

[Around meta-analysis \(9\): getting a phylogenetic tree](#)

31/5/2016

By Malgorzata Lagisz

After a long blogging break, due to moving from Dunedin, NZ, to Sydney, AU, I am re-starting my series of blogs related to meta-analysis. Earlier blog posts from this series are archived as pdf files on my personal webpage (mlagisz.weebly.com)

Meta-analysis on data from multiple species often requires taking phylogenetic relatedness among these species into account. For meta-analysis you should use a phylogenetic / taxonomic tree that is:

1. binary (strictly bifurcating = no polytomies),
2. rooted (most basal ancestor is specified),
3. branch lengths (divergence measures) are optional,
4. tip labels on the tree match exactly with species names in the meta-analytic data set.

There are several ways of obtaining a tree for a set of species in hand. Here, I present the simplest way, which should work in most cases, unless there are some rarely investigated species in your dataset.

Until recently we have usually used the [Interactive Tree of Life](#), but the service "other trees" got discontinued there. Instead, we can now use [NCBI Taxonomy Common Tree](#).

Here are some tips on how to use it:

- The input box at the top of the page allows searching for each species name separately, but it is quicker to upload a list of the binomial Latin species names from a text file (.txt). Press "Browse" button and choose your .txt file (one species name per row, with spaces between name parts, no other characters needed). Press "Add from file" button. If successful, a taxonomic tree of the uploaded species will appear after a while, depending on the tree size.
- To export the tree, select "phylip tree" from the "text tree" box and then press "Save as" button, which will prompt saving the tree as a binary file. The resulting text file "phyliptree.phy" should contain Newick-like tree with node labels and taxonomic distances.
- Using text editor remove all " ' ", and replace spaces " " with underscores "_". Now, the file can be uploaded into R using ape package.
- After obtaining the scaffold taxonomic tree from NCBI you can continue working in R to plot the tree, resolve polytomies and add missing species (e.g. by grafting), if needed.

- Polytomies can be resolved using existing phylogenetic information from other sources (e.g. published phylogenies) or at random (generally not recommended).
- When making the tree there are a few common problems with species names that you should be aware of: typos, scientific name synonyms, subspecies names. These problems may result in failure to retrieve the tree from the online databases. Thus, if any errors occur in the process, names need to be re-checked, corrected or substituted with synonyms or close relatives.

Good luck!