Response to reviewers' comments

June 18, 2025

Below is our point-by-point response to the reviewers' comments in blue.

Reviewer: 1

Comments to the Author

Summary: In this manuscript, the authors present tsbrowse, an open-source Python-based application for the interactive visualization of the components of Ancestral Recombination Graphs (ARGs). Because ARGs (in principle) contain all knowable information concerning coalescence and recombination in the history of a set of samples, there is great excitement in developing methods for inferring ARGs from empirical genomic sequence data, as well as for learning about evolutionary processes from patterns in inferred ARGs. A crucial component of both these endeavors is visualizing ARGs, which is difficult or unwieldy. Tsbrowse therefore represents an exciting contribution to the field which is likely to be of high utility and to be highly cited.

Opinion: This paper was straightforward, clear, and well written. I think it does a good job setting up the intellectual context for ARG-based research and advocating for the utility of tsbrowse. I had two minor, nitpicky comments, but overall, it was great. Thanks for a fun and easy review!

Thank you for your kind words!

1. I think the order in which the supplemental figures are cited doesn't match their numbering (i.e., S4 is cited before S1)

We have now rearranged the supplementary figures so that figure references appear in ascending order.

2. In the discussion, the authors mention that methods for visualization of tree topology can typically only handle a few hundred nodes, but it might be worth doing a "(but see…)" for some of the approaches folks have taken for visualizing much larger trees (e.g., nextstrain, onezoom).

We had already included a citation to the OneZoom project (and the Treenome browser) in the sentence that followed the one in question "… Visualisation of large-scale tree topologies with millions of nodes requires much more sophisticated approaches to capture topological features at different scales, and is an active research area". Thanks for bringing Nextstrain to our attention, and we have now Hadfield et al. as a citation for visualisation of medium-sized toplogies.

Reviewer: 2

Comments to the Author

I have only a few minor comments:
1. "and storing as a compressed .tsbrowse file. (The .tsbrowse file is also a valid input for tszip, a general utility for compressing tskit files.)" Sorry, but I don't understand this sentence. More precisely, both the compressed and the uncompressed files have the .tbrowse extension? If yes, that's a weird choice and I would like to see some justification. You should also make clear if tszip is a part of tskit or of twbrowse.
To clarify, the tszip is part of the tskit ecosystem. It is not part of the tsbrowse package but can be used to compress the tables output from the tsbrowse preprocess step. Thank you for pointing out this potentially unclear statement. We have now deleted the text in brackets, which we believe isn't crucial for understanding the data model.

2. "web-browser" - web browser
Fixed.

3. "Fig S4, for example, shows a screenshot of tsbrowse". This is a matter of personal taste, but I do not like to refer to a Figure in the supplementary materia just as if it were in the article; if the Figure is relevant, it should be in the article. If it's not relevant, that you should refer to it without hinting that it is important.
We agree with the reviewer's opinion, and believe that this figure is important to demonstrate the application of tsbrowse to large-scale datasets (which is a key design feature of the app). However, we are unable to move this figure to the main matter considering journal guidelines for page limit.

4. Looking at the github repo, I can find the installation instruction using pip only. In my experience pip is quite picky. I strongly suggest you to provide tsbrowse via conda-forge/bio-conda or a containerized solution (e.g. docker).
We have now created a Docker container in order to ensure a smooth experience across computing environments. Usage instructions have been added to the app documentation.

Reviewer: 3

Comments to the Author
This paper offers a browser-based visualisation tool for ancestral recombination graphs (ARGs), which is a data structure representing gene genealogies for a sample of DNA sequences. The complication in the p resence of recombination is that these structures are not simple trees, but - because of the possibility of

recombination - graphs with reticulated nodes, which makes exploratory data analysis difficult.

I agree with the authors that there is currently a lack of visualisation tools for ARGs, and I expect the current paper to provide a useful addition to the catalog of tools for ARG simulation and inference. Tsbrowse integrates with the widely used tskit ecosystem of software tools, developed by some of the current authors, which will facilitate a rapid uptake. The authors make a compelling argument for the need for visualisation tools - Fig S1 in particular shows that four different and popular tools for ARG inference are giving qualitatively completely different insights into their reconstructed ARGs (incidentally, it's a shame that that figure couldn't make it into the main paper, though I recognise that this journal has strict page limits).

Thank you for your valuable feedback on the manuscript, and on the utility of tsbrowse!

We completely agree that this figure would have been a compelling main display item had the journal page/figure limits allowed for it. Considering the necessity of the existing main display items in establishing the fundamental aspects of tsbrowse, we have decided to retain the figure in the supplementary.

My main comment is that tsbrowse does not actually seem to be an ARG visualisation tool. It is a useful way to visualise mutations, tree-sequence edges, and summary statistics, but nowhere will you actually see a tree (much less an ARG). On rereading the paper I see that the manuscript has been carefully worded not to include such a claim, but a casual reader might expect tsbrowse to be able to do so, as other tools such as IcyTree can. I do get fed up with reviewers tossing additional requests on top of an already polished piece of work, but in this case I think it would be a helpful addition! - would it really be too much work for tsbrowse to be able to draw at least the local trees? (For researchers with smaller datasets, being able to interrogate the fine details of an ARG would be equally valuable as the complementary selling-points around scalability.) At the very least I would like to see the manuscript at least be a little more up-front about this omission.

This is an important observation, and we agree that drawing (parts of) local trees would be a great feature. For smaller datasets, tools such as `tskit-arg-visualizer` (`https://pypi.org/project/tskit-arg-visualizer/`) and existing tskit infrastructure, for example functions `draw_text()` and `draw_svg()` already provide this functionality. A key motivation for this work is to guide large-scale inference, and implementing tree visualisation that is interpretable at this scale would require significant thought towards design, as well as considerable effort. We therefore think that implementing this feature is not feasible for this work and is better scoped as future work or as a separate endeavour. In the text, we do state that tsbrowse is a visualiser for the information structure in ARGs, which we hope implies that it is a visualiser for the properties of the ARG, rather than the ARG topology itself. We have also already explicitly stated in the Discussion that drawing local trees is an important future direction.

Minor comments:

Manuscript p1: 'it is only with recent breakthroughs in inference methods...that practical application has become possible'. I appreciate the need to emphasise the recent increase in interest in inference for ARGs, but I think this is selling older methods a little short. There are ARG inference methods going back at least as far as Griffiths and Marjoram (1996) 'Ancestral Inference from Samples of DNA Sequences' JCB 3(4):479-502. True they are not scalable in the manner of recent methods, but they *are* practical applications, and they have the advantage that they actually sample from the claimed model without introducing any model heuristics.
Thank you for pointing this out, and you are right that we are somewhat overstating the case for recent methods. We have changed the word "practical" to "widespread", which we feel is an uncontroversial statement. The history of ARG inference is a complicated and confusing subject, dealt with exhaustively by Wong et al 2024, and would be impossible to accurately summarise in the available space. We note, though, that two of the recent methods cited here (ARGweaver and SINGER) sample ARGs from the posterior defined by a coalescent model in a statistically rigorous manner

Manuscript p1: Typo 'developements'.
This is now fixed.

Manuscript p2: Where you mention 'Bokeh', can you clarify what this is (e.g. with a citation)?
Yes, Bokeh is a Python library for interactive visualisation and a citation to it has been added.

Supplement p1, Sec 1.4: "The ARG shown in Fig 4" should say "The ARG shown in Suppl. Fig 4" ?
This is now fixed.

Fig S2 caption - 'Consortium (2015)' is citing incorrectly.
Fixed (removed this citation as it was not required)

From the online documentation, Node page: `https://tskit.dev/tsbrowse/docs/stable/nodes.html` "The histograms at the bottom show the distributions of node spans over different dimensions. The leftmost histogram summarises the span of nodes on the sequence; the middle plot summarises the span of nodes over time and the rightmost plot summarises the edge "area" defined as the product of sequence span and time span for each node."

I didn't understand this. The 'nodes' section of the visualiser only seems to provide one histogram, not three.
Thank you for pointing out this documentation error. The Nodes section does indeed have one histogram and the documentation has been updated to reflect this.

Installation notes:

I had a few problems getting tsbrowse to work. There are two different sets of installation instructions:

1. https://pypi.org/project/tsbrowse/ 2. https://tskit.dev/tsbrowse/docs/stable/contributing.html

I use a Mac+Homebrew+Chrome and Option 2 seems more fitting. (”python -m pip install tsbrowse” as in option 1 will fail because Homebrew’s Python environment is externally managed.)

We apologise for the confusion surrounding installation notes. The recommended way to install tsbrowse is via pip within a virtual environment. We have fixed the documentation so that instructions to install tsbrowse as a new user are clearly stated at the beginning of the documentation, and separate from developer instructions.

Bug(?):

I frequently had problems rendering the window in the ’Mutations’ tab, which tended to get stuck at the circling wheel. When clicking the ’Mutations’ button, my terminal returned the following error:

```
2025-05-27 12:29:41,646 [5441] ERROR tornado.application: Exception in
callback functools.partial(¿, .wrapped() done, defined at
/Users/reviewer/myenv/lib/python3.13/site-packages/panel/io/server.py:153¿ exception=AbbreviatedException(, ValueError(’Supplied cmap winter not found
among matplotlib, bokeh, or colorcet colormaps.’), )¿)
```

We are grateful to the reviewer for bringing this bug to our attention. We have solved the issue and released an updated version of the package.