

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 demonstrateth 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

日期	姓名	性别	年龄	职业	住址	联系电话	备注
2018.12.01	张三	男	35	教师	北京市海淀区	13800138000	
2018.12.02	李四	女	28	程序员	上海市浦东新区	15900159000	
2018.12.03	王五	男	45	医生	广州市天河区	13900139000	
2018.12.04	赵六	女	30	设计师	深圳市南山区	15800158000	
2018.12.05	孙七	男	50	工程师	北京市昌平区	13700137000	
2018.12.06	周八	女	25	销售	浙江省杭州市	15700157000	
2018.12.07	吴九	男	40	会计师	江苏省南京市	13600136000	
2018.12.08	郑十	女	38	律师	广东省深圳市福田区	15600156000	
2018.12.09	冯十一	男	22	学生	四川省成都市	13500135000	
2018.12.10	陈十二	女	48	公务员	河南省郑州市	13400134000	
2018.12.11	林十三	男	32	作家	安徽省合肥市	13300133000	
2018.12.12	周十四	女	27	歌手	湖北省武汉市	13200132000	
2018.12.13	吴十五	男	55	教授	湖南省长沙市	13100131000	
2018.12.14	郑十六	女	33	记者	江西省南昌市	13000130000	
2018.12.15	冯十七	男	29	画家	山东省济南市	12900129000	
2018.12.16	陈十八	女	42	翻译	福建省厦门市	12800128000	
2018.12.17	林十九	男	37	科学家	广东省广州市	12700127000	
2018.12.18	周二十	女	24	模特	浙江省杭州市	12600126000	
2018.12.19	吴二十一	男	52	企业家	江苏省南京市	12500125000	
2018.12.20	郑二十二	女	36	主持人	广东省广州市	12400124000	
2018.12.21	冯二十三	男	21	运动员	四川省成都市	12300123000	
2018.12.22	陈二十四	女	46	法官	河南省郑州市	12200122000	
2018.12.23	林二十五	男	31	音乐家	安徽省合肥市	12100121000	
2018.12.24	周二十六	女	26	舞蹈家	湖北省武汉市	12000120000	
2018.12.25	吴二十七	男	51	哲学家	湖南省长沙市	11900119000	
2018.12.26	郑二十八	女	34	作家	江西省南昌市	11800118000	
2018.12.27	冯二十九	男	23	程序员	山东省济南市	11700117000	
2018.12.28	陈三十	女	44	翻译	福建省厦门市	11600116000	
2018.12.29	林三十一	男	39	科学家	广东省广州市	11500115000	
2018.12.30	周三十二	女	29	模特	浙江省杭州市	11400114000	
2018.12.31	吴三十三	男	54	企业家	江苏省南京市	11300113000	

frequency exceeds that of

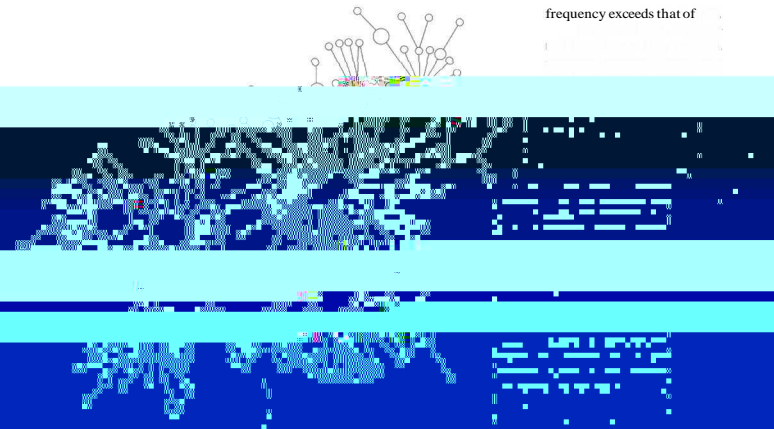


Figure 35.1. Skeletal topology of the human mtDNA haplogroup U5 hypervariable 1 phylogenetic tree for western Eurasia. Circle sizes are proportional to numbers of individuals per haplotype. 9 corresponds to a haplotype, most frequent among Saami population (HVR-1 motif 16,141; 16,189; 16,270).

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

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.

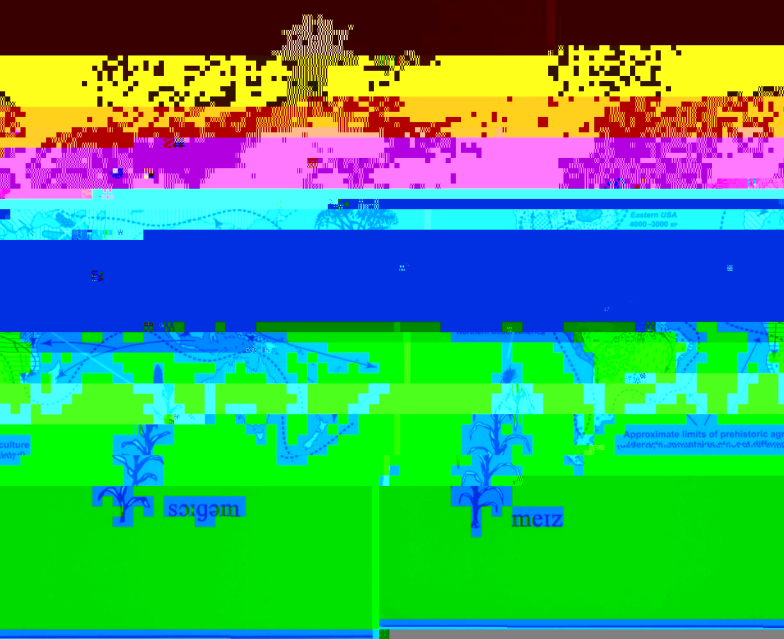
What is much less clear is how we can reach significantly better temporal resolutions. Take, for example, U5: a cluster coalescing around 40,000-50,000 BP but consisting, as we interpret it now, of a number of sub-founders coalescing about 12,000 BP. Even though U5 mtDNAs are frequent in the western 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



£4

Eastern USA
4000-3000 BP

Culture
area

Approximate limits of prehistoric agricultural dispersal, cultural differences