HUMAN MUTATION Mutation in Brief #1039, 30:E386-E394, (2008) Online

MUTATION IN BRIEF

Updated Comprehensive Phylogenetic Tree of Global Human Mitochondrial DNA Variation

Mannis van Oven* and Manfred Kayser

Department of Forensic Molecular Biology, Erasmus University Medical Center Rotterdam, The Netherlands

*Correspondence to Mannis van Oven, Department of Forensic Molecular Biology, Erasmus University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands; E-mail: m.vanoven@erasmusmc.nl

Contract grant sponsor: Netherlands Forensic Institute

Communicated by Peter J. Oefner

Human mitochondrial DNA is widely used as tool in many fields including evolutionary anthropology and population history, medical genetics, genetic genealogy, and forensic science. Many applications require detailed knowledge about the phylogenetic relationship of mtDNA variants. Although the phylogenetic resolution of global human mtDNA diversity has greatly improved as a result of increasing sequencing efforts of complete mtDNA genomes, an updated overall mtDNA tree is currently not available. In order to facilitate a better use of known mtDNA variation, we have constructed an updated comprehensive phylogeny of global human mtDNA variation, based on both coding- and control region mutations. This complete mtDNA tree includes previously published as well as newly identified haplogroups, is easily navigable, will be continuously and regularly updated in the future, and is online available at http://www.phylotree.org. © 2008 Wiley-Liss, Inc.

KEY WORDS: mitochondrial DNA, phylogenetic tree, haplogroup, mtDNA variation, human evolution

INTRODUCTION

Uniparentally inherited portions of the human genome have important applications in a wide range of fields including evolutionary anthropology and population history (Jobling and Tyler-Smith, 2003; Torroni et al., 2006; Underhill and Kivisild, 2007), medical genetics (Krausz et al., 2004; Taylor and Turnbull, 2005), genetic genealogy (Blansit, 2006; Johnston and Thomas, 2003; Shriver and Kittles, 2004) as well as forensic science (Kayser, 2007). Recently, an updated haplogroup tree of the patrilineally inherited, non-recombining part of the human Y chromosome (NRY) was published (Karafet et al., 2008). Here, we present an updated phylogenetic tree of its matrilineal counterpart, the mitochondrial (mt)DNA.

PROPERTIES OF mtDNA

MtDNA represents the small fraction of our genome that does not reside in the cell's nucleus, but is contained in the energy-producing mitochondria. It is a circular, double-stranded molecule with a length of nearly 16.6 kilobase pairs that encodes proteins required for oxidative phosphorylation. Special features of mtDNA include matrilineal inheritance, lack of recombination, high copy number, and a higher evolutionary turn over rate as compared to nuclear DNA. It is assumed that all mtDNA types in the human gene pool can ultimately be traced

Received 23 July 2008; accepted revised manuscript 11 September 2008.

DOI: 10.1002/humu.20921

back to a common matrilineal ancestor that lived approximately 200,000 years ago in Africa (Behar et al., 2008b; Macaulay et al., 2005; Mishmar et al., 2003). MtDNA sequence variation thus evolved as a result of the sequential accumulation of mutations along maternally inherited lineages, which can be represented in a tree reflecting the phylogenetic relationships of known mtDNA variants. The importance of a detailed global mtDNA phylogeny has been addressed by several authors (Bandelt et al., 2005, 2006; Kivisild et al., 2006a; Kong et al., 2006; Richards, 2004; Salas et al., 2007; Torroni et al., 2006). Recently, a global mtDNA phylogeny based on 2959 coding region sequences was published as part of the MITOMAP database (Ruiz-Pesini et al., 2007). However, this tree does not include the commonly used control region mutations and does not contain a detailed haplogroup nomenclature. Moreover, the MITOMAP tree must be considered outdated by now since many complete mtDNA sequences that have considerably improved the resolution to the mtDNA phylogeny became available only after its publication. For instance, Behar et al. (2008b) recently reported 309 complete mtDNA sequences that (combined with 315 previously published sequences) greatly increased the resolution of haplogroup L. At present (August 27, 2008), 4198 complete mtDNA genome sequences and additional 945 separate coding region sequences are available via the NCBI GenBank database (http://www.ncbi.nlm.nih.gov/GenBank). In order to facilitate a better use of the known mtDNA variation, we generated an updated phylogenetic tree of global human mtDNA variation considering all currently available sequence data.

CODING VERSUS CONTROL REGION

By convention, the nucleotide positions of the human mtDNA genome are numbered from 1 to 16569 according to the revised Cambridge Reference Sequence (rCRS) (Anderson et al., 1981; Andrews et al., 1999). Detected polymorphisms (single nucleotide polymorphisms [SNPs] and insertions/deletion polymorphisms [indels]) are then scored as differences ("mutations") to the rCRS. The mutation rate is not equal across the entire mtDNA molecule; the overall mutation rate in the 1.1 kb non-coding control region (bases 16024-576) is about 10 times higher than that of the 15.5 kb coding region (bases 577-16023) (Howell et al., 2007; Pakendorf and Stoneking, 2005). This can be explained by purifying selection filtering out from the population the often deleterious functional mutations. In contrast, the control region, which is important for mtDNA replication, is non-coding and thus accumulates more mutations. Moreover, local differences in the mutation rate among nucleotide positions within the control region exist as well (Hasegawa et al., 1993; Meyer et al., 1999). Some individual sites act as mutational hot spots (e.g. positions 146, 150, 152, 195, 16189, 16311, 16362, 16519), whereas others appear rather stable (e.g. positions 477, 493, 16108, 16219). Also within the coding region there are sites that are more prone to mutation than others (e.g. positions 709, 1719, 3010, 5460, 10398, 11914, 13105, 13708, 15884), causing homoplasy (recurrence) in the phylogeny. Although several authors have addressed the issue of site-specific (hyper)mutability (Galtier et al., 2006; Howell et al., 2007; Ingman and Gyllensten, 2007; Kivisild et al., 2006b; Malyarchuk et al., 2002), the exact mechanism accounting for it remains to be elucidated. Due to the higher overall mutation rate, the control region is relatively enriched in sequence variation, and therefore researchers typically sequence (part of) this region for various applications. However, at least some haplogroups can not be confidently assigned based on control region data only (e.g. H4, V), whereas for others (e.g. J1b1, K1a9) the accurate inference solely from control region sequence data is feasible. Therefore it is relevant to confirm predicted haplogroups based on control region sequences by use of informative SNPs from the coding region and/or to perform initial mtDNA haplogroup assignment using coding region information either solely or combined with control region data. However, ascertaining informative coding region SNPs strongly depends on the availability and quality of phylogenetic information, hence requires an updated phylogenetic global mtDNA tree.

HAPLOGROUP NOMENCLATURE

The first mtDNA haplogroups, discovered in Native Americans, were baptized A, B, C, and D (Torroni et al., 1993). Subsequently detected haplogroups were designated using other letters of the alphabet. By now, all letters of the alphabet, except O (although once proposed), have been used (Figure 1). Simple rules for mtDNA haplogroup nomenclature were proposed by Richards et al. (1998) (see also Kivisild et al., 2006a); but unfortunately, the mtDNA haplogroup nomenclature is not always used consistently. Several examples can be found where different authors coined the same name for different haplogroups: e.g. "M12" in Kong et al. (2006) versus Metspalu et al. (2006), Tanaka et al. (2004); "R1" in Malyarchuk et al. (2008a), Palanichamy et al. (2004) versus Metspalu et al. (2006); "S" in Hudjashov et al. (2007), Palanichamy et al. (2004) versus Starikovskaya et al.

(2005). Others used different names for the same haplogroup: "N12" in Hudjashov et al. (2007) versus "O" in Palanichamy et al. (2004), Pierson et al. (2006); M1 sublineages in Olivieri et al. (2006) versus Gonzalez et al. (2007); "C4" in Achilli et al. (2008), Kong et al. (2006) versus "C2" in Starikovskaya et al. (2005), Volodko et al. (2008). All these examples illustrate the need for an overall mtDNA phylogeny with universal nomenclature but also their consistent use by the mtDNA community.

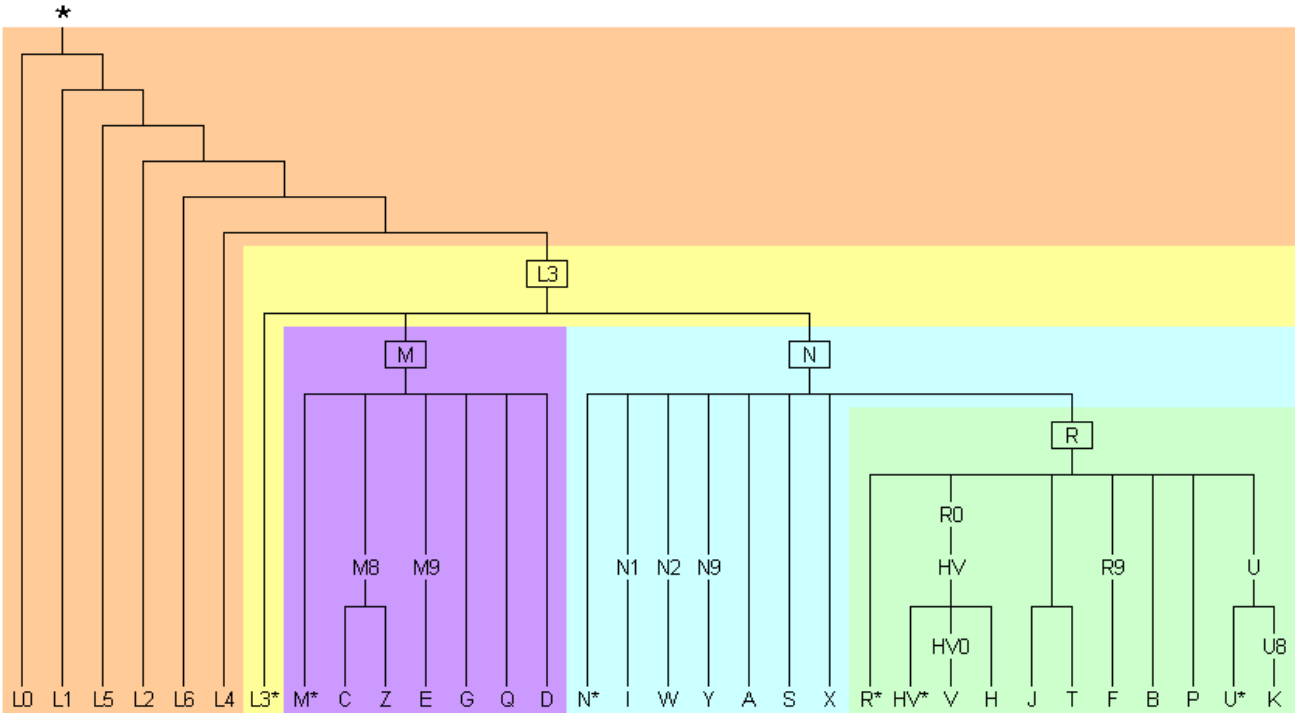

Figure 1. Simplified mtDNA phylogeny illustrating the use of alphabetical letters for haplogroup designation. All letters of the alphabet, except O (Palanichamy et al. [2004] proposed this label for a haplogroup that was later relabeled N12 by Hudjashov et al. [2007]), have been used so far. The root of the tree is indicated by a star representing the most recent common matrilineal ancestor of all humans. The L haplogroups are the most deep-rooting lineages and are African specific indicating the African origin of modern humans as well as the out-of-Africa exodus as also found based on other genetic as well as fossil data. Haplogroup L3 gave rise to macrohaplogroups M, N and R (the latter itself a subclade of N), which encompass all variation observed outside Africa. Nomenclature evolved in such a way that letters C, D, E, G, Q, and Z designate lineages belonging to M; letters A, I, S, W, X, and Y lineages within N; and B, F, HV, H, J, K, P, T, U, and V lineages within R. Haplogroup symbols followed by a star represent all other descendant lineages (besides the ones shown) of a particular clade, for which no unique alphabetical letters were reserved; e.g. N* stands for N5, N12, N13, N14, N21 and N22. Note that this tree should not be seen as a summary of global mtDNA variation (for this see the tree in the supporting information or at http://www.phylotree.org); it is merely meant to show the topology of alphabetically named haplogroups.

UPDATED COMPLETE mtDNA TREE

We have consulted 55 publications that have recently described mtDNA variation based on complete mtDNA sequences to construct an updated and complete (as per August 27, 2008) global mtDNA phylogenetic tree (the tree's references are listed in the main reference list of this article and are as separate list also available at http://www.phylotree.org). In cases of conflicting information from different papers concerning the same haplogroup, we have drawn the phylogeny ourselves by comparing the maximum number of available mtDNA sequences of that particular haplogroup, in order to get the full up-to-date picture. In addition, we made use of other recently published complete mtDNA sequences (FamilyTreeDNA; Gasparre et al., 2007, Hartmann et al., 2008; Just et al., 2008) that were not previously incorporated in phylogenetic schemes (but were available at GenBank). In some cases this allowed us to further increase the phylogenetic resolution by adding novel subbranches to the tree (i.e. haplogroups L3b1b1, L3d5, Q3a1, M33c, M52, HV0c, H2a2b and H2a2b1). We used

as a general criterion for the whole tree that a relatively stable (set of) mutation(s) must be shared by at least three complete sequences before assigning it the haplogroup status. We deviated from this rule in the following cases: some authors have defined haplogroups based on hypervariable positions (e.g. M4, X1) or on incomplete mtDNA sequence data (e.g. M21d, R22) and in those cases we decided to follow the established nomenclature; when less than three complete sequences added a relatively deep-rooting branch to the tree we also found it justified to assign it a haplogroup name (e.g. L1c1a1b, M52). For clarity, all branches that are insufficiently supported by complete mtDNA sequence data were given a preliminary status by showing their defining mutations in italic. The final mutation motifs of these haplogroups will be further refined when additional data accumulates. In total, our tree encompasses 1267 branching points, of which 657 are terminal branches. In contrast to what has become common practice for drawing mtDNA phylogenetic trees so far, we decided to use horizontal instead of vertical lines representing the branches of the tree. Not only was such representation used for most NRY trees, including the most up-to-date one by Karafet et al. (2008), but the horizontal representation results in a more compact tree, which we find easier to follow. For each branch its defining mutations were mentioned as well as the haplogroup name (newly proposed haplogroups were indicated in red). Coding region mutations were indicated in black and control region mutations were in blue, so that it becomes immediately clear which information can be derived from which part of the mtDNA. Position numbers were relative to the rCRS and the scoring method followed the guidelines outlined by Bandelt and Parson (2008). At the tip of each branch we provided references to the papers that were consulted for the particular haplogroups. The references given are not necessarily the papers in which the haplogroup was first described, but rather the papers in which the more detailed phylogenies were provided. For each terminal branch we also provided one or two NCBI accession numbers (if available) of complete mtDNA sequences belonging to the respective branch. See Figure 2 for a snapshot view of the tree we constructed. The complete updated tree is available as Supporting Information to this paper and additionally freely available at http://www.phylotree.org. We plan to regularly update this tree and keep the further updated versions available via this website. The tree can be viewed with the most common internet browsers. A convenient way to look up a particular haplogroup or a particular mutation is to use the "Find" command of the web browsers. Internet Explorer has the advantage that one can select the option "Match whole word only". This is recommended to avoid unspecific hits in cases where a searched position number also occurs as part of other position numbers (e.g. to avoid finding "11719" when searching for "1719").

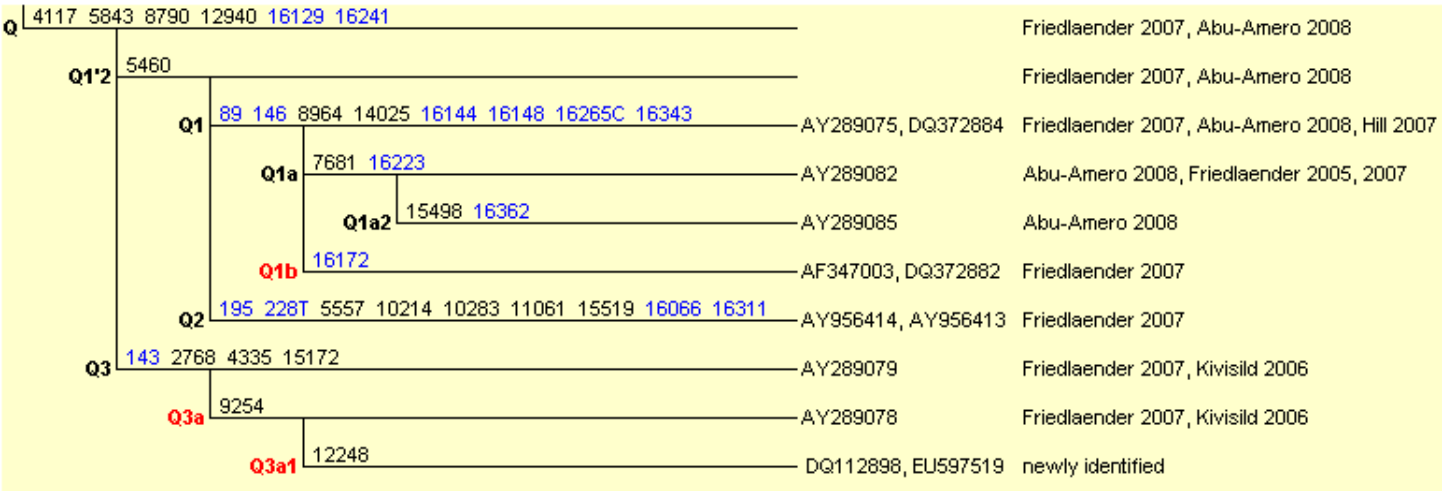

Figure 2. Snapshot of the haplogroup Q branch from the updated global mtDNA tree available as supplement of this paper as well as at http://www.phylotree.org. Coding region mutations were labeled in black, control region mutations were in blue. Mutations are transitions unless a specific base change was specified. Haplogroup names were given at the left of each branch. Haplogroups were indicated in black when previously published and in red when newly proposed by us. At the tip of each branch we provided the references that were consulted for the particular haplogroups, as well as GenBank accession numbers of underlying complete mtDNA sequences. In the snapshot example shown here we proposed a new haplogroup Q3a1 defined by a transition at position 12248, based on the fact that this mutation was shared by the sequences with accession numbers DQ112898 and EU597519.

CONCLUSIONS

We have constructed an updated, high-resolution global mtDNA phylogenetic tree based on both coding and control region data, including universal haplogroup nomenclature, and providing references to consulted papers as well as to accession numbers of underlying NCBI GenBank sequences. This tree is publicly available at http://www.phylotree.org, allows easy navigation via search functions of internet browsers and will be updated at least every six months on the website. This comprehensive mtDNA tree is meant as a framework for evolutionary anthropologists, population geneticists, medical geneticists, genealogists, forensic geneticists and other scientists interested in the description and application of human mtDNA diversity.

ACKNOWLEDGMENTS

We thank the numerous colleagues who generated mtDNA sequence data and the NCBI GenBank database to make them publicly available. This work was supported by the Netherlands Forensic Institute.

REFERENCES

Abu-Amero KK, Gonzalez AM, Larruga JM, Bosley TM, Cabrera VM. 2007. Eurasian and African mitochondrial DNA influences in the Saudi Arabian population. BMC Evol Biol 7:32.

Abu-Amero KK, Larruga JM, Cabrera VM, Gonzalez AM. 2008. Mitochondrial DNA structure in the Arabian Peninsula. BMC Evol Biol 8(1):45.

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet 75(5):910-918.

Achilli A, Rengo C, Battaglia V, Pala M, Olivieri A, Fornarino S, Magri C, Scozzari R, Babudri N, Santachiara-Benerecetti AS. 2005. Saami and Berbers—an unexpected mitochondrial DNA link. Am J Hum Genet 76(5):883-886.

Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ. 2008. The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PLoS ONE 3(3):e1764.

Álvarez-Iglesias V, Jaime JC, Carracedo Á, Salas A. 2007. Coding region mitochondrial DNA SNPs: targeting East Asian and Native American haplogroups. Forensic Science International: Genetics 1(1):44-55.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F. 1981. Sequence and organization of the human mitochondrial genome. Nature 290(5806):457-465.

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(2):147.

Bandelt HJ, Achilli A, Kong QP, Salas A, Lutz-Bonengel S, Sun C, Zhang YP, Torroni A, Yao YG. 2005. Low "penetrance" of phylogenetic knowledge in mitochondrial disease studies. Biochem Biophys Res Commun 333(1):122-130.

Bandelt HJ, Kong QP, Richards M, Macaulay V. 2006. Estimation of mutation rates and coalescence times: some caveats. In: Bandelt HJ, Macaulay V, Richards M, editors. Human Mitochondrial DNA and the Evolution of *Homo sapiens*. Berlin: Springer-Verlag. p 47–90.

Bandelt HJ, Parson W. 2008. Consistent treatment of length variants in the human mtDNA control region: a reappraisal. Int J Legal Med 122(1):11-21.

Barik SS, Sahani R, Prasad BV, Endicott P, Metspalu M, Sarkar BN, Bhattacharya S, Annapoorna PC, Sreenath J, Sun D. 2008. Detailed mtDNA genotypes permit a reassessment of the settlement and population structure of the Andaman Islands. Am J Phys Anthropol 136(1):19-27.

Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K. 2006. The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. Am J Hum Genet 78(3):487-497.

Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, Yudkovsky G, Rosengarten D, Pereira L, Amorim A. 2008a. Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. PLoS ONE 3(4).

Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D. 2008b. The dawn of human matrilineal diversity. Am J Hum Genet 82(5):1130-1140.

Blansit BD. 2006. Resources for genetic genealogy. Journal of Electronic Resources in Medical Libraries 3(2):23-35.

Carelli V, Achilli A, Valentino ML, Rengo C, Semino O, Pala M, Olivieri A, Mattiazzi M, Pallotti F, Carrara F. 2006. Haplogroup effects and recombination of mitochondrial DNA: novel clues from the analysis of Leber hereditary optic neuropathy pedigrees. Am J Hum Genet 78(4):564-574.

Chaubey G, Karmin M, Metspalu E, Metspalu M, Selvi-Rani D, Singh VK, Parik J, Solnik A, Naidu BP, Kumar A, Adarsh N, Mallick CB, Trivedi B, Prakash S, Reddy R, Shukla P, Bhagat S, Verma S, Vasnik S, Khan I, Barwa A, Sahoo D, Sharma A, Rashid M, Chandra V, Reddy AG, Torroni A, Foley RA, Thangaraj K, Singh L, Kivisild T, Villems R. 2008a. Phylogeography of mtDNA haplogroup R7 in the Indian peninsula. BMC Evol Biol 8(1):227.

Chaubey G, Metspalu M, Karmin M, Thangaraj K, Rootsi S, Parik J, Solnik A, Rani DS, KumarSingh V, Naidu BP. 2008b. Language shift by indigenous population: a model genetic study in South Asia. International Journal of Human Genetics 8(1-2):41-50.

Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Dambueva I, Perkova M, Dorzhu C, Luzina F, Lee HK, Vanecek T. 2007. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. Am J Hum Genet 81(5):1025-1041.

FamilyTreeDNA. Complete mtDNA sequences uploaded to GenBank by FamilyTreeDNA between 2006 and present.

Finnilä S, Lehtonen MS, Majamaa K. 2001. Phylogenetic network for European mtDNA. Am J Hum Genet 68(6):1475-1484.

Friedlaender J, Schurr T, Gentz F, Koki G, Friedlaender F, Horvat G, Babb P, Cerchio S, Kaestle F, Schanfield M. 2005. Expanding Southwest Pacific mitochondrial haplogroups P and Q. Mol Biol Evol 22(6):1506-1517.

Friedlaender JS, Friedlaender FR, Hodgson JA, Stoltz M, Koki G, Horvat G, Zhadanov S, Schurr TG, Merriwether DA. 2007. Melanesian mtDNA complexity. PLoS ONE 2(2):e248.

Galtier N, Enard D, Radondy Y, Bazin E, Belkhir K. 2006. Mutation hot spots in mammalian mitochondrial DNA. Genome Res 16(2):215-222.

Gasparre G, Porcelli AM, Bonora E, Pennisi LF, Toller M, Iommarini L, Ghelli A, Moretti M, Betts CM, Martinelli GN. 2007. Disruptive mitochondrial DNA mutations in complex I subunits are markers of oncocytic phenotype in thyroid tumors. Proc Natl Acad Sci US A 104(21):9001-9006.

Gonzalez AM, Garcia O, Larruga JM, Cabrera VM. 2006. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country. BMC Genomics 7:124.

Gonzalez AM, Larruga JM, Abu-Amero KK, Shi Y, Pestano J, Cabrera VM. 2007. Mitochondrial lineage M1 traces an early human backflow to Africa. BMC Genomics 8:223.

Grzybowski T, Malyarchuk BA, Derenko MV, Perkova MA, Bednarek J, Wozniak M. 2007. Complex interactions of the Eastern and Western Slavic populations with other European groups as revealed by mitochondrial DNA analysis. Forensic Science International: Genetics 1(2):141-147.

Hartmann A, Thieme M, Nanduri LK, Stempfl T, Moehle C, Kivisild T, Oefner PJ. 2008. Validation of microarray-based sequencing of 93 worldwide mitochondrial genomes. Hum Mutat 2008 Jul 11 Epub ahead of print.

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(4):347-354.

Hill C, Soares P, Mormina M, Macaulay V, Meehan W, Blackburn J, Clarke D, Raja JM, Ismail P, Bulbeck D. 2006. Phylogeography and ethnogenesis of aboriginal Southeast Asians. Mol Biol Evol 23(12):2480-2491.

Hill C, Soares P, Mormina M, Macaulay V, Clarke D, Blumbach PB, Vizuete-Forster M, Forster P, Bulbeck D, Oppenheimer S. 2007. A mitochondrial stratigraphy for Island Southeast Asia. Am J Hum Genet 80(1):29-43.

Howell N, Elson JL, Howell C, Turnbull DM. 2007. Relative rates of evolution in the coding and control regions of African mtDNAs. Mol Biol Evol 24(10):2213-2221.

Hudjashov G, Kivisild T, Underhill PA, Endicott P, Sanchez JJ, Lin AA, Shen P, Oefner P, Renfrew C, Villems R. 2007. Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. Proc Natl Acad Sci US A 104(21):8726-8730.

Ingman M, Gyllensten U. 2007. Rate variation between mitochondrial domains and adaptive evolution in humans. Hum Mol Genet 16(19):2281-2287.

Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4(8):598-612.

Johnston J, Thomas M. 2003. Summary: the science of genealogy by genetics. Dev World Bioeth 3(2):103-108.

Just RS, Diegoli TM, Saunier JL, Irwin JA, Parsons TJ. 2008. Complete mitochondrial genome sequences for 265 African American and US "Hispanic" individuals. Forensic Science International: Genetics 2(3):45-48.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome Res 18(5):830-838.

Kayser M. 2007. Uni-parental markers in human identity testing including forensic DNA analysis. Biotechniques 43(6):S16-S21.

Kivisild T, Metspalu M, Bandelt HJ, Richards M, Villems R. 2006a. The world mtDNA phylogeny. In: Bandelt HJ, Macaulay V, Richards M, editors. Human Mitochondrial DNA and the Evolution of *Homo sapiens*. Berlin: Springer-Verlag. p 149–179.

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A. 2006b. The role of selection in the evolution of human mitochondrial genomes. Genetics 172(1):373-387.

Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, Zhang YP. 2003. Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. Am J Hum Genet 73(3):671-676.

Kong QP, Yao YG, Sun C, Zhu CL, Zhong L, Wang CY, Cai WW, Xu XM, Xu AL, Zhang YP. 2004. Phylogeographic analysis of mitochondrial DNA haplogroup F2 in China reveals T12338C in the initiation codon of the ND5 gene not to be pathogenic. J Hum Genet 49(8):414-423.

Kong QP, Bandelt HJ, Sun C, Yao YG, Salas A, Achilli A, Wang CY, Zhong L, Zhu CL, Wu SF. 2006. Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations. Hum Mol Genet 15(13):2076-2086.

Krausz C, Quintana-Murci L, Forti G. 2004. Y chromosome polymorphisms in medicine. Ann Med 36(8):573-583.

Kumar S, Padmanabham PB, Ravuri RR, Uttaravalli K, Koneru P, Mukherjee PA, Das B, Kotal M, Xaviour D, Saheb SY, Rao VR. 2008. The earliest settlers' antiquity and evolutionary history of Indian populations: evidence from M2 mtDNA lineage. BMC Evol Biol 8(1):230.

Loogväli EL, Roostalu U, Malyarchuk BA, Derenko MV, Kivisild T, Metspalu E, Tambets K, Reidla M, Tolk HV, Parik J. 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. Mol Biol Evol 21(11):2012-2021.

Maca-Meyer N, González AM, Pestano J, Flores C, Larruga JM, Cabrera VM. 2003. Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography. BMC Genet 4:15.

Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. Science 308(5724):1034-1036.

Malyarchuk B, Grzybowski T, Derenko M, Perkova M, Vanecek T, Lazur J, Gomolcak P, Tsybovsky I. 2008a. Mitochondrial DNA phylogeny in Eastern and Western Slavs. Mol Biol Evol 25(8):1651-1658.

Malyarchuk BA, Rogozin IB, Berikov VB, Derenko MV. 2002. Analysis of phylogenetically reconstructed mutational spectra in human mitochondrial DNA control region. Hum Genet 111(1):46-53.

Malyarchuk BA, Derenko M, Perkova M, Grzybowski T, Vanecek T, Lazur J. 2008b. Reconstructing the phylogeny of African mitochondrial DNA lineages in Slavs. Eur J Hum Genet (advance online publication 9 April 2008; doi: 10.1038/ejhg.2008.70).

Malyarchuk BA, Perkova MA, Derenko MV, Vanecek T, Lazur J, Gomolcak P. 2008c. Mitochondrial DNA variability in Slovaks, with application to the Roma origin. Ann Hum Genet 72(2):228-240.

Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, Koki G, Friedlaender JS. 2005. Ancient mitochondrial M haplogroups identified in the Southwest Pacific. Proc Natl Acad Sci US A 102(37):13034-13039.

Metspalu M, Kivisild T, Bandelt HJ, Richards M, Villems R. 2006. The pioneer settlement of modern humans in Asia. In: Bandelt HJ, Macaulay V, Richards M, editors. Human mitochondrial DNA and the Evolution of *Homo sapiens*. Berlin: Springer-Verlag. p 181–199.

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(3):1103-1110.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD. 2003. Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci US A 100(1):171-176.

Olivieri A, Achilli A, Pala M, Battaglia V, Fornarino S, Al-Zahery N, Scozzari R, Cruciani F, Behar DM, Dugoujon JM. 2006. The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. Science 314(5806):1767-1770.

Pakendorf B, Stoneking M. 2005. Mitochondrial DNA and human evolution. Annu Rev Genomics Hum Genet 6:165-183.

Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V, Zhang YP. 2004. Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. Am J Hum Genet 75(6):966-978.

Pierson MJ, Martinez-Arias R, Holland BR, Gemmell NJ, Hurles ME, Penny D. 2006. Deciphering past human population movements in Oceania: provably optimal trees of 127 mtDNA genomes. Mol Biol Evol 23(10):1966-1975.

Rajkumar R, Banerjee J, Gunturi HB, Trivedi R, Kashyap VK. 2005. Phylogeny and antiquity of M macrohaplogroup inferred from complete mt DNA sequence of Indian specific lineages. BMC Evol Biol 5(1):26.

Reddy BM, Langstieh BT, Kumar V, Nagaraja T, Reddy AN, Meka A, Reddy AG, Thangaraj K, Singh L. 2007. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. PLoS ONE 2(11):e1141.

Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk HV, Parik J, Loogväli EL, Derenko M, Malyarchuk B. 2003. Origin and diffusion of mtDNA haplogroup X. Am J Hum Genet 73(5):1178-1190.

Richards M. 2004. The mitochondrial DNA tree and forensic science. International Congress Series 1261:91-93.

Richards MB, Macaulay VA, Bandelt HJ, Sykes BC. 1998. Phylogeography of mitochondrial DNA in western Europe. Ann Hum Genet 62(Pt 3):241-260.

Roostalu U, Kutuev I, Loogväli EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. 2007. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Mol Biol Evol 24(2):436-448.

Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Yi C, Kreuziger J, Baldi P, Wallace DC. 2007. An enhanced MITOMAP with a global mtDNA mutational phylogeny. Nucleic Acids Res 35 (Database issue):D823-D828.

Salas A, Bandelt HJ, Macaulay V, Richards MB. 2007. Phylogeographic investigations: the role of trees in forensic genetics. Forensic Sci Int 168(1):1-13.

Shen P, Lavi T, Kivisild T, Chou V, Sengun D, Gefel D, Shpirer I, Woolf E, Hillel J, Feldman MW. 2004. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. Hum Mutat 24(3):248-260.

Shlush LI, Behar DM, Yudkovsky G, Templeton A, Hadid Y, Basis F, Hammer M, Itzkovitz S, Skorecki K. 2008. The Druze: a population genetic refugium of the Near East. PLoS ONE 3(5).

Shriver MD, Kittles RA. 2004. Genetic ancestry and the search for personalized genetic histories. Nat Rev Genet 5(8):611-618.

Soares P, Trejaut JA, Loo JH, Hill C, Mormina M, Lee CL, Chen YM, Hudjashov G, Forster P, Macaulay V. 2008. Climate change and postglacial human dispersals in Southeast Asia. Mol Biol Evol 25(6):1209-1218.

Starikovskaya EB, Sukernik RI, Derbeneva OA, Volodko NV, Ruiz-Pesini E, Torroni A, Brown MD, Lott MT, Hosseini SH, Huoponen K. 2005. Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. Ann Hum Genet 69(Pt 1):67-89.

Sun C, Kong QP, Palanichamy MG, Agrawal S, Bandelt HJ, Yao YG, Khan F, Zhu CL, Chaudhuri TK, Zhang YP. 2006. The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. Mol Biol Evol 23(3):683-690.

Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK. 2007. Beringian standstill and spread of Native American founders. PLoS ONE 2(9):e829.

Tanaka M, Cabrera VM, González AM, Larruga JM, Takeyasu T, Fuku N, Guo LJ, Hirose R, Fujita Y, Kurata M. 2004. Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res 14(10a):1832-1850.

Taylor RW, Turnbull DM. 2005. Mitochondrial DNA mutations in human disease. Nat Rev Genet 6(5):389-402.

Thangaraj K, Chaubey G, Kivisild T, Selvi Rani D, Singh VK, Ismail T, Carvalho-Silva D, Metspalu M, Bhaskar LV, Reddy AG. 2008. Maternal footprints of Southeast Asians in North India. Hum Hered 66(1):1-9.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53(3):563-590.

Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. Am J Hum Genet 69(4):844-852.

Torroni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. Trends Genet 22(6):339-345.

Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, Hsu CJ, Lee ZY, Lin M. 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. PLoS Biol 3(8).

Underhill PA, Kivisild T. 2007. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. Annu Rev Genet 41:539-564.

Volodko NV, Starikovskaya EB, Mazunin IO, Eltsov NP, Naidenko PV, Wallace DC, Sukernik RI. 2008. Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am J Hum Genet 82(5):1084-1100.