

Complex Signals for Population Expansions in Europe and Beyond

Kristiina Tambets, Helle-Viivi Tolk, Toomas Kivisild, Ene Metspalu,
Jüri Parik, Maere Reidla, Michael Voevoda, Larissa Damba,
Marina Bermisheva, Elsa Khusnutdinova, Maria Golubenko,
Vadim Stepanov, Valery Puzyrev, Esien Usanga, Pavao Rudan,
Lars Beckmann & Richard Villems

François Jacob, in his brilliant 'The possible and the actual' (Jacob 1982), reminds us that 'scientific investigation begins by inventing a possible world, or a small piece of a possible world'. One may add that 1 Tz/F0 57O1Tfe piec ad

al. 1987; Vigilant *et al.* 1991), the last decade has demonstrated an increasingly better understanding of the phylogeny and phylogeography of mtDNA and of the Y chromosome. Here, the first influential achievement was a series of papers from Emory (reviewed in Wallace 1995) where, *inter alia*, it became obvious that human maternal lineages world-wide are very clearly structured geographically. This s coming from RFLP

analysis or the HVR sequence(s) alone were not informative enough to go further. Quite the opposite; it became clear that trees, based on HVR 1 sequence alone, were often phylogenetically wrong. However,

a synthesis of what is known about polymorphisms in the coding region (extensive RFLP as a tool) and HVR (direct sequencing) removes most of the ambiguities and leads to a much better understanding of the details of the topology of the phylogenetic tree of mtDNA (e.g. Macaulay *et al.* 1999). This analysis owes much to the use of median networks as an approach (Bandelt *et al.* 1995).

In this contribution we demonstrateh provides new insi

graphic processes of the past and, in particular, allows to see informative differences there, where mere haplogroup frequency calculations are able only to register flat landscapes.

General

How much further can one go in resolution? It is obvious that 'the ultimate' answer lies in analyzing, in all collected samples, all 16,500 plus nucleotides of the mtDNA genome — to carry out total (high fidelity!) re-sequencing. There are now at least a thousand fully sequenced mtDNA genomes at hand and this body of data, although rather

Year	Population	Year	Population
1980	100	1990	100
1981	100	1991	100
1982	100	1992	100
1983	100	1993	100
1984	100	1994	100
1985	100	1995	100
1986	100	1996	100
1987	100	1997	100
1988	100	1998	100
1989	100	1999	100
1990	100	2000	100
1991	100	2001	100
1992	100	2002	100
1993	100	2003	100
1994	100	2004	100
1995	100	2005	100
1996	100	2006	100
1997	100	2007	100
1998	100	2008	100
1999	100	2009	100
2000	100	2010	100
2001	100	2011	100
2002	100	2012	100
2003	100	2013	100
2004	100	2014	100
2005	100	2015	100
2006	100	2016	100
2007	100	2017	100
2008	100	2018	100
2009	100	2019	100
2010	100	2020	100
2011	100	2021	100
2012	100	2022	100
2013	100	2023	100
2014	100	2024	100
2015	100	2025	100
2016	100	2026	100
2017	100	2027	100
2018	100	2028	100
2019	100	2029	100
2020	100	2030	100

frequency exceeds that of

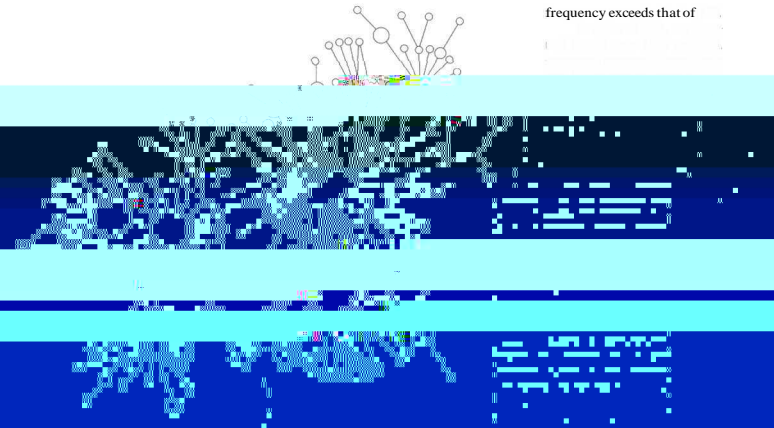


Figure 35.1. Skeletal topology of the human mtDNA haplogroup U5 hypervariable 1 phylogenetic tree for western Eurasia. Circle sizes are proportional to numbers of individuals per haplotype. 9 corresponds to a haplotype, most frequent among Saami population (HVR-1 motif 16,141; 16,189; 16,270).

with a few interesting exceptions, more frequent in eastern Europe and is either absent or very rare in the Near East and elsewhere. In the European north, an interesting exception is the Saami mtDNA pool, where U4 is virtually absent.

We have constructed a HVRI-based phylogenetic tree for U4, using information from 780 populations comprising a total of ~400 U4 genomes (Fig. 35.2). The topology of the U4 cluster is relatively simple, revealing the presence of a limited number of sub-founders. Of these, U4a and U4b are likely monophyletic, while U4c, determined by a transition at np 16,362, might be polyphyletic, at least in a pan-western Eurasian context. The highest frequencies of U4 (both in absolute terms and as a percentage of Hg U) can be observed actually not in Europe, but among Obi-Ugric Khanty and Mansi, living in northwestern Siberia. It is also frequent among the Finnic-speaking populations and in Volga Basin Turkic speakers, where, in some instances, its

U4a and U4b for the Germanic-speaking people (Germans, Norwegians, Swedes, Icelanders, Scots, German-speaking Swiss). Here the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other

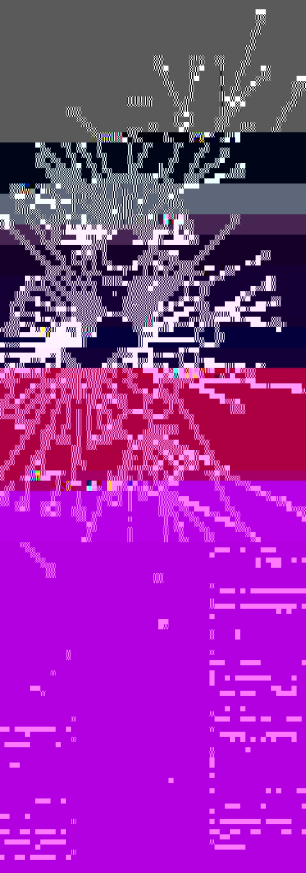
branches are for the other

branches are for the other

branches are for the other

branches are for the other

branches are for the other



branches are for the other

branches are for the other

branches are for the other

with more recent events is that they might well be

the corresponding tree (at least for western Eurasia) is concerned, and that the problems identified here will be largely solved, or shown to remain ambiguous forever because of built-in limitations resulting from the length of mtDNA.

What is much less clear is how we can reach significantly better temporal resolutions. Take, for example, U5: a cluster coalescing around 40,000-50,000 BP but consisting, as we interpret it now, of a number of sub-founders coalescing about 12,000 BP. Even though U5 mtDNAs are frequent in the western Eurasian mtDNA pool, to identify numerous sub-clusters within it one does need to operate with large sample sizes. For less frequent mtDNA varieties, only very large data bases, consisting of data about tens of thousands of mtDNAs, will allow a detailed temporal analysis, in particular for a time

- S. Santachiara-Benerecetti, R. Scozzari, A. Torroni & H.-J. Bandelt, 2000. Tracing European founder lineages in the near eastern mtDNA pool. *American Journal of Human Genetics* 67, 1251-76.
- Saillard, J. P., Forster, N., Lynnerup, H.-J. Bandelt & S. Norby, 2000. MtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *American Journal of Human Genetics* 67(3), 718-26.
- Semino, O., G. Passarino, P.J. Oefner, A.A. Lin, S.



Examining the farming/ language dispersal hypothesis

Edited by Peter Bellwood & Colin Renfrew

