**Pathway analysis in MarVis**

**MarVis Filter**

* File
* Import raw data
* Import dialog box
  + Delimiter = ,
  + Start row = 1
  + Start column = 1
  + Id column = 1
  + Id label = ID
  + Generate IDs = unticked.
  + X column = 3
  + X label = rt (retention time)
  + Y column = 2
  + Y label = m/z (mass/charge ratio)
  + Condition identifiers = name of conditions being compared e.g. 1MB, 1PMB,
  + 2MB, 2PMB
  + Additional columns = blank
  + Additional labels = blank
* Log-scale transformation = none.
* Sample based normalisation = none.
* Ranking method = none, done in previous step for ANOVA.
* Multiple testing adjustment = none as ranking method not used.
* Adduct and isotope correction.
  + Maximal number of C13 isotopes per marker candidate = 2.
  + Minimal cosine similarity = 0.75.
  + Choose adduct neg or pos file depending on the file being uploaded.
  + Rt tolerance = 0.01.
  + m/z tolerance = 10.
  + Add data set.
  + Marvis suite ➔ go to Marvis cluster.

**Marvis cluster**

* Sample aggregation = mean.
* Marker scaling = 2-norm.
* Clustering dialogue
  + Number of prototypes = left as default (e.g. in the case of ESI Neg BABA there were 117 markers (potential metabolites) and the default value here was 5.
* Sorting
  + Sort marker candidates within clusters according to 1D-SOM projection? = yes.
* Selection ➔ select all.
* Marvis suite ➔ go to Marvis pathway.

**Marvis pathway**

* Use existing 1D-SOM for marker profile ordering? = Yes.
* Local database selection.
  + Databases = KEGG.
  + Selection = KEGG Populus trichocarpa (black cottonwood) pathway.
* Load both internal library database.
  + Internal\_Library\_190914 and Internal\_Lignane\_110822 (provided by Dr Pastor’s group).
* Match entry IDs to candidates = cancel x2.
* Mass matching
  + m/z (corrected) correction = 0.
  + m/z (corrected) tolerance = 0.05.
* Ranking options = keep default.
  + Marker score = count.
  + Market hit normalisation = local.
  + Entry hit normalisation = local.
  + Overall normalisation = max.
  + Normalized by set size = unchecked.
* Delimeter = ,
* Set enrichment analysis.
  + Select file.
  + Type of analysis = entry based enrichment analysis.
  + Type of test = Hypergeometric
  + Calculation of significance = select direct p-value calculation, FDR control (Benjamini-Hochberg).
* List of significant metabolites for further analysis.