

## **A brief Introduction to Online Program iDIP (Information-based Diversity Partitioning) Under Multi-Level Hierarchical Structures for Shannon and Phylogenetic Diversities**

### Application iDIP

**Anne Chao and Chun-Huo Chiu**

*Institute of Statistics, National Tsing Hua University, Hsin-Chu, Taiwan 30043*

### **What this online tool does**

The Online Program iDIP features two analyses of decomposing information-based diversity under a specified multi-level hierarchical structure:

- (1) Shannon diversity: this analysis is used to decompose Shannon diversity (the exponential of Shannon entropy) across multiple levels given specified raw or relative species/allele (rows) abundances in each population/community (columns).
- (2) Phylogenetic diversity: this analysis is used to decompose information-based phylogenetic diversity across multiple levels given specified raw or relative species/allele (rows) abundances in each population/community (columns) and a specified phylogenetic tree (in Newick tree format) spanned by all observed species.

For each of the two analyses, Program iDIP can be applied to any arbitrary number of hierarchical levels. The output consists of a basic data summary and decomposition results, with the latter including (1) gamma (or total) diversity; alpha and beta diversity at each level; (2) proportion of total beta information (Shannon information) found at each level; (3) mean differentiation (dissimilarity) among aggregates at each level. Illustrative examples (based on demo data included in the tool) and their corresponding output interpretations are provided in the User's Guide ([link](#)). Full details about the methodology in the context of a real-world data analysis can be found in the following paper:

Gaggiotti, O. E., Chao, A., Peres-Neto, P., Chiu, C.-H., Edwards, C., Fortin, M.-J., Jost, L., Richards, C. M., and Selkoe, K. A. (2017). Diversity from genes to ecosystems: a unifying framework to study variation across biological metrics and scales. *Evolutionary Applications* Special Issue.

### Raw abundances or relative abundances

iDIP Online supports two kinds of data: *raw or relative* species/allele abundances. When raw species/allele abundances (e.g., number of individuals or copies of alleles) are uploaded, size-weights are used by default, i.e., the proportion of the total number of individuals/alleles in a population/community is taken to be the weight of that population/community.

When relative species/allele abundances are uploaded, the “size” for each population/community is unity, leading to equal-weights for all populations or communities. Typically, the two weight functions yield very close results.

### How to use this tool

Step 1. Select an analysis type ( or ) from the top menu of the iDIP Online window.

#### Step 2.

(2a) If you select , then choose either demo data (raw-abundance data or relative-abundance data) or upload your own data. The latter choice requires that the following two matrices be uploaded as txt files: (i) the abundance data matrix specifying raw or relative species/allele (rows) abundances in each population/community (columns). You can include species names as row names, though it is not necessary to do so for Shannon Diversity Analysis. (ii) You also need to upload a matrix specifying a hierarchical structure; see the simple example given in the iDIP User’s Guide ([link](#)) for how to input a hierarchical structure.

(2b) If you select , then choose either demo data (raw-abundance data or relative-abundance data) or upload your own data. For the latter, three matrices should be uploaded as txt files. In addition to the abundance data matrix and hierarchical structure matrix, as described in Step (2a), a phylogenetic tree (in Newick tree format) spanned by all observed species or

alleles must be uploaded. Note that for decomposing phylogenetic diversity, species/allele names (row names) must be included in the abundance data matrix and the species/allele names must match those in the uploaded phylogenetic tree.

Step 3. Press the **Run!** button. The output is then shown under “Data Summary” and “Decomposition analysis” along the second row menu.

### **Output, Brief Introduction and User’s Guide**

Along the second row menu, there are four output selection tabs:

- In the “Data Summary” tab panel, the output consists of the plot of the underlying hierarchical structure, the phylogenetic tree plot (only for Phylogenetic Diversity analysis) and basic data information. You can click “save as csv file” at the bottom of the output to download the basic data information.
- In the “Decomposition Analysis” tab panel, the numerical values for gamma diversity, alpha and beta components for each level, proportion of total beta information found in each level, and mean differentiation (dissimilarity) among sampling units for each level are shown. You can click “save as csv file” at the bottom of the output to download the output.
- In the “Introduction” panel, users can view a brief introduction which includes a summary of the running procedures.
- In the “User Guide” panel, a link will direct users to the iDIP User’s Guide.

To gain familiarity with the program, we suggest that users first run the demo data sets included in the program and check the output with that given in the iDIP User’s Guide ([link](#)). Part of the output for each example is also interpreted in the User’s Guide to help users understand the numerical results.