

RESEARCH ARTICLE

Vestiges of an Ancient Border in the Contemporary Genetic Diversity of North-Eastern Europe

Anu M. Neuvonen^{1*}, Mikko Putkonen¹, Sanni Översti^{2,3}, Tarja Sundell^{3,4}, Päivi Onkamo³, Antti Sajantila^{1,5}, Jukka U. Palo¹

1 Laboratory of Forensic Biology, Department of Forensic Medicine, PO Box 40 (Kytösuontie 11) FI-00014, University of Helsinki, Helsinki, Finland, **2** Department of Biology, Section of Genetics and Physiology, University of Turku, Itäinen Pitkäkatu 4, 20014, Turku, Finland, **3** Department of Biosciences, PO Box 56 (Viikinkaari 5) FI-00014, University of Helsinki, Helsinki, Finland, **4** Department of Philosophy, History, Culture and Art Studies, PO Box 59 (Unioninkatu 38 F) FI-00014, University of Helsinki, Helsinki, Finland, **5** Institute of Applied Genetics, Department of Molecular and Medical Genetics, University of North Texas Health Science Center, 3500 Camp Bowie Blvd, Fort Worth, Texas, 76107, United States of America

* anu.neuvonen@helsinki.fi



CrossMark
click for updates

OPEN ACCESS

Citation: Neuvonen AM, Putkonen M, Översti S, Sundell T, Onkamo P, Sajantila A, et al. (2015) Vestiges of an Ancient Border in the Contemporary Genetic Diversity of North-Eastern Europe. PLoS ONE 10(7): e0130331. doi:10.1371/journal.pone.0130331

Academic Editor: Luísa Maria Sousa Mesquita Pereira, IPATIMUP (Institute of Molecular Pathology and Immunology of the University of Porto), PORTUGAL

Received: May 18, 2014

Accepted: May 19, 2015

Published: July 1, 2015

Copyright: © 2015 Neuvonen et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by The Finnish Foundations' Professor Pool Grant (<http://www.professoripooli.fi>), Paulo Foundation; AS), The Finnish Population Genetics Graduate School (<http://www oulu.fi/biology/PopGenSchool/>); AMN, MP), the University of Helsinki (AMN, TS), The Finnish Concordia Fund (<http://www.konkordia-liitto.com/>), and the Academy of Finland grant no. 133056

Abstract

It has previously been demonstrated that the advance of the Neolithic Revolution from the Near East through Europe was decelerated in the northernmost confines of the continent, possibly as a result of space and resource competition with lingering Mesolithic populations. Finland was among the last domains to adopt a farming lifestyle, and is characterized by substructuring in the form of a distinct genetic border dividing the northeastern and southwestern regions of the country. To explore the origins of this divergence, the geographical patterns of mitochondrial and Y-chromosomal haplogroups of Neolithic and Mesolithic ancestry were assessed in Finnish populations. The distribution of these uniparental markers revealed a northeastern bias for hunter-gatherer haplogroups, while haplogroups associated with the farming lifestyle clustered in the southwest. In addition, a correlation could be observed between more ancient mitochondrial haplogroup age and eastern concentration. These results coupled with prior archeological evidence suggest the genetic northeast/southwest division observed in contemporary Finland represents an ancient vestigial border between Mesolithic and Neolithic populations undetectable in most other regions of Europe.

Introduction

In Europe, human history has been decisively shaped by two events: the colonization of Europe by modern humans c. 45 000 years ago (45 kya) and the spread of agricultural technology from the Fertile Crescent in Asia Minor to north of Europe 9–5 kya. The latter Neolithisation process has been a major transformation period in the history of most European populations [1–4].

Over the years, there has been much controversy regarding the mechanism by which the agricultural lifestyle advanced from the Near East to western and northern Europe (see [5] and

(<http://www.aka.fi/fi/A/>; PO, SÖ, MP, TS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

references therein). According to the current view, genetic evidence appears to favor intermediate scenarios between the opposing cultural diffusion model (advance of technology) and the demic diffusion model (advance of Neolithic people) [6]. After colonizing the Balkans, the Neolithic advance continued westwards by two main routes, along the Mediterranean coast and through the Central European plains following the Danube [4, 7]. Eventually, this led to the admixture of Mesolithic hunter-gatherer and Neolithic farmer gene pools (see e.g. [8]). Studies based on analyses of contemporary genetic diversity have estimated widely varying proportions for genes traceable to the Paleolithic and Neolithic gene pools [9–11]. This picture has been further elucidated by ancient DNA (aDNA) studies that have helped to identify Mesolithic (hunter-gatherer) and Neolithic (farmers) genetic elements [6, 12–14]. Recently, ancient DNA and isotope analyses have shown that in Central Europe the two livelihoods, and also gene pools, existed in parallel for extended periods after the Neolithic influx of people and technology [15].

The contribution of Neolithic genes in the present European gene pool dominate, but in various degrees, in different parts of Europe [6]. Although variation between different ancient or contemporary DNA data sets can be partly explained by differences in quantity and in interpretive methodology (see [9, 16]), there appears to be regional variation that probably traces back to the local Neolithisation processes. In this context, the events in the western and northern extremes of Europe may have been totally different compared to the Central Europe (*cf.* [7]). Archaeological studies based on radiocarbon dating show that when reaching northern parts of Europe, the speed of the Neolithic advance has slowed down. Although this might be explained by the time required for development of crops adapted to northern latitudes, Isern et al. (2012) [17] suggested, based on simulation studies, that this has mainly been due to niche occupation by a sizeable population of Mesolithic people who based their subsistence on foraging. Competition on resources has reduced Neolithic population growth and delayed the advance and admixture [18]. If this hypothesis holds, it would entail that the Mesolithic and Neolithic genetic signatures could be more discernible in the marginal regions of Europe, as suggested by [7].

Finland has been a region that was among the last to adopt the sedentary agricultural way of life in Europe, perhaps even as late as 500 BC [19] although the issue is contentious (see e.g. [20]). Intriguingly, a genetic border separating South-Western and Eastern Finland has been identified in a number of human genetic studies, especially in the male-mediated Y-chromosomes [21–23]. This is rather exceptional, when contrasted against the largely clinal variation elsewhere in Europe [24]. The border localized by genes also appears to coincide with a medieval political border and, more importantly, with a border in many cultural features—but not with any apparent geographical borders [23].

Here, we present and explore a hypothesis that the geographic patterns actually denote a still existing border—genetic and cultural—between Neolithic farmers and Mesolithic hunter-gatherers. We focus on the haplogroup-level diversity of mtDNA and Y-chromosomes in Finland today. Obviously, ancient DNA data would hold a capacity for more direct analysis of the proposed scenario, but samples for these analyses are unavailable. Local conditions rarely allow survival of ancient biological samples over a thousand years, and one is forced to focus on contemporary genetic patterns.

Materials and Methods

The data comprises mitochondrial DNA and Y-chromosomal data, with the main focus on haplogroup distribution and diversity in different parts of Finland. The samples were collected from voluntary donors with written informed consents, limiting use of the samples to the

characterization of geographical patterns of neutral genetic variation in Finland. The consent forms were signed upon sampling and included information on study design and an option to decline the use of samples at any time in the future. It also specified publication of the results in anonymized form in scientific series (ethics committee approval: Helsinki University Central Hospital's Ethical Committee Dnro 329/13/03/00/2013). The samples were assigned based on donors' place of residence to 13 different subpopulations (Fig 1). For both mtDNA and Y-chromosomal (below) data basic diversity indices were estimated using ARLEQUIN v. 3.5.1.3 [25]. Haplogroup frequencies, number of haplotypes (A) and haplotype diversity (\hat{H}) were estimated for different haplogroups on the different geographic levels (all samples, regions and subpopulations). Differentiation was assessed by estimating pairwise F_{ST} and Φ_{ST} indices. The statistical dispersion values associated with the diversity indices was determined through 10 000 randomization steps.

Mitochondrial DNA data

Two different Finnish mitochondrial DNA sequence data sets were analyzed in this study:

1. Hypervariable regions 1 and 2 data (HVR1+2; $N = 832$), consisting of 643 base pairs (aligned length; sites 16024–16385 and 73–340) defining 384 unique haplotypes. The data was obtained from Palo et al. 2009 [23].
2. Complete mitochondrial sequences from Finland ($N = 367$) were obtained through GenBank searches ($N = 274$, the majority of these are published in [26, 27]) and from the 1000Genomes-project ($N = 93$; [28]). The accession numbers, references and associated information are listed in S1 Table.

The mtDNA haplogroup information was inferred from sequence data using mtDNA tree Build 15 (PhyloTree.org, 30 Sep 2012). For each haplotype, the haplogroup definition was projected with HAPLOGREP [29] [30]. For the HVR1+2 sequences, the phylogenetic sense of the inferred haplogroups was also assessed by inspecting the haplotype position in a phylogenetic tree by eye. For this end, unrooted maximum likelihood (ML) trees were constructed using MEGA v. 5.05 [31] assuming Tamura-Nei+ Γ substitution model with shape parameter $\alpha = 0.7$. As this merely aimed at checking the overall topology, the fit of the mutation model parameters were not formally tested, but the robustness of the topology was checked by constructing the trees also assuming a simpler Jukes-Cantor model.

The HVR1+2 mtDNA data was subdivided into two haplogroup clusters based on their inferred association with 1) Mesolithic hunter-gatherers (HUNT) including haplogroups U and V, and 2) Neolithic farmers (FARM) including haplogroups H, J, T and K. This clustering was based on results from [3]. A conservative assignment strategy was adopted, and all sequences that could not be assigned unambiguously were excluded (see Results).

For the HVR1+2 data haplotype and haplogroup frequencies as well as basic diversity indices (number of haplotypes A , haplotype diversity \hat{H} and nucleotide diversity π) were estimated within haplogroups and within haplogroup clusters HUNT and FARM.

Geographical differences in HUNT and FARM haplogroup frequencies within Finland were visualized by extrapolated maps constructed with MapInfo, and the patterns were tested by linear regression assuming a generalized linear model (GLM). Here, logistic regression analysis in R v.2.14.1, assuming binomial distribution and logit link function, was used to find a bipartition that minimized the product of the P -values of combined Hg-frequency differences of HUNT and FARM clusters. As a control for the validity of the partition, the results were contrasted with random non-continuous bipartitions.



Fig 1. Map of Northern Europe with Finland and the different subpopulations analyzed.

doi:10.1371/journal.pone.0130331.g001

In order to explore the demographic dynamics within the HUNT and FARM haplogroup clusters in Finland, complete genome mtDNA data was analyzed to infer past changes in (female) effective population sizes in haplogroups H ($N = 94$) and U ($N = 86$). These groups were chosen as they are considered to represent the hunter-gatherers (U) and farmers (H) well and because similar analyses in these groups have recently been published [32].

Bayesian Skyline Plots (BSPs) were constructed with the BEAST v. 1.7.4 software package programs [33]. The method uses Bayesian coalescence inference from sequence data and Markov Chain Monte Carlo (MCMC) sampling to produce posterior probability distributions for effective population sizes. In order to better fit the mutation model used in the genealogy building, only the coding region (bases 577–16023) sequence data was used in the analyses.

BEAUTi was used to generate input files for the BEAST runs. Six mutation models were fitted to the data, General Time Reversible (GTR) and Hasegawa-Kishino-Yano (HKY; [34]) with three alternative parameter combinations (with proportion of invariants p_{inv} , gamma distributed substitutions Γ , and both). The strongest support was obtained for GTR with proportion of invariants set to $p_{inv} = 0.80$, base frequencies estimated from the data, and a lognormal relaxed clock. A clock rate 1.69×10^{-8} substitutions⁻¹ site⁻¹ year⁻¹ [32] was assumed in the analyses.

For each sequence group, three independent MCMC chains were run for 40 or 80 million steps after a burn-in period with length of 10% of the actual run. Values were recorded every 10,000 steps. These parameters were chosen based on a number of trial runs.

LOGCOMBINER was used to combine the results from multiple runs and the results examined using TRACER v. 1.5 [35]. Proper mixing and convergence of runs were confirmed by evaluating effective sampling sizes (ESS) reported for the different parameters as well as by comparing results from independent runs. The ESS values denote the number of independent samples accepted, and $ESS > 200$ was used as a cut-off for acceptance.

To determine if the BSP model was the best method for reconstructing past population sizes it was compared to the constant population model using Bayes Factors calculated from the marginal likelihoods for each model, as implemented in TRACER. Strength for the BSP model was determined using guidelines on Bayes Factors provided in [36].

Y-chromosomal DNA data

As with the mtDNA, the Y-chromosomal data set combined Y-STR haplotype (17-locus AmpFI STR Yfiler) and haplogroup data ($N = 584$) from two different sources:

1. In-house data set ($N = 330$). For this data, sample collection, DNA extraction and Y-STR typing was performed as described in Palo et al. 2009 [23]. Haplogroup information was obtained through SNP typing (see below) with emphasis on the two main haplogroups in Finland, haplogroups N1c1 and I1 and their subhaplogroups. These two haplogroups represent approximately 90% of Finnish Y-chromosomes.
2. Y-chromosomal data obtained by data mining on the Family Tree website (data mining 12.9.2012, <http://www.familytreedna.com/>; $N = 254$). This data included 16 out of the 17 Y-STR loci in the AmpFI STR Yfiler set (barring DYS635) as well as the haplogroup designation defined by SNP-information.

For the Y-chromosomal haplogroup definitions, we follow the International Society of Genetic Genealogy (ISOGG; <http://www.isogg.org>) nomenclature published in 2015. For the in-house data set, the samples were first preclassified as belonging to haplogroup N1c1, I1 or other based on the Y-STR information using the Haplogroup Predictor algorithm (<http://www.hprg.com/hapest5/>) with “Northwest Europe” as a metapopulation prior. Haplogroup N-

predicted samples were genotyped for M46, M178 and L550, and haplogroup I1 samples were further genotyped for SNPs L22, L258 and L300. Some samples were also designated through comparison of haplotypes with previously typed samples (N1c1 and I1 subhaplogroups).

Real Time PCR genotyping of SNPs M46 was performed using TaqMan technology. TaqMan SNP Genotyping Master Mix, with one custom-ordered Genotyping Assay (rs34442126) including sequence-specific primers from Life Technologies (Carlsbad, CA, USA), was used to set up 13 μ l reactions including SNP Assay with 5.625 μ l of template DNA. Reactions were analyzed on ABI 7500 RT-PCR machine using cycling conditions consisting of a 10-min activation at 95°C, followed by 40 cycles of denaturing at 95°C for 15 s, and extension at 60°C for 1 min.

SNPs M178, L550, L22, L258, L300 were sequenced, with amplification conducted using 1x PCR buffer II (Life Technologies), 1.5 mM MgCl₂ (Promega, Madison WI, USA), 200 μ M dNTPs (Biofellow Oy via Oligomer, Helsinki, Finland), 2.5 U AmpliTaq Gold polymerase (Life Technologies), 6.5 μ g bovine serum albumin (Thermo Fisher Scientific, Waltham, MA, USA), and 0.2 μ M of each primer. Approximately 10 ng of genomic template DNA was added to the master mix in each reaction. Cycling conditions consisted of a 7-minute denaturation step at 95°C, followed by annealing and extension at 56°C and 68°C respectively for 33 cycles. For SNP M178 a lower annealing temperature of 51.7°C was used. Amplified fragments were purified enzymatically and sequenced using the PCR primers and BigDye Terminator v1.1 chemistry (Life Technologies). The sequencing reactions were purified using XTerminator Purification Kit (Life Technologies) and analysed on ABI Prism Genetic Analyzer 3130xl. Data were compiled using SEQUENCHER v.4.10 software (GeneCodes Inc., Ann Arbor, MI, USA).

Results

mtDNA: HVR1 and HVR2

The HVR1+2 data set consisted of 832 sequences and 384 unique haplotypes, showing an overall haplotype diversity of $\hat{H} = 0.993 \pm 0.001$. The haplogroup frequencies, deduced from the sequence data, are presented in in [Table 1](#). The overall haplogroup distribution in Finland was similar than in Western Europe, with H as a dominant haplogroup but also with relatively high occurrence of >20% for haplogroup U, and especially U5.

Out of the 832 haplotypes, 232 (27.9%) and 419 (50.4%) fell in the hunter-gatherer (HUNT; Hgs U, V) and farmer (FARM; H, T, K, J) groups, respectively. The remaining 173 samples represented haplogroups D, HV, I, N, R, W, X and Z.

Results from the GLM regression analysis showed greatest HUNT/FARM frequency differences between southern/western subpopulations AL, TU, HA, VA, UU, LMO and northern/eastern subpopulations MI, CF, KU, KY, NC, OU and LA (lowest product of p -values $P_{\text{HUNT}} * P_{\text{FARM}} = 9.85\text{E-}08$). This bipartition was several orders better than for any of the ten random non-continuous bipartitions tested ($P_{\text{HUNT}} * P_{\text{FARM}} = 0.002 \dots 0.313$), suggesting validity.

In the HUNT group, logistic regression estimate for U was -0.54, i.e. showing 54% lower occurrence in SW compared NE ($p = 0.0008$; [Fig 2](#)). However, in the subhaplogroup level, the pattern gets more diverse. The overall NE affinity for U can be attributed to strong eastern bias in subhaplogroup U5b, especially U5b1. This subhaplogroup shows clearly higher frequency in Finland than in most other European populations. However, haplogroup U5a shows a lower frequency but also contrasting geographical pattern. Haplogroup V and both its main subhaplogroups show eastern bias.

In the FARM group, the most significant geographical disparity was observed in haplogroup J (164% more in SW, $p = 1.05\text{E-}03$). All the other main FARM haplogroups show minor SW

Table 1. MtDNA haplogroup frequencies and basic diversity indices.

Hg	FINLAND						SOUTHWEST					NORTHEAST				
	N	f	A	Ĥ	π		N	f	A	Ĥ	π	N	f	A	Ĥ	π
HUNT	U	202	0.243	85	0.957 ±0.011	0.010 ±0.005	70	0.180	42	0.968 ±0.010	0.009 ±0.005	132	0.298	60	0.949 ±0.011	0.010 ±0.005
	V	30	0.036	14	0.830 ±0.063	0.004 ±0.003	12	0.031	10	0.955 ±0.057	0.006 ±0.004	18	0.041	7	0.726 ±0.096	0.003 ±0.002
	U+V	232	0.279	99	0.965 ±0.006	0.011 ±0.006	82	0.211	52	0.976 ±0.008	0.011 ±0.006	150	0.339	67	0.957 ±0.009	0.011 ±0.006
FARM	H	276	0.332	117	0.976 ±0.004	0.023 ±0.013	133	0.342	71	0.978 ±0.005	0.007 ±0.004	143	0.323	68	0.971 ±0.006	0.007 ±0.004
	J	46	0.055	27	0.941 ±0.024	0.009 ±0.005	33	0.085	22	0.911 ±0.044	0.009 ±0.005	13	0.029	9	0.936 ±0.051	0.010 ±0.006
	T	51	0.061	33	0.977 ±0.009	0.010 ±0.005	28	0.072	21	0.976 ±0.016	0.009 ±0.003	23	0.052	18	0.976 ±0.020	0.011 ±0.006
	K	46	0.055	22	0.917 ±0.027	0.004 ±0.003	24	0.062	16	0.931 ±0.039	0.004 ±0.003	22	0.050	10	0.883 ±0.041	0.004 ±0.003
	H+J +T+K	419	0.504	199	0.987 ±0.002	0.011 ±0.006	218	0.560	130	0.989 ±0.002	0.011 ±0.006	201	0.454	105	0.983 ±0.003	0.011 ±0.006
ALL	-	832	1.000	384	0.993 ±0.001	0.012 ±0.006	389	0.468	236	0.994 ±0.001	0.013 ±0.007	443	0.532	225	0.990 ±0.001	0.012 ±0.006

Hg = haplogroup, N = number of samples, f = haplogroup frequency, A = number of unique haplotypes, Ĥ = within-Hg haplotype diversity, π = within-Hg nucleotide diversity.

doi:10.1371/journal.pone.0130331.t001

bias, but with more fragmented subhaplogroup patterns. Haplogroups H1a and H1f show relatively strong NE bias, but H2 a western bias.

Interestingly, the strength of haplogroups NE bias in Finland correlates with the difference of haplogroup ages in Near Eastern and European populations estimated in [37]. Haplogroups with older estimated ages in Europe than in Near East show stronger eastern bias in Finland, and vice versa (Fig 3), with a clear-cut overall correlation $R^2 = 0.983$. Spearman's rank correlation gives the same signal, with $r_s = 0.9643$ ($p < 0.01$).

mtDNA: complete genomes

Altogether 367 complete mitochondrial genomes were obtained from the Genbank and 1000Genomes project. This data showed in general similar overall haplogroup distribution than the control region data.

The Bayesian skyline plots show substantially smaller effective population sizes for haplogroups in the HUNT than in the FARM group in Finland, as well as for individual haplogroups in these groups (Fig 4). The FARM haplogroups do show a relatively early population growth: assuming a mutation frequency of 1.69×10^{-8} substitutions⁻¹ site⁻¹ year⁻¹ this occurred c. 9 kya, slightly later than estimated in e.g. [32, 38] for Western European H haplotypes. However, the HUNT group BSPs do show a population growth, but relatively late c. 4 kya. In relative terms, the approximate start of population growth in the HUNT haplogroups occur at 0.2 times the FARM group growth start time.

Y-chromosomes

Y-chromosomal STR haplotype and haplogroup data was obtained for altogether 584 Finnish males. Among this data 294 unique Y-STR haplotypes were observed corresponding an overall

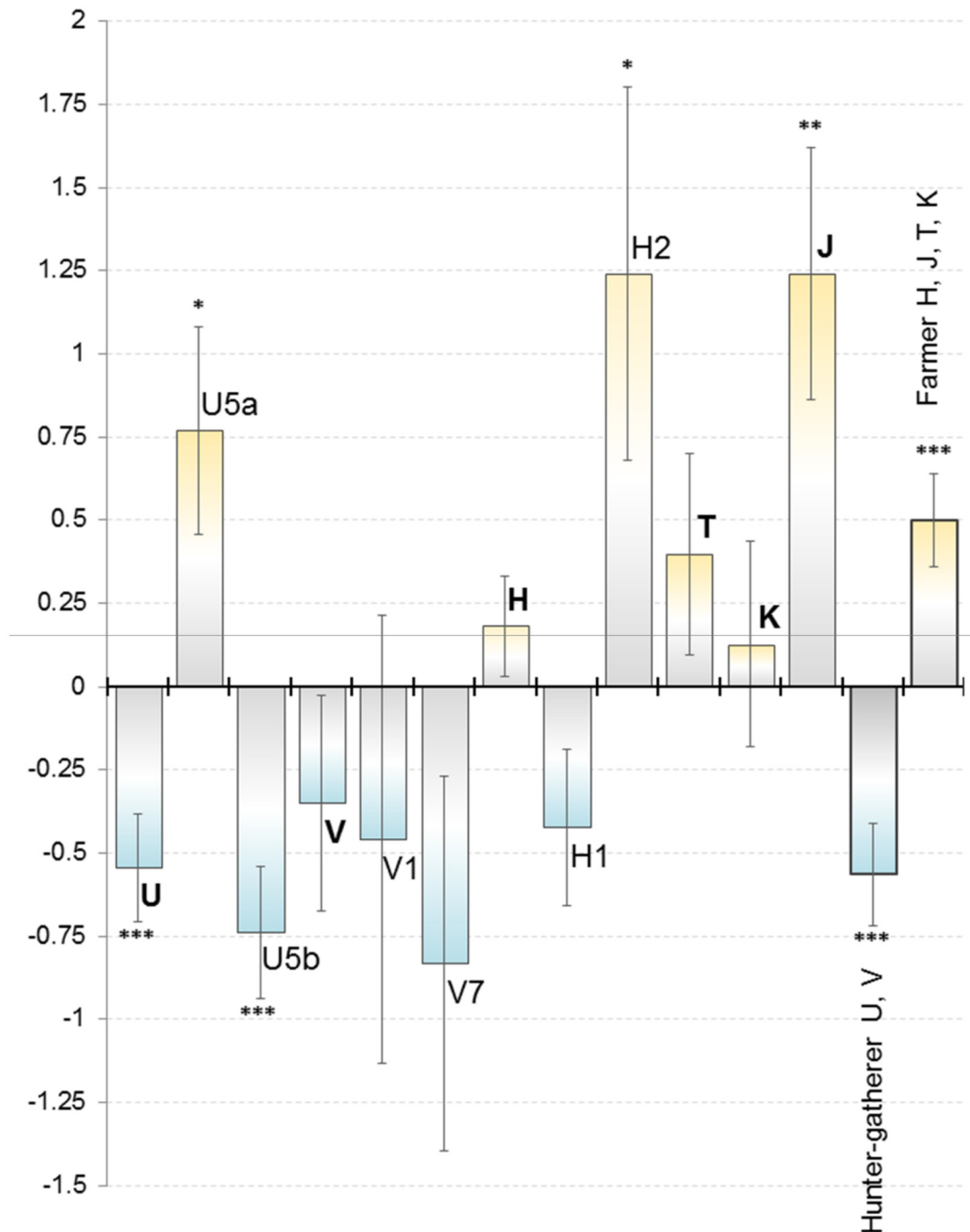


Fig 2. Logistic regression estimates representing the difference in haplogroup frequencies between the SW & NE subpopulations. Above X-axis: SW dominance, below: NE dominance. The results are shown for division (cf. Fig 2) that maximized the difference. Error bars denote standard deviation, statistical significance is marked with stars. No statistically significant values were obtained in randomized, non-continuous divisions.

doi:10.1371/journal.pone.0130331.g002

haplotype diversity $\hat{H} = 0.9863 \pm 0.0019$ (SD) (Table 2). Ninety-one percent of the samples fell into haplogroups N1c1 ($N = 289$) and I1 ($N = 242$).

The samples were assigned to regions NE and SW, which were defined based on haplotype information in Palo et al. 2009 [23]. Regional haplogroup frequencies are consistent with previous studies [22], with N-frequencies highest in eastern subpopulations, and decreasing moving

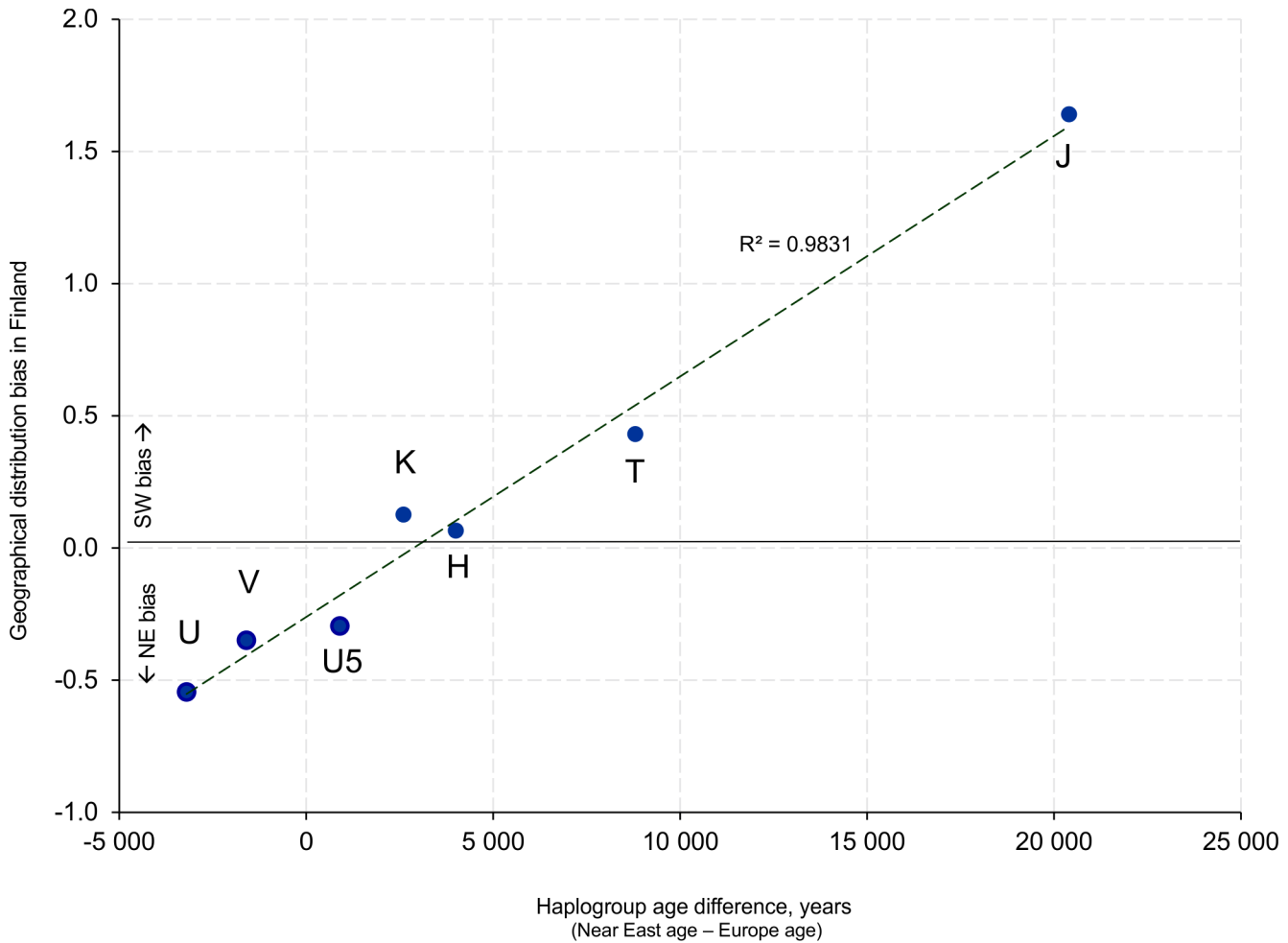


Fig 3. The strength of NE bias for the main haplogroups plotted against the difference of haplogroup ages in Near Eastern and European populations estimated in (Richards, et al. 2000). Haplogroups with older estimated ages in Europe than in Near East show stronger eastern bias in Finland, and vice versa.

doi:10.1371/journal.pone.0130331.g003

West. The opposite pattern is seen in the I-haplogroup; high frequency on the western coast and low in the east. The ratio of N1c1/I1 frequencies show strikingly similar spatial pattern with the ratio of mtDNA HUNT/FARM haplogroup frequencies (Fig 5).

The overall haplogroup distribution disparity is reflected in the differentiation estimates: considering the N1c1 and I1 data (representing 91% of the total data), significant differentiation between the NE and SW was observed on the allelic ($\Phi_{ST} = 0.107$), but not on the haplotypic level ($F_{ST} = 0.010$). Nevertheless, no significant differences were observed in the diversity between the regions NE and SW ($\hat{H} = 0.9733 \pm 0.0061$ vs. $\hat{H} = 0.9867 \pm 0.0024$). Within the haplogroups the haplotype diversity was similar as well: N1c1-haplogroup showed 147 ($\hat{H} = 0.9657 \pm 0.0067$) and I1-haplogroup 106 ($\hat{H} = 0.9699 \pm 0.0048$) unique haplotypes. No differences in the haplotype diversities within the N- and I-haplogroups were observed in NE ($\hat{H}_{N1c1} = 0.9473 \pm 0.0132$; $\hat{H}_I = 0.9589 \pm 0.0139$) or in SW ($\hat{H}_{N1c1} = 0.978 \pm 0.0062$ and $\hat{H}_I = 0.9669 \pm 0.0069$ respectively).

However, there was a clear difference in the number of observed subhaplogroups between the N1c1 and I1. Within these two major clades, most Finnish Y-chromosomes fall into subhaplogroups of N1c1a1a-L1026 and I1a-DF29 (Tables 2 and S3). Of the other observed haplogroups,

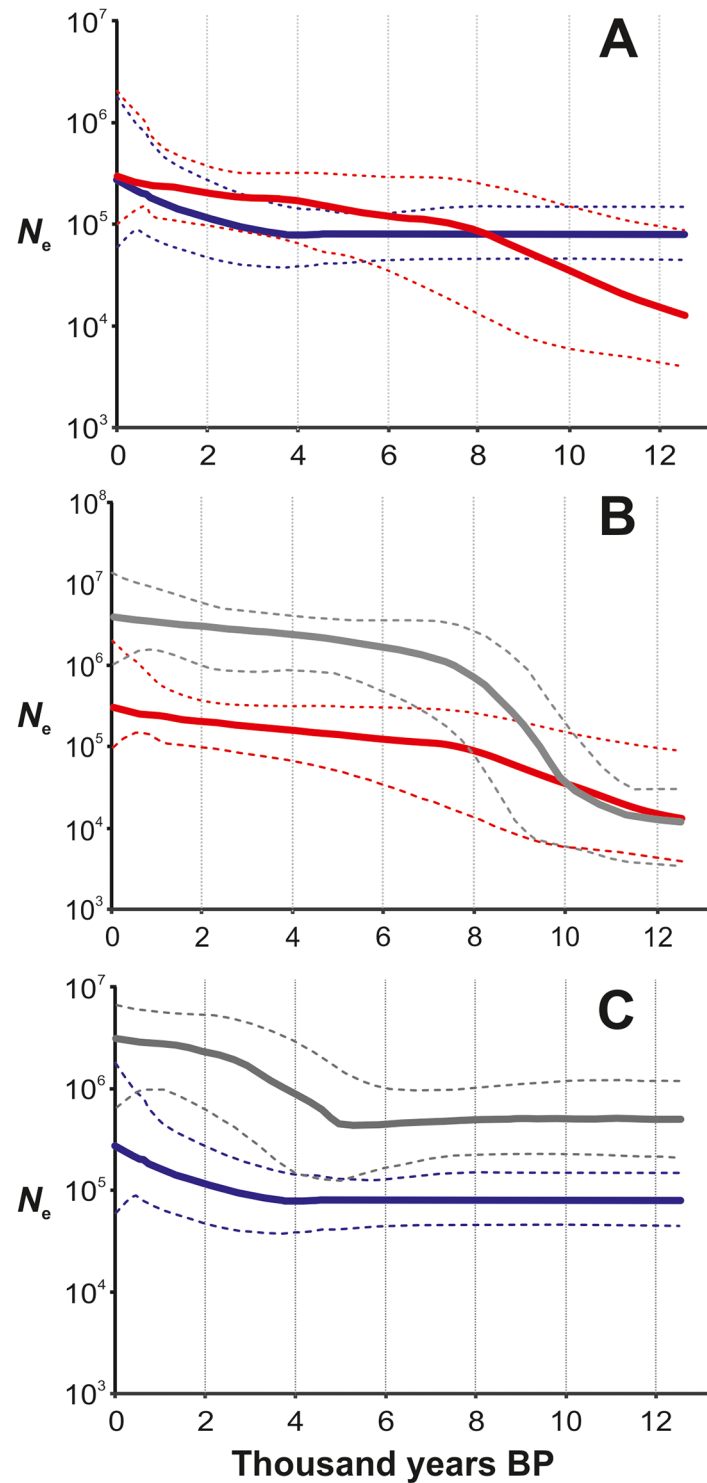


Fig 4. Bayesian Skyline Plots for mtDNA haplogroups H and U in Finland, with European reference data from Fu et al. 2012. The hatched lines denote 95% confidence intervals. A: MtDNA haplogroups H (red) and U (blue) in Finland. B: Haplogroup H in Finland (red) and in Europe (grey). C: Haplogroup U in Finland (blue) and in Europe (grey).

doi:10.1371/journal.pone.0130331.g004

Table 2. Y-chromosomal haplogroup frequencies and basic diversity indices. Note that 35 haplotypes for which the sampling location in Finland was unknown are included in “Finland”.

Hg	FINLAND				SOUTHWEST				NORTHEAST			
	<i>N</i>	<i>f</i>	<i>A</i>	\hat{H}	<i>N</i>	<i>f</i>	<i>A</i>	\hat{H}	<i>N</i>	<i>f</i>	<i>A</i>	\hat{H}
N1c	289	0.495	147	0.966 ± 0.007	115	0.376	71	0.978 ± 0.006	162	0.667	84	0.947 ± 0.013
I1	242	0.414	106	0.970 ± 0.005	170	0.556	84	0.967 ± 0.007	58	0.239	33	0.959 ± 0.014
I2	10	0.017	10	1.000 ± 0.045	2	0.007	2	1.000 ± 0.500	6	0.025	6	1.000 ± 0.096
R1a	22	0.038	16	0.931 ± 0.046	11	0.036	11	1.000 ± 0.039	9	0.037	3	0.556 ± 0.165
Q	1	0.002	1	1	1	0.003	1	1	0	0.000	0	0
R1b	16	0.027	10	0.867 ± 0.079	6	0.020	5	0.933 ± 0.122	8	0.033	3	0.464 ± 0.200
E1b1	3	0.005	3	1.000 ± 0.272	1	0.003	1	1	0	0.000	0	0
J	1	0.002	1	1	0	0.000	0	0	0	0.000	0	0
ALL	584	1.000	294	0.986 ± 0.002	306	0.524	175	0.987 ± 0.002	243	0.416	129	0.973 ± 0.006

Hg = haplogroup, *N* = number of samples, *f* = haplogroup frequency, *A* = number of unique haplotypes, \hat{H} = within-Hg haplotype diversity.

doi:10.1371/journal.pone.0130331.t002

I1a1b-L22, with an overall frequency of 29% (71% of the I1 haplotypes), is commonly considered the major “Nordic” branch. In the Family Tree data I1a2a1b-Z73 is found mostly in Finland and Northern Scandinavia, and haplogroups found almost exclusively in Finland include I1a1b3a-L287, I1a1b3a1-L258, I1a1b3a1a-L296 and I1a1b4-L300.

Geographically some Finnish terminal haplogroups show some regional association within Finland; I1a2a-Z59, I1a2a1-Z60, and I1a3-Z63, as well as I1a1b3a1a-L296 and I1a1b4-L300 were found in the Southwestern part of the country. However, two of the lineages unique to Finland, I1a1b3a-L287 and I1a1b3a1-L258, could be observed throughout the country.

Discussion

In contrast to Y-chromosome, mtDNA haplotype-level assessments have until now failed to identify clear geographical differences within Finland (see e.g. [23]). Here a haplogroup-level analysis revealed spatial patterns that are very similar for both uniparental markers (Fig 5). In the mtDNA data, the SW/NE divergence can be accounted to stem from the frequency differences of haplogroups that have been associated with farmers (H,J,T,K; more common in the SW) and with hunter-gatherers [3]. There is not enough ancient DNA data to allow such an association for the Y-chromosomal data, largely due to poorer preservation of Y-chromosomes in archaeological material. However, haplogroup I has probably arrived later in Finland, and can be thus associated with farmers, whereas the opposite is true for the N haplogroup (see below).

The genetic border in Finland, similar in both the mtDNA and Y-chromosomal data, is in its sharpness rather exceptional in Europe, and cannot be explained by any observable migration barriers. Roewer et al. [39] observed a similar Y-chromosomal border in Central Europe, and interpreted this to stem from political events in Europe since the Middle Ages. Instead, here we propose that the Finnish genetic border represents vestiges of an ancient border between two modes of subsistence, farming and hunter-gathering. It is very likely that this signal has been dampened by internal migration especially during the last century, but its survival until the present day speaks for its strength in the past. In what follows we elaborate this from the viewpoint of the two marker classes.

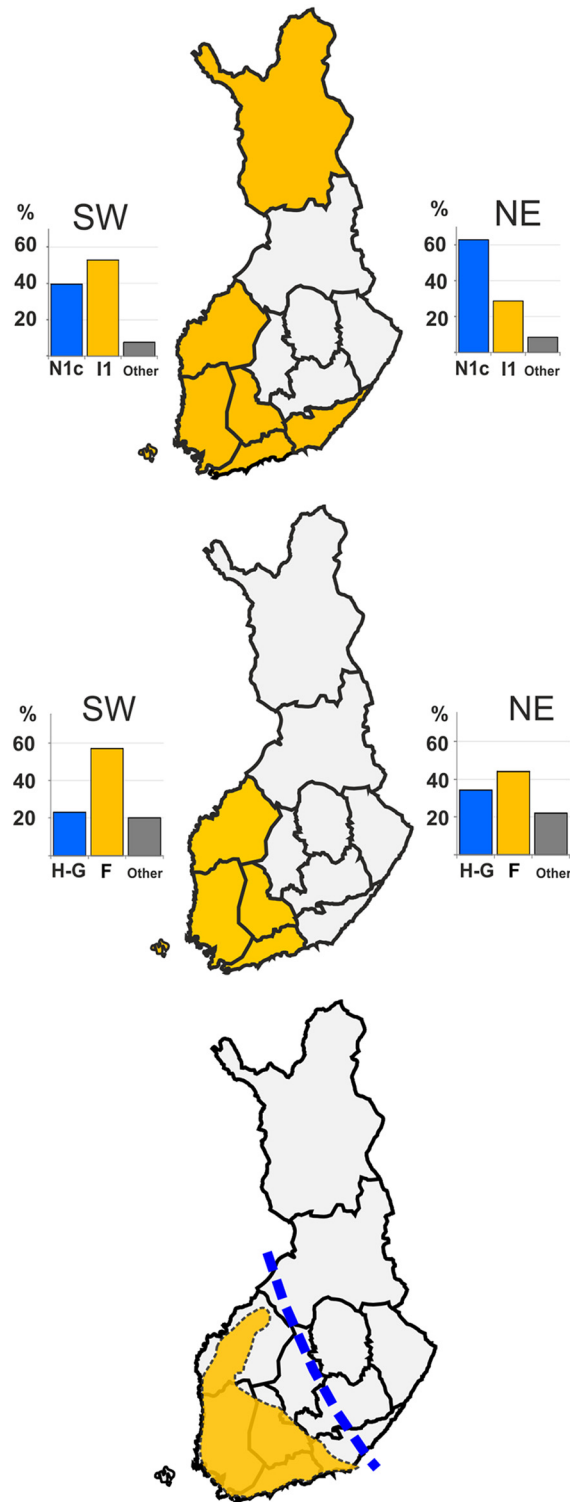


Fig 5. Spatial patterns in different marker classes and in archaeological evidence for Combed Ware Culture in Finland. A: Division maximizing Y-STR haplotype differences, and frequencies of the main Y-haplogroups in Finland. B: Division maximizing the difference between Hunter-Gatherer (H-G: hgs U & V) and Farmer (F: hgs H, T, J & K) mtDNA haplogroups and their frequencies. C: The extent of Corded-Ware Culture (CWC; data from www.nba.fi) in Finland, and the approximate location of the first political border between Sweden and Novgorod (AD 1323; hatched blue line).

doi:10.1371/journal.pone.0130331.g005

mtDNA

Overall, the mtDNA haplogroups associated with hunter-gatherers and farmers show opposite frequency trends along a SW-NE axis in Finland. Hunter-associated hgs are more common in Eastern and Northern Finland. This applies to mitochondrial hgs U and V (especially U5b). Haplogroup U5, together with U8, is an old haplogroup that arrived to Europe at least 30 kya [40] and, as of yet, practically all ancient DNA studies have proven its prevalence in pre-Neolithic hunter-gatherers in Europe [3], [41], [42], also in Scandinavia [43], see also [3].

The farmer haplogroups H, J, T and K show, in turn, a significant SW-bias in Finland. This is especially strong in haplogroup J. The absolute majority (85%) of J haplotypes belong to sub-haplogroup J1c2. J1 is a relatively young European haplogroup, which has been observed in Neolithic remains [44] and references therein. Specifically, J1c has been reported from 5–5.5 ky old Neolithic remains from Iberian peninsula [6].

In our data, haplogroup H shows conflicting subhaplogroup patterns. H as a whole is a very diverse group, and its evolutionary history is complex. The most common subhaplogroups in Finland, H1 and H2, show opposing SW-NE trends, with H1 increasing towards NE. Although the current haplogroup H diversity in Europe has been associated with Neolithic cultures [38] and is very rare in Northern European hunter-gatherers [41, 43], the age of the basal H1 subhaplogroup has been dated back to Pleistocene/Holocene boundary c. 11 ky [45], similar to H3 and V. It has also, unlike the other H subgroups, shown continuity from the early Neolithic to the present in Europe [38]. According to Achilli et al. H1 spread from the Franco-Cantabrian refugium with post-glacial expansion of hunter-gatherers [45]. Analyzing the data more detailed in a subhaplogroup level could elucidate the patterns even further, but was out of the scope of this paper.

Y-chromosomes

The focal parent haplogroups N-M231 and I-M170 both have their origins in the Paleolithic era, and both are strongly associated with the hunter-gatherer lifestyle. N-M231 originated in Southeast Asia approximately 20 kya, while its subgroup N1c1-M46 arose 12 kya [46], [47], and may have appeared in eastern Finland as early as the immediate post-glacial era [47] as well as in later waves associated with the Finno-Ugric speakers; the Mesolithic hunter-gatherer Kunda and Comb Ceramic cultures [48]. [22] [49] [50]. In Finland, N1c1 is the most common Y-haplogroup [51] [50] with an overall occurrence of 58% and with highest concentration [70.9%] in Northern Carelia in the east [20].

I-M170 arose in the Balkans approximately 22 kya [52] and the splitting of this parental clade into subhaplogroups I1-M253, I2a2a-P37.2, and I2a1-M223 also occurred on the European continent [53] [52]. The subhaplogroup I1-M253 and its further branch I1a-DF29 are most prominent in the Scandinavian countries and western Finland, with greatest frequency of I1-M253 in central Sweden (52%) [54] [55], ISOGG. In Finland I1 shows highest concentration in the western provinces (40%) and lowest in eastern Finland (19%) [22]. The arrival of later I1 subhaplogroups to Finland seems to coincide temporally with the arrival of domesticated animals, according to osteological evidence [56].

Signal of ancient genetic border

The haplogroup distribution in both markers implies that the SW-NE difference still retained in the contemporary genetic diversity in Finland represents an ancient edge of Neolithic farmer advance. In Central Europe, the amalgamation of the Neolithic and Mesolithic gene pools has been more thorough, most likely due to longer time and environment more favourable for the immigrating farming technology. This is in line with the recent ancient DNA results, showing

that the genome of 7 000-year-old Mesolithic individual from Northern Spain associated closer to present day genomes from Finland than to any other Europeans included in the study [57].

Like in many other European populations, the recent demographic history of Finns entails strong population growth and, especially in the 20th century, internal migration. The fact that vestiges of an ancient boundary are still discernible suggests that this pattern has in the past been very distinct and that it endured to a later date in the North of Europe. Indeed, archaeological evidence alludes to prolonged coexistence of farmer and hunter-gatherer populations in the north [43, 58]. This is probably due to the later transition to farming in NE Europe (e.g. [59]). Considering the Near Eastern origin and spreading along a southeast-northwest axis into Europe [4], the Fennoscandian region has been the “Ultima Thule”, the northern fringe colonized last by the Neolithic farmers. In Northern Europe the Neolithic advance slowed down. There are probably a number of reasons for this (such as time needed for development of locally adapted crops, e.g. [60]), but one important reason is space competition between farmers and the indigenous Mesolithic populations [17]. Space competition is a process restricting colonization, well-known from a post-glacial recolonization of Europe by wide variety of taxa [61] [62].

These processes gain support from the effective population sizes in haplogroups H, representing the Neolithic farmers, and U representing the hunter-gatherers (see e.g. [32], [43], [38]). When compared to the European averages in these groups [32], the Bayesian Skyline Plots show overall similarity but with two exceptions. Firstly, the estimated sizes are an order of magnitude lower in Finland, and lack patterns of rapid growth, which reflects the lower carrying capacities of the northern latitudes. In fact, the U haplogroup effective size remains rather constant throughout most of the post-glacial. Secondly, and perhaps more importantly, the hunter-gatherer haplogroup U N_e starts to grow only c. 3000 years ago—some 2000 years later than in Europe. As speculated also for Europe, this growth probably denotes the adoption of agricultural technology by the hunter-gatherers either independently or after population admixture. Note that although the absolute values of effective size and time of events can be questioned, the relative difference between results here and in [32] should hold as the same mutation rate has been assumed. Unfortunately, the available complete mtDNA genome data did not allow comparisons between different regions of Finland.

In the North-East of Europe, a short growing season and relatively unproductive soil were not favorable for farming. In the same time, the western edge of the taiga zone offered plenty of game and fish. Thus, unlike in Western Europe, the colonizing Neolithic farming communities remained relatively small, and probably were assimilated to hunter communities rather than vice versa. Gradual admixture between the arriving Neolithic farmers and relatively numerous Mesolithic hunters in a limited area (SW Finland) could plausibly explain the fact that, unlike most other European populations, Finns do not speak an Indo-European language (*cf.* [63]). Note that the Neolithization process has been connected to the spread of Indo-European languages into Western Europe, although the process might have been complex [64, 65].

The region identified with mtDNA and Y-chromosomal data (see also [66]) matches spatially with the extent of archaeological finds associated to the Corded-Ware Culture (CWC) in Finland. The CWC flourished in a wide area south of the Baltic Sea c. 4.9–4.3 kya. The CWC people based their subsistence on pastoralism and sedentary farming and spoke Indo-European languages (see [48] and references therein).

The CWC spread into SW Finland c. 4.5 kya, which temporally coincides with the advent of farming in Finland. The geographical NE edge of the CWC in Finland has been sharp, dividing Finland into two cultural spheres (Halinen P, In: Suomen historian kartasto). While causality is hard to prove directly, the geographical boundary patterns between the genes and culture are strikingly similar, and can also be seen in a number of cultural features, some of which have

persisted into modern times [67]. Interestingly, the first political border in Finland, the 1323 AD agreement between Sweden and Novgorod, also roughly followed the CWC NE edge and the genetic boundary identified in Finland. The reasons for the localization of this boundary may be ecological: the soil most amenable to field farming can be found in SW Finland. Also the vegetation zones and length of the thermic growing seasons changes along a SW-NE axis in Finland (for maps see Finnish Meteorological Service <http://ilmatieteentaitos.fi/terminen-kasvukausi>, in Finnish).

While there is plenty of circumstantial evidence suggesting that the CWC has had a strong influence in SW Finland, the identity of the indigenous hunters is more enigmatic. Finland was colonized soon after deglaciation, probably by “converging human groups gradually taking over deglaciated territories” [42]. The most prominent culture in Finland at the time of the CWC arrival was however the Comb Ceramic Culture (CCC) that has been dated back to c. 6.0 kya and extended to the whole of Finland. The extent of the CCC, matching with the extent of Finno-Ugric-speaking populations, suggests that they spoke Finno-Ugric language. Despite the use of ceramics, their subsistence was based on hunting and foraging.

The scenario postulated here is to some extent similar to the ‘language replacement’ theory [68], suggesting that the invading Neolithic population assimilated the local Fenno-Ugric language. However, questions arise especially of the role of the Saami of northern Finland, Sweden, Norway and Russia (Lapland), which also show high frequencies of HUNT haplogroups of this study (esp. mtDNA U5b and Y N1c1) but a clearly distinct overall genetic composition. The similarities between the Saami and Finns today stem probably from admixture. [69] have suggested that the Saami have contributed to the Finnish gene pool especially in the regions directly south of Lapland. An alternative intriguing possibility, fitting well with the uniparental marker data, is that the present-day Saami in fact represent a population admixture between the Palaeolithic people that colonized the north of Europe via the Norwegian coastal corridor as early as c. 11 kya, and the Mesolithic Finno-Ugric Comb Ceramic Culture. This would plausibly explain the conflicting Franco-Cantabrian/ Asian genetic signals, especially the high frequencies of mtDNA U5b1 and Y-chromosomal N1c1 in the Saami gene pool [50, 70–72] (S2 and S3 Tables) and the Finno-Ugric language spoken by the Saami. Formal investigation of this question, however, is out of the focus of this article.

In conclusion, the haplogroup-level analysis of mtDNA and Y-chromosomal markers indicates a contemporary genetic boundary that most likely denotes the limes of Neolithic advance. The persistence of these genetic signals complies with archaeological evidence and simulation studies showing the late arrival farmers in the north of Europe, and subsequent extended coexistence of farmers and hunters in this area.

Supporting Information

S1 File. Original template for Fig 1, from (Palo et al. 2009). Palo JU, Ulmanen I, Lukka M, Ellonen P, Sajantila A. Genetic markers and population history: Finland revisited. *Eur J Hum Genet.* 2009;17:1336–46.

(PDF)

S1 Table. Complete Finnish mitochondrial sequences (N = 367) analyzed in the study.

(PDF)

S2 Table. Mitochondrial regional association and haplogroup classification. Population abbreviations coincide with Finnish counties Mikkeli (MI), Central Finland (CF), Kuopio (KU), Kymenlaakso (KY), Northern Carelia (NC), Oulu (OU), Lappi (LA), Åland (AL), Turku (TU),

Häme (HA), Vaasa (VA), Uusimaa (UU) and Larsmä (LMO).
(XLSX)

S3 Table. Y-chromosomal regional association, haplogroup classification, and Y-haplotypes. Samples designated SG and L are in-house data, while FT samples have been obtained from the Family Tree website (www.familytreedna.com). Identification numbers for FT samples are taken from the individual Family Tree testing kit number. Population abbreviations are congruent to those in mitochondrial Table S2. The abbreviation “ht” refers to samples in which the haplogroup has been predicted from the haplotype.

(XLSX)

S4 Table. Main haplogroup frequencies in Finland estimated from complete mtDNA and HV1+2 data sets.

(PDF)

Acknowledgments

The authors wish to thank Prof. Jukka Corander for his invaluable help with the BEAST analyses, Dr. Minttu Hedman for her help during all phases of the study and Dr. Jari Haukka for his help in the statistical multivariate analysis. We are also indebted to the two anonymous referees for their helpful comments.

Author Contributions

Conceived and designed the experiments: AMN MP SÖ TS PO AS JUP. Performed the experiments: AMN MP SÖ TS PO AS JUP. Analyzed the data: AMN MP SÖ TS PO AS JUP. Contributed reagents/materials/analysis tools: AMN MP SÖ TS PO AS JUP. Wrote the paper: AMN MP SÖ TS PO AS JUP.

References

1. Ammermann AJ, Cavalli-Sforza LL. The Neolithic transition and the genetics of populations of Europe. Princeton: Princeton University Press; 1984. 176 p.
2. Cavalli-Sforza LL, Menozzi P, Piazza A. Demic expansions and human evolution. *Science*. 1993; 259(5095):639–46. PMID: [8430313](https://pubmed.ncbi.nlm.nih.gov/8430313/)
3. Pinhasi R, Thomas MG, Hofreiter M, Currat M, Burger J. The genetic history of Europeans. *Trends Genet*. 2012; 28:496–505. doi: [10.1016/j.tig.2012.06.006](https://doi.org/10.1016/j.tig.2012.06.006) PMID: [22889475](https://pubmed.ncbi.nlm.nih.gov/22889475/)
4. Lemmen C, Gronenborn D, Wirtz KW. A simulation of the Neolithic transition in Western Eurasia. *J Archaeol Sci*. 2011; 38(12):3459–70.
5. Fort J. Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. *P Natl Acad Sci USA* 2012; 109(46):18669–73. doi: [10.1073/pnas.1200662109](https://doi.org/10.1073/pnas.1200662109) PMID: [23112147](https://pubmed.ncbi.nlm.nih.gov/23112147/)
6. Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Marti M, et al. Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proc R Soc Lond B Biol Sci*. 2007; 274(1622):2161–7. PMID: [17609193](https://pubmed.ncbi.nlm.nih.gov/17609193/)
7. Tresset A, Vigne J-D. Last hunter-gatherers and first farmers of Europe. *C R Biol*. 2011; 334(3):182–9. doi: [10.1016/j.crv.2010.12.010](https://doi.org/10.1016/j.crv.2010.12.010) PMID: [21377612](https://pubmed.ncbi.nlm.nih.gov/21377612/)
8. Bentley RA, Chikhi L, Price TD. The Neolithic transition in Europe: Comparing broad scale genetic and local scale isotopic evidence. *Antiquity*. 2003; 77(295):63–6.
9. Chikhi L, Nichols RA, Barbujani G, Beaumont MA. Y genetic data support the Neolithic demic diffusion model. *P Natl Acad Sci USA*. 2002; 99(17):11008–13. PMID: [12167671](https://pubmed.ncbi.nlm.nih.gov/12167671/)
10. Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, et al. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet*. 2001; 69(4):844–52. PMID: [11517423](https://pubmed.ncbi.nlm.nih.gov/11517423/)

11. Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, et al. The genetic legacy of paleolithic Homo sapiens sapiens in extant Europeans: A Y chromosome perspective. *Science*. 2000; 290(5494):1155–9. PMID: [11073453](#)
12. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, et al. Ancient DNA from European early Neolithic farmers reveals their near eastern affinities. *Plos Biol*. 2010; 8(11). e1000536 doi: [10.1371/journal.pbio.1000536](#) PMID: [21085689](#)
13. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, et al. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*. 2005; 310(5750):1016–8. PMID: [16284177](#)
14. Sanchez-Ouinto F, Schroeder H, Ramirez O, Avila-Arcos MC, Pybus M, Olalde I, et al. Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr Biol*. 2012; 22(16):1494–9. doi: [10.1016/j.cub.2012.06.005](#) PMID: [22748318](#)
15. Bollongino R, Nehlich O, Richards MP, Orschiedt J, Thomas MG, Sell C, et al. 2000 years of parallel societies in Stone Age Central Europe. *Science*. 2013; 342(6157):479–81. doi: [10.1126/science.1245049](#) PMID: [24114781](#)
16. Currat M, Excoffier L. The effect of the Neolithic expansion on European molecular diversity. *Proc R Soc Lond B Biol Sci*. 2005; 272(1564):679–88. PMID: [15870030](#)
17. Isern N, Fort J. Modelling the effect of Mesolithic populations on the slowdown of the Neolithic transition. *J Archaeol Sci*. 2012; 39(12):3671–6.
18. Isern N, Fort J, Vander Linden M. Space competition and time delays in human range expansions. Application to the Neolithic transition. *PLoS ONE*. 2012; 7(12). doi: [10.1371/journal.pone.0051106](#)
19. Lahtinen M, Rowley-Conwy P. Early farming in Finland: was there cultivation before the Iron Age (500 BC)? *Eur J Archaeol*. 2013; 16(4):660–84.
20. Alenius T, Mökkönen T, Lahelma A. Early farming in the northern boreal zone: reassessing the history of land use in southeastern Finland through high-resolution pollen analysis. *Geoarchaeology*. 2013; 28(1):1–24.
21. Kittles RA, Perola M, Peltonen L, Bergen AW, Aragon RA, Virkkunen M, et al. Dual origins of Finns revealed by Y chromosome haplotype variation. *Am J Hum Genet*. 1998; 62(5):1171–9. PMID: [9545401](#)
22. Lappalainen T, Koivumäki S, Salmela E, Huoponen K, Sistonen P, Savontaus ML, et al. Regional differences among the Finns: a Y-chromosomal perspective. *Gene*. 2006; 376(2):207–15. PMID: [16644145](#)
23. Palo JU, Ulmanen I, Lukka M, Ellonen P, Sajantila A. Genetic markers and population history: Finland revisited. *Eur J Hum Genet*. 2009; 17:1336–46. doi: [10.1038/ejhg.2009.53](#) PMID: [19367325](#)
24. Lao O, Lu TT, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, et al. Correlation between genetic and geographic structure in Europe. *Curr Biol*. 2008; 18(16):1241–8. doi: [10.1016/j.cub.2008.07.049](#) PMID: [18691889](#)
25. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010; 10(3):564–7. doi: [10.1111/j.1755-0998.2010.02847.x](#) PMID: [21565059](#)
26. Moilanen JS, Finnilä S, Majamaa K. Lineage-specific selection in human mtDNA: lack of polymorphisms in a segment of MTND5 gene in haplogroup J. *Mol Biol Evol*. 2003; 20(12):2132–42. PMID: [12949126](#)
27. Soini HK, Moilanen JS, Finnilä S, Majamaa K. Mitochondrial DNA sequence variation in Finnish patients with matrilineal diabetes mellitus. *BMC Res Notes*. 2012; 5:350. doi: [10.1186/1756-0500-5-350](#) PMID: [22780954](#)
28. Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491(7422):56–65. doi: [10.1038/nature11632](#) PMID: [23128226](#)
29. Kloss-Brandstatter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, et al. HaploGrep: A fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat*. 2011; 32(1):25–32. doi: [10.1002/humu.21382](#) PMID: [20960467](#)
30. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009; 30(2):E386–E94. doi: [10.1002/humu.20921](#) PMID: [18853457](#)
31. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28(10):2731–9. doi: [10.1093/molbev/msr121](#) PMID: [21546353](#)
32. Fu Q, Rudan P, Paeaebo S, Krause J. Complete mitochondrial genomes reveal Neolithic expansion into Europe. *PLoS ONE*. 2012; 7(3).
33. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012; 29(8):1969–73. doi: [10.1093/molbev/mss075](#) PMID: [22367748](#)

34. Hasegawa M, Kishino H, Yano T. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol.* 1985; 22:160–74. PMID: [3934395](#)
35. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 2007; 7. doi: [10.1186/1471-2148-7-214](#)
36. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* 1995; 90(430):773–95.
37. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, et al. Tracing European founder lineages in the near eastern mtDNA pool. *Am J Hum Genet.* 2000; 67(5):1251–76. PMID: [11032788](#)
38. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, et al. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun.* 2013; 4:1764. Epub 2013/04/25. doi: [10.1038/ncomms2656](#) PMID: [23612305](#)
39. Roewer L, Croucher PJP, Willuweit S, Lu TT, Kayser M, Lessig R, et al. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Hum Genet.* 2005; 116(4):279–91. PMID: [15660227](#)
40. Soares P, Achilli A, Semino O, Davies W, Macaulays V, Bandelt H-J, et al. The archaeogenetics of Europe. *Curr Biol.* 2010; 20(4):R174–R83. doi: [10.1016/j.cub.2009.11.054](#) PMID: [20178764](#)
41. Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, et al. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science.* 2009; 326(5949):137–40. doi: [10.1126/science.1176869](#) PMID: [19729620](#)
42. Nunez M. A model for the early settlement of Finland. *Fennoscandia Archaeologica.* 1987; 4:3–18.
43. Malmström H, Gilbert MTP, Thomas MG, Brandström M, Storå J, Molnar P, et al. Ancient DNA reveals lack of continuity between Neolithic hunter-gatherers and contemporary Scandinavians. *Curr Biol.* 2009; 19(20):1758–62. doi: [10.1016/j.cub.2009.09.017](#) PMID: [19781941](#)
44. Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, et al. Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am J Hum Genet.* 2012; 90(5):915–24. doi: [10.1016/j.ajhg.2012.04.003](#) PMID: [22560092](#)
45. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, et al. 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.* 2004; 75(5):910–8. PMID: [15382008](#)
46. Rootsi S, Zhivotovsky LA, Baldovic M, Kayser M, Kutuev IA, Khusainova R, et al. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet.* 2007; 15(2):204–11. PMID: [17149388](#)
47. Shi H, Qi XB, Zhong H, Peng Y, Zhang XM, Ma RZL, et al. Genetic evidence of an East Asian origin and Paleolithic northward migration of Y-chromosome haplogroup N. *Plos ONE.* 2013; 8(6): e66102. doi: [10.1371/journal.pone.0066102](#) PMID: [23840409](#)
48. Laitinen V, Lahermo P, Sistonen P, Savontaus ML. Y-chromosomal diversity suggests that Baltic males share common Finno-Ugric-speaking forefathers. *Hum Hered.* 2002; 53(2):68–78. PMID: [12037406](#)
49. Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus ML, et al. Migration waves to the Baltic Sea region. *Ann Hum Genet.* 2008; 72:337–48. doi: [10.1111/j.1469-1809.2007.00429.x](#) PMID: [18294359](#)
50. Lahermo P, Savontaus ML, Sistonen P, Beres J, de Knijff P, Aula P, et al. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. *Eur J Hum Genet.* 1999; 7(4):447–58. PMID: [10352935](#)
51. Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, et al. Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet.* 1997; 60(5):1174–83. PMID: [9150165](#)
52. Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 2008; 18(5):830–8. doi: [10.1101/gr.7172008](#) PMID: [18385274](#)
53. Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, et al. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet.* 2004; 75(1):128–37. PMID: [15162323](#)
54. Karlsson AO, Wallerström T, Götherström A, Holmlund G. Y-chromosome diversity in Sweden—a long-time perspective. *Eur J Hum Genet.* 2006; 14:963–70. PMID: [16724001](#)
55. Lappalainen T, Hannelius U, Salmela E, von Döbeln U, Lindgren CM, Huoponen K, et al. Population structure in contemporary Sweden—a Y-chromosomal and mitochondrial DNA analysis. *Ann Hum Genet.* 2009; 73:61–73. doi: [10.1111/j.1469-1809.2008.00487.x](#) PMID: [19040656](#)
56. Blauer A, Kantanen J. Transition from hunting to animal husbandry in southern, western and eastern Finland: new dated osteological evidence. *J Archaeol Sci.* 2013; 40(4):1646–66.

57. Olalde I, Allentoft ME, Sanchez-Quinto F, Santpere G, Chiang CWK, DeGiorgio M, et al. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature*. 2014; 13; 507(7491):225–8. doi: [10.1038/nature12960](https://doi.org/10.1038/nature12960) PMID: [24463515](https://pubmed.ncbi.nlm.nih.gov/24463515/)
58. Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, et al. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science*. 2012; 336(6080):466–9. doi: [10.1126/science.1216304](https://doi.org/10.1126/science.1216304) PMID: [22539720](https://pubmed.ncbi.nlm.nih.gov/22539720/)
59. Zvelebil M, Dolukhanov P. The transition to farming in eastern and northern Europe. *J World Prehist*. 1991; 5(3):233–78.
60. Shennan S. Demographic continuities and discontinuities in Neolithic Europe: evidence, methods and implications. *J Archaeol Method Th*. 2013; 20(2):300–11.
61. Hewitt GM. Post-glacial re-colonization of European biota. *Biol J Linn Soc Lond*. 1999; 68:87–112.
62. Waters JM, Fraser CI, Hewitt GM. Founder takes all: density-dependent processes structure biodiversity. *Trends Ecol Evol*. 2013; 28(2):78–85. doi: [10.1016/j.tree.2012.08.024](https://doi.org/10.1016/j.tree.2012.08.024) PMID: [23000431](https://pubmed.ncbi.nlm.nih.gov/23000431/)
63. Zvelebil M. On the transition to farming in Europe, or what was spreading with the Neolithic—a reply. *Antiquity*. 1989; 63(239):379–83.
64. Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, et al. Mapping the origins and expansion of the Indo-European language family. *Science*. 2012; 337(6097):957–60. doi: [10.1126/science.1219669](https://doi.org/10.1126/science.1219669) PMID: [22923579](https://pubmed.ncbi.nlm.nih.gov/22923579/)
65. Gray RD, Atkinson QD. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*. 2003; 426(6965):435–9. PMID: [14647380](https://pubmed.ncbi.nlm.nih.gov/14647380/)
66. Hannelius U, Salmela E, Lappalainen T, Guillot G, Lindgren CM, von Döbeln U, et al. Population substructure in Finland and Sweden revealed by the use of spatial coordinates and a small number of unlinked autosomal SNPs. *BMC Genet*. 2008; 9:54. doi: [10.1186/1471-2156-9-54](https://doi.org/10.1186/1471-2156-9-54) PMID: [18713460](https://pubmed.ncbi.nlm.nih.gov/18713460/)
67. Vuorela T. Atlas of Finnish folk culture. Porvoo: Finnish Literature Society; 1976. pp. 151.
68. Sajantila A, Pääbo S. Language replacement in Scandinavia. *Nat Genet*. 1995; 11(4):359–60. PMID: [7493010](https://pubmed.ncbi.nlm.nih.gov/7493010/)
69. Meinilä M, Fennilä S, Majamaa K. Evidence for mtDNA admixture between the Finns and the Saami. *Hum Hered*. 2001; 52(3):160–70. PMID: [11588400](https://pubmed.ncbi.nlm.nih.gov/11588400/)
70. Achilli A, Rengo C, Battaglia V, Pala M, Olivieri A, Fornarino S, et al. Saami and Berbers—an unexpected mitochondrial DNA link. *Am J Hum Genet*. 2005; 76(5):883–6. PMID: [15791543](https://pubmed.ncbi.nlm.nih.gov/15791543/)
71. Denisova GA, Derenko MV, Malyarchuk BA. A partial central Asian/eastern Siberian origin of the Saami mtDNAs. *Am J Hum Genet*. 1999; 65(4):1101.
72. Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogvali EL, et al. The western and eastern roots of the Saami—The story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet*. 2004; 74(4):2004–682.