Tarja Sundell, Martin Heger, Juhana Kammonen & Päivi Onkamo MODELLING A NEOLITHIC POPULATION BOTTLENECK IN FINLAND: A GENETIC SIMULATION

Abstract
Population continuation versus replacement has been debated for many years in the Finnish archaeological literature. The controversy surrounding the question of dwelling voids, whether they indicate population bottlenecks or not, is ongoing. We propose to utilize information as well as methods from a different discipline, genetics, in order to assess a possible population bottleneck. Ideally this could be done by analysing ancient DNA (aDNA) samples. Unfortunately, Finland almost entirely lacks well-preserved ancient organic remains which would enable extensive aDNA studies. Here, our approach is to evaluate effects of a Neolithic population bottleneck on the assumed Finnish gene pool by simulations. In this first attempt, we focus on uniparental markers: mitochondrion and Y chromosome. We aim to assess the effects that prehistoric bottlenecks could have on the diversity of the ancient and the current gene pool. The simulations were carried out with the population genetic simulation environment simuPOP. The results indicate that, as expected, bottlenecks have a fundamental effect on genetic diversity, even today, if the population under consideration is a rather isolated one. Immigration from neighbouring populations, even if very limited but constant over prolonged time periods, can have drastic effects on a population's genetic composition. In the future, simulations will be further refined with internal subpopulations and gender specific migration, as well as autosomal markers.

Keywords: prehistoric population, population bottleneck, population simulation, simuPOP, Neolithic, Stone Age, Finland.

Tarja Sundell, Department of Philosophy, History, Culture and Art Studies, Archaeology and Department of Biosciences, PO Box 56, FI-00014 University of Helsinki, Finland, E-mail: tarja.sundell@helsinki.fi; Martin Heger, Department of Biosciences, PO Box 56, FI-00014 University of Helsinki, Finland, E-mail: martin.heger@helsinki.fi; Juhana Kammonen, Department of Computer Science and Department of Biosciences, PO Box 56, FI-00014 University of Helsinki, Finland, E-mail: juhana.kammonen@helsinki.fi; Päivi Onkamo, Department of Biosciences, P.O. Box 56, FI-00014 University of Helsinki, Finland, E-mail: paivi.onkamo@helsinki.fi

INTRODUCTION

Prehistoric population events have long intrigued both archaeologists and geneticists. The past settlement patterns, as well as possible population movements, have been extensively studied by archaeologists (e.g., Siiriäinen 1981; Meinander 1984; Nuñez 1987; Huurre 1990, 2001; Lavento 1997; 1998; 2001; Carpelan 1999a; 1999b; Edgren 1999; Halinen 1999; 2005; Nuñez & Okkonen 1999; Mökkönen 2002; Pesonen 2002; Takala 2004; Saipio 2008). The earliest postglacial habitation in Finland dates back over 10,000 years (Takala 2004; Pesonen 2005).

With the development of new genetic methods and tools (PCR in 1980s, high-throughput genotyping in the 2000s, etc.) it has become possible to study genetic differences between populations in detail, and thus open views to for example the genetic origins of Finns. An abundance of such genetic literature published in the 1990s and 2000s exist (Cavalli-Sforza et al. 1994; Sajantila et al. 1995; 1996; Lahermo et al. 1996; 1998; De la Chapelle & Wright 1998; De la Chapelle 1999; Kittles et al. 1998; 1999; Savontaus & Lahermo 1999; Norio 2000; 2004; Peltonen et al. 2000; Varilo et al. 2000; 2003; Kere 2001; Uimari et al. 2005; Lappalainen et al.

2006; 2008; Salmela et al. 2006; 2008; Service et al. 2006; Hedman et al. 2007; Hannelius et al. 2008; Jakkula et al. 2008; McEvoy et al. 2009; Palo et al. 2009). Several of these studies show that there are significant differences between the western and eastern parts of Finland. Whereas the mitochondrial and autosomal (see Box 1 on page 18) data are relatively homogenous across the country, the Y-chromosome distributions show substantial differences between East and West. Moreover, there is reduced genetic diversity in the Y chromosome in Eastern Finland.

Here, our approach is to reconstruct demographic events with a series of hypotheses relating to prehistory and thus, the genetic history of the population. One goal is to evaluate how clear traces a prehistoric bottleneck would leave in the current gene pool. We do this by combining two different elements: archaeological evidence and population genetic simulations, thus bringing together rather disparate research areas. This paper is a part of a multidisciplinary research project, Argeopop (http://www.helsinki.fi/bioscience/argeopop/), aiming to shed new light on Finnish population prehistory (Sundell 2008; 2009; Sundell & Onkamo 2010).

BACKGROUND

Finland was settled during the Mesolithic and the Neolithic by prehistoric hunter-gatherers and is assumed to have been continuously inhabited ever since. However, in the putative demographic continuum, there are periods in which archaeological material displays a change so remarkable that it poses the question of new population arriving (e.g., Meinander 1984; Carpelan 1999a) vs. intensification of trade activities (e.g., Nuñez 1987). The spread of Typical Comb Ware culture is clearly seen in archaeological finds e.g. new style in ceramics, the significant rise in frequency of novel materials such as flint and amber as well as new house building tradition (e.g., Meinander 1984; Huurre 1990; Carpelan 1999a; Edgren 1999; Halinen 1999; Pesonen 2002). Also, the number of dwelling sites increases considerably from the previous era. The climate was at its temperature maximum, which probably contributed positively to the resources available to the inhabitation. However, around 5750 BP the number of sites turns into a decline (Pesonen & Tallavaara 2008; Tallavaara et al. 2010), which continues over 1600 years, to a minimum around 4100–3800 BP, where the number of inhabitants might have been as low as few thousands. Thus, archaeologically, there is evidence for a population bottleneck. Such an event could have also had a profound effect on the gene pool.

From the genetic point of view, the reduced genetic diversity in the present day Finnish population (Lahermo et al. 1996; 1998; Lappalainen et al. 2006; 2008; Palo et al. 2009) and the specific 'Finnish Disease Heritage' could be explained by bottleneck(s) or founder effect(s) (Nevanlinna 1972; Sajantila et al. 1996; De la Chapelle 1999). Population bottlenecks and founder effects are factors that have a major impact on the population gene pool (see Box 2 on page 18 and Fig. 1). A population bottleneck does not need to be a sudden event; on the contrary, it may take dozens of generations to reach the minimum population size and several generations before the population starts growing again. The tighter and longer the bottleneck, the more genetic variation is lost. Accordingly, the amount of genetic variation lost can also be used as a rough measure of the intensity of the population event.

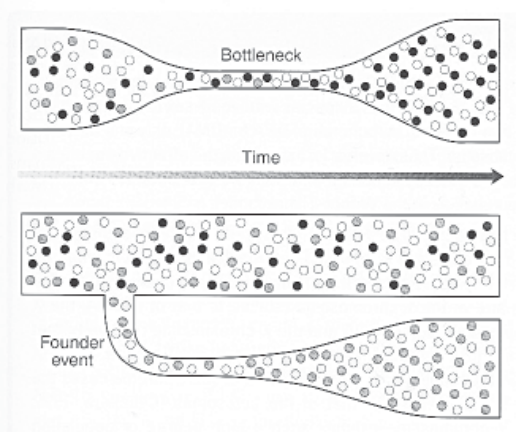

Fig. 1. Bottleneck and founder effect. Different alleles are illustrated by differently coloured circles (Jobling et al. 2004).

The loss of genetic variation should ideally be measured directly after the bottleneck event, since the more time passes from the bottleneck event, the more immigration and mutation may have compensated the loss. The best case scenario would be a sample of ancient DNA from the time period immediately after the bottleneck event.

Alternatively, when lacking access to aDNA, we have to turn to modern DNA extracted from contemporary individuals. The latter approach is based on the assumption that the present day population are direct descendants of the ancestors who went through the bottleneck. The present genetic diversity in Finland is very well known, as stated in the beginning of this article.

Due to the naturally acidic soil in Finland, bones from archaeological sites show poor biomolecular preservation. Ancient bone material, that is, the source of aDNA, is scarcely found and almost non-existing from the Stone Age. Thus, direct measurement of past population seems rather unlikely in the foreseeable future. However, there is a wealth of data on the present day population.

Therefore, our approach is to simulate possible scenarios and compare them with the observed variation in the extant population. The different scenarios are based on the conclusions we have on prehistoric demographic events.

The ancient population sizes, estimated based on radiocarbon datings, suggest that the demographic fluctuations have been more significant in the eastern part of Finland, based on a paper by Tallavaara and others (2010; also Pesonen & Tallavaara 2008). The population size reached its peak around 5750 BP in the Typical Comb Ware when it might have been as much as 25,000, only to decline in the following centuries.

The feasibility of simulation studies in retracing prehistoric population events

We use computer simulations to track genetic changes over generations. Each individual carries a mitochondrial genome and the males also a Y chromosome. These chromosomes are subject to the same evolutionary principles as those of true genomes and can be sampled and compared to genuine population data to determine if simulated scenarios could lead to observed results.

When studying population history centuries or even millennia in the past, obtaining samples in large quantities is nearly certain to prove impossible. Simulation tools are one method to circumvent such limitations, though they must be used in conjunction with other methods in order to reach conclusions with any reliability. One can run simulations of populations under different models and assumptions, calculate statistics from the results, and compare these with real world data.

Simulation methods are relatively cost-effective; essentially they only require sufficient computing power and time to set up and run the simulations and analyze the results. Thousands of years – or generations – can pass literally in the blink of an eye, and large datasets can be generated rapidly once a method has been developed.

Simulations necessarily rely on many simplifying assumptions to be feasible. On the other hand, modelling with too many variable parameters can make a simulation produce any kind of output from any kind of input rendering all conclusions meaningless. Of course, the real world data that simulation results are compared with has to come from somewhere, so simulation methods are hardly a true replacement for laboratory work, only a valuable supplement.

Changes in the number of dwelling sites between the Stone Age and the Early Metal Period

There seems to be a clear difference between the number of settlement sites in inland culture when turning to the Early Metal Period (see e.g., Lavento 1997; 2001; Saipio 2008). There are over 800 known settlement sites in the Ancient Lake Saimaa Area, of which ca. 700 have been classified as belonging to the Stone Age and only 72 either purely or partly to the Early Metal Period, respectively (Lavento 1997). According to the Register of the Ancient Monuments (http://kulttuuriymparisto.nba.fi/netsovellus/rekisteriportaali/mjreki/read/asp/r_default.aspx) provided by the National Board of Antiquities, there are (June 2010) 9522 dwelling sites on mainland Finland belonging to the Stone Age and only 190 belong to the Early Metal Period. It can be argued that Early Metal Period dwelling sites are more difficult to find whereas quartz findings constitute a Stone Age dwelling site.

Nevertheless, there is a profound difference in the number of sites. Furthermore, taphonomic processes (Surovell & Brantingham 2007; Surovell et al. 2009) would affect the older Stone Age sites harder than the younger Early Metal Period ones, thus causing a sampling bias in favour of Early Metal Period graves.

Whereas Stone Age sites are situated on eskers, foots of moraine hills on sandy or gravel soil and glacifluvial deltas, many Early Metal Period sites

are situated on fine sand or silts (albeit few are still on glacifluvial deltas). The shift of dwelling sites from pine forests into birch forests can be followed in the Saimaa area where shore displacement has profoundly affected the topography.

When comparing the Early Metal Period dwelling sites in relation to the Neolithic ones it can be observed that the Early Metal Period sites are usually considerably smaller. The change in size, type and geographical location may, according to Lavento (2001), reflect the decrease in population size and, perhaps, a transition to a mobile way of life instead of semi-sedentary tradition. The number of finds also indicates either a smaller population or a shorter period of habitation. Lavento (2001) argues that there might have been a bottleneck in eastern Finland between the Late Neolithic and the Early Metal Period. Large areas of Finland were possibly almost uninhabited and the number of local inhabitants decreased very drastically from those periods when large villages were settled in Saimaa (Lavento 2001) or north Ostrobotnia (Nuñez & Okkonen 1999). However, the inhabitation has been continuous in the coastal areas.

The size of a given site and the small number of structures found in excavations may easily lead to the interpretation that the site was temporary. It is difficult to say if sites were used only seasonally, that is, particularly during a certain period of the year. Furthermore, hunter-gatherers are presumed to have left signs on numerous occupation sites due to their seasonal settlement model. Despite the shortcomings and inadequacies due to the classification criteria, there are considerable differences between the numbers of Stone Age and Early Metal Period sites (Lavento 1997; 2001; Saipio 2008).

MATERIALS AND METHODS

Population genetic simulations can be divided into two categories. Coalescent simulations simulate backward in time into the coalescent (most recent common ancestor) from present genetic structure, and forward simulations go forward in time from a defined starting point. In this study we have employed the forward simulation method.

Forward simulations can be used to create entire virtual populations with various parameters. These populations are then simulated through their entire histories, creating a simplified model of whatever real world situation is being studied. The simulated populations can then be sampled and analyzed and the results compared with real data to support conclusions from other types of research, such as archaeology. It is to be noted that generally simulations are not perfectly suited to proving the validity of one specific population history or parameter set, but are more useful as evidence that a certain model is unlikely to be the true one.

Simulator / Simulation tool

SimuPOP (Peng & Kimmel 2005), a forward time population genetics simulation environment, was used in this study. SimuPOP consists of a number of Python objects, such as populations, chromosomes and individuals, and functions and operators, such as mating schemes, migration operators and mutation models. Users assemble a suitable simulator from these various components by writing a Python script file. Python is a popular programming language in many fields, including bioinformatics. This modularity makes simuPOP one of the most flexible genetic simulation tools

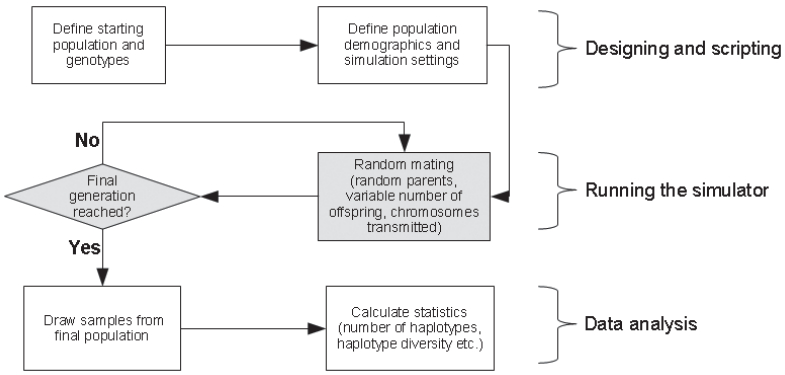

Fig. 2. A Flowchart describing the course of a single simulation run with simuPOP. To get equal sample sizes of mitochondria and Y-chromosomes collected, female individuals are removed from the final population before drawing samples from the population (lower left process box). Thus the samples will consist only of male individuals.

currently available, but at the cost of necessitating some skill in Python programming. The simulations themselves were run on the Murska server cluster of the IT Center for Science Ltd (CSC).

Population simulations

Every individual carries a mitochondrial genome which is transmitted from mother to offspring. Mitochondrial DNA (mtDNA) is simulated as a 631 bp DNA sequence. Any nucleotide can mutate into the other three creating new variation. The length and the mutation model used mimic the HVS-I and HVS-II of the mtDNA control region. Here, our aim is not to imitate the actual haplotype composition but instead observe whether simulations arrive at similar diversity measures.

Prehistoric Finnish population was simulated with two bottlenecks, the first one at 4100–3800 and the second at 1500–1300 BP. Simulation moves forward in ten-year-steps with reproductive ages being different for males and females.

Background populations

A few migration waves have been alleged to have directed to Finland from neighbouring populations (see e.g., Carpelan 1999; Lavento 2001). In addition to these, there has probably been minor, more or less constant gene flow into the country. To follow the assumed demographic events as realistically as possible we decided to form three background populations from which individuals migrate into the Finnish population: Archaic European, Archaic Scandinavian and Saami. We have used information on actual ancient mtDNA haplotypes to form the general background of these populations (Appendix 1). The haplotypes have been obtained from published aDNA studies of prehistoric European populations (Haak et al. 2005; Rudbeck et al. 2005; Melchior et al. 2007; 2008; Bramanti et al. 2009; Malmström et al. 2009). According to recent studies, there are considerable differences in haplotype distributions of hunter-gatherers and early farmers (e.g., Haak et al. 2005; Bramanti et al. 2009; Malmström et al. 2009) which was carefully taken into consideration for the haplotype compositions.

The Archaic European background population was considered to include 50,000 individuals, the Archaic Scandinavian 25,000 individuals and the Saami 5000 individuals. With extensive migration waves the first (Typical Comb Ware) replaces ca. 20 % of the Finnish population with Archaic European and the second (Corded Ware) replaces ca. 9 % of the Finnish population with Archaic European. With smaller migration waves the percentages are divided by four, that is, approximately 5 % and 2 %, respectively.

Extensive constant gene flow replaces 0.1 % of the population every 10 years with Archaic European during the entire simulation. Additionally, beginning at 4000 BP, 0.05 % of the population is replaced with Saami every 10 years and after 3500 BP 0.5 % of the population is replaced with Archaic Scandinavian every 10 years. Lower constant gene flow is reduced to one tenth, that is, 0.01 %, 0.005 % and 0.05 %, respectively.

Each background population evolves separately for 12,000 years. A snapshot is saved every 2000 years. The snapshot haplotype frequencies are then used as sources of migration in the bottleneck simulations over a period of 2000 years each, so that an approximation of evolving, dynamic background populations is reached without having to simulate them in all 15,000 runs.

The Saami background population forms an exception here. Ancient Saami DNA has not yet been found nor studied, but nevertheless we wanted to include the effect of a small population occasionally changing genes with the neighbouring Finnish population. Thus the Saami background population is not based on actual aDNA samples but instead on the extant Saami haplotype composition. Therefore the Saami influence begins somewhat later, even though the Saami areas of

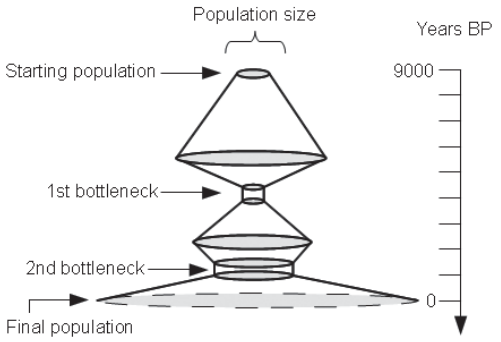

Fig. 3. The general demographic model. The width of the cone represents the proportional population size at the time before present (BP) depicted on the y axis.

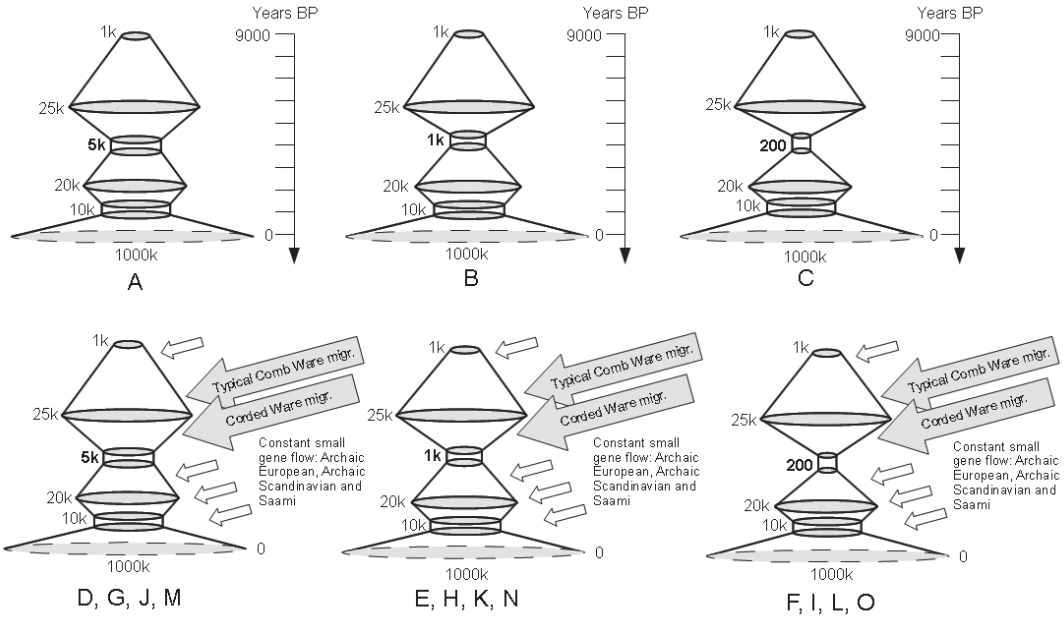

Fig. 4. Bottleneck population models (A–O) used in this study.

Model (A): a mild bottleneck without migration waves or constant gene flow: bottleneck population size 5000 individuals
Model (B): a moderate bottleneck without migration waves or constant gene flow: bottleneck population size 1000 individuals
Model (C): a severe bottleneck without migration waves or constant gene flow: bottleneck population size 200 individuals
Models (D, E, F): small migration waves and small constant gene flow: migration from the Archaic European, Archaic Scandinavian and Saami background populations. Bottleneck population sizes 5000 (D), 1000 (E) and 200 (F) individuals
Models (G, H, I): extensive migration waves and extensive constant gene flow: migration from the Archaic European, Archaic Scandinavian and Saami background populations. Bottleneck population sizes 5000 (G), 1000 (H) and 200 (I) individuals
Models (J, K, L): extensive migration waves and small constant gene flow: migration from the Archaic European, Archaic Scandinavian and Saami background populations. Bottleneck population sizes 5000 (J), 1000 (K) and 200 (L) individuals
Models (M, N, O): small migration waves and extensive constant gene flow: migration from the Archaic European, Archaic Scandinavian and Saami background populations. Bottleneck population sizes 5000 (M), 1000 (N) and 200 (O) individuals

Genetic data is collected from five predetermined checkpoints:
1. 9000 BP
2. 4650 BP (before the first bottleneck)
3. 4090 BP (immediately after the population has reached its minimum)
4. 1400 BP (in the middle of the second bottleneck)
5. 0 BP (present time, final generation)

In a checkpoint, a random sample of 1500 individuals is extracted from the population extant at the time. If total population size at this point is smaller than 1500, all individuals in the population are included in the sample.

northern Finland have been inhabited for over 10,000 years and, most probably, themselves received gene flow from several sources.

The rationale for adding background populations into the simulation model is to see how the very presence of background gene flow affects the simulated Finnish population. Most important at this stage was to model, compared with the Finnish population, a very small Saami population, an intermediate archaic Scandinavian population and a larger archaic European population. The specific sizes here are only approximations, as we only need to see if such gene flow produces changes in the results.

Bottleneck simulations

Our model simulation begins at 9000 BP with a population of 500 females and 500 males. The simulated population is age-structured so that generations overlap. The simulation proceeds in steps of 10 years. In every step individuals with age higher than maximum allowed age (60 years), are removed from the population (die). Additionally, a population mortality rate of 15 percent / step was introduced. Individuals in mating ages (males from 20 to 60 years in age, females from 20 to 40 years in age) are allowed to mate randomly. The offspring of each couple (0 to N) and the already existing individuals together fill the population at the next simulation step.

The initial population size is set to grow exponentially up until around 5750 BP, after which the population begins to gradually decline over 1600 years towards the bottleneck (Tallavaara et al. 2010). The population maximum is ca. 25,000 individuals. In our model there are two bottlenecks, first at 4100–3800 BP and a second, less severe one at 1500–1300 BP (Tallavaara et al. 2010). The main hypothesis in this simulation study is that the first bottleneck is genetically more significant. We have designed different scenarios concerning the severity of this bottleneck as described below. The final population size is set to 1,000,000 individuals instead of the present census size of circa 5,000,000. The rationale behind this is that in an expanding population, the genetic diversity does not change relevantly in the course of a few generations (see e.g., Jobling et al. 2004).

The simulations were carried out with 15 separate scenarios (A–O). The first three are so-called reference simulations without any migration, whereas the others include migration. In model (A), the population size in the first bottleneck is 5000 individuals, after which the population starts to recover. In model (B) the population size in the first bottleneck is 1000 individuals and in model (C) the population size in the first bottleneck reduces the total population to a mere 200 individuals. This first bottleneck endures for 300 years in all settings. The second, less severe, bottleneck is assumed to have consisted of 10,000 individuals for 200 years in all three models (A–C). In models (D–O), the pattern of bottleneck severities continues in groups of three as in simulations (A–C) but is accompanied by varying magnitudes of migration waves and gene flow (see Fig. 4).

Population genetic analyses

The genetic changes were measured with two basic indicators of diversity: the number of haplotypes present and haplotype diversity ($\hat{H}$) in a sample. The first is simply a direct count of the number of different haplotypes (i.e. differing in at least one nucleotide position or microsatellite locus). $\hat{H}$ (Nei 1987) is based on haplotype frequencies in a population and measures the probability of observing different haplotypes when sampling two random chromosomes or, as in this case, haploid individuals, from a population. Haplotype diversity can be calculated with the formula $\hat{H} = n(1-\sum x_i^2)/(n-1)$ where n is the number of individuals and x_i the haplotype frequency of the *i*th haplotype. When $\hat{H}$ is low it is likely that two randomly drawn chromosomes are identical, and vice versa. Many more population genetic measures exist, and as we refine our simulation methods in future studies we will also include additional measures.

RESULTS

Bottleneck severity

The first bottleneck substantially reduces genetic diversity, at least in the two models with the narrowest bottlenecks (see checkpoint 3 in Figs. 5–8 and models C and L in Figs. 5–10). Especially the most severe bottleneck (simulation model C, 200 individuals) drastically reduces genetic diversity. This is, obviously, in compliance with population genetic principles. The least severe

bottleneck (simulation model A, 5000 individuals) allows the population to sustain existing diversity. The second bottleneck (10,000 individuals) has a much smaller effect on genetic diversity (see checkpoint 4 in Figs. 5–8).

Migration waves vs. constant gene flow

Most importantly, the effects of constant gene flow from background populations clearly outweigh the effects of larger, short-term migration waves. Extensive constant gene flow (simulation models G–I and M–O) also completely eclipses the effects of bottleneck severity in the mitochondrial data, in other words the simulated populations appear nearly identical in the final generation independent of the first bottleneck size (200, 1000 or 5000 individuals, see Figs. 11–12). Migration waves at any setting have barely any effect if constant gene flow is extensive.

On the other hand, if constant gene flow is small or absent, the different bottlenecks result in clearly different diversity distributions even in the present day population (see Fig. 9 and 10). Mitochondrial diversities in models with extensive constant gene flow (simulation models G–I and M–O) look distinctly different from the others, and almost identical to one another (see models H and N in Fig. 13). Y chromosomal diversities, however, are nearly identical no matter what migrations are applied (see Fig. 14).

Comparison to true present day genetic diversity

None of these somewhat simplified simulation models produces an exact fit with diversity measures calculated from present-day Finnish sample data. Simulated mitochondrial diversity remains below observed diversity with any bottleneck size

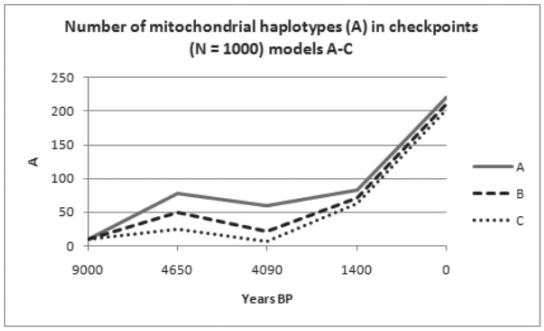

Fig. 5. Overall development of the number of mitochondrial haplotypes throughout the simulations, models A–C.

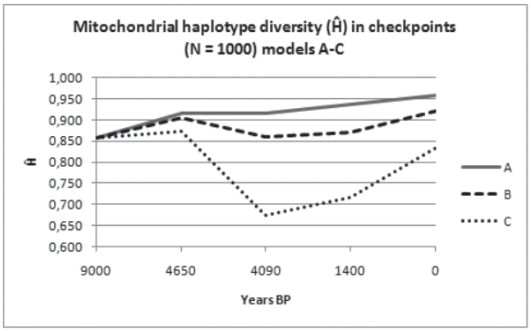

Fig. 6. Overall development of the average mitochondrial genetic diversity throughout the simulations, models A–C.

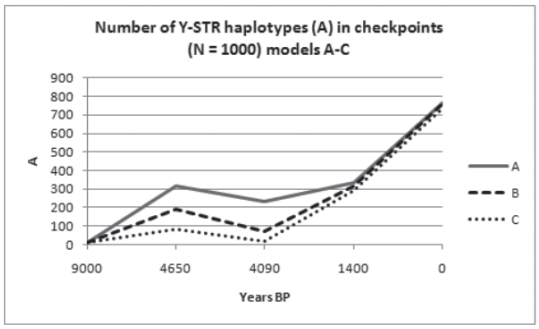

Fig. 7. Overall development of the number of Y-STR haplotypes throughout the simulations, models A–C. The Y-STR and mitochondrial data differ from each other due to the different mutation models used.

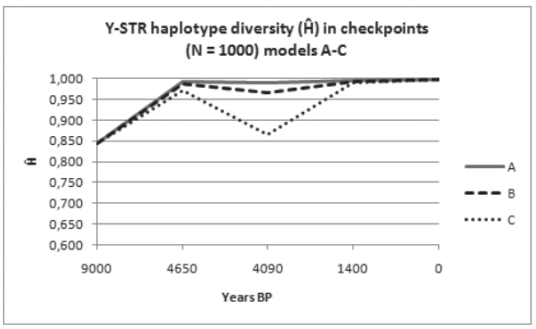

Fig. 8. Overall development of the average Y-STR genetic diversity throughout the simulations, models A–C. The Y-STR figures differ from the mitochondrial figures due to the different mutation models used.

(see Figs. 9–13), and Y chromosome diversity, somewhat confusingly, ends up above observed diversity (see Fig. 14). In our opinion, this is most probably due to the initial settings and should not be considered in too much detail at this stage of development of the methodology.

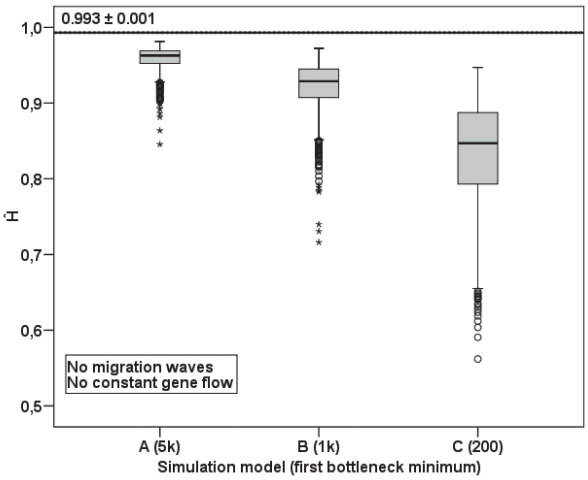

Fig. 9. Mitochondrial haplotype diversity ($\hat{H}$) in 1000 replicate samples of 832 individuals in the final generation. The first bottleneck minimum is 5000 (A), 1000 (B) and 200 (C) individuals, respectively. There are neither migration waves nor continuous gene flow to the population. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

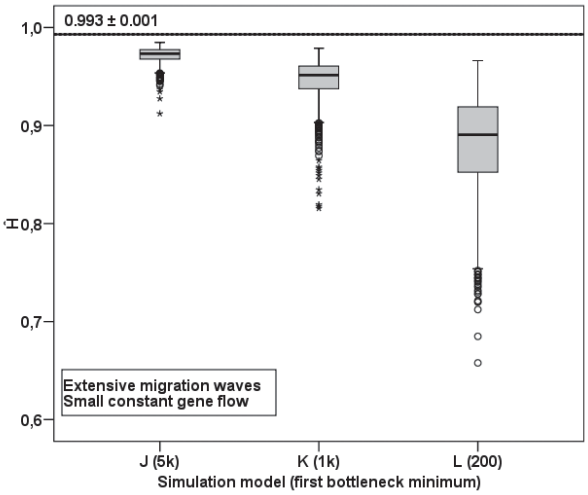

Fig. 10. Mitochondrial haplotype diversity ($\hat{H}$) in 1000 replicate samples of 832 individuals in the final generation. The first bottleneck minimum is 5000 (J), 1000 (K) and 200 (L) individuals, respectively. Migration waves are extensive and constant gene flow is small. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

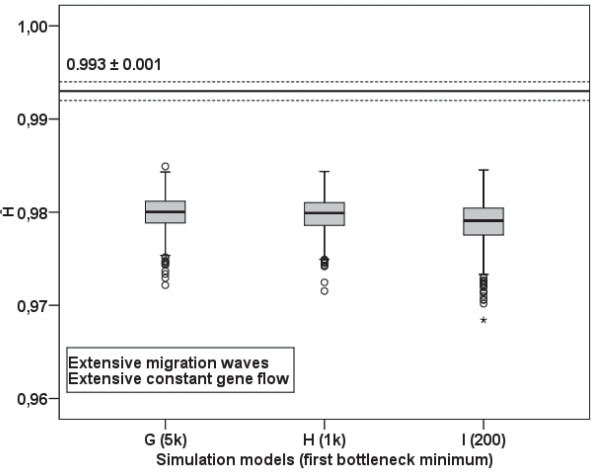

Fig. 11. Mitochondrial haplotype diversity ($\hat{H}$) in 1000 replicate samples of 832 individuals in the final generation. The first bottleneck minimum is 5000 (G), 1000 (H) and 200 (I) individuals, respectively. Both migration waves and constant gene flow are extensive. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

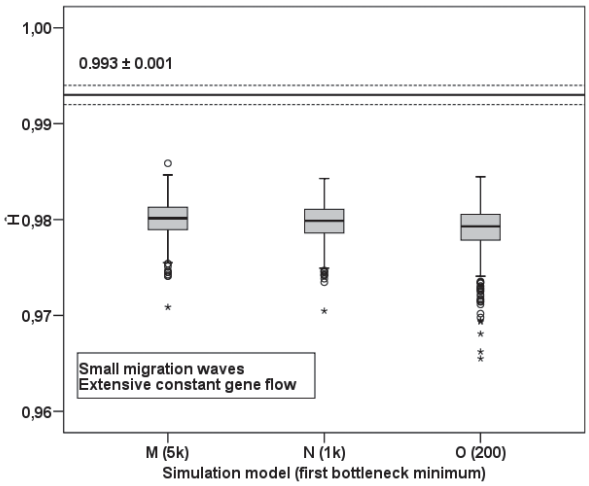

Fig. 12. Mitochondrial haplotype diversity ($\hat{H}$) in 1000 replicate samples of 832 individuals in the final generation. The first bottleneck minimum is 5000 (M), 1000 (N) and 200 (O) individuals, respectively. Migration waves are small and constant gene flow is extensive. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

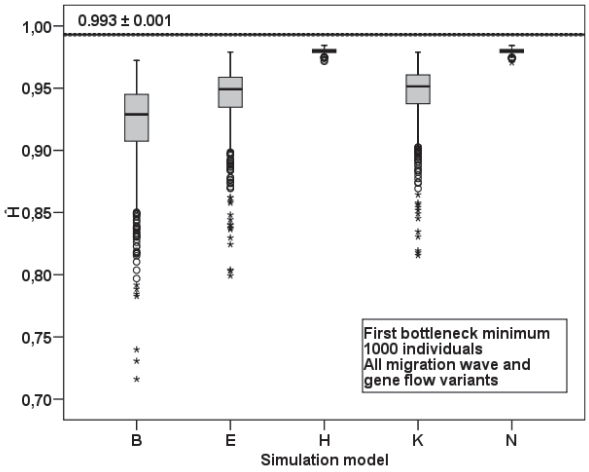

Fig. 13. Mitochondrial haplotype diversity ($\hat{H}$) in 1000 replicate samples of 832 individuals in the final generation with the first bottleneck minimum set at 1000 individuals and all migration wave and constant gene flow variants. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

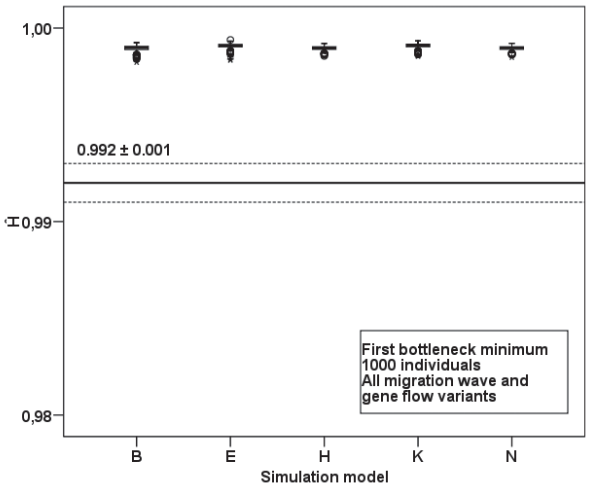

Fig. 14. Y-chromosome microsatellite haplotype diversity ($\hat{H}$) in 1000 replicate samples of 907 individuals in the final generation with the first bottleneck minimum set at 1000 individuals and all migration wave and gene flow variants. The horizontal reference line shows $\hat{H}$ in the present-day Finnish people with its 95 % confidence interval. Note the difference in scale between Figs. 9–14.

DISCUSSION

Our simulations show that the severity of a prehistoric bottleneck may still affect current genetic diversity. Somewhat surprisingly, the constant small gene flow seems to be a much more important factor than migration waves. While the migration waves have barely any effect, moderate constant gene flow can cause great differences to genetic diversity. It probably forces the Finnish population's diversity to change to the background population's. Finland has always been viewed as a genetic isolate, and arguably even lower gene flow rates than used here should be explored. Especially the extensive constant gene flow seems far too high for Finnish prehistory.

In conclusion, we show that current genetic diversity cannot be reproduced with these rather simplified models. The Y chromosome especially highlights the need to refine our simulations. Differences in the Y chromosome gene pool between eastern and western Finland are evident in the current population (Lappalainen 2006; Palo 2009). It is conceivable that modelling Finland as an eastern and western subpopulation with limited internal migration should considerably improve the validity of simulation results.

Limitations of the methodology

This study considers only uniparental markers, the Y chromosome and mitochondrion. Autosomal loci would also be highly interesting: adding them will bring an extra dimension to the simulations, bridging the gap between the effects of male and female. Moreover, we have not yet incorporated internal or gender specific migration.

It can be argued that simulation models entail a fundamental weakness: the fact that the observed data can be produced by a particular model does not guarantee that the model represents the reality. Also, different models created by different simulation settings may replicate the observed patterns of diversity equally well. It can however be argued that the point of simulations is not to prove one model to describe the absolute truth but rather to exclude the least compatible ones.

Considering the archaeological readership, we wish to emphasize the complexity of simulations. For every parameter, a numerical value has to be set exactly in order to be able to run simulations at all. The often circumspect information (containing conditionals and diminutive expressions) from (archaeological) studies cannot be incorporated as such into a simulation but must first be converted into numerical values. Perhaps and circa are not expressions that simuPOP understands. There is always uncertainty concerning dates and population sizes etc. If we knew enough to set every parameter with certainty, simulations would be quite unnecessary. As it is, we attempt to base all the parameters required by the simulations on facts, studies, and lacking these, on best academic guesses.

Developing the simulation tool

SimuPOP has an open source code and the software is under constant development. New versions are released in intervals of approximately two months. This has given us a great opportunity to contribute to the development of simuPOP. Although simuPOP has built-in tools for calculating statistics of interest, we decided to include our own methods in the actual simulation script to do this (Heger, in prep.). This way, we got data output in a format that suited our needs better. Simulating millions of mtDNA sequences, even though each is only 631 base pairs long, requires a huge amount of computer memory. To increase the length of the simulated DNA sequences and further improve the output of our simulations, a forthcoming study is investigating the memory loading properties of simuPOP (Kammonen, in prep.).

What happens when populations meet?

Exchange of genes does not have to occur, although it presumably often happens. Interactions between given populations might have been solely cultural. An archaic population, supposed to have left behind a certain style of ceramics, might not have left any descendants, that is, these people do not have any contribution to present-day DNA. Genetic admixture could not be observed in archaeological evidence. However, it may be observed in genetic data; this might be one possible direction of future research.

Causes for population bottlenecks

Population bottlenecks can occur rapidly or be slower events taking centuries or millennia.

Slowly developing bottlenecks may be caused by for example gradually deteriorating climate conditions. Other causes for a sudden population bottleneck have to be considered, such as famine, epidemic and war (Sundell 2008).

Epidemics/infections

Historically epidemics have been by far the most crucial cause to high mortality, rather than violence and hunger (Kallioinen 2005). However, straightforward proof of existence of infectious diseases before the historical times or development of agriculture and animal husbandry are scarce and the reconstruction of the early history of infectious diseases is very difficult, almost impossible. Pathogenic organisms have probably not been a significant cause in restraining the number of early hunters outside tropical regions (McNeill 1979). On the contrary, prehistoric hunter-gatherers living in cool climates have probably been very healthy despite their relatively short lifespan due to fatal accidents.

In farming cultures the alimentation became more one-sided and dependent on just few types of grain. Domesticated animals replaced wild game, but also introduced various pathogens. Several contagious diseases found in people today have their origins most likely in animals. When population density grows, the likelihood for an infection to transfer from one host to another multiplies. After a population exceeds the critical threshold, the infection can suddenly burst into a hyper infection which can be very damaging to a community since they harm foraging as well as child breeding significantly. In big populations, containing more than 500,000 individuals, epidemics usually occur in five to ten-year-cycles in order for an adequate population without immunity to develop (Kallioinen 2005). In small populations epidemics occur more irregularly instead. It has not been before the Early Middle Ages that Finland has had the sufficient population density for epidemics to spread other than locally only. Deficiency diseases may also have had a significant effect.

Conflicts

Could pre-historical wars have had a major impact on demography? The old research tradition is grounded on the assumption that Comb Ware cultures were peace seeking hunter-gatherers whereas Corded Ware represents a warlike conquest (see e.g., Huurre 2001; Salo 2003). Furthermore, there might have been a conflict situation between the contemporaneous cultures Pöljä Ware and Corded Ware (Halinen 1999). A book by L.H. Keeley, *War before Civilization* (1996), has greatly influenced the formation of a new research area, conflict archaeology. Whether there have or have not been any wars in Stone Age Finland is very hard to prove. This is especially difficult with prehistoric contexts. The majority of the most common arms and weapons are also the ones that are worst preserved or the most difficult to observe and categorize such as bows, wooden clubs, slings, and so forth (Lahelma & Sipilä 2004).

Even though we could find bone material that show clear evidence of violence, it would be hard to tell which wounds are consequences from conflicts between people and which are caused purely by accidents. Furthermore, due to our naturally acidic soil in Finland, we lack almost completely all prehistoric bone material which would allow us to do any research regarding warfare in Stone Age Finland.

Despite the fact that violence between humans is not necessarily easy to prove archaeologically, Raninen (2006) argues that it is evident that chronic lethal violence has been a central means of organization and experience to some prehistoric communities. This does not mean that all prehistoric communities would have lived under conditions labelled with pandemic violence. Instead it is probable that there have been reasonably peaceful and actually pacifistic periods in Finland's prehistory (Raninen 2006). Therefore prehistoric wars are unlikely to have had a major influence on demography.

Changes in climatic conditions

Changes in climate could have caused successive shortages in resources and therefore famine. People weakened by famine, on the other hand, have been more vulnerable to diseases. In the Boreal climate zone, where livelihood preconditions are limited, already small changes in climatic conditions have been significant in determining the success of different subsistence models. Total

abandonment of dwelling sites have probably been caused by drastic changes in the environment. Access to resources has been challenged or compromised thus significantly affecting the subsistence strategies and maintenance of livelihood. However, the effect of prehistoric climatic conditions on the demographics of past populations is beyond the scope of this paper, and deserves further research in paleoecology and perhaps consideration in future simulation studies.

Future aspects

Internal and gender specific migration will be one focus of our future simulations, as well as incorporation of autosomal markers in order to see effects independent of gender. The latter requires large numbers of autosomal loci to be simulated. Moreover, population admixture might be studied in more detail. Genetic consequences of population admixture, given that the source populations are somewhat different, can be observed in allele frequencies, for example, mtDNA and Y chromosome distributions and marker-to-marker linkage disequilibrium (LD) patterns. The amount of LD is not constant but depends on the size of the population, the distance between genes, mutation frequency and natural selection. By studying the amount and distribution of LD in populations it is possible to make estimates of for example the size of the founding population.

ACKNOWLEDGEMENTS

The authors would like to thank the sample donors, as well as Prof. Antti Sajantila and Dr. Jukka Palo who kindly gave us permission to use their Finnish mitochondrial and Y STR data. We also thank Prof. Mika Lavento, Dr. Petri Halinen and BSc. Teija Ojala for their helpful comments. This work was supported by grants from the Academy of Finland, the University of Helsinki Funds and Jenny and Antti Wihuri Foundation.

REFERENCES

Unpublished sources

Heger, M. *Population Genetic Simulations of the Savonian Expansion in Finland.* MSc thesis in preparation). Department of Biosciences, University of Helsinki.

Kammonen, J. in preparation. *Using simuPOP to Model Finnish Demography Relative to Large Neighbouring Populations*. MSc thesis in preparation. Department of Computer Science, University of Helsinki.

Pesonen, P. & Tallavaara, M. 2008. *Suomen esihistoriallinen väestömäärä 8500 calBC – 1000 calAD*. Argeopop-project report.

Sundell, T. 2008. *Populaatiopullonkaulat kivikauden Suomessa: arkeologinen ja geneettinen tulkinta*. MA thesis. Institute for Cultural Research, Department of Archaeology, University of Helsinki.

Internet sources

Argeopop. http://www.helsinki.fi/bioscience/argeopop/

Register of the Ancient Monuments. http://kulttuuriymparisto.nba.fi/netsovellus/rekisteriportaali/mjreki/read/asp/r_default.aspx

Literature

Bramanti, B., Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., M.N., Kind, C.-J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P. & Burger, J. 2009. Genetic discontinuity between local hunter-gatherers and Central Europe's first farmers. *Science* 326: 137–40.

Burger, J., Kirchner, M., Bramanti, B., Haak, W. & Thomas, M.G. 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proceedings of the National Academy of Sciences USA* 104(10): 3736–41.

Carpelan, C. 1999a. Käännekohtia Suomen esihistoriassa aikavälillä 5100...1000 eKr. In P. Fogelberg (ed.), *Pohjan poluilla: suomalaisten juuret nykytutkimuksen mukaan*: 249–80. Bidrag till kännedom av Finlands natur och folk 153. Societas Scientarum Fennica, Helsinki.

Carpelan, C. 1999b. On the postglacial colonisation of eastern Fennoscandia. In M. Huurre (ed.), *Dig It Al: Papers Dedicated to Ari Siiriäinen*: 151–71. The Finnish Antiquarian Society & The Archaeological Society of Finland, Helsinki.

Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A. 1994. *The History and Geography of Human Genes*. Princeton University Press, Princeton.

De la Chapelle, A. & Wright, F.A. 1998. Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. *Proceedings of the National Academy of Sciences USA* 95(21): 12416–23.

De la Chapelle, A. 1999. Geenimutaatiot Suomen väestön juurten tutkimuksissa. In P. Fogelberg (ed.), *Pohjan poluilla: suomalaisten juuret nykytutkimuksen mukaan*: 307–10. Bidrag till kännedom av Finlands natur och folk 153. Societas Scientarum Fennica, Helsinki.

Edgren, T. 1999. Käännekohtia Suomen kivikaudessa. In P. Fogelberg (ed.), *Pohjan poluilla: suomalaisten juuret nykytutkimuksen mukaan*: 281–93. Bidrag till kännedom av Finlands natur och folk 153. Societas Scientarum Fennica, Helsinki.

Haak, W., Forster, P., Bramanti, B., Matsumura, S., Brandt, G., Tänzer, M., Villems, R., Renfrew, C.,

Gronenborn, D., Alt, K.W. & Burger, J. 2005. Ancient DNA from the first European farmers in 7500-Year-Old Neolithic sites. *Science* 310: 1016–8.

Halinen, P. 1999a. Burial practices and the structure of societies during the Stone Age in Finland. In M. Huurre (ed.), *Dig It All: Papers Dedicated to Ari Siiriäinen*: 173–9. The Finnish Antiquarian Society & The Archaeological Society of Finland, Helsinki.

Halinen, P. 2005. *Prehistoric Hunters of Northernmost Lapland: Settlement Patterns and Subsistence Strategies.* Iskos 14. Finnish Antiquarian Society, Helsinki.

Hannelius, U., Salmela, E., Lappalainen, T., Guillot, G., Lindgren, C.M., von Döbeln, U., Lahermo, P. & Kere, J. 2008. Population substructure in Finland and Sweden revealed by the use of spatial coordinates and a small number of unlinked autosomal SNPs. *BioMedCentral Genetics* 9(54): 1–12.

Hedman, M., Brandstätter, A., Pimenoff, V., Sistonen, P., Palo, J.U., Parson, W. & Sajantila, A. 2007. Finnish mitochondrial DNA HVS-I and HVS-II population data. *Forensic Science International* 172: 171–8.

Huurre, M. 1990. *9000 vuotta Suomen esihistoriaa*. Otava, Helsinki.

Huurre, M. 2001. *Kivikauden Suomi.* Otava, Helsinki.

Jakkula, E., Rehnström, K., Varilo, T., Pietiläinen, O.P.H., Paunio, T., Pedersen, N.L., deFaire, U., Järvelin, M-R., Saharinen, J., Freimer, N., Ripatti, S., Purcell, S., Collins, A., Daly, M.J., Palotie, A. & Peltonen, L. 2008. The genome-wide patterns of variation expose significant substructure in a founder population. *The American Journal of Human Genetics* 83: 787–94.

Jobling, M.A., Hurles, M.E. & Tyler-Smith, C. 2004. *Human Evolutionary Genetics: Origins, Peoples and Disease.* Garland Science, New York.

Kallioinen, M. 2005. *Rutto ja rukous: tartuntataudit esiteollisen ajan Suomessa*. Atena, Jyväskylä.

Keeley, L.H. 1996. *War before Civilization*. Oxford University Press, New York.

Kere, J. 2001. Human population genetics: lessons from Finland. *Annual Review of Genomics and Human Genetics* 2: 103–28.

Kittles, R.A., Perola, M. & Peltonen, L. 1998. Dual origins of Finns revealed by Y chromosome haplotype variation. *The American Journal of Human Genetics* 62: 1171–9.

Kittles, R.A., Bergen, A.W., Urbanek, M., Virkkunen, M., Linnoila, M., Goldman, D. & Long, J.C. 1999. Autosomal, mitochondrial and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. *American Journal of Physical Anthropology* 108: 381–99.

Lahelma. A & Sipilä, J. 2004. Pasifistiset pyyntikulttuurit? Sodan paradigmateoria ja kivikauden Suomi. *Muinaistutkija* 4/2004: 2–21.

Lahermo, P., Sajantila, A., Sistonen, P., Lukka, M., Aula, P., Peltonen, L., Savontaus, M.-L. 1996. The genetic relationship between the Finns and the Finnish Sami (Lapps): analysis of nuclear DNA and mtDNA. *The American Journal of Human Genetics* 58: 1309–22.

Lahermo, P., Savontaus, M-L., Sistonen, P., de Knijff, P., Aula, P., Sajantila, A. 1998. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Sami. In P. Lahermo, *Mitochondrial and Y Chromosomal Variations in the Finno-Ugric- Speaking Peoples*. Turun yliopiston julkaisuja, sarja D, osa 315. Turun yliopisto, Turku.

Lappalainen, T., Koivumäki, S., Salmela, E., Huoponen, K., Sistonen, P., Savontaus, M.-L. & Lahermo, P. 2006. Regional differences among Finns: a Y-chromosomal perspective. *Gene* 376: 207–15.

Lappalainen, T., Laitinen, V., Salmela, E., Andersen, P., Huoponen, K., Savontaus, M.-L. & Lahermo, P. 2008. Migration waves to the Baltic Sea region. *Annals of Human Genetics* 72(3): 337–348.

Lavento, M. 1997. Geoarchaeological observations on the Early Metal Period dwelling sites in the ancient Lake Saimaa area. *Slavjane i Finno-ugri: Arheologija, istorija, kultura*: 38–46. St. Petersburg.

Lavento M. 1998. Sisämaan vanhemman metallikauden väestö tutkimusongelmana. *Muinaistutkija* 4/1998: 46–55.

Lavento, M. 2001. *Textile Ceramics in Finland and on the Karelian Isthmus: Nine Variations and Fugue on a Theme of C.F. Meinander.* Suomen Muinaismuistoyhdistyksen Aikakauskirja 109. Finnish Antiquarian Society, Helsinki.

Malmström, H., Gilbert, M.T.P., Thomas, M.G., Brandström, M., Storå, J., Molnar, P., Andersen, P.K., Bendixen, C., Holmlund, G., Götherström. A. & Willerslev, E. 2009. Ancient DNA reveals lack of continuity between Neolithic hunter-gatherers and contemporary Scandinavians. *Current Biology* 19: 1758–62.

McEvoy, B.P., Montgomery, G.W., McRae, A.F., Ripatti, S., Perola, M., Spector, T.D., Cherkas, L., Ahmadi, K.R., Boomsma, D., Willemsen, G., Hottenga, J.J., Pedersen, N.L., Magnusson, P.K.E., Kyvik, K.O., Christensen, K., Kaprio, J., Heikkilä, K., Palotie, A., Widen, E., Muilu, J., Syvänen, A.-C., Liljedahl, U., Hardiman, O., Cronin, S., Peltonen, L., Martin, N.G. and Visscher, P.M. 2009. Geographical structure and differential natural selection among North European populations. *Genome Research* 19: 804–14.

McNeill, W.H. 1979. *Plagues and Peoples*. Penguin, London.

Meinander, C.F. 1984. Kivikautemme väestöhistoria. In Åström, S-E. (ed). *Bidrag till kännedom av Finlands natur och folk* H. 131. Suomen väestön esihistorialliset juuret. Societas Scientiarum Fennica. Helsinki. 21-48.

Melchior, L., Gilbert, M.T.P., Kivisild, T., Lynnerup, N. & Dissing, J. 2007. Rare mtDNA haplogroups and genetic differences in rich and poor Danish Iron-Age villages. *American Journal of Physical Anthropology* 135: 206–15.

Melchior, L., Kivisild, T., Lynnerup, N. & Dissing, J. 2008. Evidence of authentic DNA from Danish Viking Age skeletons untouched by humans for 1,000 Years. *The Public Library of Science, PLOSone* 3(5).

Mökkönen, T. 2002. Chronological variation in the locations of hunter-gatherer occupation sites vis-à-vis the environment. In H. Ranta (ed.), *Huts and Houses: Stone Age and Early Metal Age Buildings in Finland*: 53–64. National Board of Antiquities, Helsinki.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nevanlinna, H.R. 1972. The Finnish population structure: a genetic and genealogical study. *Hereditas* 71: 195–236.

Norio, R. 2000. *Suomi-neidon geenit*. Otava, Helsinki.
Norio, R. 2004. Suomalaisten mutkalliset geenijuuret. In R. Grünthal (ed.), *Ennen, muinoin: miten menneisyyttämme tutkitaan?*: 174–87. Suomalaisen kirjallisuuden seura, Helsinki.
Nuñez, M. 1987. A model for the early settlement of Finland. *Fennoscandia archaeologica* 4: 3–18.
Nuñez, M. & Okkonen, J. 1999. Environmental background for the rise and fall of villages and megastructures in North Ostrobotnia 4000–2000 calBC. In M. Huurre (ed.), *Dig It All: Papers Dedicated to Ari Siiriäinen*: 105–15. The Finnish Antiquarian Society & The Archaeological Society of Finland, Helsinki.
Palo, J.U., Ulmanen, I., Lukka, M., Ellonen, P. & Sajantila, A. 2009. Genetic markers and population history: Finland revisited. *European Journal of Human Genetics* 17: 1336–46.
Peng, B. & Kimmel, M. 2005. SimuPOP: a forward-time population genetics simulation environment. *Bioinformatics* 21: 3686–7.
Peltonen, L., Palotie, A. & Lange, K. 2000. Use of population isolates for mapping complex traits. *Nature Reviews Genetics* 1: 182–90.
Pesonen, P. 2002. Semisubterranean houses in Finland: a review. In H. Ranta (ed.), *Huts and Houses: Stone Age and Early Metal Age Buildings in Finland*: 53–64. National Board of Antiquities, Helsinki.
Pesonen, P. 2005. Sarvingin salaisuus: Enon Rahakankaan varhaismesoliittinen ajoitus. *Muinaistutkija* 2/2005: 2–13.
Raninen, S. 2006. Tuskan teatteri Turun Kärsämäessä, II osa: väkivalta varhaisrautakauden kontekstissa. *Muinaistutkija* 3/2006: 2–22.
Rudbeck, L., Gilbert, M.T.P., Willerslev, E., Hansen, A.J., Lynnerup, N., Christensen, T. & Dissing, J. 2005. MtDNA analysis of human remains from an early Danish Christian cemetary. *American Journal of Physical Anthropology* 128: 424–9.
Saipio, J. 2008. Hävisikö sisämaan asbestikeraaminen väestö pronssikaudella? *Muinaistutkija* 2/2008: 2–18.
Sajantila, A., Lahermo, P., Anttinen, T., Lukka, M., Sistonen, P., Savontaus, M.-L., Aula, P., Beckman, L., Tranebjaerg, L., Gedde-Dahl, T., Isser-Tarver, L., DiRienzo, A. & Pääbo, S. 1995. Genes and languages in Europe: an analysis of mitochondrial lineages. *Genome Res* 5: 42–52.
Sajantila, A., Salem, A.-H., Savolainen, P., Bauer, K., Gierig, C., Pääbo, S. 1996. Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *Proceedings of the National Academy of Sciences* USA 93: 12035–9.
Salmela, E., Taskinen, O., Seppänen, J.K., Sistonen, P., Daly, M.J., Lahermo, P., Savontaus, M.-L. & Kere, J. 2006. Subpopulation difference scanning: a strategy for exclusion mapping of susceptibility genes. *Journal of Medical Genetics,* 43(7): 590–7.
Salmela, E., Lappalainen, T., Fransson, I., Andersen, P.M., Dahlman-Wright, K., Fiebig, A., Sistonen, P., Savontaus, M.-L., Schreiber, S., Kere, J. & Lahermo, P. 2008. Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in Northern Europe. *The Public Library of Science, PLos ONE* 3(10): e3519.
Salo, U. 2003. *Kauas missä katse kantaa: Unto Salon puheita ja kirjoituksia.* Satakunta 21. Emil Cedercreuzin säätiö, Helsinki.
Savontaus, M.-L. & Lahermo, P. 1999. Uralilainen muinaisuutemme väestögenetiikan valossa. In P. Fogelberg (ed.), *Pohjan poluilla: suomalaisten juuret nykytutkimuksen mukaan*: 60–3. Bidrag till kännedom av Finlands natur och folk 153. Societas Scientarum Fennica, Helsinki.
Service, S., DeYoung, J., Karayiorgou, M., Roos J.L., Pretorious, H., Bedoya, G., Ospina, J., Ruiz-Linares, A., Macedo, A., Palha, J.A., Heutink, P., Aulchenko, Y., Oostra, B., van Duijn, C., Jarvelin, M.R., Varilo, T., Peddle, L., Rahman, P., Piras G., Monne, M., Murray, S., Galver, L., Peltonen, L., Sabatti, C., Collins, A. & Freimer, N. 2006. Magnitude and distribution of linkage disequilibrium in population isolates and implifications for genome-wide association studies. *Nature Genetics* 38: 556–60.
Siiriäinen, A. 1981. On the cultural ecology of the Finnish Stone Age. *Suomen Museo* 87: 5–40.
Sundell, T. 2009. Kadonneet geenit: Suomen kivikautisten populaatiopullonkaulojen tutkiminen geneettisin menetelmin. *Muinaistutkija* 4/2009: 10–8.
Sundell, T. & Onkamo, P. 2010. Argeopop: Keramiikasta geenivirtoihin? *Muinaistutkija* 1/2010: 3–8.
Surovell, T.A. & Brantingham, P.J. 2007. A note on the use of temporal frequency distributions in studies of prehistoric demography. *Journal of Archaeological Science* 34: 1868–77.
Surovell, T.A., Finley, J.B., Smith, G.M., Brantingham, P.J. & Kelly, R., 2009. Correcting temporal frequency distributions for taphonomic bias. *Journal of Archaeological Science* 36: 1715–24.
Takala, H. 2004. *The Ristola Site in Lahti and the Earliest Postglacial Settlement of South Finland*. Lahti City Museum, Lahti.
Tallavaara, M., Pesonen, P. & Oinonen, M. 2010. Prehistoric population history in eastern Fennoscandia. *Journal of Archaeological Science* 37: 251–60.
Uimari, P., Kontkanen, O., Visscher, P.M., Pirskanen, M., Fuentes, R. & Salonen, J.T. 2005. Genome-wide linkage disequilibrium from 100,000 SNP's in the East Finland founder population. *Twin Research and Human Genetics* 8(3): 185–97.
Varilo, T., Laan, M., Hovatta, I., Wiebe, V., Terwilliger, J.D. & Peltonen, L. 2000. Linkage disequilibrium in isolated populations: Finland and a young subpopulation of Kuusamo. *European Journal of Human Genetics* 8: 604–12.
Varilo, T., Paunio, T., Parker, A., Perola, M., Meyer, J., Terwilliger, J.D. & Peltonen, L. 2003. The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Human Molecular Genetics* 12: 51–9.

BOX 1. Concepts of genetics

Allele: one of two or more alternative forms of a gene or DNA sequence.
Autosome: one of the 22 biparentally inherited chromosomes, each present in two copies in each cell.
Effective population size (Ne): the number of individuals in a given generation, whose gametes contribute to next generations; almost always considerably lower than the actual census size.
Gamete: eggs containing an X chromosome and sperm containing either an X or a Y chromosome.
Haplotype: the sequence of alleles of a set of polymorphic markers on the same chromosome.
Linkage disequilibrium (LD): non-random association between closely linked loci due to their tendency to be co-inherited.
Mitochondrion: a cellular organelle primarily concerned with energy generation, contains its own circular genome (mtDNA), maternally inherited.
Nucleotide: The molecular component of the polymers DNA or RNA. DNA and RNA sequence length is measured in nucleotides (also known as base pairs, bp)
Y chromosome: one of the sex chromosomes, only present in males.

BOX 2.

In genetics, a **population bottleneck** is defined as an event in which a considerable part of the population is prevented from reproduction. The population undergoes a considerable decrease in size (or the number of reproducing individuals), which may happen as a sudden incident or over a long time period. A sudden bottleneck may occur due to famine, epidemic, war or other reason. Only the survivors will pass their genes on to the next generations, which leads to a reduction of genetic diversity compared to the situation which prevailed before the bottleneck.

A **founder effect** takes place when a small number of individuals from a larger base population colonize a new site and no significant gene flow occurs thereafter between these two. The genetic consequences are very similar to a bottleneck. The theoretical effects of a genetic bottleneck or founder effect are well known in population genetics (see e.g. Jobling 2004), but specific ones taking into account all the archeologically or otherwise known prehistoric demographic events are better traced with simulations.

APPENDIX 1

Ancient Human mtDNA Haplogroups from Northern and Central Europe

Country	Site	Culture/Period	Dating ca.	Haplogroup	mtDNA	Reference
Sweden	Gökhem	TRB farmers	5500–4500 BP	H, J, T	16051–16383	Malmström
	Gotlanti	PWC, hunter-gatherers	4800–4000 BP	J, ?(2), T, U4/H1b(8), U5(3), U5a(3), V		et al. 2009
Denmark	Bøgebjergsgård	Roman Iron-Age	1 AD	H(2), I(2), R0a, U2e, U5b1	16064–16405	Melchior
	Skovgaarde	Roman Iron-Age	AD 200–270	H(4), H1, J(2), K(2), U3a, V		et al. 2007
	Kongemarken	early Christian	AD 1000–1250	H(3), I(2), J, T, T2, U7	16055–16408	Rudbeck
						et al. 2005
	Galgedil	Viking Age	AD 700–1100	H(5), I, K, T2, U5a1a, X2c	16064–16405	Melchior
						et al. 2008
Lithuania	Spiginas	hunter-gatherers	6350 cal BC	U4	15997–16409	Bramanti
	Donkalnis	Mesolithic	*	U5b2		et al. 2009
	Kretuonas	hunter-gatherers	4450 cal BC	U5b2		
			4200 cal BC	U5b2		
Poland	Dudka	hunter-gatherers	3650 cal BC	U5b1	15997–16409	Bramanti
			4000–3000 cal BC*	U5b1		et al. 2009
	Drestwo	hunter-gatherers	2250 cal BC	U5a		
Russia	Chekalino	hunter-gatherers	7800 cal BC	U5a	15997–16409	Bramanti
	Lebyazhinka	hunter-gatherers	8000–7000 cal BC*	U5a1		et al. 2009
Germany	Bad Dürrenberg	hunter-gatherers	6850 cal BC	U4	15997–16409	Bramanti
	Hohlenstein-Stadel	hunter-gatherers	6700 cal BC	U5a1, U5a2		et al. 2009
	Hohler Fels	Magdalenian	13400 cal BC	U(2)		
	Falkensteiner Höhle	hunter-gatherers	7200 cal BC	U5b2		
	Ostorf	hunter-gatherers	3200 cal BC	K, U5		
			3100 cal BC	U5		
			3000 cal BC	J, T2e(2)		
			2950 cal BC	U5a		
	Derenburg	LBK/AVK	7500–7000 BP*	HV(2), K, N1a(2)	15997–16409	Haak et al. 2005
	Eilsleben	LBK/AVK	7500–7000 BP*	H*		
	Flomborn	LBK/AVK	7500–7000 BP*	H*, K(2), T, N1a		
	Halberstadt	LBK/AVK	7500–7000 BP*	N1a, V, T		
	Schwetzingen	LBK/AVK	7500–7000 BP*	H*, T(3)		
	Seehausen	LBK/AVK	7500–7000 BP*	J*		
	Unterwiederstedt	LBK/AVK	7500–7000 BP*	K, N1a		
	Vaihingen an der Enz	LBK/AVK	7500–7000 BP*	U3		
Austria	Asparn Schletz	LBK	7500–7000 BP*	H*	15997–16409	Haak et al. 2005
Hungary	Ecsegfalva	AVK	7500–7000 BP*	N1a	15997–16409	Haak et al. 2005

Key. Culture/Period: AVK= Alföldi Vonaldiszes Kerámia, LBK= Linear Pottery Culture, PWC= Pitted Ware Culture , TRB= Funnel Beaker Culture; Dating: *= date based on cultural context; Haplogroup: the number of more than one individuals indicated in brackets