

Your perspective

The absence of a formal reporting standard for phylogenetic analyses is a major impediment for digital access and reuse of published gene trees and species trees. Efforts are underway to develop a standard for Minimal Information About Phylogenetic Analyses (MIAPA). An important part of this process is community input on metadata - what is important for use and evaluation, and what is reasonable to expect from producers of trees?

Results from this survey will inform two efforts: the collection of digital phylogenetic data for Open Tree of Life (<http://opentreeoflife.org>) and the development of a minimum information standard for reporting phylogenetic analyses (MIAPA, <http://www.evoio.org/wiki/MIAPA>). If you have any questions, please contact Karen Cranston, National Evolutionary Synthesis Center (karen.cranston@nescent.org).

Are you a producer of trees or a user of trees, or both?

- ☐ Producer of trees: you infer phylogenies, publish trees
- ☐ User of trees: you view phylogenies, interpret them, or use them as inputs for computations, e.g., ancestral reconstruction
- ☐ Both producer and user of trees

What is your career stage?

- ☐ Graduate student
- ☐ Postdoctoral fellow
- ☐ Faculty
- ☐ Other

Select 1-2 terms that best describe your field of study

- | | |
|---|---|
| <input type="checkbox"/> Systematics | <input type="checkbox"/> Biodiversity informatics |
| <input type="checkbox"/> Phylogenetic methods | <input type="checkbox"/> Taxonomy |
| <input type="checkbox"/> Biogeography | <input type="checkbox"/> Genomics / metagenomics |
| <input type="checkbox"/> Ecology | <input type="checkbox"/> Other <input type="text"/> |

INTERPRETING and EVALUATING TREES

Please order the following metadata by how important it is to interpret, evaluate, and use trees produced by others.

CRITICAL = without this information, you would not (or could not) use a tree

IMPORTANT = it would make use much easier, or allow use with more confidence

USEFUL = could think of uses for this data, but I would use a tree without it

	Critical	Important	Useful
Informative text labels that identify the entity at the tips of the tree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identifiers that link a tip to an entity in some external taxonomy or database	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

or database

The source of biological material (tissue, specimen, etc) for data at the tips

☐☐☐

Rooted vs unrooted topology

☐☐☐

Type of tree estimate: maximum-likelihood, majority-rule consensus, etc.

☐☐☐

Geographic coordinates indicating where specimens were collected

☐☐☐

Branch lengths

☐☐☐

Units for branch lengths (time, expected # of substitutions, etc)

☐☐☐

Data matrix used to infer the tree

☐☐☐

What method was used to infer the tree (likelihood, parsimony, distance, bayesian, etc)

☐☐☐

Name and version of software used to infer the tree

☐☐☐

Input parameters used to infer the tree

☐☐☐

Support values

☐☐☐

Type of support values (bootstrap, bayesian, etc)

☐☐☐

Full list of samples used to calculate support values

☐☐☐

Alignment method (progressive, iterative, HMM, etc)

☐☐☐

Name and version of software used to align sequences

☐☐☐

Input parameters used to align the data

☐☐☐

Was alignment manually adjusted?

☐☐☐

Comments about any of the above elements for evaluating trees? Are there elements missing from this list that you would want with a tree?

PUBLISHING and ARCHIVING TREES

Thinking about archiving or databasing your trees to make them accessible to other scientists, please categorize the following metadata in terms of how difficult it would be for you to provide.

EASY = included in the standard input or output files for your analysis.

MEDIUM = data readily available, but you would need to enter this separately from the input / output files

DIFFICULT = you would have to search for this information and / or would not generally be available for you data

	Easy	Medium	Difficult
--	------	--------	-----------

Informative text labels that identify the entity at the tips of the tree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identifiers that link a tip to an entity in some external taxonomy or database	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The source of biological material (tissue, specimen, etc) for data at the tips	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rooted vs unrooted topology	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Type of tree estimate: maximum-likelihood, majority-rule consensus, etc.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Geographic coordinates indicating where specimens were collected	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Branch lengths	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Units for branch lengths (time, expected # of substitutions, etc)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data matrix used to infer the tree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What method was used to infer the tree (likelihood, parsimony, distance, bayesian, etc)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Name and version of software used to infer the tree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Input parameters used to infer the tree	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Support values	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Type of support values (bootstrap, bayesian, etc)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Full list of samples used to calculate support values	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Alignment method (progressive, iterative, HMM, etc)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Name and version of software used to align sequences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Input parameters used to align the data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Was alignment manually adjusted?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments about providing any of the above elements when databasing / archiving trees? Are there elements missing from the list that you would want to provide?