Stationary Models of Population Structure Fail to Explain Observed Declines in Neanderthal and Denisovan N_e

Alan R. Rogers

October 16, 2023

Abstract

In a subdivided population, effective population size (N_e) reflects not only the number of individuals, but also the rate and pattern of gene flow. If we sample two haploid genomes from the same deme, estimates of N_e will be small in the recent past but large in the distant past, even if there were no changes in census size or in the rate or pattern of gene flow. This article asks whether this effect can plausibly explain observed declines in effective sizes of archaic populations.

Archaic human populations (Neanderthals and Denisovans) had very low heterozygosity. If this value represents a stationary equilibrium and migration is conservative, then heterozygosity alone allows us to estimate that each of these populations had a total effective size (including all subdivisions) of about 3600 individuals.

Both archaic populations also exhibit a pronounced decline in N_e over a period of about 400 thousand years. Under the finite island model of population structure, these data, together with a total population size of 3600, imply an implausible level of separation among archaic demes. Convergence is slower under models of isolation by distance. Thus, we might expect such models to provide a better fit to the slow observed decline. Yet even the most extreme form of isolation by distance, which restricts gene flow to a single dimension, implies an implausible level of separation among demes. Furthermore, these models all imply that the ratio of early to late N_e is much greater than that observed in archaic humans.

It seems unlikely that any model assuming stationary equilibrium can explain the pattern observed in archaic DNA. That pattern apparently reflects change either in the number of individuals or in the rate or pattern of gene flow.

1 Introduction

A variety of statistical methods use genetic data to estimate the history of effective population size, N_e [8, 16, 17, 20, 27]. In a subdivided population, these estimates depend not only on the number of individuals but also on gene flow between subdivisions [14, 29, 31]. Not only does N_e change in response to changes in gene flow [12, 15, 28], it may also exhibit a prolonged decline even when there has been no change either in the number of individuals or in the rate or pattern of gene flow [13, 15].

This article will ask whether this phenomenom can explain the estimated histories of archaic population sizes shown in Fig. 1. These show an apparent decline in population size between roughly 700 kya and 300 kya—an interval of 400 ky. Is it possible that this decline might result from subdivision alone, without requiring any changes in deme sizes or in the rate or pattern of gene flow?

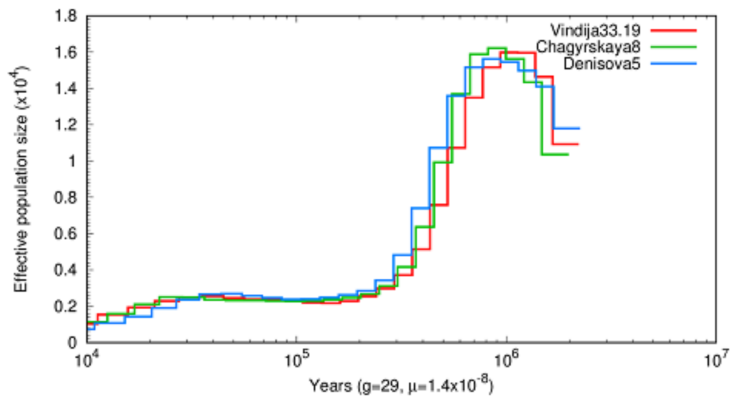

Figure 1: PSMC estimates of the history of population size based on three archaic genomes: two Neanderthals and a Denisovan [9].

2 Methods

2.1 Coalescent hazard and effective population size

Suppose that we sample two genes from a population at time 0 and trace the ancestry of each backwards in time. Eventually, the two lineages will *coalesce* into a common ancestor [5, 26]. The *hazard*, $h(t)$, of a coalescent event at time t is the conditional coalescent rate at t given that the two lineages did not coalesce between 0 and t [6, p. 6]. In a randomly-mating population of constant size N, this hazard is $h = 1/2N$ per generation. In a population of more complex structure, this formula doesn't hold. Nonetheless, we can define an *effective population size* as half the reciprocal of the hazard [19, p. 144].

$$N_e(t) = \frac{1}{2h(t)} \tag{1}$$

This quantity has also been called the inverse instantaneous coalescent rate (IICR) [2, 12, 13, 15]. That alternative name emphasizes that change in N_e (or IICR) need not imply change in the size of the population or even in the rate or pattern of mobility within it. Furthermore, this quantity depends not only on the characteristics of the population but also on those of the sample. Thus, the conventional term—effective population size—may be misleading. Nonetheless, I adhere to that convention in this article.

2.2 The finite island model

I will consider two models, both of which assume that population sizes and the rates and pattern of gene flow are constant. The first of these is the finite island model of population structure, which assumes that demes of equal size exchange migrants at equal rates [1, 12, 15]. In more realistic models, a gene beginning in one deme may need to traverse many intervening demes to reach a deme that is far away. There is no such necessity under the island model: a gene can move from one deme to any other in a single generation. Consequently, this model converges rapidly toward its asymptote for any given level of gene flow. This rapid convergence will exacerbate the difficulty of fitting the slow decline in Fig. 1. For comparison, I then study a model at the opposite end of the range of possibilities: one involving isolation by distance in a one-dimensional habitat. The hope is that these extreme models will bracket the range of realistic possibilities.

The finite island model assumes a population comprising d demes, each with N diploid individuals. Migration occurs at rate m per gene per generation. When a migration occurs (as we look backwards into the past), a lineage in one deme moves to one of the $d - 1$ other demes, which is

chosen at random. We study a pair of homologous genes sampled from a single deme. Coalescent events occur with hazard $1/2N$ per generation when the two ancestral lineages are in the same deme; they don't occur at all when the lineages are in different demes.

2.2.1 $N_e(t)$ under the finite island model

To study this model, let us formulate it as a Markov chain in continuous time. At a given time in the past, the two sampled lineages will be in one of three states: S, they are in the same deme; D, they are in different demes; and C, they have coalesced. Once the process reaches state C, it never leaves, so the probability of this state increases through time. Such states are said to be *absorbing*. States S and D, on the other hand, are *transient*; their probabilities will eventually decline toward zero. Let q_{ij} denote the rate of transitions from state i to state j. In other words, $q_{ij}h$ is approximately the probability that the process is in state j at time $t+h$ given that it is in state i at time t, provided that h is small. The matrix of transition rates looks like

$$\begin{array}{c} \\ S \\ D \\ C \end{array} \begin{array}{c} \begin{array}{ccc} S & D & C \end{array} \\ \begin{pmatrix} -2m - 1/2N & 2m & 1/2N \\ 2m/(d-1) & -2m/(d-1) & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}$$

as shown by Rodríguez et al. [15, p. 668]. Here, $2m$ is the rate of transition from S to D (because there are two lineages, each of which migrates at rate m), and $1/2N$ is the rate of transition from S to C. In other words, it's the rate of coalescence for two lineages in the same deme. The rate of transition from D to S is $2m/(d-1)$, because although $2m$ is the rate at which migrations occur, only a fraction $1/(d-1)$ of these events result in one lineage joining the other in the same deme. The diagonal entries in a transition rate matrix are the negative of the sum of the other entries in that row. (See Sukhov and Kelbert [24, sec. 2.1] for a discussion of transition rate matrices.)

We can simplify this model in two ways. First, because the absorbing state, C, does not contribute to the probabilities of the other states, we can delete the third row and column to form what Hobolth, Siri-Jégousse, and Bladt [4] call a "sub-intensity matrix." Second, we can reexpress all rates using $2N$ generations as the unit of time. This involves multiplying all entries of the matrix above by $2N$. The resulting sub-intensity matrix is

$$\mathbf{Q} = \begin{array}{c} \\ S \\ D \end{array} \begin{array}{c} \begin{array}{cc} S & D \end{array} \\ \begin{pmatrix} -M-1 & M \\ M/(d-1) & -M/(d-1) \end{pmatrix} \end{array}$$

where $M = 4Nm$ is twice the expected number of immigrant genes per deme per generation. Let $\tau = t/2N$ represent time in units of $2N$ generations, and let $\mathbf{p}(\tau) = (p_1(\tau), p_2(\tau))$ represent the row vector of probabilities that the process is in states S and D at time τ. It equals

$$\mathbf{p}(\tau) = \mathbf{p}(0)e^{\mathbf{Q}\tau} \tag{2}$$

where $\mathbf{p}(0)$ is the vector of initial probabilities, and $e^{\mathbf{Q}\tau}$ is a matrix exponential [3, p. 182]. Because we have sampled two genes from the same deme, $\mathbf{p}(0) = (1, 0)$, and $\mathbf{p}(\tau)$ is the first row of $e^{\mathbf{Q}\tau}$.

As τ increases, both entries of $\mathbf{p}(\tau)$ will eventually decline toward zero, because it becomes increasingly unlikely that the two lineages have not yet coalesced. We are interested, however, in the conditional probability that the two are in the same deme, given that they have not yet coalesced. At time τ, this conditional probability is

$$s(\tau) = \frac{p_1(\tau)}{\sum p_i(\tau)},$$

where the sum is across the transient states, excluding state C. The coalescent hazard is the product of $s(\tau)$ and the hazard for two lineages in the same deme. Returning now to a time unit of one generation, that product is $h(t) = s(t/2N)/2N$, and effective population size is

$$N_e(t) = N/s(t/2N) \tag{3}$$

This is equivalent to previously published results on the IICR ([12, Eqn. 16]; [15, p. 669]).

2.2.2 F_{ST} under the island model

To constrain guesses about the value of M, I make use of the fact that under the island model, F_{ST} (a measure of differentiation among demes [32]) is

$$F_{ST} \approx \frac{1}{1 + Md^2/(d-1)^2} \tag{4}$$

where the approximation is good if the mutation rate is small [22, 25].

2.3 The circular stepping stone model

One-dimensional stepping stone models describe a population whose demes are arranged in a line and exchange genes only with their immediate neighbors [7]. They are abstractions of real-world situations in which demes are arrayed across a landscape, and neighboring pairs of demes are less isolated from each other than are pairs separated by large distances. I will follow the common practice of assuming that the ends of the line of demes are joined to form a circle of d demes, each of size N, and that each pair of neighboring demes exchanges migrants at the same rate [1, 10, 11]. The circular arrangement is for convenience only. It simplifies things, because it implies that no deme is more central or peripheral than any other.

Each lineage migrates at rate m per generation in backwards time, and when it does so it is equally likely to move one step clockwise or one step counterclockwise around the circle of demes. Because we are studying the history of a pair of genes, the migration rate is $2m$ per generation or $M = 4Nm$ per unit of $2N$ generations, provided that the two lineages have not yet coalesced.

2.3.1 N_e under the circular stepping stone model

Chikhi et al. [2] used computer simulations to estimate IICR (or $N_e(t)$) under a linear stepping-stone model. Here I use the method introduced above. As in the island model, this process has one absorbing state, C, in which the ancestors of the two sampled genes have coalesced. Two lineages are separated by 0 steps if they are in the same deme, by 1 step if they're in adjacent demes, and so on. The maximum separation is $\lfloor d/2 \rfloor$, the largest integer less than or equal to $d/2$. For example, $\lfloor d/2 \rfloor = 3$ if d is either 6 or 7. The transient states in the model correspond to these distances: $0, 1, \ldots, \lfloor d/2 \rfloor$. These states label the rows and columns of the sub-intensity matrices below.

$$\mathbf{Q} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \overbrace{\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -M-1 & M & 0 & 0 \\ M/2 & -M & M/2 & 0 \\ 0 & M/2 & -M & M/2 \\ 0 & 0 & M & -M \end{pmatrix} \end{array}}^{d=6} \qquad \mathbf{Q} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \overbrace{\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -M-1 & M & 0 & 0 \\ M/2 & -M & M/2 & 0 \\ 0 & M/2 & -M & M/2 \\ 0 & 0 & M/2 & -M/2 \end{pmatrix} \end{array}}^{d=7}$$

In each matrix, row 0 refers to the case in which the two lineages are in the same deme. The only positive entry in that row equals M, because any migration in state 0 will move the process to state 1. The "–1" in the left-most entry accounts for coalescent events. In rows 1 and 2, the two positive entries equal $M/2$, because in states 1 and 2, migration is equally likely to increase by 1 or to reduce by 1 the distance between lineages. The entries in row 3 depend on whether d is even or odd. If it is even, there is only one deme that is $\lfloor d/2 \rfloor$ steps away from any given deme. Consequently, any migration event must reduce the distance by 1 step, and the transition rate is M. On the other hand, if d is odd, there are two demes $\lfloor d/2 \rfloor$ steps away. Half of the migration events will move a lineage from one of these to the other without changing the distance between demes. The other half reduce that distance by 1, so the transition rate is $M/2$. $N_e(t)$ is calculated from $\mathbf{Q}$ as before, using Eqns. 2–3.

2.3.2 F_{ST} under the circular stepping stone model

Under this model

$$F_{ST} \approx \frac{1}{1 + 6M/(d - d^{-1})} \tag{5}$$

provided that mutation is weak [30, p. 582]. This approximation breaks down when d is very large (compare [30, Eqn. 34]) but is quite accurate in the context of the models discussed below.

2.4 The magnitude of change in $N_e(t)$

For both of the models discussed above, $N_e(t)$ increases toward an asymptote in backwards time. The asymptotic value, $N_e(\infty)$, can be obtained from the left eigenvector of $\mathbf{Q}$ associated with the largest eigenvalue. This eigenvector is proportional to the asymptotic value of $\mathbf{p}(\tau)$, so we can use it to calculate $N_e(\infty)$, just as we used $\mathbf{p}(\tau)$ to calculate $N_e(t)$. If $\tilde{p}_i$ is the i'th entry of this eigenvector, then the asymptotic value of N_e is

$$N_e(\infty) = \frac{N}{\tilde{p}_1 / \sum \tilde{p}_i}$$

Mazet et al. [12] derived an explicit formula for the asymptote under the island model. The expression above holds for either model. The maximum proportional increase as we move backwards in time is $N_e(\infty)/N_e(0) = \sum \tilde{p}_i / \tilde{p}_1$.

2.5 Heterozygosity

In evaluating models of history, we can exclude those that imply values of within-deme heterozygosity, H_W, that are far from those observed in archaic genomes. Expected heterozygosity at equilibrium is approximately

$$E[H_W] \approx 4Ndu \tag{6}$$

when u, the mutation rate per site per generation, is small [21, 23]. This is the same as the heterozygosity of a randomly-mating population whose size equals that of the metapopulation. This result holds with remarkable generality, being affected neither by variation in the sizes of demes nor by the rate or pattern of migration among them, provided that the number of demes is finite, that migration is does not tend to change deme sizes, and that a gene beginning in any deme may eventually reach any other [30, p. 580]. In the real world, some demes may have suffered recent bottlenecks, which reduced their heterozygosity. But we are interested in whether the pattern in Fig. 1 can be explained without recourse to changes in population size and migration rates. Under

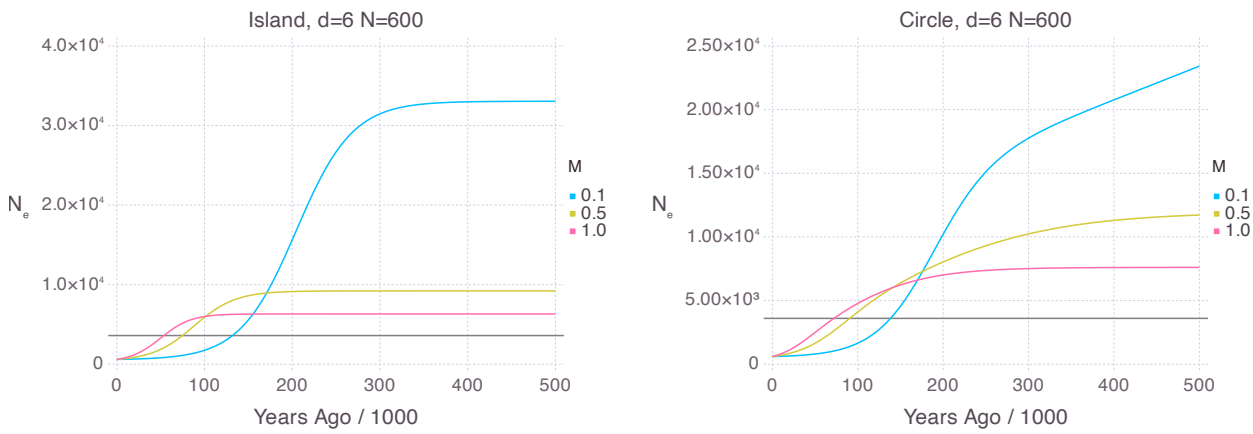

Figure 2: Time path of effective population size, assuming that two genes are sampled from the same deme and that the generation time is 29 y. The horizontal gray line shows the metapopulation size, Nd. F_{ST} values under the finite island model are 0.87, 0.58, and 0.41 for $M = 0.1$, 0.5, and 1. Under the circular stepping stone model, the corresponding F_{ST} values are 0.91, 0.66, and 0.49.

this hypothesis all demes—no matter how large or small, how central or peripheral—have the same heterozygosity.

3 Results

Archaic genomes consistently exhibit lower heterozygosities than those of modern humans. For example, Mafessoni et al. [9, table S8.2] list values between 1.5×10^{-4} and 2×10^{-4} for four archaic genomes: three Neanderthal and one Denisovan. To be conservative, I will take the upper end of this range as representative of archaic heterozygosity. I also assume a mutation rate of 1.4×10^{-8} per nucleotide site per generation. If we assume equilibrium under conservative migration, then Eqn. 6 implies that the entire population had effective size 3571. This small size seems implausible in view of the wide geographic extent of the Neanderthal population and the richness of its archaeological record. Nonetheless, let us ask what it implies in the context of two specific models of geographic population structure.

To stay within our budget of about 3600 individuals, Fig. 2 considers several models in which $d = 6$ and $N = 600$. None of the curves for the island model decline as slowly as archaic estimates in Fig. 1, even though Fig. 2 allows for extreme levels of isolation between demes. The blue curve, for example, assumes that demes receive an immigrant every 40th generation on average. For the island model, this implies that $F_{ST} = 0.87$, too high to be credible for a population of large mammals. With the circular stepping-stone model, the yellow curve (for $M = 0.5$) declines over an interval of about 400 ky, in agreement with the archaic data in Fig. 1. However this model also assumes an implausible degree of isolation: demes receive an immigrant every 8th generation, and $F_{ST} = 0.66$. I have done similar experiments (not shown) with other values of d and N but have been unable to find models in which N_e declines as slowly as the archaic data do, while keeping $Nd \approx 3600$, as required by the low archaic heterozygosity.

These results also depart from reality in another way. In Fig. 1, the early value of N_e is about five times as large as the recent one. Table 1 shows the ratio of early to late N_e as predicted by each of the models in Fig. 2. These values are all much larger than the five-fold change seen in the archaic data. The decline in N_e is slow only if M is small. But when M is small, the ratio of

Table 1: Predicted magnitude of change in N_e for the models in Fig. 2.

	$N_e(\infty)/N_e(0)$	
M	Island	Circle
0.1	55.1	79.4
0.5	15.3	19.9
1.0	10.5	12.7

$N_e(\infty)$ to $N_e(0)$ is large. Thus, the decline in N_e must either be too fast or too large and cannot fit the archaic human data.

4 Discussion

This article has asked when it is reasonable to interpret a decline in effective population size as the consequence of geographic population structure, without positing change in census population size or in the rate or pattern of gene flow among demes.

The expected decline is illustrated for two models of population structure in Fig. 2. These declines reflect the fact that we have sampled two genes from a single deme. In the recent past, it is likely that the two ancestors of these genes are still neighbors. Coalescent hazard is therefore high and N_e is small. As we move farther into the past, the ancestors are less likely to be neighbors, so the coalescent hazard declines and N_e increases, eventually reaching an asymptote.

At this asymptote, $N_e(\infty)$ is even larger than Nd, the size of the metapopulation. (See the horizontal gray line in Fig. 2.) Why should the asymptote be so large? Among all possible histories of a pair of genes, the subset that has not yet coalesced at time t will be enriched with histories in which the two lineages haven't spent much time together in the same deme. This implies that when t is large, the two lineages are less likely to be in the same deme than are two genes drawn at random from the population as a whole. Consequently, the coalescent hazard is less than $1/2Nd$, and $N_e(\infty) > Nd$. This excess is pronounced if M is small but disappears as M grows large.

Previous authors have fit estimated trajectories of $N_e(t)$ to models in which migration rate changed but population size did not [12, 15]. Because they were forced to posit changes in migration rate, it seems fair to assume that they were unable to fit models in which migration and population size were both stationary. The current results explain why some sort of change was needed.

5 Conclusions

The island model is a poor fit to archaic human data. Under that model, the low archaic heterozygosity together with the slow decline in archaic N_e (Fig. 1) imply an implausible degree of isolation among archaic demes. This suggests that a model of isolation by distance, which should converge more slowly, might provide a better fit. But because archaic heterozygosity was so low, there cannot have been a large number of demes of substantial size. This limits us to small models, and in such models we discover once again that the archaic data imply an improbable degree of isolation.

It therefore seems unlikely that any model assuming stationary equilibrium can account for the declines in N_e observed in Neanderthals and Denisovans. These declines presumably reflect changes either in census size or in the rate or pattern of gene flow.

Acknowledgements

I am grateful for comments from Simon Boitard, Lounès Chikhi, Josué Corujo, Daniel Tabin, and Olivier Mazet. The first draft of this manuscript [18] used an incorrect argument, which this version corrects.

Data, script, and code availability

Source code is available at `doi:10.17605/OSF.IO/S9WT8`.

Conflict of interest disclosure

The author declares that he has no financial conflict of interest with the content of this article.

Funding

This work was supported by a grant from the National Science Foundation, USA: BCS 1945782.

References

[1] D. Carmelli and L.L Cavalli-Sforza. "Some models of population structure and evolution". *Theoretical Population Biology* 9.3 (1976), pp. 329–359. DOI: `10.1016/0040-5809(76)90052-6`.

[2] Lounès Chikhi et al. "The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice". *Heredity* 120.1 (2018), pp. 13–24. DOI: `10.1038/s41437-017-0005-6`.

[3] D.R. Cox and H.D. Miller. *The Theory of Stochastic Processes*. London: Chapman and Hall, 1965.

[4] A. Hobolth, A. Siri-Jégousse, and M. Bladt. "Phase-type distributions in population genetics". *Theoretical Population Biology* 127 (June 2019), pp. 16–32. DOI: `10.1016/j.tpb.2019.02.001`.

[5] Richard R. Hudson. "Gene genealogies and the coalescent process". In: *Oxford Surveys in Evolutionary Biology*. Ed. by Douglas Futuyma and Janis Antonovics. Vol. 7. Oxford: Oxford University Press, 1990, pp. 1–44.

[6] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*. New York: Wiley, 1980. DOI: `10.2307/3315078`.

[7] Motoo Kimura and George H. Weiss. "The stepping stone model of population structure and the decrease of genetic correlation with distance". *Genetics* 49 (1964), pp. 561–576.

[8] Heng Li and Richard Durbin. "Inference of human population history from individual whole-genome sequences". *Nature* 475.7357 (2011), pp. 493–496. DOI: `10.1038/nature10231`.

[9] Fabrizio Mafessoni et al. "A high-coverage Neandertal genome from Chagyrskaya Cave". *Proceedings of the National Academy of Sciences, USA* 117.26 (2020), pp. 15132–15136. DOI: `10.1073/pnas.2004944117`.

[10] Takeo Maruyama. "On the rate of decrease of heterozygosity in circular stepping stone models". *Theoretical Population Biology* 1.1 (1970), pp. 101–119. DOI: 10.1016/0040-5809(70)90044-4.

[11] Takeo Maruyama. *Stochastic Problems in Population Genetics*. New York: Springer-Verlag, 1977.

[12] O. Mazet et al. "On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference?" *Heredity* 116.4 (2016), pp. 362–371. DOI: 10.1038/hdy.2015.104.

[13] Olivier Mazet and Camille Noûs. "Population genetics: coalescence rate and demographic parameters inference". *Peer Community Journal* 3.e53 (2023). DOI: 10.24072/pcjournal.285.

[14] Masatoshi Nei and Naoyuki Takahata. "Effective population size, genetic diversity, and coalescence time in subdivided populations". *Journal of Molecular Evolution* 37 (1993), pp. 240–244. DOI: 10.1007/BF00175500.

[15] Willy Rodríguez et al. "The IICR and the non-stationary structured coalescent: towards demographic inference with arbitrary changes in population structure". *Heredity* 121.6 (2018), pp. 663–678. DOI: 10.1038/s41437-018-0148-0.

[16] Alan R. Rogers. "Legofit: estimating population history from genetic data". *BMC Bioinformatics* 20 (2019), p. 526. DOI: 10.1186/s12859-019-3154-1.

[17] Alan R. Rogers. "An efficient algorithm for estimating population history from genetic data". *Peer Community Journal* 2 (2022), e32. DOI: 10.24072/pcjournal.132.

[18] Alan R. Rogers. "Declines in Neanderthal and Denisovan population sizes are probably not an artifact of subdivision". *bioRxiv* 551046 (2023). DOI: 10.1101/2023.07.28.551046v1.

[19] François Rousset. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, 2004. ISBN: 0691088179. DOI: 10.1515/9781400847242.

[20] Stephan Schiffels and Richard Durbin. "Inferring human population size and separation history from multiple genome sequences". *Nature Genetics* 46.8 (Aug. 2014), pp. 919–925. DOI: https://doi.org/10.1038/ng.3015.

[21] Montgomery Slatkin. "The average number of sites separating DNA sequences drawn from a subdivided population". *Theoretical Population Biology* 32.1 (1987), pp. 42–49. DOI: 10.1016/0040-5809(87)90038-4.

[22] Montgomery Slatkin. "Inbreeding coefficients and coalescence times". *Genetical Research, Cambridge* 58 (1991), pp. 167–175.

[23] Curtis Strobeck. "Average number of nucleotide differences in a sample from a single subpopulation: a test for population subdivision". *Genetics* 117.1 (1987), pp. 149–153. DOI: 10.1093/genetics/117.1.149.

[24] Yuri Sukhov and Mark Kelbert. *Probability and Statistice by Example: II: Markov Chains: A Primer In Random Processes and their Applications*. Cambridge: Cambridge University Press, 2008.

[25] Naoyuki Takahata. "Gene identity and genetic differentiation of populations in the finite island model". *Genetics* 104.3 (1983), pp. 497–512. DOI: 10.1093/genetics/104.3.497.

[26] Simon Tavaré. "Line-of-descent and genealogical processes, and their applications in population genetics models". *Theoretical Population Biology* 26 (1984), pp. 119–164. DOI: 10.1016/0040-5809(84)90027-3.

[27] Jonathan Terhorst, John A Kamm, and Yun S Song. "Robust and scalable inference of population history from hundreds of unphased whole genomes". *Nature Genetics* 49.2 (2017), pp. 303–309. DOI: 10.1038/ng.3748.

[28] John Wakeley. "Nonequilibrium migration in human history". *Genetics* 153.4 (1999), pp. 1863–1871. DOI: 10.1093/genetics/153.4.1863.

[29] M. C. Whitlock and Nicholas H. Barton. "The effective size of a subdivided population". *Genetics* 146 (1997), pp. 427–441.

[30] Hilde M. Wilkinson-Herbots. "Genealogy and subpopulation differentiation under various models of population structure". *Journal of Mathematical Biology* 37.6 (1998), pp. 535–585. DOI: 10.1007/s002850050140.

[31] Sewall Wright. "Isolation by distance". *Genetics* 28.2 (1943), pp. 114–138. DOI: 10.1093/genetics/28.2.114.

[32] Sewall Wright. "The genetical structure of populations". *Annals of Eugenics* 15 (1951), pp. 323–354. DOI: 10.1111/j.1469-1809.1949.tb02451.x.