Assignment 8

Code ▾

- Nadejda Boev (20056079)
- Due Date - 2022/03/16
- Github user - 16nbb1
- Github link - https://github.com/16nbb1/Biol432_A8 (https://github.com/16nbb1/Biol432_A8)

Loading in libraries

Hide

```
library(dplyr)
library(tibble)
library(ggplot2)
library(ape)
library(ggtree)
library(vegan)
library(cowplot)
library(tidyverse)
```

Background

Garlic mustard is an invasive species found near QUBS. We know invasive species, like garlic mustard, can heavily influence their environment, so plant diversity decreases, allowing for less competition and therefore are "greedy". Recently, a community ecology survey sampled regions in and outside of patches of garlic mustard patches. Our goal is to understand:

1. What effect (if any) does garlic mustard have on the plant community?
2. What has a stronger effect on plant communities: the presence/absence of garlic mustard (in/out) or sampling population?

Methods

We will create table of species found at each sampling site, this will represent our "plant community."

Loading in data

Hide

```
(df<-read.csv("./Data/FloristicSurvey.csv"))
```

Quadrate <chr>	Population <int>	Location <chr>	Rosettes <int>	Bolting <int>	Budding <int>	Bud_Flw <int>	Flower <int>	Flw_Sil <int>	Sillique <int> ▶
7o3	7	o	0	0	0	0	0	0	0
7o1	7	o	0	0	0	0	0	0	0
7o2	7	o	0	0	0	0	0	0	0
7i3	7	i	14	8	157	0	0	0	0
7i2	7	i	3	18	184	0	0	0	0
7i1	7	i	0	3	122	0	0	0	0
3i1	3	i	14	3	11	0	0	0	0
3i2	3	i	0	12	23	0	0	0	0
3i3	3	i	0	26	19	0	0	0	0
3o1	3	o	0	0	0	0	0	0	0

1-10 of 30 rows | 1-10 of 44 columns   Previous **1** 2 3 Next

Creating an "OTU" table using the species which are found in the quadrate site.

Hide

```
otu = df %>%
  # making row names, the names of each Quadrate
  column_to_rownames('Quadrate') %>%
  # Claytonia_virginiana onwards..
  select(11:length(df)-1)
```

We will create calculate the pairwise distance for all sampled regions

Calculating binary distance

Hide

```
OTU_bin<-otu
for (i in 1:nrow(OTU_bin)){
  for (j in 1:ncol(OTU_bin)){
    if(OTU_bin[i,j]>0){

      # calculating raw differences, using a binary 0 vs 1 for if a difference is present
      OTU_bin[i,j]<-1
    }
  }
}

OTU_bin<-otu
OTU_bin[OTU_bin>0]<-1
```

Calculating distance

Hide

```
OTU_bin_dist<-dist(OTU_bin,method='binary')
```

We will use the pairwise distance, to visualize phylogenetic trees and color by garlic mustard- associated features.

I could visualize the trees using neighbour-joining to show how samples cluster together. However, since we have "abundance" data from the plan community, we can also directly calculate Bray-Curtis dissimilarity.

Hide

```
OTU_dist<- vegdist(otu,method="bray",binary=F)
OTU_tree2<-nj(OTU_dist)
```

Visualizing if Location influences branching

Hide

```
tree1 = ggtree(OTU_tree2,layout="rectangular") %<+%
  df +
  geom_tiplab(aes(colour=Location)) +
  scale_color_manual(values = c('red', 'blue'), labels = c("Inside patch", "Outside patch"))
```

Visualizing if the number of plants bolting influences branching.

Hide

```
tree2 = ggtree(OTU_tree2,layout="rectangular") %<+%
  df +
  geom_tiplab(aes(colour=Bolting))
```

We will calculate NMDS to visualize clustering by by garlic mustard- associated features, including a comparison of the presence/absence of garlic mustard (in/out) and sampling population.

Calculating MDS, which includes only 2 dimensions (garlic mustard in vs out of population sampled)

Hide

```
set.seed(13)
NMDSdat<-metaMDS(OTU_dist,k=2,trymax = 100)
```

```
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
```

Hide

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

Hide

```
# Mergin so we can use categories from df
PDat<-merge(PDat,df,by="Quadrate",all.x=T,all.y=F)
```

Visualizing if Population influences clustering

Hide

```
plot3 = qplot(x=NMDS1,y=NMDS2,colour=as.factor(Population),alpha=I(0.6),data=PDat)+
  theme_bw() +
  scale_color_discrete(name = "Population")
```

Visualizing if Location influences clustering

Hide

```
plot4 = qplot(x=NMDS1,y=NMDS2,colour=Location,alpha=I(0.6),data=PDat)+
  theme_bw()+
  scale_color_manual(values = c('red', 'blue'), labels = c("Inside patch", "Outside patch"))
```

Conclusions: What effect (if any) does garlic mustard have on the plant community?

To begin, we wanted to identify if the sampling location could explain distance branching. However, from Figure 1, there is no clear pattern.

Hide

```
ggdraw(add_sub(tree1, size =10,
              'Figure 1. Tree made from Bray-Curtis dissimilarity, depicting
              branching from inside and outside garlic patches.'))
```
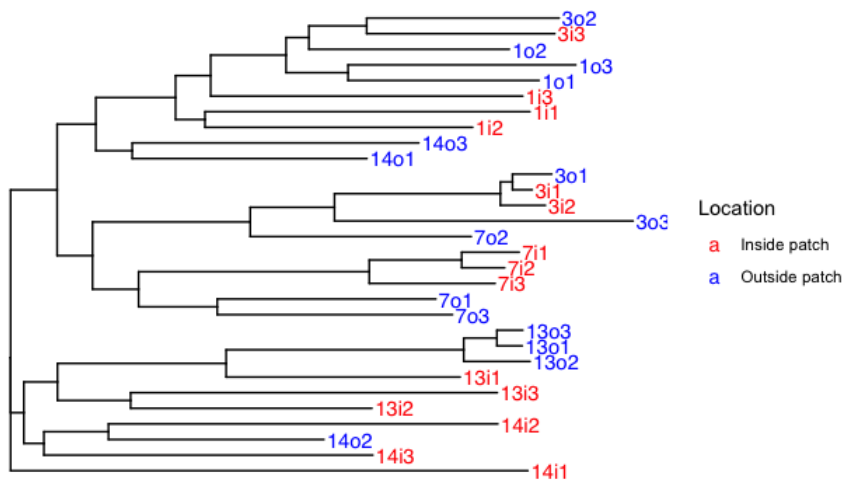
Figure 1. Tree made from Bray-Curtis dissimilarity, depicting branching from inside and outside garlic patches.

With regards to garlic mustards' potential strategy, to germinate and create many plants prepared to bolt, we tested to see if there was clear branching. However, from Figure 2, there appears to be some patterning, where regions sampled with many bolting garlic mustards somewhat cluster together (bottom).

Hide

```
ggdraw(add_sub(tree2, size =10,
              'Figure 2. Tree made from Bray-Curtis dissimilarity, depicting
              branching from the number of garlic mustard bolting'))
```
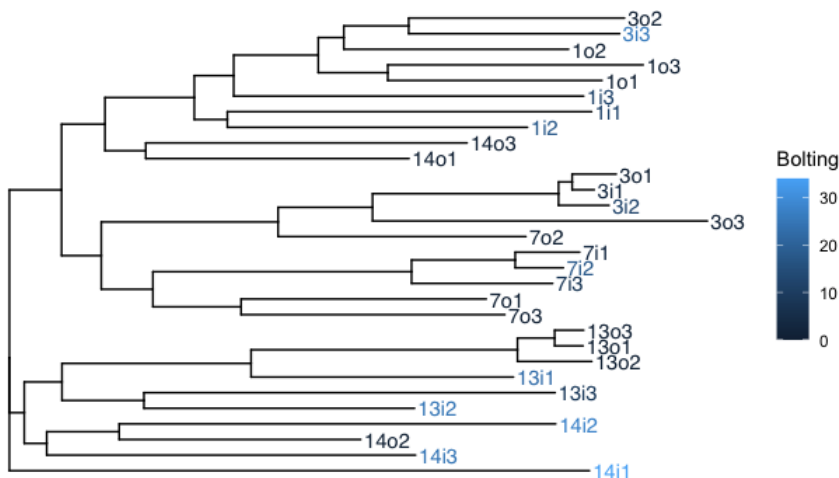

Figure 2. Tree made from Bray-Curtis dissimilarity, depicting branching from the number of garlic mustard bolting

Note: Bud_Flw, Flower, Flw_Sil and Sillique all have only 0s documented in the survey. Therefore, we cannot test if garlic mustard is "influential" during later stages of germination.

Therefore, there appears to be greater similarity among populations with greater numbers of bolting plants, not necessarily just the presence of garlic mustard. We may conclude, that garlic mustard, does use bolting/germination as an invasive strategy.

Conclusions: What has a stronger effect on plant communities: the presence/absence of garlic mustard (in/out) or sampling population?

We compared the clustering using NMDS, of plant communities, where clusters were coloured by population (Figure 3) or the presence/absence of garlic mustard (Figure 4). In Figure 3, we can see clear groupings among populations, for example, population 7 is tightly and uniquely clustered together. In contrast, in Figure 4, both groups are dispersed and clusters overlap.

Hide

```
ggdraw(add_sub(plot3, size =10,
              'Figure 3. NMDS plot clustering visualizing clustering of
              plant community based on sampled population'))
```
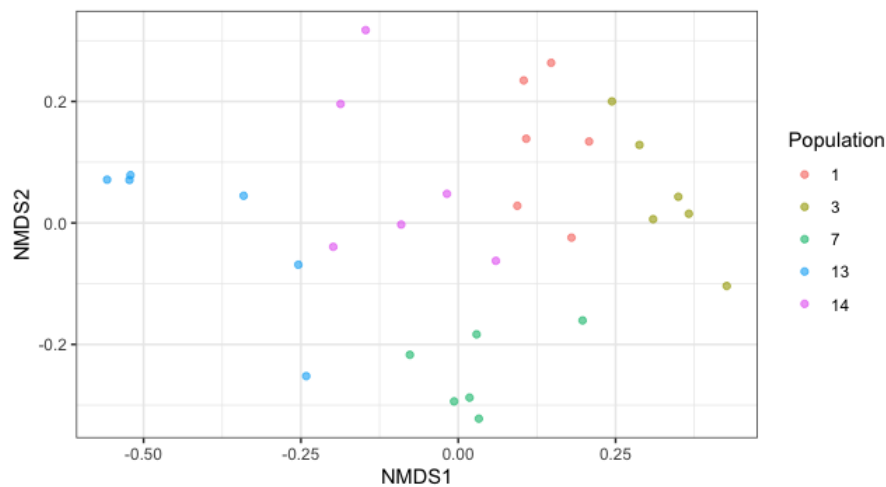
Figure 3. NMDS plot clustering visualizing clustering of plant community based on sampled population

Hide

```
ggdraw(add_sub(plot4, size =10,
               'Figure 4. NMDS plot clustering visualizing clustering of
               plant community based on sampled location'))
```
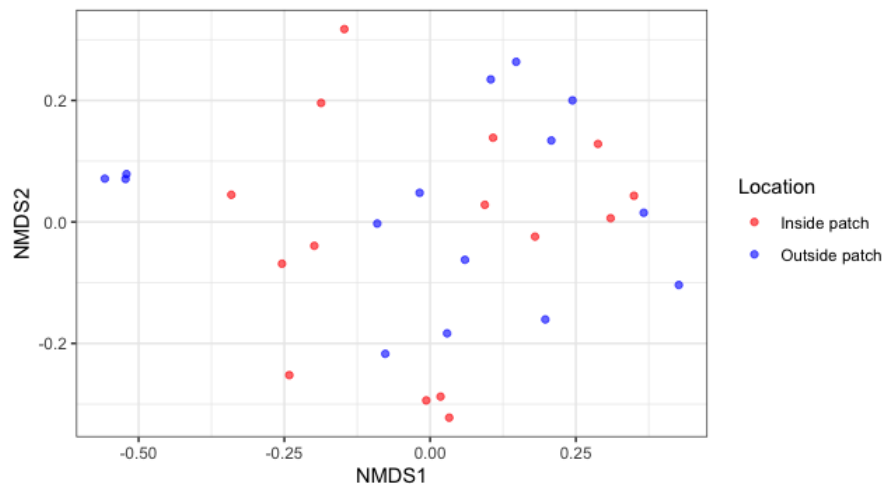

Figure 4. NMDS plot clustering visualizing clustering of plant community based on sampled location

Therefore, it appears the location sampled is more influential to the similarity in plant community compared to the presence/absence of garlic mustard.

Appendix:

• Quadrate – this is a code for a 1m x 1m square in which samples were measured

• Population – this is a code for a specific sampling location at QUBS, determined by presence of garlic mustard nearby

• Location – whether quadrate measurements are taken 'i' for 'inside' or 'o' for 'outside' of garlic mustard patches

• Rosettes – the number of garlic mustard rosettes

• Bolting – the number of bolting garlic mustard plants (i.e. getting ready to flower)

• Budding – the number of garlic mustard plants with buds

• Bud_Flw, Flower, Flw_Sil, Silique – These are different stages of flowering

• Claytonia, Anemone, … Unknown_9, mianthenum_racemosum – these are the 34 different species present at all the sites. 'Unknown' means that the species hasn't been definitively identified, but all the sites with the same Unknown number have that species.