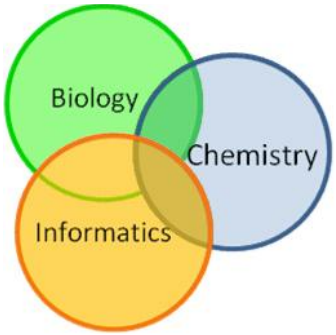


Comparison of pumpkin and tomatillo leaf primary metabolites

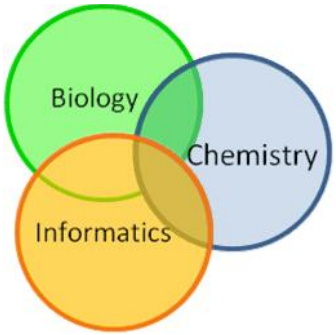
Comparison of Leaf Metabolites

Goal:

Carry out a statistical, HCA, PCA and O-/PLS-DA analyses comparing leaf primary metabolite profiles (Used DATA: Pumpkin and Tomatillo 1.csv)

Cucurbita pepo

Physalis philadelphica



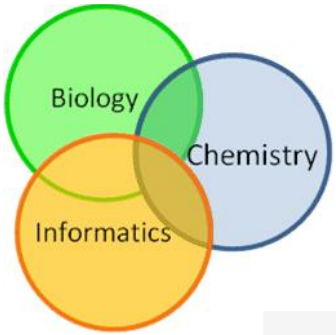
Comparison of pumpkin and tomatillo leaf primary metabolites

Comparison of Leaf Metabolites

Used DATA: Pumpkin and Tomatillo 1.csv

Steps:

1. Identify analysis strategy (hint: use HCA and PCA)
2. Conduct statistical comparison
3. Identify top multivariate discriminants



Data Exploration

Species	Treatment
pumpkin :12	MeOH:CHCl3:H2O (5:2:2) - fresh frozen:12
tomatillo:12	MeOH:CHCl3:H2O (5:2:2) - lyophilized :12

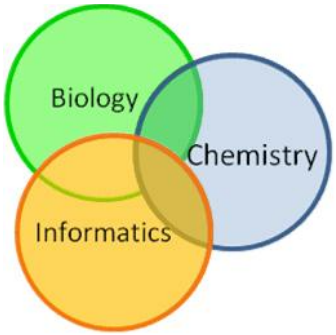
Comparison of Leaf Metabolites

Steps:

1. Identify the effect of treatment on species differences
    - Use HCA
    - PCA

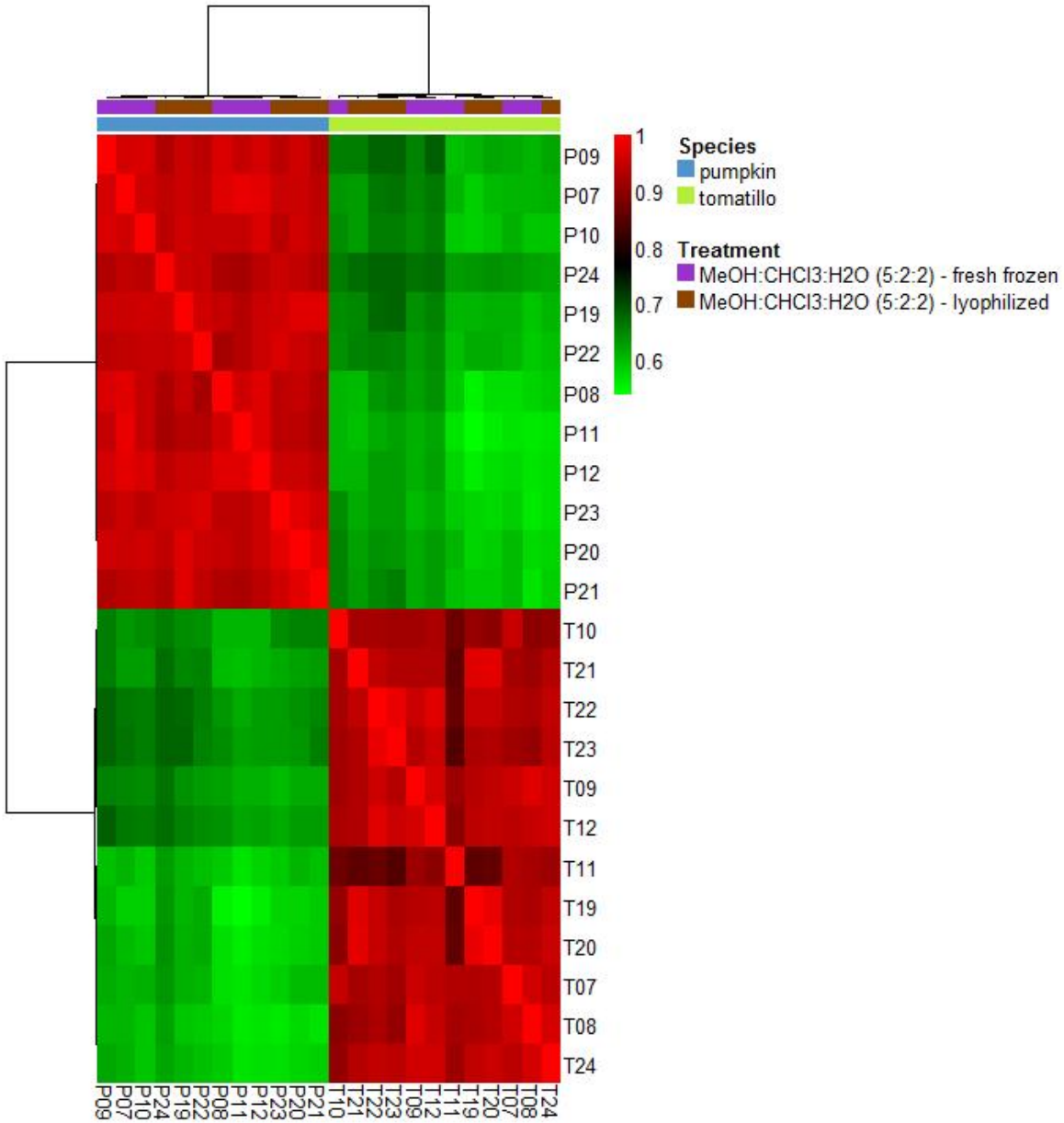
Exercise:

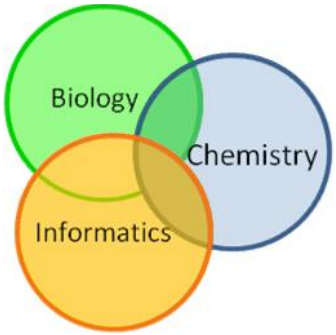
Can different treatments be analyzed together to identify species differences?



HCA clustering of samples

Comparison of Leaf Metabolites



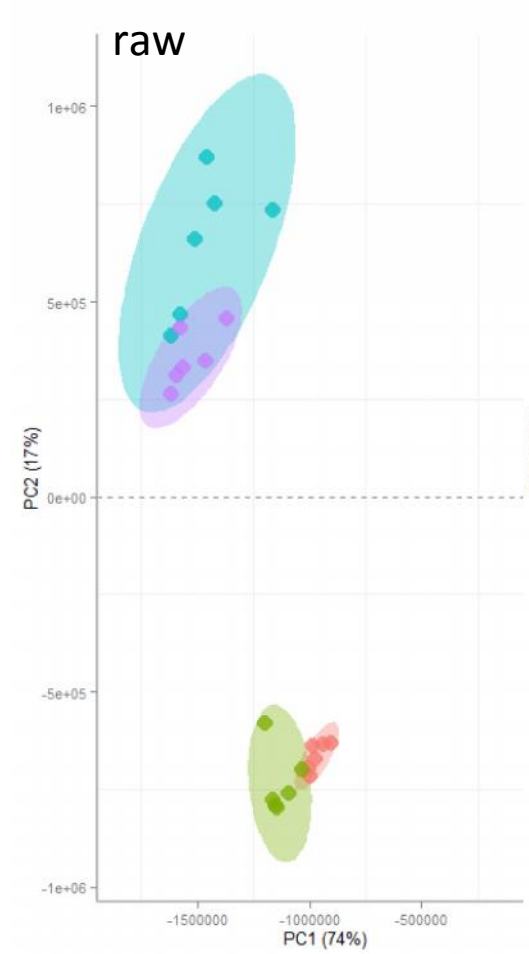


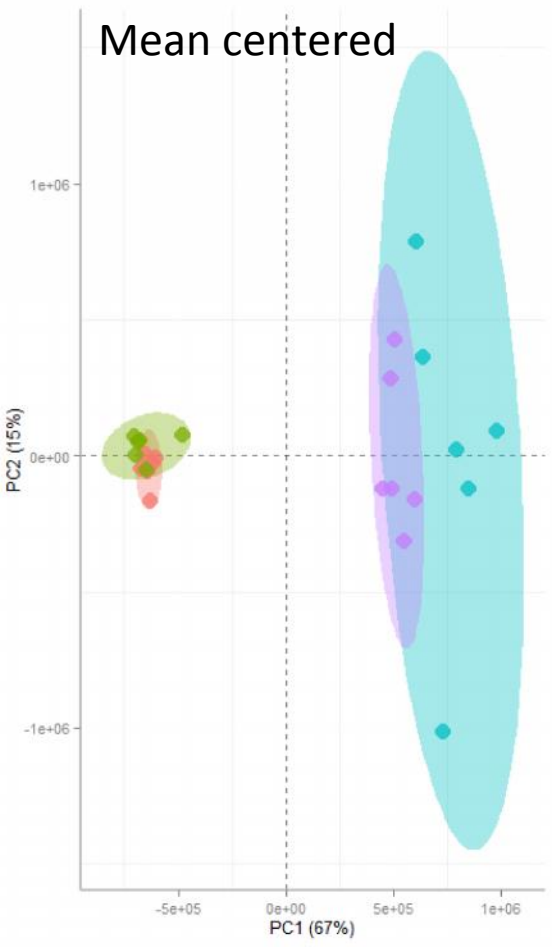
PCA: comparison of pretreatments

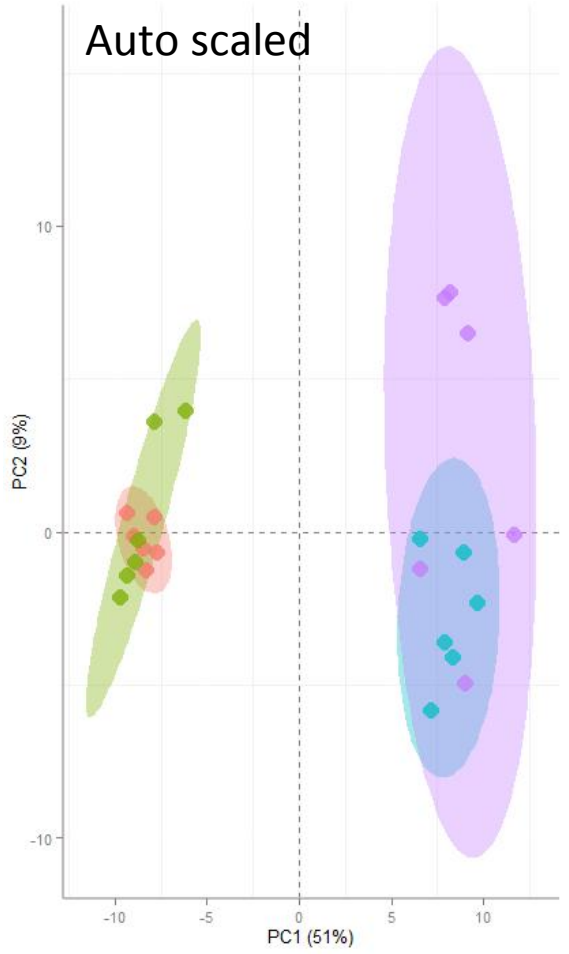
Species|Treatment

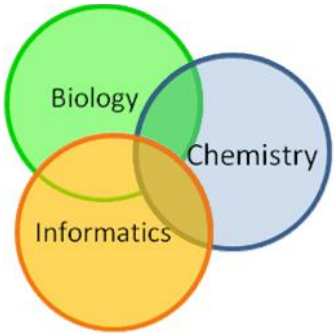
- pumpkin|MeOH:CHCl3:H2O (5:2:2) - fresh frozen
- pumpkin|MeOH:CHCl3:H2O (5:2:2) - lyophilized
- tomatillo|MeOH:CHCl3:H2O (5:2:2) - fresh frozen
- tomatillo|MeOH:CHCl3:H2O (5:2:2) - lyophilized

Comparison of Leaf Metabolites

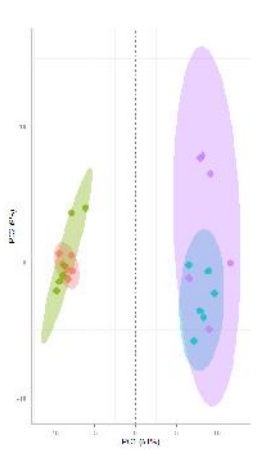


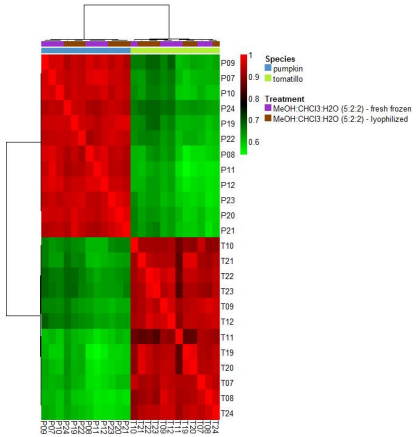






Identify Analysis Strategy





Comparison of Leaf Metabolites

Analysis Options:

If the treatment is a minor effect compared to species differences:

- two-sample t-Test for Species

If the treatment has a considerable effect compared to species differences:

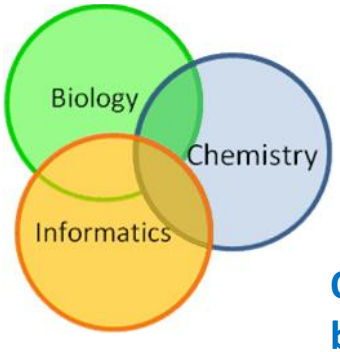
- two-way ANOVA for Species + treatment + interaction (species/treatment)

If the treatment has a similar effect size to species differences:

- Eliminate one treatment type from analysis and use t-Test

Conclusions:

- Both PCA and HCA analyses suggests the treatment effect is minor
- Using 12 compared to 6 samples per group will increase study power

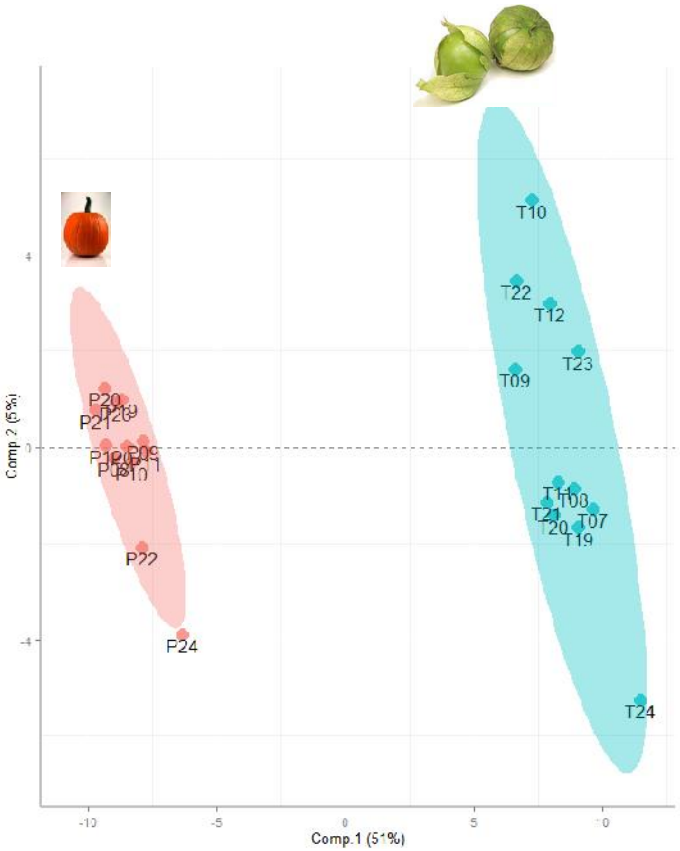


Comparison PLS-DA to O-PLS-DA

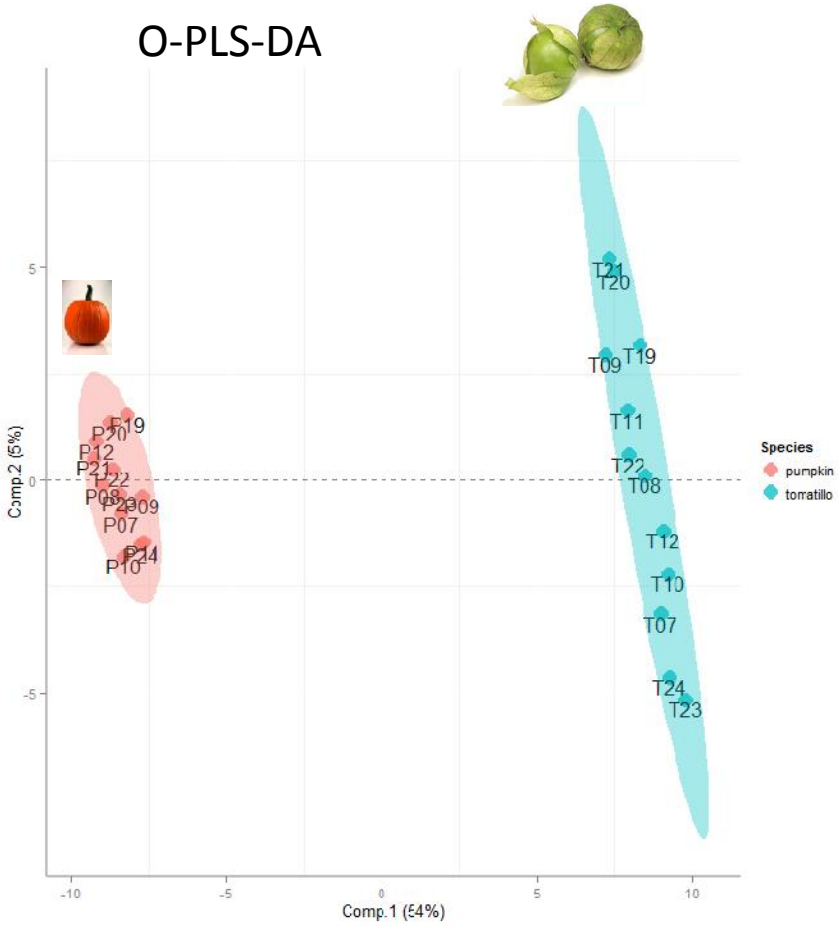
O-PLS-DA is only useful over PLS-DA when the axis of separation between two groups spans >1 dimension

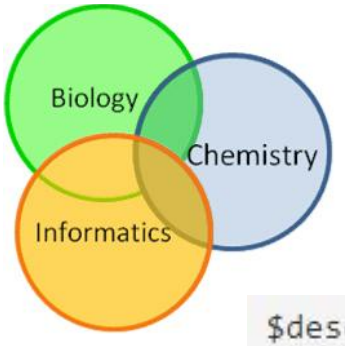
Comparison of Leaf Metabolites

PLS-DA



O-PLS-DA



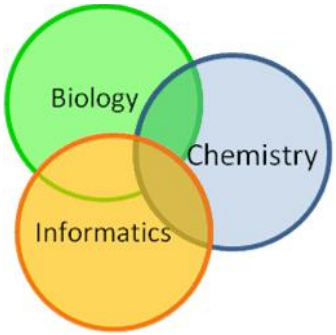


Validation of PLS-DA model for discrimination between pumpkin and tomatillo leaf metabolites

Comparison of Leaf Metabolites

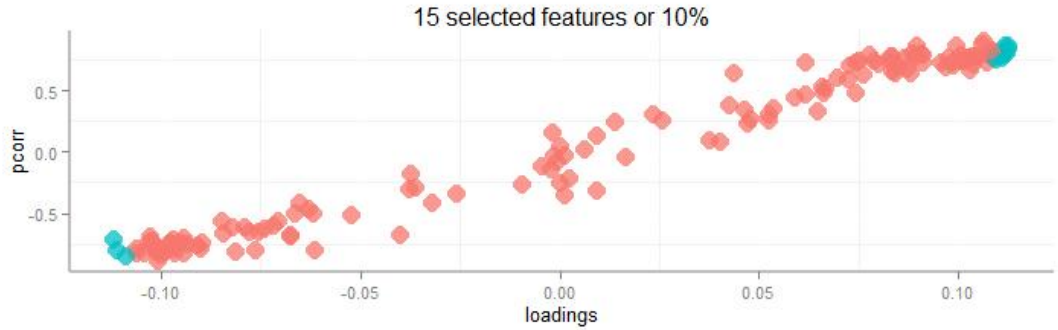
```
$description

                                                          selection
Dependent Variables                                     (1)Species
Latent variables (LVs)                                           2
Orthogonal latent variables (OLVs)                               0
model cross-validation                                         LOO
method                                                  oscorespls
Internal train/test index          50 repetitions generated by random

$statistics
          Xvar      Q2  RMSEP
intercept  0.0 -0.0888 0.5217
LV 1      51.1  0.9754 0.0780
LV 2      56.5  0.9736 0.0803

$`Validated Model Performance (Y1)`
                       Xvar                Q2             RMSEP
model          59.12 ± 2.65 0.9685 ± 0.00664 0.07844 ± 0.0179
permuted model 32.44 ± 7.44   0.1281 ± 0.312    0.625 ± 0.129
p-value           9.863e-33        2.551e-24        8.184e-34
```

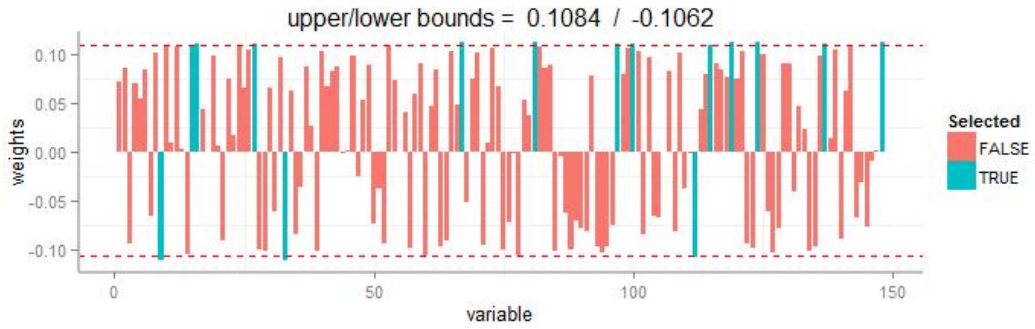
Outstanding model performance, highly unlikely by random chance



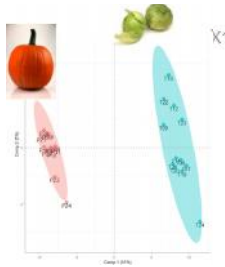
Identification of top multivariate discriminants between pumpkin and tomatillo leaf primary metabolites

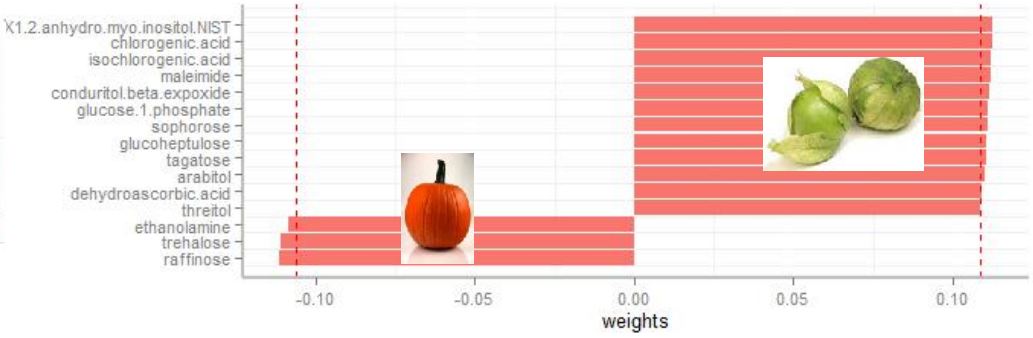
Comparison of Leaf Metabolites





p.value	0.05
FDR	TRUE
cut.type	quantile
top	0.9
separate	FALSE
type	spearman





Could also select from increasing and decreasing metabolites separately