern and central Asian

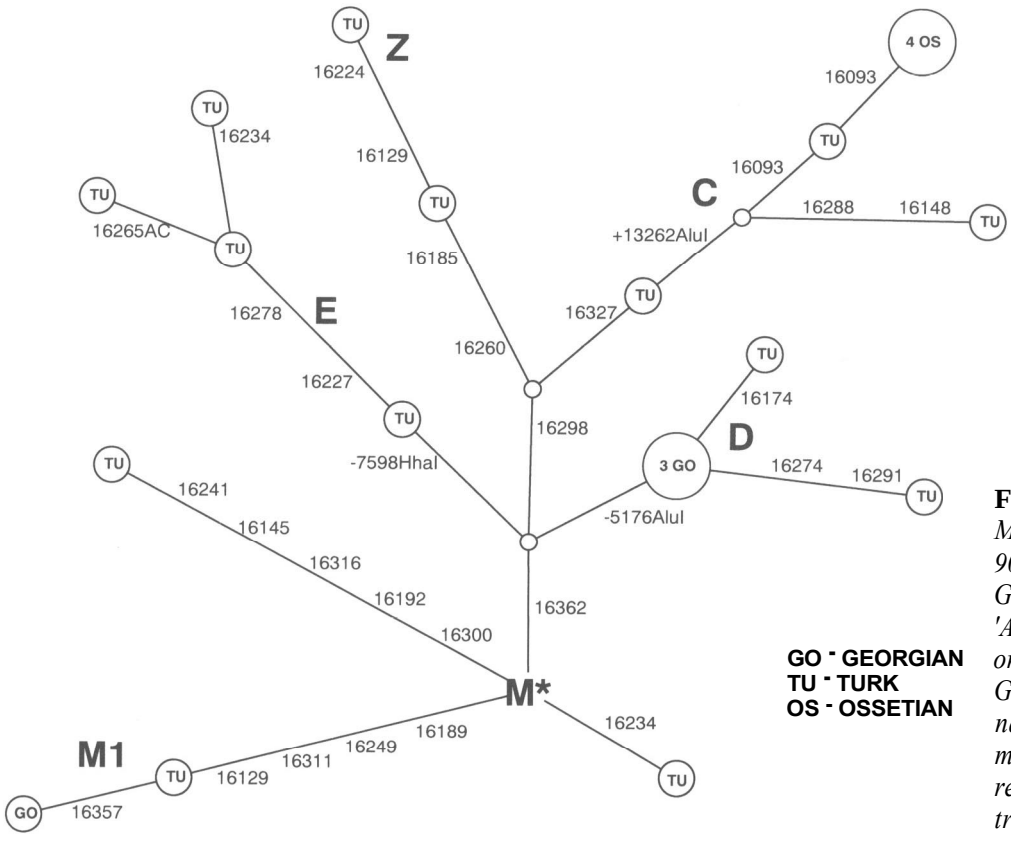


Figure *25.2.22 haplogroup M mtDNAs found among 906 Turks, Ossetes, Georgians and Armenians. 'African' M1 is present in one Turkish and one Georgian individual. The nature of mutations in mtDNA hypervariable region I is specified onlyfor transitions.*

populations.

We consider it likely that haplogroup M arose some 50,000–70,000 years ago (Chen *et al.* 1995) and suggest that its carriers migrated from southern to southeastern Asia before its further diversification and underwent the first expansion/diversification event simultaneously in southern and southeastern Asia around 50,000–60,000 years ago (see Kivisild *et al.* Chapter 31). Time depth of sub-haplogroup M1 is uncertain to us because of the topology of this lineage cluster. However, diversification at two putative internal sub-founders: after transitions at np 16,129 and at np 16,359 (Fig. 25.3), exhibit very similar expansion times around 12,000–13,000 BP (12,800±3400 and 11,900±2300 BP, with $n = 22$ and 44, respectively). Interestingly, a similar value (about 12,800 BP) can be obtained for a yet another subfounder of a branch characterized by nine base-pair deletion in the intergenic tRNALys/COII region (Fig. 25.3). On the other hand, M1 is probably much older than these apparent expansion phases: accumulation of an array of transitions takes time. In a recent paper by Quintana-Murci *et al.* (1999) the authors argue that the presence of haplogroup M1 in northern Africa/ Ethiopia is best interpreted as an evidence for an African origin of haplogroup M. We feel that the question is still unsettled: a widespread presence of the nodal variant for haplogroup M in southern, eastern and central Asia and its lack in Africa plus a vast diversity of this haplogroup in the former regions compared to a single 'ripe' branch in Africa and in Mediterranean basin, sub-haplogroup M1, offer several different scenarios.

However, as far as Anatolians are concerned, the important point is that the frequency of haplogroup M in this Turkish-speaking population is an order of magnitude lower than among the Central Asian Turkish speakers. Here, it is interesting to add that Turkish-speaking Tartars in Russia, living predominantly in the Volga basin, possess this haplogroup at frequencies around 13 per cent: intermediate between the low in Turks and the high in Central Asians (V. Orekhov pers. comm.).

From R to O*: a tiny but phylogeographically interesting internal node P**

As far as Caucasians are concerned, R* and O* are the two important internal nodes of the mtDNA phylogenetic tree. At the level of the present resolution of the topology of the tree, these two nodes are

separated by two mutations. Looking from the direction of the African root, the first is the loss of the *Mse*I restriction site at np 14,766 and the next is a G to A transition at np 00073 (see Fig. 25.2a in the accompanying paper by Kivisild *et al.)*. We have now found a number of haplotypes that radiate from this node which we propose to identify as P* (Fig. 25.4), all of them harbouring an additional transition at np 16,217. Their phylogenetic position linked to geographic distribution is interesting and, we think, important: it extends from the western Mediterranean to India, including the Trans-Caucasus.

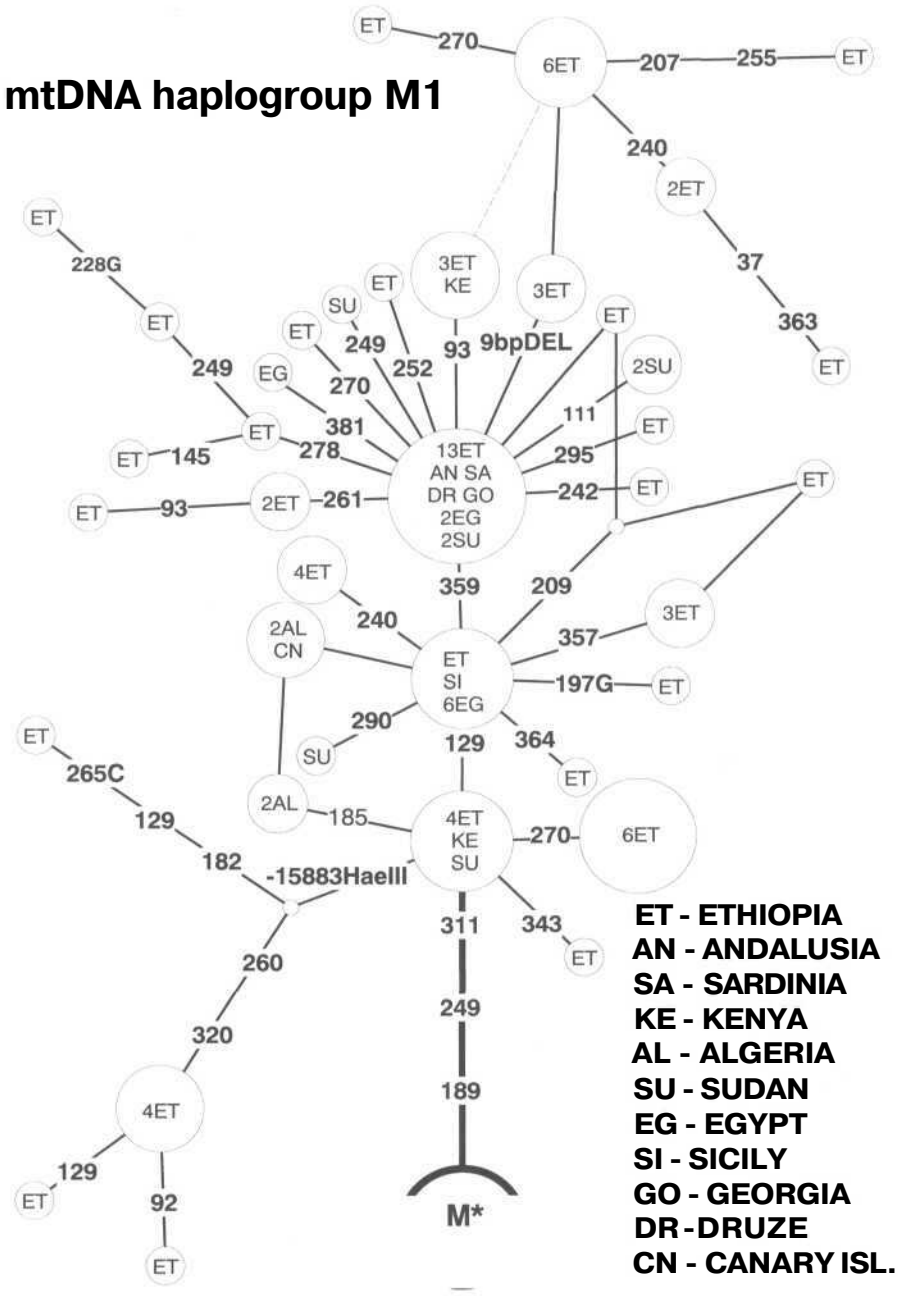
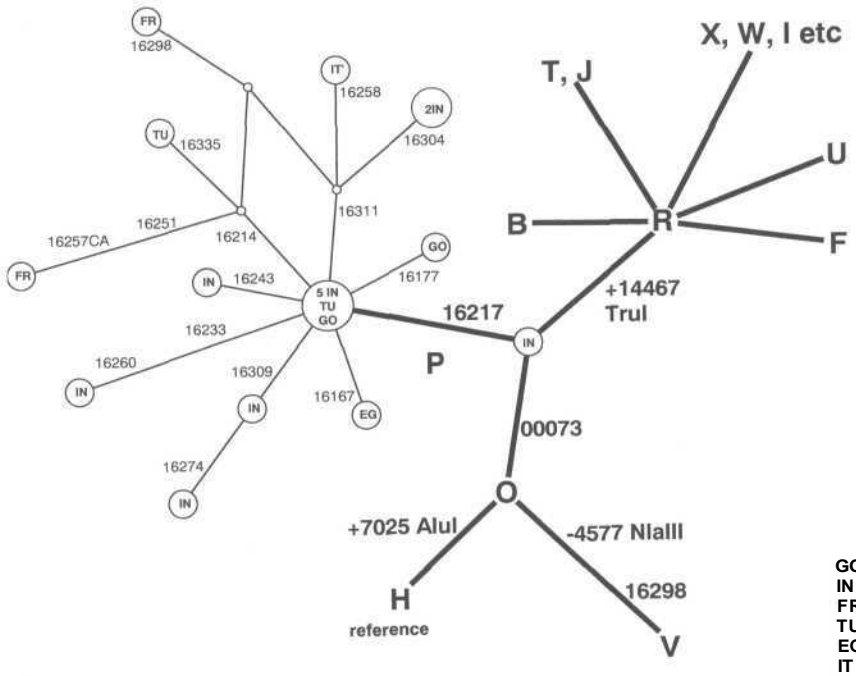


Figure 25.3. *Topology and phylogeography of sub-haplogroup M1 in Africa and western Eurasia.*

The internal node O, ancestral to about a half of Europeans*

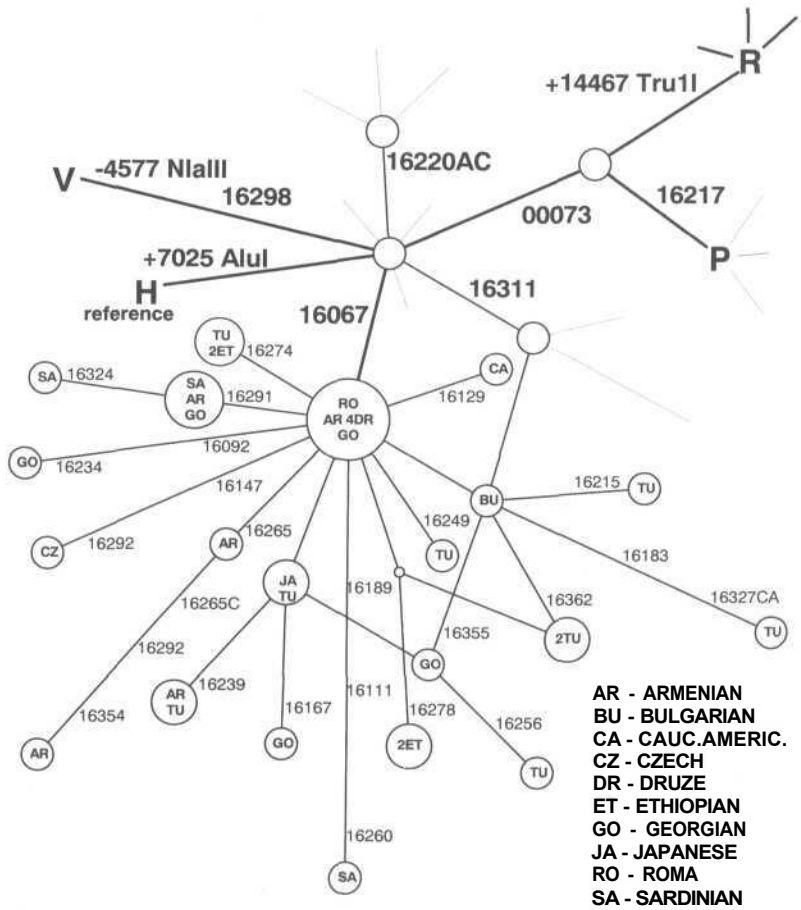
Metspalu *et al.* (1999) have shown that the mtDNA pools in Armenians, Georgians and Ossetes are particularly rich in haplotypes that belong to neither haplogroup H nor V, but are nevertheless derivatives of (or coincide with) the internal node O* (HV* in Richards *et al.* 1998). With new data on Turks available, this picture is even richer and proves that this internal node, important for the understanding of the history of maternal lineages of the present-day Europeans, is particularly diverse and frequent in Anatolians-Trans-Caucasians (Fig. 25.5). Maternal lineages of four populations studied by us — Turks, Armenians, Georgians and Ossetes — are represented in all lineage clusters, deriving from internal node O*, while Indians, although represented, are distributed less widely and an overall frequency of such lineages in Indian populations studied by us is around 1.5 per cent (Kivisild *et al.* 1999b) compared to 7-13 per cent in Trans-Caucasus area populations and around 4 per cent in Turks.

One of the several lineage clusters additional to H and V, deriving from node O*, is defined by a transition at np 16,067 (Fig. 25.5). The coalescence age for this cluster is approximately 30,000±4000 BP (n = 38; ρ = 1.5), suggesting that its expansion has started before the Last Glacial Maximum (LGM). This time estimate is somewhat older than similarly calculated coalescence age for haplogroup H (see below). The spread of this cluster is of particular



GO - GEORGIAN
IN - INDIA
FR- FRENCH
TU- TURK
EG - EGYPTIAN
IT - ITALIAN

Figure 25.4. *Phylogenetic position of an internal node P*, occupied by one Indian, and topology and phylogeography of haplogroup* P. *'Reference' indicates Cambridge Reference Sequence (Anderson* et al. *1981).*



AR - ARMENIAN
BU - BULGARIAN
CA - CAUC.AMERIC.
CZ - CZECH
DR - DRUZE
ET - ETHIOPIAN
GO - GEORGIAN
JA - JAPANESE
RO - ROMA
SA - SARDINIAN
TU - TURK

Figure 25.5. *Phylogenetic position of an internal node O* (HV* in Richards* et al. *1998) and topology and phylogeography of a lineage cluster with a characteristic founder transition at np 16,067. 'Reference' indicates Cambridge Reference Sequence (Anderson* et al. *1981).*

interest: it is most frequent in the Trans-Caucasus populations, seems to be absent in Indians (none out of about 1200 Indian mtDNAs studied) and is rare in Europe, where it seems to be relatively frequent only among eastern-central Mediterranean islands. Its presence in northeastern African populations is also evident. However, in western, northern and north-eastern Europe its frequency is below 0.5 per cent. Thus, we see here a pre-LGM, Upper Palaeolithic mtDNA lineage cluster that has an epicentre of frequency and diversity in the Trans-Caucasus area populations, but it is virtually absent in most of Europe as well as in India, while present in northwestern Africa and central and eastern Mediterranean (Druzes, Sardinians, Cretans). Although one can always blame random genetic drift, it nevertheless raises a question about selectivity in possible westward migrations: not all of the mtDNA pool which one would be expecting to find in Anatolia at the beginning of Neolithic, can be sampled among the present-day Europeans, except in a restricted Mediterranean area.

Sub-haplogroups U1, U3 and U7 are more frequent in Anatolians-Trans-Caucasians than in Europeans

Table 25.2 shows that three sub-clusters of haplogroup U (Ul, U3 and U7), are more frequent in Anatolians-Trans-Caucasians than in Europeans. Sub-cluster U7 is present among Turks, Armenians, Georgians and Ossetes (Table 25.2; Fig. 25.6), as well as in Druzes (Macaulay *et al.* 1999). However, this sub-cluster has hardly penetrated the European mtDNA pool: we did not find U7 among French, Russian, Polish, Czech, Croatian, Cretan, and Estonian populations (*N* ~1400), and we could infer only a single U7 from the data published by others on German, Swiss, Austrian, Iberian, British, Finnish, Karelian and Saami HVR-I sequences (N ~ 2300). Interestingly enough, like O*-16,067, U7 seems to be present in Sardinians (Di Rienzo & Wilson 1991; identification by HVR-I data only). Furthermore, while Indian and western Eurasian (including Europe, Trans-Caucasus and Anatolia) varieties of U2 differ profoundly (Kivisild *et al.* 1999b), with U7 we see a different picture: Indian and Anatolian-Trans-Caucasus haplotypes even coincide partially (Fig. 25.6). It suggests that the cluster has not only a common founder but the carriers of this particular variety of mtDNA seem not to have been isolated from each other until relatively recently. Nevertheless, a nodal haplotype of a sub-founder with motif 16,318AT — 16,309 is abundant only in Indians (Fig. 25.6).

On the other hand, we have not found sub-haplogroup U3 in Indians, whereas it forms a major

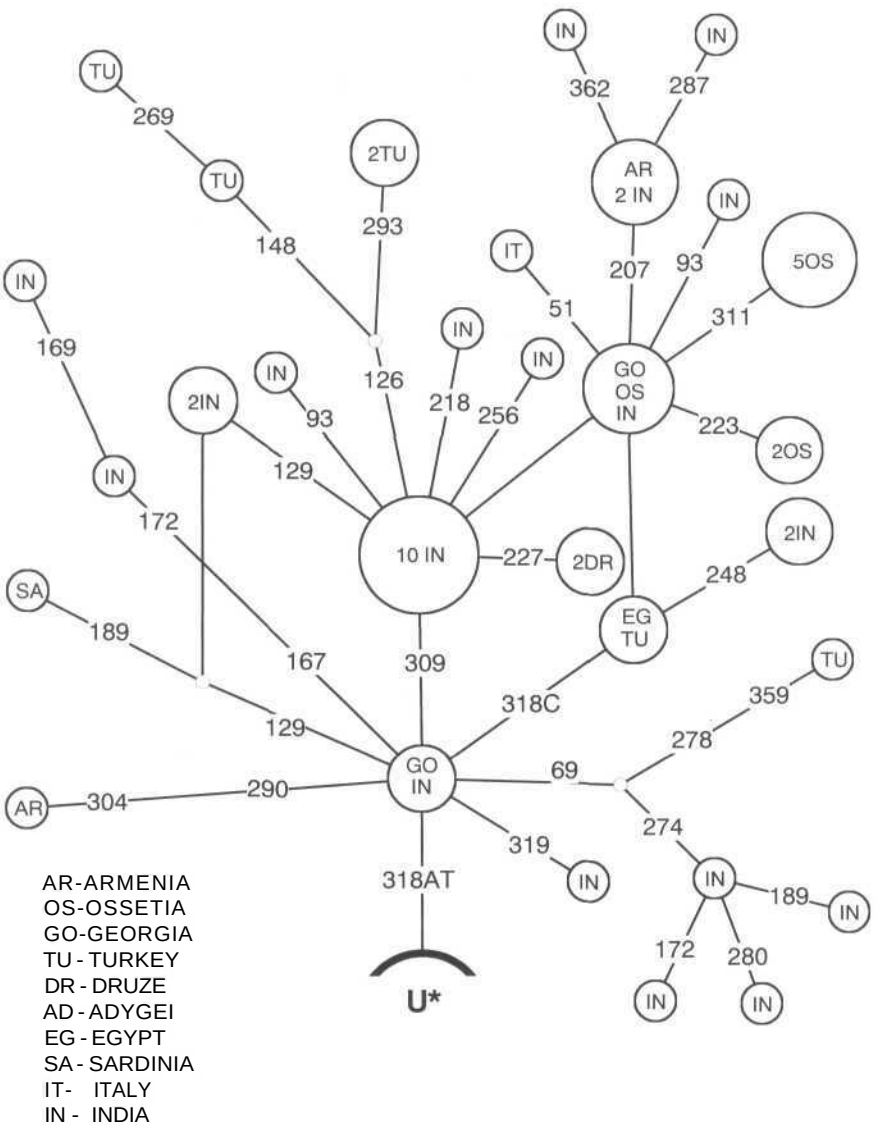


Figure 25.6. *Topology and phylogeography of mtDNA sub-haplogroup U7. Mutations in mtDNA hypervariable region I sequence are shown less 16,000.*

branch of the Armenian, Georgian, Ossetian and Turkish haplogroup U sub-clusters (Table 25.2). This sub-cluster is less frequent in European populations, except in some eastern Mediterranean islands (Metspalu *et al.* in prep.). It is also present in Ethiopians and, presumably, in Nile Valley populations outside Ethiopia.

U1, in turn, although most frequent among Anatolians–Trans-Caucasians, can be found at lower frequencies also in Indians and Europeans as well as in Ethiopians (Table 25.2).

Hence, a different east–west polarity of sub-haplogroups U1, U3 and U7, as well as of 'non-canonical' derivatives of O*(see, e.g. Metspalu *et al.* 1999), shows that populations living in Anatolia and the Trans-Caucasus area share maternal lineages with southern Asians, northeastern Africans and with Europeans in a complex, not yet fully understood manner.

Several sub-clusters of typically western Eurasian mtDNA haplogroups display an Early Neolithic beginning of expansion in Europe compared to 20,000 BP or earlier in Anatolians–Trans-Caucasians.

In our previous paper (Metspalu *et al.* 1999) we described several sub-clusters of western Eurasian mtDNA haplogroups which are shared between Europeans and Anatolians–Trans-Caucasians, but exhibit profoundly different expansion times. Examples for such sub-clusters can be found in many haplogroups, most evidently among T, J and U. Here we bring just one example among haplogroup T (Fig. 25.7). As a rule, the inferred coalescence times of such sub-cluster in Anatolians–Trans-Caucasians lie around 20,000–25,000 BP, while their equivalents in Europe around 8000–9000 BP.

16,126 – 00073A variants of mtDNA in Anatolians–Trans-Caucasians, which are virtually absent in Europeans but morefrequent in Ethiopians and likely also in the Nile Valley populations

Haplogroups T and J derive from an internal node R* (Fig. 25.4). Both these haplogroups are common in Europeans (Torroni *et al.* 1994; 1996; Richards *et al.* 1998; Kivisild *et al.* 1999b)and also among Anatolians–Trans-Caucasians (Table 25.1). They are sister groups because they share a number of common mutations compared to the Cambridge Reference Sequence, not shared by other derivatives of the internal node R* (Macaulay *et al.* 1999). The combination of two mutations: transition at np 16,126 and the gain of *Nla*III site at np 4216, are most informative. We have found in Anatolians–Trans-Caucasians, albeit at low frequencies, mtDNA variants which possess a characteristic for T and J transition at np 16,126 but lack the indicated RFLP site at np 4216 and we classified them as pre-JT (Table 25.1). In fact, they are not ´pre-JT´ and the variants found by us are also not monophyletic. There are several such sub-clusters, some of them specific to Indians only; the most frequent among them is a sub-cluster possessing an additional transition at np 16,362 and A at np 00073, characteristic for O* derivatives. Such mtDNA variants were earlier described among Druze population as pre JT*/preHV* (Macaulay *et al.* 1999). We found such mtDNA lineages in Georgians and Turks (Table 25.1). These lineages are rare in Europe: for example, out of more than 1100 Germans (Pfeiffer *et al.* 1999; Richards *et al.* 1996; Lutz *et al.* 1998) only a single HVR I haplotype can be tentatively (no RFLP data available) assigned to this lineage cluster, suggesting frequencies in northern Europe of around 0.1 per cent. However, we found that this type of lineage is rather frequent in Ethiopian populations. Furthermore, the same types of HVR I motifs are also present in the Nile Valley populations (Krings *et al.* 1999). In our Ethiopian sample their frequency is 10 per cent ($N = 270$), exceeding significantly that in Anatolians–Trans-Caucasians (Table 25.1) as well as the frequency of haplogroups T and J in Ethiopians, which in our sample is 2.3 and 1.1 per cent, respectively. Inferred from HVR I data of Krings *et al.* (1999), the frequency of the 16,126–16,362 lineage in the Nile Valley populations is approximately 4 per cent: somewhat higher than in Anatolians but still more than twice as low as in Ethiopians.

We feel that this finding might be important in further understanding ancient migrations from and to eastern–northeastern Africa. Lack of similar data on many crucial in this context populations does not allow further speculations but we remind here that similar phylogeographic spread is true for mtDNA sub-cluster M1, discussed above: found in Turks and Georgians, very rare in Europeans and frequent in Ethiopians and likely present also in the Nile Valley (defined as Egypt and Sudan: see Krings *et al.* 1999) populations.

Sub-cluster of haplogroup H with a motif of transitions at nps 16,293 and 16,311 is present in Europeans but absent in Anatolians–Trans-Caucasians

We started to pay attention to this motif because we found it at a relatively high frequency among Estonians and did not find it at all among nearly a thousand Turks, Armenians, Ossetes and Georgians. Searches among further populations showed that this sub-lineage and its derivatives are indeed spread quite unevenly among the Caucasoids of western

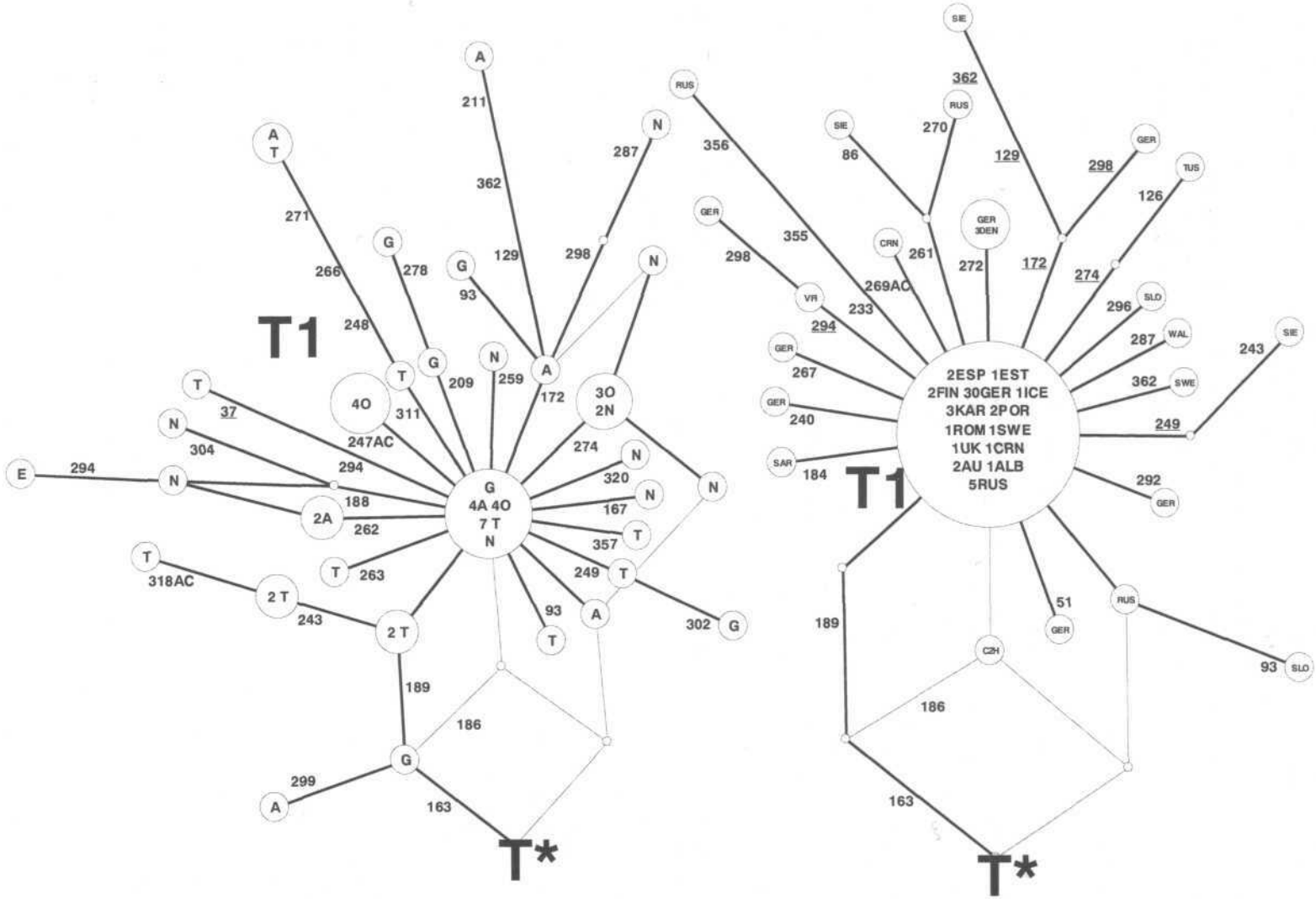


Figure 25.7. *Sub-haplogroup T1 in Anatolians-Trans-Caucasians, Nile Valley populations, Ethiopians and in Europeans. Mutations in mtDNA hypervariable region I sequence are shown less 16,000 and T* differs from the CRS by transitions at nps 16,126 and 16,294. Abbreviations: A = Armenians; G = Georgians; O = Ossetes; T = Turks; E = Ethiopians; N = Nile Valley populations; ALB = Albanians; AU = Austrians; CRN = Croatians; DEN = Danish; ESP = Spanish; EST = Estonians; FIN = Finnish; GER = Germans; ICE = Icelanders; KAR = Karelians; FOR = Portugese; ROM = Gypsies; RUS = Russians; SAR = Italians (Sardinia); SIE = Italians (Siena); SLO = Slovaks; SWE = Swedes; SZH = Czechs; TUS = Italians (Tuscany); UK = British; VFI = Volga-Finnish; WAL = Welsh. References to HVR 1 sequences published elsewhere (HVR I and RFLP) are given in the text.*

Eurasia (Table 25.3). Although frequent among Estonians, it cannot be considered specific for Finno-Ugric-speaking populations since it is low among Finns and undetected so far in Saamis and Karelians (Sajantila *et al.* 1995). Table 25.3 indicates that this motif has not reached either western or eastern Mediterranean, Anatolian or the Trans-Caucasian populations. Elsewhere in Europe its frequency is generally low; an interesting exception is the eastern Mediterranean Albanian population, where the sub-lineage is well represented. Because of its topology, the time depth of the entire sub-cluster is unclear. However, we have calculated approximate ages of the beginning of expansion for the two very similar sub-founders: 16,293, 16,311 and 16,293, 16,311, 16,278, respectively. The result suggests that this geographically localized expansion started about 5200±1700 BP, i.e. at the end of Neolithic and was almost simultaneous with an expansion of several cultures all over Europe, such as Linear Pottery, Impressed Ware and early pottery-bearing sites in the eastern Baltic area and in central and southern Russia. However, the reason why we have chosen to present this case-study is to show an example of how sub-founders inside haplogroup H may possibly be informative for determining demographic expansion ages that differ substantially from the estimate of the coalescence time of the haplogroup as a whole. And not only in haplogroup H. For example, haplogroup U contains yet another additional minor branch, characterized by a transition at np 16,146. This branch can be found among populations from Iberia to Karelia, but we have not sampled it in Anatolians-Trans-Caucasians or in Africans.

General discussion

The first massive investigation of the European mtDNA pool (Richards *et al.* 1996) re-initiated an intensive scientific debate about the peopling of Europe. Here, the central issue became the age and the place of origin of the extant European gene lineages, particularly with respect to mtDNA. Coalescence age calculations in that paper were used to challange the demic diffusion hypothesis, which stresses a Neolithic gene flow from the Middle East and/or Anatolia to Europe, as a result of food production, i.e. the spread of agriculture. The demic diffusion hypothesis, in its genetic form, is supported by the east-west gradient of the distribution of gene frequencies

Table 25.3. *Distribution of the* 26,293-26,322 *motif of haplogroup H.*

Population	*N*	*n*	Frequency	Reference
Estonians	148	9	6.1	this work
Finns	73	2	2.7	Sajantila *et al.* 1995; Richards *et al.* 1996
Karelians	82	0	0	Sajantila *et al.* 1995
Saami	373	0	0	Sajantila *et al.* 1995; Dupuy & Olaisen 1996; Delghandi *et al.* 1998
Slovaks	130	5	3.8	this work
Poles	93	2	2.1	this work
Czechs	94	2	2.1	this work
Russians	192	4	2.1	this work and Orekhov *et al.* 1999
Croats	400	2	0.5	this work
Albanians	129	6	4.7	this work
Swiss	74	1	1.3	Pult *et al.* 1994
UK	161	2	1.2	Richards *et al.* 1996
Germans	854	9	1.1	Pfeiffer *et al.* 1999; Richards *et al.* 1996
Sardinians	69	1	1.4	Di Rienzo & Wilson 1991
Sienese	150	1	0.7	this work
French	88	0	0	this work
Turks	388	0	0	this work
Iberians	269	0	0	Corte-Real *et al.* 1996; Bertranpetit *et al.* 1995; Richards *et al.* 1996 Comas *et al.* 1998
Armenians	192	0	0	this work
Ossetes	201	0	0	this work
Georgians	132	0	0	this work
Uigurs	50	1	2	Comas *et al.* 1998
Indians	1008	0	0	this work
Ethiopians	270	0	0	this work

explaining the first principal component (for a review see Cavalli-Sforza *et al.* 1994). In its strong form, the demic diffusion hypothesis suggests that the 'original European' hunter-gatherers perhaps survived only in geographically isolated areas (in the Pyrenees, in Sardinia, the Saami etc.). The critique points out that: a) most mtDNA lineage clusters in Europe seem to be older than 10,000 years (more precisely: display signs of expansion before the Neolithic period); b) gene frequency gradients do not indicate the direction of the movement; c) several anthropological findings suggest continuity of morphological markers throughout the critical period. In contrast, recent autosomal DNA marker studies prompted proponents of the Neolithic migration to suggest that at best a small fraction of the present-day genetic lineages of Europeans is older than the Neolithic (e.g. Chikhi *et al.* 1998). The debate is ongoing and a forthcoming paper (Simoni *et al.* 2000) argues against a Late Upper Palaeolithic re-expansion hypothesis as suggested by Torroni *et al.* (1998).

Discussing the impact of the Neolithic demic diffusion from Anatolia/ Middle East, we first need to remind ourselves that the first principal component and the synthetic map derived from it explains about a quarter of classical genetic markers out of those included in the analysis (Cavalli-Sforza *et al.* 1994). That, of course, is far from any nearly total replacement scenario.

It is clear now that from the point of view of mtDNA haplogroups, European and western Asian populations can be viewed as a largely homogeneous phylogeographic entity of Palaeolithic origin. Even the extremes like Saami mtDNA, are only seemingly different from typical for the western Eurasian populations mtDNA pool: the apparent differences are perhaps caused by very strongly pronounced random genetic drift, which has drastically reduced the number of mtDNA haplogroups and lineages in this population. However, only a few Mongoloid-specific mtDNA varieties are present in Saami, in average around 5 per cent of their total mtDNA pool. Hence, the Saami mtDNA pool is not an outlier relative to that of the European Caucasoids, as it has been suggested (Sajantila *et al.* 1995; Simoni *et al.* 2000) but a set of a subset: we think that this distinction is very important for the understanding of biological history of their maternal lineages and of the Saami population in general (Villems *et al.* 1998).

The second conclusion derives from the same often stressed homogeneity of the western Eurasian mtDNA pool: any contemporary analysis should, to be informative, make at least an attempt to analyze mtDNA data phylogenetically. Much has been done already (e.g. Torroni *et al.* 1994; 1996; 1997; Richards *et al.* 1996; Macaulay *et al.* 1999), but much is still left to be done. The wealth of the data makes it evident that some mtDNA lineage clusters and sub-clusters, for example haplogroup V, were 'born' in Europe after the LGM, but before the Neolithic (Torroni *et al.* 1998). However, there is no definite answer available at present about the proportion of such lineage clusters/sub-clusters in the European mtDNA pool. Above, we used one sub-cluster of haplogroup H as an example of a 'European-born' varieties. There are likely a number of other sub-clusters like that.

We have, however, found that several sub-clusters of the common Eurasian mtDNA haplogroups, which are present both in the Anatolian-Trans-Caucasus populations studied by us, as well in the eastern and western European populations, show systematically much later signs of expansion in Europe than in western Asia (Metspalu *et al.* 1999). This difference is large: in average 18,000–22,000 BP for Anatolians compared to 7000–9000 BP for Europeans. This significant time difference can be interpreted as supporting an Early Neolithic migration of the carriers of these lineages to Europe, coinciding with their expansion in new surroundings. Such a dissection of composite lineage clusters is not always straightforward and here we have limited our analysis to sub-clusters which exhibit star-like expansion patterns. Meanwhile, it is interesting to note that these sub-clusters, taken together, may well explain 25 per cent of the existing mtDNA variety in the extant Europeans. However, this is a summary quantitative estimate and the analogy with the above-mentioned gradient seen in the frequency of the first principal component is superficial. On the other hand, if we believe (and we do) that the identical (coinciding in their characteristic HVR polymorphic sites and diagnostic RFLP sites) sub-clusters in western Asian and in European mtDNA pools are indeed monophyletic and have not arisen independently, then we must accept that identical sub-clusters have expanded at widely different times in different localities. If this conclusion were true, one would have to be careful in the interpretation of the coalescence age calculations based on haplotype diversity of lineage clusters covering large geographic areas: obtained estimates may result from many discrete expansion events and the net outcome would be significantly less informative than it might seem at the first glance.

Since haplogroup H is the most frequent mtDNA lineage cluster in Europe, we feel that fur-

ther progress in the reconstruction of the demographic history of the Caucasoid maternal lineages depends on progress in better understanding the topology of this haplogroup. So far, however, we can say that the coalescence age of haplogroup H appears to be slightly older for Anatolian-Trans-Caucasus than for western European populations. For Georgians, the estimate is 22,500±4500 BP (with n = 23; p = 1.1); for Armenians 23,500±3000 BP (n = 53; p = 1.17) and for Turks 25,000±2300 BP (n = 97; p = 1.24). Similar calculations gave an even earlier date for Ossetes — about 29,000 BP — but since the topology of the phylogenetic tree of haplogroup H among Ossetes is not star-like, its coalescence age might be overestimated. One may also consider the different demographic history of Ossetes, discussed above. Otherwise, we see that coalescence ages for haplogroup H lineages for the three major Anatolian–Trans-Caucasus populations coincide within the limits of error and are close to coalescence ages for several discussed above mtDNA sub-haplogroups. The last finding is reasonable: one expects that demographic expansions during the Palaeolithic took place in a given geographic region for various reasons, expected to affect simultaneously a population as a whole. The time estimates here are slightly younger than those suggested by other authors (Torroni *et al.* 1998) for the Near Eastern populations (28,000–30,000 BP) and somewhat older than for the western European populations (–13,000–15,000 BP). For Estonians (a Finno-Ugric language-speaking population) we obtained the corresponding value as 30,000 BP, but this higher estimate is partially caused by a relatively high frequency among Estonians of derivatives of the discussed above haplogroup H sub-lineage with a motif of transitions at nps 16,293 and 16,311.

Summary

MtDNA lineages of all four Anatolian-Trans-Caucasus populations of this study contain Mongoloid-specific varieties at surprisingly low frequencies and an absolute majority of mtDNA haplogroups found are those present also in Europe. Furthermore, the frequencies of the major haplogroups are largely similar to those observed for Europeans;

MtDNA lineages which derive from an internal node **O** (HV in Richards *et al.* 1998) are significantly more divergent in Armenians, Georgians and Ossetes than they are among European populations;

3. Contrary to the former, haplogroup V, a sister group to haplogroup H, is virtually absent in Anatolian–Trans-Caucasus populations. Taking the last two observations together and bearing in mind that haplogroup H itself is frequent and diverse among Turks, Armenians, Georgians and Ossetes, one can assume that haplogroup V arose in Europe and has never migrated back to western Asia, while the latter area and Europe not only share its sister-group H, but H is also a dominant variety of mtDNA in both these regions;
4. A substantial fraction of the shared lineages between European and Anatolian–Trans-Caucasus populations exhibit much earlier expansion times in the latter area. The differences are large: around LGM for Anatolians and Early Neolithic for Europeans. Since we assume that these sub-clusters are monophyletic, this finding may be explained in terms of demic diffusion, but this hypothesis leaves several questions to be answered;
5. Anatolian–Trans-Caucasus populations share sub-haplogroup U7 with Indians. This lineage cluster, easily identified by characteristic transversion at np 16,318, is virtually absent in Europeans, but relatively frequent in Indians (Kivisild *et al.* 1999b). Furthermore, the Trans-Caucasus and Anatolian populations possess mtDNA lineages deriving from the internal node R", but do not classify as any of the common haplogroups such as B, F, H, J, K, T, U or V. Frequencies of such lineages are low, but the lineages overlap with those which are frequent in Indian populations, suggesting some limited westward migration of people from southeastern Asia;
6. The presence of only about 4 per cent predominantly Mongoloid-type haplogroup M lineages in the Anatolian mtDNA pool and their absence among Armenians suggests a limited maternal lineage gene flow from the original Altaic language group people to present-day Turkey. This finding is in sharp contrast to the Central-Asian Turkish-speaking populations, where such variants make up about a half of all mtDNA lineages.

In conclusion one may say that except haplogroup V, the European mtDNA haplogroups are a representative subset of the set of haplogroups present in Anatolian and the Trans-Caucasus area populations. In addition, the mtDNA pool of the latter possesses lineages shared by Indians and Central Asians. Occasional additional lineages (e.g. varieties of L) are present, in this geographic area, in frequencies much below polymorphic frequencies. Some other region-

specific sub-clusters such as predominantly northern African U6 and M1 — 'an African M' — rare in Mediterranean Europe, are equally absent or rare in Anatolian populations. We think that this excludes models which predict recent substantial flow of maternal lineages from northern Africa and/or Ethiopia to western Eurasia. On the other hand, low frequency of haplogroup H in Ethiopians tells us that any backwards migration must have been either very selective or predating the spread of this major haplogroup in western Asia.

Unless new and significantly earlier archaeological dates for the presence of modern humans in Anatolia, Mesopotamia and Iran emerge, it seems to us that the occupation of western Eurasia by these modern humans took place 40,00045,000 years ago, covering most of the area, including possibly Scandinavia and other parts of northern Europe, which were subsequently under ice during the LGM. Perhaps the glaciers pushed these early Nordic humans south because south became north: much of the faunal material from the southwest French refugium is reindeer. It is quite likely that during global warming after the end of the LGM reindeer gradually returned north as did the people who hunted them. It is not migration in its typical later sense, creating gene frequency gradients because of admixture with locals: there were no northern locals present. Random drift (including bottlenecks) was probably the main force adjusting mtDNA haplogroup frequencies, but it worked only by reducing possible initial diversity. Proto-Saamis carried their maternal lineages back to the north and this is why we can see that more than 90 per cent of Saami mtDNA lineages are typically European, consisting predominantly of one very old branch, U5, and of another, much more recent branch, V, which is nevertheless a phylogenetic sister group to that most common among western Eurasians-haplogroup H. And because of that, the identity rather than relative frequencies of particular haplogroups in the Saami mtDNA pool is an important issue to consider.

Molecular genetics is only starting to contribute towards the understanding of the reconstruction of demographic history of modern human populations. As far as maternal lineages are concerned, further priorities as we can see them at present include more refined analysis of the topology of the mtDNA tree (additional markers in the coding region) and, of course, covering of hitherto un- and under-sampled regions, specifically Afghanistan, Iran, Iraq and Syria in the 'southern front' and the Ukraine and the steppe belt between Black Sea and Altai Mountains. Parallel studies of paternal lineages and autosomal markers are essential since the information obtained from the mtDNA alone, however valuable it may be, is only a part of the full genetic picture that must be digested together with archaeological, paleoclimatic and linguistic data.

Acknowledgements

We are grateful to Dr A. Torosjan for his help in collecting the Trans-Caucasian samples. We thank Dr P. Rudan for providing Croatian samples, Dr I. Mikerezi for collecting Albanian samples, Dr V. Ferak for providing Czech and Slovakian samples, Dr M. Claustres for the French samples and Dr S. Koziel for Polish samples; analyses of these data will be published in full length elsewhere. We thank Dr M. Stoneking and Dr S. Meyer for sending the Nile Valley sequences. We thank I. Hilpus and J. Lind for technical assistance.

This work was supported by Citrina Foundation UK and by Estonian Science Fund grants. R.V. wishes to thank Wellcome Trust for an award to participate in Human Genome Diversity meeting, September 1999.

References

Anderson, S., A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden & I.G. Young, 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290,457-65.

Bertranpetit, J.,J. Sala, F. Calafell, P.A. Underhill, P. Moral & D. Comas, 1995. Human mitochondrial DNA variation and the origin of Basques. *Annals of Human Genetics* 59, 63–81.

Calafell, F., P. Underhill, A. Tolun, D. Angelicheva & L. Kalaydieva, 1996. From Asia to Europe: mitochondrial DNA variability in Bulgarians and Turks. *Annals of Human Genetics* 60, 35-49.

Cavalli-Sforza, L.L., P. Menozzi & A. Piazza, 1994. *The History and Geography of Human Genes.* Princeton (NJ):Princeton University Press.

Chen, Y.S., A. Torroni, L. Excoffier, AS. Santachiara-Benerecetti & D.C. Wallace, 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *American Journal of Human Genetics* 57, 133-49.

Chikhi, L., G. Destro-Bisol, G. Bertorelle, V. Pascali & G. Barbujani, 1998. Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proceedings of the National Academy of Sciences of the USA* 95, 9053–8.

Comas D., F. Calafell, E. Mateu, A. Perez-Lezaun & J.

Bertranpetit, 1996. Geographic variation of human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations. *Molecular Biology and Evolution* 13, 1067-77.

Comas, D., F. Calafell, E. Mateu, A. Perez-Lezuan, E. Bosch, R. Martinez-Arias, J. Clarimon, F. Facchini, G. Fiori, D. Luiselli, D. Pettener & J. Bertranpetit, 1998. Trading genes along the Silk Road: mitochondrial DNA sequences and the origin of Central Asian populations. *American Journal of Human Genetics* 63, 1824–38.

Corte-Real, H.B.S.M., V.A. Macaulay, M.B. Richards, G. Hariti, M.S. Issad, A. Chambon-Thomsen, S. Papiha, J. Bertranpetit & B.C. Sykes, 1996. Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Annals of Human Genetics* 60,331–50.

Deka, R. & S.S. Papiha (eds.), 1999. *Genome Diversity*. New York (NY): Kluwer Academic/ Plenum Publishers.

Delghandi, M., E. Utsi & S. Krauss, 1998. Saami mitochondrial DNA reveals deep maternal lineage clusters. *Human Heredity* 48, 108-14.

Di Rienzo, A. & A.C. Wilson, 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA* 88,1567–601.

Dupuy, B.M. & B. Olaisen, 1996. MtDNA sequences in the Norwegian Saami and main populations, in *Advances in Forensic Haematogenetics,* vol. 6, eds. A. Carracedo, B. Brinkmann & W. Bar. Berlin: Springer-Verlag, 23-5.

Gamkrelidze T.V. & V.V. Ivanov, 1984. *Indo-European Language and Indo-Europeans.* Tbilisi: Tbilisi University Press.

Kivisild, T., M.J. Bamshad, K. Kaldma, M. Metspalu, E. Metspalu, M. Reidla, S. Laos, J. Parik, W.S. Watkins, M.E. Dixon, S.S. Papiha, S.S. Mastana, M.R. Mir, V. Ferak & R. Villems, 1999a. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Current Biology* 9, 13314.

Kivisild T., K. Kaldma, M. Metspalu, J. Parik, S.S. Papiha & R. Villems, 1999b. The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the old world, in Deka & Papiha (eds.), 135-52.

Krings, M., A.H. Salem, K. Bauer, H. Geisert, A.K. Malek, L. Chaix, C. Simon, D. Welsby, A. Di Rienzo, G. Utermann, A. Sajantila, S. Paabo & M. Stoneking, 1999. MtDNA analysis of Nile Valley populations: a genetic corridor or a barrier to migration. *American Journal of Human Genetics* 64, 1166-76.

Lahermo, P., A. Sajantila, P. Sistonen, M. Lukka, P. Aula, L. Peltonen & M-L. Savontaus, 1996. The genetic relatonship between the Finns and Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. *American Journal of Human Genetics* 58, 1309–22.

Lutz, S., H.J. Weisser & J.P. Heizmann, 1998. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. *International Journal of Legal Medicine* 111, 67–77.

Macaulay, V.A., M.B. Richards, E. Hickey, E. Vega, F. Cruciani, V. Guida, R. Scozzari, B. Bonne-Tamir, B. Sykes & A. Torroni, 1999. The emerging tree of the West Eurasian mtDNAs: a synthesis of control region sequences and RFLPs. *American Journal of Human Genetics* 64, 232-49.

Metspalu, E., T. Kivisild, K. Kaldma, J. Parik, M. Reidla, K. Tambets & R. Villems, 1999. The Trans-Caucasus and the expansion of the Causasoid-specific human mtDNA kineages, in Deka & Papiha (eds.), 121–34.

Orekhov, V., A. Poltoraus, L.A. Zhivotovsky, P. Ivanov & N. Yankovsky, 1999. Mitochondrial DNA sequence diversity in Russian. *FEBS Letters* 19, 197-201.

Passarino, G., **O.** Semino, L. Quintana-Murci, L. Excoffier, M. Hammer & AS. Santachiara-Benerecetti, 1998. Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *American Journal of Human Genetics* 62,420-34.

Pfeiffer, H., B. Brinkman, J. Hiihne, B. Rolf, A.A. Morris, R. Steigner, M.M. Holland & P. Forster, 1999. Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. *International Journal of Legal Medicine* 112, 291–8.

Pult, I., A. Sajantila., J. Simanainen, **O.** Georgiev, W. Schaffner & S. Pääbo, 1994. Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biological Chemistry Hoppe-Seyler* 375, 83740.

Quintana-Murci, L., O .Semino, H-J. Bandelt, G. Passarino, K. McElreavey & A.S. Santachiara-Benerecetti, 1999. Genetic evidence for an exit of *Homo sapiens* from Africa vie East-Africa. *Nature Genetics* 23, 43741.

Rando, J.C., F. Pinto, A.M. Gonzalez, M. Hernandez, J.M. Larruga, V.M. Cabrera & H-J. Bandelt, 1998. Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, near-eastern and sub-Saharan populations. *Annals of Human Genetics* 62, 531-50.

Rando, J.C., V.M. Cabrera, J.M. Larruga, M. Hernandez, A.M. Gonzalez, F. Pinto & H-J. Bandelt, in press. Phylogeographic patterns of mtDNA reflecting the colonisation of the Canary Islands.

Richards, M.B., H. Corte-Real, P. Forster, V. Macaulay, H. Wilkinson-Herbots, A. Demaine, S. Papiha, R. Hedges, H-J. Bandelt & B. Sykes, 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics* 59,185-203.

Richards, M.B., V.A. Macaulay, H-J. Bandelt & B.C. Sykes, 1998. Phylogeography of mitochondrial DNA in western Europe. *Annals of Human Genetics* 325,241-60.

Rybakov, B.A., R.M. Munchayev, V.A. Bashilov & P.G. Gaidukov (eds.), 1984. *Archaeology of the USSR: Paleolithic of USSR*. Moscow: Nauka. [In Russian.]

Rybakov, B.A., R.M. Munchayev, V.A. Bashilov & P.G. Gaidukov (eds.), 1989. *Archaeology of the USSR:*

Mesolithic of USSR. Moscow: Nauka. [In Russian.]

Sajantila, A., P. Lahermo, T. Lukka, P. Sistonen, M-L. Savontaus, P. Aula, L. Beckman, L. Tranebjaerg, T. Gedde-Dahl, L. Issel-Tarvel, A. DiRienzo & S. Pääbo, 1995. Genes and languages in Europe: an analysis of mitochondria1 lineages. *Genome Research* 5, 42–52.

Simoni, L., F. Calafell, D. Pettener, J. Bertranpetit & G. Barbujani, 2000. Geographic patterns of mtDNA diversity in Europe. *American Journal of Human Genetics* 66, 262–78. [Electronically published in *American Journal of Human Genetics* December 1, 1999.1

Torroni, **A.,** T.G. Schurr, M.F. Cabell, M.D. Brown, J.V. Neel, M. Larsen, C.M. Vullo & D.C. Wallace, 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *American Journal of Human Genetics* 53, 563–90.

Torroni, A., M.T. Lott, M.F. Cabell, Y-S. Chen, L. Lavergne & D.C. Wallace, 1994. MtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *American Journal of Human Genetics* 55, 760–76.

Torroni, A., K. Huoponen, P. Francalacci, M. Petrozzi, L. Morelli, R. Scozzari, D. Obidu, M-L. Savontaus & D.C. Wallace, 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144, 1835–50.

Torroni, A,, M. Petrozzi, L. D'Urbino, D. Sellitto, M. Zaviani, F. Carrara, C. Carducci, V. Leuzzi, V. Carelli, P. Barboni, A. De Negri & R. Scozzari, 1997. Haplotype and phylogenetic analysis suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *American Journal of Human Genetics* 60, 1107–21.

Torroni, A., H-J. Bandelt, L. D'Urbino, P. Lahermo, P. Moral, D. Sellitto, C. Rengo, P. Forster, M-L. Savontaus, B. Bonne-Tamir & R. Scozzari, 1998. mtDNA analysis reveals a major Late Paleolithic population expansion from southwestern to northeastern Europe. *American Journal of Human Genetics* 62, 1137-52.

Villems, R., M. Adojaan, T. Kivisild, E. Metspalu, J. Parik, G. Pielberg, S. Rootsi, K. Tambets & H-V. Tolk, 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,* vol. *l,* eds. K. Julku & K. Wiik. Jyvaskyla: Gummerus Kirjapaino Oy, 180–200.