

# Jewitt\_Ally\_A8

Ally Jewitt

2023-03-07

## Part I

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

#### Load libraries

```
library(ggplot2)
library(ape)
library(ggtree)
```

```
## ggtree v3.6.2 For help: https://yulab-smu.top/treedata-book/
##
## If you use the ggtree package suite in published research, please cite
## the appropriate paper(s):
##
## 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
##
## LG Wang, TTY Lam, S Xu, Z Dai, L Zhou, T Feng, P Guo, CW Dunn, BR
## Jones, T Bradley, H Zhu, Y Guan, Y Jiang, G Yu. treeio: an R package
## for phylogenetic tree input and output with richly annotated and
## associated data. Molecular Biology and Evolution. 2020, 37(2):599-603.
## doi: 10.1093/molbev/msz240
##
## Guangchuang Yu. Data Integration, Manipulation and Visualization of
## Phylogenetic Trees (1st edition). Chapman and Hall/CRC. 2022,
## doi:10.1201/9781003279242
```

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

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

```
library(vegan)
```

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

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

```
## This is vegan 2.6-4
```

```
library(dplyr)
```

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

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

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

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

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages ————— tidyverse 2.0.0 —  
## ✓ forcats 1.0.0 ✓ stringr 1.5.0  
## ✓ lubridate 1.9.2 ✓ tibble 3.1.8  
## ✓ purrr 1.0.1 ✓ tidyr 1.3.0  
## ✓ readr 2.1.4
```

```
## — Conflicts — tidyverse_conflicts() —
## ✖ tidyr::expand() masks ggtree::expand()
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag() masks stats::lag()
## ✖ dplyr::where() masks ape::where()
## i Use the conflicted::conflict_warn() to force all
conflicts to become errors
```

```
library(data.table)
```

```
##
## Attaching package: 'data.table'
##
## The following objects are masked from 'package:lubridate':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday, week,
##     yday, year
##
## The following object is masked from 'package:purrr':
##
##     transpose
##
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

## Import OTU table

```
OTU_table<-data.table::fread("OTU.csv", header=T,data.table = F)
tibble::column_to_rownames(OTU_table, "Quadrant") -> OTU_table
```

## Bray-Curtis dissimilarity

```
OTU_dist<-vegdist(OTU_table,method="bray",binary=F)
```

```
OTUtree<-nj(OTU_dist)
```

## Cluster by similarity

```
ggtree(OTUtree,layout="rectangular")
```

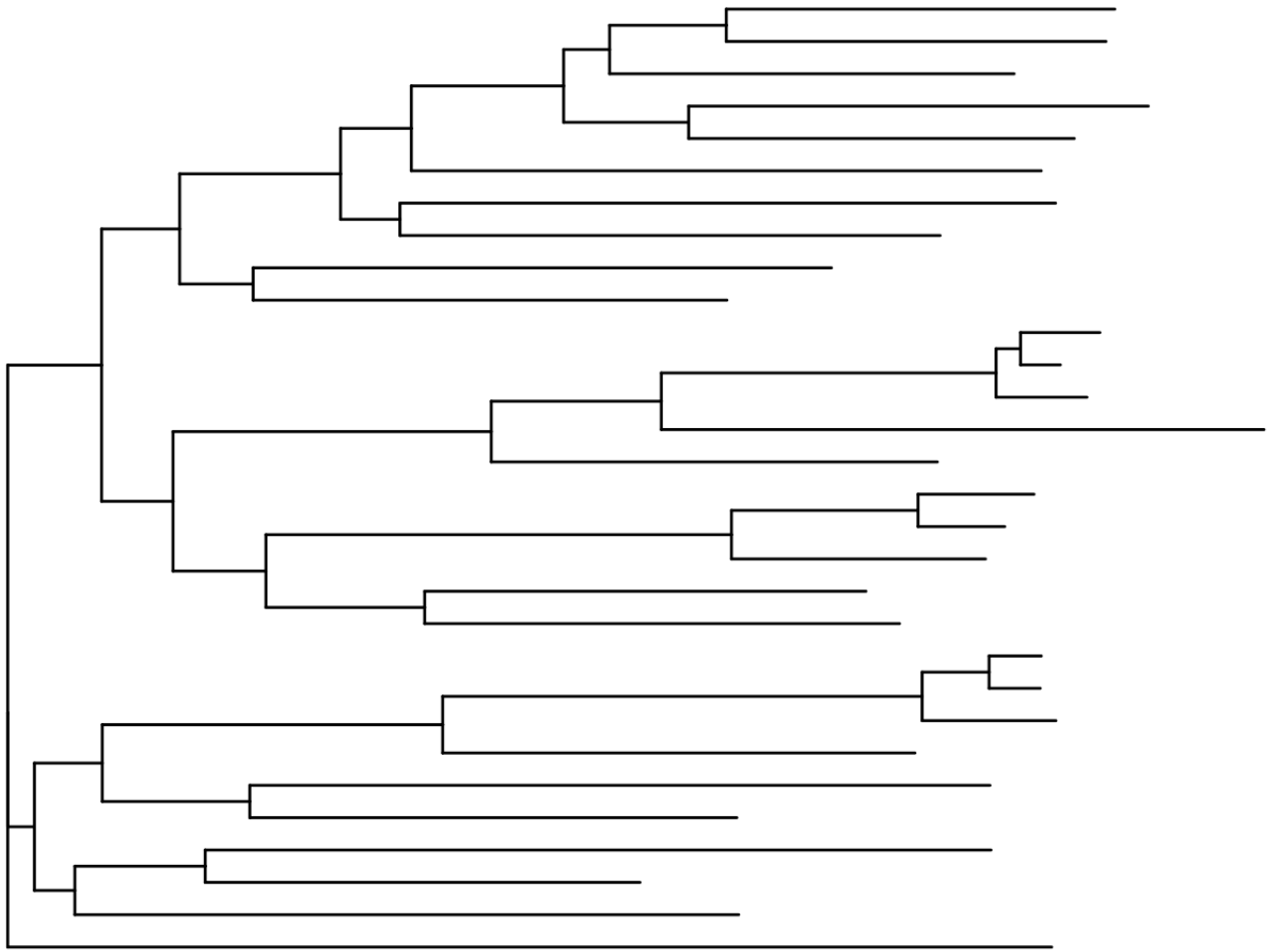


Figure 1. Neighbor-joining tree that clusters samples based on similarity of plant communities

```
Samples <- read.csv("Sample.csv", header=T)
```

## Colour code and annotate tree

```
ggtree(OTUtree,layout="rectangular") %<+% Samples +  
  geom_tiplab(aes(colour=Location)) + theme(legend.position = "right") + scale_colour_  
_discrete(name = "Location", labels = c("Inside garlic mustard patch", "Outside garli  
c mustard patch"))
```

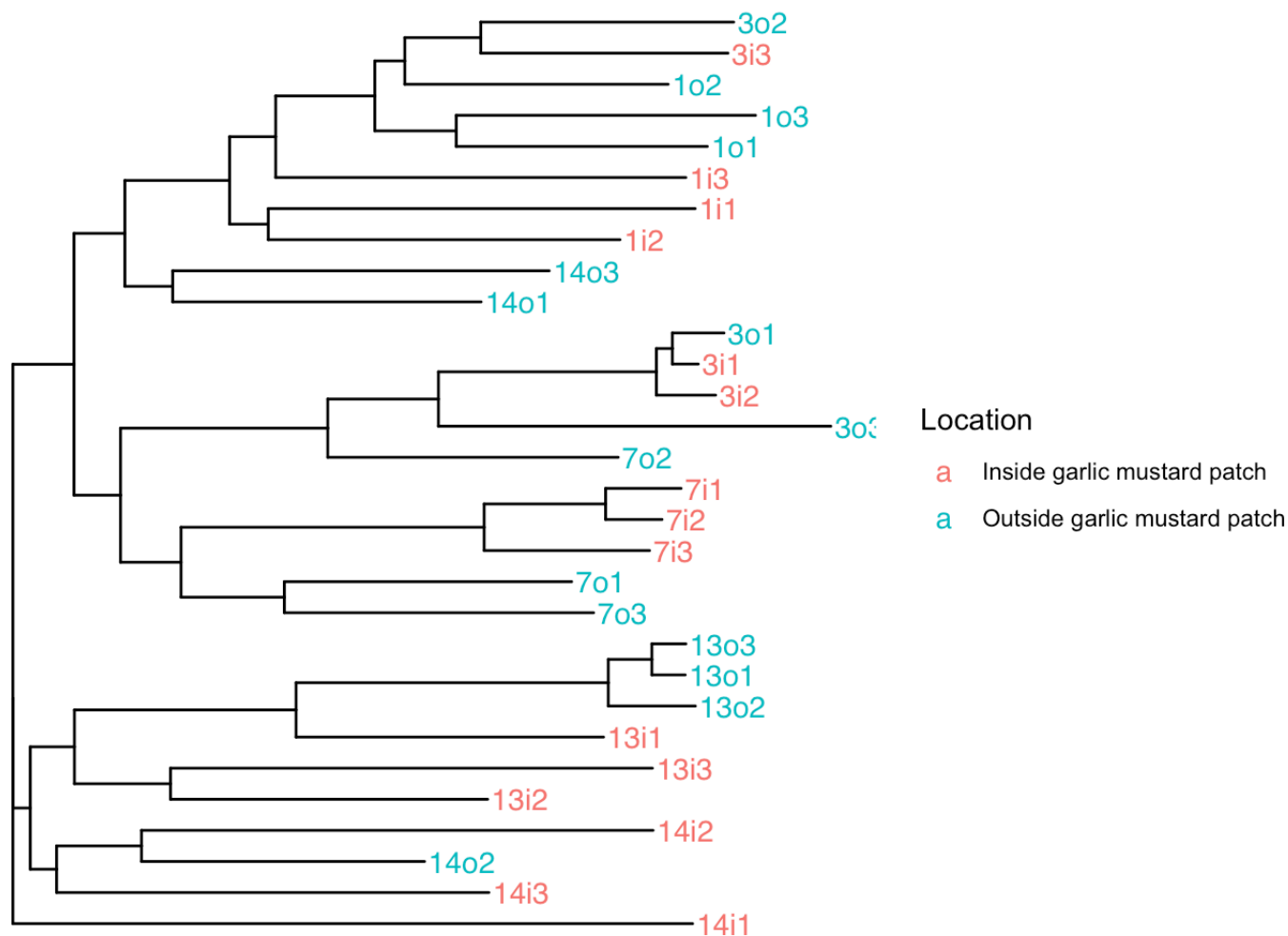


Figure 2. Neighbour-joining tree that clusters samples based on similarity of plant communities. The tip labels indicate the quadrat that the measurements were taken from. The colour code is based on location. Location is whether quadrat measurements are taken inside (red) or outside (blue) of garlic mustard patches.

Based on Figure 2, there is no clear differentiation based on location of the quadrat.

## NMDS

```
set.seed(123)
NMDSdat<-metaMDS(OTU_dist,k=2)
```

```
## Run 0 stress 0.1461135
## Run 1 stress 0.1461135
## ... Procrustes: rmse 0.0001326056 max resid 0.0003731696
## ... Similar to previous best
## Run 2 stress 0.1489956
## Run 3 stress 0.1414389
## ... New best solution
## ... Procrustes: rmse 0.06684359 max resid 0.309278
## Run 4 stress 0.1560521
## Run 5 stress 0.1414391
## ... Procrustes: rmse 0.0005635653 max resid 0.00213791
## ... Similar to previous best
## Run 6 stress 0.1461135
## Run 7 stress 0.1534441
## Run 8 stress 0.1414388
## ... New best solution
## ... Procrustes: rmse 0.0001057004 max resid 0.0004177689
## ... Similar to previous best
## Run 9 stress 0.1414389
## ... Procrustes: rmse 7.874346e-05 max resid 0.0002693626
## ... Similar to previous best
## Run 10 stress 0.1461135
## Run 11 stress 0.153444
## Run 12 stress 0.1666367
## Run 13 stress 0.1461135
## Run 14 stress 0.1560521
## Run 15 stress 0.1589506
## Run 16 stress 0.165655
## Run 17 stress 0.141439
## ... Procrustes: rmse 9.649305e-05 max resid 0.0003586765
## ... Similar to previous best
## Run 18 stress 0.1461136
## Run 19 stress 0.2068458
## Run 20 stress 0.1620716
## *** Best solution repeated 3 times
```

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

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

```
qplot(x=NMDS1, y=NMDS2, colour=Location, alpha=I(0.6), data=PDat) + theme_bw() + scale_
_colour_discrete(name = "Location", labels = c("Inside garlic mustard patch", "Outsid
e garlic mustard patch"))
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
```

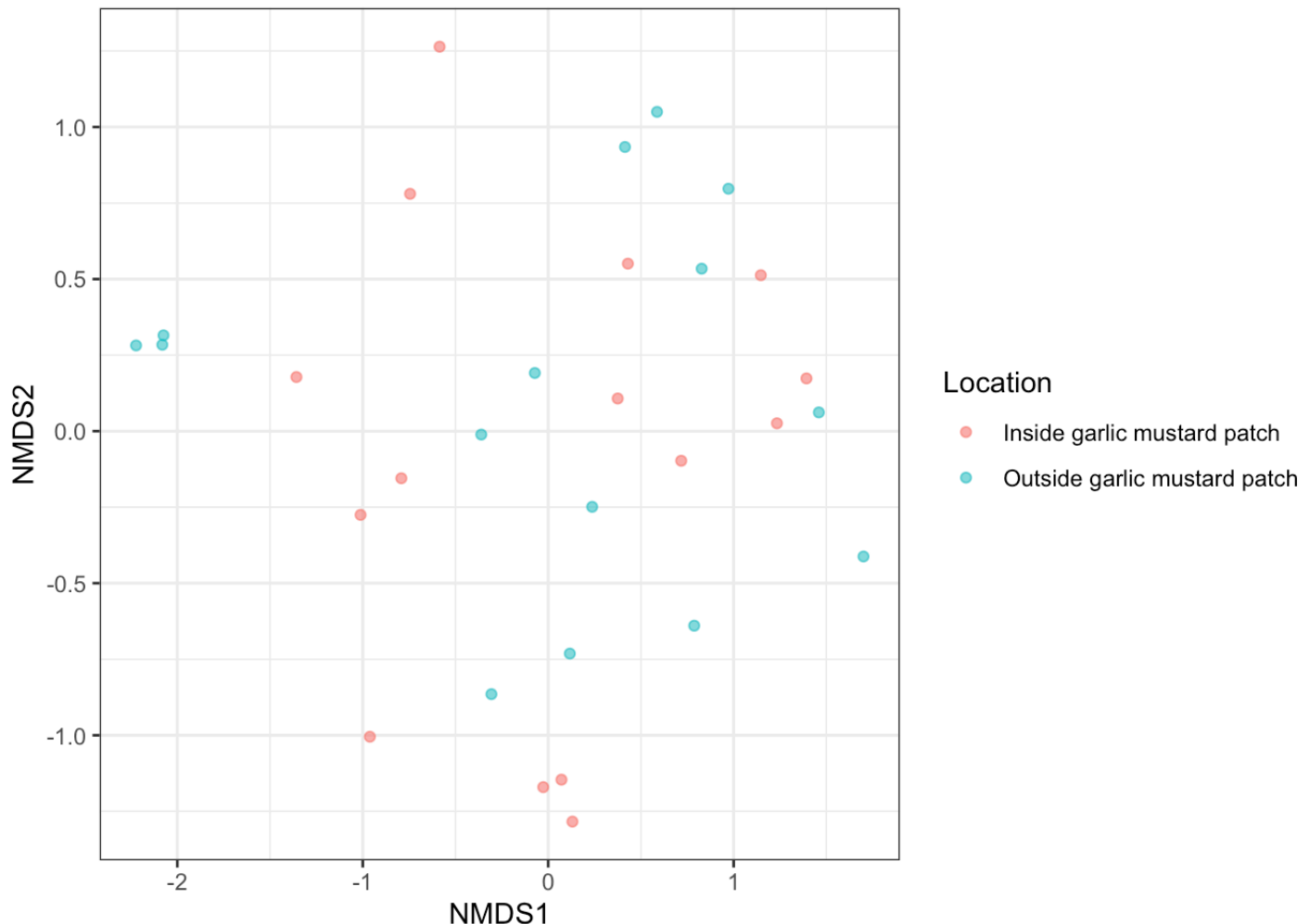


Figure 3. NMDS bivariate plot indicating how the plant communities are clustered based on their location. Location is whether quadrature measurements are taken inside (red) or outside (blue) of garlic mustard patches.

## Part II

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

## Cluster by similarity

```
ggtree(OTUtree,layout="rectangular") %<+% Samples +
  geom_tiplab(aes(colour=as.factor(Population))) + theme(legend.position = "right") +
  scale_colour_discrete(name = "Population", labels = c("Population 1", "Population 3",
"Population 7","Population 13","Population 14"))
```

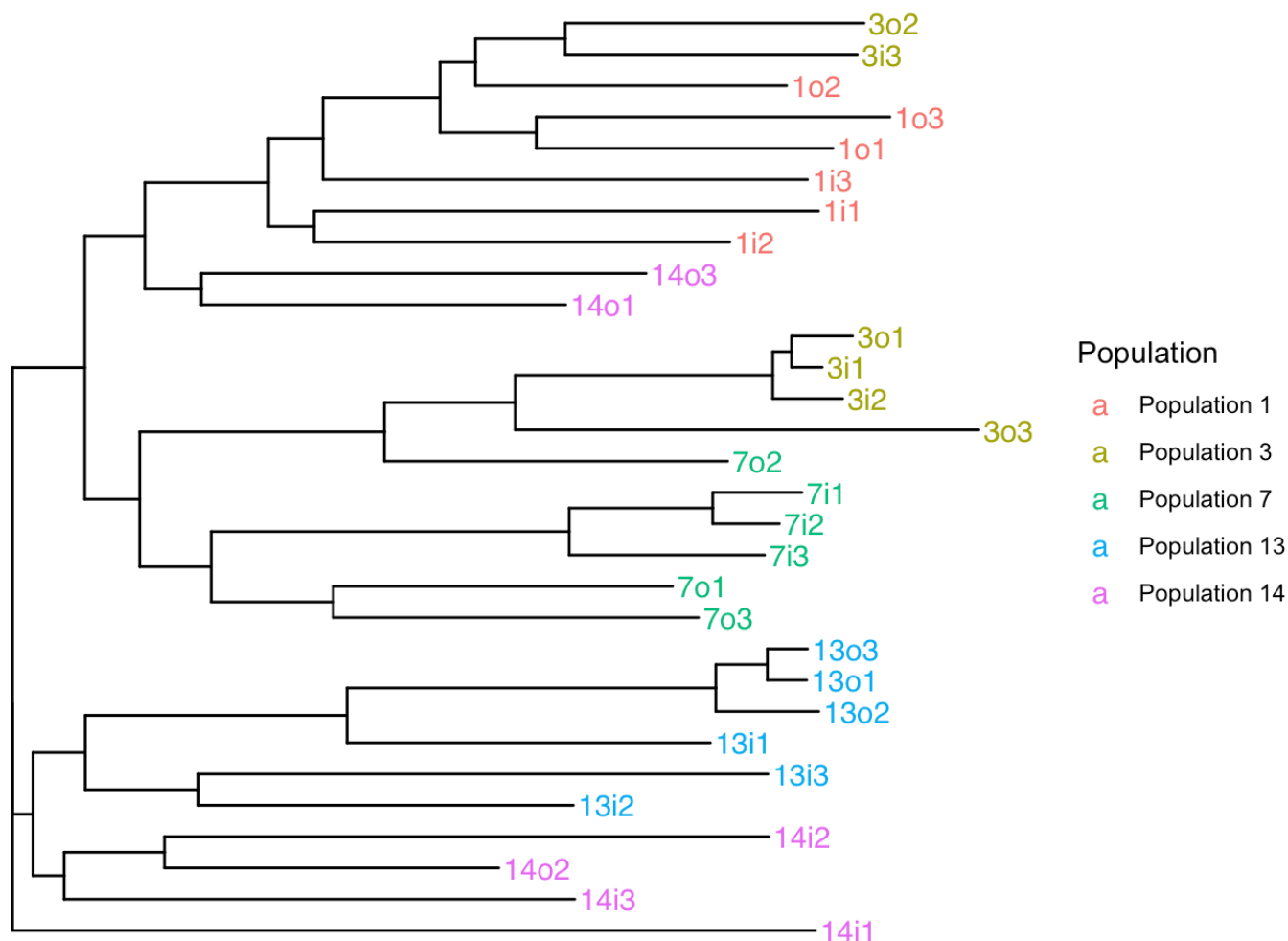


Figure 4. Neighbour-joining tree that clusters samples based on similarity of plant communities. The tip labels indicate the quadrat that the measurements were taken from. The colour code is based on sample population. Population is a code for a specific sampling location at QUBS, determined by presence of garlic mustard nearby. There are 5 different populations in this data set.

## NMDS

```
qplot(x=NMDS1, y=NMDS2, colour=as.factor(Population), alpha=I(0.6),data=PDat) + theme
_bw() + scale_colour_discrete(name = "Population", labels = c("Population 1", "Popula
tion 3", "Population 7","Population 13","Population 14"))
```



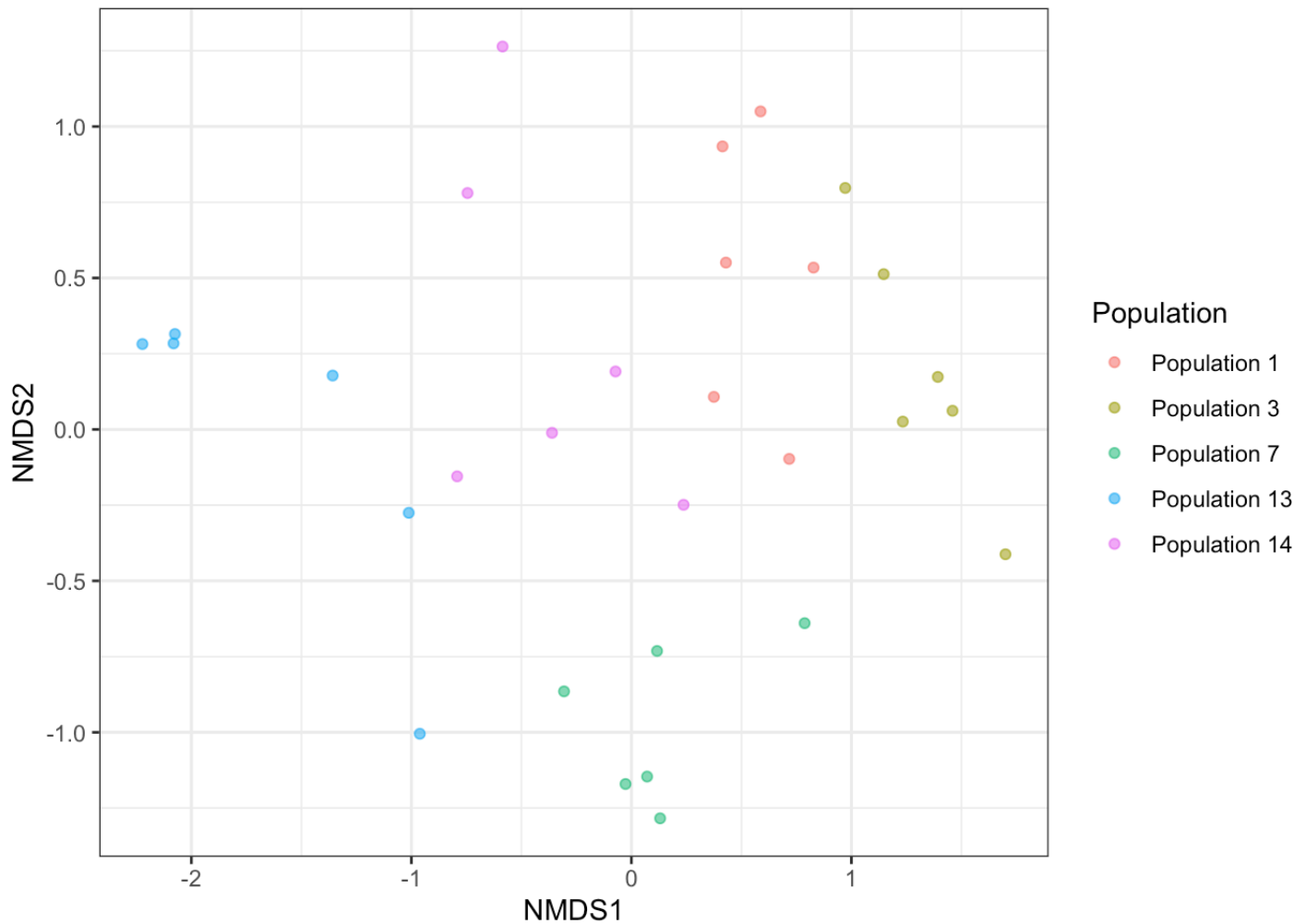


Figure 5. NMDS bivariate plot indicating how the plant communities are clustered based on their population. Population is a code for a specific sampling location at QUBS, determined by presence of garlic mustard nearby. There are 5 different populations in this data set.

When comparing Figure 3 and Figure 5, there is better clustering among the populations in Figure 5, than the locations in Figure 3. This suggests that sampling population has a stronger effect on plant communities than the location.