Chapter 5

Genetic demography: What does it mean and how to interpret it, with a case study on the Neolithic transition

Michela Leonardi1, Guido Barbujani2, Andrea Manica1

Abstract

The present work describes the basic principles underlying demographic reconstructions from genetic data, and reviews the studies using such methods with respect to the Neolithic Demographic Transition. It is intended as a tool for scholars outside the field of population genetics (e.g., archaeologists, anthropologists, etc.) to better understand the significance and intrinsic limitations of genetic demography, and to help integrate its results within the broader context of the reconstruction of the human past.

THE PRINCIPLES OF GENETIC DEMOGRAPHY: NULL MODELS, GENETIC DRIFT AND EFFECTIVE POPULATION SIZE

Genetic demography is the branch of population genetics aiming at inferring changes in the size of one or more given populations from genetic data. It is widely used to reconstruct the demographic trajectories through time (e.g., Leonardi, Barbujani and Manica 2017; Miller, Manica and Amos 2018) or to test which demographic model within a set of explicit ones better fits the observed data (e.g., Leonardi et al. 2018; Vai et al. 2019).

It is important to highlight that, in population genetics, demography is always considered in terms of effective population size ($N_e$), which, as we will discuss, is not a proxy for census size.

---

1 Department of Zoology, University of Cambridge, United Kingdom.
2 Dipartimento di Scienze della Vita e Biotecnologie, Università di Ferrara, Italy.

© 2021, Kerns Verlag / https://doi.org/10.51315/9783935751377.005
This section will answer the question “what is effective population size?” by putting the concept within its historical context. This will help to highlight its importance in the field of population genetics and its limitations when borrowed by other fields.

Population genetics can be seen, in a broad way, as an attempt to create models as precise as possible of the way a population evolves, i.e. its genetic diversity is transmitted from a generation to the following one.

The basis for this process started at the beginning of the 20th century, when a (very unrealistic – but extremely useful) “null model” was created, defining how populations behave when no evolutionary pressure is acting (Hardy 1908; Weinberg 1908).

The so-called Hardy-Weinberg equilibrium shows how a population is in equilibrium, i.e., does not change its allele frequencies from one generation to the following one, when it meets the following assumptions:

- Sexual reproduction and diploidy (two copies of the genome in each cell);
- Non-overlapping generations;
- Random mating;
- (Infinitely) large population size;
- Negligible migration and mutation;
- Mortality and fertility independent from genotype.

This model shows that, for evolution to occur, at least one of the assumptions is not met. The next steps, therefore, were to find the way each of these assumptions influences the evolution of a population.

A few years later, Wright and Fisher expanded this model (developing the Wright-Fisher model) to estimate what happens to allele frequencies when only the assumption of infinitely large population size is not met (Fisher 1923; Wright 1931). In a finite population, at each generation the individuals may or may not leave descendants, who then in turn may or may not survive: this results in each generation being a random sample of the previous one. This effect of random sampling is called genetic drift and is stronger (i.e., causes broader changes between generations) in smaller population relative to larger ones.

Now it is possible to get back to our original question: effective population size is defined, somewhat tautologically, as “the size of a Wright-Fisher population experiencing the same genetic drift as the one under study” (Jobling et al. 2014) which, as we have seen, can also be translated into “the number of individuals contributing to the following generation” (Ayala 1982).

In a nutshell, it measures the effect of random sampling on the population analyzed: the larger the population, the smaller the expected changes between generations due to random sampling of gametes.

Effective population size can be estimated on the basis of many different types of genetic markers with different methods, reconstructing its
Genetic demography: What does it mean and how to interpret it

variation through time. It must be highlighted, though, that such methods either tend to assume a single unstructured population evolving in isolation, or very simple effects of migration and other processes (e.g., just a few changes in population sizes through long lapses of time). It is important to consider that all the evidence discussed in this chapter is produced with methods of the first class, i.e., assuming negligible migration and geographic/cultural/social structure.

INTERPRETING EFFECTIVE POPULATION SIZE

It is very important to stress that estimates of effective population size hardly correspond to a something that, taken at face value, can be meaningful for other disciplines (Hawks 2008).

Effective population size can be regarded as a measure of the number of individuals actually contributing to the gene pool of the following generation. As a rule of thumb, in many cases it has been approximated as one-third of the census size, but there is no guarantee this assumption holds in every specific case. Therefore, census size is only one of the determinants of effective population size, together with a host of other factors such as gene flow (i.e., immigration/emigration), geographic structuring, age structure, mating/marriage patterns, sex ratio, breeding practices (for domesticates), social structure (for humans), etc. In the absence of genetic data from other populations, as well as additional information from other lines of evidence such as archaeology and historical records, it is impossible to disentangle the effect of these factors.

As an example, a population that has some form of geographic, social or cultural structuring (i.e., where individuals are more likely to have offspring within a specific subset of the population), will tend to have a larger effective size than a population with the same census size but where mating is completely random (Waples 2010). The reason is that, in such a scenario, each sub-population is relatively isolated from the others, and thus affected by drift in an independent manner; this effect is compounded by the fact that each sub-population is also smaller in size than the whole metapopulation, leading to stronger drift. The result is that sub-populations will end up differing from each other in their allele frequencies, potentially even developing private variants. In a random mating scenario, this level of genetic variability would require a much larger population size (Charlesworth, Charlesworth, and Barton 2003).

Another important point is that from a genetic point of view it is not possible to define the area inhabited by a population and its density, which add another level of complexity when trying to link population size with census size. There have been attempts at solving this problem by comparing effective population size with the potential range of species reconstructed based on climatic data. While there seems to be some level of correlation between the two in some cases (Lorenzen et al. 2011), this is not always true (Miller et al. 2021).
For all the reasons mentioned above, when we observe an increase in effective population size through time in our data it may be the result of many different scenarios. In the following list, as an example, we will detail some such possibilities both as general concepts and in a few situations that may occur in a human population (in italic).

- Increased population density when occupying the same geographic area: following a climatic amelioration, the environmental productivity increases and the occupied region can sustain more individuals; or, the development of new technologies allows better use of the available resources.
- An increase in fertility: the transition from a nomadic to a sedentary way of life changes the group behavior, allowing an increase in the number of children that a family can sustain at a given time.
- An increase in the geographic area habitable: climatic changes give access to a previously unoccupied region.
- Beginning or increase in immigration from one or more populations, genetically different from the original one: a military conquest; or, the development of new trading routes; or, immigration of a new population in the area.
- A change in marriage rules: after a military conquest foreign men tend to marry local women; or, because of new commercial connections there is an increase in marriages between the two populations involved, to strengthen such economical relationships.
- Any other situation leading to an increase of the genetic diversity within the population, including different combinations of the proposed scenarios.

Similarly, below we present some examples of processes that lead to a decrease in effective population size:

- Strong reduction in census size (in genetic terms defined as a bottleneck): an epidemic; or, a war; or, a decrease in food availability.
- A small group from the parental population moving away from the latter: the occupation of a new area by a subset of individuals (in genetic terms defined founder event);
- Reduction or interruption of genetic exchange with other populations (i.e., isolation): degradation and finally unavailability of a previously developed route or trade network.
- Population fragmentation (e.g., because of the emergence of geographical, cultural or social barriers within it): the development of a caste-like social system that excludes marriage between different social classes.
- Increase in mortality: increase of child mortality due to famine.
- Decrease in fertility: the needs of more frequent or longer migrations reduces the number of young children that the group can
carry at the same time; or, migrating in a harsher climate increases the time to reach sexual maturation in women.

- Reduction or geographical shift of the area environmentally habitable by the population (leading to the survival of only a part of it): a climatic change reducing the productivity of the region inhabited (e.g., Sahara).
- Increase of marriage between relatives: to keep power or wealth.
- Reduction in the number of individuals of a specific sex (there is a stronger effect if this affects females): a war.
- For domesticates: starting or change in breeding practices.
- Any other situation leading to a decrease of the genetic diversity within the population including combinations of the scenarios above.

CASE STUDY: GENETIC DEMOGRAPHY AND THE NEOLITHIC DEMOGRAPHIC TRANSITION

The advent of the Neolithic, i.e. the transition from foraging to farming, led to an increase in population density in many different regions of the world (Armelagos, Goodman, and Jacobs 1991; Bellwood et al. 2007; Kılınç et al. 2016; J.-P. Bocquet-Appel 2011; J. P. Bocquet-Appel and Bar-Yosef 2008).

From a genetic point of view, many studies have approached the subject by reconstructing the genetic demography in the present and through time. It is well established that the effective population size significantly differs between modern-day food producers on the one hand and foragers on the other, with the latter showing much smaller values (Excoffier and Schneider 1999; Destro-Bisol et al. 2004; Pilkington et al. 2008; Aimé et al. 2014; Patin et al. 2014; Leonard, Barbujani, and Manica 2017; Gopalan et al. 2019).

Several studies used different portions of the genome to reconstruct the demographic profile in populations with contrasting subsistence strategies to date the onset of these differences. The expectation was that food producers and foragers shared similar trajectories until more or less the Neolithic, when the former started increasing in numbers while the latter did not (Menozzi, Piazza, and Cavalli-Sforza 1978; Sokal, Oden, and Wilson 1991). However, contrary to this hypothesis, a large number of studies suggest a significant difference between the two groups and an increase in effective population size in farmers starting before the Neolithic (Leonardi, Barbujani, and Manica 2017; Atkinson, Gray, and Drummond 2008; Batini et al. 2015, 2011; Maisano Delser et al. 2017; Chaix et al. 2008; Zheng et al. 2011, 2012; Aimé et al. 2013; Patin et al. 2014; Miller, Manica, and Amos 2018). Since all populations relied on a hunting-gathering lifestyle before the Neolithic transition, this finding calls for an explanation.

A first possibility is that the observed evidence could be due to climate: an uneven distribution of the natural resources and/or the climatic
amelioration following the Last Glacial Maximum may have allowed some environments to sustain larger groups (Bar-Yosef 1998; Berger and Guilaine 2009). For this reason, both the differences between lifestyles and the early increase in the effective population size in farmers could be the result of the ancestors of modern-day food producers living in more productive environments than the ancestors of modern foragers. Such a more favorable climate could also have facilitated the development of food production, increasing the environmental productivity for foragers and allowing them to sustain even larger groups.

Was that the case? To answer this question, our recent paper (Leonardi, Barbujani, and Manica 2017) analyzed the trajectories of effective population sizes of present-day foragers and food producers from Sub-Saharan Africa, South-Eastern Asia and Siberia over the last twenty thousand years. Changes in effective populations sizes were then compared with estimates of Net Primary Productivity through time. In all three regions, food producers systematically show higher numbers of effective individuals than foragers, even after correcting for environmental productivity. Furthermore, the trajectories also indicate higher effective growth rates in the farmers, and this difference can only be attributed to farmers living in climatically more favourable regions for Siberian populations (but not for Sub-Saharan Africa and South-East Asia).

Another possibility is that we are observing the result of behavioral and social differences among Paleolithic hunter-gatherers. Present-day foragers exhibit large variability in term of complex behaviors that may influence their genetic diversity (and hence the estimation of effective population size), e.g., sedentism, storage activity and social stratification (Rowley-Conwy 2001). Those three aspects are all typical of food producers, and appear to be all linked to stationality. Indeed, it has been suggested that the transition to farming happened in foraging populations that were already sedentary, because they would have had the resources to sustain themselves while experimenting with agriculture and the continuity to check the cultivations as they developed (Sauer 1952). It has also been shown that, in modern-day hunter-gatherers, a larger ratio of population size over ecological productivity is positively correlated to a sedentary lifestyle, storage of goods and hierarchy, which means that populations that exhibit such behaviors are likely to have larger effective population size than other groups (Rowley-Conwy 2001).

The observed differences in effective population size are then coherent with a scenario where the Paleolithic ancestors of modern food producers were more sedentary/socially stratified/interconnected/etc. than the contemporary ancestors of modern-day foragers.

It is also possible that food producers originated from populations with larger effective size as both the main factors influencing it (higher population density or higher rates of gene flow and so cultural connectivity with neighboring groups) could facilitate technical innovation, possibly leading to an improvement in subsistence technologies. Indeed, a link between higher effective size and increase of cultural complexity
has been observed in different contexts (Powell, Shennan, and Thomas 2009).

It is important to stress that, as discussed above, the methods used to obtain the evidence presented rely on the assumption of a single population with negligible migration and population structure (e.g., geographic, social, cultural). In contrast, Neolithization involved the spread of human groups from a different region and subsequent mixing with local foragers was more than likely (e.g., Lazaridis et al. 2014; Prendergast et al. 2019). Whether, and to what extent, admixture between expanding early farmers and local hunter-gatherers affected the results of retrospective studies is difficult to tell. However, the observed gradual changes in effective population size and the consistent pattern all over the world do not suggest that admixture, which likely took place at different rates in different areas, has been the main cause of this phenomenon.

CONCLUSIONS

In short, changes in effective population sizes inform us that some process has been affecting the genetic diversity of the population analyzed, and identifying such process(es) should be done carefully and in the light of the archaeological/historical context.

The terms demography and effective population size suggest a variation in census size as the main reason to explain such evidence. However, and to the contrary, the results of genetic demographic analyses may be linked to a variety of processes. Luckily in many cases, archaeology, anthropology and other sister disciplines have the data and methods to test different possibilities and find the correct one. Modern Bayesian methods of demographic inference associated with simulation studies (Beaumont, Zhang, and Balding 2002) allow one to model in greater detail demographic changes, selecting the model best fitting the data among various options.

For example, immigration and introgression from a different population may be tested through the analysis of anthropological, isotopic and cultural data; the development (or collapse) of a trade network can be shown by studying the distribution through time of the sources of raw materials; social stratification can be suggested by looking at mortuary practices; the distribution of archaeological sites through space and time could be informative about population density, etc.

The reconstructions of genetic demography should therefore not be considered as providing definitive answers, but rather as a starting point for debate. A debate that can only be resolved through interdisciplinarity.

ACKNOWLEDGMENTS

ML and AM have been funded by the ERC Consolidator Grant 647787 “LocalAdaptation”. All Authors thank the two anonymous reviewers whose comments helped improving the quality of this manuscript.
REFERENCES


Genetic demography: What does it mean and how to interpret it


