

HTT haplogroups in Finnish patients with Huntington disease

Susanna Ylönen, MSc, Jussi O.T. Sipilä, MD, DMSc, Marja Hietala, MD, DMSc, and Kari Majamaa, MD, DMSc

Neurol Genet 2019;5:e334. doi:10.1212/NXG.0000000000000334

Correspondence

Dr. Sipilä
jussi.sipila@utu.fi

Abstract

Objective

To study genetic causes of the low frequency of Huntington disease (HD) in the Finnish population, we determined *HTT* haplogroups in the population and patients with HD and analyzed intergenerational Cytosine-Adenosine-Guanosine (CAG) stability.

Methods

A national cohort of patients with HD was used to identify families with mutant *HTT* (*mHTT*). *HTT* haplogroups were determined in 225 archival samples from patients and from 292 population samples. CAG repeats were phased with *HTT* haplotypes using data from parent-offspring pairs and other *mHTT* carriers in the family.

Results

The allele frequencies of *HTT* haplotypes in the Finnish population differed from those in 411 non-Finnish European subjects ($p < 0.00001$). The frequency of haplogroup A was lower than that in Europeans and haplogroup C was higher. Haplogroup A alleles were significantly more common in patients than in controls. Among patients with HD haplotypes A1 and A2 were more frequent than among the controls ($p = 0.003$). The mean size of the CAG repeat change was +1.38 units in paternal transmissions being larger than that (−0.17) in maternal transmissions ($p = 0.008$). CAG repeats on haplogroup A increased by 3.18 CAG units in paternal transmissions, but only by 0.11 units in maternal transmissions ($p = 0.008$), whereas haplogroup C repeat lengths decreased in both paternal and maternal transmissions.

Conclusions

The low frequency of HD in Finland is partly explained by the low frequency of the HD-associated haplogroup A in the Finnish population. There were remarkable differences in intergenerational CAG repeat dynamics that depended on *HTT* haplotype and parent gender.

From the Division of Clinical Neuroscience (S.Y., K.M.), Neurology, University of Oulu; Department of Neurology and Medical Research Center (S.Y., K.M.), Oulu University Hospital; Department of Neurology (J.O.T.S.), North Karelia Central Hospital, Siun Sote, Joensuu; Division of Clinical Neurosciences (J.O.T.S.), Turku University Hospital; Neurology (J.O.T.S.), University of Turku; Department of Clinical Genetics (M.H.), Turku University Hospital; and Institute of Biomedicine (M.H.), University of Turku, Finland.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by the University of Oulu.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Glossary

CAG = cytosine-adenosine-guanosine; **DNA** = deoxyribonucleic acid; **HD** = Huntington's disease; **LNA** = locked nucleic acid; **mHTT** = mutant HTT; **SNP** = single-nucleotide polymorphism; **THL** = National Institute for Health and Welfare.

Huntington disease (HD) is caused by an expansion of a Cytosine-Adenosine-Guanosine (CAG) repeat in the *HTT* gene.¹ The expansion is driven by paternal transmissions, where the repeat instability correlates with repeat length.^{2–5} The prevalence of HD varies markedly between populations. The disease is most common among white populations, whereas the prevalence is only a 10th or less in Asian populations.⁶ Among whites, the Finns and Icelanders have low prevalence, whereas the prevalence rates in other Scandinavian countries are similar to those in other whites.^{7,8}

Shortly after the identification of the *HTT* gene, it was established that the risk of CAG expansion is associated with certain chromosome 4 haplotypes.⁹ Indeed, differences in the frequency of HD between populations correlate with differences in the frequency of *HTT* haplogroups defined by single-nucleotide polymorphisms (SNPs). Haplogroup A is the most frequent haplogroup in white patients with HD,¹⁰ while haplogroup C is the most common one in white general populations. On the other hand, East Asian populations lack haplotypes A1 and A2 and HD cases harbor haplogroup C.¹¹ *HTT* haplogroups have been suggested to differ in the rate of new mutations,¹¹ implying differences in the intergenerational stability of the CAG repeats.

We have previously estimated that the frequency of haplogroup A is lower in the Finnish general population than that in most European populations.¹² Here we report the frequencies of *HTT* haplogroups in Finnish patients with HD and in the general population and assess intergenerational stability of CAG repeat expansion across *HTT* haplogroups.

Methods

Patients and controls

A previously reported national cohort of patients with HD¹² was used to identify families with mutant HTT (*mHTT*). Archival deoxyribonucleic acid (DNA) samples (N = 225) remaining from diagnostic or predictive *HTT* analyses were obtained from the 2 national laboratories that perform HD diagnostics. Family relationships were determined using data from the Population Register Centre, a governmental authority keeping records of citizens, the archives of Family Federation of Finland, a nongovernmental nonprofit welfare organization that has provided genetic counselling to Finnish families with HD, and hospital records. Phasing of CAG repeat and *HTT* haplotype was conducted using data from parent-offspring pairs and other *mHTT* carriers in the family. As controls, we had 292 healthy Finnish blood donors with a mean age of 40.5 years and range 19–64 years.¹³

The study has been approved by the ethics committee of Hospital District of Southwestern Finland (Dnro ETMK 19/180/2010) and received the national study permits from the National Institute for Health and Welfare (THL, Dnro THL/1456/5.05.00/2010) and the National Supervisory Authority for Welfare and Health, Valvira (Dnro 1195/06.01.03.01/2012). The study involved no contact with patients. Hence, no informed consent was stipulated.

Molecular methods

HTT haplogroups A, B, and C were defined using 2 intragenic SNPs, rs762855 and rs4690073, and haplotype A variants were further defined using 4 additional intragenic SNPs, rs2857936, rs363096, rs2276881, and rs362307.¹⁰ The SNPs were determined with restriction fragment analysis (FastDigest; Thermo Fisher Scientific, Waltham, MA), allele-specific amplification using locked nucleic acid (LNA) primers (Exiqon, Vedbaek, Denmark), or by sequencing (ABI3500xL Genetic Analyzer, Applied Biosystems, Foster City, CA) as appropriate (table e-1, links.lww.com/NXG/A153). The PCRs were performed using Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific) according to the standard procedures. The LNA amplifications were conducted in duplicates or confirmed by sequencing. Additional SNPs (table e-1, links.lww.com/NXG/A153) were analyzed in samples, where haplogroup A variant remained undetermined.^{10,14} Haplotypes were manually annotated and then phased to CAG repeat length by familial relationship or by homozygosity of haplotype.

Statistical analysis

Exact test of population differentiation as implemented in Arlequin 3.5¹⁵ was used to compare frequencies of haplogroups between cases and controls. Chi-square test or *t* test was used to compare 2 groups as appropriate.

Data availability statement

Access to data was regulated by the Finnish law, THL, and Valvira. Permission to conduct this study prohibits disclosing data to third parties without explicit permission from THL and Valvira. Those fulfilling the requirements for viewing confidential data as required by the Finnish law and who receive the permission from THL and Valvira are able to access the data.

Results

Frequencies

The allele frequencies of *HTT* haplotypes in the Finnish population differed from those in 411 non-Finnish European subjects¹¹ ($p < 0.00001$, exact test of population differentiation). The frequency of haplogroup A was lower ($p = 0.0003$, χ^2 test)

than that in Europeans, and haplogroup C was higher ($p = 0.046$, χ^2 test), but there was no difference in the proportion of haplotypes A1 and A2 out of all haplogroup A alleles.

The allele frequencies of *HTT* haplotypes among Finnish patients with HD and asymptomatic *mHTT* carriers differed from those in the controls ($p < 0.00001$, exact test of population differentiation). Haplogroup A alleles were significantly more common in patients than in controls, while haplogroup B and C alleles were more common in population controls (table 1). Frequencies of haplotypes A1–A5 differed between controls and patients with HD ($p < 0.00001$, exact test of population differentiation). Among patients with HD (including asymptomatic *mHTT* carriers), haplotypes A1 and A2 were more frequent and haplotypes A4 and A5 less frequent than among the controls ($p = 0.003$, exact test of population differentiation).

Transmissions

We were able to analyze the change in the size of *mHTT* CAG repeat in 65 transmissions (table 2). The mean size of the change was +1.38 CAG units in paternal transmissions being larger than that in maternal transmissions (−0.17 CAG units) ($p = 0.008$, t test). *mHTT* haplotype SNPs and CAG repeat could be phased in 50 transmissions (haplogroup A, 38; haplogroup C, 10; other haplogroup, 2; figure 1). *mHTT* CAG repeats in haplogroup A (42.6 CAG units) were shorter than those in haplogroup C (48.4 CAG units). In haplogroup C, the *mHTT* CAG repeats decreased in paternal as well as maternal transmissions ($p = 0.74$ for gender difference), whereas the mean increase of the repeats in haplogroup A was considerable in paternal transmissions (3.18 CAG units) but meager in maternal transmissions (0.11 CAG units) ($p = 0.008$ for difference) (figure 2). The gender difference in *mHTT* CAG change was significant in haplotype A1 inheritances ($p = 0.022$), but not in haplotype A2 or A3 inheritance.

Discussion

We found remarkable differences in intergenerational CAG repeat dynamics that depended on *HTT* haplotype and parent

Table 1 Frequencies of *HTT* haplotypes in Finnish patients with HD and controls

Haplogroup	Patients with HD, N (%)	Controls, N (%)	Non-Finnish Europeans, N (%)
A	291 (64.7)	209 (35.8)	197 (47.9)
B	27 (6.0)	63 (10.8)	17 (4.1)
C	128 (28.4)	310 (53.1)	190 (46.2)
O	4 (0.9)	2 (0.3)	7 (1.7)
Total	450	584	411

Abbreviation: HD = Huntington disease. Patients with HD includes asymptomatic carriers. Data on non-Finnish Europeans are from Warby et al.¹¹

Table 2 The change in the size of *mHTT* CAG repeat in 65 transmissions

	Total	Mother	Father
Unchanged alleles, N (%)	18 (28)	14 (34)	4 (17)
Contractions, N (%)	21 (32)	15 (37)	6 (25)
Expansions, N (%)	26 (40)	12 (29)	14 (58)
Mean change, N (%)	+0.4 (+1.0)	−0.2 (−0.3)	+1.4 (+3.1)
Range of change, N	−7 to +12	−3 to +4	−7 to +12
Range of change, %	−14 to +27	−7 to +10	−14 to +27

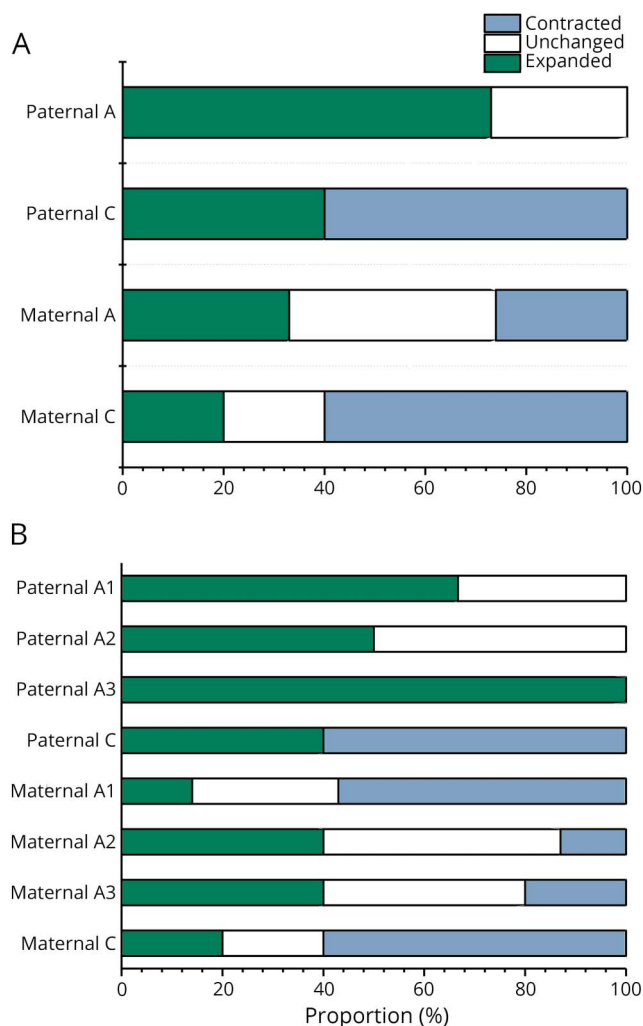
Abbreviation: SNP = single-nucleotide polymorphism. The median size of the parent CAG repeat was 43 triplets (percentiles 41, 46) in each group. The 65 transmissions include 50 transmissions, where *mHTT* haplotype SNPs and CAG repeat could be phased.

gender. *HTT* haplotype A1 was not only associated with HD, but also showed a predilection to CAG repeat expansion in paternal transmissions. Repeat tracts on haplotypes A2 and A3 showed propensity to expand, whereas those on haplogroup C showed a preference for contraction irrespective of parental gender.

Previous studies have shown that paternal transmission drives intergenerational CAG expansion in *HTT*.^{2–5} We identified 65 Finnish parent-child pairs, where paternal transmission was associated with repeat expansion (+1.38 CAG units), whereas maternal transmission mainly resulted in repeat contraction (−0.17 CAG units). Strikingly similar figures have been obtained in 337 parent-offspring transmissions in Dutch patients, where paternal transmission was associated with CAG repeat expansion (+1.76 CAG units) and maternal transmissions resulted in contraction (−0.07 CAG units).³ These findings suggest that intergenerational stability of *mHTT* in Finnish HD families does not differ from that in other white populations. Interestingly, the intergenerational stability of the CAG repeat differed between *mHTT* haplotypes. On average, repeat tracts on haplotype A1 gained length in paternal transmissions, while maternally inherited tracts decreased in length. Haplotype A1 seems to be a major contributor to paternal anticipation in HD, as 48% of all the pathogenic expansions were on haplotype A1.

We found that repeat tracts on haplogroup C were larger than those on haplogroup A and that repeat tracts on haplogroup C showed a tendency to contract. Indeed, the ratio between the number of contraction events and the number of expansion events was 2.0 in haplogroup C transmissions, while the corresponding ratio was 0.4 in haplogroup A transmissions (figure 1A). The skew toward contraction suggests that *HTT* haplogroup C harbors DNA sequence elements that favor contraction in previously elongated repeat tracts. These findings should be interpreted with caution, however, as *HTT* haplotype and CAG repeat could be phased in only 10 haplogroup C samples. The haplogroup C samples showed

Figure 1 Proportions of changes in *mHTT* CAG repeat tracts in transmission events in 48 Finnish families with Huntington disease

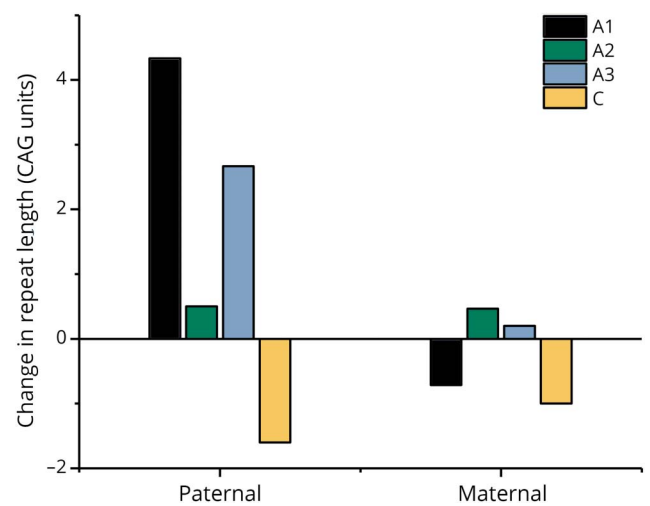


Two families belonging to haplogroups other than A or C are not included. (A) Changes in haplogroups A and C. (B) Changes in haplogroup A haplotypes.

a tendency for contraction in both paternal and maternal transmissions, which resembles the findings in some Cretan HD families.¹⁶ The repeat tracts are stable or contract in transmission in these families and, clinically, the families do not show anticipation.

Remarkable ethnic differences in the prevalence of HD have suggested a genetic contribution. The frequency is highest among European populations and a recent systematic review compiled 14 studies and found that the median prevalence is 5.7/100,000 (range 0.96–11.8).⁶ The lowest prevalence rates included in the review were 0.96/100,000 in Iceland¹⁷ and 2.14/100,000 in Finland.¹² In the European population, CAG expansions occur predominantly on haplotypes A1 and A2, which are present in 20% of the individuals with <27 CAG repeats.^{10,11} We found that haplogroup frequencies in the

Figure 2 Mean change in *mHTT* CAG repeat length in maternal and paternal transmission events in 48 Finnish families with Huntington disease



Two families belonging to haplogroups other than A or C are not included. Changes are shown for haplogroup A haplotypes and for haplogroup C.

Finnish general population differ from those among non-Finnish Europeans. In particular, haplogroup A was low, and thus the low frequency of the susceptibility haplogroups partly explains the low HD prevalence in the Finnish population. Similar to other white populations, the majority of HD cases in Finland harbored haplotype A1 or A2 and the 2 haplotypes constituted similar proportions of haplogroup A chromosomes.

In addition to *HTT* haplotype, other possible contributors to variation in prevalence rates include differences in the rate of new mutations and founder effect. A recent analysis has suggested that a higher frequency of haplotype A1 in the general population yield a higher rate of new HD mutations.¹⁸ Hence, the relatively low rate of new mutations would be the mechanism producing low prevalence of HD in the Finnish population by virtue of the low proportion of *HTT* haplotypes A1 and A2. In Finland, the haplotype distribution may be because of small founder population and geographic, cultural, and linguistic isolation have shaped the gene pool of the Finnish population.¹⁹ These events have yielded Finnish disease heritage, where certain Mendelian diseases are enriched in the population, while some others are lacking or occur at a low prevalence, such as HD.¹² The population genomics also shows some ancient features that are undetectable in most other regions of Europe.²⁰

Taking into account the low frequency of HD in Finland, our sample set of 225 is rather comprehensive. The archival nature of the samples posed some limitations, as lack of samples precluded phasing in many families. However, homozygosity of the *HTT* haplotype or results from 79 out of the 225 samples enabled us to phase *HTT* haplotype and CAG repeat in 50 transmissions.

Here we found that the intergenerational stability of the CAG repeat differed between *mHTT* haplotypes. Repeat tracts on haplotype A1 gained length only in paternal transmissions, while paternally as well as maternally inherited tracts decreased in length on haplogroup C. In addition, we showed that the frequency of the HD-associated *HTT* haplogroup A was lower in the Finnish population than that in the non-Finnish European population and that the frequency of haplogroup A was higher in patients with HD than that in the population. These findings at least partly explain the low prevalence of HD in Finland.

Disclosure

S. Ylönen has nothing to declare. J.O.T. Sipilä has received honoraria (Merck, Pfizer, Sanofi Genzyme), has received a consultancy fee (Rinne Foundation), has received travel grants and congress sponsorship (Orion Corporation, Merck Serono, Sanquin, Lundbeck, Novartis), and holds shares (Orion Corporation). M. Hietala has nothing to declare. K. Majamaa has received travel support from MSD Finland and honoraria from Sanofi Genzyme and UCB. This study was not industry-sponsored. Go to Neurology.org/NG for full disclosure.

Acknowledgment

The expert technical assistance of Ms. Anja Heikkinen is gratefully acknowledged. The authors would like to thank Dr. Annukka M. Tuiskula, PhD, and Mr. Antti-Jussi Kortevara for valuable assistance in obtaining the DNA samples archived by HUSLAB and the staff of the genetic laboratory of the Hospital District of Southwest Finland for gathering and dispatching the samples stored there.

Study funding

This study was supported in part by grants from the Sigrd Juselius Foundation, Turku University Foundation, and the Finnish Parkinson Foundation.

Publication history

Received by *Neurology: Genetics* February 8, 2019. Accepted in final form March 20, 2019.

Appendix Authors

Name	Affiliation	Role	Contribution
Susanna Ylönen, MSc	University of Oulu, Oulu	Author	Performed all laboratory work; analyzed the data; drafted the manuscript for intellectual content
Jussi Sipilä, MD, DMSc	North Karelia Central Hospital, Joensuu and University of Turku and Turku University Hospital, Turku	Author	Design and conceptualized study; major role in the acquisition of data; analyzed the data; revised the manuscript for intellectual content

Appendix (continued)

Name	Affiliation	Role	Contribution
Marja Hietala, MD, DMSc	Turku University Hospital, Turku	Author	Helped in data acquisition; interpreted the data; revised the manuscript for intellectual content; supervised the study
Kari Majamaa, MD, DMSc	University of Oulu, Oulu and Oulu University Hospital	Author	Design and conceptualized study; interpreted the data; performed the statistical analyses; analyzed the data; revised the manuscript for intellectual content; supervised the study

References

- MacDonald ME, Ambrose CM, Duyao MA, et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 1993;72:971-983.
- Norremølle A, Sørensen SA, Fenger K, Hasholt L. Correlation between magnitude of CAG repeat length alterations and length of the paternal repeat in paternally inherited Huntington's disease. *Clin Genet* 1995;47:113-117.
- Aziz NA, van Belzen MJ, Coops ID, Belfroid RDM, Roos RAC. Parent-of-origin differences of mutant HTT CAG repeat instability in Huntington's disease. *Eur J Med Genet* 2011;54:e413-e418.
- Ramos EM, Cerqueira J, Lemos C, Pinto-Basto J, Alonso I, Sequeiros J. Intergenerational instability in Huntington disease: extreme repeat changes among 134 transmissions. *Mov Disord* 2012;27:583-585.
- Semaka A, Kay C, Doty C, et al. CAG size-specific risk estimates for intermediate allele repeat instability in Huntington disease. *J Med Genet* 2013;50:696-703.
- Baig SS, Strong M, Quarrell OW. The global prevalence of Huntington's disease: a systematic review and discussion. *Neurodegener Dis Manag* 2016;6:331-343.
- Gilling M, Budtz-Jørgensen E, Boonen SE, et al. The Danish HD Registry—a nationwide family registry of HD families in Denmark. *Clin Genet* 2017;92:338-341.
- Roos AK, Wiklund L, Laurell K. Discrepancy in prevalence of Huntington's disease in two Swedish regions. *Acta Neurol Scand* 2017;136:511-515.
- Andrew SE, Hayden MR. Origins and evolution of Huntington disease chromosomes. *Neurodegeneration* 1995;4:239-244.
- Warby SC, Montpetit A, Hayden AR, et al. CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. *Am J Hum Genet* 2009;84:351-366.
- Warby SC, Visscher H, Collins JA, et al. HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia. *Eur J Hum Genet* 2011;19:561-566.
- Sipilä JO, Hietala M, Siitonen A, Päiväranta M, Majamaa K. Epidemiology of Huntington's disease in Finland. *Parkinsonism Relat Disord* 2015;21:46-49.
- Meinilä M, Finnilä S, Majamaa K. Evidence for mtDNA admixture between the Finns and the Saami. *Hum Hered* 2001;52:160-170.
- Kay C, Collins JA, Skotte NH, et al. Huntingtin haplotypes provide prioritized target panels for allele-specific silencing in Huntington disease patients of European ancestry. *Mol Ther* 2015;23:1759-1771.
- Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2007;1:47-50.
- Tsagourisakis M, Fesdjian OC, Shashidharan P, Plaitakis A. Stability of the Huntington disease (CAG)_n repeat in a late onset form occurring on the island of Crete. *Hum Mol Genet* 1996;4:2239-2243.
- Sveinsson O, Halldórsson S, Olafsson E. An unusually low prevalence of Huntington's disease in Iceland. *Eur Neurol* 2012;68:48-51.
- Kay C, Collins JA, Wright GEB, et al. The molecular epidemiology of Huntington disease is related to intermediate allele frequency and haplotype in the general population. *Am J Med Genet B Neuropsychiatr Genet* 2018;177:346-357.
- Norio R. Finnish Disease Heritage I: characteristics, causes, background. *Hum Genet* 2003;112:441-456.
- Neuvonen AM, Putkonen M, Översti S, et al. Vestiges of an ancient border in the contemporary genetic diversity of north-eastern Europe. *PLoS One* 2015;10:e0130331.