OPINION ARTICLE

Animating and exploring phylogenies with fibre plots
[version 2; referees: 3 approved with reservations]

William D. Pearse
Department of Biology & Ecology Center, Utah State University, Logan, UT, USA

Abstract
Despite the progress that has been made in the visualisation of information since Haeckel's time, phylogenetic visualisation has moved forward remarkably little. In this brief essay, I give a short review of what I consider to be some recent major advances, and outline a new kind of phylogenetic visualisation. This new graphic, the fibre plot, uses the metaphor of sections through a tree to describe change in a phylogeny. As an animation, I suggest it is a powerful method to help interpret large phylogenetic hypotheses, although snapshots of it can also be displayed. As we enter the Anthropocene, I argue there has never been a greater need to know humanity's true place in the world, as depicted in the tree of life.

This article is included in the Phylogenetics channel.
A new generation of phylogeneticists are piecing together the entire tree of life, making vast phylogenies of millions of taxa\textsuperscript{1-4}. Many have produced tree-like depictions of the relationships among species, both before and after Darwin described the origin of species\textsuperscript{5}, but Haeckel’s drawings are perhaps the most well-known. As our phylogenies become larger, a problem has emerged: humans cannot easily interpret phylogenies with millions of tips. In this brief essay, I will describe recent progress in the visualisation of phylogenies, and outline a new kind of plot in development (the fibre plot). My aim is not to write a review [c.f. 5], but rather to provide an opinionated commentary on some major milestones in the progress of phylogenetic visualisation.

Haeckel’s phylogenies are beautiful to look at, and convey the overall structure of a phylogeny well. Each minor branch rarely maps onto a particular species, but their presence reminds the reader of the ever-changing nature of diversification. Both Haeckel and Darwin convey two kinds of information in their visualisations: time through depth on the page, and relatedness through the branching structure itself. Haeckel is also notable for producing a series of phylogenies, each examining a finer phylogenetic scale. Haeckel grasped that humans cannot process the fine details of all species without becoming lost, and that a series of phylogenies provides the same information in a more digestible format than a single, large, fully-resolved tree.

The last one hundred years have seen transformative changes to phylogenetic inference [see 6, but the same is not true of phylogenetic visualisation. The pace of change of phylogenetic visualisation has not matched that of other aspects of statistical visualisation. A time-traveller from 1859 could decipher a phylogeny from 2016 with On the Origin of Species as a guide, but the box-plots\textsuperscript{7} and histograms\textsuperscript{8} we rely on today would be foreign to them. Circular (“radial”) phylogenies are sometimes preferred when space is limited, and “magnifiers” in some computer programs highlight certain parts of the tree in more detail [e.g. 9], but for the most part any advances have been relatively minor.

To my knowledge, Paloverde\textsuperscript{10} was the first openly-released 3D phylogeny viewer. While it offered the user the ability to explore the most tree-like phylogenetic depiction to date, it also permitted the user to fly around a circular phylogeny rendered in a virtual space. Paloverde is notable for its claim that phylogeny is something to be explored, not merely viewed, and that “a 3D world, offers visual cues that aid in navigation and display that is unavailable in strictly 2D versions of the same layout.”\textsuperscript{10} The author of Paloverde, like Haeckel, recognised that scientists need to shift between finer and coarser phylogenetic scales when examining data, and so allowed users to collapse nodes at will. The program was a major advance in helping phylogeneticists conceptualise their own phylogenetic hypotheses. At least as transformative was the release of OneZoom\textsuperscript{11}: a fractal phylogeny representation capable (theoretically) of displaying the entire tree of life on one page. OneZoom also requires the user to explore the tree, scanning up and down between finer and coarser details to make sense of the entire tree. Critically, OneZoom’s authors recognised that we are reaching the limits of what can be displayed in books: “[w]e now need to take the next step with a transition to data visualization that is optimized for interactive displays rather than printed paper.”\textsuperscript{11} They suggest that the way to display the next generation of data is to use the next generation of technology. A common thread running through these developments is their capacity to change the information displayed to the viewer, to better emphasise difference in structure across different phylogenetic depths. Consequently, I suggest the use of a new visualisation, the “fibre plot”, which is intended to leverage our natural ability to detect visual change through time. The fibre plot may be considered a horizontal slice through the tree of life, taken at whatever height (depth) the viewer requires (Figure 1). By moving along the tree, from the root to the tip, viewers will see the relative width of each fibre, and so gauge the number of terminal tips subtending that clade. I suggest an animation, with frames recorded at equal intervals along that trunk, which will provide the viewer with an intuitive sense of the timing of the diversification of major clades. I emphasise that, while Figure 1 shows the underlying logic behind the plot, the “plot” should really be called an animation - it is most readily interpretable when the user watches a video composed of successive slices through the trunk of the tree. I have included a toy implementation of this preliminary idea in R (Supplementary File 1), and an example of how it can be used to visualise the entire mammal tree of life\textsuperscript{11} (Supplementary File 2).

![Figure 1. An explanation of a fibre plot.](image-url)
Picking an ideal layout of the tips in the two-dimensional plane at the top of the tree is complex; a colleague has suggested using Hilbert curves based on the phylogeny’s distance matrix as an alternate solution to the one I have implemented.

Despite humanity being closer than ever to a true tree of all life on Earth\(^1,2\), phylogenetic visualisation may seem like a niche topic. I strongly feel that phylogenetic visualisation is critical to grasp the full extent of our planet’s biodiversity. Human activity has carelessly altered almost every aspect of our planet, and we must now live with the shame and hubris of a geologic age we named after ourselves\(^3\). There has never been a greater need to find a way to show humanity our true place in the world. In whatever sense a phylogeneticist can have a duty, I believe it is ours to show the world that we are nothing more than a twig on a tree that we are cutting down.

**Author contributions**

WDP was responsible for all aspects of this work.

**Competing interests**

The author has no competing interests.

**Grant information**

The author(s) declared that no grants were involved in supporting this work.

**Acknowledgements**

I am grateful to Bonnie G. Waring for constructive comments and suggestions, and Eric Talevich for suggesting using Hilbert curves.

**Supplementary files**

**Supplementary File 1:** R code for fibre plot. Click here to access the data.

**Supplementary File 2:** Mammals.gif. Animated fibre plot in Graphics Interchange Format (GIF) of a phylogeny of all extant mammals\(^12\). Click here to access the data.

**References**

Open Peer Review

Current Referee Status:  ?  ?  ?

Version 2

Referee Report 17 February 2017
doi:10.5256/f1000research.11550.r19938

Diego San Mauro, Ainhoa Agorreta
Department of Zoology and Physical Anthropology, Complutense University of Madrid, Madrid, Spain

This paper presents a brief overview on phylogenetic visualization and introduces a novel approach for visualizing phylogenies (timetrees) using fibre plots. Given the rapid accumulation of phylogenetic information over the last years that has enabled the construction of massive trees (mega-phylogenies) containing millions of branches and leaves (taxa), the new visualization method appears to be interesting and with some potential. However, it is outlined only succinctly in the paper, and we feel that there are a few issues that require further attention. More discussion is needed on the specific applications and/or implementations of phylogenetic fibre plots compared to the other visualization approaches already available. For instance, what are the advantages of fiber plots over conventional phylogenetic plots in terms of comparing e.g. different topology sets? (as used for example in hypothesis testing). Also, what is the applicability (if any) of fiber plots for visualizing phylogenetic trees whose branches represent rate of evolution (e.g., substitutions/site) instead of time? (as in phylograms). Or, how do fibre plots deal with extinct branches? (as those displayed by extinct fossil lineages). Discussing these issues (among others) more in detail would make it easier for the reader to assess the breadth of novelty and usefulness of the new method for the general field of phylogenetics, and its applicability beyond the reconstruction of the timetree of life. As described in the current paper, it seems that fibre plots could be a complement, but not substitute of the other (more conventional) phylogenetic visualization approaches. The output of the fibre plot is colorful, but in general very difficult to interpret. In fact, interpreting the fibre plot output of very large phylogenies or even the tree of all life would be more difficult than interpreting more conventional approaches (those zooming in and out the phylogeny). Implementing some sort of labeling/cross-referencing with lists of taxa or even conventional phylogenetic trees live on the side could help in the precise interpretation of what is being displayed at each timeframe.

There are also some additional issues that we want to mention:

- First paragraph: The sentence beginning "Many have..." needs some rewording... It is true that many have produced tree-like depictions of the relationships among species, but certainly not many before Darwin. So, please reword.

- Second paragraph: Please provide a reference for the sentence beginning "Haeckel grasped that humans...".

- Third paragraph: Besides Dendroscope, it would be fair to cite FigTree (http://tree.bio.ed.ac.uk/software/figtree/) as well in the last sentence.
Fourth paragraph: It would be good to cite and discuss Walrus (http://www.caida.org/tools/visualization/walrus/) here as well. It appeared in 2001 (earlier than Paloverde) and allowed interactive 3D visualization of hierarchical graphs.

Fifth paragraph: Please add references and expand the last statement about using Hilbert curves.

Last paragraph: The last paragraph of the paper appears unnecessary and probably should be removed. Only the first sentence could be kept as part of the previous paragraph (as closing statement). If this sentence is retained, please keep in mind that phylogenies (e.g., the tree of life) are hypotheses. Therefore, it would be more appropriate to say "...being closer than ever to a reliable tree of all life", rather than "...being closer than ever to a true tree of all life".

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 13 February 2017

doi:10.5256/f1000research.11550.r19876

Stephen A. Smith
Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

Pearse presents a new means for visualizing large phylogenies called the fibre plot. The purpose of this plot is to better represent splits by different colors and shades. This is an interesting idea and is demonstrated with an accompanying animated gif. However, I am left wondering if there are significant insights gathered from this view of the mammal tree. The animation proceeds and areas of the graph change color. I understand why they change but don’t know where I am in the tree and what the significance of the change is in speed or area of the graph. The figure presented (Figure 1) shows a somewhat different view of the fibre plot as presented alongside the phylogeny. This makes me think that perhaps a more informative presentation would be the view of the phylogeny along with the fibre plot. Then the animation would follow a line that moves in a preorder fashion from the root to the tips. This would allow for a more direct comparison of the tree and the plot. Without this additional guide, I am not sure what to make of the animation. I don’t know where I am in the tree (in time or place) and I can’t "move around" in any particular way. I can also envision any number of statistics presented with the plot.

This is an interesting start of an idea but I think it needs a little more development before it would be useful for navigating the size of the tree intended by the author. However, there may be some interesting uses for this or something like it in the future.

Editorial comments
I recommend that the author edit the abstract. For example, the sentence “Despite the progress that has been made in the visualisation of information since Haeckel's time, phylogenetic visualisation has moved forward remarkably little.” seems to suggest that Haeckel was the first person to try and visualize data. While this may be accurate for some biological data, it is not true for data in general as cartographers have been trying to visualize information and data for centuries. The final sentence in the paragraph could
also use some adjustments. While the statement is trying to convey a general sense of the importance of phylogenies, I am not certain that “our place” in the tree of life will dramatically change as a result of visualization of the data. I would also recommend changes to the intervening sentences.

I would recommend some changes to the remaining text but won’t outline all of those here.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Competing Interests:** No competing interests were disclosed.

Referee Report 10 February 2017

doi:10.5256/f1000research.11550.r19939

Rafael Zardoya

Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales, Madrid, 28006, Spain

The paper entitled “Animating and exploring phylogenies with fibre plots” by Pearse is an interesting contribution that proposes a new and distinct way to visualize phylogenetic trees. The new method propose by the author uses fibre plots to slice a phylogenetic tree from root to tips and visualize, as an animation, the cladogenetic process in time.

As the author correctly argues, while it is now possible to reconstruct phylogenetic trees involving tens of thousands of species, visualization of such trees is complex and has not advanced at the same pace as probabilistic inference methodology. Hence, the challenge is set. There are many programs for visualizing trees but few have explored the need of dealing with large phylogenies. Different strategies have been proposed to represent phylogenies including the collapse of certain nodes, distortion of the view, and representation in 3D, but thus far, the most popular approach probably consists on zooming in and out the phylogeny (OneZoom, Rosindell et al. 2012) using appropriate tools (e.g., a tablet). These viewers are complemented with others that allow incorporating other information pertinent to the phylogeny (e.g., iTOL, Letunic and Bork 2016).

The proposal here presented explores in a very different direction. While the idea of looking at different temporal slices in the phylogeny to get a feeling of the timing of diversification of the different clades is original, I think it is too preliminary in the present contribution. The video composed of successive slices shows in different colors how a single (ancestral) lineage is successively split into many but the viewer is unable to discern to which exact descendant lineages is looking at, as there are no labels. Moreover, at some point the number of splits (and colors) is too large to obtain useful information from the animation. As presently devised, the analysis of different clades will render very similar plots, which will be difficult to interpret (beyond seen an increase in the number of lineages) and compare. If the author wants this tool to be widely used, he should make the final outcome more appealing and understandable (e.g., perhaps a grid plot with labels of each lineage in the corresponding axis would help following which lineages and their ancestors are diverging) by peers from other fields than phylogenetics and by the general public.

**Minor changes:**
• The author mentions that PaloVerde in 2006 was the first 3D phylogeny viewer to his knowledge. I think he should check the Walrus graph visualisation tool by Hughes et al. 2004
• Close brackets after [see 6]

References

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

**Competing Interests:** No competing interests were disclosed.