Editor Comments to Author:

Three reviewers have provided thorough assessments on this opinion manuscript. All authors are in agreement regarding the value of this contribution. I concur with this assessment. Additionally, as R1 highlights, the manuscript is timely given the increased reliance on long read data. However, all three reviewers also bring up important points that should be addressed before submitting a revision. While I agree with all points of criticism, I will highlight a few of them here. First, more background is needed on theoretical population genetic ideas that are relevant in this context (R2). In providing this information, the authors should aim to improve the readability of their paper and making it accessible to the diverse readership of Molecular Ecology. Also, the paper could reach a broader audience if a worked example would be provided on how an empirical dataset would be used in this context (including currently available tools and limitations; see comments from R1, R2, and R3). For example, how exactly might this definition of haplotype blocks benefit studies of selective sweeps? More details on the simulations used here are needed (R1), as are clarifications of the figures (R2). Lastly, the manuscript would benefit from increased discussion on the tradeoffs between the feasibility of analyses for realistic datasets, and how informative and analytically tractable results are likely to be (R3). I congratulate the authors for this contribution, and look forward to reading a revised version.

Dear Dr. Sibbett,

We are happy that our study is of interest for Molecular Ecology and we thank you for the opportunity to revise the manuscript and address the issues highlighted by the reviewers. We are happy to submit a revised manuscript that has been improved throughout, along with a worked example on empirical data. Throughout this text our responses are highlighted in green.

We highlight here the main improvements addressing the major points:

1. Theoretical background

We agree with the reviewers' that the question of haplotype block definition requires previous knowledge of coalescent theory. In order to make the text more accessible we implemented the following changes: we clarified our statements in writing and added a glossary of key terms. As suggested by referees (R1, R2) we expanded on key details of the simulation in the text and in the supplementary and clarified figures.

Additionally, we switched the key terms from “branch” to “edge”. We agree that the term “edge” relates more directly to graph and ARG usage and helps to clarify the text.

1. Empirical dataset and further practical implications

A key new addition in our revision is the incorporation of an empirical example. We thank the Editor and the Reviewer for the suggestion: this example serves to clarify how haplotype blocks can be visualized, highlights many exciting avenues for further investigation, and demonstrates the utility of the ARG approach. We believe this example demonstrates feasibility and utility, and we discuss possible areas of improvements for ARG-based analysis along the lines we propose here. The provided example refers to a selective sweep example, which clearly motivates many readers, but the same principle is generally applicable to other scenarios (see also our example on an additional background locus in the Supplementary Information).

Reviewer Comments to Author:

Reviewer: 1

Comments to the Author

Shipilina, Stankowski, Paul, Chan, and Barton make the case for using haplotype blocks, rather than SNPs, as the core focus of locus-specific population genomic studies. They argue that technological advances that enable experimental phasing, together with computational tools for ancestral recombination graph (ARG) inference, should allow studies to track haplotype blocks rather than individual mutations, and that this change in focus could help address shortcomings of current methods for inferring selection, population structure, and gene flow. The authors then propose a definition for haplotype blocks based on ARG topology and explore properties of haplotype blocks, thus defined, using simulated data under neutral evolution as well as a selective sweep.

This manuscript arrives at a time when big changes in population genomic inference are likely, due to advances in both data collection (i.e. long read sequencing) and data analysis (i.e. scalable ARG or tree sequence inference). Furthermore, the suggestion to move away from SNPs as the core focus of studies is a good one (the authors could cite the Ebert 2021 paper in Science that found a large indel in an intron of LCT that may be the causative mutation behind lactase persistence, rather than the previously reported causal SNP). This paper also provides some useful figures for understanding and visualizing properties of the ARG, which will assist many readers new to ARG literature. The 3D visualizations of haplotype blocks together with the branches they arose on are particularly helpful. I have several broad concerns about the manuscript as written, however.

We are very grateful for the thorough and constructive comments. The reviewer’s perspective and suggestions have greatly helped to improve our manuscript. As this manuscript was meant to give a perspective on the question of haplotype block definition and provoke discussion, we are grateful to the reviewer for suggesting multiple further directions and we have discussed some of them. While in our view, some of these directions may constitute separate studies in their own right, we nevertheless try to point them out in the text. We respond to specific comments below.

The core problem this paper seeks to solve – confusing definitions of haplotype blocks – seems incompletely addressed, as the proposed definition depends on a choice of branch in the ARG. This allows for there to be many different overlapping haplotype blocks, some present in the same genomes, at any given locus. Without an inferred ARG, any given branch could still manifest as a set of SNPs in perfect LD, making it unclear why the ARG is necessary. The authors seem to suggest that one benefit knowledge of the ARG provides over SNP data alone is the ability to include regions of haplotype blocks that are not tagged by SNPs. Without informative SNPs, though, no ARG inference method could be relied on to get the history of these regions correct in the first place (this is alluded to on lines 224-227). And given the authors’ concerns about potential biases in existing scalable ARG inference strategies (line 391-392, 837-840), it is unclear how the authors believe they should be identified. As it stands, choosing a branch in an inferred ARG does not seem much less arbitrary than choosing a set of linked SNPs that tags that branch (similar to, for example, S\* from Plagnol and Wall 2006) – unless the authors, for example, mean to highlight that the ARG also tells us how far along the chromosome that haplotype persists. In addition to proposing this new definition for haplotype blocks, the manuscript should clearly state why and how ARG-based haplotype identification can benefit specific types of analyses.

Our paper is primarily about showing how “haplotype blocks'' arise from the structure of the ARG, and showing how they can be clearly defined; we do not undertake the much larger task of developing practical inference methods from actual data (which would be a major new study in its own right!). However, we now implemented a particle example to illustrate how choice of edges can be performed in practice using a *Heliconious* linked read data set, and suggest ways to describe and visualize them.

Choosing a particular edge (formerly termed “branch”) would indeed be arbitrary (as we showed in the example of selective sweep) - rather, data exploration and subsequent inferences should ideally be made from the full set of branches (that is, from the full ARG). In practice, we may consider the most substantial branches, tagged by sufficient SNPs; how many are informative will depend on what question is asked. It may be that in certain types of analysis, we can use SNP that are shared by a particular set of sampled genomes as surrogates for a branch in the ARG - but finding how accurate that is, compared with using an inferred ARG, is beyond our scope.

We do have in mind that the extent of a branch along the genome may give important information, as the reviewer suggests. Again, finding how much extra information this gives, beyond the genealogy at the sweep (for example) is an open question. We argue that length of the haplotype block may contain additional information about dating of evolutionary events and provide insight on inference of incomplete or multiple co-occurring sweeps.

In our new example of a swept locus, the choice of edges were fairly obvious, given our particular interest in a subset of sampled genomes (butterflies, in this case). Still, selecting these particular edges highlight features of the sweep that may otherwise not be obvious. We thank the Reviewer for suggesting this analysis, which goes a long way towards focusing and refining our thinking.

The authors discuss selective sweeps – a topic of interest to many readers – but do not clearly explain how their definition of haplotype blocks will benefit future studies. Selective sweeps are often population-specific, but the example given shows an allele that swept to fixation in all lineages. In Figure 3D, the TMRCA (of all lineages?) is shown to be lower near the selected mutation, but in the case of population-specific selection, the TMRCA of all lineages is less likely to be a useful metric (the Rasmussen 2014 paper defined the “relative TMRCA halflife” as a way to identify regions where some but not all clades might have an abnormally recent TMRCA). In the example given, the authors highlight ARG branches with large numbers of descendants as indicative of the simulated sweep, but they do not elaborate. A new, ARG-based selection statistic would be of use to many and could be explored further here (to complement, for example, those from Speidel 2019 and Stern 2019). As it stands, it is not clear exactly what the authors are proposing for the detection or study of selective sweeps, and the sentence on line 322-323 ends this discussion without making a clear point. Moving in the opposite direction – seeking to identify the selected site using the ARG topology – could be a useful exercise and could illustrate the power of using ARG-defined haplotype blocks in selection scans.

We again emphasise that we aim to find ways to better define and describe the ARG, rather than develop new statistical methods for specific estimation problems. We provide one illustration of the very simplest case, of a complete sweep, from one copy to fixation, in a single population. There are very many other scenarios, and we cannot investigate their consequences for the structure of the ARG here. Rather, we point out that on the one hand, the ARG contains much more information than is currently used, but on the other, there is considerable variability amongst different realisations, even under the simplest model.

We chose the very simplest case of a hard sweep that fixes within a single population. There are of course many other scenarios - but the purpose of our paper is not to propose new methods for (say) detecting selective sweeps, but to clarify the more fundamental question of how we should define "haplotype blocks", and how we can use them to help visualise the ancestral structure.

The Reviewer may welcome our empirical example, which shows a clear selective sweep in a subset of the samples. Here, we show that similar concepts like the Reviewer has proposed can be very powerful: subset TMRCA, long internal edges, etc. In the future, we would be keen to see if a thorough exploration of these details may yield useful statistics.

The simulations are a large part of this manuscript but are not described in much detail in the text, aside from mentioning the infinite sites assumption and some of the parameters chosen. The supplementary files are very detailed and contain some beautiful visualizations, but a separate writeup (or reference to another publication in which the simulator is described) would be helpful. Readers might be interested, for example, how the simulator here compares to others with which they might be more familiar.

We added an introduction to the supplement which explains the simulation in detail, and substantially extended its summary in the main text.

The authors mention that Relate and tsinfer, two recently-described ARG inference methods that scale to large data sets, both use the Li and Stephens haplotype-copying model and express concern that this could lead to bias (line 391-392, 837-840). This potential bias would arise from the existence of disjunct haplotype blocks, or specific haplotype blocks that exist in multiple, non-contiguous places in the genome, because the Li and Stephens HMM only considers the source/ancestral haplotype of the immediately preceding site when modeling a sampled haplotype. It seems to me, however, that the HMM could easily transition back to the same source haplotype again, if SNPs are encountered further down the chromosome in the sampled haplotype that again resemble that source haplotype closely enough. In other words, the model should work well enough if there are enough informative SNPs. If there are not enough observed SNPs, it is also hard to imagine any ARG inference strategy that could accurately assign ancestry in a given region. Do the authors have a suggestion on how to overcome this potential bias?

We have elaborated on the potential limitations of the Li and Stephens HMM, pointing out situations where the algorithm may be able to accurately detect transitions back to previous states. We have also pointed out that future work is needed to understand how problematic this would be for accurate inference of the ARG.

This bias could also be explored using the simulated data sets – since disjunct haplotype blocks have already been identified, one could infer ARGs using tsinfer and Relate on the simulated data and investigate how often disjunct haplotype blocks are successfully recovered. The authors could also explore the effects of the mutation/recombination rate ratio on the ability to identify such haplotype blocks. This could have implications for other studies – for example, there must be a tradeoff between the wide availability of reference data but low heterozygosity in humans, versus the scarcity of reference data but higher heterozygosity in some other species.

We have pointed out that more work is needed to evaluate and understand the performance of different methods for inferring the ARG under a variety of circumstances. However, this will take a substantial effort, well beyond the scope of our paper. (Brandt et al. (Genetics, 2022) consider the performance of these methods in recovering pairwise coalescence times, which is the simplest feature one can examine). Additionally, we have touched on the need to test the application of definition to the species varying heterozygosity and genetic data availability.

Finally, the conclusion (on lines 424-426) introduces the concept of graph genomes and alludes to problems with reference-guided assembly. Although structural variants are clearly an important and overlooked class of mutation, it is not clear how or whether the more complex classes of variants that graph genomes can uncover can be encoded into ARGs. Some structural variants, for example, could change the physical distance between neighboring mutations in some but not other genomes, which could affect the probability of recombination in the region and thus impair an ARG inference method’s ability to infer recombination events. Additionally, methods for assigning ages to mutations depend on assumptions about mutation rates, which are well studied for SNPs and less so for other classes of variants. I have a similar concern with the statement about recombination events providing an alternative ‘clock’ to mutations (line 62-63), as it is not clear how or whether these two different pieces of information can be (or have been) reconciled to produce a single time estimate. More discussion of graph genomes and the complex variants they can describe would be welcome here, if the topic is raised.

There are two issues here. First, the ARG describes the ancestry of a set of genomes, which is defined independently of allelic state: all that is required is homology. Structural variants could change recombination rates, but the same is true of SNPs; this might raise practical difficulties, but doesn’t affect the fundamental principles. Second, mutation and recombination both act as “clocks” and both are used in existing methods, though often implicitly (e.g. in PSMC). Use of recombination as a ‘clock’ is perhaps most obvious for sharing of large blocks of genome, which indicates recent ancestry, at a time the inverse of the map length (e.g. Ralph & Coop, 2013).

Line-specific comments:

Line 50-56: Several problems with selection scans and GWAS are identified, but it is not made clear (i.e. in the following paragraph) how the new definition of haplotype blocks proposed here would address these problems. For example, it seems like a focus on haplotype blocks rather than SNPs could broaden the list of candidate causal mutations for a selective sweep. But selecting a variant from this list would likely require some kind of functional data.

Yes, and this is a key point: SNP are an imperfect reflection of the underlying haplotype structure, making it clear that it may be impossible in principle to locate causal mutations without functional information or, ultimately, experimental manipulation.

Line 106: Panel (A) indicator is missing from the beginning of the caption

This has been added

Line 137: This is shown later in Fig B1B, but mentioning that coalescences join lineages, while recombination events split them going back in time, might be helpful to readers.

We have made this more explicit in the text. This also relates to reviewer 2s request for clearer description and more background on population genetic processes.

Line 189-190: “The genome is divided into 34 non-recombining intervals…” – it is unclear to me whether this was a decision that went into the simulation, or a result observed after the simulation was run. More information about the simulations would help clarify this.

In addition to elaborating on the details of the simulation, we have expanded on this description to highlight that there are 34 regions that are defined by recombination events.

Line 211: “(C & D) Examples of regions of blocks that, by chance, are revealed by mutations…” should be “are not revealed by mutations”

Done

Line 247-250: The concept of nested haplotype blocks is interesting. It also seems to exclude haplotype blocks defined by child branches – i.e. the block defined by branch i would exclude the block defined by branch ii. This was not immediately clear, though, and could be stated more explicitly.

We have discussed it more clearly in relation to blocks rather than just edges (branches) both in the main text and provided deeper insight in SI. Practical choice of edges is now illustrated in the empirical example.

Line 252: “a particular point on the map” uses different language than the rest of the sentence -- does this mean a particular site on the chromosome?

We have revised this to make our language use more consistent.

Line 254: “The rate of recombination is proportional to the branch length” – should this be changed to the “incidence of recombination,” since the underlying rate should depend on genomic span / local recombination rate, whereas branch length indicates the amount of time over which recombination events can occur?

Yes, this is correct–the length of time is proportional to the incidence or probability that recombination will occur. We have clarified this in the text.

Line 260-261: “In simulation, deep branches tend to be wider than expected…” – it is not clear exactly what the expectation was, or how much the simulated data differed from the expectation.

We have now included a more detailed explanation in the supplementary material for the interested reader, but opt to keep an easier flow in the main text.

Line 305-306: “9 of the most substantial branches are shown” – unclear here exactly what “substantial” means, and whether these are the only branches that passed the criteria that follow in parenthesis, or if these were chosen from that set. Could these branches also be highlighted somehow in panel B?

We have clarified the details of how the edges were chosen: they are defined in the caption of Fig 4: “(These have more than 8 descendants, formed by coalescence more recently than the sweeping mutation, and have areas >0.5). “. We considered highlighting these specific branches in B but we think it makes more sense to highlight all of the branches that coalesce before the sweeping mutation rather than just specific branches. However, the coloured branches are now shown in the supplement.

Line 365: “hidden Markov model” appears twice

Fixed

Line 390-392: If haplotype blocks are poorly recovered by tsinfer and Relate, does ArgWeaver do a better job? Evidence of the proposed bias would be helpful, and ArgWeaver could be tested as a non-Li and Stephens-based alternative.

We now used ARGweaver in the practical example and discussed its applicability and limitations. The relative performance of these methods is really something of an open question, we refer to an excellent comparative study (Brandt et al., 2022) and we encourage readers to explore this in the future.

Line 397: “…but infer haplotype blocks only as an incidental output.” In my understanding, the output of Relate tells which branches existed over which genomic intervals and maps specific mutations to specific branches. The shortcoming of this type of output data for identifying haplotype blocks, as defined in this manuscript, is unclear.

To our knowledge, Relate initially infers genomic intervals in a discrete fashion, therefore the possibility to include nested blocks extending through the breakpoints may be a limiting factor for some analysis. We clarified this point in the text.

Figure B2 top panel: axis label “Reference Haplotypes” is misspelled

Fixed

Reviewer: 2

Comments to the Author

In this paper, the authors propose a new population genetic definition of “haplotype block” based on the ancestral recombination graph (ARG). The authors outline the idea behind this definition, illustrated examples using schematics, simulations, and empirical data. They discuss the utility of this new definition relative to the ad-hoc definitions used before, and outline how it might be used in the context of currently available tools (e.g. for inferring the ARG).

I’m going to be honest: I struggled to understand a lot of the finer points being made in this paper. However, I would argue that I share a lot of the background and interests of a typical Mol Ecol reader. I work with empirical population genetic data every day, and I am familiar with the ideas behind the ARG, the coalescent, etc. Anyway, in my opinion, if Mol Ecol is now the target venue, this paper would greatly benefit from more basic explanations of many of the theoretical population genetic ideas being discussed here. It seems likely that Mol Ecol was probably not the author's first choice, but I digress. In particular, the figures, while visually attractive, are very hard to understand. I realize that the authors were likely specifically attempting to use these visuals (and the boxes) as an aid to understanding the underlying theory, but I don’t think they have been completely successful. Overall, it seems like an important contribution, but I fear it will be largely impenetrable to the merely mortal readership of Mol Ecol. I’ve tried to outline some ideas for places to address this issue below.

We thank the reviewer for the detailed comments and suggested improvements. We do believe the Mol Ecol readership cares a lot about these concepts we bring up and it is our view that engaging empiricists will help strengthen our respective research and the field. We sympathize with the conceptual challenges highlighted by this reviewer, because we ultimately give a very different perspective on the ways one might approach empirical sequence datasets in light of new sequencing methods, theory and exciting new analytical tools. We have tried to include more explanation on some of the fundamental principles. Equally, we want to emphasise that the manuscript is also intended to motivate the reader to think about things beyond the status quo.

We have made significant revisions to improve clarity and readability: added explanations, references and glossary of the fundamental terms. Some of these changes coincide with detailed comments from all 3 reviewers.

The Figures

Figure 2 was interesting, but ultimately too complex for me to derive much understanding from. I did not understand the following:

- What is the connection between the colors of the branches and the colors on the DNA sequences. For example, why are the top branches on the tree on the right side of “A” orange, but they lead to red sequences (this is not the case elsewhere?).

Thanks for pointing out this detail. There was in fact an error in the coloring of the branches. This has been fixed. Hopefully it clarifies the confusion. We further extended the caption to make understanding of the figure easier:

- What specifically is meant by “light grey branch in both trees shows the effect of recombination in the genealogy” – how does this show the effect of recombination? Recombination between the sequences depicted, or some other hypothetical sequences?

We have clarified this in the captions, but now that the coloring is fixed (see comment above), hopefully this is much clearer – this must have been very confusing!

Figure 3 again was interesting but again was difficult for me to understand. Notes below:

- What are the axes in the “SNPs” section? (x and y?)

Added.

- “The trees (a - o) show all of the unique topologies that coincide with the genomic spans shown in the central panel (also labeled a- o)”. What is meant by “genomic span” in this case? For example, “O” seems to not be a span, but a single (end) point on a haplotype.

This has been clarified in the caption and main text. O is actually a span, but just a very short one!

- The “L” span appears to contain no SNPs, but has a corresponding tree with internal structure. Is the tree the genealogy (independent of allelic content), or something else?

This has been clarified in the caption, but yes: the tree shows the genealogical history of this region of the genome, but just by chance, no mutations happened to occur in this region.

How the definition would actually work in practice. The authors briefly discuss some of the practical applications of their ideas. I think it would be really useful to actually provide a worked example of how an empirical dataset could have its haplotype blocks, as defined here, and using the existing tools discussed, identified.

We now added an empirical example illustrating haplotype blocks in *Heliconius* butterflies. Our walkthrough demonstrates how ARGweaver and a combination of widely used tools may apply to a given definition.

Reviewer: 3

Comments to the Author

Shipilina et al. provide a new definition for the concept of “haplotype block”, which is based on the ancestral recombination graph (ARG). They suggest that the proposed definition better captures genomic variation than other available definitions and use simulations to highlight this.

I found the manuscript to be fairly well written (although I list a few comments below) and the authors raise several interesting points. The main contribution of this work is the proposed ARG-based definition of haplotype block. I agree with the authors that this definition is meaningful and deserves to be highlighted, although perhaps my main reservation is that this seems a bit limited in scope.

We are grateful to the reviewer for the thorough evaluation and valuable comments on our manuscript. We acknowledge that without a practical example, our manuscript may have given an impression that the scope of the term is limited. We hope that with the new practical example and more detailed discussion of further applications, it is now clear to the readers that haplotype blocks are generally applicable in population genetics.

One high-level thought about this work is that working with summaries of the data usually involves a trade-off between how easy these summaries are to compute (for instance, the frequency of an observed allele vs the ARG) and how informative and analytically tractable these summaries are for a particular analysis (see for instance PMID:26395773 for an example of “LD-blocks” that are informative about haplotype structure while remaining fairly easy to compute). The need to work directly with inferred ARGs seems to shift the effort from defining and computing meaningful ARG summaries to the non-trivial task of inferring the ARG itself. It is probably worth further highlighting this trade-off, perhaps in the discussion. Having said that, several methods are being developed to perform accurate ARG inference, so I agree with the authors that the proposed definition is becoming of practical interest. A related comment is that I find the relevance attributed to the proposed definition a bit overstated at times, for instance “haplotypes and haplotype blocks should be the core concepts through which we understand population genetic processes”.

We agree that the trade-off between effort and gains in power to make inferences will play a role in practical applications. For exactly that reason, we have highlighted efficient new approaches like tskit and RELATE to compute ARG proxies that overcome previous limitations. We have now also included a practical example that goes some way towards demonstrating feasibility. However, we should also emphasise that a core focus is on how to describe and understand the fundamental structure of the ARG, which we believe should be made clear before we think about specific applications. Thus, we stand by the sentence quoted above, that haplotype blocks (or more precisely, edges) should be core concepts for understanding population genetic processes.

I have the following specific suggestions and thoughts:

- I found the paragraph starting at L43 a bit confusing. The authors list several examples of challenging analyses where haplotype structure is important (selection scans, polygenic selection, polygenic scores, gene flow), but the implied connection with haplotype blocks is a bit vague.

We now reworked this paragraph to demonstrate, how these of type analysis could benefit from knowing haplotype structure/haplotype blocks

- The authors provide a definition of IBD sharing that is based on the existence of a founder population (L131). However, other coalescent-based definitions of IBD exist (reviewed for instance in Wakeley and Wilton, “Coalescent and Models of Identity by Descent”, 2016). The Carmi et al paper cited, for instance, uses this kind of definition.

We have added these references to the text mentioned that more wide use of the IBD approach.

- On a related point, the key element of a haplotype block in the definition provided by the authors is the presence of a coalescence event involving several individuals in a specific region, corresponding to a shared ancestor in the ARG. I suspect an equivalent definition to that proposed by the author may be provided using IBD sharing, as all individuals involved in a haplotype block will be IBD in the region. I don’t think the authors need to try to come up with this kind of definition, but might want to highlight a connection.

Yes, there is a connection, but the way we define blocks would not be the same as a definition based on IBD sharing. The key point (illustrated in Fig 2) is that the IBD definition depends on an arbitrary decision about where to cut the genealogy, which subsequently fixes the number of haplotype blocks. If “IBD sharing”, in this sense, is defined independently of a fixed time-point, then it is equivalent to coalescence and thus ARG. The differences with the ARG approach are that (a) the haplotype blocks are determined objectively from the ARG (b) and that ARG blocks can be nested, such that the multiple blocks can be defined based on subbranches of the ARG. This should become clear as we explore the implications of the definition and a practice example.

- Fig 2A: I found it confusing that haplotypes for the yellow tree are red.

This was a mistake with the coloring. We have fixed this.

“This branch is defined by a unique coalescent event” it took a few readings of the definition to understand that this coalescent event refers to the lower end of the ARG branch (I assumed it referred to the upper end, which is not unique). Some rephrasing to clarify this earlier on in the definition might help.

We have elaborated on the relationship between coalescence and the edges (branches) in the ARG. This also relates to reviewer 2's request for more explanation about basic population generic processes.

- L183: specify N is the effective size.

Done

- Fig 3: labels of tree leaves are very small.

We have now made these more legible.

- The authors may want to restrict the definition of haplotype block to non-trivial, non-singleton ARG branches (or ARG branches with lower end > 0). Or perhaps not plot them in Figure 3.

We could remove some singleton branches from the figure, but we actually think it will be harder to understand the trees if we remove some of the branches. We have therefore decided to leave this as is. Also, these branches are real features of the data even if they are not especially interesting. In practical inference, one would focus on the more substantial branches that carry multiple SNP.

- L260: the authors observe that deep branches are longer (Supp 1). I believe this is also likely due to re-coalescence events, which are less likely on short branches at the bottom (see difference between SMC and SMC’ models by McVean & Cardin and Marjoram & Wall).

We have looked at this question more closely, and expanded the corresponding section in the Supplement (“Inverse relation..”). It is true that re-coalescence events will make deep branches longer than otherwise. However, on re-examining the question, we see that the most important feature is that the distribution of spanned map length (“width”) and spanned time (“depth”) are both highly random, so that a small fraction of branches carry most of the ancestry (“area”), and therefore, of SNP. We have reworded the text accordingly.

- The discussion section felt a bit rushed and could probably be improved. For instance the paragraph starting at L401 may be moved to the discussion instead.

We have spent some effort to tidy this up.

- L826: is Li and Stephens used in ARGWeaver?

Li Stephens isn't used in argweaver. ARGweaver is based on the assumptions of sequentially Markov coalescent and a discretization of time.