nature portfolio

Peer Review File

Improved Allele Frequencies in gnomAD through Local Ancestry Inference

Corresponding Author: Dr Elizabeth Atkinson

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This paper described a computational pipeline that parses the local ancestry inference (LAI) results and collect ancestry-population-specific allele frequencies. It would augment the popular resource gnomAD by having higher resolution ancestry-population-specific allele frequencies, which would help interpretations of alleles in a higher resolution ancestry population context. Their main contribution is the resource. The improvement over previous version of gnomAD is however not quantified but I suspect is incremental. Their conceptual contribution and method is straightforward and minimal.

Below are my specific comments:

Major:
The work seems to have an unspoken assumption: The genetic ancestry is actually discrete. It is the recent admixture events that making them seemingly continuous. By deconvoluting genome into local ancestry tracks, one can reveal its pre-admixture ancestry. I am keen to see if the author would articulate their assumptions.

If such assumption is indeed held by the authors, I am curious where does this assumption breaks. One can always deconvolute populations even into intra-continental groups. Will this method still holds there? They admitted that “genetic ancestry is truly continuous8”. I am wondering how would they reconciliate this contradiction.

I can see in Figure 2 that LAI deconvolution can have identifying causal sub-populations of a GWAS hit. But translating it quickly to improevd clinical interpretation might need some additional considerations?

I can see the authors hope to reveal some allele frequency divergence trend between populations. However, the plots are mostly simple stacked barplots. It would be more informative if some population genetics theory informed discussions can be made.

The authors claimed “The large sample size of gnomAD enables us to examine ancestral allele frequency trends at an unprecedented scale, and with the integration of LAI, at an unprecedented resolution.” I would love to see this unprecedentedness be supported by some data, e.g., some before-after contrast analysis.

What is “full gnomAD collection” shown in Figure 1a? It is not surprising to see two individuals b and c with different local ancestry got collapsed in this PC plot, as the PC plot for the top 2 PCs are only able to show very crude levels of global ancestry. It might be interesting to show if the two are separable in a higher PC plots (e.g., PC7 vs PC8).

Minor:
Somehow Figure 1 panel b and c shown as blank.
Figure 2 has low density of information and can be probably more effectively shown as a table.
Figure 4: again low information density. Unclear what is AF_nfe.
Extended Data Figure 1 is actually more informative and can be merged with Figure 1 and 2 into a multi-panel figure.

(Remarks on code availability)

The code is basically a dump of their own pipeline. No attempt to making a public-facing README.md. Also, their pipeline is presented in a way that is only useful to re-produce their LAI calls. Not even sufficient to reproduce the figures in the paper. Not to mention to be a useful tool to the community for other purposes.

Reviewer #2

(Remarks to the Author)
Paper summary: The authors analyse two gnomAD genetic ancestry groups with high degrees of recent admixture - "Admixed American" (n=7,612) and "African/African American" (n=20,250). By conducting local ancestry inference, they derive more precise ancestry-specific allele frequencies of genomic variants. The results of this work are made publicly available through gnomAD and will potentially further our understanding of human genomics.

Below are my comments.

Introduction
Probably need more detailed description of the two admixed populations ("Admixed American" and "African/African American") in the Introduction. There is some information in the first paragraph in the Methods, but it needs to be expanded and brought forward.

Results
"Here we release local ancestry-informed frequency data for 14,804,207 bi-allelic single nucleotide polymorphisms (SNPs) within the Admixed American (n=7,612) and 24,204,574 bi-allelic SNPs within the African/African American genetic ancestry groups (n=20,250), both from gnomAD v3.1." (lines 120-123)
This statement suggests to me that the variants reported are derived from WGS data, is that correct? Maybe worth stating it explicitly in the manuscript. For which variants local ancestry data is not available and why? What proportion of the human genome was successfully LAI resolved for "Admixed American" and "African/African American" groups; what proportion of the variants identified for "Admixed American" and "African/African American" groups have LAI data?

The term "Amerindigenous" would also benefit from expanded description at first reference (line 136, and in Fig1) – only in Methods they are described as the harmonized "AMR" samples from HGDP and 1kGP3.

In Figure 1a, I cannot see the "African reference samples" (line 134), only "African/African American"? In an attempt to derive a more eye-friendly colour scheme for Figure 1a, I would also suggest creating a "simplified" Figure 1a, containing only "relevant" populations – i.e., "Admixed American", "African/African American", "African reference", "East Asian", "South Asian" and "European" and move the current version to the Supplementary Figures. For consistency, please use either "Admixed American" or "Latino/Admixed American" (Figure 1a, lines 160 and 273, Supp Figure 3).

Line 141: should not "homogenous AMR" be "homogenous AFR or AMR or EUR"?

Figure 2: Maybe the right-hand side legends in each of the plots are not necessary; removing them would allow for larger font size to be used for the axis labels.

Figure 2a: Should not the reference for the variant in the SLC16A11 gene be to 14, rather than 16? Am I correct that for the zoom-in plot (and line 163), "LAI Amerindigenous" would be more appropriate/consistent label than "LAI Continental American"? These are the only 2 places where the term "LAI Continental American" is used.

Figure 2b: Should not the reference for the variant in the SLC16A11 gene be to 15 or 16, rather than 17? Please check the references throughout the manuscript.

Lines 170-177: Providing a link to the gnomAD page for this variant would be illustrative.

Lines 177-182: Using the online version of gnomAD v3.1.2 shows from the "gnomAD tab": Total AF = 0.0002768 and African/African American AF = 0.001018. From the "Local Ancestry" tab: Total/Overall AF = 0.001043 (due to the slightly smaller AN in Local Ancestry compared to gnomAD), LAI African AF = 0.001297 and LAI European AF = 0.000 (with AC = 0). Can the discrepancy with the reported "0.007% gnomAD-wide ("global") frequency (line 180) be explained with the different gnomAD versions – in the manuscript v3.1, while currently available online is gnomAD v3.1.2?

Figure 3: I would suggest using consistent labelling throughout the manuscript to distinguish more easily between "ancestry-specific allele frequencies" derived in this work (e.g., "LAI EUR", "LAI AFR", "LAI Amerindigenous") and reference population allele frequency as reported in gnomAD (e.g., "gnomAD EUR", etc.). Also, from the main text (and the figure title) I understand that Figure 3 depicts data about variants with LAI frequencies at least 2 times more frequent in one of the constituent populations compared to the LAI frequencies in other constituent populations, but the use of the word "different" in the figure text is confusing (it suggests also at least 2 times less frequent).

I have several comments related to Figure 4 and the corresponding text (lines 200-208).
First, for what proportion of the LAI variants discovered in the "Admixed American" and "African/African American" populations a ClinVar annotation is available?
Second, the statement "The results indicate that most variants across all ClinVar categories, including pathogenic/likely

pathogenic, benign/likely benign, uncertain significance, and conflicting classifications of pathogenicity are rare, with AFs concentrated between 0 and 0.1 in all genetic ancestry groups" should not be made without providing the site frequency spectrum distributions considering *all* variants discovered in these two cohorts and contrasting to the variants with ClinVar annotation only, thus demonstrating different frequency distributions.
Third, while ClinVar is an invaluable resource, it should be used with care. The authors do not specify any criteria for including/excluding a ClinVar annotation. If no such criteria were applied, it can be expected that a significant proportion of the ClinVar annotations are of questionable quality (e.g., those with single submitter, no criteria provided, etc.). Furthermore, even a robust "benign/likely benign" ClinVar annotation of a variant for condition A does not imply much for the same variant in the context of another condition B.
Fourth, I completely agree with the need for more diversity in genomic studies. However, I am not convinced that "The broader distribution of benign variants in African/African American groups …." (line 205) is not, at least to some degree, a result of variants being labelled as "benign/likely benign" in ClinVar simply due to their higher AF in gnomAD reference populations (e.g., for rare disorders dominant variants with AF > 0.0001 and recessive variants with AF > 0.005 are usually excluded from further analysis and follow-up). Maybe add zoom-in to variants with AF < 0.01/0.1?
For consistency, I would suggest to use LAI_AFR, LAI_EUR and LAI_Amerindigenous for the labels in Figure 4; could you please improve the readability of Figure 4 by increasing the font size for example?

Lines 210-224: I find this paragraph difficult to follow. Given that it describes one of the major contributions of this work, clearer rephrasing would be beneficial. Please define grpmax and popmax (grpmax first mentioned in the Abstract, line 46). My understanding of what is described is as follows. Performing a LAI based analysis of the two recently admixed cohorts results in higher values for the grpmax metric for the vast majority (81.49%) of the variants found in them, compared to the same metric computed based on other (non-bottlenecked) populations available in gnomAD. These updated grpmax values in turn will allow for correctly excluding more non-disease-causing variants when assessing an individual from the underrepresented "Admixed American" and "African/African American" groups, supported by the observation that 87.69% and 89.32% of the variants in the two cohorts with higher corrected grpmax values are found in ClinVar labelled as "benign / likely benign" (with the ClinVar use caveat as discussed above). Is my understanding correct?

Discussion
Investigating the reverse question – are there any variants, which based on the conducted LAI analysis of the two cohorts, should not be overlooked/discarded when assessing their pathogenicity? - may provide another important contribution of this work. This is somehow described (lines 268-277), focusing on a single example of a variant associated with type 2 diabetes risk, with global AF of 2%, but much more frequent in Admixed American population, driven by its frequency in Amerindigenous individuals. Are there any other such variants? In these two groups, are there any otherwise generally rare protein-coding variants with predicted functional impact (i.e., nonsense/missense) and potentially associated with a rare/common disorder, which exhibit much higher LAI derived AF in one of their constituent ancestries? Given that the sets of variants enriched at least twice in one LAI ancestry compared to the other(s) for the two groups is available (Figure 3), the suggested systematic analysis could be relatively easily performed using some carefully designed filtering steps.

References
Missing authors names for reference 8.

Methods
No comments

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
In this work, the authors calculate, report and share allele frequencies stratified by local ancestry, across biallelic gnomAD variants. To accomplish this, they ran local ancestry deconvolution on the genotypes of African/African American and Admixed American participants from gnomAD. To justify the impact of their work, the authors begin by demonstrating the usefulness of local ancestry deconvolution for capturing between-individual heterogeneity in genetic variation. Subsequently, they demonstrate several ways in which "local ancestry"-stratified allele frequencies are useful. These include: (1) characterizing the impact of common variants; (2) characterizing the impact of rare variants; (3) improving on existing methodology for filtering variants in clinical variant interpretation tasks.

I consider this work important for multiple reasons. First, it demonstrates how studies on admixed populations can support interpretation of variant effects. The authors list multiple well-studied examples of how (local) ancestry-specific allele frequencies of functional variants in disease-linked genes could better account for the distribution of disease incidence between ancestrally diverse groups of individuals. While it would be ideal to report ancestry-specific variant associations with the trait alongside the corresponding "pooled" variant association (i.e., trait-variant correlations with/without conditioning on local ancestry, or even just admixture mapping P-values or trait prevalence in the same gnomAD study cohort), it is still reassuring to see that the examples do broadly show a better concordance between ancestry-specific allele frequency and the disease incidence heterogeneity between ancestrally diverse individuals.

Second — and I consider this a broader scope of impact — this work provides new ideas for investigating complex traits within admixed individuals. We know that variants with highly differentiated ancestry-specific allele frequencies (e.g., the

Duffy null polymorphism rs2814778 at chromosome 1q23.2) can explain differential trait expression. We also know that rare mutations are more likely to be pathogenic, but the notion of rare depends on how the allele frequency or allele count is computed. The authors explored these ideas briefly in their paper, by reporting overall statistics of allele frequency heterogeneity by local ancestry, and by comparing distribution of allele frequencies in variants stratified by ClinVar annotations. I consider their finding, that benign variants seem to span a larger range of allele frequencies when tagged by African local ancestry, reasonable for justifying the use of local ancestry-stratified allele frequencies as a preliminary filter of variant pathogenicity. This is because (at least to my understanding) a strongly pathogenic variant should be selected against, given any population history that is not also prominently marked by the potentially confounding effects of bottleneck and genetic drift. So, by considering allele frequencies across as many population histories possible — including both recently admixed populations and historically admixed populations subsequently isolated long enough to lose any signature of ancestral mosaicism — our confidence of an allele that segregates at a moderate frequency being benign should increase. Arguments of this sort can also be provided to evaluate the impact of a variant in environments specific to a group of individuals, potentially enhancing our understanding of the effect of a mutation across different environmental contexts.

Finally, some comments on methodology and analytical approach. The use of Eagle for phasing, followed by RFMix (v2) for local ancestry inference (LAI) and Tractor for allele dosage calling all sound reasonable to me. The reference populations used for LAI also sound reasonable to me. I do not see any glaring issues with their analyses, though I have a small suggestion (see below).

In summary, I congratulate the authors for performing these important analyses and for making their allele frequency data publicly available. I have some comments, one of which I consider a necessary action item (Major Comments).

Major Comments
1. The authors write in Lines 217-218 that “87.69% and 89.32% of ClinVar variants with a higher tract AF than grpmax AF were classified as ‘Benign/Likely Benign’...” and conclude that LA-informed AFs have a substantial influence such that their exclusion could lead to inaccurate variant classification. I find their evaluation not very convincing. I also looked at Supplementary Table 1, and was not sure where the change in counts is shown. If the original grpmax/popmax AF was sufficiently large, then I would not expect the “Benign/Likely Benign” label to change from the inclusion of LA-informed AFs to its calculation. I would be more convinced if one looked at variants with original grpmax/popmax AF near the threshold AF for filtering (below, not above it). Show that among these variants, those less pathogenic are likely to have larger LA-informed AF (which would be more convincing for showing the power of LA-informed AFs in “correcting” a previously inaccurate classification).
2. Building on my previous point, one could use the same ClinVar variant classification table to label variants here, but I would also like to see (and highly encourage) the use of Variant Effect Predictors (VEPs), for example CADD. (CADD scores are continuous and can be downloaded from https://cadd.gs.washington.edu/.) I do not expect the authors to perform substantial comparisons using multiple VEPs, but given the growing interest in VEPs (https://www.varianteffect.org/veps) I consider this type of analysis crucial for broadening the impact of computational variant effect predictors to diverse groups of individuals — echoing a point raised by some of the present authors in their earlier work (see, e.g., first two sentences of Atkinson et al., 2021 “Reply to: On powerful GWAS in admixed populations” in Nature Genetics).

Minor Comments
1. In Lines 205-206, the authors write that “the broader distribution of benign variants in African/African American groups reinforces the need to include diverse ancestry groups in genomic studies”. I am curious if the site frequency spectra (SFS) are significantly different. Could the authors run a simple non-parametric test and report p-values (say, by converting the SFS into an empirical probability distribution and then running a Kolmogorov-Smirnov test)?
2. In Lines 210-211, I was initially a bit confused by what grpmax meant, though in retrospect, from the name I was able to make a good guess. It would be helpful to add a short definition of grpmax to remind the reader exactly what it measures.

(Remarks on code availability)
From looking at the Github repository, the authors' Python scripts all have basic documentation and so are interpretable. However, I did not execute their code to evaluate its functionality, because I do not have the input data files to test their code.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors addressed all my concerns.

(Remarks on code availability)
Code is acceptable.

Reviewer #2

(Remarks to the Author)
I believe the authors addressed my questions related to the original version of their manuscript; I have no further comments.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Overall feedback
The authors have addressed most of my concerns. Overall, I find that the revised manuscript has a clearer and broader scope of impact, thanks to the expanded list of examples and systematic analyses comparing allele frequencies across clinically and computationally relevant benchmarks. I have a few comments, arranged in approximately decreasing order of issue seriousness.

Comments
1. In Lines 274-275, the authors write that "ancestry-specific allele frequency information can better assist in the identification of likely benign variants." I am not so sure that is the conclusion of the analysis of tract-specific AF exceeding grpmax values — how is tract-specific AF exceeding 5% assisting with benign variant identification? The BA1 guideline seems to have been developed from various lines of evidence across analyses of continental populations. (In Ghosh et al., 2018 PMID: 30311383, there is no mention of admixed population or genetic ancestry group.) It is not clear to me that the variants identified with tract-specific AF > 5% have already been validated as likely benign. There is also potential noise in the estimated AFs given the smaller effective sample size of these local ancestry-conditioned AFs. In my view, a more assumption-lean conclusion is that these variants potentially complicate interpretations of variant benignness based on the BA1 guideline. Or perhaps there is already orthogonal evidence (e.g., functional depletion using VEPs) demonstrating that these variants are indeed likely benign, in which case the authors should report and cite it. I would like the authors to clarify this point in their current manuscript.
2. In Lines 276-277, the authors write that "CADD PHRED score thresholds (≥10 and ≥15, corresponding to the top 10 and 5 percent...". I could be mistaken, but given the formula -10*log10(rank/total) for PHRED scaling, shouldn't a cutoff of 15 correspond to the top 3 percent (since -10*log10(0.031) = 15.1)? If so, please change it.
3. Which CADD version was used to score variants? Computational VEPs like CADD are regularly updated, so please specify the version used here to improve reproducibility of the analysis.

(Remarks on code availability)
The revised Github repository appears more well-documented. I was also able to access the toy dataset in the Tractor wiki.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
The authors have addressed all of my concerns. There is only a small issue, which the authors should be able to address quickly. On p.15, the authors write "...the CHIT1 variant of uncertain significance...". I was not sure which variant this is, and could not find any other reference to it in the Main Text. It would be informative to specify the variant.

(Remarks on code availability)
-

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This paper described a computational pipeline that parses the local ancestry inference (LAI) results and collect ancestry-population-specific allele frequencies. It would augment the popular resource gnomAD by having higher resolution ancestry-population-specific allele frequencies, which would help interpretations of alleles in a higher resolution ancestry population context. Their main contribution is the resource. The improvement over previous version of gnomAD is however not quantified but I suspect is incremental. Their conceptual contribution and method is straightforward and minimal.

We thank the reviewer for their careful evaluation of our manuscript. We have made efforts to address all of your suggestions in detail, described point-by-point below, and believe that they have helped improve the quality of our work. Notably, we now release a generalized software deliverable, complete with tutorial-style usage instructions, alongside the scripts we used to actually complete our analyses. We believe this is a valuable resource for the community to compute local ancestry specific allele frequencies from their own cohort data to complement the browser resource that we describe in detail here. We have also done a number of additional analyses to refine and expand our results, and better highlight the improvement of our LAI-informed frequencies to the aggregate frequency data in gnomAD. We hope that you agree that these revisions to our manuscript better illustrate the value of this work and the resources we release here.

Below are my specific comments:

Major:
The work seems to have an unspoken assumption: The genetic ancestry is actually discrete. It is the recent admixture events that making them seemingly continuous. By deconvoluting genome into local ancestry tracks, one can reveal its pre-admixture ancestry. I am keen to see if the author would articulate their assumptions.

If such assumption is indeed held by the authors, I am curious where does this assumption breaks. One can always deconvolute populations even into intra-continental groups. Will this method still holds there? They admitted that “genetic ancestry is truly continuous8”. I am wondering how would they reconciliate this contradiction.

We appreciate the reviewer's insightful comments regarding the assumptions underlying our approach. We fully acknowledge that genetic ancestry is inherently continuous rather than discrete, as populations have experienced complex histories of gene flow and admixture over time. However, for the purposes of local ancestry inference (LAI), we adopt a model in which we approximate genetic ancestry as originating from a set of discrete source populations. This is a necessary simplification for inference, as LAI methods rely on reference panels of haplotypes assigned to specific ancestral populations to deconvolute ancestry at a local level.

Our approach does not assume that present-day populations are strictly discrete entities but rather that the major sources of genetic variation in recently admixed populations can be well-modeled by a small number of continental or subcontinental reference populations. This assumption holds particularly well for cases of recent admixture (e.g., within the past 10-20 generations), where distinct ancestry components remain identifiable along the genome. However, we recognize that as one moves beyond recent admixture events, ancestry becomes increasingly diffuse, and the distinction between discrete ancestry sources becomes less clear.

We also acknowledge the reviewer's point that genetic ancestry can be deconvoluted into finer-scale subcontinental groups. Our method is not inherently limited to broad continental ancestries, and in principle, it could be extended to analyze intra-continental variation given appropriate reference data. However, the resolution of such analyses depends on the availability and comprehensiveness of reference panels (which are currently insufficient for many areas of the globe), as well as the degree of genetic differentiation between subpopulations (that is, deconvolving more recently diverged tracts is more challenging and has lower accuracy). This is a fundamental limitation of all ancestry inference approaches and reflects the underlying complexity of human population structure. As reference resources continue to expand, we would love to resolve intra-continental level ancestral components in the future.

In summary, while we approximate genetic ancestry as discrete for the purpose of local ancestry inference, we do not assume that human populations are inherently discrete. Instead, we use a pragmatic framework that captures the major ancestry components relevant to recently admixed populations, while acknowledging that this framework has limitations in contexts of deeper or more continuous population structure. We have worked to clarify this point further in the manuscript to ensure that our assumptions and their limitations are explicitly stated. Please see *Page 16, Lines 362-376*:

"We recognize that genetic ancestry is inherently continuous and not discrete. For the purposes of local ancestry inference, we approximate ancestry as originating from discrete source populations, a necessary simplification for modeling recently

admixed populations. Our approach does not assume that present-day populations are strictly discrete entities but rather leverages the fact that the major sources of genetic variation in recently admixed populations can be well-modeled by a small number of continental reference populations. This assumption holds well for cases of recent admixture, however, as one moves beyond recent admixture events such ancestry becomes increasingly diffuse. We additionally appreciate that continental regions are not monoliths, but contain numerous ancestries therein. Our method is not inherently limited to broad continental ancestries, and in principle, it could be extended to analyze intra-continental variation given appropriate reference data. However, the resolution of such analyses depends on the availability and comprehensiveness of reference panels as well as the degree of genetic differentiation between modeled ancestries. For this reason, we are currently limited in resolving LAI to the continental rather than subcontinental level. We look forward to resolving finer-scale ancestries in the future as reference panels supporting such analyses continue to expand."

I can see in Figure 2 that LAI deconvolution can have identifying causal sub-populations of a GWAS hit. But translating it quickly to improevd clinical interpretation might need some additional considerations?

Thank you for this thoughtful comment. We agree that while Figure 2 demonstrates the potential of LAI for pinpointing sub-population-specific signals, translating these insights into improved clinical interpretation does require further steps. We now clarify this in the manuscript and emphasize that integrating LAI-informed allele frequencies with clinical databases and functional annotations can enhance our understanding of variant pathogenicity in ancestry-specific contexts, ultimately improving clinical decision-making.

To further illustrate this point, we've added new text to the manuscript: "Allele frequency data is used for more than just clinical interpretation and pathogenicity. Population allele frequency (AF) databases have recently been utilized to estimate aggregate carrier frequencies and genetic prevalence, providing insights into the potential number of affected individuals. LAI data enhances these estimations, enabling more precise predictions of populations with potentially higher genetic prevalence. For instance, the pathogenic variant 2-168970131-G-A (p.Arg575Ter) in the *ABCB11* gene, which causes Bile Salt Export Pump (BSEP) deficiency resulting in cholestasis, has a frequency of 0.04845% in the Admixed American genetic ancestry group, the highest observed across all gnomAD genetic ancestry groups. The cumulative carrier frequency of all likely pathogenic/pathogenic (LP/P) variants in *ABCB11* across genetic ancestry groups suggests that the 2-168970131-G-A variant leads to a slightly higher carrier frequency in this group when compared to the total estimate. However, analysis of LAI data

revealed that this variant was exclusively observed in the LAI-EUR regions of Admixed American samples. Consequently, the elevated carrier frequency, and thus the estimated genetic prevalence, primarily pertains to individuals of higher LAI-EUR ancestry within the broader genetic ancestry inferred label." (*Page 8, Lines 176-190*).

We believe this example highlights how LAI not only identifies substructure relevant to GWAS but also has direct utility for improving the accuracy of variant interpretation and disease risk estimation in diverse populations.

I can see the authors hope to reveal some allele frequency divergence trend between populations. However, the plots are mostly simple stacked barplots. It would be more informative if some population genetics theory informed discussions can be made.

We thank the reviewer for this thoughtful suggestion. Our primary goal in this work was to generate and release local ancestry-informed allele frequencies in gnomAD, providing a practical resource for the broader research and clinical genetics communities. While we do observe extensive allele frequency divergence across ancestry components, we intentionally refrained from making strong claims about the evolutionary origins of these differences, as we did not perform explicit analyses of demography, drift, or selection in this study.

We agree, however, that some additional population genetics framing can be helpful for contextualizing the observed patterns. In the revised manuscript, we have added a brief discussion noting that allele frequency differences across ancestry components may reflect a combination of drift, historical demographic events (e.g., population bottlenecks or founder effects), or local adaptation. Although we are not able to distinguish among these possibilities with the current data and analyses, we believe that future work building on this resource may be well-positioned to explore the evolutionary history of population-enriched variants more deeply. Our new text conveying this is reproduced here:

"Allele frequency differences observed across local ancestry backgrounds may arise from a combination of evolutionary and demographic forces, including genetic drift, founder effects, and natural selection. While we do not attempt to infer the specific historical or selective pressures underlying the observed patterns in this study, the extensive allele frequency divergence we report highlights the importance of modeling ancestry at finer scales. Future work leveraging these local ancestry-informed frequencies may provide insight into the history of population-enriched variants, particularly as more detailed demographic models and functional data become available." (*Page 14, Lines 315-321*).

To further illustrate the extent of allele frequency divergence across ancestry tracts, we directly quantified the variance in AFs for each tract in the Admixed American cohort. We computed site-level AF variances and observed a higher level of dispersion in the LAI-AMR component, as detailed below:

"Allele frequency variance was highest within the LAI-AMR tract (0.0447), compared to the LAI-AFR (0.0367) and LAI-EUR tracts (0.0356), suggesting greater heterogeneity among variants assigned to this ancestry. This elevated variance may reflect historical bottlenecks or founder effects that have shaped allele distributions in LAI-AMR haplotypes. Moreover, pairwise Levene's tests for equality of variances between ancestry tracts were all highly significant ($p < 1 \times 10^{-16}$)." (*Page 9, Lines 200-205*).

The authors claimed "The large sample size of gnomAD enables us to examine ancestral allele frequency trends at an unprecedented scale, and with the integration of LAI, at an unprecedented resolution." I would love to see this unprecedentedness be supported by some data, e.g., some before-after contrast analysis.

We have rephrased the language in the manuscript to clarify our intention to indicate that this is the first time LAI frequencies across global data on WGS data have been reported in this manner. Our revised language is: "The large sample size of gnomAD enables us to examine allele frequency trends on a more global scale, and with the integration of LAI, we achieve a finer resolution of ancestry-specific frequency estimates, improving the utility of the resource" (*Page 14, Lines 302-304*). This revision removes the term "unprecedented" while maintaining the intended emphasis on the novelty and scale of our analysis.

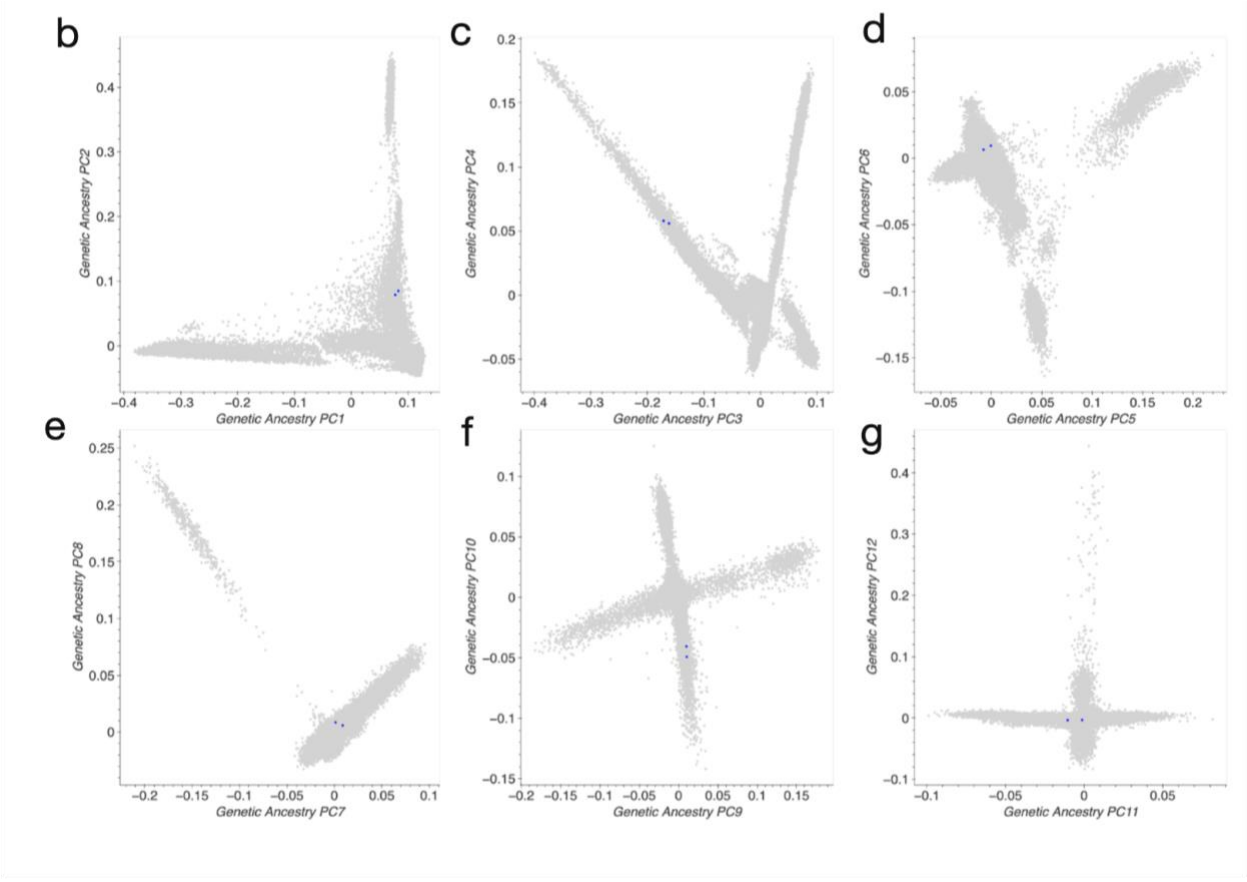
What is "full gnomAD collection" shown in Figure 1a? It is not surprising to see two individuals b and c with different local ancestry got collapsed in this PC plot, as the PC plot for the top 2 PCs are only able to show very crude levels of global ancestry. It might be interesting to show if the two are separable in a higher PC plots (e.g., PC7 vs PC8).

Thank you for bringing up this interesting point. The term "full gnomAD collection" refers to the complete set of samples from the current Genome Aggregation Database (gnomAD), which we used to contextualize genetic ancestry in our analysis.

We appreciate your suggestion to investigate whether individuals 'b' and 'c', who have different local ancestry yet appear overlapping in the initial PC plot (PC1 vs PC2), can be distinguished at higher principal components. To address this, we examined the location of these 2 samples in principal components beyond the ones showcased in the

main text (e.g., PC1 to PC12) and have included these plots in Supplementary Figure 2b-g. Interestingly, these individuals (highlighted with blue dots) remained in close proximity across higher-order PCs as well. This observation suggests that, despite differences in their local ancestry, their overall genetic profiles remain similar even when finer-scale genetic variation is considered. We have now clarified this point in the main text (*Page 7, Lines 143-146*) and reproduce these additional higher-PC figures below.

“Notably, even across higher principal components (e.g., PC1 to PC12), individuals 'b' and 'c' remain closely clustered, indicating that despite differences in their local ancestry, their overall genetic profiles are still similar when considering finer-scale genetic variation (**Supplementary Fig. 2b-g**).”



Minor:
Somehow Figure 1 panel b and c shown as blank.

Thank you for pointing this out. We have re-uploaded Figure 1 panels b and c, which display admixture plots of the Admixed American and African/African American genetic ancestry groups, in a different format to ensure they display correctly (*Page 22*).

Figure 2 has low density of information and can be probably more effectively shown as a table.

We thank you for your thoughtful feedback on Figure 2. We understand the concern about this figure being a low-density figure, but we believe that the visual comparison between global and local ancestry allele frequencies is crucial for effectively illustrating the impact of local ancestry inference. To address your suggestion while preserving the visual representation of our findings, we have expanded Figure 2 by:

1. Adding two additional rare variant examples to increase the amount of information presented.
2. Merging Extended Data Figure 1 into Figure 2 (as this reviewer suggests below) to create a more comprehensive multi-panel figure, which now includes both the LAI pipeline schematic and an illustrative example of local ancestry inference at the haplotype level.

These updates provide a broader and more informative visualization of how local ancestry inference refines allele frequency estimates compared to aggregate genetic ancestry groups. We have also revised the figure caption accordingly to reflect these additions (*Page 23*).

Figure 4: again low information density. Unclear what is AF_nfe.

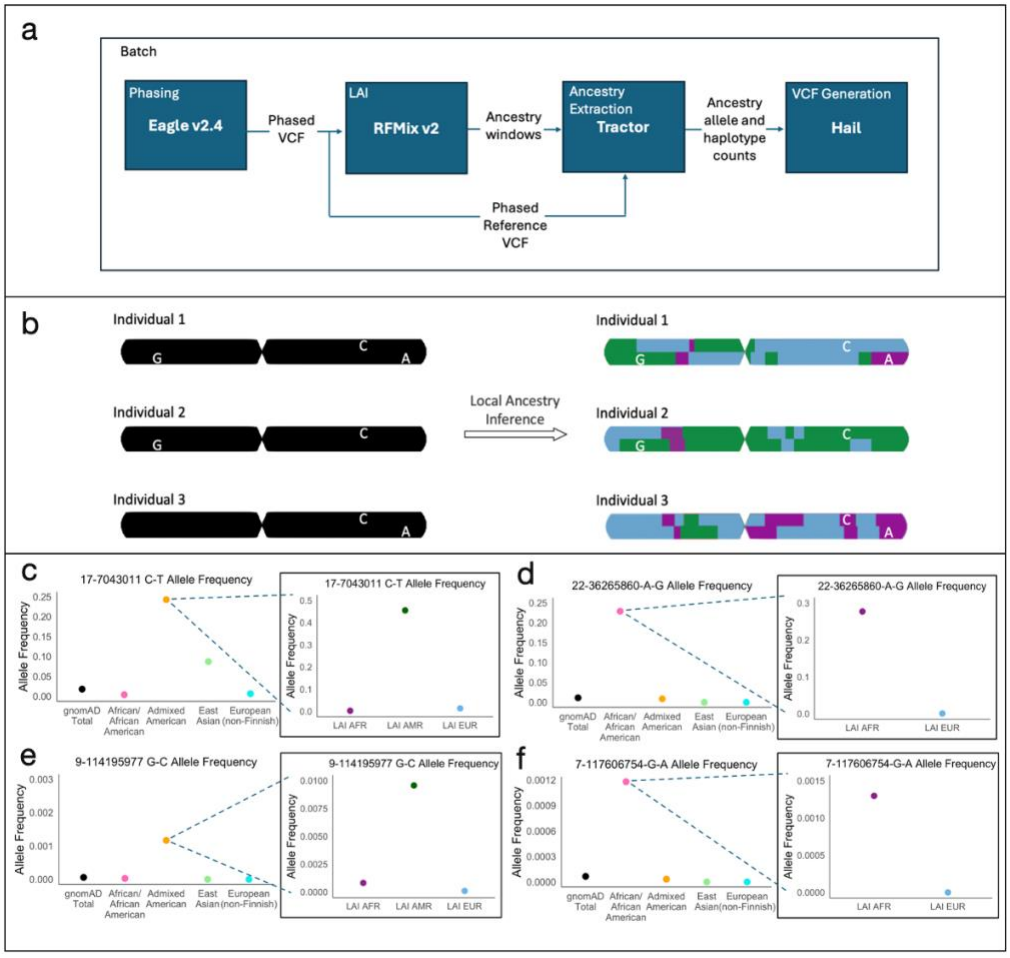
Thank you for your feedback. We agree that "AF_nfe" was unclear and have clarified our labeling throughout this revised manuscript for clarity and consistency:

- AF_nfe is now LAI-EUR (Local Ancestry–Inferred European AF)
- AF_afr is now LAI-AFR (Local Ancestry–Inferred African AF)
- AF_amr is now LAI-AMR (Local Ancestry–Inferred Amerindigenous AF)

We considered your point and believe the original eight-panel SFS figure allows readers to directly compare allele frequency distributions across ancestry tracts for each ClinVar category, highlighting important differences between ancestries. To further support these findings, we have included an additional figure showing the site frequency spectrum filtered by ClinVar variants with two or more review stars, categorized by ClinVar category (Supplementary Figure 4). This new supplementary analysis confirms similar allele frequency trends across each clinical significance category for highly confidently assigned variants.

Extended Data Figure 1 is actually more informative and can be merged with Figure 1 and 2 into a multi-panel figure.

We thank the reviewer for their thoughtful suggestion. We agree that merging Extended Data Figure 1 with Figure 2 provides a more comprehensive view of the data and better illustrates the derivation of local-ancestry-informed allele frequencies. To enhance this figure further, we have also added two additional examples of rare variants to complement our prior common variant examples, making it a more informative multi-panel figure. We have decided to keep Figure 1 as is, as it is already a dense figure that we feel effectively conveys the population structure of gnomAD. These changes help streamline the presentation of our results while maintaining consistency in their visual representation. We appreciate your feedback and believe these changes strengthen the manuscript.



Reviewer #1 (Remarks on code availability):

The code is basically a dump of their own pipeline. No attempt to making a public-facing README.md. Also, their pipeline is presented in a way that is only useful to re-produce their LAI calls. Not even sufficient to reproduce the figures in the paper. Not to mention to be a useful tool to the community for other purposes.

We thank the reviewer for their useful suggestion on creating a comprehensive README. We have made significant improvements and expansions to our README to ensure it is more clear and accessible for the broader research community. The updated README now provides step-by-step instructions for running our Local Ancestry Inference (LAI) pipeline, along with a structured overview of the pipeline workflow. It includes detailed guidance on installation and setup, explicit command-line instructions for each stage of the pipeline, and links to relevant resources such as the Tractor wiki for additional context. These updates make it easier for users to follow and apply the pipeline to their own datasets. You can find the updated README here: https://github.com/broadinstitute/gnomad_local_ancestry. (*Page 29, Lines 587-588*).

Reviewer #2 (Remarks to the Author):

Paper summary: The authors analyse two gnomAD genetic ancestry groups with high degrees of recent admixture - “Admixed American” (n=7,612) and “African/African American” (n=20,250). By conducting local ancestry inference, they derive more precise ancestry-specific allele frequencies of genomic variants. The results of this work are made publicly available through gnomAD and will potentially further our understanding of human genomics.

We thank the reviewer for their appreciation of our work and their thoughtful suggestions that have helped to improve our manuscript. In response, we expanded the description of the two admixed populations, clarified the data sources and QC criteria for LAI variants, and revised figures to improve consistency and readability. We also addressed ClinVar-related concerns by applying stricter quality filters to all variants and additionally conducting a stratified analysis of only 2-star and above variants, adding site frequency spectrum comparisons, and performing statistical tests to examine allele frequency distributions. We believe these revisions substantially strengthen our conclusions and provide clearer support for the value of ancestry-informed analyses – thank you again for your careful and thoughtful review.

Below are my comments.

Introduction

Probably need more detailed description of the two admixed populations (“Admixed American” and “African/African American”) in the Introduction. There is some information in the first paragraph in the Methods, but it needs to be expanded and brought forward.

We thank the reviewer for this helpful suggestion and we agree that this context is important. In response, we have expanded the start of the Methods to provide a brief overview of the genetic composition of the two gnomAD genetic ancestry groups analyzed in this study: African/African American and Admixed American. The new text, reproduced below, describes the general ancestral backgrounds of these groups, shaped by recent admixture involving African, European, and Indigenous American source populations, and highlights the relevance of applying local ancestry inference in these contexts. We additionally have clarified the implementation of the ancestry classifier used in gnomAD.

Our added text (*Page 27, Line 534-545*): “In this work, we focus on two admixed genetic ancestry groups represented in gnomAD: the African/African American and Admixed American groups. The underlying composition of the genomes of individuals from these groups contain complex mosaics of ancestries shaped by recent admixture events involving African, European, and Indigenous American sources. African/African American individuals typically carry a majority of West African ancestry, with variable proportions of European and, to a lesser extent, Indigenous American ancestry. Admixed American individuals generally derive their genomes from a mixture of Indigenous American, European (primarily Iberian), and African ancestries, with substantial regional and individual-level variation. These groups have historically been underrepresented in genomic studies, and their complex ancestry patterns make them particularly well-suited for local ancestry-informed analyses.”

Results

“Here we release local ancestry-informed frequency data for 14,804,207 bi-allelic single nucleotide polymorphisms (SNPs) within the Admixed American (n=7,612) and 24,204,574 bi-allelic SNPs within the African/African American genetic ancestry groups (n=20,250), both from gnomAD v3.1.” (lines 120-123)

This statement suggests to me that the variants reported are derived from WGS data, is that correct? Maybe worth stating it explicitly in the manuscript. For which variants local ancestry data is not available and why? What proportion of the human genome was successfully LAI resolved for “Admixed American” and “African/African American” groups; what proportion of the variants identified for “Admixed American” and “African/African American” groups have LAI data?

We appreciate your suggestion regarding explicitly stating the data type and providing further clarification on the availability of local ancestry inference (LAI) data. You are correct that the LAI data reported in our study is derived from whole genome sequencing data. To clarify this in the manuscript, we have revised the methods section to explicitly state: “The 7,612 Admixed American and 20,250 African/African American whole genome sequencing samples from gnomAD v3.1 used here include all individuals in the database except those from the 1000 Genomes Project (1kGP) and Human Genome Diversity Project (HGDP) joint callset, which were used to curate reference panels for LAI in each genetic ancestry group.” (*Page 26, Lines 521-525*).

Additionally, we have clarified the criteria that determine whether a variant has LAI data. Once samples were subset based on inferred genetic ancestry groups, our QC process involved limiting to 1) bi-allelic variants with low genotype missingness (i.e., call rate > 0.9), 2) an allele frequency greater than 0.1% within the Admixed American and African/African American gnomAD genetic ancestry groups, and 3) a minimum allele count (AC) cutoff of 7 within each genetic ancestry group to ensure that frequency estimates are more reliable.

To address the your request for the proportion of variants with LAI data, we now report the following statistics here:

- For the African/African American group, there are 290,938,182 variants with at least one alternate allele observed (AFR AC >0) in the raw dataset, of which 246,557,024 are SNPs. After applying our QC threshold (AC > 7), 56,159,957 SNPs remain. After limiting to bi-alleleic variants with low genotype missingness and an allele frequency of greater than 0.1%, 24,204,574 SNPs remain and now have local ancestry information. Overall, 9.82% of all SNPs have LAI data.
- For the Admixed American group, there are 153,254,643 variants with at least one alternate allele observed (AMR AC >0) in the raw dataset, of which 127,940,698 are SNPs. After applying our QC threshold (AC > 7), 43,385,434 SNPs remain. After limiting to bi-alleleic variants with low genotype missingness and an allele frequency of greater than 0.1%, 14,804,206 SNPs remain and now have local ancestry information. Overall, 11.57% of all SNPs have LAI data.

These numbers reflect the high proportion of rare variation in gnomAD, which is why only a subset of variants meets the criteria for LAI annotation. To better clarify this in the methods section, we have now added the phrase "Additionally, an allele count (AC) threshold of AC > 7 was applied for each genetic ancestry group." (*Page 26, Lines 520-521*). We appreciate the reviewer’s request to quantify the proportion of the genome successfully resolved with LAI. Given that our analysis focuses on SNPs rather than the

entire genome, we have now added these proportions relative to total SNPs in the results section: “Of the total SNPs in each genetic ancestry group, 9.82% of SNPs in the African/African American group and 11.57% of SNPs in the Admixed American group passed QC and were assigned local ancestry calls.”(*Page 5, Lines 108-110*).

The term “Amerindigenous” would also benefit from expanded description at first reference (line 136, and in Fig1) – only in Methods they are described as the harmonized “AMR” samples from HGDP and 1kGP3.

We appreciate the reviewer’s suggestion to clarify our use of the term “Amerindigenous.” In response, we have revised the manuscript to provide a more explicit description at its first mention. Specifically, we now clarify that we use “Amerindigenous” to refer to local ancestry tracts (and the ancestry-specific frequencies derived from said tracts) inferred to originate from Indigenous American populations, based on reference haplotypes harmonized from the continental American-designated samples in the Human Genome Diversity Project (HGDP) and 1000 Genomes Project (1kGP3). This is in contrast to the use in some prior work of the term ‘AMR’ to designate a continental label. We hope this difference is now clearer in our manuscript text.

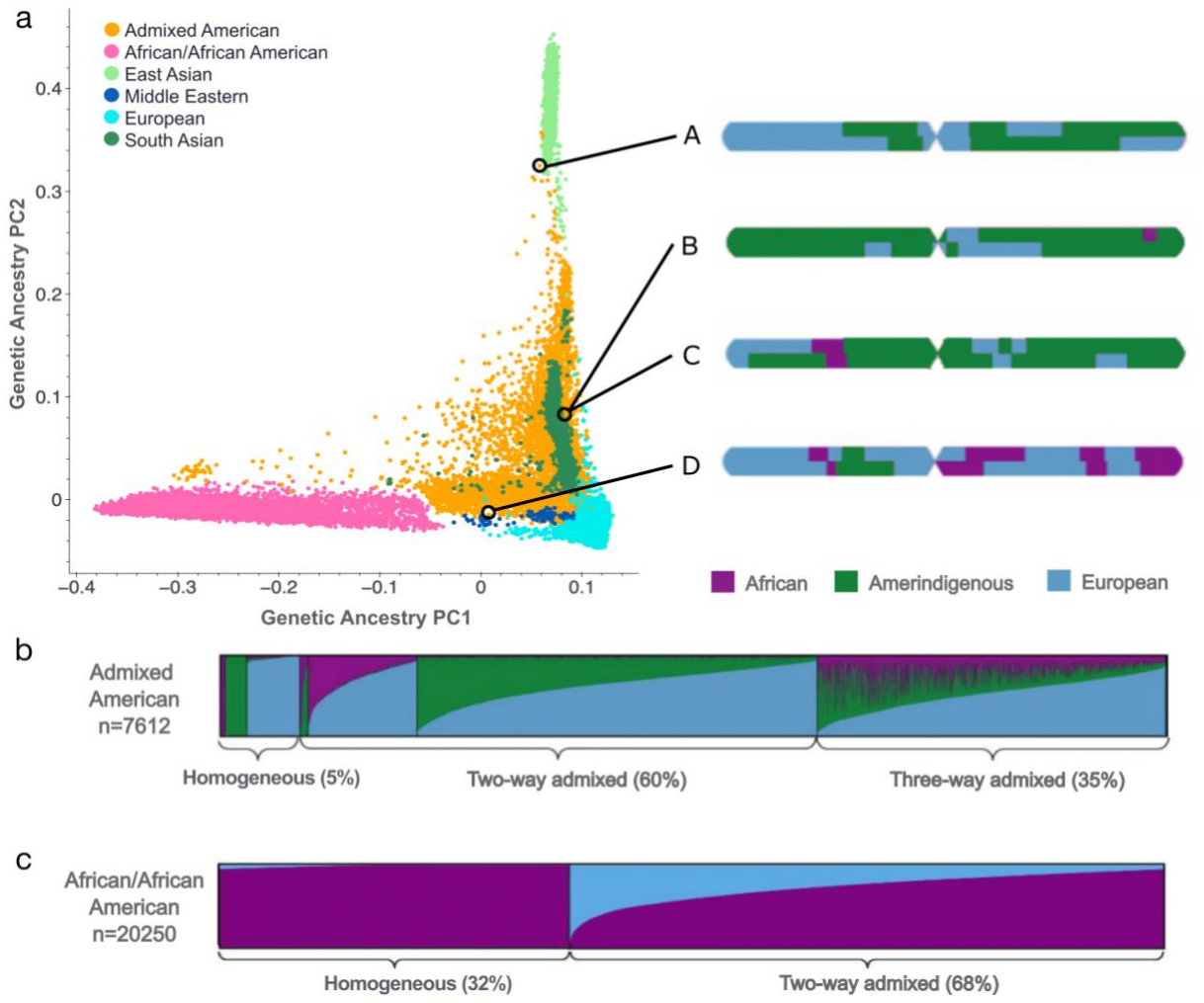
In addition, we have worked throughout the manuscript to clarify the distinction between genetic ancestry groups and local ancestry tracts. We now consistently use separate terminology for frequencies derived from gnomAD’s genetic ancestry groupings (African/African American and Admixed American, assigned at the individual level) versus those derived from painted local ancestry segments within those individuals (now referred to as LAI-AFR, LAI-EUR, and LAI-AMR). We believe this revision improves the clarity of the manuscript and avoids potential confusion in interpreting the ancestry-specific frequency data.

Our added text (*Page 4, Lines 85-99*): ”We wish to be clear that in all analyses we exclusively examine genetic ancestry. Assignment of an individual’s genetic ancestry is not equivalent to and does not negate how an individual self-identifies. To refer to the samples included in these analyses, we adhere to the current nomenclature in gnomAD, namely “Admixed American” and “African/African American”. We acknowledge that this terminology is not ideal, however for consistency with broader efforts utilizing these data, and to be clear on the individuals included in these analyses, we retain it here. We look forward to continuing to further refine gnomAD allele frequency presentations in future releases to better reflect the continuous nature of genetic ancestry. Throughout this article, we use the term “Amerindigenous” to refer to local ancestry tracts inferred to originate from Indigenous American populations, based on reference haplotypes harmonized from AMR-designated samples in the Human Genome Diversity Project (HGDP) and 1000 Genomes Project Phase 3 (1kGP3). To distinguish between

individual-level genetic ancestry groupings (e.g., Admixed American) and ancestry-specific segments within individuals, we refer to frequencies derived from painted tracts as LAI-AMR (Amerindigenous), LAI-AFR (African), and LAI-EUR (European), respectively."

In Figure 1a, I cannot see the "African reference samples" (line 134), only "African/African American"? In an attempt to derive a more eye-friendly colour scheme for Figure 1a, I would also suggest creating a "simplified" Figure 1a, containing only "relevant" populations – i.e., "Admixed American", "African/African American", "African reference", "East Asian", "South Asian" and "European" and move the current version to the Supplementary Figures. For consistency, please use either "Admixed American" or "Latino/Admixed American" (Figure 1a, lines 160 and 273, Supp Figure 3).

Thank you for your insightful comments regarding Figure 1a. In response to your suggestions, we have revised Figure 1a to simplify it and highlight only the relevant genetic ancestry groups (*Page 22*), also pasted below. The revised figure now includes the following populations: "Admixed American", "African/African American", "East Asian", "Middle Eastern", "European", and "South Asian". Additionally, for consistency throughout the manuscript, we have standardized the terminology to "Admixed American". The original Figure 1a, the PCA plot of all gnomAD samples which is consistent with past gnomAD color schemes, has been moved to the supplemental (Supplementary Figure 2a) as suggested. We believe these modifications substantially improve the clarity and readability of the figure.



Line 141: should not “homogenous AMR” be “homogenous AFR or AMR or EUR”?

We thank the reviewer for catching this inconsistent phrasing. Our revised language is: “After applying a 5% ancestry inclusion threshold, 5% of individuals derive their genetic ancestry primarily from a single continent (homogeneous AMR or EUR or AFR), 60% from two continents (AFR/AMR or AFR/EUR or AMR/EUR), and 35% from three continents (AMR/EUR/AFR) (**Fig. 1b**).” (*Page 6, Lines 128-132*)

Figure 2: Maybe the right-hand side legends in each of the plots are not necessary; removing them would allow for larger font size to be used for the axis labels.

We completely agree with your suggestion and have removed the right-hand side legends from each plot in Figure 2. This adjustment allows us to use a larger font size for the axis labels, improving readability. Additionally, we have merged Extended Data

Figure 1 with Figure 2 to create a more comprehensive representation of the data, better highlighting the derivation of local-ancestry-informed allele frequencies. To further develop this figure, we have also added two additional rare variant examples, making the figure a more informative multi-panel visualization.

Figure 2a: Should not the reference for the variant in the SLC16A11 gene be to 14, rather than 16? Am I correct that for the zoom-in plot (and line 163), “LAI Amerindigenous” would be more appropriate/consistent label than “LAI Continental American”? These are the only 2 places where the term “LAI Continental American” is used.

We appreciate your attention to detail and your helpful suggestions. You are correct that the reference for the variant in the *SLC16A11* gene should be to 14 instead of 16 (currently labeled as 11), and we have updated it accordingly. We also agree that "LAI Amerindigenous" is the more appropriate and consistent label. We have relabeled the zoom-in plot (Figure 2c-f) and updated the term used in line 163 to "LAI AMR", which serves as a shorthand for "LAI Amerindigenous" to maintain consistency throughout the manuscript.

Figure 2b: Should not the reference for the variant in the SLC16A11 gene be to 15 or 16, rather than 17? Please check the references throughout the manuscript.

Thank you for pointing this out. You are correct that the reference for the current Figure 2d variant in the *APOL1* gene should be 15 (currently labeled as 13), and we have updated it accordingly. We have also carefully double-checked all references throughout the manuscript to ensure accuracy. We appreciate your thorough review.

Lines 170-177: Providing a link to the gnomAD page for this variant would be illustrative.

We thank the reviewer for this suggestion and completely agree that including the exact genomic variant or a direct link to the gnomAD page for this variant (https://gnomad.broadinstitute.org/variant/9-114195977-G-C?dataset=gnomad_r4) would be helpful for readers. We have now explicitly referenced the variant of interest in the text : “Within gnomAD, the causal variant 9-114195977-G-C has an allele frequency of 0.1% in the Admixed American genetic ancestry group, while gnomAD’s gnomAD-wide allele frequency is 0.006% (**Fig. 2e**).” (*Page 8, Lines 165-167*).

Lines 177-182: Using the online version of gnomAD v3.1.2 shows from the “gnomAD tab”: Total AF = 0.0002768 and African/African American AF = 0.001018. From the “Local Ancestry” tab: Total/Overall AF = 0.001043 (due to the slightly smaller AN in Local Ancestry compared to gnomAD), LAI African AF = 0.001297 and LAI European AF = 0.000 (with AC = 0). Can the discrepancy with the reported “0.007% gnomAD-wide (“global”) frequency (line 180) be explained with the different gnomAD versions – in the manuscript v3.1, while currently available online is gnomAD v3.1.2?

We thank the reviewer for pointing out the allele frequency (AF) discrepancy across gnomAD versions. After giving it some more thought, we’ve decided to report the total AF and genetic ancestry group AFs from gnomAD v4.1.0 for all example variants. Since most gnomAD users will be referencing the most up-to-date dataset, we felt this was the best choice for consistency and accessibility. We agree that citing genomes AF from v3.1.2 would allow for a more direct comparison, but we wanted to ensure we were being as inclusive as possible by incorporating the expanded diversity and improved population representation in v4.1.0. We’ve also clarified this in the Methods section, noting that while LAI data itself doesn’t change between versions, the aggregate AFs (total and by genetic ancestry group) have improved with v4.1.0 due to the inclusion of additional samples.

Our added text is: “To ensure consistency with the most up-to-date dataset available to users, we report total allele frequencies and genetic ancestry group AFs from gnomAD v4.1.0 for all example variants. While LAI calls remain unchanged between gnomAD versions, the aggregate allele frequencies (both total and stratified by genetic ancestry group) have been updated in v4.1.0 due to the inclusion of additional diverse samples.” (*Page 28, Lines 572-576*).

For the specific variant example you mentioned (7-117606754-G-A), the “gnomAD tab”: Total AF = 0.00006622 and African/African American AF = 0.001179 in gnomAD v4.1.0. The “Local Ancestry” tab remains the same: Total/Overall AF = 0.001043 (due to the slightly smaller AN in Local Ancestry compared to gnomAD), LAI African AF = 0.001297 and LAI European AF = 0.000 (with AC = 0).

Figure 3: I would suggest using consistent labelling throughout the manuscript to distinguish more easily between “ancestry-specific allele frequencies” derived in this work (e.g., “LAI EUR”, “LAI AFR”, “LAI Amerindigenous”) and reference population allele frequency as reported in gnomAD (e.g., “gnomAD EUR”, etc.). Also, from the main text (and the figure title) I understand that Figure 3 depicts data about variants with LAI frequencies at least 2 times more frequent in one of the constituent populations compared to the LAI frequencies in other constituent populations, but the use of the

word “different” in the figure text is confusing (it suggests also at least 2 times less frequent).

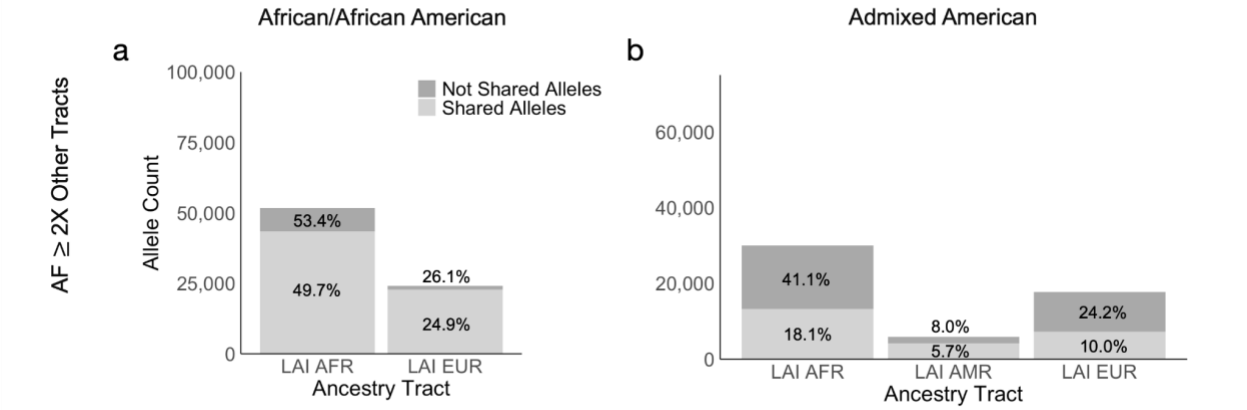
We appreciate your thoughtful suggestions and agree that maintaining consistent labeling throughout the manuscript will improve clarity. We have revised the manuscript to distinguish ancestry-specific allele frequencies derived in this work (e.g., "LAI-EUR", "LAI-AFR", "LAI-AMR") from reference population allele frequencies reported in gnomAD. We have kept the original gnomAD genetic ancestry group terms to stay consistent with previous gnomAD nomenclature (e.g., "African/African American", “Admixed American”, “East Asian”, “European (non-Finnish)”etc.).

Additionally, we appreciate your feedback on the wording in Figure 3, as the use of the term "different" in the figure caption may be misleading. To ensure clarity and added analyses, we revised the figure text as follows: "**Figure 3: Distribution of variants with enriched or depleted ancestry tract allele frequencies across genetic ancestry groups.** The total percentage for each local-ancestry background is indicated on the bars. **a,b.** Variants with ancestry-specific allele frequencies at least 2X higher than the other tracts within the African/African American and Admixed American groups, respectively. **c.d.** Variants with ancestry-specific allele frequencies at most 0.5X lower than the other tracts within the African/African American and Admixed American groups, respectively." (*Page 24*). These changes aim to accurately convey the intended meaning.

I have several comments related to Figure 4 and the corresponding text (lines 200-208). First, for what proportion of the LAI variants discovered in the “Admixed American” and “African/African American” populations a ClinVar annotation is available?

Thank you for your thoughtful question regarding the proportion of LAI variants with ClinVar annotations in the “Admixed American” and “African/African American” populations. For the Admixed American population, of the 14,804,206 SNPs evaluated through local ancestry inference, 131,886 had annotations in ClinVar, corresponding to approximately 0.9%. For the African/African American population, 183,164 of 24,204,574 LAI SNPs had ClinVar annotations, corresponding to approximately 0.8%. When applying a more stringent quality control—retaining only those annotations with 2 or more ClinVar stars—the counts are reduced to 73,087 for the Admixed American population and 96,851 for the African/African American population. We additionally examined the number of variants with a tract-specific allele frequency at least twice as high as the other ancestral tract frequencies: the African/African American group has 75,756 variants when including all shared alleles (66,100 variants when considering only monomorphic variants), and the Admixed American group has 53,588 variants when including all shared alleles (24,683 variants when considering only monomorphic

variants). We display these additional ClinVar 2+ star-filtered results in Supplementary Figure 5 - the trends are similar with Figure 3 a and b when looking at all variants.

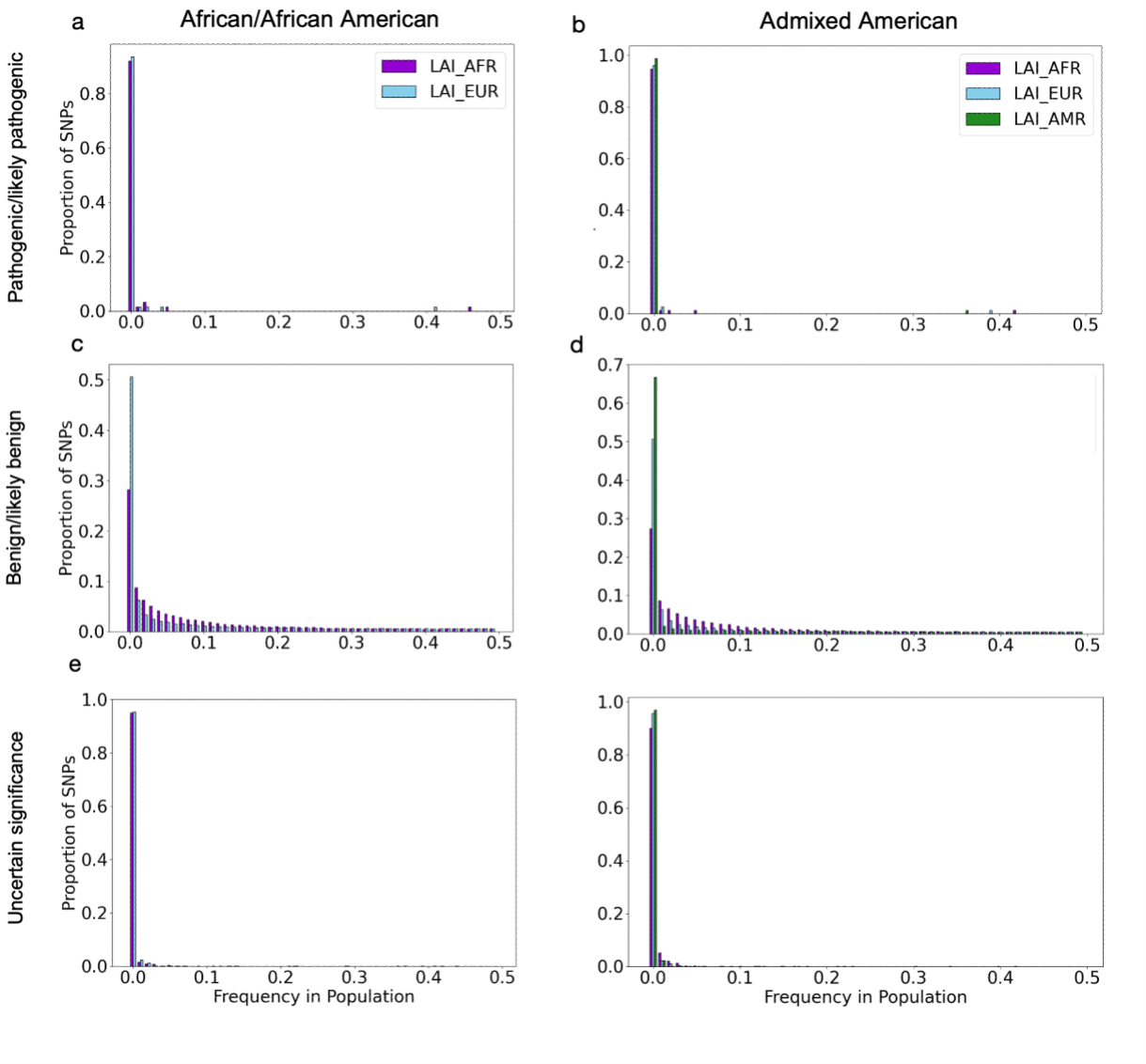


We added text that details these numbers: "Of the LAI variants, only a small fraction have corresponding ClinVar annotations: 131,886 out of 14,804,206 (0.9%) in the Admixed American and 183,164 out of 24,204,574 (0.8%) in the African/African American genetic ancestry groups. Because single-submitter and no-criteria entries may be less reliable, we applied an additional QC filter, retaining only annotations with ≥ 2 ClinVar stars, resulting in variant counts of 73,087 and 96,851, respectively. We additionally examined the number of variants with a tract-specific allele frequency at least twice as high as the other ancestral tract frequencies: the African/African American group has 75,756 variants when including all shared alleles (66,100 variants when considering only monomorphic variants), and the Admixed American group has 53,588 variants when including all shared alleles (24,683 variants when considering only monomorphic variants)(**Supplementary Fig. 5)**." (*Page 10, Lines 225-235*).

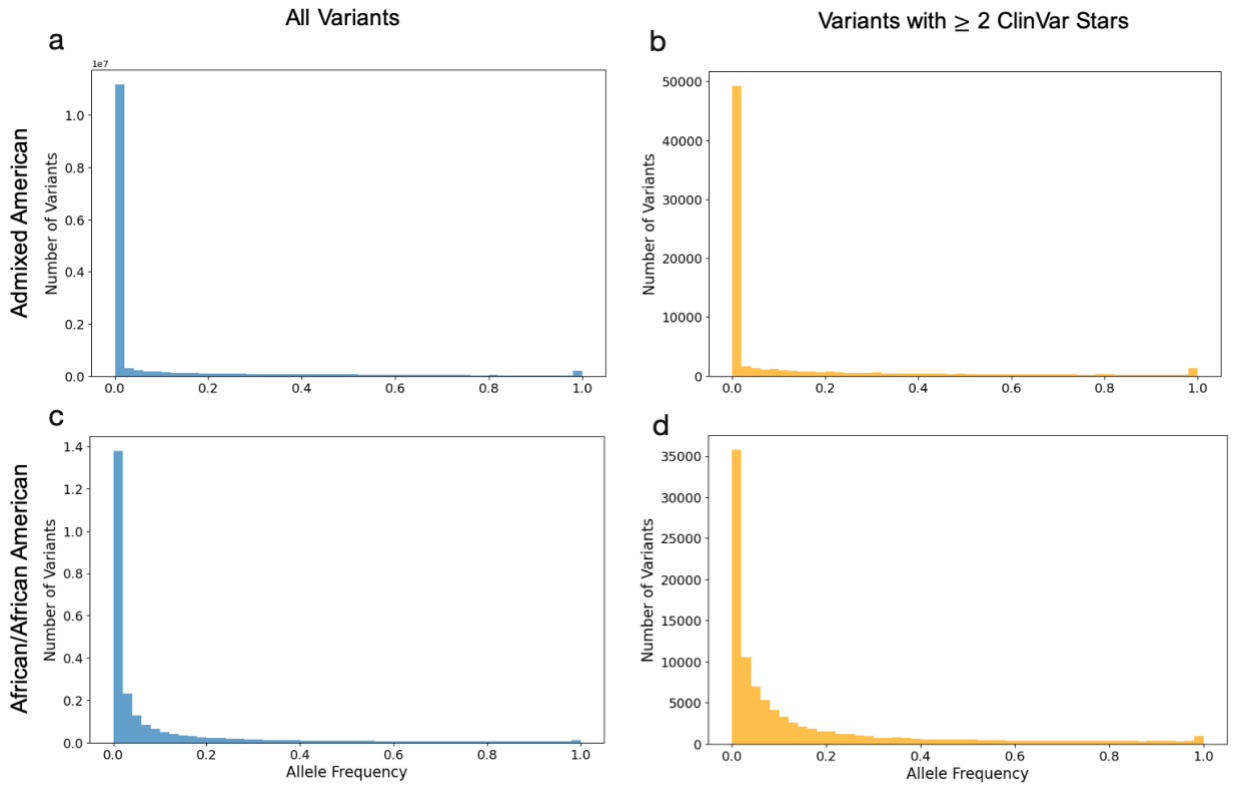
Second, the statement “The results indicate that most variants across all ClinVar categories, including pathogenic/likely pathogenic, benign/likely benign, uncertain significance, and conflicting classifications of pathogenicity are rare, with AFs concentrated between 0 and 0.1 in all genetic ancestry groups” should not be made without providing the site frequency spectrum distributions considering *all* variants discovered in these two cohorts and contrasting to the variants with ClinVar annotation only, thus demonstrating different frequency distributions.

Thank you for the useful suggestion – we agree that our statement is best supported by contrasting the site frequency spectrum (SFS) distributions for the entire dataset against the subset of ClinVar-annotated variants (particularly those passing a more thorough quality control). We have now included six new SFS plots as Supplementary Figure 4 (pasted below) for the Admixed American and African/African American genetic

ancestry group. This subset of variants have ClinVar annotations, which are further filtered to 2+ ClinVar stars (indicating higher confidence).



We also added a Supplementary Figure 6 (pasted below), which demonstrates that the SFS of ClinVar-annotated variants differs from the overall variant distribution, and that most variants—irrespective of whether they are ClinVar-annotated—indeed exhibit lower allele frequencies. We believe these comparisons strengthen our conclusion regarding the rarity of ClinVar variants in both populations while simultaneously highlighting the different frequency distributions for clinically annotated versus unannotated variants.



Third, while ClinVar is an invaluable resource, it should be used with care. The authors do not specify any criteria for including/excluding a ClinVar annotation. If no such criteria were applied, it can be expected that a significant proportion of the ClinVar annotations are of questionable quality (e.g., those with single submitter, no criteria provided, etc.). Furthermore, even a robust “benign/likely benign” ClinVar annotation of a variant for condition A does not imply much for the same variant in the context of another condition B.

We thank the author for their valid concern. ClinVar is an invaluable, but heterogeneous, resource. You are correct that there are ClinVar entries that come from a single submitter or lack rigorous curation (e.g., missing submission criteria), which can reduce their reliability. For example, annotations based on a single submitter or annotated as “not_provided” may be of questionable quality, and some annotations include ambiguous terms that do not consistently inform a variant’s clinical significance across different conditions. Thus, in our first submission ClinVar analyses, we excluded variants with the following clinical significance tags:

- Affects
- Affects|risk_factor
- Likely_risk_allele

- Association
- Association_not_found
- Association|drug_response|risk_factor
- Confers_sensitivity
- Confers_sensitivity|other
- Drug_response
- No_classification_for_the_single_variant
- Not_provided
- Other
- Protective
- Protective|risk_factor
- Risk_factor

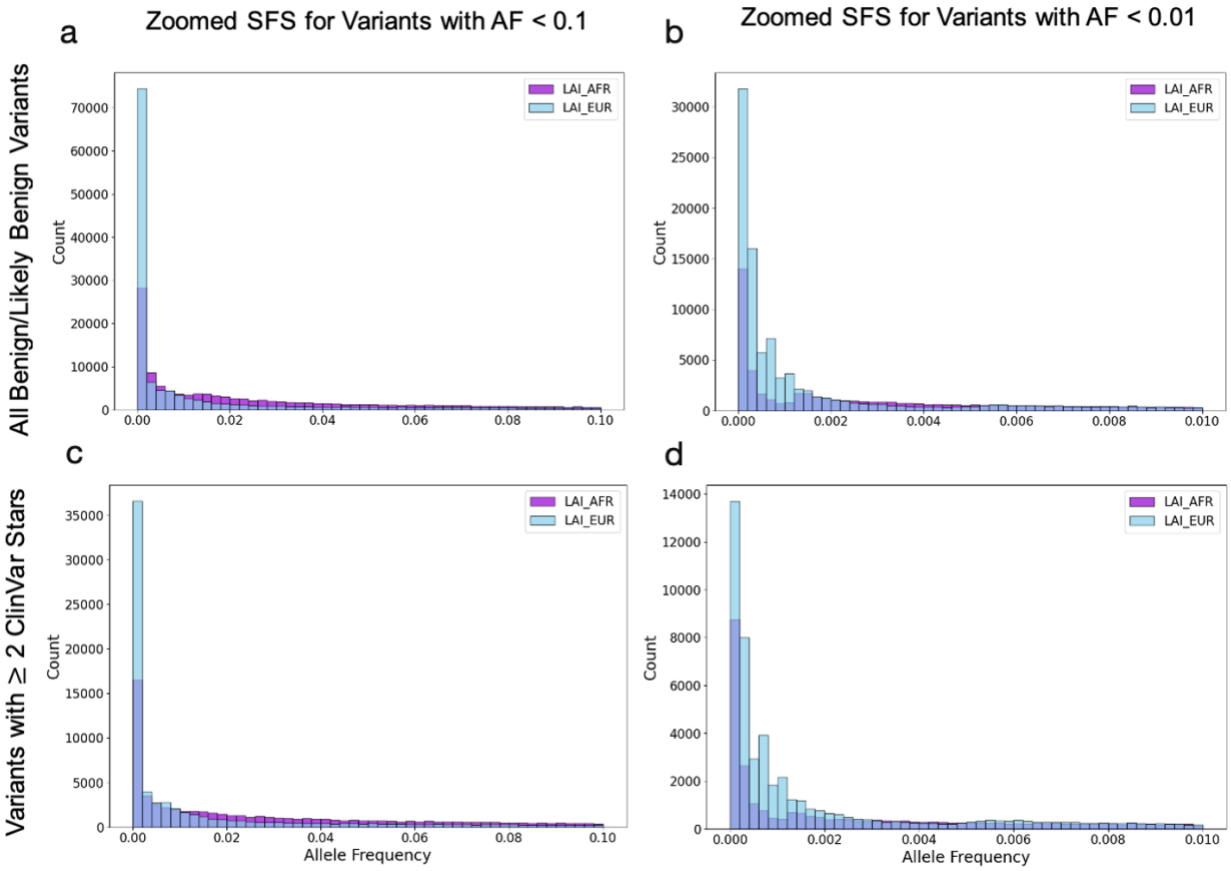
However, to further mitigate these issues, we have now applied a stricter QC filter that only retains ClinVar annotations with 2 or more stars, a commonly used threshold to indicate higher confidence (e.g., multiple independent submissions or detailed evaluation criteria). This filtering reduces the noise from lower-quality annotations and allows us to make more confident inferences regarding the distribution of clinically annotated variants. For instance, while the overall number of ClinVar-annotated variants in the Admixed American group is 131,886, applying the 2+ star criterion yields 73,087 high-confidence SNPs; similarly, in the African/African American group, the number is reduced from 183,164 to 96,851. We now report these results in the revised manuscript: “Because single-submitter and no-criteria entries may be less reliable, we applied an additional QC filter, retaining only annotations with ≥ 2 ClinVar stars, resulting in variant counts of 73,087 and 96,851, respectively.” (*Page 10, Lines 227-230*).

Fourth, I completely agree with the need for more diversity in genomic studies. However, I am not convinced that “The broader distribution of benign variants in African/African American groups ….” (line 205) is not, at least to some degree, a result of variants being labelled as “benign/likely benign” in ClinVar simply due to their higher AF in gnomAD reference populations (e.g., for rare disorders dominant variants with AF > 0.0001 and recessive variants with AF > 0.005 are usually excluded from further analysis and follow-up). Maybe add zoom-in to variants with AF < 0.01/0.1?

Thank you for your suggestion. We have now generated zoomed-in SFS plots for benign variants and those variants with 2 or more ClinVar stars with AF below 0.1 (Supplementary Figure 7 a and c) and below 0.01 (Supplementary Figure 7 b and d). These histograms (see below) compare the LAI_AFR and LAI_EUR allele frequency distributions in the African/African American cohort. Consistent with our broader observations, the LAI_AFR distribution remains shifted relative to LAI_EUR even in these lower-frequency ranges. Additionally, to address the possibility that benign

variants in African/African American groups are labeled as “benign/likely benign” primarily due to higher allele frequencies, we performed a two-sample Kolmogorov–Smirnov test specifically for benign variants (i.e., labeled “Benign/Likely_benign”) in LAI_AFR versus LAI_EUR:

- **Full set of benign variants**: KS statistic = 0.278 (p = 10^-100), strongly rejecting the null hypothesis of identical distributions.
- **Subset of benign variants with 2+ ClinVar stars**: KS statistic = 0.225 (p = 10^-100), again indicating a significant difference in allele frequency distributions.

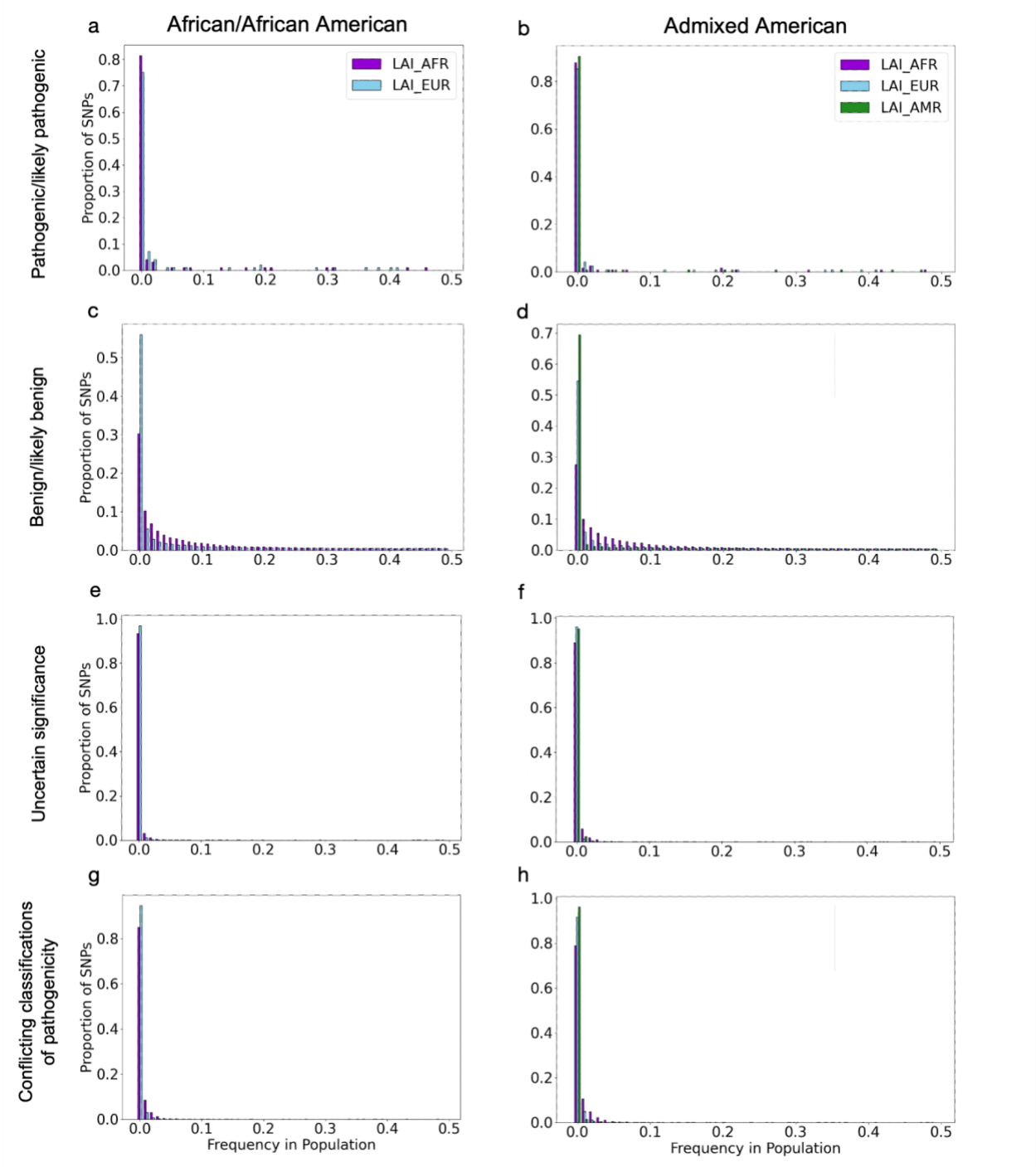


We also added text to refer to this figure: “To further investigate if benign classifications may be driven primarily by frequency thresholds, we zoomed-in on benign variants and again on benign high-confidence ClinVar variants (≥2 stars) with allele frequency <0.1 and <0.01 (**Supplementary Fig. 7**). We conducted a two-sample Kolmogorov–Smirnov test on benign/likely benign ClinVar variants in the African/African American group, comparing the SFS between LAI-AFR and LAI-EUR. The test yielded a KS statistic of 0.278 with a p-value near zero (p = 10^-100), indicating a significant difference between

the two distributions." (*Page 11, Lines 238-245*). These findings demonstrate that the broader distribution of benign variants in the African/African American cohort is not solely driven by ClinVar's frequency-based filtering, as the shift is still somewhat consistently observed even when we focus on high-confidence ClinVar annotations and/or examine lower allele frequency ranges.

For consistency, I would suggest to use LAI_AFR, LAI_EUR and LAI_Amerindigenous for the labels in Figure 4; could you please improve the readability of Figure 4 by increasing the font size for example?

We thank the reviewer for the valuable suggestions regarding the labeling and readability of Figure 4. In the revised manuscript, we have updated the population labels for consistency so that they now read **LAI-AFR, LAI-EUR,** and **LAI-AMR.** In addition, we have increased the font sizes of the title, axis labels, tick labels, and legend to improve the overall readability of the figure. These modifications indeed enhance the visual presentation of Figure 4 and also ensure consistency with the nomenclature used throughout our work (pasted below, and also on *Page 25*).



Lines 210-224: I find this paragraph difficult to follow. Given that it describes one of the major contributions of this work, **clearer rephrasing would be beneficial.** Please define grpmax and popmax (grpmax first mentioned in the Abstract, line 46). My understanding of what is described is as follows. Performing a LAI based analysis of the two recently admixed cohorts results in higher values for the grpmax metric for the vast

majority (81.49%) of the variants found in them, compared to the same metric computed based on other (non-bottlenecked) populations available in gnomAD. These updated grpmax values in turn will allow for correctly excluding more non-disease-causing variants when assessing an individual from the underrepresented “Admixed American” and “African/African American” groups, supported by the observation that 87.69% and 89.32% of the variants in the two cohorts with higher corrected grpmax values are found in ClinVar labelled as “benign / likely benign” (with the ClinVar use caveat as discussed above). Is my understanding correct?

We appreciate your detailed feedback and have rephrased the manuscript to define grpmax more clearly. Our revised language is: “We additionally investigated the frequency of ancestry-enriched AFs exceeding those of grpmax (formerly popmax), which measures the highest allele frequency observed in any gnomAD genetic ancestry group. As grpmax is often used to inform a variant classification of benign/likely benign, better resolving allele frequencies across ancestral haplotypes with LAI provides a more precise estimate of frequencies that may change clinical interpretation.” (*Page 11, Lines 250-254*). Your understanding is correct—grpmax, derived from local ancestry inference, replaces popmax as a more precise metric based on ancestry rather than self-reported populations. This updated metric enables better exclusion of non-disease-causing variants in underrepresented groups, as demonstrated by the ClinVar benign/likely benign findings (87.69% and 89.32%).

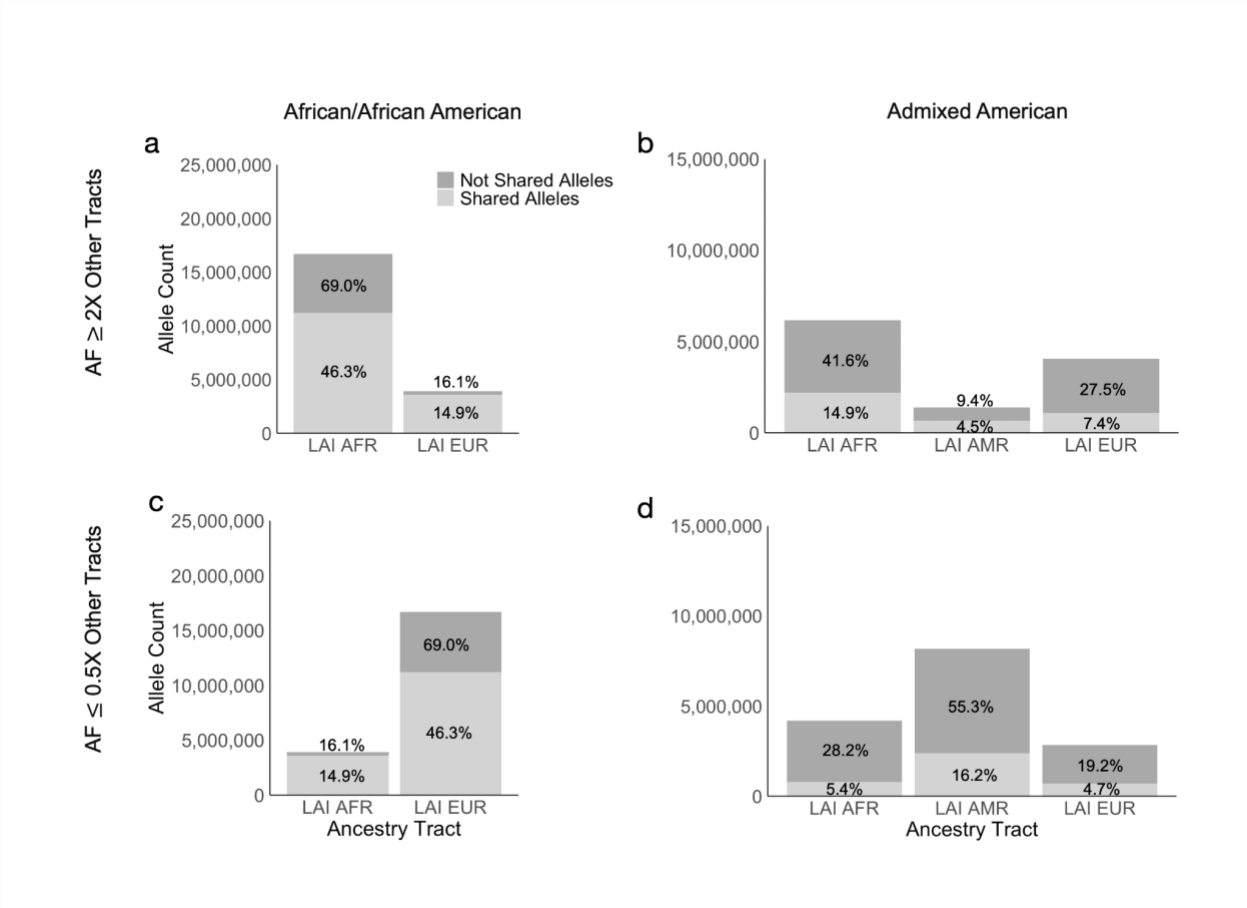
Discussion

Investigating the reverse question – are there any variants, which based on the conducted LAI analysis of the two cohorts, should not be overlooked/discarded when assessing their pathogenicity? - may provide another important contribution of this work. This is somehow described (lines 268-277), focusing on a single example of a variant associated with type 2 diabetes risk, with global AF of 2%, but much more frequent in Admixed American population, driven by its frequency in Amerindigenous individuals. Are there any other such variants? In these two groups, are there any otherwise generally rare protein-coding variants with predicted functional impact (i.e., nonsense/missense) and potentially associated with a rare/common disorder, which exhibit much higher LAI derived AF in one of their constituent ancestries? Given that the sets of variants enriched at least twice in one LAI ancestry compared to the other(s) for the two groups is available (Figure 3), the suggested systematic analysis could be relatively easily performed using some carefully designed filtering steps.

Thank you for your insightful suggestion. We agree that systematically identifying variants whose ancestry-informed AFs would have been “overlooked” by global AF

filters is an important complement to our primary analysis. In the revised manuscript we have:

1. Expanded the Results to describe a total of five protein-coding variants—including *SLC16A11* (Figure 2c)—whose LAI-derived AFs reveal strong ancestry-specific enrichment, despite being classified as "rare" by global gnomAD frequencies:
    a. *APOL1* 22-36265860-A-G (focal segmental glomerulosclerosis, HIV-associated nephropathy, and hypertensive end-stage kidney disease risk): 27% in African tracts vs. 1% global (Figure 2d)
    b. *COL27A1* 9-114195977-G-C (Steel syndrome founder mutation): 0.1% Admixed American global vs. ~1% in Amerindigenous tracts (Figure 2e)
    c. *CFTR* 7-117606754-G-A (cystic fibrosis): 0.007% global vs. 0.1% in African tracts (Figure 2f)
    d. *ABCB11* 2-168970131-G-A (BSEP deficiency): 0.048% Admixed American global, but exclusively European-tract in LAI
2. Added two new panels to Figure 3—one for each genetic-ancestry group. Originally, in Figure 3 a and b, we show the number of variants that have a tract AF is **at least 2× higher** ("enriched"), but we have now added panels c and d that show the number of variants that have a tract AF is **≤ 0.5×** of the other tracts ("depleted"). Our added text is: "To complement this analysis, we also evaluated ancestry-specific depletion - variants whose AF in one ancestry tract is less than or equal to half that of the others (**Fig. 3c-d**) providing a more comprehensive genome-wide view of how allele frequencies diverge across ancestries. When restricting to monomorphic sites in the depleted tracts, 62.2% of variants in the African/African American group and 26.2% in the Admixed American group met this depletion criterion." (*Page 10, Line 215-220*). The figure is shown below.



Together, these additions demonstrate both through individual variant examples and a genome-wide tract-enrichment scan that LAI-informed AFs can help uncover clinically relevant variants that would otherwise be overlooked.

References
Missing authors names for reference 8.

Thank you for catching the missing author names in reference 8 (currently reference 4). We have updated the citation as follows: National Academies of Sciences, Engineering, and Medicine. 2023. *Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26902.

Methods
No comments

Reviewer #3 (Remarks to the Author):

In this work, the authors calculate, report and share allele frequencies stratified by local ancestry, across biallelic gnomAD variants. To accomplish this, they ran local ancestry deconvolution on the genotypes of African/African American and Admixed American participants from gnomAD. To justify the impact of their work, the authors begin by demonstrating the usefulness of local ancestry deconvolution for capturing between-individual heterogeneity in genetic variation. Subsequently, they demonstrate several ways in which “local ancestry”-stratified allele frequencies are useful. These include: (1) characterizing the impact of common variants; (2) characterizing the impact of rare variants; (3) improving on existing methodology for filtering variants in clinical variant interpretation tasks.

I consider this work important for multiple reasons. First, it demonstrates how studies on admixed populations can support interpretation of variant effects. The authors list multiple well-studied examples of how (local) ancestry-specific allele frequencies of functional variants in disease-linked genes could better account for the distribution of disease incidence between ancestrally diverse groups of individuals. While it would be ideal to report ancestry-specific variant associations with the trait alongside the corresponding “pooled” variant association (i.e., trait-variant correlations with/without conditioning on local ancestry, or even just admixture mapping P-values or trait prevalence in the same gnomAD study cohort), it is still reassuring to see that the examples do broadly show a better concordance between ancestry-specific allele frequency and the disease incidence heterogeneity between ancestrally diverse individuals.

Second — and I consider this a broader scope of impact — this work provides new ideas for investigating complex traits within admixed individuals. We know that variants with highly differentiated ancestry-specific allele frequencies (e.g., the Duffy null polymorphism rs2814778 at chromosome 1q23.2) can explain differential trait expression. We also know that rare mutations are more likely to be pathogenic, but the notion of rare depends on how the allele frequency or allele count is computed. The authors explored these ideas briefly in their paper, by reporting overall statistics of allele frequency heterogeneity by local ancestry, and by comparing distribution of allele frequencies in variants stratified by ClinVar annotations. I consider their finding, that benign variants seem to span a larger range of allele frequencies when tagged by African local ancestry, reasonable for justifying the use of local ancestry-stratified allele frequencies as a preliminary filter of variant pathogenicity. This is because (at least to my understanding) a strongly pathogenic variant should be selected against, given any population history that is not also prominently marked by the potentially confounding effects of bottleneck and genetic drift. So, by considering allele frequencies across as many population histories possible — including both recently admixed populations and historically admixed populations subsequently isolated long enough to lose any

signature of ancestral mosaicism — our confidence of an allele that segregates at a moderate frequency being benign should increase. Arguments of this sort can also be provided to evaluate the impact of a variant in environments specific to a group of individuals, potentially enhancing our understanding of the effect of a mutation across different environmental contexts.

Finally, some comments on methodology and analytical approach. The use of Eagle for phasing, followed by RFMix (v2) for local ancestry inference (LAI) and Tractor for allele dosage calling all sound reasonable to me. The reference populations used for LAI also sound reasonable to me. I do not see any glaring issues with their analyses, though I have a small suggestion (see below).

In summary, I congratulate the authors for performing these important analyses and for making their allele frequency data publicly available. I have some comments, one of which I consider a necessary action item (Major Comments).

We sincerely thank the reviewer for their thoughtful and encouraging comments regarding the importance and potential impact of our study. We appreciate your recognition of the utility of local ancestry-informed allele frequencies in improving variant interpretation, particularly in the context of admixed populations. Your summary of our approach and findings was both accurate and insightful, and we are glad to hear that the methodology and reference panel choices used for local ancestry inference were deemed appropriate.

We especially appreciate your point regarding how stratifying allele frequencies by local ancestry can support both rare and common variant interpretation, and the broader potential for these data to inform complex trait analysis and trait prevalence patterns in diverse populations. As you noted, understanding the distribution of allele frequencies within specific ancestry tracts can strengthen the evidence for or against variant pathogenicity—especially for alleles that hover near critical filtering thresholds.

To further support this interpretation, we expanded our analysis in the revised manuscript to examine local ancestry-informed allele frequencies for variants near the 5% filtering threshold, which is commonly used in clinical variant interpretation guidelines (e.g., the ACMG BA1 rule). We hope these revisions more clearly highlight the clinical utility of our analyses and underscore the value of our local ancestry-informed resource.

Thank you again for your generous and thoughtful feedback. We are excited to have this local ancestry-informed frequency resource publicly available and look forward to its

continued use in clinical and research settings to improve the resolution and equity of variant interpretation.

Major Comments

1. The authors write in Lines 217-218 that “87.69% and 89.32% of ClinVar variants with a higher tract AF than grpmax AF were classified as ‘Benign/Likely Benign’...” and conclude that LA-informed AFs have a substantial influence such that their exclusion could lead to inaccurate variant classification. I find their evaluation not very convincing. I also looked at Supplementary Table 1, and was not sure where the change in counts is shown**. If the original grpmax/popmax AF was sufficiently large, then I would not expect the “Benign/Likely Benign” label to change from the inclusion of LA-informed AFs to its calculation**. I would be more convinced if one looked at variants with original grpmax/popmax AF near the threshold AF for filtering (below, not above it). Show that among these variants, those less pathogenic are likely to have larger LA-informed AF (which would be more convincing for showing the power of LA-informed AFs in “correcting” a previously inaccurate classification).

Thank you for your insightful comment. In our original manuscript, we stated that “87.69% and 89.32% of ClinVar variants with a higher tract AF than grpmax AF were classified as ‘Benign/Likely Benign’,” suggesting that LAI-informed allele frequencies substantially influence variant classification. As per your suggestion, to further scrutinize this point, we focused our analysis on variants with global (grpmax/popmax) allele frequencies below the 5% threshold—the range where filtering decisions critically impact classification as clinical guidelines treat AF > 5% as stand-alone benign evidence (BA1).

Specifically, we compared grpmax allele frequencies to the LAI-informed allele frequencies for the Admixed American and African/African American groups. For the Admixed American group:

- 563,042 of 8,527,238 variants (~6.6%) showed this pattern.

In contrast, for the African/African American group the corresponding proportion was lower:

- 475,698 of 15,775,667 variants (~3.0%) showed this pattern.

To further address concerns about ClinVar filtering, we repeated this analysis using only high-confidence variants (i.e. those with 2 or more ClinVar stars). For the Admixed American group:

- 3,164 out of 31,485 variants (~10.0%) had a tract AF > 5% that was higher than the grpmax AF.

For the African/African American group:

- 2,684 out of 42,461 variants (~6.3%) met these conditions.

These results demonstrate that even when restricting to variants with high-confidence ClinVar annotations, a notable—and in some cases higher—proportion of variants near the grpmax AF threshold exhibit increased LAI-informed allele frequencies compared to all variants. We have added these analyses (with detailed counts and proportions) and text to the revised manuscript: “To further explore the clinical relevance of LAI-informed allele frequencies, we focused on variants below the commonly used 5% allele frequency threshold, which is treated as stand-alone benign evidence under ACMG’s “BA1” guideline (**Supplementary Table 3**). Comparing the former grpmax AF to LAI-informed tract AFs in this critical range revealed that for the Admixed American group, 563,042 of 8,527,238 variants (~6.6%) had a tract AF exceeding 5%. In contrast, 475,698 of 15,775,667 variants (~3.0%) in the African/African American group showed the same pattern. To determine whether these shifts persisted among clinically curated variants, we repeated this analysis using only ClinVar variants with ≥ 2 stars. Among high-confidence variants, 3,164 of 31,485 variants (~10.0%) in the Admixed American group and 2,684 of 42,461 variants (~6.3%) in the African/African American group had tract AFs above 5% that exceeded their grpmax values. These findings highlight that ancestry-specific allele frequency information can better assist in the identification of likely benign variants.”(*Page 12, Lines 264-275*).

Regarding Supplementary Table 1 (current Supplementary Table 2) , the “Allele Count” column reports the variant count that has a higher tract AF when compared to the grpmax value, categorized, by clinical significance. The “Percentage” column corresponds to the “Allele Count” number. We hope this clarifies the table layout.

2. Building on my previous point, one could use the same ClinVar variant classification table to label variants here, but I would also like to see (and highly encourage) the use of Variant Effect Predictors (VEPs), for example CADD. (CADD scores are continuous and can be downloaded from https://cadd.gs.washington.edu/.) I do not expect the authors to perform substantial comparisons using multiple VEPs, but given the growing interest in VEPs (https://www.varianteffect.org/veps) I consider this type of analysis crucial for broadening the impact of computational variant effect predictors to diverse groups of individuals — echoing a point raised by some of the present authors in their

earlier work (see, e.g., first two sentences of Atkinson et al., 2021 “Reply to: On powerful GWAS in admixed populations” in Nature Genetics).

Thank you for your suggestion to incorporate VEPs such as CADD into our analysis. We agree that continuous metrics from predictors like CADD can provide important insights into variant pathogenicity and help broaden the impact of our work. To address this, we stratified our analyses by different CADD PHRED score thresholds (≥10 and ≥15, corresponding to the top 10 and 5 percent of deleterious variants in the genome) and compared the proportion of variants for which the LAI-informed allele frequency (tract AF) exceeded both 5% and the grpmax AF.

Specifically, we compared grpmax allele frequencies to the LAI-informed allele frequencies for the Admixed American and African/African American groups. For the Admixed American group:

- Under a CADD PHRED score cutoff ≥10, 32,267 of 529,971 variants (~6.1%) exhibited a tract AF (LAI AF) exceeding 5% and higher than the global AF.
- With a more stringent cutoff (CADD PHRED ≥15), 10,279 of 179,525 variants (~5.7%) did so.

In contrast, for the African/African American group the corresponding proportions were lower:

- ~2.7% for CADD PHRED ≥10 (24,336/916,611),
- ~2.4% for CADD PHRED ≥15 (7,394/303,688).

This data indicates that even when incorporating CADD scores—which provide an independent assessment of variant deleteriousness—the inclusion of LAI-informed allele frequencies may impact variant classification. In other words, variants near the filtering threshold for grpmax AF often have LAI-informed AFs that increase their likelihood of being classified as benign. We have reported these statistics in Supplementary Table 3 and added text to detail this in the manuscript: “We additionally conducted a similar analysis stratifying by CADD PHRED score thresholds, comparing the frequency of variants where LAI-informed AF exceeded 5% and grpmax. For the Admixed American group, 6.1% of variants with CADD ≥10 and 5.7% with CADD ≥15 met this condition. For the African/African American group, corresponding proportions were 2.7% and 2.4%, respectively. These trends reinforce the observation that local ancestry resolution meaningfully alters the interpretation of both potentially benign and functionally consequential variants. Using grpmax AF's without consideration of LAI

could lead to inaccurate variant classification in clinical studies, particularly when evaluating the pathogenicity of variants in underrepresented populations." (*Page 13, Lines 275-284*)

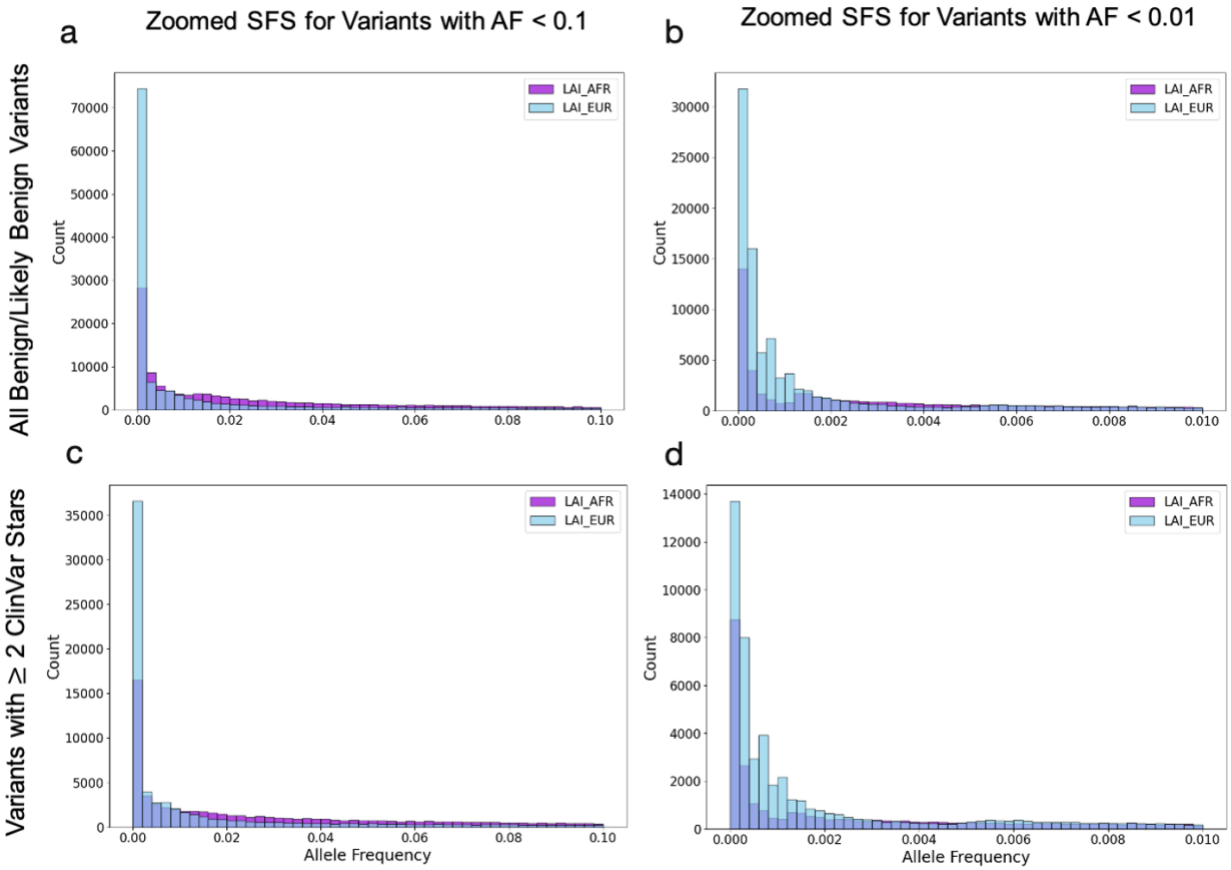
Minor Comments
1. In Lines 205-206, the authors write that "the broader distribution of benign variants in African/African American groups reinforces the need to include diverse ancestry groups in genomic studies". I am curious if the site frequency spectra (SFS) are significantly different. Could the authors run a simple non-parametric test and report p-values (say, by converting the SFS into an empirical probability distribution and then running a Kolmogorov-Smirnov test)?

We appreciate the reviewer's suggestion to formally test whether the site frequency spectra of benign variants are significantly different between LAI_AFR and LAI_EUR distributions in the African/African American group. We performed a two-sample Kolmogorov–Smirnov (KS) test to compare the allele frequency distributions of benign variants (i.e., variants labeled as "Benign/Likely_benign") between the LAI_AFR and LAI_EUR groups in the African/African American cohort. The KS test yielded a statistic of 0.278 and a p-value of near 0 (p = 10^-100), thereby strongly rejecting the null hypothesis that the two distributions are identical. We conducted the same test after we filtered for variants with 2 or more ClinVar stars: KS statistic = 0.225 (p = 10^-100), again indicating a significant difference in allele frequency distributions. These results demonstrate that the SFS of benign variants differ significantly between the African and European ancestry tracts, suggesting that the broader distribution observed in African/African American populations is not solely an artifact of ClinVar filtering based on allele frequency.

Reviewer 2 had a similar comment and suggested zooming in to the same SFS plot. Thus, we have additionally generated a zoomed-in SFS plots for benign variants with AF below 0.1 and below 0.01. These histograms (see below) compare the LAI AFR and LAI EUR allele frequency distributions in the African/African American cohort. Consistent with our broader observations, the LAI_AFR distribution remains shifted relative to LAI_EUR even in these lower-frequency ranges.

We have now included these results in the main text and the corresponding SFS plots (zooming in on variants with AF < 0.1 and AF < 0.01) as a Supplementary Figure 7. Our revised text is: "To further investigate if benign classifications may be driven primarily by frequency thresholds, we zoomed-in on benign variants and again on benign high-confidence ClinVar variants (≥2 stars) with allele frequency <0.1 and <0.01 (**Supplementary Fig. 7**). We conducted a two-sample Kolmogorov–Smirnov test on

benign/likely benign ClinVar variants in the African/African American group, comparing the SFS between LAI-AFR and LAI-EUR. The test yielded a KS statistic of 0.278 with a p-value near zero (p = 10^-100), indicating a significant difference between the two distributions." (*Page 11, Lines 239-245*).



2. In Lines 210-211, I was initially a bit confused by what grpmax meant, though in retrospect, from the name I was able to make a good guess. It would be helpful to add a short definition of grpmax to remind the reader exactly what it measures.

We completely agree with your suggestion and have rephrased the manuscript to define grpmax more clearly. Our revised language is: "We additionally investigated the frequency of ancestry-enriched AFs exceeding those of grpmax (formerly popmax), which measures the highest allele frequency observed in any gnomAD genetic ancestry group. As grpmax is often used to inform a variant classification of benign/likely benign, better resolving allele frequencies across ancestral haplotypes with LAI provides a more precise estimate of frequencies that may change clinical interpretation." (*Page 11, Lines 250-254*).

Reviewer #3 (Remarks on code availability):

From looking at the Github repository, the authors' Python scripts all have basic documentation and so are interpretable. However, I did not execute their code to evaluate its functionality, because I do not have the input data files to test their code.

We appreciate your assessment of our GitHub repository and the documentation within our Python scripts. To further improve usability, we have updated our README to include a clear example of how to calculate Local Ancestry Inference (LAI) data using our pipeline. This step-by-step guide ensures that users can easily follow along and better understand our workflow. Additionally, we have included a reference to a toy dataset available in the Tractor wiki, which provides an example cohort dataset. Specifically, this unphased whole genome sequencing dataset consists of chromosome 22 from 61 African American individuals from the Thousand Genomes Project, who are two-way admixed with ancestry components from continental Africa (AFR) and Europe (EUR). This will allow users to explore the functionality of our pipeline without requiring access to gnomAD data. You can find the updated README here: https://github.com/broadinstitute/gnomad_local_ancestry. (*Page 29, Lines 587-588*).

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors addressed all my concerns.

Reviewer #1 (Remarks on code availability):

Code is acceptable.

We appreciate your time and insight throughout the review process. We're glad to hear that the revised manuscript fully addresses your comments, and we truly appreciate your help in improving the clarity and quality of the work.

Reviewer #2 (Remarks to the Author):

I believe the authors addressed my questions related to the original version of their manuscript; I have no further comments.

Thank you for your careful review and helpful suggestions. We're pleased that the revised manuscript meets your expectations, and we appreciate your feedback throughout the revisions process, which has significantly strengthened the work.

Reviewer #3 (Remarks to the Author):

Overall feedback
The authors have addressed most of my concerns. Overall, I find that the revised manuscript has a clearer and broader scope of impact, thanks to the expanded list of examples and systematic analyses comparing allele frequencies across clinically and computationally relevant benchmarks. I have a few comments, arranged in approximately decreasing order of issue seriousness.

Comments
1. In Lines 274-275, the authors write that "ancestry-specific allele frequency information can better assist in the identification of likely benign variants." I am not so sure that is the conclusion of the analysis of tract-specific AF exceeding grpmax values — how is tract-specific AF exceeding 5% assisting with benign variant identification? The BA1 guideline seems to have been developed from various lines of evidence across analyses of continental populations. (In Ghosh et al., 2018 PMID: 30311383, there is no mention of admixed population or genetic ancestry group.) It is not clear to me that the variants identified with tract-specific AF > 5% have already been validated as likely benign. There is also potential noise in the estimated AFs given the smaller effective sample size of these local ancestry-conditioned AFs. In my view, a more assumption-lean conclusion is that these variants potentially complicate interpretations

of variant benignness based on the BA1 guideline. Or perhaps there is already orthogonal evidence (e.g., functional depletion using VEPs) demonstrating that these variants are indeed likely benign, in which case the authors should report and cite it. I would like the authors to clarify this point in their current manuscript.

Genetic Ancestry Group	Filter	Total variants	Upgraded variants	Upgraded variants w/ ClinVar annotations	Benign/Likely benign	Pathogenic/ Likely pathogenic	Conflicting classifications	Uncertain significance
Admixed American	No filters	8,527,238	563,042	5,038	4,964 (98.5 %)	1 (0.0 %)	48 (1.0 %)	25 (0.5 %)
	CADD PHRED ≥ 10	529,971	32,267	1,243	1,219 (98.1 %)	1 (0.1 %)	19 (1.5 %)	4 (0.3 %)
	CADD PHRED ≥ 15	179,525	10,279	754	735 (97.5 %)	0 (0.0 %)	16 (2.1 %)	3 (0.4 %)
	ClinVar stars ≥ 2	31,485	3,164	3,164	3,159 (99.8 %)	1 (0.0 %)	0 (0.0 %)	4 (0.1 %)
African/ African American	No filters	15,775,667	475,698	4,504	4,466 (99.2 %)	1 (0.0 %)	23 (0.5 %)	14 (0.3 %)
	CADD PHRED ≥ 10	916,611	24,336	911	898 (98.6 %)	1 (0.1 %)	9 (1.0 %)	3 (0.3 %)
	CADD PHRED ≥ 15	303,688	7,394	536	520 (98.9 %)	0 (0.0 %)	5 (1.0 %)	1 (0.2 %)
	ClinVar stars ≥ 2	42,461	2,684	2,684	2,676 (99.7 %)	1 (0.0 %)	0 (0.0 %)	7 (0.3 %)

We thank the reviewer for this insightful comment. Among LAI-upgraded variants (grpmax AF ≤ 5% but tract AF > 5%) with high-confidence ClinVar annotations (≥ 2 stars), 99.8 % in the Admixed American group and 99.7 % in the African/African American group were already classified as benign or likely benign. This near-perfect

concordance confirms that our local ancestry–informed strategy reliably identifies benign variants.

To clarify the connection between tract-specific AF elevations and variant benignness, we have revised our interpretation by adding the following text: “Notably, when limiting to high-confidence variants almost all upgraded variants were already classified as benign or likely benign, with 99.8% in the Admixed American group and 99.7% in the African/African American group. This strong concordance provides compelling evidence that local ancestry–informed frequency refinement is a practical approach for identifying truly benign variants. However, there remain eight high-confidence ClinVar variants of uncertain significance that exceed the 5% threshold in one or more ancestry-specific tracts and may similarly benefit from reevaluation.” (Page 12, Line 274-280).

“Almost all of these “upgraded” ClinVar variants were already classified as benign or likely benign, demonstrating the specificity and clinical utility of local ancestry refinement. By uncovering haplotype-specific frequency shifts, this approach can streamline variant interpretation and reduce unnecessary follow-up across diverse patient populations. For instance, the *CHIT1* variant of uncertain significance surpasses 5% in both African and Amerindigenous tracts, meeting the BA1 criterion for benign classification. Given *CHIT1*’s links to Chitotriosidase Deficiency and inflammation-related traits, recognizing this local enrichment could help avoid unwarranted clinical evaluations.” (Page 15, Line 339-346).

2. In Lines 276-277, the authors write that "CADD PHRED score thresholds (≥10 and ≥15, corresponding to the top 10 and 5 percent...". I could be mistaken, but given the formula -10*log10(rank/total) for PHRED scaling, shouldn’t a cutoff of 15 correspond to the top 3 percent (since -10*log10(0.031) = 15.1)? If so, please change it.

We thank the reviewer for catching this. You are correct that a PHRED score of 15 actually corresponds to the top ~3 % of variants ($-10 \cdot \log_{10}(0.031) \simeq 15.1$), not 5 %. We have updated the manuscript text: “We additionally conducted a similar analysis stratifying by CADD PHRED score thresholds (≥10 and ≥15, corresponding to the top 10 and ≈ 3 percent of deleterious variants in the genome, respectively), comparing the frequency of variants where LAI-informed AF exceeded 5% and grpmax.” (Page #13 Line 281-284).

3. Which CADD version was used to score variants? Computational VEPs like CADD are regularly updated, so please specify the version used here to improve reproducibility of the analysis.

We used CADD version 1.6 to score all variants in the gnomAD v4 dataset and thus used these in our analyses. We have added text to specify this in the methods section: “All gnomAD v4 variants were annotated with Combined Annotation Dependent Depletion (CADD) version 1.6 scores and used in all downstream analyses.” (Page 28, Line 578-580).

Reviewer #3 (Remarks on code availability):

The revised Github repository appears more well-documented. I was also able to access the toy dataset in the Tractor wiki.

Thank you for evaluating our expanded repository and wiki. We are glad to hear that it is now more user-friendly and will well support replication of the work.

REVIEWER COMMENTS

Reviewer #3 (Remarks to the Author):

The authors have addressed all of my concerns. There is only a small issue, which the authors should be able to address quickly. On p.15, the authors write "...the CHIT1 variant of uncertain significance...". I was not sure which variant this is, and could not find any other reference to it in the Main Text. It would be informative to specify the variant.

We're grateful for the reviewer's careful reading and for flagging this point. To clarify the *CHIT1* variant in question, we've added the following to the Results section: "For example, the ClinVar "uncertain significance" variant chr1:203229707 (C/T) in *CHIT1*, associated with inflammatory or tissue-remodeling phenotypes, has a grpmax AF of below 5% but exceeds 5% in both LAI-AFR and LAI-AMR tracts. Such enrichment pushes it past the ACMG BA1 benign threshold in multiple ancestral contexts, suggesting it may warrant reevaluation as benign/likely benign." (Page 13, Line 280-284).