Dear Authors,

Thank you for submitting the revised version of your manuscript titled “Estimating Allele Frequencies, Ancestry Proportions, and Genotype Likelihoods in the Presence of Mapping Bias”. Both the reviewers and I are pleased with the improvements made in this version. The manuscript is now nearly ready for recommendation, pending a few minor revisions suggested by two of the reviewers.

Congratulations,

We are very pleased with the positive assessment of the revised manuscript. Our responses to the reviewer comments are written in italics.

Reviewer 1

add "(to allow for more mismatches and gaps)"

Added.

Very good you add all the justifications and details for parameters but I think it could be better to only mention relevant ones and maybe remove or keep the other for supplementary materials.

We believe that having all parameters and their descriptions in the same place maximizes reproducibility, therefore, we did not change this part.

SInce it is the last step, I think it oculd be better to put this sentence at the end of the paragraph.

Thank you for the suggestion, we have reorganized the paragraph.

Reviewer 2

Thank you for the detailed revision, it is much clearer what has been done now and it makes more sense.

Thank you very much for this assessment and your constructive feedback!

I only have a few remaining comments.

Note the line numbers refer to the track changes document.

137: Typo + I thought the doGeno -4 was to not print genotypes?

Thank you for catching this! We changed the description.

193: Based on your reviewer response I think this refers to the haploid fasta mentioned above but I don't think the reader wouldn't know that just from stating "different reference genomes" here

To clarify this, we have changed the sentence to "All reads (merged and the small proportion of unmerged) were then mapped to the haploid FASTA files representing reference genomes from the three populations (S1, S2 and S3) using bwa aln v0.7.17…"

251-253: While the short read lengths reflect aDNA, I think it is more likely the aDNA damage that influences mappability of the reads the most (as more mismatches mean less probablity of mapping) so I am not sure just shortening the reads produces results comparable to empirical aDNA damaged reads

The reason why mapping bias is exacerbated in aDNA data is probably a combination of the two: fragmentation leads to short DNA reads for mapping and deamination increases the number of mismatches proportional to their length. We have added the following sentence to explain that: "In addition to fragmentation, deamination is a major factor contributing to mapping bias in ancient DNA due to the resulting excess of mismatches, which we did not explore here."

Figure 3: What exactly is the Y-axis showing? I assume it is the different ancestry proportion but of which population, S2 or S3? Also, the dotted line across the figure is a little confusing as it only corresponds to one X-axis value

We added a description to the figure caption that this is the proportion of gene flow from S3. We have also reduced the length of the dashed lines so they do not cover the full width of the plots.

404-405: We found a similar result when mapping empirical data with simulated aDNA damage to different reference genomes and running Dstatistics with mapping to an ingroup producing the most reliable results - https://doi.org/10.1016/j.cub.2024.04.050 Supplementary figure S2 (Feel free not to cite it, I just thought it may be relevant)

Thank you for this suggestion, we were not aware of this supplementary result. We have added the citation. We are happy to cite this study since we assume your constructive feedback was partly based on the experiences you made during those analyses.

Reviewer 3

Thank you for addressing my comments and suggestions so thoroughly. I am pleased with the changes made to the manuscript, including the clarifications provided in your responses. The adjustments to the discussion, the inclusion of the JPT population, and the detailed explanation of your reproducibility measures significantly strengthen the paper.

I also appreciate the effort to correct typographical errors and improve the overall presentation of the manuscript. I am happy with the revisions and have no further comments. Thank you for your thoughtful work on this.

We appreciate your feedback and thank you for this positive assessment.