CrossMark

ORIGINAL ARTICLE

Whole mitochondrial genome genetic diversity in an Estonian population sample

Monika Stoljarova[1,2] · Jonathan L. King[2] · Maiko Takahashi[2] · Anu Aaspõllu[1] · Bruce Budowle[2]

Abstract Mitochondrial DNA is a useful marker for population studies, human identification, and forensic analysis. Commonly used hypervariable regions I and II (HVI/HVII) were reported to contain as little as 25 % of mitochondrial DNA variants and therefore the majority of power of discrimination of mitochondrial DNA resides in the coding region. Massively parallel sequencing technology enables entire mitochondrial genome sequencing. In this study, buccal swabs were collected from 114 unrelated Estonians and whole mitochondrial genome sequences were generated using the Illumina MiSeq system. The results are concordant with previous mtDNA control region reports of high haplogroup HV and U frequencies (47.4 and 23.7 % in this study, respectively) in the Estonian population. One sample with the Northern Asian haplogroup D was detected. The genetic diversity of the Estonian population sample was estimated to be 99.67 and 95.85 %, for mtGenome and HVI/HVII data, respectively. The random match probability for mtGenome data was 1.20 versus 4.99 % for HVI/HVII. The nucleotide mean pairwise difference was 27±11 for mtGenome and 7±3 for HVI/HVII data. These data describe the genetic diversity of the Estonian

✉ Monika Stoljarova
monika.stoljarova@ttu.ee

1 Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15A-604, Tallinn 12618, Estonia

2 Institute of Applied Genetics, Department of Molecular and Medical Genetics, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107, USA

population sample and emphasize the power of discrimination of the entire mitochondrial genome over the hypervariable regions.

Introduction

The analysis of mitochondrial DNA (mtDNA) has been implemented in molecular anthropology [1], evolutionary biology [2], medical genetics [3, 4], and human identity testing [5]. Strictly maternal inheritance, lack of recombination and high mutation rate make mtDNA a viable marker system for assessing genetic relationships among individuals or groups. Although having less discrimination power compared to autosomal DNA markers, the high copy number of mtDNA per cell increases the success rate of DNA typing of damaged, degraded, and low-quantity samples that fail to yield nuclear DNA profiles.

Forensic scientists have focused mainly on hypervariable regions I and II (HVI and HVII, respectively) due to high concentration of variants in those mitochondrial genome (mtGenome) regions. However, it has been reported that 75 % of the total variation within the mtGenome resides outside the control region, and therefore sequencing of the entire mtGenome increases the discrimination power and the value of generated data [6]. Sanger-type sequencing (STS) is still the main mtDNA typing technique in case work laboratories, and sequencing beyond the control region is attempted rarely as the well-established methodology is labor intensive, costly, and time consuming. In recent years, massively parallel sequencing (MPS) has been shown to be a feasible alternative to

STS. With the utility of bench-top sequencers like Illumina MiSeq and Ion Torrent PGM (Personal Genome Machine), generation of whole mtGenome data is feasible for the application-orientated laboratory [6–10].

Population data are essential for haplotype frequency estimation. While more than 25,000 forensic mtDNA sequences and almost 35,000 mtDNA sequences in total are available currently in the EMPOP database [11, 12], these data include information only from the mtDNA control region. In addition to STS, a number of other technologies permit practical access to mtDNA coding region data through single nucleotide polymorphism assays, sequence-specific oligonucleotide probes, mass spectroscopy, and MPS technology that emphasize the importance of a whole mtGenome database [13]. Recently, 588 forensic-quality whole mtGenomes from three major US populations have been determined with Sanger sequencing and will be available for query [14]. However, with MPS, population data can be generated far more expeditiously and at a lower cost per nucleotide. Thus, more whole genome data can be generated to exploit the full power of mtDNA for forensic identity testing.

The current population size of Estonia is slightly above 1.3 million. Throughout history, the native Estonian population has been affected by migration from both east and west due to numerous conquests. Estonians have served under German, Danish, Polish, Swedish, and Russian rule. Thus, there is some expectation of genetic admixture from these populations. A European genetic map of >1500 individuals and based on ~270,000 single nucleotide polymorphism data divided the European population into four groups and placed Estonians into the Baltic region, Poland, and Western Russia group [15]. It has been reported that Estonians have a higher Y-haplotype diversity, and based on their mtDNA HVI sequence, they have higher mean pairwise differences compared with other populations in the Baltic region [16].

Despite that fact that the Estonian population data has been used for migration studies based on its mtDNA HVI region and a number of coding region single nucleotide polymorphisms [17, 18, 16], to the best of our knowledge there are no published whole mtGenome data. The objective of this study was to describe the genetic variability of mtGenome in an Estonian population sample and to compare the discrimination power of mtGenome with solely HVI/HVII data.

Materials and methods

Sample preparation and target amplification

Buccal swabs were collected from 114 unrelated Estonian volunteers according to protocols approved by the Tallinn Medical Research Ethics Committee and have been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. In addition, all samples were collected in accordance with the University of North Texas Health Science Center Institutional Review Board. Buccal samples were collected with sterile Eurotubo® collection swabs (Deltalab, Rubí, Spain). The swabs were allowed to air dry and were stored at ambient temperature. DNA extraction was performed with the QIAamp® DNA Blood Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. The quantity of extracted DNA was determined using the Qubit® dsDNA HS (High Sensitivity) Quantification Kit and a Qubit® 2.0 Fluorometer (Life Technologies, Foster City, CA, USA). Amplification of the mtGenome was accomplished as described by King et al. [6].

Nextera XT library preparation

For library preparation, 1.0 ng of DNA was used. Libraries were prepared using Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol [19] except for the library normalization after bead cleanup and library preparation for sequencing. Following PCR cleanup, the libraries were quantified using the Qubit dsDNA BR (Broad Range) kit (Life Technologies, Foster City, CA, USA) and evaluated for fragment size using the High Sensitivity D1000 ScreenTape and Tape Station 2200 (Agilent Technologies, Santa Clara, CA, USA). Library normalization and preparation for sequencing was performed according to an in-house protocol. Purified libraries were normalized to 2 nM and pooled. Illumina PhiX control v3 (Illumina) was diluted to 2 nM with resuspension buffer (RSB, Illumina), and 2 μl of diluted PhiX were mixed with 14 μl of pooled library (2 nM). Further, 10 μl of PhiX-library pool were mixed with 10 μl of freshly made 0.1 N NaOH and vortexed. The library-PhiX-NaOH mix was incubated for 5 min at room temperature. Pre-chilled HT1 (980 μl, hybridization buffer, Illumina) was added to the mix. Then, 600 μl of library were mixed with 400 μl of HT1 for a final 12 pM sequencing library.

Sequencing and data generation

The 12 pM pooled library was sequenced with MiSeq v2 (2× 250 bp) chemistries (Illumina). The MiSeq re-sequencing protocol for small genome sequencing was followed according to the manufacturer's recommendations. Criteria used for variant calling (quality threshold, heteroplasmy threshold, coverage threshold) were as described by King et al. [6]. On-board software (i.e., Real-TimeAnalysis and MiSeq Reporter) converted raw data to Binary Alignment/Map (BAM) [20] and Variant Call Format (VCF) v4.1 files [21] using Genome Analysis Toolkit (GATK) [22]. During this process, the

sequenced regions of interest (ROIs) were aligned to the revised Cambridge Reference Sequence (rCRS) [23].

Data analysis

Generated VCF files were converted into haplotypes using MitoSAVE [24]. mtDNA variants were confirmed manually using BAM files and Integrative Genomic Viewer (IGV) software. Indels at the positions 309, 315, and 16193 were not included in the analysis. Random match probability (RMP) and genetic diversity (GD) were calculated according to the methods described by Stoneking et al. [25] and Tajima [26], respectively. Mean pairwise comparison was calculated using MEGA [27]. HaploGrep software based on Phylotree 16 [28, 29] was used for haplogroup assignment.

Results and discussion

All 114 samples resulted in whole mtGenome sequences obtained by MPS (Supplemental Table 1). Of these mtDNA profiles, 100 (87.7 %) were unique within the data set, and 12.3 % of sequenced mtGenomes were observed twice. Compared to the previous whole mtGenome population studies [6, 14], the number of shared haplotypes in the current study is relatively high. This might be a reflection of lower mtGenome diversity in Estonia or just may arise from sampling variance. The majority of the samples were collected at two educational institutions, and although an effort was made to ensure that the sampled individuals were unrelated, a long-distance kinship between these individuals cannot be excluded. In total, 2663 positions were reported as variants in relation to rCRS. These variants were distributed across 512 mtDNA positions. Variants 263G, 750G, 1438G, 4769G, 8860G, and 15326G were seen in 111 of 114 samples. The detection of these variants in majority of the samples is the reflection of reference used. The remaining three samples (EST-9, EST-19, and EST-40) exhibited few differences with respect to the rCRS (≤3), of which 1–2 were local or global private mutations as defined by HaploGrep. The low haplogroup assignment quality score for sample EST-19 was noted. The limited number of variants within the sample could indicate bias previously observed in reference to Phylotree and the rCRS [30]. Therefore, the true haplogroup for EST-19 individual (haplogroup H5e with the quality score 53 % assigned by HaploGrep) is likely to be in the H2 lineage. Six variants (73G, 2706G, 7028T, 11719A, 14766T, and 16519C) were found in ≥50 % of the samples. From these variants, 73G and 11719A were haplogroup nodes for haplogroup R, variants 2706G and 7028T for haplogroup H, and variant 14766T for haplogroup HV. Position 16519 is considered a hotspot and thus not useful for haplogroup assignment. Point heteroplasmy was detected in 14 samples (12.3 %). Observed point heteroplasmy with position

coverage and heteroplasmy percentage is listed in Supplemental Table 2. Three samples (EST-33, EST-81, and EST-106) exhibited two point heteroplasmy positions each. A similar extent of heteroplasmy (16.2 %), along with heteroplasmy at multiple positions per sample, has been observed previously in buccal cell mtDNA [31].

Haplogroup assignment using HaploGrep software resulted in 11 major clades and 87 distinct haplogroups. Major clades were D, HV (including haplogroup H), I, J, M, N, R, T, U (including haplogroup K), W, and X. Two clades were dominant: 54 samples (47.4 %) belonged to the haplogroup HV (including H) and 27 samples (23.7 %) pertained to the haplogroup U (including K). Haplogroups D, R, and X were seen once. Haplogroups HV and U are the most represented haplogroups in the European population with the estimated frequency of ≥50 and ≥20 %, respectively [17]. Our results are concordant with previous mtDNA control region reports conforming of the prevalence of haplogroup H and U in the Estonian population [32, 33]. While haplogroup H was introduced to Europe from the Franco-Cantabrian region, haplogroup U5, which was observed in 13 of our samples (11.4 %), is thought to have evolved in situ [32, 17]. Haplogroup U4 has been reported in the Eastern Baltic Sea region with a frequency up to 8.8 % and associated with Volga-Ural influence [16]. While haplogroup D is the second most common haplogroup in Northern Asia, haplogroup D5 has been found with a very low frequency in several European populations including Estonians [34]. The rare subhaplogroup D4e4b, observed in one of our samples, has been reported in Tatars and Russians [35].

Out of the observed 2663 nucleotide variants, 607 (22.8 %) were identified in HVI/HVII regions; accordingly, 77.2 % of variation resided in the coding region. These results are in accordance with results reported by King et al. on 283 individuals from Caucasian, Hispanic, and African-American populations [6]. As in the case of whole mtGenome data, the proportionally smaller level of variation in HVI/HVII may be a reflection of lower mtGenome diversity in Estonia or just may arise from sampling variance. Whereas 100 unique mtGenome haplotypes and 87 distinct haplogroups were observed with whole mtGenome data, only 66 unique haplotypes and 79 distinct haplogroups were observed using HVI/HVII data. Haplogroup comparison between full mtGenome and HVI/HVII data resulted in a haplogroup clade change according to HaploGrep for 1 sample (EST-59) that changed from haplogroup U5b2a1a2 (HVI/HVII; quality score 86.8 %) to haplogroup H (mtGenome; quality score 81 %). Haplogroup assignment based on mtGenome data resulted in a quality score increase for the majority of the samples that yielded a quality score less than 100 % with HVI/HVII. A lower quality score of sample EST-59 can be explained with an abundance of local and global private mutations that were not present in HVI/HVII data. GD for mtGenome and HVI/

HVII was 99.67 and 95.85 %, respectively. RMP for mtGenome data was 1.20 versus 4.99 % for HVI/HVII. Compared to RMP values of other populations [14, 6, 36], the RMP results presented herein for HVI/HVII data are higher. This finding might be explained with 24 HVI/HVII haplotypes having ≥4 identical matches in the population sample. Mean pairwise difference within the Estonian population was 27±11 for mtGenome data, which is slightly lower than reported by King et al. [6] for Caucasians. Mean pairwise difference for HVI/HVII data within the Estonian population sample was 7±3. These results support the power of discrimination of entire mtGenome over HVI/HVII.

Conclusion

In this study, 114 mtGenome profiles from the Estonian population were generated. The results show that the use of the entire mtGenome compared to HVI/HVII data substantially improves the discrimination power of the quality of haplotypes.

References

1. King TE, Fortes GG, Balaresque P, Thomas MG, Balding D, Delser PM, Neumann R, Parson W, Knapp M, Walsh S, Tonasso L, Holt J, Kayser M, Appleby J, Forster P, Ekserdjian D, Hofreiter M, Schurer K (2014) Identification of the remains of King Richard III. Nat Commun 5:5631. doi:10.1038/ncomms6631

2. Rito T, Richards MB, Fernandes V, Alshamali F, Cerny V, Pereira L, Soares P (2013) The first modern human dispersals across Africa. PLoS One 8(11), e80031. doi:10.1371/journal.pone.0080031

3. Li CT, Bai YM, Hsieh JC, Lee HC, Yang BH, Chen MH, Lin WC, Tsai CF, Tu PC, Wang SJ, Su TP (2015) Peripheral and central glucose utilizations modulated by mitochondrial DNA 10398A in bipolar disorder. Psychoneuroendocrinology 55C:72–80. doi:10.1016/j.psyneuen.2015.02.003

4. Hagen CM, Aidt FH, Hedley PL, Jensen MK, Havndrup O, Kanters JK, Moolman-Smook JC, Larsen SO, Bundgaard H, Christiansen M (2013) Mitochondrial haplogroups modify the risk of developing hypertrophic cardiomyopathy in a Danish population. PLoS One 8(8), e71904. doi:10.1371/journal.pone.0071904

5. Tokutomi T, Takada Y, Kanetake J, Mukaida M (2009) Identification using DNA from skin contact: case reports. Leg Med (Tokyo) 11(Suppl 1):S576-577. doi:10.1016/j.legalmed.2009.02.004

6. King JL, LaRue BL, Novroski NM, Stoljarova M, Seo SB, Zeng X, Warshauer DH, Davis CP, Parson W, Sajantila A, Budowle B (2014) High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. Forensic Sci Int Genet 12:128–135. doi:10.1016/j.fsigen.2014.06.001

7. Seo SB, Zeng X, King JL, Larue BL, Assidi M, Al-Qahtani MH, Sajantila A, Budowle B (2015) Underlying data for sequencing the mitochondrial genome with the massively parallel sequencing platform ion torrent() PGM (). BMC Genomics 16(1):6938. doi:10.1186/1471-2164-16-S1-S4

8. Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delport R, Langit R, Wootton S, Lagace R, Irwin J (2013) Reprint of: evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7(6):632–639. doi:10.1016/j.fsigen.2013.09.007

9. Bodner M, Iuvaro A, Strobl C, Nagl S, Huber G, Pelotti S, Pettener D, Luiselli D, Parson W (2015) Helena, the hidden beauty: resolving the most common West Eurasian mtDNA control region haplotype by massively parallel sequencing an Italian population sample. Forensic Sci Int-Gen 15:21–26. doi:10.1016/j.fsigen.2014.09.012

10. McElhoe JA, Holland MM, Makova KD, Su MS, Paul IM, Baker CH, Faith SA, Young B (2014) Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet 13:20–29. doi:10.1016/j.fsigen.2014.05.007

11. European DNA Profiling Group (EDNAP) (1999–2015) EMPOP. www.empop.org. Accessed 19 Mar 2015

12. Parson W, Dur A (2007) EMPOP—a forensic mtDNA database. Forensic Sci Int Genet 1(2):88–92. doi:10.1016/j.fsigen.2007.01.018

13. Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Rock AW, Huber GE, Nagl S, Strobl C, Zimmermann B, Parson W, Irwin JA (2014) Development of forensic-quality full mtGenome haplotypes: success rates with low template specimens. Forensic Sci Int Genet 10:73–79. doi:10.1016/j.fsigen.2014.01.010

14. Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Rock AW, Bush JM, Higginbotham JL, Peck MA, Ring JD, Huber GE, Xavier C, Strobl C, Lyons EA, Diegoli TM, Bodner M, Fendt L, Kralj P, Nagl S, Niederwieser D, Zimmermann B, Parson W, Irwin JA (2015) Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three U.S. populations. Forensic Sci Int Genet 14:141–155. doi:10.1016/j.fsigen.2014.09.021

15. Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, Karachanak S, Piskackova T, Balascak I, Peltonen L, Jakkula E, Rehnstrom K, Lathrop M, Heath S, Galan P, Schreiber S, Meitinger T, Pfeufer A, Wichmann HE, Melegh B, Polgar N, Toniolo D, Gasparini P, D'Adamo P, Klovins J, Nikitina-Zake L, Kucinskas V, Kasnauskiene J, Lubinski J, Debniak T, Limborska S, Khrunin A, Estivill X, Rabionet R, Marsal S, Julia A, Antonarakis SE, Deutsch S, Borel C, Attar H, Gagnebin M, Macek M, Krawczak M, Remm M, Metspalu A (2009) Genetic structure of Europeans: a view from the North-East. PLoS One 4(5), e5472. doi:10.1371/journal.pone.0005472

16. Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus ML, Lahermo P (2008) Migration waves to the Baltic Sea region. Ann Hum Genet 72(Pt 3):337–348. doi:10.1111/j.1469-1809.2007.00429.x

17. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Golge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Norby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozari R, Torroni A, Bandelt HJ (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67(5):1251–1276

18. Pliss L, Tambets K, Loogvali EL, Pronina N, Lazdins M, Krumina A, Baumanis V, Villems R (2006) Mitochondrial DNA portrait of

Latvians: towards the understanding of the genetic structure of Baltic-speaking populations. Ann Hum Genet 70(Pt 4):439–458. doi:10.1111/j.1469-1809.2005.00238.x

19. Illumina (2015) Nextera XT DNA Library preparation guide
20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079. doi:10.1093/bioinformatics/btp352
21. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, Genomes Project Analysis G (2011) The variant call format and VCFtools. Bioinformatics 27(15):2156–2158. doi:10.1093/bioinformatics/btr330
22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20(9):1297–1303. doi:10.1101/gr.107524.110
23. 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. doi:10.1038/13779
24. King JL, Sajantila A, Budowle B (2014) mitoSAVE: mitochondrial sequence analysis of variants in Excel. Forensic Sci Int Genet 12: 122–125. doi:10.1016/j.fsigen.2014.05.013
25. Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, Erlich HA (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am J Hum Genet 48(2):370–382
26. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–595
27. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729. doi:10.1093/molbev/mst197
28. Kloss-Brandstatter A, Pacher D, Schonherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F (2011) HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum Mutat 32(1):25–32. doi:10.1002/humu.21382

29. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30(2):E386–E394. doi:10.1002/humu.20921
30. Bandelt HJ, Salas A (2012) Current next generation sequencing technology may not meet forensic standards. Forensic Sci Int Genet 6(1):143–145. doi:10.1016/j.fsigen.2011.04.004
31. Naue J, Horer S, Sanger T, Strobl C, Hatzer-Grubwieser P, Parson W, Lutz-Bonengel S (2015) Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. Mitochondrion 20:82–94. doi:10.1016/j.mito.2014.12.002
32. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogvali EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, Semino O, Torroni A (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. doi:10.1086/425590
33. Ottoni C, Ricaut FX, Vanderheyden N, Brucato N, Waelkens M, Decorte R (2011) Mitochondrial analysis of a Byzantine population reveals the differential impact of multiple historical events in South Anatolia. Eur J Hum Genet: EJHG 19(5):571–576. doi:10.1038/ejhg.2010.230
34. Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogvali EL, Tolk HV, Reidla M, Metspalu E, Pliss L, Balanovsky O, Pshenichnov A, Balanovska E, Gubina M, Zhadanov S, Osipova L, Damba L, Voevoda M, Kutuev I, Bermisheva M, Khusnutdinova E, Gusar V, Grechanina E, Parik J, Pennarun E, Richard C, Chaventre A, Moisan JP, Barac L, Pericic M, Rudan P, Terzic R, Mikerezi I, Krumina A, Baumanis V, Koziel S, Rickards O, De Stefano GF, Anagnou N, Pappa KI, Michalodimitrakis E, Ferak V, Furedi S, Komel R, Beckman L, Villems R (2004) The western and eastern roots of the Saami—the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. Am J Hum Genet 74(4):661–682. doi:10.1086/383203
35. Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Rogalla U, Perkova M, Dambueva I, Zakharov I (2010) Origin and post-glacial dispersal of mitochondrial DNA haplogroups C and D in northern Asia. PLoS One 5(12), e15214. doi:10.1371/journal.pone.0015214
36. Lembring M, van Oven M, Montelius M, Allen M (2013) Mitochondrial DNA analysis of Swedish population samples. Int J Legal Med 127(6):1097–1099. doi:10.1007/s00414-013-0908-6