Decision

All three reviews of this preprint are favorable, and I foresee a decision to recommend it. However, the reviews also had suggestions and complaints. I want to give the authors a chance to respond to these, so the present decision will be "revise and resubmit."

The first section of my own review (below) complains about the authors' suggestion that we use the term IICR rather than N_E. I won't insist that the authors abandon their own terminology, but I will probably make this point in my recommendation.

All three reviewers asked for more attention to previous literature. I think this will greatly strengthen the manuscript.

The first reviewer asks whether "typical natural populations [are] sufficiently structured for the problem to be serious." An answer to this question would strengthen the manuscript. I think the answer is tied up with the point about how genomes are sampled. Because PSMC looks at two haploid genomes from a single individual, those two genomes must have belonged to the same local population. Consequently, the coalescent rate in the recent past is that of the local population. In the distant past, however, the coalescent rate is that of the metapopulation. Because these two rates can be very different, the potential effect is very large.

There were two suggestions about sections of the manuscript that might usefully be deleted. One of these suggested deleting the section on the site frequency spectrum, and another suggested either deleting much of the math or moving it to an appendix. I think these would both improve the manuscript.

In my own review, I expressed puzzlement (verging on incredulity) at the large effect of selection on N_E. I hope the reviewers can help me intuit their result.

All three reviewers had concerns about the clarity of the manuscript and suggested ways to improve the exposition.

None of these problems are major, and I anticipate a positive response to a revised manuscript.

My own review

Major comments

Mazet and Noûs [4] review their own recent work on the interpretation of estimates obtained from the "partially-sequential Markovian coalescent" [PSMC, 3]. Such results are typically described in terms of the history of "effective population size," N_E. However, Mazet and his colleagues point out that they are also affected by geographic population structure and by selection.

Their review makes two substantive points, which I think are important. It also makes a semantic one that I find unconvincing. The semantic point is this: the quantity estimated by PSMC should not be called N_E because it is affected not only by the census size of the population but also by subdivision, rates of gene flow, and the way subdivisions are sampled. One can mimic observed PSMC estimates with models in which rates of gene flow vary but census size does not.

The trouble with this critique is that it could also be made of all other forms of N_E. For example, Sewall Wright derived a formula for the effective size of a population in which the number of males differs from that of females. In that case, N_E varies in response to the sex ratio, and changes in N_E need not involve any change in census size. It has always been true that N_E depends on things other than N. Consequently, population geneticists are unlikely to be convinced that we need a new term—the "inverse instantaneous coalescent rate"—just because change in N_E need not imply change in N.

Mazet and his colleagues are not the first to point out that the effective size of a subdivided population differs from that of an otherwise-equivalent population that mates at random. Nei and Takahata [5] derived a simple formula describing this effect nearly 30 years ago. In their model, subpopulations never go extinct, and subdivision inflates N_E. Whitlock and Barton [8] argued that when one allows for extinction of local groups, subdivision is more likely to decrease than to increase N_E. It would be useful to discuss the new work within the context of this previous work.

Previous work has also pointed out that coalescent rate depends on how the population is sampled. This goes back at least to Slatkin [7], who derives a formula for F_{ST} from the coalescence times of (a) two genes sampled from the same population, and (b) two drawn from different populations. Whitlock and Barton [8] also point out the effect of sampling. Mazet et al should describe their own contributions in the context of this previous work. It would also be useful to emphasize that we might learn something from PSMC by comparing results from (a) two chromosomes sampled from the same population to (b) two sampled from different populations. Li and Durbin [3] did this in their original publication on PSMC, using the X chromosome.

In my view, the novelty in the work of Mazet and Noûs [4] lies in its emphasis on transient dynamics. Previous authors have emphasized the steady-state equilibrium. This emphasis on the transient case is what allows Mazet et al to reinterpret the wiggles of a PSMC plot in entirely novel ways. Most existing work assumes that these wiggles reflect changes in census size, so Mazet et al provide an entirely new perspective for interpreting PSMC plots.

Mazet and Noûs [4] also make a second substantive point: that selection produces substantial effects on estimates of N_E. This makes qualitative sense (see below), but I'm puzzled by quantitative effect. I understand that selection may inflate the rate of coalescent events across a small fraction of the genome. But because this fraction is small, I would expect a correspondingly small effect on genome-wide estimates of N_E. This is what Nathan Harris and I found [2]. I hope the authors can provide some intuition. At present, their result seems so counterintuitive that I have trouble believing it.

I would suggest shortening or eliminating most of the mathematical material in the first part of this manuscript. A review such as this will be most valuable if it is accessible to a wide audience, and interested readers can presumably find the math in the original papers. Alternatively, one could move the math to an appendix. I would not insist on this; it's just a suggestion.

Minor comments

p. 2: "Each individual independently generates a number of descendants following a Poisson distribution." This is only approximately correct. The marginal distribution of an individual is binomial with parameters $1/2N$ and $2N$. The joint distribution of all individuals is multinomial.

Section "Structure and IICR: sampling strategy": This one-paragraph section is cryptic. It needs an example, or a fuller explanation, to make clear what is going on.

p. 13: "One way to model selection on genomic sequences is to assume that the portions of genomes under selection have an effective size different from the neutral areas." It would be useful to expand a bit about why this makes sense and what it leaves out. Charlesworth's model (which the authors cite) referred only to purifying selection, but the authors are interested also in directional selection. That makes sense too, because when one allele is favored, haplotypes carrying that allele will have short gene genealogies. The higher the frequency of the favored allele, the shorter the average genealogy. So directional selection also makes the average genealogy short, and this can be modelled as an effect on N_E within a restricted region of the chromosome. This is a simplification, because it doesn't capture the way selection makes gene genealogies unbalanced: some branches lead to lots of twigs and some to only a few. This principal was used by Seger et al. [6] also by

Field et al. [1] in their paper on the SDS statistic. Nonetheless, the authors' argument relies only on the average length of a gene genealogy, and so it makes sense to use this simplified model.

p. 13: It took me a few minutes to figure out what was going on here. I came up with the following prose to explain this to myself. Here it is, in case the authors want to paraphrase it. If the genome is divided into several zones, each with a different effective population size, then the coalescence rate in the recent past will be high because of all the coalescent events happening in zones of small N_E. Consequently, N_E will be small in the recent past. In the distant past, all lineages within zones of small N_E will have coalesced into a single ancestral lineage, and multiple lineages will remain only in zones of large N_E. Consequently N_E will be large in the distant past. This will look like a decline in population size. As explained above, however, it seems strange that this effect would be large, in view of the small fraction of the genome that is likely to be affected by selection.

p. 14: I was unable to understand the paragraph after Fig. 10.

References

[1] Yair Field et al. "Detection of Human Adaptation during the Past 2,000 Years". In: *Science* 354.6313 (2016), pp. 760–764. DOI: 10.1126/science.aag0776.

[2] Nathan S. Harris and Alan R. Rogers. "Genomic Regions Linked to Soft Sweeps Approximate Neutrality When Inferring Population History from Site Pattern Frequencies". In: *bioRxiv* (2020). DOI: 10.1101/2020.04.21.053629.

[3] Heng Li and Richard Durbin. "Inference of Human Population History from Individual Whole-Genome Sequences". In: *Nature* 475.7357 (2011), pp. 493–496.

[4] Olivier Mazet and Camille Noûs. "Population Genetics: Coalescence Rate and Demographic Parameters Inference". In: (2022). DOI: 10.48550/ARXIV.2207.02111.

[5] Masatoshi Nei and Nayouki Takahata. "Effective Population Size, Genetic Diversity, and Coalescence Time in Subdivided Populations". In: *Journal of Molecular Evolution* 37 (1993), pp. 240–244.

[6] Jon Seger et al. "Gene Genealogies Strongly Distorted by Weakly Interfering Mutations in Constant Environments". In: *Genetics* 184.2 (2010), pp. 529–545.

[7] Montgomery Slatkin. "Inbreeding Coefficients and Coalescence Times". In: *Genetical Research, Cambridge* 58 (1991), pp. 167–175.

[8] M. C. Whitlock and Nicholas H. Barton. "The Effective Size of a Subdivided Population". In: *Genetics* 146 (1997), pp. 427–441.