**R code for constructing the figures in the main text.**

First, please install and library the following packages

library(chaoUtility)

library(readr)

library(ape)

library(dplyr)

library(ggplot2)

library(Rcpp)

library(reshape2)

library(ggpubr)

library(gg.gap)

library(devtools)

the main functions used in main text are implemented in R language and depend on the unpublished package “iNEXT3D”.

install\_github("AnneChao/iNEXT3D")

library(iNEXT3D)

and upload some specific functions to make the figures in the main text

source("figure-making.txt")

####################################################################

There are three main functions of iNEXT.3D package used to construct the figures in the main text. They are “estimate3D”, “Asy3D” and “Obs3D”.

1. “estimate3D” to performs interpolation and extrapolation of Hill number with order q.

estimate3D(data, class = 'TD', q = c(0,1,2), datatype = "abundance", base = "coverage", level = NULL, nboot=50, conf = 0.95, tree, nT, reftime = NULL, PDtype = 'PD', distM, threshold = NULL)

1. “Asy3D” to performs the estimated (asymptotic) diversity of order q.

Asy3D(data, class = " TD", q = seq(0, 2, 0.2), datatype = "abundance", nboot = 50, conf = 0.95, tree, nT, reftime = NULL, PDtype = "PD", distM, threshold = NULL)

1. “Obs3D” to performs the empirical (observed) diversity of order q.

Obs3D(data, class = "TD", q = seq(0, 2, 0.2), datatype = "abundance", nboot = 50, conf = 0.95, tree, nT, reftime = NULL, PDtype = "PD", distM, threshold = NULL)

The functions have the following usage:

|  |  |
| --- | --- |
| **parameter** | **explanation** |
| data | a matrix, data.frame (species by sites), or list of species abundances or incidence frequencies. If datatype = "incidence\_freq", then the first entry of the input data must be total number of sampling units in each column or list. |
| class | a choice of three-level diversity: ’TD’ = ’Taxonomic’, ’PD’ = ’Phylogenetic’, and ’FD’ = ’Functional’ under certain threshold. Besides, ’AUC’ is the fourth choice which integrates several threshold functional diversity to get diversity. |
| q | a numerical vector of the order of Hill number. |
| datatype | data type of input data: individual-based abundance data (datatype = "abundance"), sampling-unit-based incidence frequencies data (datatype = "incidence\_freq") or species by sampling-units incidence matrix (datatype = "incidence\_raw"). |
| base | comparison base: sample-size-based (base="size") or coverage-based (base="coverage"). |
| level | a sequence specifying the particular sample sizes or sample coverages(between 0 and 1). If base="size" and level=NULL, then this function computes the diversity estimates for the minimum sample size among all sites extrapolated to double reference sizes. If base="coverage" and level=NULL, then this function computes the diversity estimates for the minimum sample coverage among all sites extrapolated to double reference sizes. |
| conf | a positive number < 1 specifying the level of confidence interval, default is 0.95. |
| nboot | an integer specifying the number of replications. |
| tree | a phylo object describing the phylogenetic tree in Newick format for all observed species in the pooled assemblage. It is necessary when class = 'PD'. |
| nT | needed only when datatype = "incidence\_raw", a sequence of named nonnegative integers specifying the number of sampling units in each assemblage. If names(nT) = NULL, then assemblage are automatically named as "assemblage1", "assemblage2",..., etc. It is necessary when class = 'PD' and datatype = "incidence\_raw". |
| reftime | is set to be the tree depth of the phylogenetic tree, which is spanned by all the observed species in the pooled assemblage. Default is NULL. It will be use when class = 'PD'. |
| PDtype | desired phylogenetic diversity type: PDtype = "PD" for Chao et al. (2010) phylogenetic diversity and PDtype = "meanPD" for mean phylogenetic diversity (phylogenetic Hill number). It will be use when class = 'PD'. Default is "PD". |
| distM | a pair wise distance matrix for all pairs of observed species in the pooled assemblage. It will be use when class = 'FD' or 'AUC'. |
| threshold | a sequence between 0 and 1 specifying tau. If NULL, threshold = dmean. Default is NULL. It will be use when class = 'FD'. |

If you want to get more details, use '?iNEXT3D', '?estimate3D', '?Asy3D', '?Obs3D' to obtain help.

#####################################################################

Application on the fish data set:

1. **Figure1**

The yearly temporal diversity patterns for three types of diversity with order q=0, 1, and 2 based on yearly fish abundance data at four different levels which including asymptotic estimates, standardized estimates under a coverage value of *Cmax*, observed diversity, and standardized estimates under a coverage of *Cmin*.

**1a. Construct Figure 1a (yearly taxonomic diversity pattern)**

One type of dataset is needed: yearly species abundance data (“Fish Abundance data.csv”).

Here, species abundance data is recorded as a species(row) by plot(colnum) data matrix, and

each element in the matrix is the number of individuals.

First, copy the file “Fish Abundance data.csv” in your working directory.

and load(or import) the data set in your R console:

Abun <- read.csv("Fish Abundance data.csv", row.names = 1, header= TRUE)

and then execute following instructions to obtain Cmin and Cmax, where

Cmin is the smallest observed coverage among all assemblages, and

Cmax is the smallest coverage after extrapolating double sample size among all assemblages.(see Chao et al.(2019) for detail)

Cmax <- apply(Abun, 2, function(x) iNEXT3D:::Chat.Ind(x, 2\*sum(x))) %>% min %>% round(., 4)

Cmin <- apply(Abun, 2, function(x) iNEXT3D:::Chat.Ind(x, sum(x))) %>% min %>% round(., 4)

Then separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of taxonomic diversity at different levels.

1. Use function “estimate3D” to calculate the taxonomic diversity with order q=0, 1, and 2 at the sample coverage of Cmax and Cmin.

TD\_est <- estimate3D(data = Abun, class = 'TD', q = c(0, 1, 2), datatype = 'abundance', base = 'coverage', level = c(Cmin, Cmax), nboot = 0)

1. Use function “Obs3D” to calculate the observed taxonomic diversity with order q=0, 1,and 2.

TD\_obs <- Obs3D(data = Abun, class = 'TD', q = c(0, 1, 2), datatype = 'abundance', nboot = 0)

1. Use function “Asy3D” to calculate the asymptotic taxonomic diversity with order q=0, 1,and 2.

TD\_asy <- Asy3D(data = Abun, class = 'TD', q = c(0, 1, 2), datatype = 'abundance', nboot = 0)

Finally, execute the following two instructions to obtain the Figure 1a.

out\_TD <- rbind(TD\_est %>% select(Assemblage, Order.q, qD, goalSC),

TD\_obs %>% select(Assemblage, Order.q, qD) %>% mutate(goalSC = 'Observed') ,

TD\_asy %>% select(Assemblage, Order.q, qD) %>% mutate(goalSC = 'Asymptotic') )

fig\_13(out\_TD, y\_label = 'Taxonomic diversity')

**1b. Construct Figure 1b (yearly phylogenetic diversity pattern)**

Two types of dataset are needed: yearly species abundance data (“Fish Abundance data.csv”), and species phylogenetic tree data (“Fish PhyTree.txt”).

Here, species abundance data is recorded as a species(row) by plot(colnum) data frame, and

each element in the data frame is the number of individuals.

Species phylogenetic tree data should be in Newick format.

First, copy the files “Fish Abundance data.csv” and "Fish PhyTree.txt" in your working directory and load(or import) the data sets in your R console:

Abun <- read.csv("Fish Abundance data.csv", row.names = 1, header= TRUE)

tree <- read.tree("Fish PhyTree.txt")

Then, separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of phylogenetic diversity at different levels.

1. Use function “estimate3D” to calculate the phylogenetic diversity with order q=0, 1, and 2 at the sample coverage of Cmin and Cmax.

PD\_est <- estimate3D(data = Abun, class = 'PD', tree = tree, datatype = 'abundance', PDtype = 'meanPD', q = c(0, 1, 2), level = c(Cmin, Cmax), nboot = 0, reftime = 1)

1. Use function “Obs3D” to calculate the observed phylogenetic diversity with order q=0, 1, and 2.

PD\_obs <- Obs3D(data = Abun, class = 'PD', tree = tree, datatype = 'abundance', PDtype = 'meanPD', q = c(0, 1, 2), nboot = 0, reftime = 1)

1. Use function “Asy3D” to calculate the asymptotic phylogenetic diversity with order q=0, 1,and 2.

PD\_asy <- Asy3D(data = Abun, class = 'PD', tree = tree, datatype = 'abundance', PDtype = 'meanPD', q = c(0, 1, 2), nboot = 0, reftime = 1)

Finally, execute the following two instructions to obtain the Figure 1b.

out\_PD <- rbind(PD\_est %>% select(Assemblage, Order.q, qPD, goalSC),

PD\_obs %>% select(Assemblage, Order.q, qPD) %>% mutate(goalSC = 'Observed'),

PD\_asy %>% select(Assemblage, Order.q, qPD) %>% mutate(goalSC = 'Asymptotic') )

fig\_13(out\_PD, y\_label = 'Phylogenetic diversity')

**1c. Construct Figure 1c (yearly functional diversity pattern)**

Two types of dataset are needed: yearly species abundance data(“Fish Abundance data.csv”), and species functional trait data (“Fish Traits.csv”).

Here, species abundance data is a species(row) by plot(colnum) data frame, and

each element in the data frame is the number of individuals.

Species functional trait data is a species(row) by trait(colnum) data frame and each element in the data frame is the mean trait value of species.

First, copy the files “Fish Abundance data.csv” and " Fish Traits.csv " in your working directory and load(or import) the data sets in your R console:

Abun <- read.csv("Fish Abundance data.csv", row.names = 1, header= TRUE)

traits <- read.csv("Fish Traits.csv", row.names = 1, header= TRUE)

and execute following instructions to obtain functional distances among species by using Gower metrics.

for (i in 1:ncol(traits)) { if (class(traits[,i]) == "character") traits[,i] <- factor(traits[,i], levels =unique(traits[,i]))}

distM <- cluster::daisy(x = traits, metric = "gower") %>% as.matrix()

Then, separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of functional diversity at different levels.

1. Use function “estimate3D” to calculate the functional diversity with order q=0, 1, and 2 at the sample coverage of Cmin and Cmax, separately.

FD\_est <- estimate3D(data = Abun, class = 'AUC', distM = distM, q = c(0, 1, 2), level = c(Cmin, Cmax), nboot = 0, datatype = 'abundance')

1. Use function “Obs3D” to calculate the observed functional diversity with order q=0, 1, and 2.

FD\_obs <- Obs3D(data = Abun, class = 'AUC', distM = distM, q = c(0, 1, 2), datatype = 'abundance', nboot = 0)

1. Use function “Asy3D” to calculate the asymptotic functional diversity with order q=0, 1,and 2.

FD\_asy <- Asy3D(data = Abun, class = 'AUC', distM = distM, datatype = 'abundance', q = c(0, 1, 2), nboot = 0)

Finally, execute the following two instructions to obtain the Figure 1c.

out\_FD <- rbind(FD\_est %>% select(Assemblage, Order.q, qFD = qAUC, goalSC),

FD\_obs %>% select(Assemblage, Order.q, qFD = qAUC) %>% mutate(goalSC = 'Observed') ,

FD\_asy %>% select(Assemblage, Order.q, qFD = qAUC) %>% mutate(goalSC = 'Asymptotic'))

fig\_13(out\_FD, y\_label = 'Functional diversity')

1. **Figure 2**

Since Figure1 and Figure2 use identical running outputs, and have different types of presentation.

Figure1 presents the yearly diversity pattern with order q=0,1,2 in the same figure for TD, PD, and TD.

Figure2 presents the yearly patterns of TD, PD, and FD in the same figure for each order q.

Construct Figure 2a by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 0)

Construct Figure 2b by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 1)

Construct Figure 2c by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 2)

1. Figure 3

The three yearly temporal diversity patterns for three types of diversity with order q=0, 1, and 2 based on three-year fish incidence data at four different levels which including asymptotic estimates, standardized estimates under a coverage value of *Cmax*, observed diversity, and standardized estimates under a coverage of *Cmin*.

**3a. Construct Figure 3a (three yearly taxonomic diversity pattern)**

One type of dataset is needed: three yearly species incidence data ("Fish Incidence frequency data.csv").

First, copy the file "Fish Incidence frequency data.csv" in your working directory.

and load(or import) the data set in your R console:

Inci\_freq <- read.csv("Fish Incidence frequency data.csv", row.names = 1, header= TRUE)

and then execute following instructions to obtain Cmin and Cmax, where

Cmin is the smallest observed coverage among all assemblages, and

Cmax is the smallest coverage after extrapolating double sample size among all assemblages.(see Chao et al.(2019) for detail)

Cmax <- apply(Inci\_freq, 2, function(x) iNEXT3D:::Chat.Sam(x, 2\*x[1])) %>% min %>% round(., 4)

Cmin <- apply(Inci\_freq, 2, function(x) iNEXT3D:::Chat.Sam(x, x[1])) %>% min %>% round(., 4)

Then separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of taxonomic diversity at different levels.

1. Use function “estimate3D” to calculate the taxonomic diversity with order q=0, 1, and 2 at the sample coverage of Cmax and Cmin.

TD\_est <- estimate3D(data = Inci\_freq, class = 'TD', q = c(0, 1, 2), datatype = 'incidence\_freq', base = 'coverage', level = c(Cmin, Cmax), nboot = 0)

1. Use function “Obs3D” to calculate the observed taxonomic diversity with order q=0, 1,and 2.

TD\_obs <- Obs3D(data = Inci\_freq, class = 'TD', q = c(0, 1, 2), datatype = 'incidence\_freq', nboot = 0)

1. Use function “Asy3D” to calculate the asymptotic taxonomic diversity with order q=0, 1,and 2.

TD\_asy <- Asy3D(data = Inci\_freq, class = 'TD', q = c(0, 1, 2), datatype = 'incidence\_freq', nboot = 0)

Finally, execute the following two instructions to obtain the Figure 1a.

oout\_TD <- rbind(TD\_est %>% select(Assemblage, Order.q, qD, goalSC),

TD\_obs %>% select(Assemblage, Order.q, qD) %>% mutate(goalSC = 'Observed'),

TD\_asy %>% select(Assemblage, Order.q, qD) %>% mutate(goalSC = 'Asymptotic'))

fig\_13(out\_TD, y\_label = 'Taxonomic diversity')

**3b. Construct Figure 3b (theee yearly phylogenetic diversity pattern)**

Two types of dataset are needed: three yearly species incidence raw data ("Fish Incidence raw data.csv"), and species phylogenetic tree data ("Fish PhyTree.txt").

First, copy the files "Fish Incidence raw data.csv" and "Fish PhyTree.txt" in your working directory and load(or import) the data sets in your R console:

Inci\_raw <- read.csv("Fish Incidence raw data.csv", row.names = 1, header= TRUE)

tree <- read.tree("Fish PhyTree.txt")

Moreover, the number of sampling units in each three yearly incidence data can be obtained by

nT <- as.matrix(Inci\_freq)[1,]

Then, separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of phylogenetic diversity at different levels.

1. Use function “estimate3D” to calculate the phylogenetic diversity with order q=0, 1, and 2 at the sample coverage of Cmin and Cmax.

PD\_est <- estimate3D(data = Inci\_raw, class = 'PD', tree = tree, datatype = 'incidence\_raw', PDtype = 'meanPD', nT = nT, q = c(0, 1, 2), level = c(Cmin, Cmax), nboot = 0, reftime = 1)

1. Use function “Obs3D” to calculate the observed phylogenetic diversity with order q=0, 1, and 2.

PD\_obs <- Obs3D(data = Inci\_raw, class = 'PD', tree = tree, datatype = 'incidence\_raw', PDtype = 'meanPD', nT = nT, q = c(0, 1, 2), nboot = 0, reftime = 1)

1. Use function “Asy3D” to calculate the asymptotic phylogenetic diversity with order q=0, 1,and 2.

PD\_asy <- Asy3D(data = Inci\_raw, class = 'PD', tree = tree, datatype = 'incidence\_raw', PDtype = 'meanPD', nT = nT, q = c(0, 1, 2), nboot = 0, reftime = 1)

Finally, execute the following two instructions to obtain the Figure 3b.

out\_PD <- rbind(PD\_est %>% select(Assemblage, Order.q, qPD, goalSC),

PD\_obs %>% select(Assemblage, Order.q, qPD ) %>% mutate(goalSC = 'Observed'),

PD\_asy %>% select(Assemblage, Order.q, qPD) %>% mutate(goalSC = 'Asymptotic'))

fig\_13(out\_PD, y\_label = 'Phylogenetic diversity')

**3c. Construct Figure 3c (three yearly functional diversity pattern)**

Two types of dataset are needed: three yearly species incidence data("Fish Incidence frequency data.csv"), and species functional trait data ("Fish Traits.csv").

First, copy the files "Fish Incidence frequency data.csv" and "Fish Traits.csv" in your working directory and load(or import) the data sets in your R console:

Inci\_freq <- read.csv("Fish Incidence frequency data.csv", row.names = 1, header= TRUE)

traits <- read.csv("Fish Traits.csv", row.names = 1, header= TRUE)

and execute following instructions to obtain functional distances among species by using Gower metrics.

for (i in 1:ncol(traits)) {if (class(traits[,i]) == "character") traits[, i] <- factor(traits[,i], levels = unique(traits[, i]))}

distM <- cluster::daisy(x = traits, metric = "gower") %>% as.matrix()

Then, separately use “estimate3D”, “Obs3D” and “Asy3D” to obtain the estimates of functional diversity at different levels.

1. Use function “estimate3D” to calculate the functional diversity with order q=0, 1, and 2 at the sample coverage of Cmin and Cmax, separately.

FD\_est <- estimate3D(data = Inci\_freq, class = 'AUC', distM = distM, datatype = 'incidence\_freq', q = c(0, 1, 2), level = c(Cmin, Cmax), nboot = 0)

1. Use function “Obs3D” to calculate the observed functional diversity with order q=0, 1, and 2.

FD\_obs <- Obs3D(data = Inci\_freq, class = 'AUC', distM = distM, datatype = 'incidence\_freq', q = c(0, 1, 2), nboot = 0)

1. Use function “Asy3D” to calculate the asymptotic functional diversity with order q=0, 1, and 2.

FD\_asy <- Asy3D(data = Inci\_freq, class = 'AUC', distM = distM, datatype = 'incidence\_freq', q = c(0, 1, 2), nboot = 0)

Finally, execute the following two instructions to obtain the Figure 3c.

out\_FD <- rbind(FD\_est %>% select(Assemblage, Order.q, qFD = qAUC, goalSC),

FD\_obs %>% select(Assemblage, Order.q, qFD = qAUC) %>% mutate(goalSC = 'Observed'),

FD\_asy %>% select(Assemblage, Order.q, qFD = qAUC) %>% mutate(goalSC = 'Asymptotic'))

fig\_13(out\_FD, y\_label = 'Functional diversity')

1. Figure 4

Since Figure3 and Figure4 use identical running outputs, and show different types of presentation.

Figure3 presents the three yearly diversity pattern with order q=0,1,2 in the same figure for TD, PD, and FD.

Figure4 presents the three yearly patterns of TD, PD, and TD in the same figure for each order q.

Construct Figure 4a by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 0)

Construct Figure 4b by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 1)

Construct Figure 4c by executing the instruction.

fig\_24(TD.output = out\_TD, PD.output = out\_PD, FD.output = out\_FD, q = 2)