

Notes S3: Main analysis code

Natacha Bodenhausen, Vincent Somerville and Klaus Schlaeppli (ENTER YEAR and JOURNAL HERE)

2018-08-22

Contents

Abstract	2
Method	2
Data analysis	3
Data exploration	4
Results:	7
Alpha Diversity	7
Filter data	11
Using edgeR to normalize data (TMM) and to identify differentially abundant OTU	11
Beta diversity	12
Taxonomy	16
Identification of P-sensitive OTUs (EDGE R analysis)	19
Display of all P-sensitive OTUs per species using Ternary plots	20
Comparing P-sensitive OTUs between in Arabidopsis and Petunia	22
Network analysis	26
Keystone OTUs	37

Abstract

In this study, we investigate the interplay between bacteria and fungi that are associated with the roots of *Petunia* plants as a function of the P availability in soil. Additionally, we examine *Arabidopsis* plants, which secrete minute amounts of strigolactones and which are not colonized by AMF. Finally, we included soil samples to control for the effect of P levels on the microbial community in the absence of plants. The goals of this study are to understand the dynamics of the root microbiome in response to the P availability in soil and to clarify the role of strigolactones as signaling molecules in this process.

Method

Petunia and *Arabidopsis* were grown in pots filled with Soil from next to the FAST trial. Pots were watered with a P gradient: low, medium and high. Plants were harvested after 10 weeks. Two fractions for each sample were harvested: 1) the roots were washed three times with phosphate buffer and dried on filter paper, this fraction is called the 'root' community; 2) the three washes of each sample were pooled and centrifuged, the pellet is called the 'rhizosphere'. The roots and the rhizosphere samples were stored at -80C until DNA extraction. For this data set, only the root-associated community was analyzed. In addition, soil from unplanted pot was also collected. Together, there were 108 samples (72 *Petunia* samples, 24 *Arabidopsis* samples and 12 soil samples).

DNA was extracted from the roots by Natacha with NucleoSpin Soil kit from Macherey-Nagel. For the bacterial community, a fragment of the 16S rRNA gene was amplified with primers 799F and 1193R. For the fungal community, a fragment of the intergenic spacer region 1 was amplified with primers ITS1F and ITS2 which were found in a previous comparison to perform better than other primers pairs (higher richness, little contamination with plant DNA, high percentage of Glomeraceae) . Alain prepared the fungal library and Natacha prepared the bacterial library. After clean-up, the PCR products were pooled in equimolar concentrations and sequenced at the Functional Genomics Center in Zurich with MiSeq.

Data analysis

Data upload










Sample table

Table 1: samples

	low	medium	high
Arabidopsis	8	7	9
Petunia	7	8	9
Soil	4	4	4

Define colours

Arabidopsis: red Petunia: blue Soil: brown P-levels (low, medium & high) are indicated with color tone from light to dark

	low	medium	high
Arabidopsis			
Petunia			
Soil			

MySeq (16S): loading data from bioinformatic analysis

MySeq (ITS): loading data from bioinformatic analysis

Pacbio (ITS): loading data from bioinformatic analysis

Data exploration

Number of sequences per sample (data for beginning of results)

- 1) distribution of sequence counts
- 2) total amount of sequences
- 3) number of OTUs

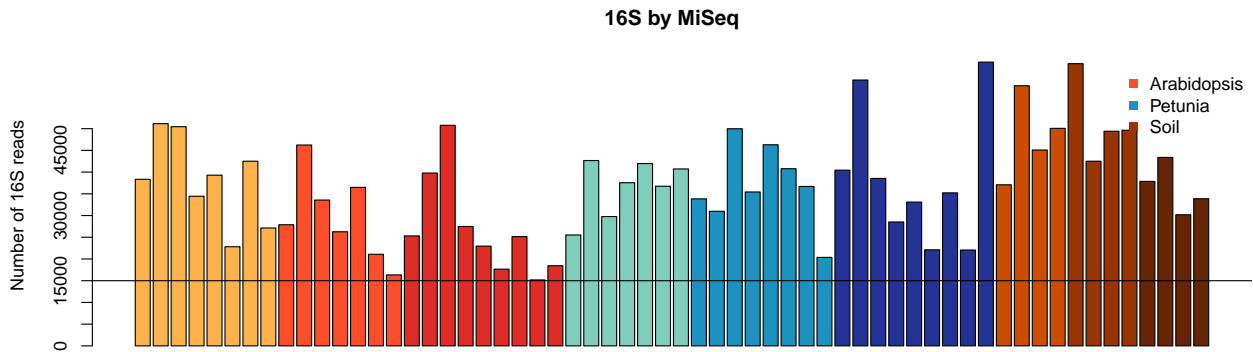
16S by MiSeq

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 15177 27396 36718 36605 42840 65319
```

```
## [1] 2196310
```

```
## [1] 3701
```

Number of sequences per sample. Black horizontal line show the threshold used for rarefying (15'000).



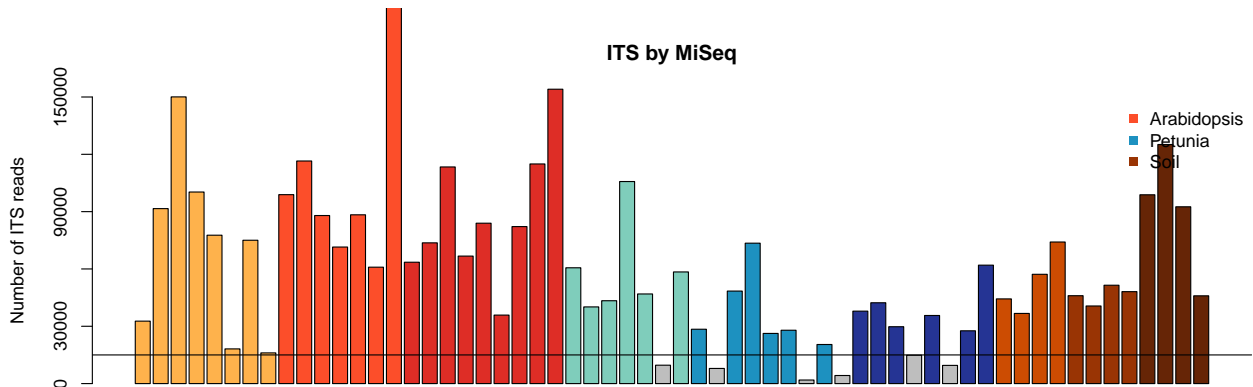
ITS by MiSeq

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 1867 34906 54337 63489 84950 311084
```

```
## [1] 3809350
```

```
## [1] 1688
```

Number of sequences per sample. Black horizontal line show the threshold used for rarefying (15'000). Five samples colored in grey were removed from the analyses of rarefied data.

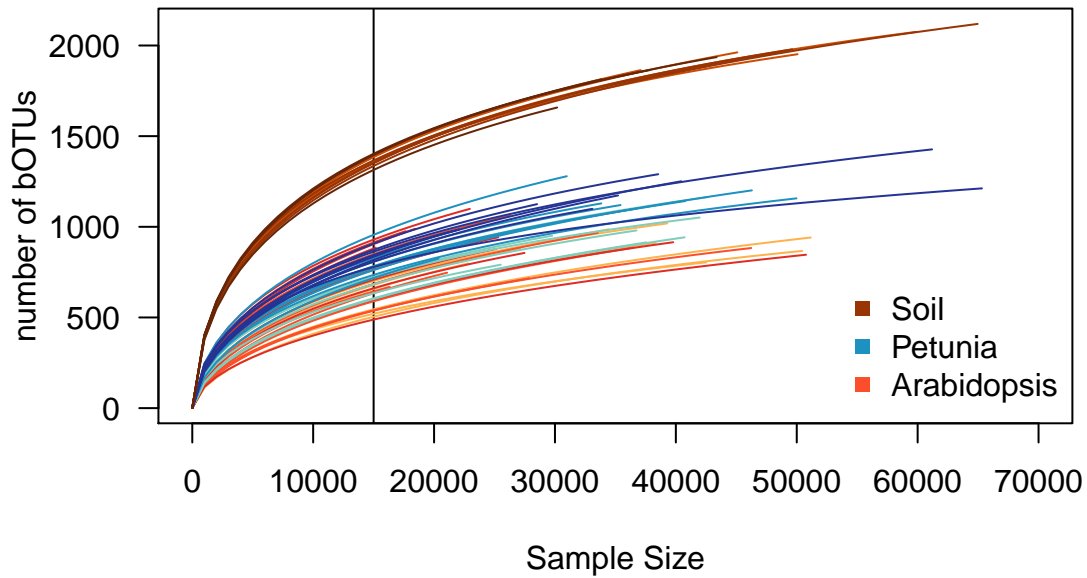


Rarefaction analysis

16S bacteria data

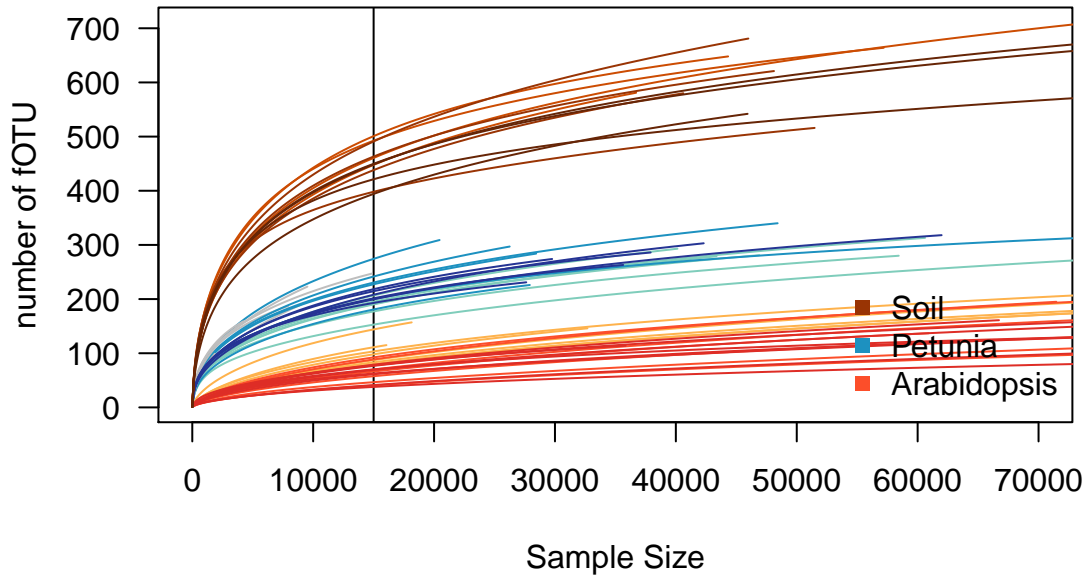
16S Rarefaction analysis. Diversity is highest in the soil followed by Petunia and then Arabidopsis:

16S



ITS fungi data

ITS Rarefaction analysis. Diversity is highest in the soil followed by Petunia and then Arabidopsis:



Results:

Alpha Diversity

Subset roots samples

Alpha Diversity: how it is done

- concentrate on plant samples (remove soil samples)
- For 16S, there are no samples below threshold (15000)
- For ITS, there are 6 samples below threshold (15000)
- rarefy to 15000 multiple times (at the moment set to 500)
- for each subsample, estimate richness, diversity and evenness
- richness=number of OTUs
- from Jost 2006: diversity is defined as $\exp(\text{shannon})$, shannon is calculated with diversity function of vegan package
- sheldon evenness is $\exp(H)$ over richness (from Sheldon 1069)
- 2-factor ANOVA with plant species and P treatment (order matters)
- diagnostic plots are saved into plots folder
- after checking diagnostic plots, data are log-transformed for 4/6 cases so that assumptions are fulfilled
- in the end decide to present and discuss in the text only S and E (not $\exp(H)$), because $E = \exp(H)/S$.

Compare general fungal and bacterial richness

- 1) average bOTU richness
- 2) average fOTU richness
- 3) ratio fOTU/bOTU richness

```
## Arabidopsis      Petunia
##      75.95958    205.70078
```

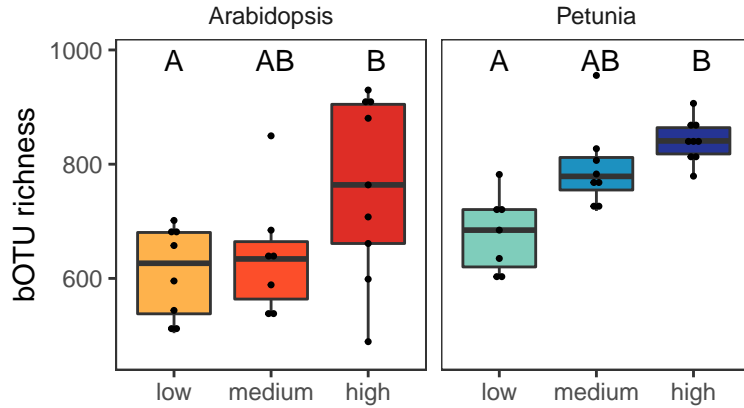
```
## Arabidopsis      Petunia
##      675.5473     778.2564
```

```
## Arabidopsis      Petunia
##      8.893510     3.783439
```

Richness 16S

```
##                Df Sum Sq Mean Sq F value    Pr(>F)
## Species         1  0.2827  0.28267   14.759 0.000406 ***
## Treatment       2  0.3659  0.18297    9.554 0.000380 ***
## Species:Treatment 2  0.0315  0.01575    0.822 0.446433
## Residuals      42  0.8044  0.01915
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

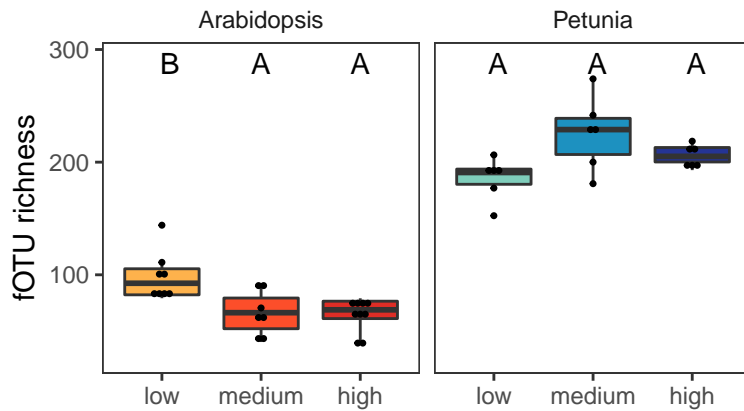
S: *** T: *** SxT:
```



Richness ITS

```
##                Df Sum Sq Mean Sq F value    Pr(>F)
## Species         1 11.078   11.078 242.625 <0.0000000000000002 ***
## Treatment       2  0.332    0.166   3.634   0.0365 *
## Species:Treatment 2  0.768    0.384   8.415   0.0010 **
## Residuals      36  1.644    0.046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

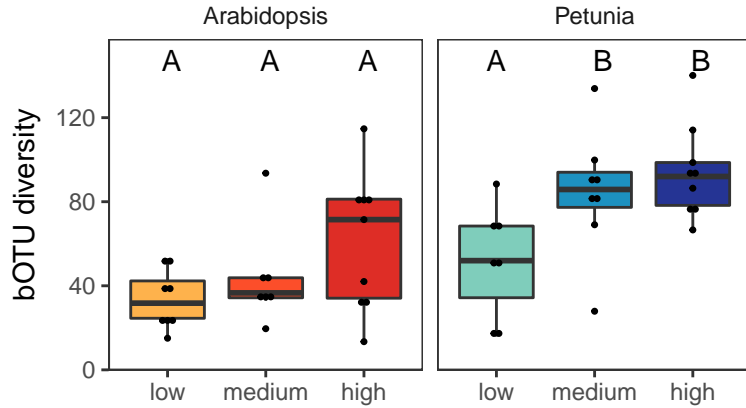
S: *** T: * SxT: **
```



Shannon diversity 16S

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Species          1  11989   11989  18.241 0.000109 ***
## Treatment        2   9768   4884   7.431 0.001726 **
## Species:Treatment 2    982    491   0.747 0.480109
## Residuals       42 27605    657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

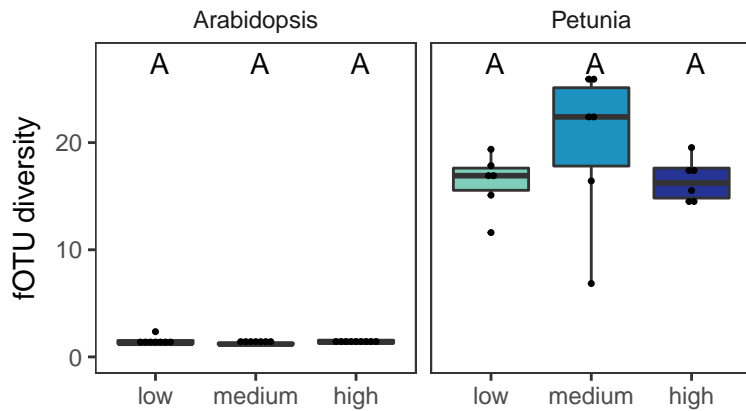
S: *** T: ** SxT:
```



Shannon diversity ITS

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Species          1  65.22   65.22 1078.485 <0.0000000000000002 ***
## Treatment        2   0.00    0.00   0.004    0.996
## Species:Treatment 2   0.12    0.06   1.002    0.377
## Residuals       36   2.18    0.06
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

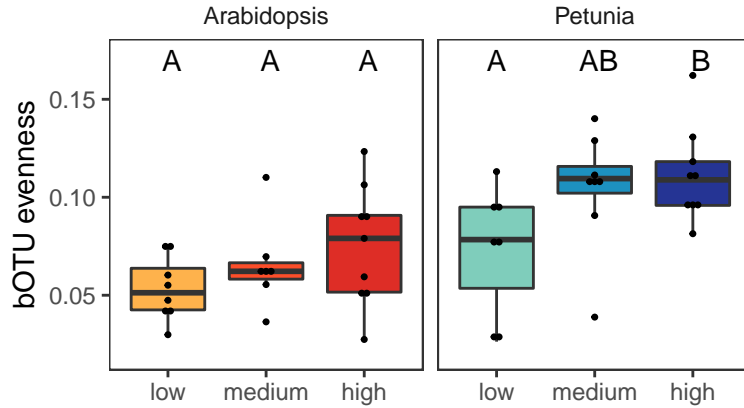
S: *** T: SxT:
```



Sheldon evenness 16S

```
##              Df  Sum Sq Mean Sq F value  Pr(>F)
## Species      1  0.012990 0.012990  18.248 0.000109 ***
## Treatment    2  0.007412 0.003706   5.206 0.009552 **
## Species:Treatment  2  0.000768 0.000384   0.539 0.587132
## Residuals    42  0.029898 0.000712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

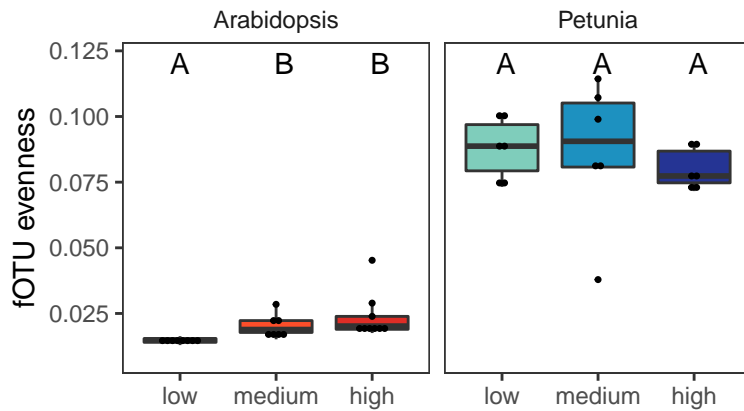
S: *** T: ** SxT:
```



Sheldon evenness ITS

```
##              Df Sum Sq Mean Sq F value      Pr(>F)
## Species      1 22.414  22.414 406.571 <0.0000000000000002 ***
## Treatment    2   0.338   0.169   3.067   0.0589 .
## Species:Treatment  2   0.499   0.250   4.527   0.0176 *
## Residuals    36   1.985   0.055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

S: *** T: . SxT: *
```



Filter data

- for beta diversity
- to identify differentially abundant OTUs (edgeR)
- Charlotte's filtering strategy: We exclude all OTUs that don't have at least 4 assigned reads in at least 4 samples (the size of the smallest group in our data).
- For 16S, there are 1456 OTU left for the analysis. This corresponds to about a third of the OTUs: 39.34 of the initial OTUs (3701)
- For ITS, there are 508 OTU left for the analysis. This corresponds to 30.09% of the initial OTUs (1688)
- For Pacbio, there are 109 OTU left for the analysis. This corresponds to 32.54% of the initial OTUs (335)
- we do NOT rarefy for beta diversity and the rest of the analysis
- we use TSS normalization for beta-diversity
- we use TMM normalization for edgeR
- what about barplots, ternary plot and plot individual OTUs, do we use TSS or TMM ? to be decided.

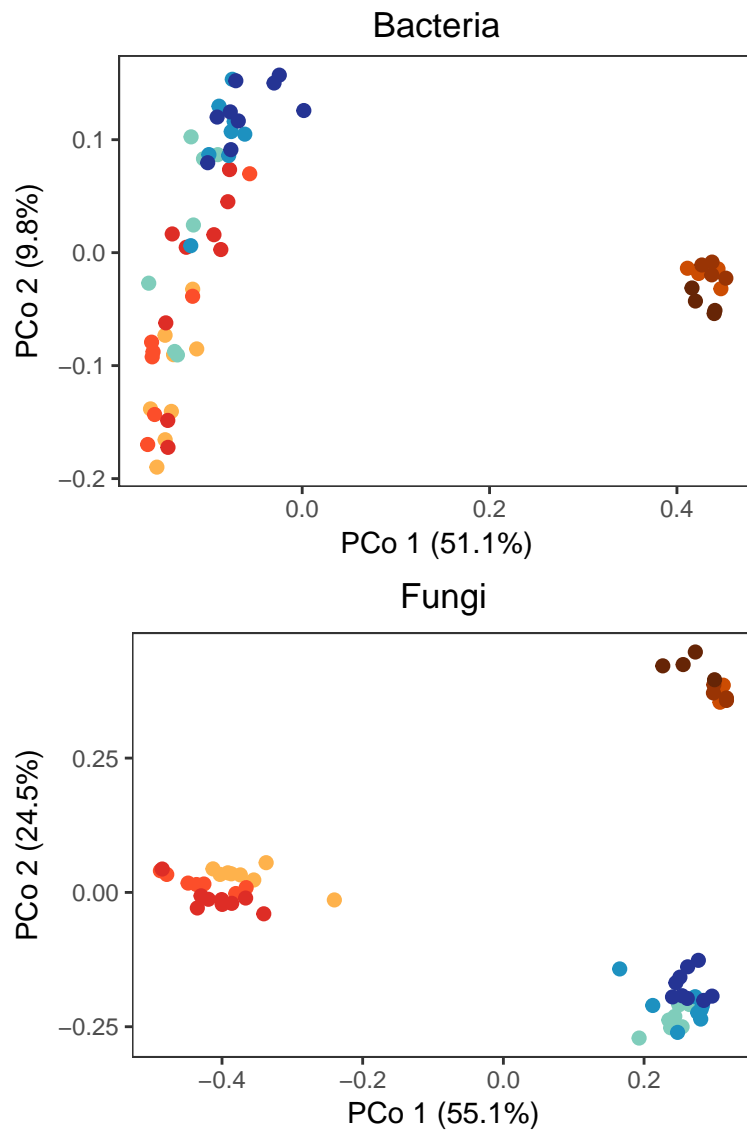
Using edgeR to normalize data (TMM) and to identify differentially abundant OTU

- work also with soil samples for this analysis
- filtered data

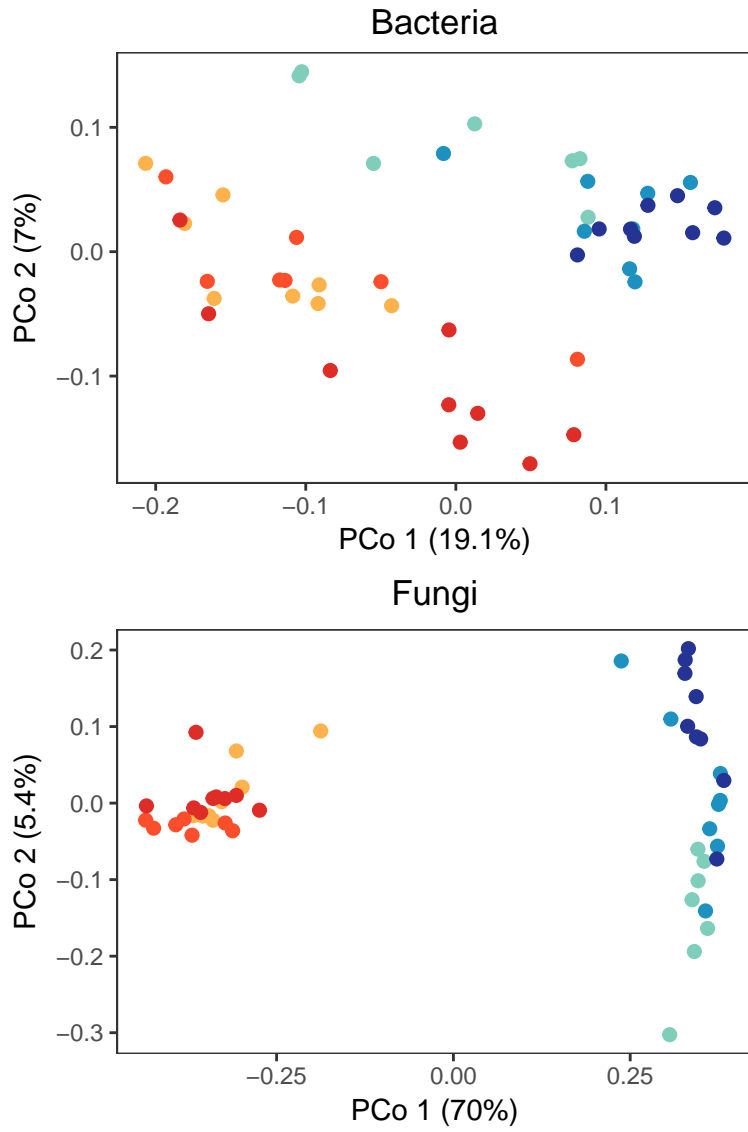
Beta diversity

- two plots for each 16S and ITS, first with soil (supplementary) and without soil samples (main text)
- data normalized (TSS, log=FALSE)
- like Leff et al., use square-root transformed data.
- Bray-Curtis dissimilarity (binary=FALSE)
- PCoA
- Phyloseq package

Principle Coordinate Analysis, WITH SOIL SAMPLES (Figure S5)



Principle Coordinate Analysis, ROOT SAMPLES (Figure 4)



PERMANOVA (Tables S2 and S3)

adonis plant 16S

Table 2: Table continues below

	R2
sampleTable_plant_16S\$Species	0.1412
sampleTable_plant_16S\$Treatment	0.07188
sampleTable_plant_16S <i>Species : sampleTable_plant_16S</i> Treatment	0.05336
Residuals	0.7336
Total	1
NA	NA

	Pr(>F)
sampleTable_plant_16S\$Species	0.001
sampleTable_plant_16S\$Treatment	0.003
sampleTable_plant_16S <i>Species : sampleTable_plant_16S</i> Treatment	0.028
Residuals	NA
Total	NA
NA	NA

adonis only Arabidopsis 16S

	R2	Pr(>F)
sampleTable_plant_16S_ara\$Treatment	0.1502	0.001
Residuals	0.8498	NA
Total	1	NA

adonis only Petunia 16S

	R2	Pr(>F)
sampleTable_plant_16S_pet\$Treatment	0.1412	0.004
Residuals	0.8588	NA
Total	1	NA

adonis plant ITS

Table 6: Table continues below

	R2
sampleTable_plant_ITS\$Species	0.5306
sampleTable_plant_ITS\$Treatment	0.03846
sampleTable_plant_ITS\$Species : sampleTable_plant_ITS\$Treatment	0.03748
Residuals	0.3934
Total	1
NA	NA

	Pr(>F)
sampleTable_plant_ITS\$Species	0.001
sampleTable_plant_ITS\$Treatment	0.07
sampleTable_plant_ITS\$Species : sampleTable_plant_ITS\$Treatment	0.064
Residuals	NA
Total	NA
NA	NA

adonis only Arabidopsis ITS

	R2	Pr(>F)
sampleTable_plant_ITS_ara\$Treatment	0.217	0.002
Residuals	0.783	NA
Total	1	NA

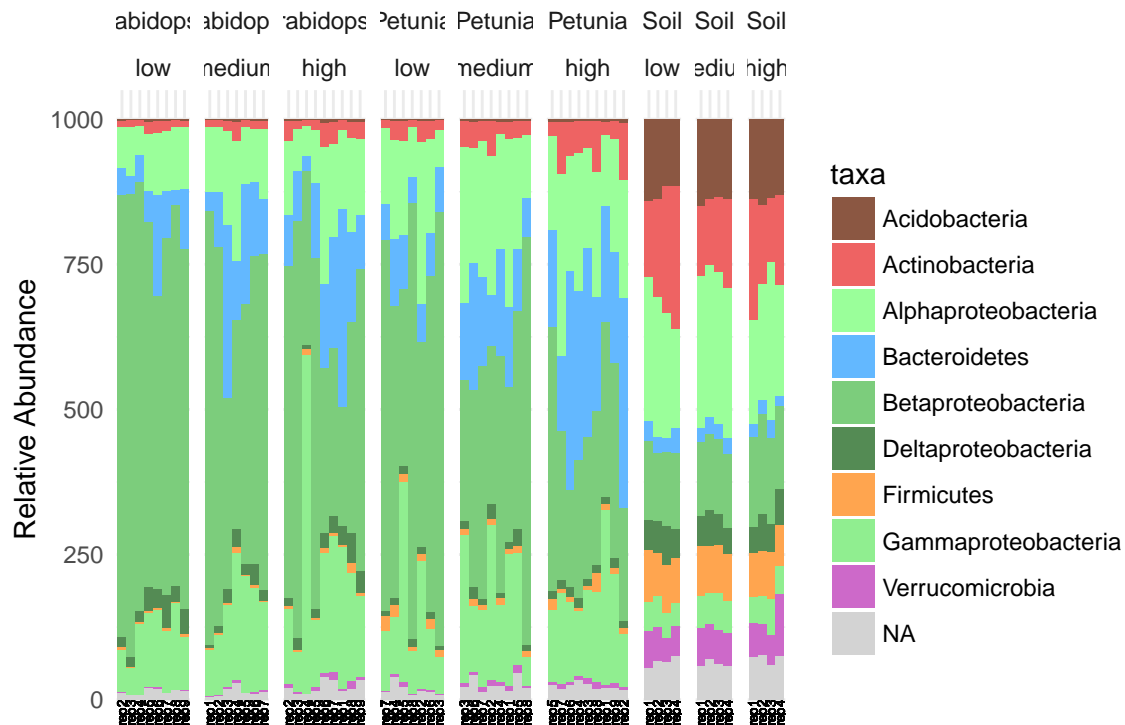
adonis only Petunia ITS

	R2	Pr(>F)
sampleTable_plant_ITS_pet\$Treatment	0.1327	0.053
Residuals	0.8673	NA
Total	1	NA

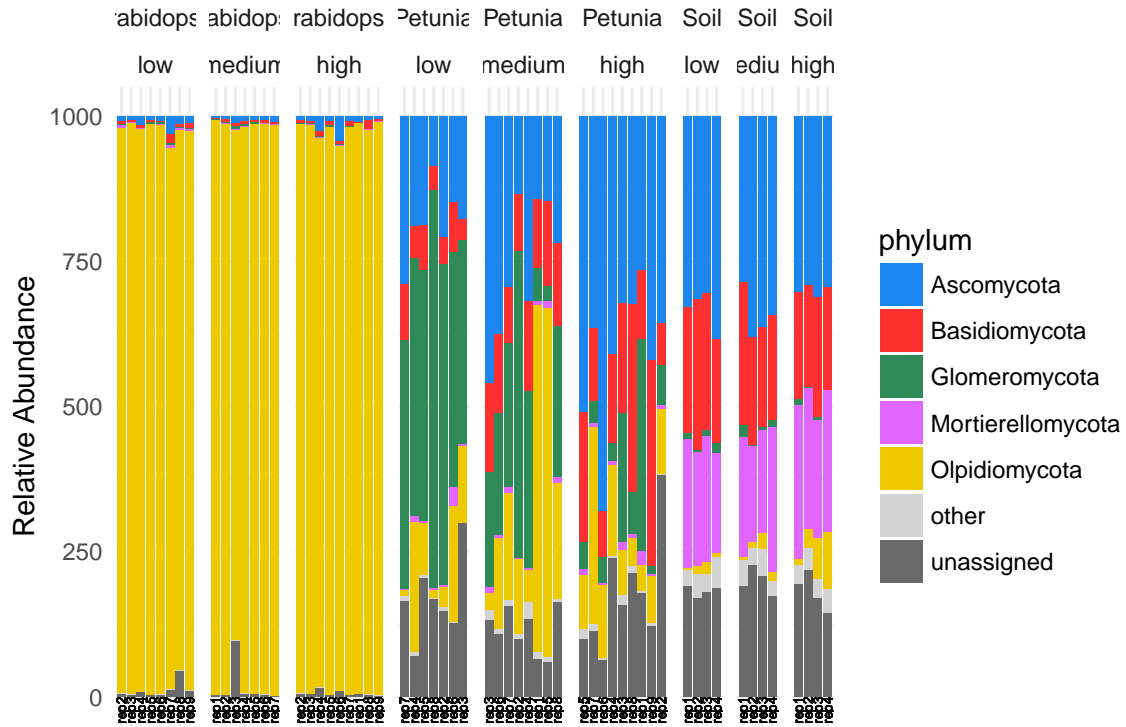
Taxonomy

16S Barplot at phylum level

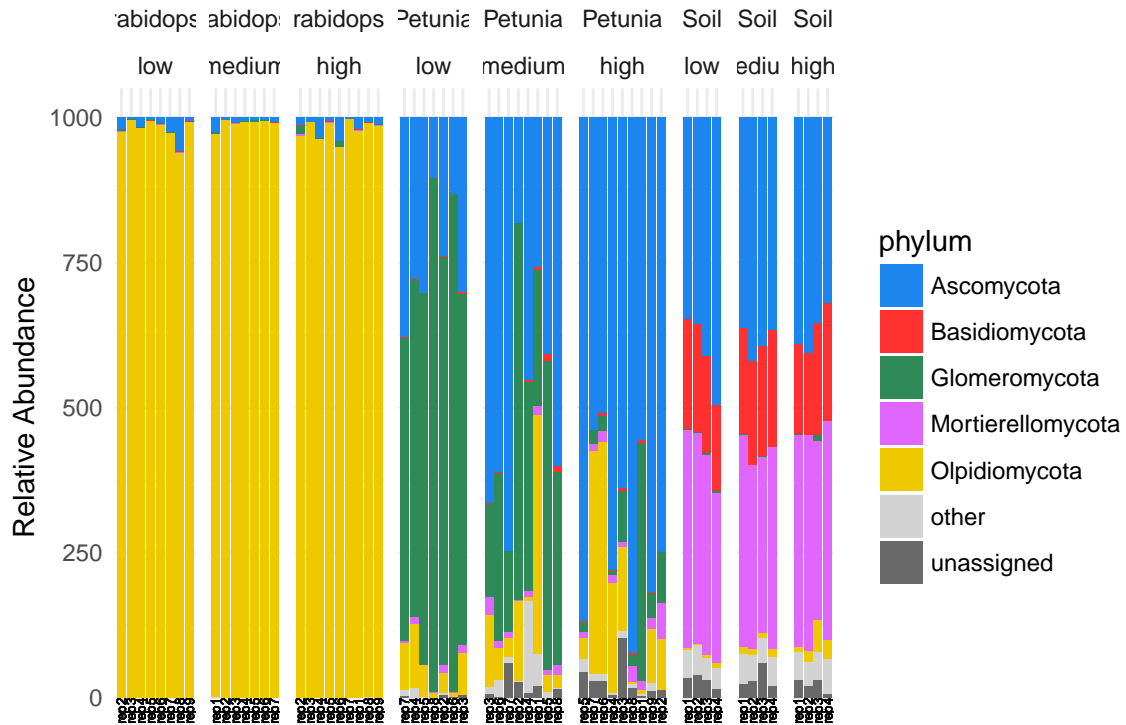
- We use TSS normalized counts that are expressed as permilles for taxonomic data display.



ITS Barplot at phylum level



PACBIO Barplot at phylum level



Abundant *Olpidium* in fungal data

MiSeq data

```
##           Phylum      Order      Genus
## fOTU_2  Olpidiomyota Olpidiales  Olpidium
## fOTU_1  Olpidiomyota unassigned unassigned
## fOTU_527 Olpidiomyota Olpidiales  Olpidium
## fOTU_318 Olpidiomyota Olpidiales unassigned
##                                           Species
## fOTU_2  Olpidium_brassicae_SH216672.07FU
## fOTU_1                                     unassigned
## fOTU_527                                    unassigned
## fOTU_318                                    unassigned

##      OTU  Arabidopsis      Petunia      Soil
## 1  fOTU_2  21.837255409  147.166892805  9.08474461
## 2  fOTU_1  946.385653888   8.256252202  16.81503811
## 3  fOTU_527  0.007009403   0.166890898  0.04063825
## 4  fOTU_318  0.012087718   0.008079231  0.37838979
```

Pacbio data

```
##           Phylum      Order      Genus                                           Species
## pOTU_1  Olpidiomyota Olpidiales  Olpidium  Olpidium_brassicae_SH194417.07FU
## pOTU_3  Olpidiomyota Olpidiales  Olpidium  Olpidium_brassicae_SH216672.07FU

##      OTU  Arabidopsis  Petunia      Soil
## 1  pOTU_1  961.78861  43.79451  14.2364506
## 2  pOTU_3  20.26482  61.62742  0.4885352

##                                           species  Arabidopsis  Petunia      Soil
## 1  Olpidium_brassicae_SH194417.07FU  961.78861  43.79451  14.2364506
## 2  Olpidium_brassicae_SH216672.07FU   20.26482  61.62742  0.4885352
```

Identification of P-sensitive OTUs (EDGE R analysis)

- We use TMM transformed data
- Same filtering parameter as for beta diversity

high P vs low P in Arabidopsis 16S

	up in low P	unchanged	up in high P
Treatmenthigh	19	1411	26

A total of 1456 bOTUs were analyzed for differential abundance in Arabidopsis, of which 19 were increased in abundance under low P and 26 were increased in abundance under high P, corresponding to 3.1% of P responsive OTUs.

high P vs low P in Arabidopsis ITS

	up in low P	unchanged	up in high P
Treatmenthigh	49	445	14

A total of 508 fOTUs were analyzed for differential abundance in Arabidopsis, of which 49 were increased in abundance under low P and 14 were increased in abundance under high P, corresponding to 12% of P responsive OTUs.

high P vs low P in Petunia 16S

	up in low P	unchanged	up in high P
Treatmenthigh	17	1424	15

A total of 1456 bOTUs were analyzed for differential abundance in Petunia, of which 17 were increased in abundance under low P and 15 were increased in abundance under high P, corresponding to 2.2% of P responsive OTUs.

high P vs low P in Petunia ITS

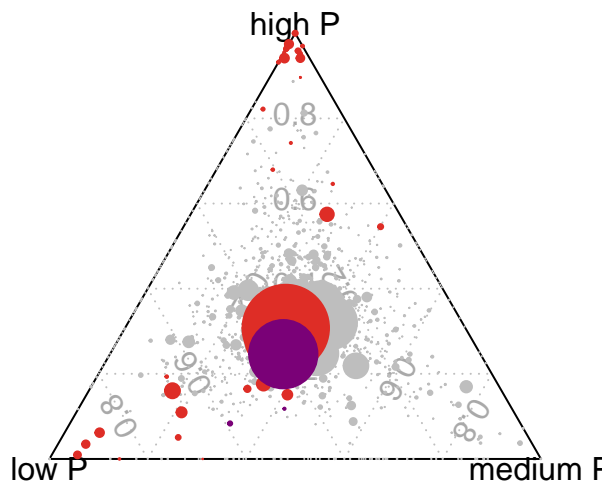
	up in low P	unchanged	up in high P
Treatmenthigh	23	477	8

A total of 508 fOTUs were analyzed for differential abundance in Petunia, of which 23 were increased in abundance under low P and 8 were increased in abundance under high P, corresponding to 6.1% of P responsive OTUs.

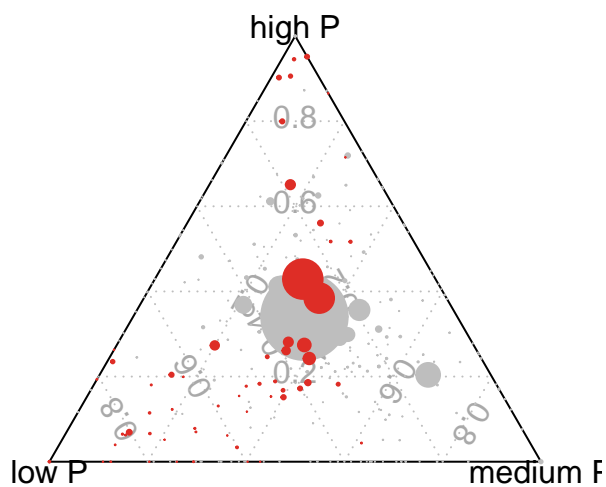
Display of all P-sensitive OTUs per species using Ternary plots

- For the display of the statistic results, we used TMM normalized counts (as used for edgeR) and express them as permilles.
- Ternary plots used scaled OTU abundances ($\log_2(x + 1)$)

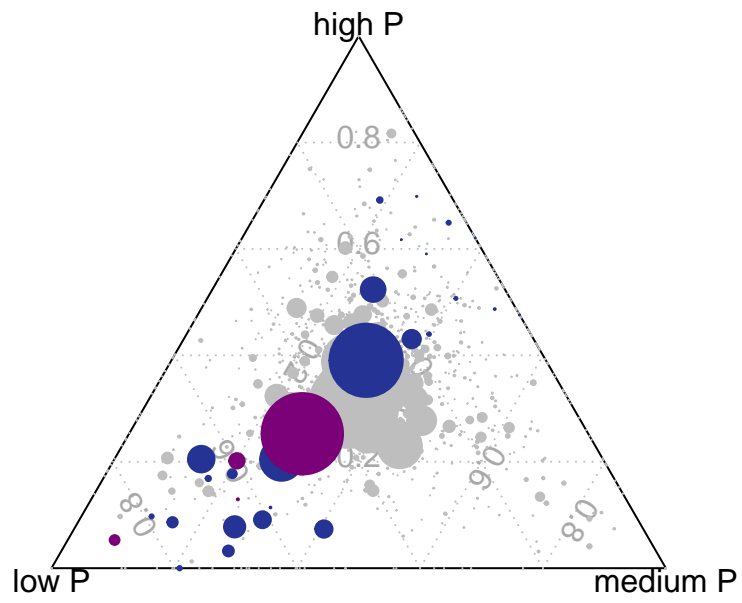
Arabidopsis 16S



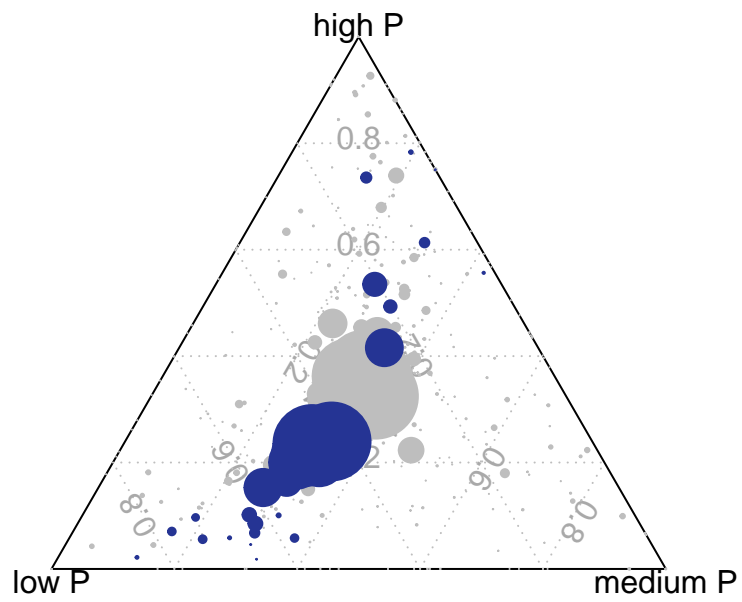
Arabidopsis ITS



Petunia 16S



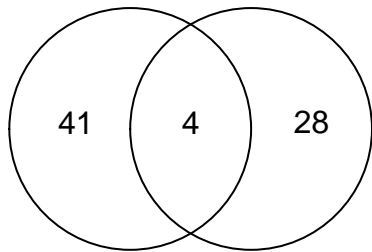
Petunia ITS



Comparing P-sensitive OTUs between in Arabidopsis and Petunia

Bacteria (16S)

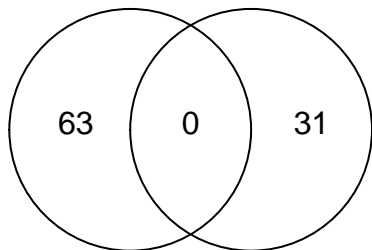
Arabidopsis Petunia



```
## [1] "bOTU_2" "bOTU_234" "bOTU_1787" "bOTU_3346"
```

fungi (ITS)

Arabidopsis Petunia



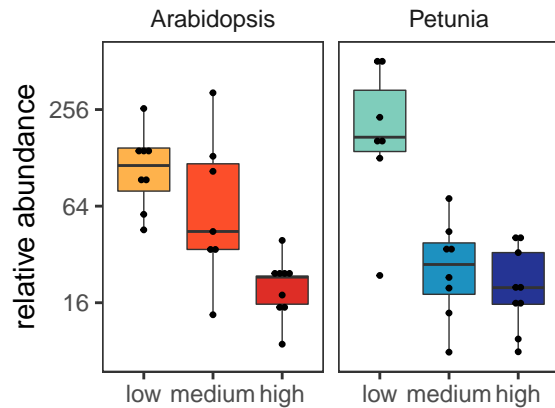
A few representative OTUs

- For the display of the P-sensitive OTUs, we used TSS normalized counts and express them as permilles.
- Y-axis of the Boxplots present the permille values scaled by log2.

bOTU_2

```
##                Phylum                Class                Order                Family
## b0TU_2 Proteobacteria Betaproteobacteria Rhodocyclales Rhodocyclaceae
##                Genus                Species
## b0TU_2 Dechloromonas unassigned
```

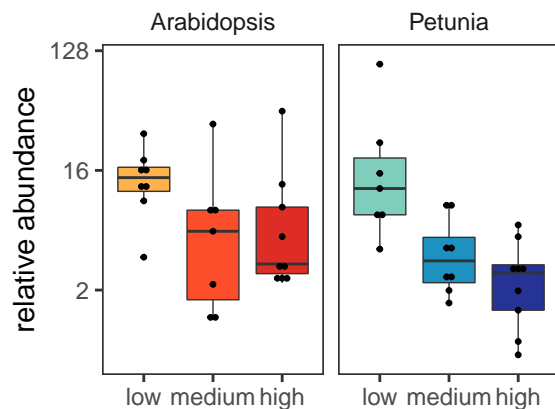
bOTU2 Dechloromonas sp.



bOTU_13

```
##                Phylum                Class                Order                Family
## b0TU_13 Proteobacteria Betaproteobacteria Rhodocyclales Rhodocyclaceae
##                Genus                Species
## b0TU_13 unassigned uncultured_bacterium
```

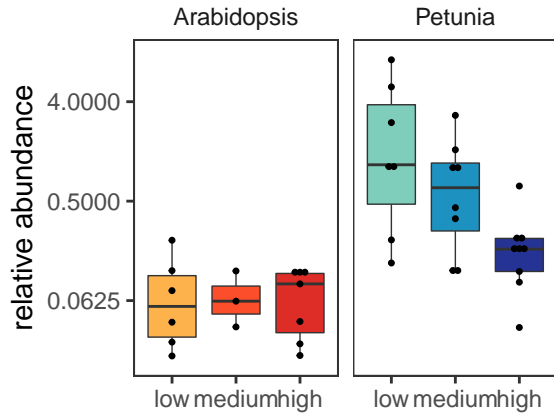
bOTU13 Rhodocyclaceae sp.



bOTU_134

```
##                Phylum                Class                Order
## b0TU_134 Proteobacteria Betaproteobacteria Burkholderiales
##                Family                Genus                Species
## b0TU_134 Burkholderiaceae unassigned unassigned
```

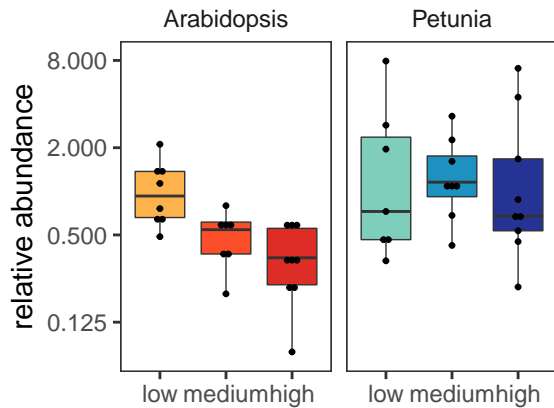
bOTU134 Candidatus Glomeribacter



fOTU_7

```
##                Phylum                Class                Order                Family
## f0TU_7 Mortierellomycota Mortierellomycetes Mortierellales Mortierellaceae
##                Genus                Species
## f0TU_7 Mortierella unassigned
```

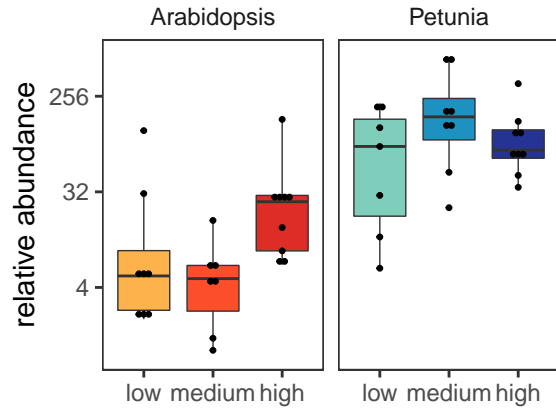
fOTU7 Mortierella sp.



fOTU_2

```
##           Phylum           Class           Order           Family           Genus
## f0TU_2  Olpidiomycota  Olpidiomycetes  Olpidiales  Olpidiaceae  Olpidium
##                                     Species
## f0TU_2  Olpidium_brassicae_SH216672.07FU
```

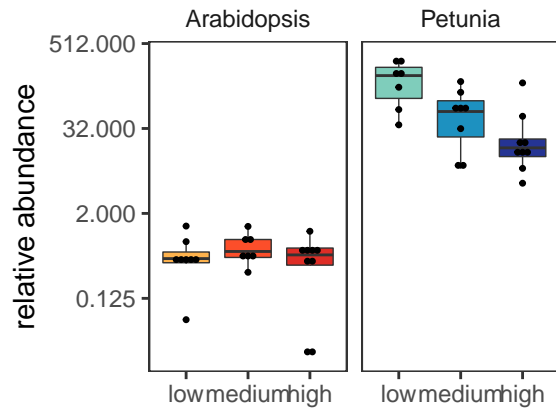
fOTU2 *Olpidium brassicae*



fOTU_6 (Example for P-sensitive OTU in Petunia)

```
##           Phylum           Class           Order           Family           Genus
## f0TU_6  Glomeromycota  Glomeromycetes  Glomerales  Glomeraceae  unassigned
##                                     Species
## f0TU_6  unassigned
```

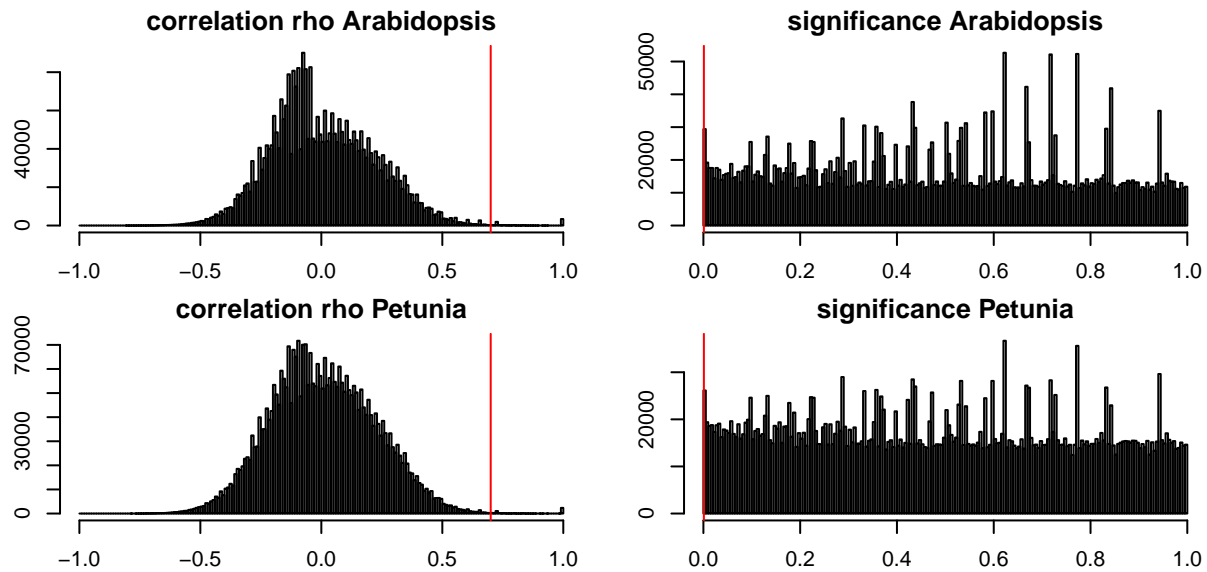
fOTU6 *Glomeraceae* sp.



Summary tables for identified P-sensitive OTUs (EDGE R analysis)

Network analysis

Perform Spearman rank correlation



Network construction is based on Spearman rho correlation cutoff >0.7 and fdr adjusted p.value cutoff < 0.001 .

In Arabidopsis 2652 links between 995 OTUs (nodes) fulfill the cutoffs. In Petunia 1467 links between 918 fulfill the cutoffs.

Number of bacteria and fungi nodes

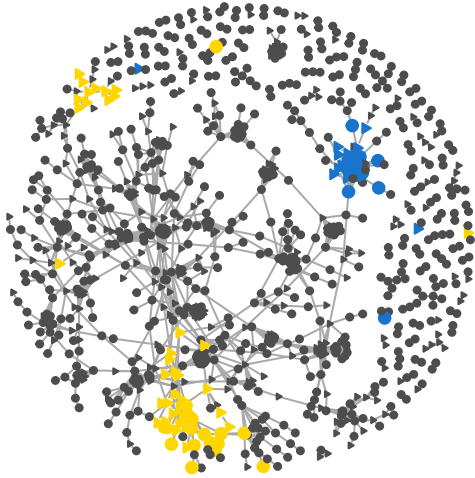
The Arabidopsis co-occurrence network is composed of 670 bacteria and 325 fungal OTUs (nodes). The Petunia co-occurrence network is composed of 656 bacteria and 262 fungal OTUs (nodes).

Defining node shape and color

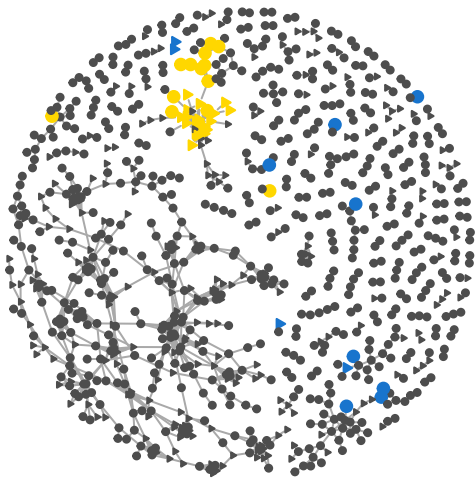
Defining P behaviour of microbes to be used as node attributes (response to phosphate treatments based on three categories: positive, negative and no P-dependence)

Co-occurrence network plots

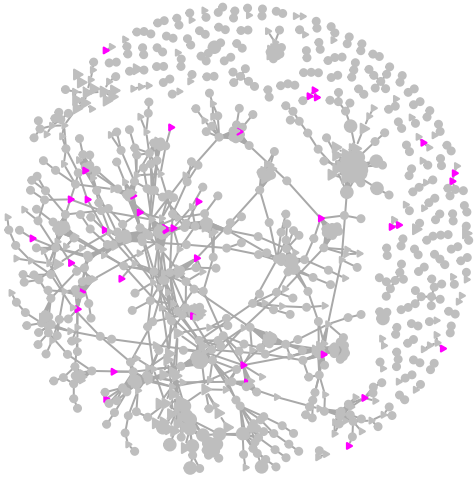
Arabidopsis



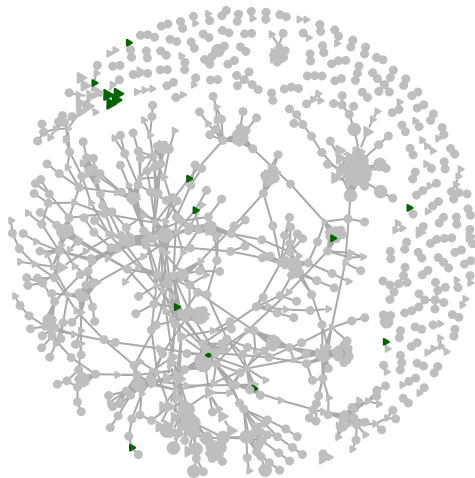
Petunia



AMF fOTUs in Arabidopsis network

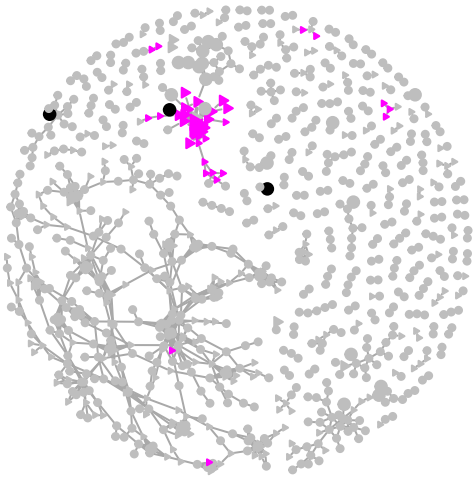


Mortierellomycota fOTUs in Arabidopsis

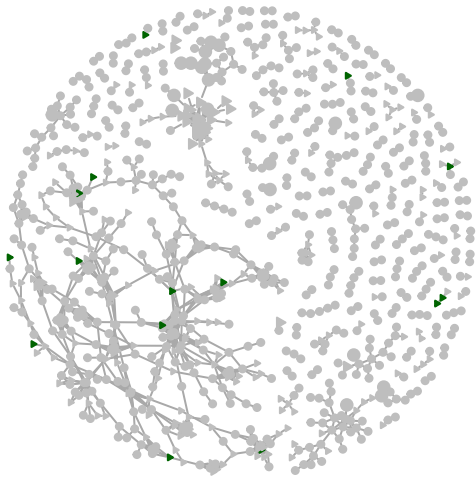


AMF fOTUs (pink) and endobacteria bOTUs (black) in Petunia network

AMF fOTUs in Petunia network



Mortierellomycota fOTUs in Petunia

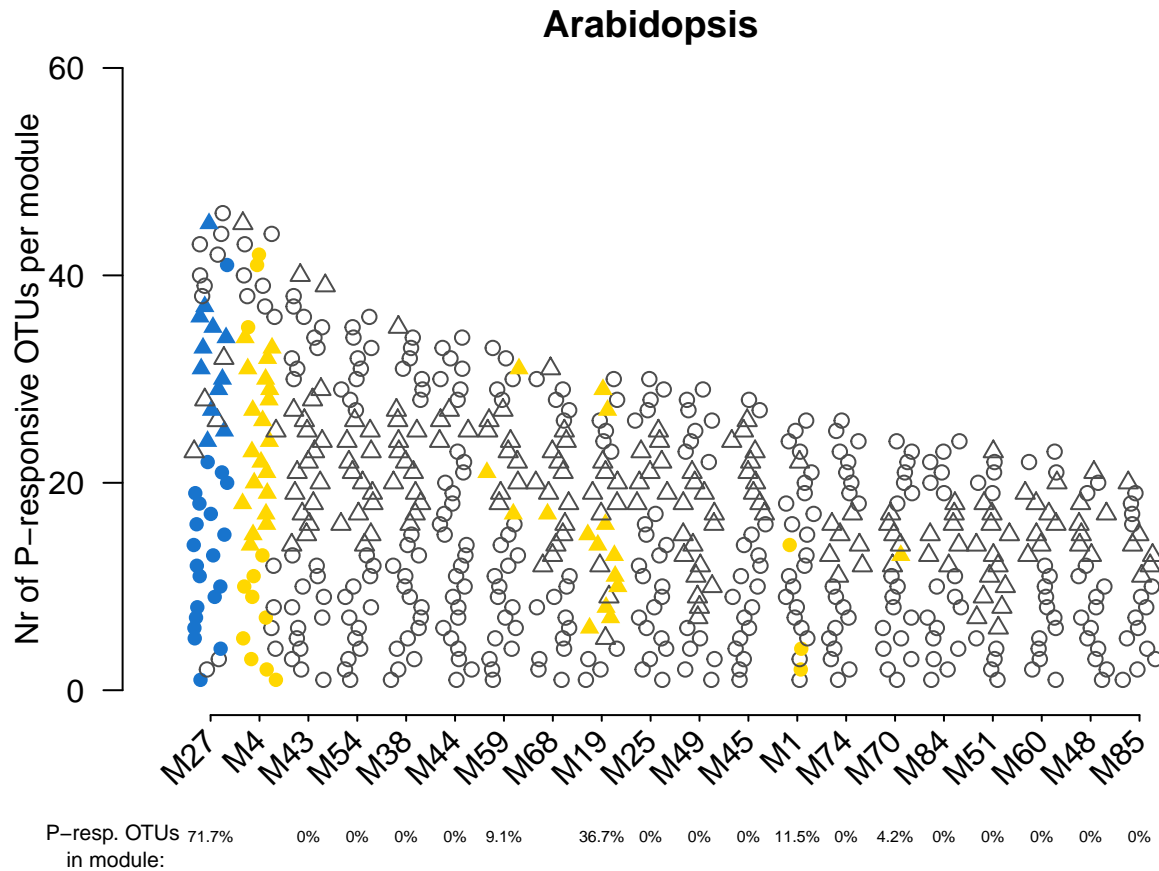


Module detection using hclust (www.kateto.net/networks-r-igraph)

Next, we statistically partition the network into discrete community modules using the *edge betweenness by Newman-Girvan*. This approach identifies the modularity of these network, in which we then want to identify the modules containing the p-dependent microbes.

Arabidopsis

```
## [1] "133 modules assigned for Arabidopsis from 995 OTUs"  
## [1] "The modularity score for Arabidopsis is 0.916376965277351"  
## [1] "Module 27 contains P responsive b/fOTUs in Arabidopsis"  
## [2] "Module 4 contains P responsive b/fOTUs in Arabidopsis"  
## [3] "Module 19 contains P responsive b/fOTUs in Arabidopsis"
```



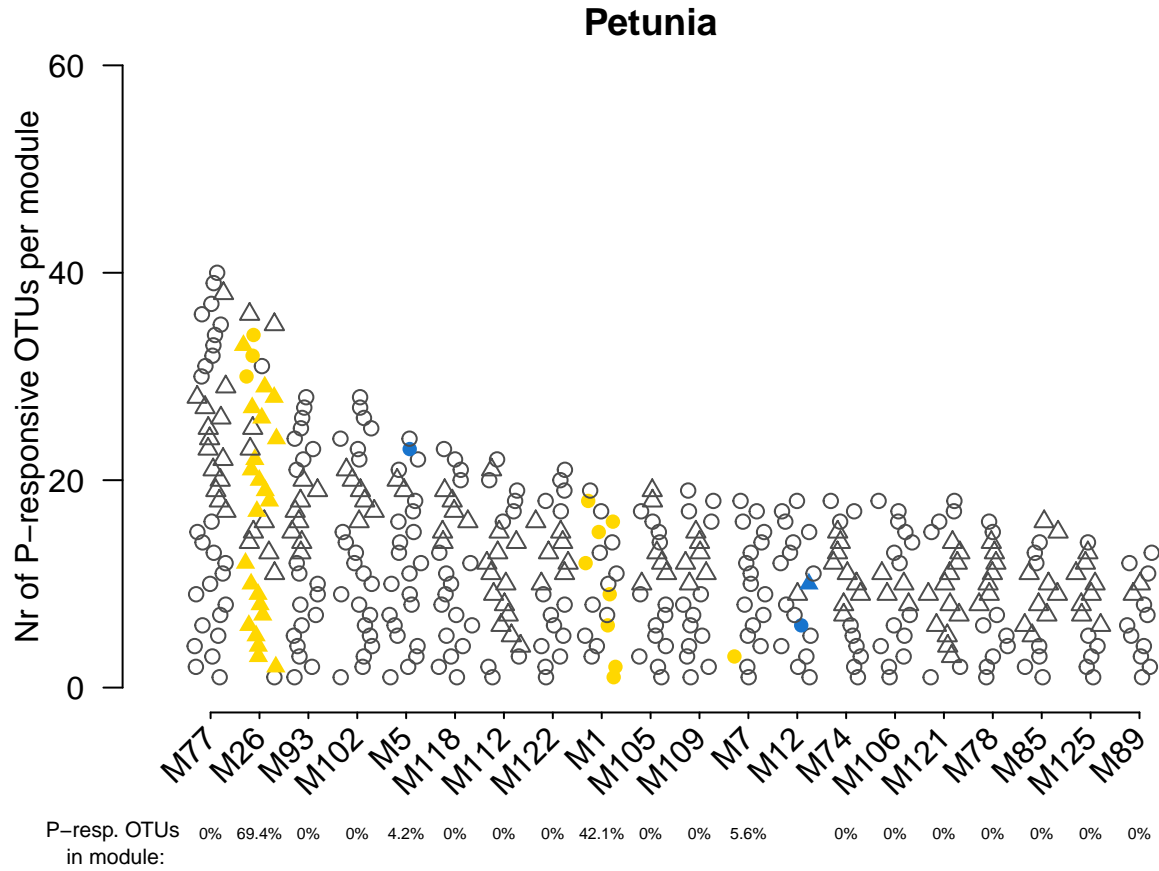
Petunia

[1] "189 modules assigned for Arabidopsis from 995 OTUs"

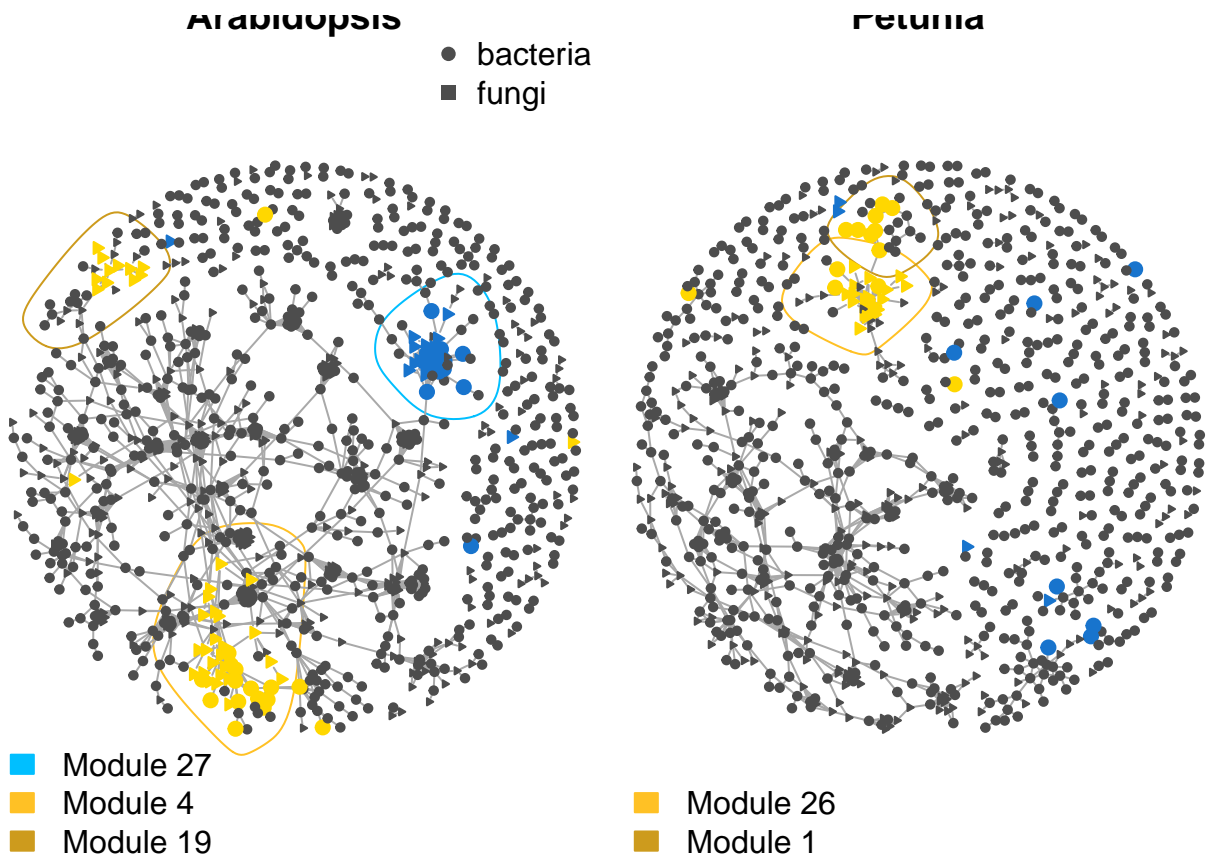
[1] "The modularity score for Arabidopsis is 0.916232088914539"

[1] "Module 26 contains P responsive b/fOTUs in Petunia"

[2] "Module 1 contains P responsive b/fOTUs in Petunia"



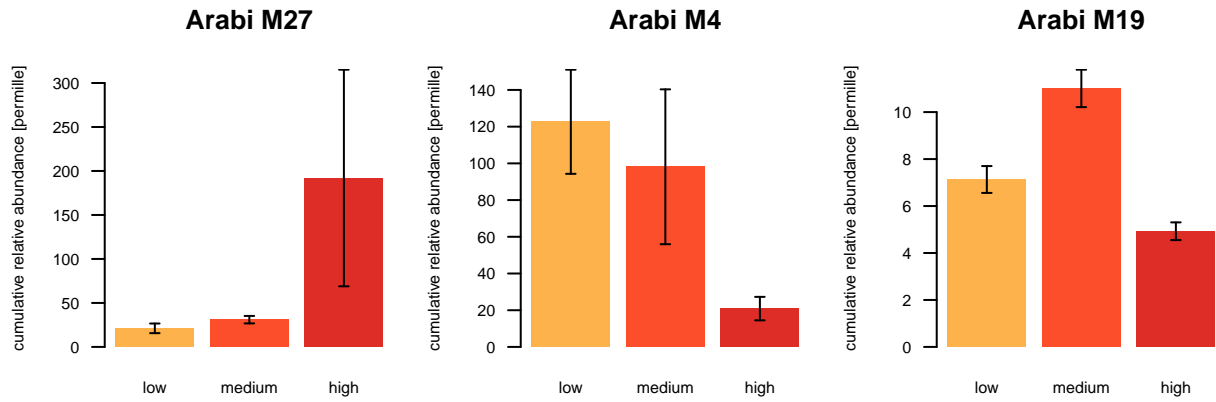
Network figure with highlighted modules



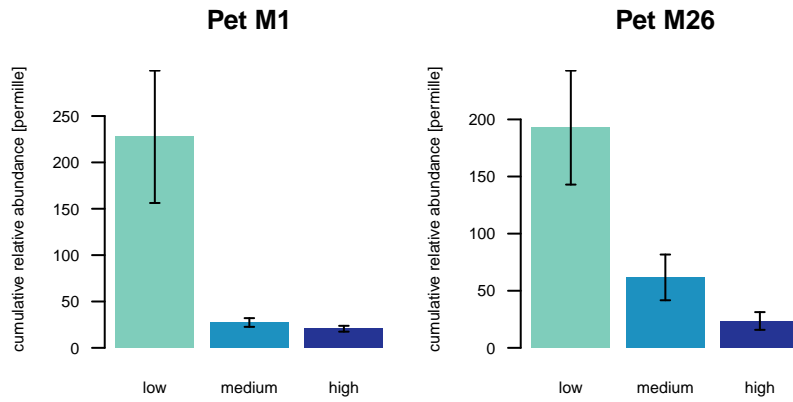
Printing Figure 6 (variant with top20 modules highlighted)

Visualizing abundance behaviour of modules

Arabidopsis

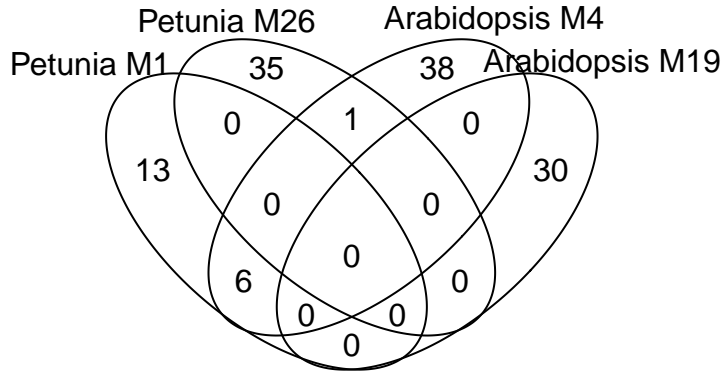


Petunia



Shared OTU occurrence in Arabidopsis and Petunia modules?

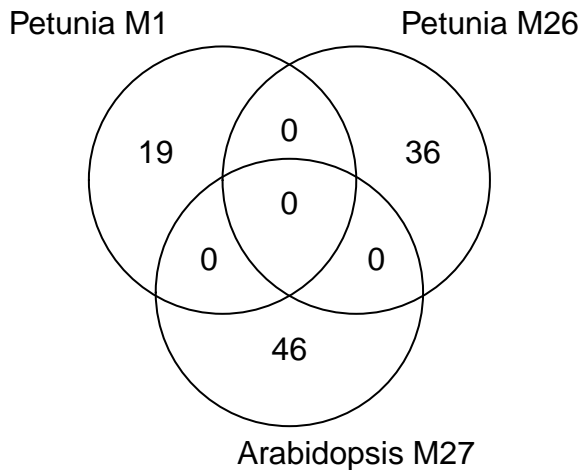
1) between low-P responsive modules of Arabidopsis and Petunia?



##	Phylum	Order	Genus
## bOTU_3346	Proteobacteria	Sphingomonadales	unassigned
##	Phylum	Order	Genus
## bOTU_1787	Proteobacteria	unassigned	unassigned
## bOTU_1874	Proteobacteria	unassigned	unassigned
## bOTU_1802	Proteobacteria	Burkholderiales	Massilia
## bOTU_234	Proteobacteria	Bdellovibrionales	Bdellovibrio
## bOTU_2	Proteobacteria	Rhodocyclales	Dechloromonas
## bOTU_12	Proteobacteria	Rhodocyclales	Methyloversatilis

2) between low-P Petunia and high-P Arabidopsis modules?

low P Petunia to high P Arabi OTUs



Keystone OTUs

We identified keystone OTUs separately for the Arabidopsis and Petunia networks and defined them as those OTUs within the top 1% values of the degree of co-occurrence.

Arabidopsis

	Phylum	Order	Genus
bOTU_584	Chloroflexi	uncultured_bacterium	uncultured_bacterium
bOTU_538	Acidobacteria	unassigned	unassigned
bOTU_1345	Planctomycetes	Planctomycetales	uncultured
bOTU_1548	Acidobacteria	unassigned	unassigned
bOTU_1655	Planctomycetes	Planctomycetales	uncultured
bOTU_258	Verrucomicrobia	Chthoniobacterales	uncultured_bacterium
bOTU_2251	Verrucomicrobia	uncultured_bacterium	uncultured_bacterium
bOTU_624	Verrucomicrobia	Chthoniobacterales	uncultured_bacterium

In the Arabidopsis network 8 keystone OTUs were found. These were 8 bacteria and 0 fungi. The following keystone OTUs occur in p-responsive module M27: bOTU_584, bOTU_538, bOTU_1345, bOTU_1548, bOTU_1655, bOTU_258, bOTU_2251, bOTU_624.

Petunia

	Phylum	Order	Genus
bOTU_920	Firmicutes	Clostridiales	unassigned
bOTU_3183	Acidobacteria	Acidobacteriales	uncultured
bOTU_1391	Firmicutes	Clostridiales	unassigned
bOTU_899	Actinobacteria	Acidimicrobiales	unassigned

	Phylum	Order	Genus
fOTU_6	Glomeromycota	Glomerales	unassigned
fOTU_109	unassigned	unassigned	unassigned
fOTU_382	Ascomycota	Hypocreales	Trichoderma
fOTU_341	unassigned	unassigned	unassigned
fOTU_111	Glomeromycota	Glomerales	unassigned

In the Arabidopsis network 9 keystone OTUs were found. These were 4 bacteria and 5 fungi. The following keystone OTUs occur in p-responsive modules: fOTU_6, fOTU_109, fOTU_111.

Vizualization of keystone OTUs in network graphss

keystone OTUS in Arabidopsis

keystone OTUS in Petunia

