

# Tutorial vignette for package ‘specificity’

*John L Darcy*

*30 August 2019*

## Introduction

In this example analysis, our goal is to analyze the extent to which microbes have specificity to different aspects of their environment. We will use the ‘endophyte’ data set (included with the specificity R package), which contains data from a survey of foliar endophytic fungi living within the leaves of native Hawaiian plants. In this context, a fungus with strong specificity would be one that occupies a narrow range of sampled environment. For example, a fungus that preferentially associates with a narrow clade of host plants (e.g. asteraceae), or prefers a narrow range of elevation (e.g. between 700 and 1200 m.a.s.l). The previous two examples were for unimodal specificity, but specificity as calculated with our software can be multimodal as well (e.g. phylogenetic specificity to both asteraceae and malvaceae). It is worth noting that a strength of this approach is that specificity is agnostic to the ideal habitat of a species: a species with strong specificity to high elevations may be just as specific to elevation as a species specific to low elevations. Indeed, a strength of this approach is that the ideal habitat does NOT need to be modeled in order to calculate specificity. Instead, the standardized effect size (SES) of specificity is a statistic that quantifies the degree to which a species occupies a more narrow range of some environmental variable than would be expected by random chance.

## Software Requirements

- R (run on version 3.6.0, but likely works fine on earlier versions)
- Other dependencies will be automatically installed with the package.

## Installation

To install specificity, open up R and run:

```
library(devtools)
install_github("darcyj/specificity")
```

## Run specificity analysis on included ‘endophyte’ example data set

The endophyte dataset consists of foliar endophytic fungi sampled from leaves of native Hawaiian plants across the Hawaiian archipelago. These data can be loaded into R with:

```
# load specificity R package
suppressMessages(library(specificity))
# load endophyte data set from specificity
data(endophyte)
# check what objects are within endophyte
names(endophyte)
```

```
## [1] "metadata" "zotutable" "supertree"
```

```
# add those objects to R namespace so we aren't typing "endophyte$" every time
attach(endophyte)
```

As you can see, there are 3 objects inside ‘endophyte’.

- metadata: table of data where each row corresponds to a sample, and each column is a different metadata category (e.g. PlantGenus, Elevation, Lon=longitude, Lat=latitude)
- zotutable: table of zOTU (species) observation data, where each row is a sample, and each column is a different fungal zOTU.
- supertree: a phylogenetic tree of host plants.