

Analysis of scale-dependent biodiversity changes with mobr

Felix May and Dan McGlinn

2017-08-06

Installing mobr

The package `mobr` is currently available on GitHub and you can freely download the source code here [mobr on GitHub](#)

The easiest option is to install the package directly from GitHub using the package `devtools`. If you do not already have `devtools` installed then need to install it.

```
install.packages('devtools')
library(devtools)
```

The package also requires the `dplyr` package which has many dependencies. If you do not already have the `dplyr` package installed we suggest you install it and all of its dependencies using :

```
install.packages(c('bindrcpp', 'glue', 'pkgconfig', 'tibble', 'plyr', 'dplyr'))
```

Then check that `dplyr` can be loaded with `library(dplyr)`. Now you should be ready to go with the `mobr` install

```
install_github('MoBiodiv/mobr')
```

If you receive an error we would love to receive your bug report [here](#)

Data structure required by mobr

To work with `mobr` you need two matrix-like data tables:

1. Species abundance data in a community matrix (rows are sites and columns are species)
2. A site attributes table (rows are sites and columns are site attributes)

The community matrix must include the number of individuals of each species.

The table with the plot attributes can include for instance spatial coordinates of plots, experimental treatments and/or environmental variables. If spatial coordinates are supplied they must be named “x” and “y” for the easting and northing respectively. If temporal trends are of interest rather than spatial you simply set either x or y to a single number and put the temporal measurement for the other coordinate.

Throughout this vignette we use an example of a study on the effects of an invasive plant *Lonicera maackii* on understory plant biodiversity in a Missouri woodland (Powell et al. 2003)

This data set is included in `mobr` and available after loading the library

```

library(mobr)

## Warning: package 'pracma' was built under R version 3.3.3

## Warning: package 'dplyr' was built under R version 3.3.3

data(inv_comm)      # Community matrix
data(inv_plot_attr) # Plot attributes data.frame

str(inv_comm)

##  num [1:100, 1:111] 0 0 0 0 0 0 0 0 0 0 ...

head(inv_plot_attr)

##           group x y
## 1 uninverted 1 0
## 2 uninverted 2 0
## 3 uninverted 3 0
## 4 uninverted 4 0
## 5 uninverted 5 0
## 6 uninverted 6 0

```

The plot attributes include the information if a plot is located in invaded or uninverted sites as well as the spatial xy coordinates of a plot.

Preparing data

In order to analyse the data with `mobr` the two data tables have to be combined into on single object

```

inv_mob_in <- make_mob_in(inv_comm, inv_plot_attr)
inv_mob_in

## Only the first 6 rows of any matrices are printed
##
## $tests
## $N
## [1] TRUE
##
## $SAD
## [1] TRUE
##
## $agg
## [1] TRUE

```

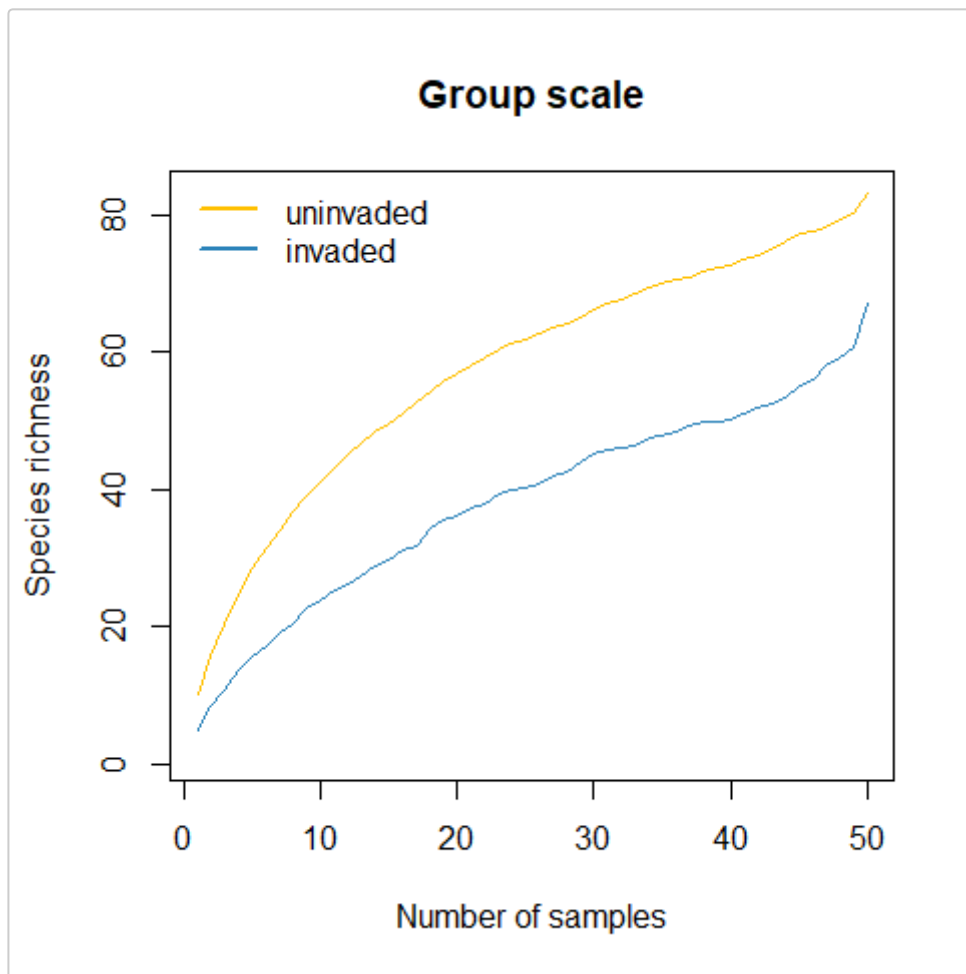
```
##
##
## $comm (Only first 5 species columns are printed)
##   X1 X2 X3 X4 X5
## 1  0  0  1  0  0
## 2  0  0  1  0  0
## 3  0  0  6  0  0
## 4  0  1  2  0  0
## 5  0  0  0  0  0
## 6  0  0  0  0  0
##
## $env
##      group
## 1 uninverted
## 2 uninverted
## 3 uninverted
## 4 uninverted
## 5 uninverted
## 6 uninverted
##
## $spat
##   x y
## 1 1 0
## 2 2 0
## 3 3 0
## 4 4 0
## 5 5 0
## 6 6 0
```

Exploratory data analysis

The package `mobr` offers functions for exploratory data analysis and visualization.

First let's look at the spatial rarefaction curve in which samples are added depending on their spatial proximity.

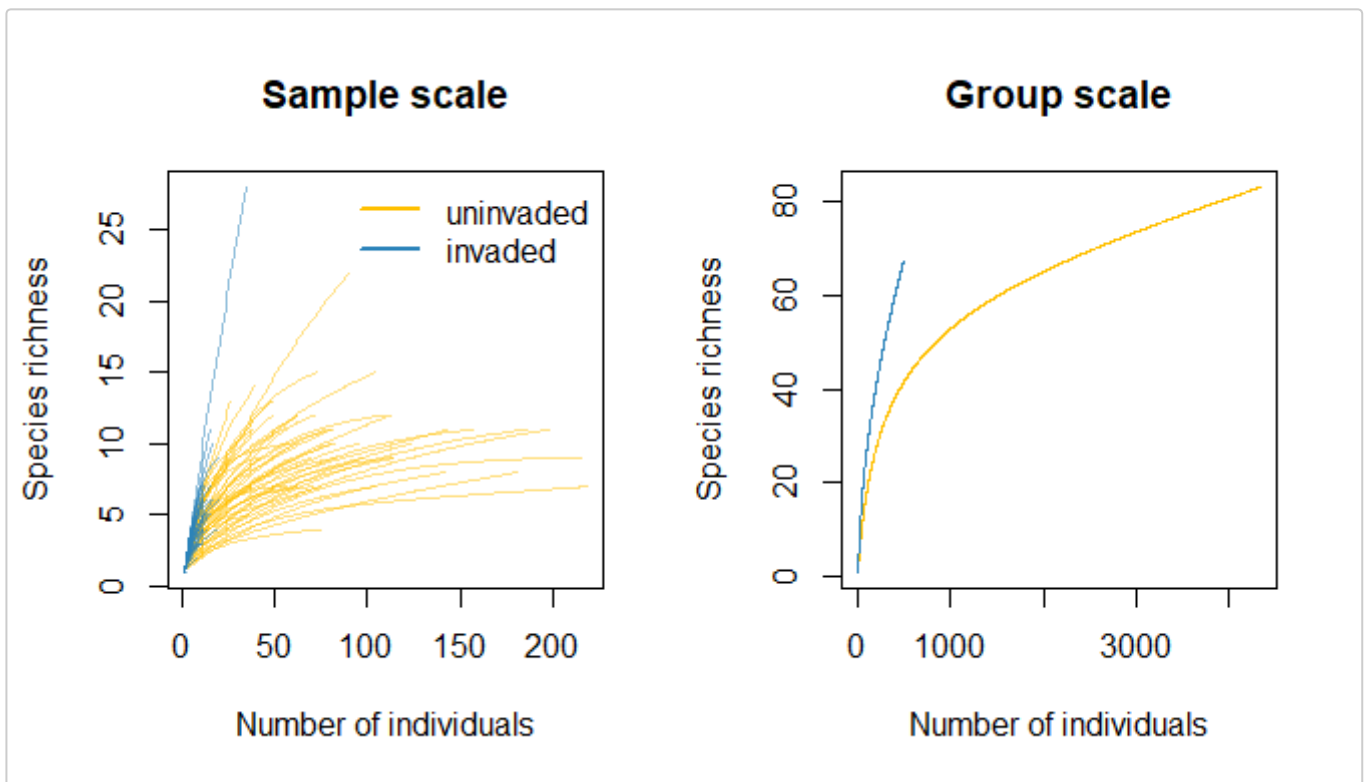
```
plot_rarefaction(inv_mob_in, 'group', 'uninverted', 'spat', lwd=4)
```



We can see that invasion decreases richness and that the magnitude of the effect depends on scale. Let's dig in further to see if we can better understand exactly what components of the community are changing due to invasion.

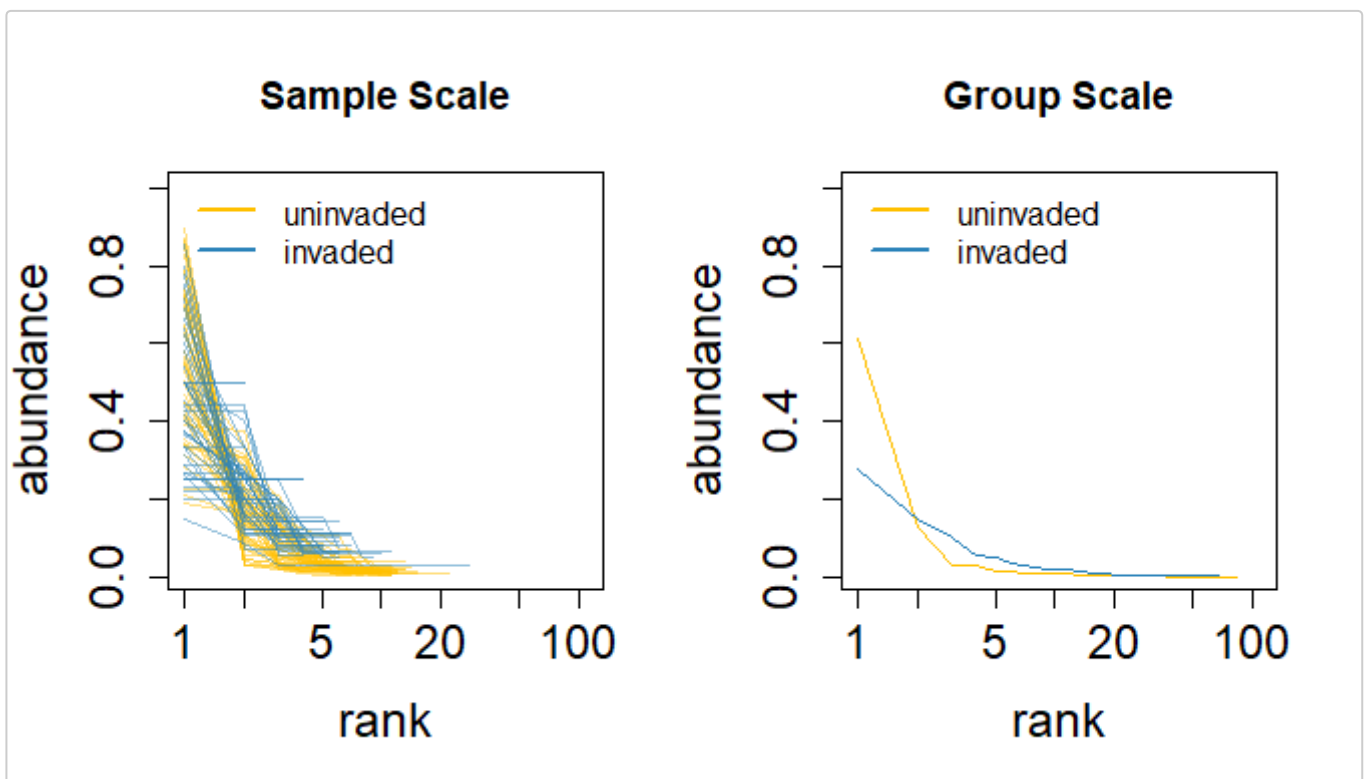
First we look at the individual rarefaction curves for each treatment which only reflect the shape of the SAD (i.e., no individual density or aggregation effects):

```
par(mfrow=c(1,2))
plot_rarefaction(inv_mob_in, 'group', 'uninvaded', 'indiv', pooled=F, lwd=2,
                 leg_loc='topright')
plot_rarefaction(inv_mob_in, 'group', 'uninvaded', 'indiv', pooled=T, lwd=4,
                 leg_loc=NA)
```



Visually you can see there are some big differences in total numbers individuals (i.e., how far on the x-axis the curves go). We can also see that for small numbers of individuals the invaded sites are actually more diverse! This is a bit surprising and it implies that the invaded sites have greater evenness. We can directly examine the species abundance distribution (SAD):

```
par(mfrow=c(1,2))
plot_abu(inv_mob_in, 'group', 'uninvaded', type='rad', pooled=F, log='x')
plot_abu(inv_mob_in, 'group', 'uninvaded', type='rad', pooled=T, log='x')
```



The SADs suggest that there are differences in the SADs where the invaded site has greater evenness in its common species (i.e., flatter left hand-side of rank curve).

Two-scale analysis

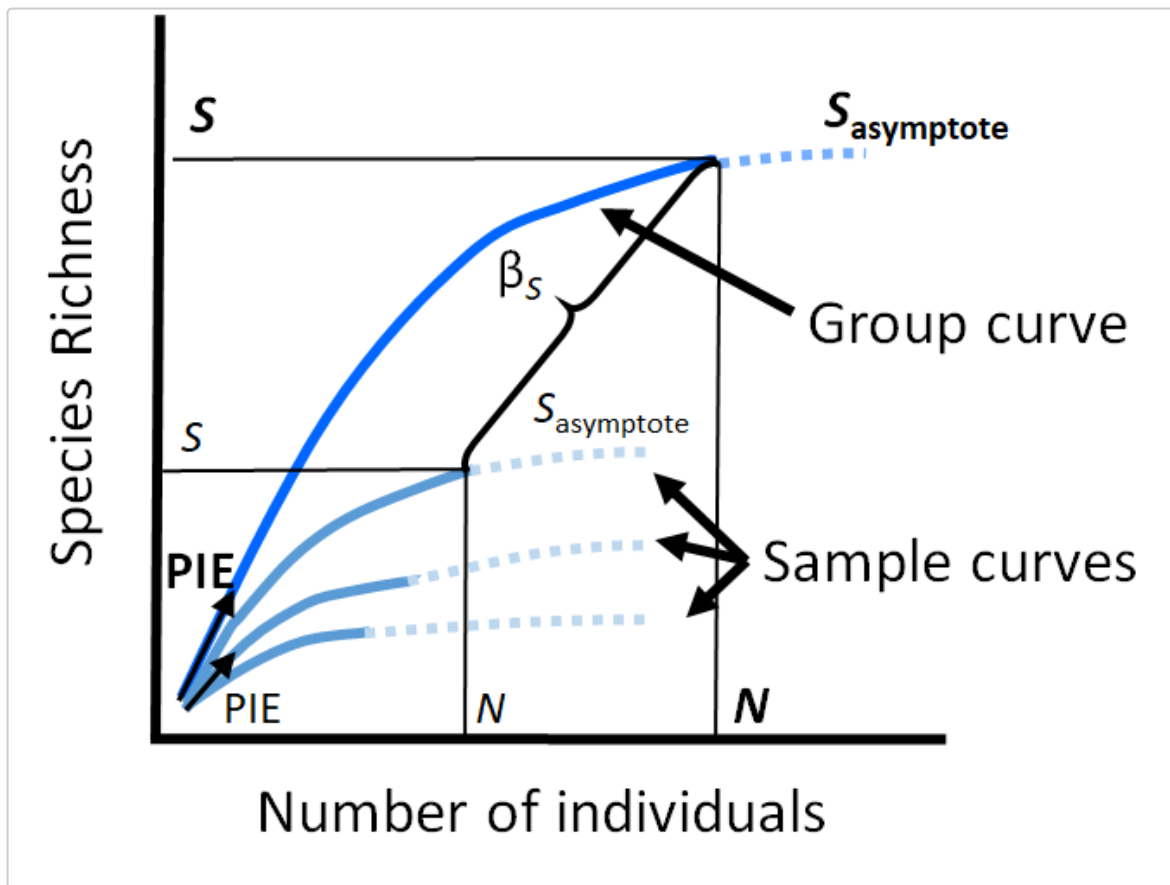
There are a myriad of biodiversity indices. We have attempted to chose a subset of metrics that can all be derived from the individual rarefaction curve which capture the aspects of biodiversity we are most interested in, namely:

1. Numbers of individuals (i.e., density effects)
2. The distribution of rarity and commonness (i.e., the SAD)
3. The spatial patchiness or aggregation of conspecifics.

The metrics we have selected are:

- **N** - Number of individuals
- **S** - Observed species richness
- **S_rarefied** - Rarefied species richness (Hurlbert 1971)
- **S_asymptote** - Estimated asymptotic species richness using biased corrected method (Chiu et al. 2014)
- **PIE** - Probability of Interspecific Encounter (Hurlbert 1971)
- **S_PIE** - Effective number of species based on PIE (Jost 2007)

Each of these metrics can be computed for either the sample or group scale individual rarefaction curves as shown in the figure below:



The ratio of a given biodiversity metric at the group and sample scales (i.e., $\beta_S = S_{group}/S_{sample}$) can provide simple measures of species turnover or β -diversity. Depending on which metric the β -diversity is derived from

will determine what aspects of species turnover it is most sensitive too (e.g., $\beta_{asymptote}$ is more sensitive to turnover in rare species).

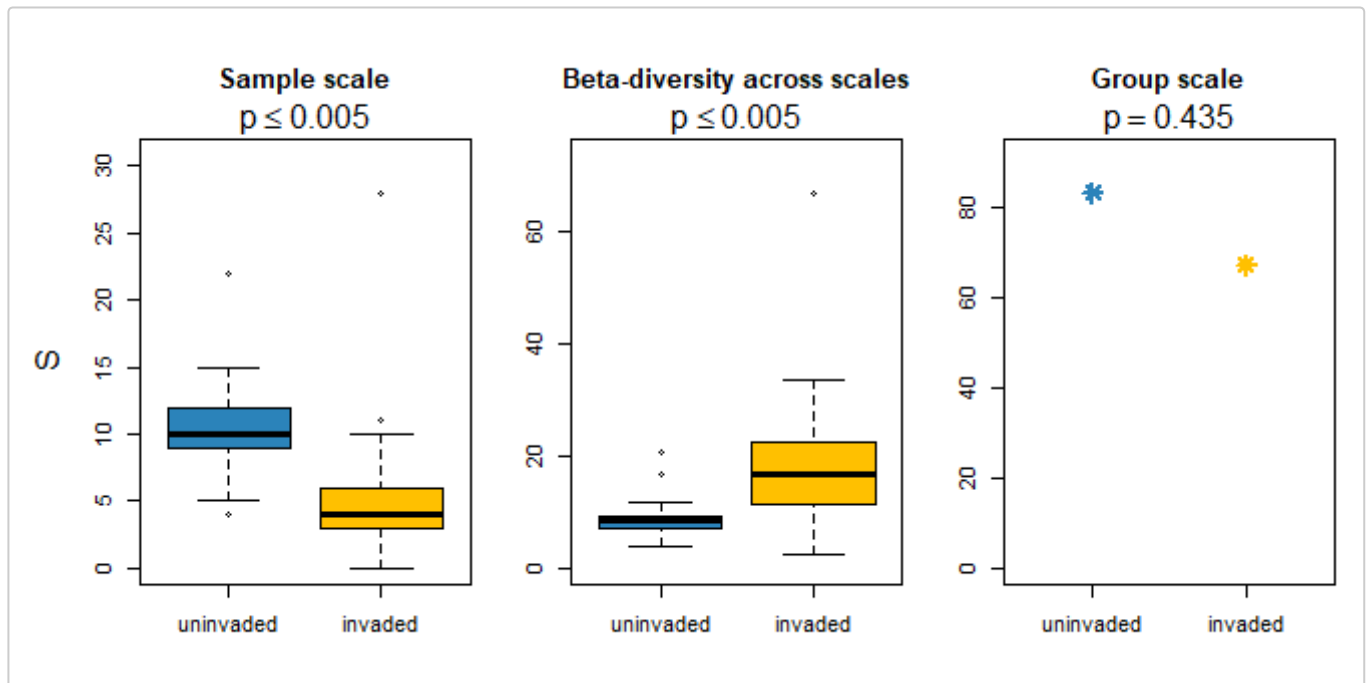
```
inv_stats <- get_mob_stats(inv_mob_in, group_var = "group",  
                           ref_group = 'uninvaded', n_perm = 200)
```

We can examine the `inv_stats` object

```
names(inv_stats)  
  
## [1] "samples_stats" "groups_stats" "samples_pval" "groups_pval"  
## [5] "p_min"
```

There are also functions for plotting the indices at the sample and group levels. First let's examine species richness:

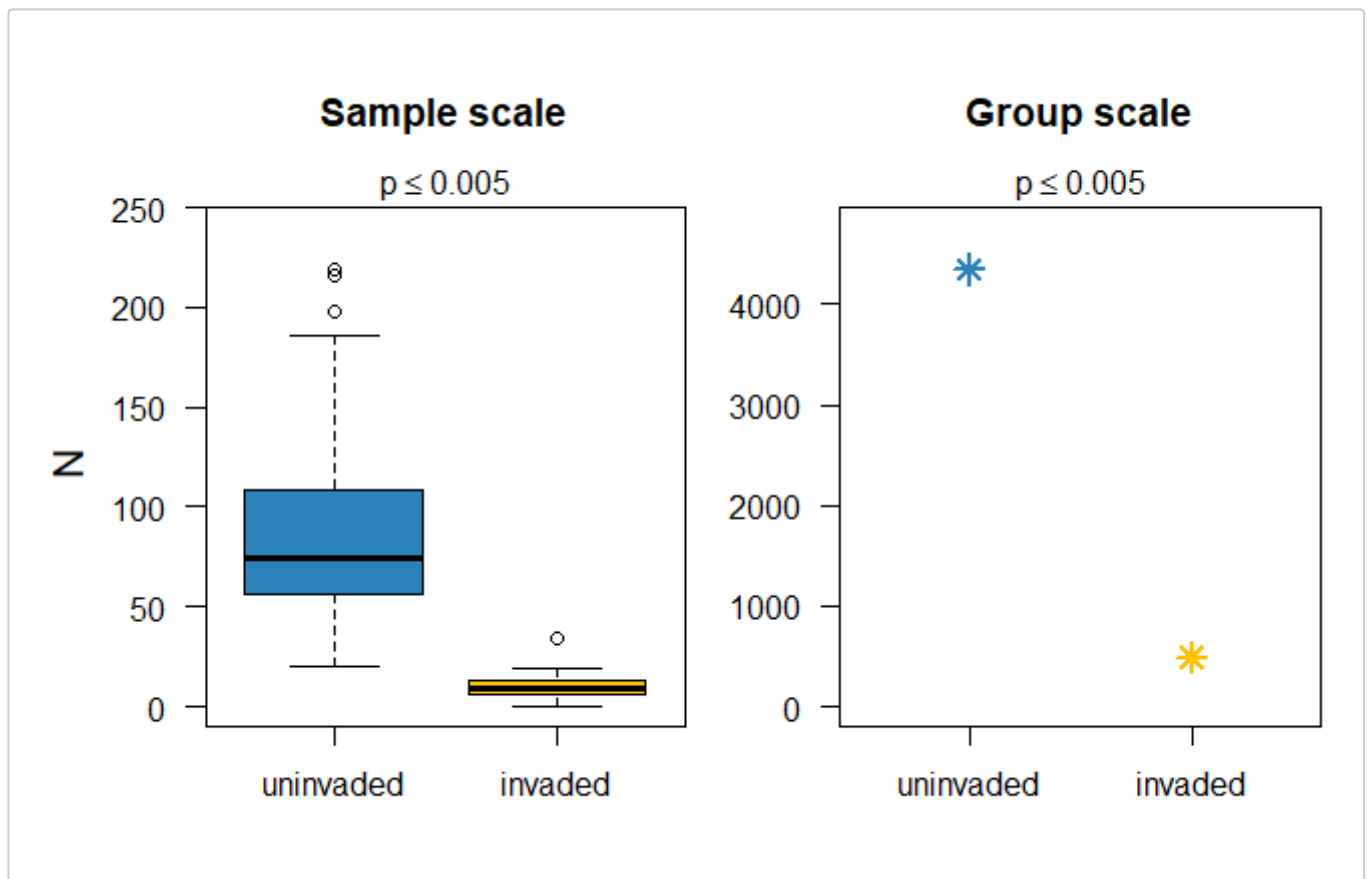
```
plot(inv_stats, 'S')  
  
## Warning: package 'bindrcpp' was built under R version 3.3.3
```



Invasion appears to decrease local sample diversity but not gamma diversity. Somewhat surprisingly it appears to increase β diversity.

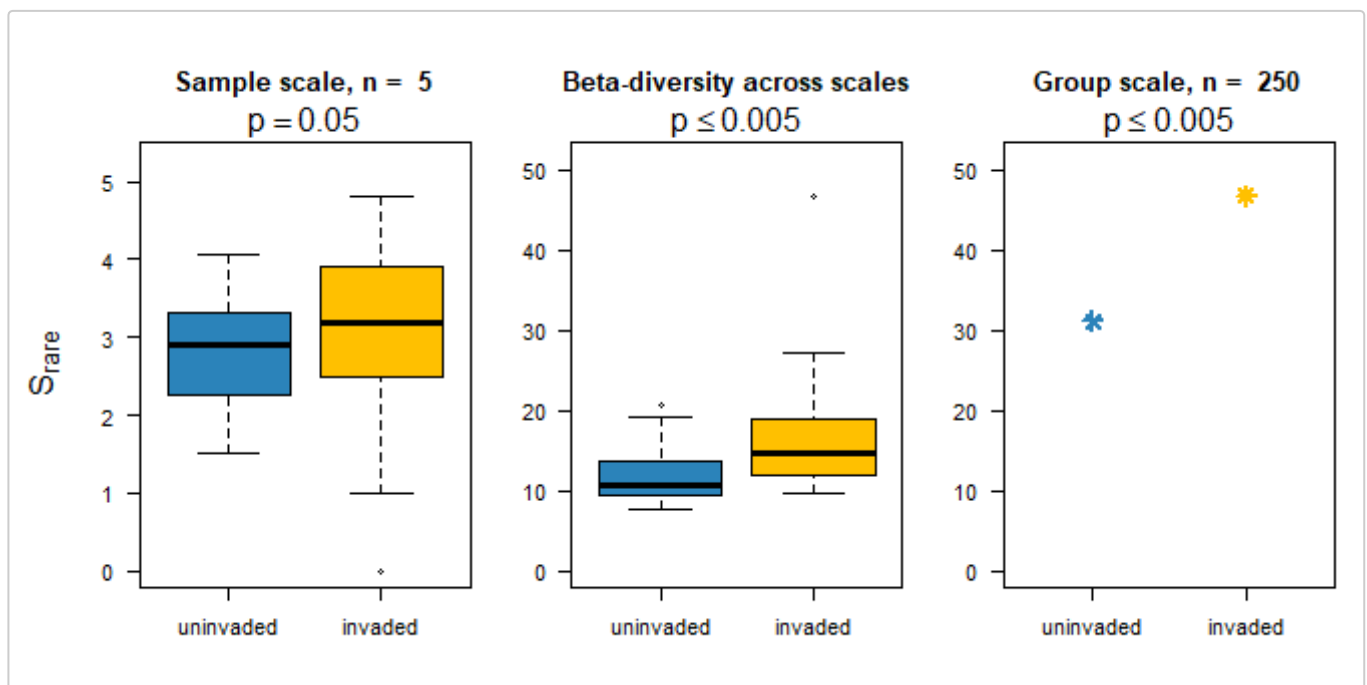
One of the major effects we observed in the individual rarefaction curve was that the invaded sites appeared to have fewer individuals, let's examine the test of that:

```
plot(inv_stats, 'N')
```

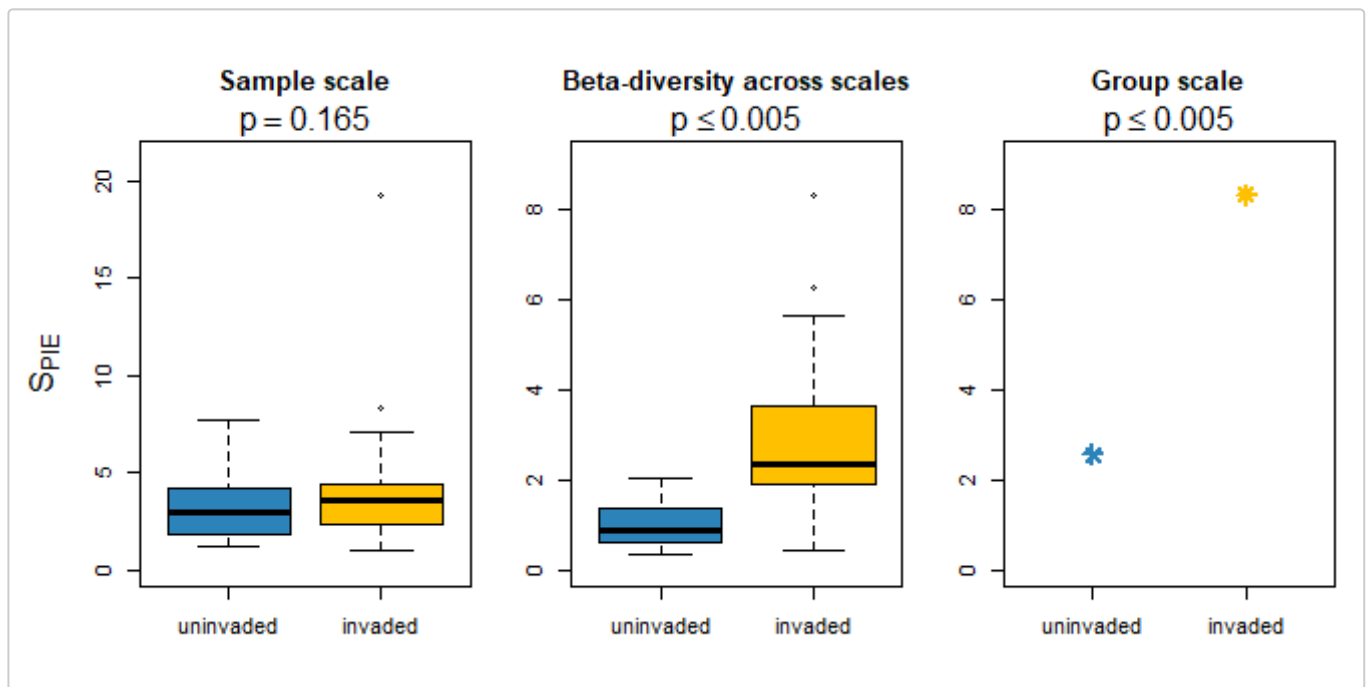


Clearly our intuition was correct there is a very strong negative effect of invasion on N . So it appears that the changes we observed in S may be due to the negative effect of invasion on N . Let's examine S_{rare} to test this:

```
plot(inv_stats, 'S_rare')
```



```
plot(inv_stats, 'S_PIE')
```

We can also plot S_{asympt} but for this dataset this metric does not show strong patterns so we'll stop here for now. If you want to plot all the biodiversity metrics at once you can simply use:

```
plot(inv_stats)
# alternatively
plot(inv_stats, multi_panel = TRUE)
```

Continuous scale analysis

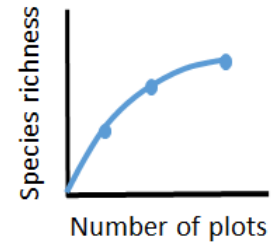
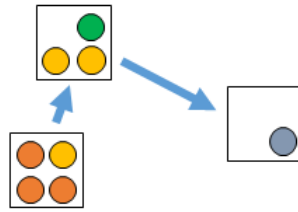
The continuous scale analysis using `mobr` aims at disentangling the consequences of three biodiversity components on observed changes in species richness

1. **Species abundance distribution (SAD)**
2. **Number of individuals (N)**
3. **Aggregation (clumping) of conspecific individuals**

To accomplish this we use three different rarefaction curves which each capture different aspects of community structure:

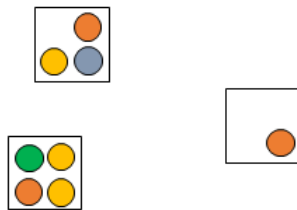
a) Spatial rarefaction curve

Sample nearest neighbors



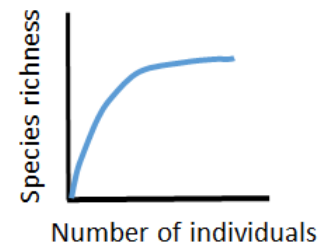
b) Nonspatial rarefaction curve

Shuffle individuals maintain density

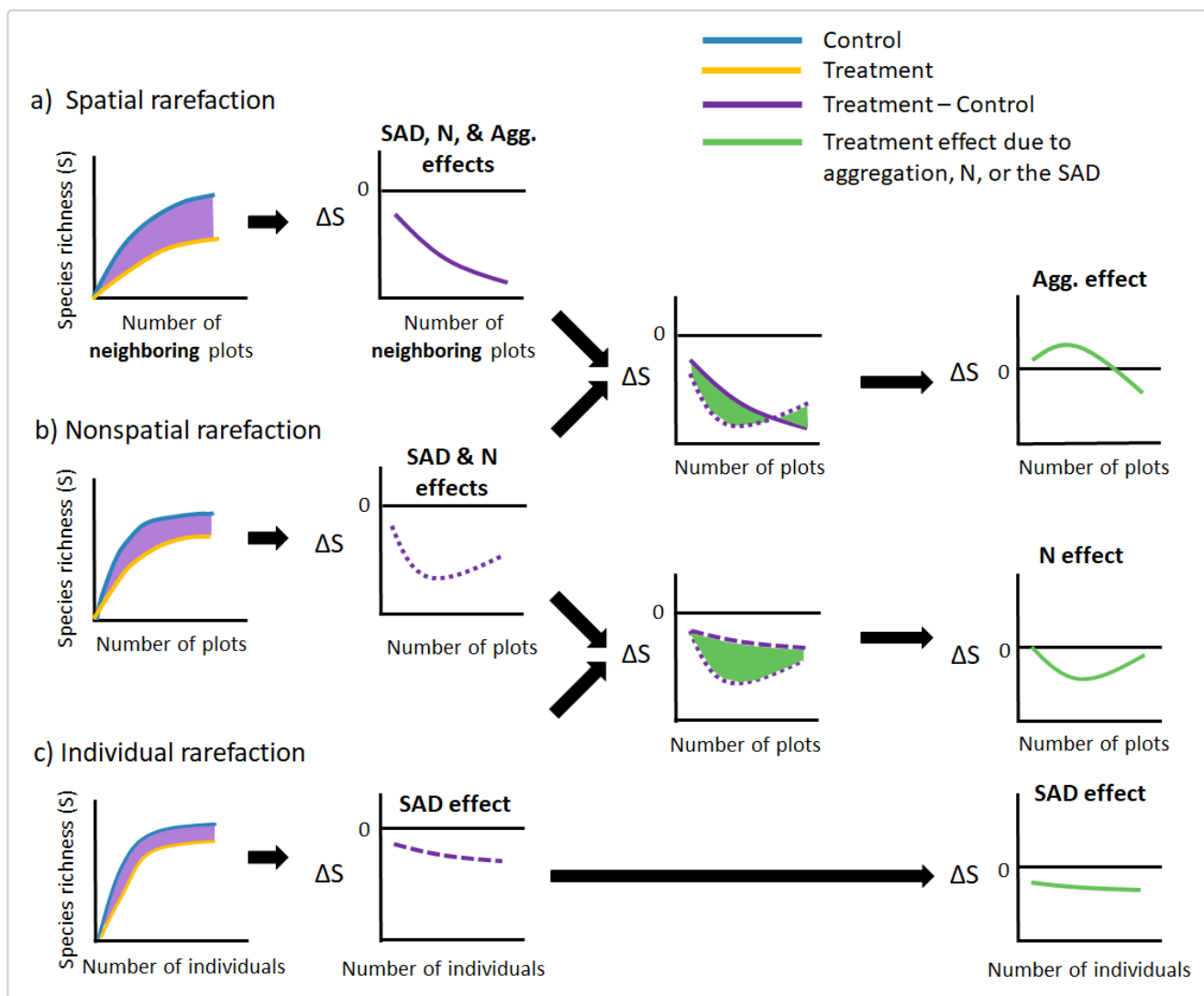


c) Individual rarefaction curve

Randomly sample individuals



If we examine the difference between each of these curves in our two treatments we can learn how the treatment influences richness via its effects on different components of community structure.

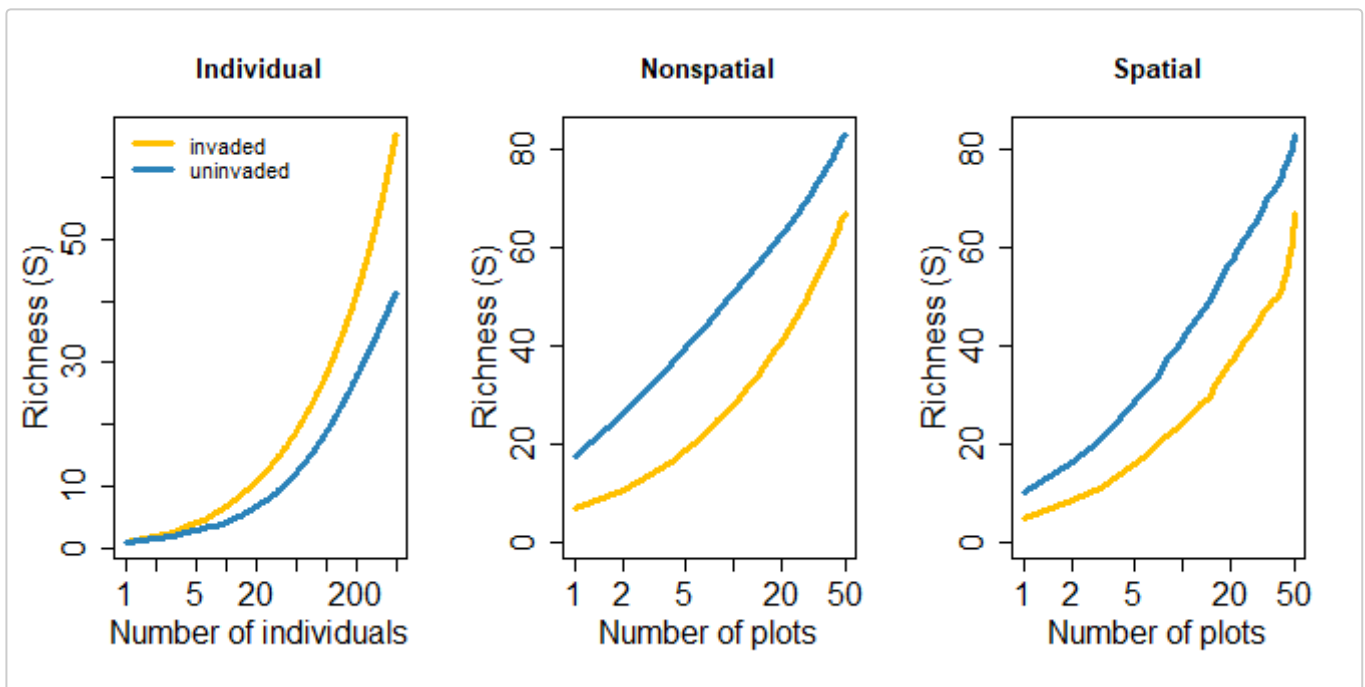


We can carry out this analysis in `mobr` using the function `get_delta_stats`. For the sake of speed we'll run the analysis with just 20 permutations but at least 200 are recommended for actual applications.

```
inv_deltaS = get_delta_stats(inv_mob_in, 'group', ref_group='uninvaded',
                             type='discrete', log_scale=TRUE, n_perm = 20)
```

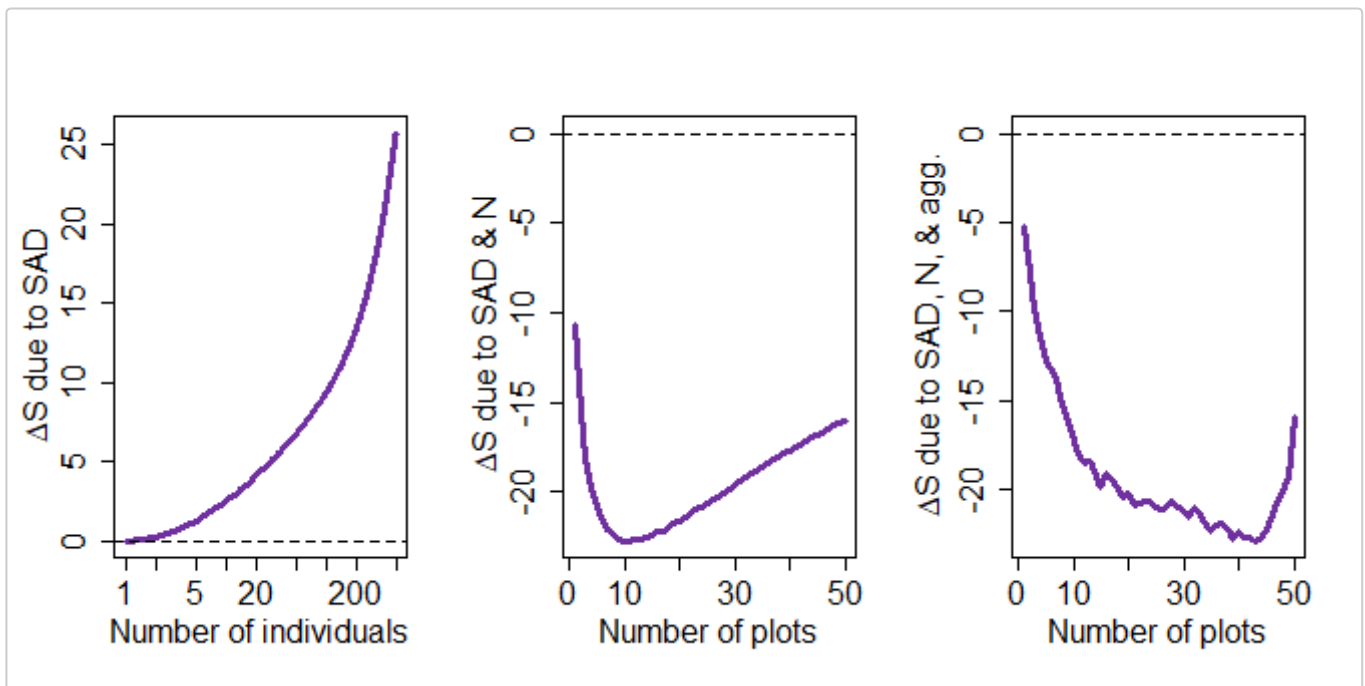
The best way to examine the contents of this object is to plot it. First let's examine the three rarefaction curves:

```
plot(inv_deltaS, 'invaded', 'uninvaded', display='rarefaction')
```



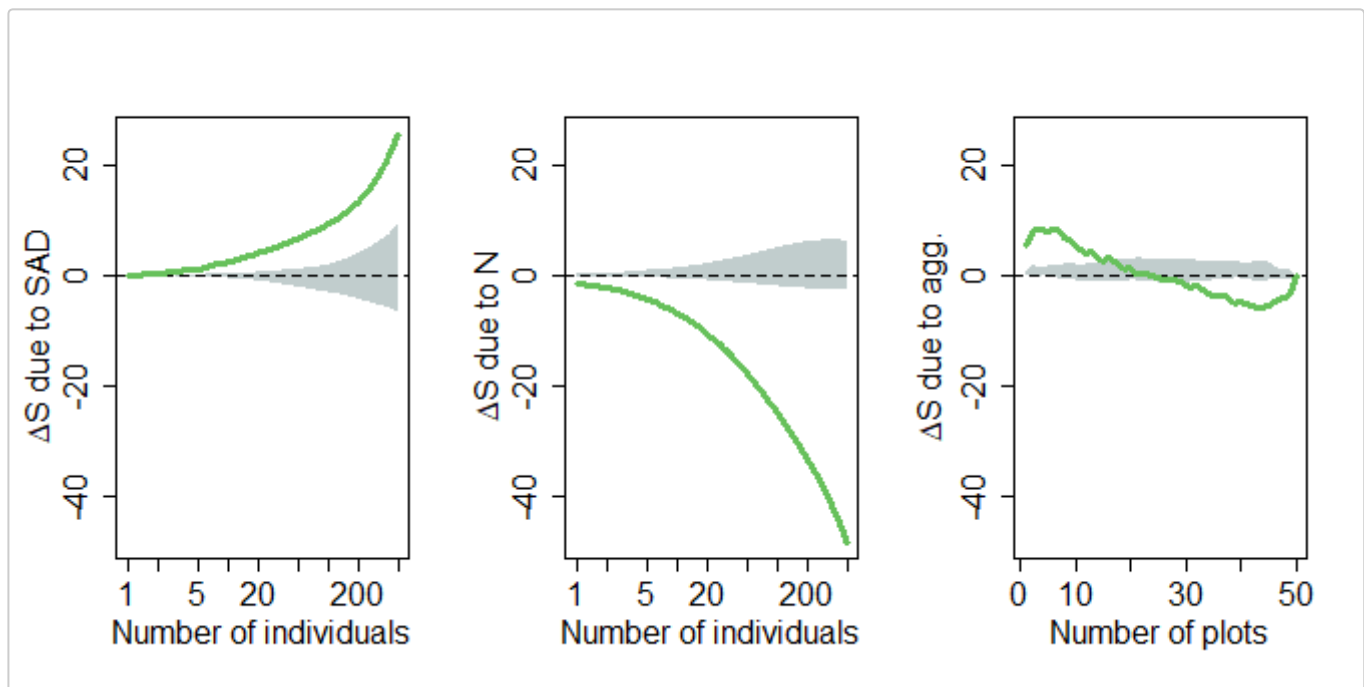
Now let's consider the differences between each set of curves:

```
plot(inv_deltaS, 'invaded', 'uninvaded', display='delta S')
```



Lastly, we can isolate the individual effects of each component by taking one more difference between the curves:

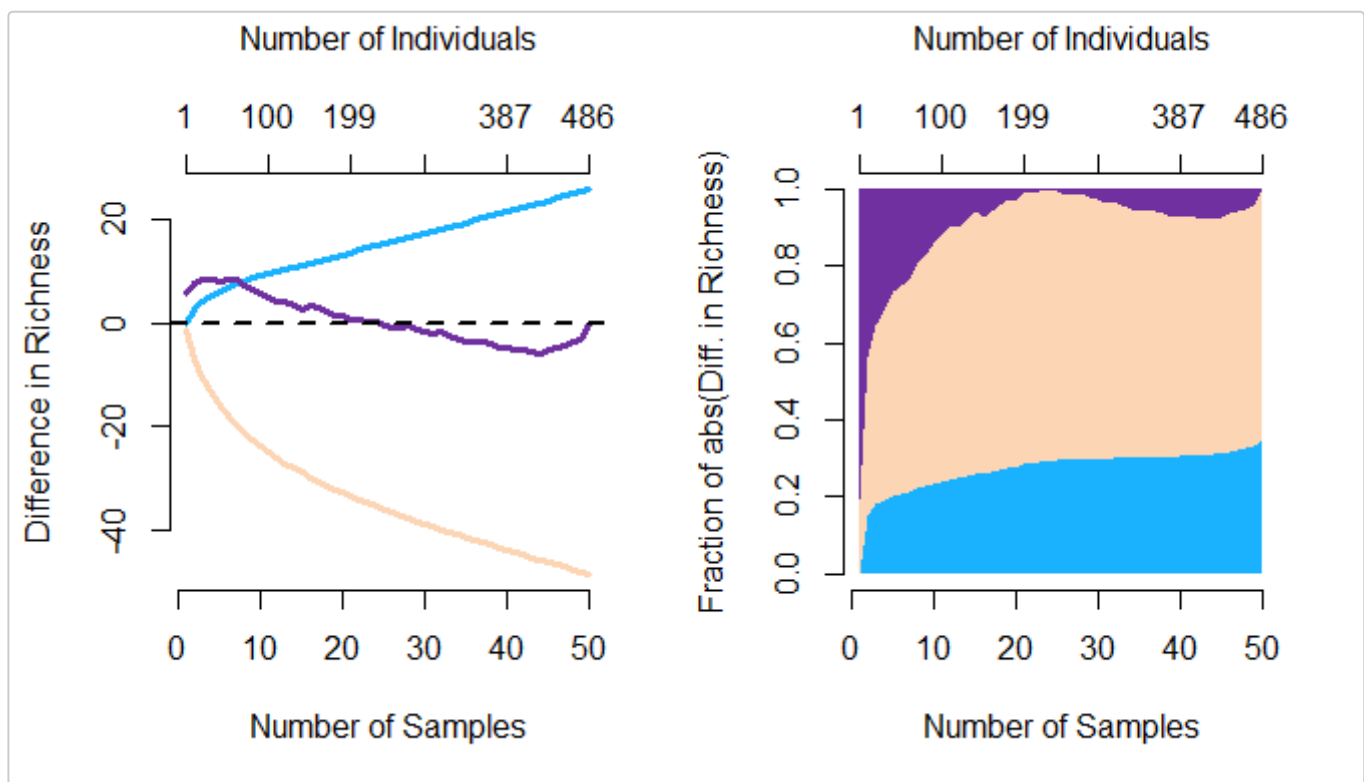
```
plot(inv_deltaS, 'invaded', 'uninvaded', display='ddelta S')
```



The grey polygons above represent the 95% quantile for the null models of no treatment effect.

Let's examine these individual effects across in a way that is easier to compare their effects using the function `overlap_effects`

```
par(mfrow=c(1,2))
overlap_effects(inv_deltaS, 'invaded', display='raw')
overlap_effects(inv_deltaS, 'invaded', display='stacked', prop=T)
```



You can use these same kinds of analyses on your own datasets or on some of other datasets we have included with the package:

```
# plant community in response to a prescribed fire treatment in a  
# central US woodland  
data(fire_comm)  
data(fire_plot_attr)  
  
# aquatic invertebrates in experimental cattle tanks where nutrients  
# were manipulated  
data(tank_comm)  
data(tank_plot_attr)
```

References

1. Powell, K.I., Chase, J.M. & Knight, T.M. (2013). Invasive Plants Have Scale-Dependent Effects on Diversity by Altering Species-Area Relationships. *Science*, 339, 316–318.
2. Gotelli, N.J. & Colwell, R.K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, 4, 379–391
3. Chiu, C.-H., Wang, Y.-T., Walther, B.A. & Chao, A. (2014). An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics*, 70, 671-682.
4. Hurlbert, S.H. (1971). The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology*, 52, 577–586
5. Jost, L. (2007). Partitioning Diversity into Independent Alpha and Beta Components. *Ecology*, 88, 2427-2439.