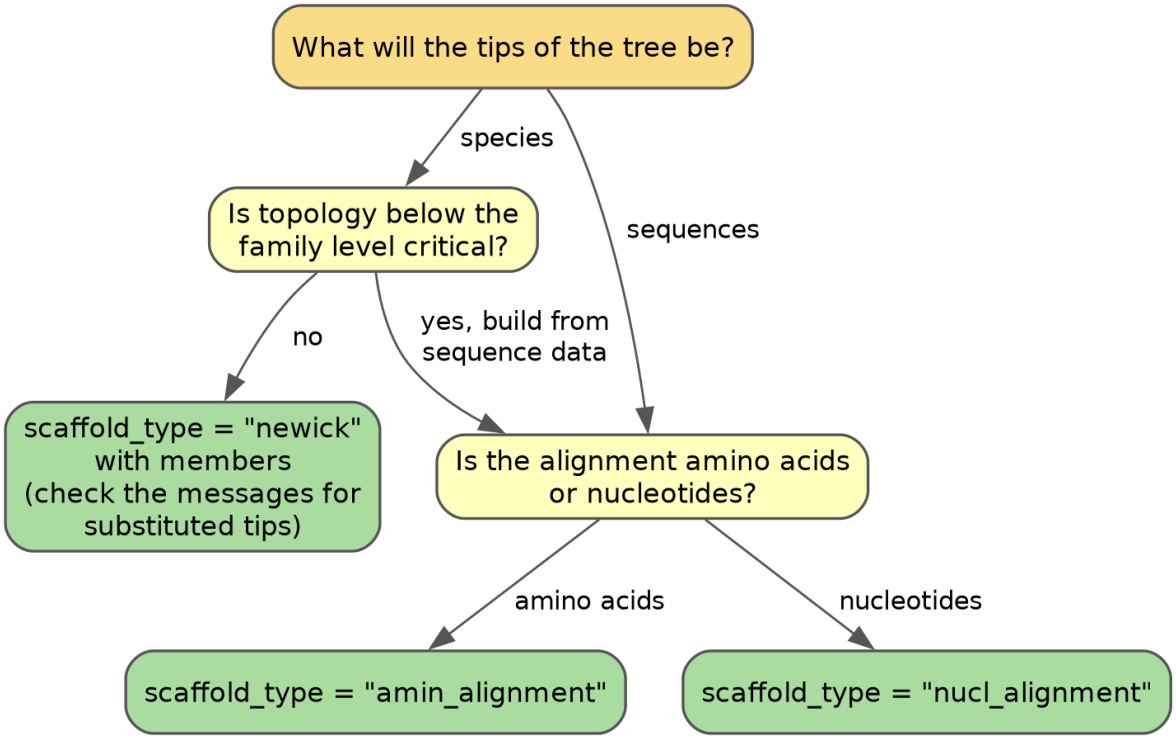
phylogenies: summary

Integrated Bioanalytical Chemistry · thebustalab.github.io/integrated_bioanalytics



Building a tree starts with one decision: **what the tree is built from.** A newick scaffold is quick for species when only the relationships above the family level matter, and `buildTree()` grafts a species missing from the scaffold onto a relative's tip, so its messages are worth reading. When genus-level topology matters, or when the tips are genes, the tree should come from a sequence alignment instead, by neighbor joining or by maximum likelihood. Once built, a tree is plotted with `ggtree()`, joined to data after `fortify()`, collapsed to a higher level with `collapseTree()`, and drawn beside its traits, where the trait axis must be reordered to match the tree and the tree tips and trait labels checked by eye to confirm they line up.

Function	What it is for	Settings students vary
`buildTree()` from a newick scaffold	a tree of species pruned from an existing phylogeny	`scaffold_type = "newick"`, `scaffold_in_path`, `members`
`buildTree()` from an alignment	a tree built from aligned sequences	`scaffold_type` (`"amin_alignment"` or `"nucl_alignment"`), `scaffold_in_path`, `ml` (TRUE for maximum likelihood, FALSE for neighbor joining), `model_test`, `bootstrap`, `ancestral_states`, `root`
`plot()`	a quick look at a small tree	none
`ggtree()`	plotting a tree with ggplot layers	add `geom_tiplab()` with `offset`, `align`, `size`
`fortify()` + `left_join()`	turning a tree into a data frame and adding metadata by tip label	`by = c("label" = ...)`
`collapseTree()`	one tip per higher group (for example, per family)	`tree`, `associations` (tip labels, then group)
`reorder(label, y)`	putting trait rows in the same order as the tree tips	the column to reorder and the tree's y
`plot_grid()`	the tree and the trait plot side by side	`nrow = 1`, `align = "h"`, `axis = "tb"`