

# Beyond Metabarcoding: Community Ecology

Jayde MacMillan 20068250

github-jaydealexandra (<https://github.com/jaydealexandra/Metabarcoding>)

March 16 2022

## Introduction

Metabarcoding (<https://en.wikipedia.org/wiki/Metabarcoding>) is the barcoding of DNA or RNA that allows taxonomic identification on a large scale. This method provides a fast and efficient way of identifying many organisms within a specific area Westfall et al., 2019 (<https://pubmed.ncbi.nlm.nih.gov/31657513/>). As invasive species continue to spread, the use of metabarcoding as a biosurveillance tool can be utilized to detect and manage these species.

Garlic mustard, *Alliaria petiolata* is an invasive species native to Europe. If left unmanaged, it can displace other native wildflowers and interfere with the growth of species of fungi (Garlic Mustard | Ontario's Invading Species Awareness Program, 2021) (<http://www.invadingspecies.com/invaders/plants/garlic-mustard/>)

Using a floristic survey from Queen's University Biological Station (QUBS), we will address two questions:

1. Does garlic mustard effect the plant community?
2. Does either the presence or absence of garlic mustard have a stronger effect on the sampling population?



Figure 1. A photo of garlic mustard.(Garlic Mustard - Ontario Invasive Plant Council, 2021)  
(<https://www.ontarioinvasiveplants.ca/invasive-plants/species/garlic-mustard/>)

## Setup

```
#Loads the required packages
```

```
library(ggplot2)
```

```
library(ape)
```

```
library(ggtree)
```

```
## ggtree v3.2.1 For help: https://yulab-smu.top/treedata-book/
```

```
##
```

```
## If you use ggtree in published research, please cite the most appropriate paper(s):
```

```
##
```

```
## 1. Guangchuang Yu. Using ggtree to visualize data on tree-like structures. Current Protocols in Bioinformatics. 2020, 69:e96. doi:10.1002/cpbi.96
```

```
## 2. Guangchuang Yu, Tommy Tsan-Yuk Lam, Huachen Zhu, Yi Guan. Two methods for mapping and visualizing associated data on phylogeny using ggtree. Molecular Biology and Evolution. 2018, 35(12):3041-3043. doi:10.1093/molbev/msy194
```

```
## 3. Guangchuang Yu, David Smith, Huachen Zhu, Yi Guan, Tommy Tsan-Yuk Lam. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 2017, 8(1):28-36. doi:10.1111/2041-210X.12628
```

```
##  
## Attaching package: 'ggtree'
```

```
## The following object is masked from 'package:ape':  
##  
##      rotate
```

```
library(vegan)
```

```
## Warning: package 'vegan' was built under R version 4.1.3
```

```
## Loading required package: permute
```

```
## Warning: package 'permute' was built under R version 4.1.3
```

```
## Loading required package: lattice
```

```
## This is vegan 2.5-7
```

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##      filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##      intersect, setdiff, setequal, union
```

```
library(reshape2)
```

```
#imports data into an object  
floristicSurvey <- read.csv("FloristicSurvey.csv", header = T, fileEncoding="UTF-8-BOM")  
  
#selects only the species  
floristicSpec <- select(floristicSurvey, Claytonia_virginiana:maianthenum_racemosum)  
  
#creates a data frame of the species and sets the row names to the  
floristicSub <- data.frame(floristicSpec, row.names = floristicSurvey$Quadrat)
```

## Bray-Curtis Dissimilarity

The Bray-Curtis Dissimilarity ([https://en.wikipedia.org/wiki/Bray%E2%80%93Curtis\\_dissimilarity](https://en.wikipedia.org/wiki/Bray%E2%80%93Curtis_dissimilarity)) is used to count the ecological distance between two sites based on the number of occurrences at each site.

```
florDist <- vegdist(florigenicSub, method = "bray", binary = F)
```

## Neighbour Joining Method

```
#performs a neighbour joining tree estimation of the dissimilarity index
florTree <- nj(florigenicDist)
```

```
#creates a tree
ggtree(florigenicTree, layout="rectangular") %<+% florigenicSurvey +
  geom_tiplab(aes(colour= Location)) +
  theme(legend.position="right") +
  scale_colour_discrete(name= "Garlic Mustard", labels = c("Present", "Absent"))
```

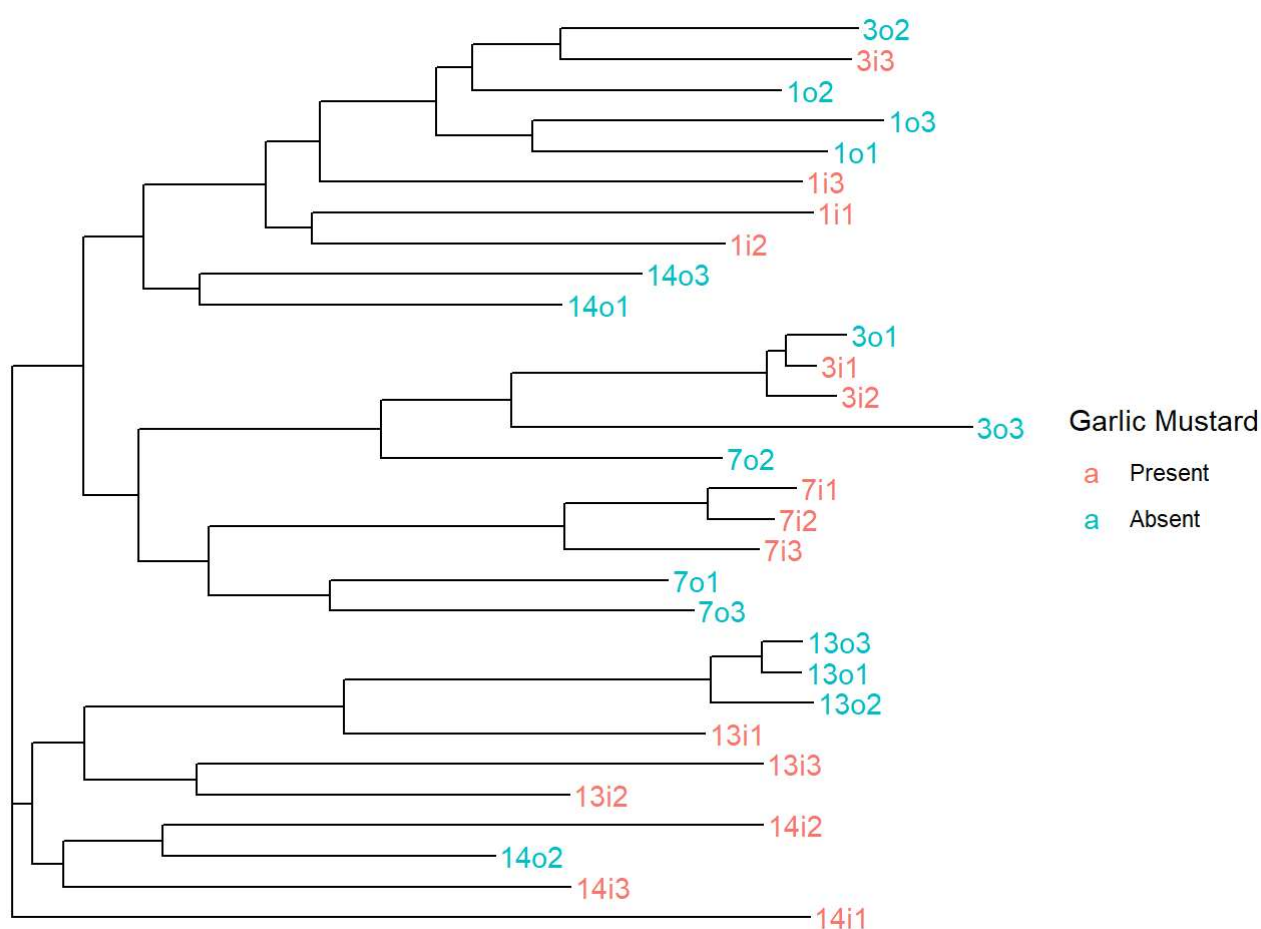


Figure 2. A tree showing species that cluster together in the presence and absence of garlic mustard. Garlic mustard has an effect only within the same population.

## NMDS

```
set.seed(13)
NMDSdat <- metaMDS(florigenicDist, k=2)
```

```
## Run 0 stress 0.1461135
## Run 1 stress 0.153444
## Run 2 stress 0.153444
## Run 3 stress 0.141439
## ... New best solution
## ... Procrustes: rmse 0.06682265  max resid 0.3092955
## Run 4 stress 0.1489959
## Run 5 stress 0.1489955
## Run 6 stress 0.1560521
## Run 7 stress 0.161178
## Run 8 stress 0.1414389
## ... New best solution
## ... Procrustes: rmse 0.0001063247  max resid 0.0003944835
## ... Similar to previous best
## Run 9 stress 0.141439
## ... Procrustes: rmse 0.0004805731  max resid 0.001783307
## ... Similar to previous best
## Run 10 stress 0.153444
## Run 11 stress 0.1589506
## Run 12 stress 0.1560521
## Run 13 stress 0.1461135
## Run 14 stress 0.2101893
## Run 15 stress 0.1489955
## Run 16 stress 0.206437
## Run 17 stress 0.1589506
## Run 18 stress 0.1665459
## Run 19 stress 0.1676016
## Run 20 stress 0.1414388
## ... New best solution
## ... Procrustes: rmse 0.0002399126  max resid 0.0008810417
## ... Similar to previous best
## *** Solution reached
```

*#creates data for plotting*

```
PDat<-data.frame(NMDS1=NMDSdat$points[,1],
                 NMDS2=NMDSdat$points[,2],
                 Quadrate=row.names(florigisticSub))
```

*#adds labels using the merge function*

```
PDat<-merge(PDat,florigisticSurvey,by="Quadrate",all.x=T,all.y=F)
```

*#creates a plot*

```
qplot(x=NMDS1,y=NMDS2,colour=Location,alpha=I(0.6),data=PDat)+theme_bw()
```

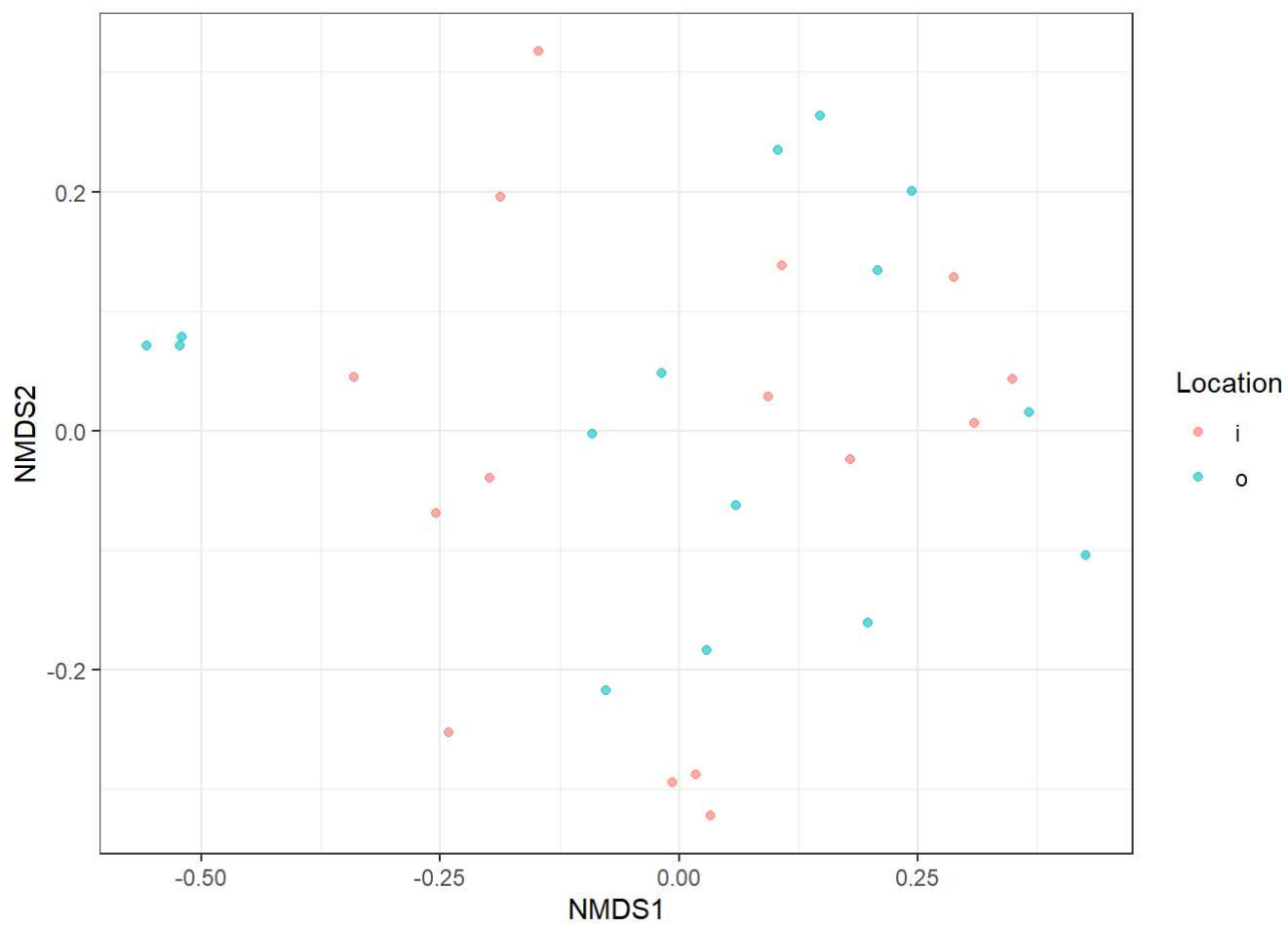


Figure 3.