Taxonomic distinctness/ aggregation files

A later section (15) discusses univariate diversity indices that can be computed from each sample, including biodiversity measures that are based on the relatedness of the species making up a simple species list (P/A data), see Chapter 17 of CiMC. Though the supplied relatedness could be genetic, phylogenetic or functional – through suitable provision of a distance/dissimilarity matrix among the species, perhaps (but not necessarily) their pairwise distances apart through some hierarchical arrangement of species – PRIMER 7 implements the idea mainly in terms of *taxonomic distinctness* (see Section 15). These are the distances travelled in connecting every pair of species through a tree with a fixed set of levels (typically, a Linnaean taxonomy). If, on average, these distances are large, then the sample is considered biodiverse. A necessary input is a *variable information* sheet, which (for historic reasons) PRIMER calls an *aggregation* file (see the end of Section 2), defining the taxonomy – which species belong to which genera, families, orders, etc. From this, path weights ω_{ij} are calculated between every pair of species, i and j. Always, ω_{ij} takes the value 100 for two species that are connected at the most distant level; e.g. if the final column heading in the taxonomy file is *phylum* then two species in different phyla are defined to be 100 units apart (do not add a final column, say *kingdom*, for which all species have the same entry, Animalia; you could then only attain the value 100 for species in different kingdoms). By default, intervening levels are considered to be equally-spaced. For example, for a hierarchy of species from different classes all in the same phylum, with the five levels of species, genus, family, order and class, two species in the same genera are 20 units apart, in different genera but the same family are 40 units apart, etc. This can be overruled in two ways: either a user can define his/her own step branch-lengths, which will again be rescaled to a maximum of 100 for two species in different top-level groups, whatever scale is input for the absolute steps; or the information in the aggregation matrix about taxon richness at each hierarchical level can be used (a level in the tree which has almost as many taxa as the level below it gives rise to a step of shorter branch-length).

Revision #5
Created 10 June 2024 23:17:06 by Arden
Updated 15 January 2025 02:52:58 by Abby Miller