

For the WCGNA analysis, I used the Omics Playground v3 from BigOmics Analytics.

Selections:

- RNA-seq
- Canis familiaris

Inputs:

- Gene_count_matrix.csv which was cleaned and reformatted.
- PHENOTYPES.txt that I converted to a csv file for reading into the program.

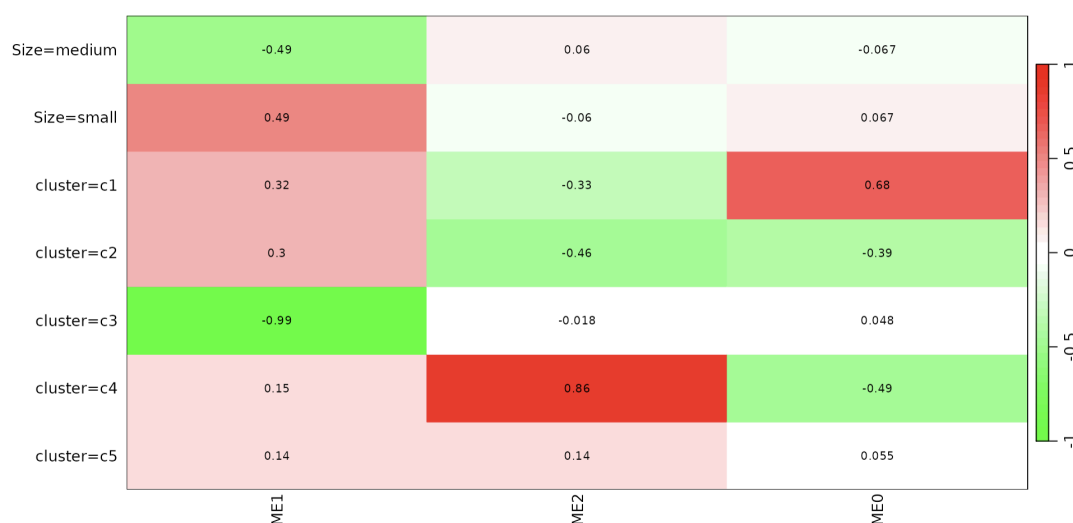
Definitions:

- Weighted Gene Co-expression Network Analysis (WCGNA): a method used to find groups of genes that behave similarly and may work together in biological processes
- Modules: clusters of genes that turn on or off together. These are found by looking at the gene expression similarity tree and then using dynamic branch cutting to figure out where to split the tree into meaningful modules.
- Module eigengene: a summary of the whole module, the average behavior of all genes within that module.

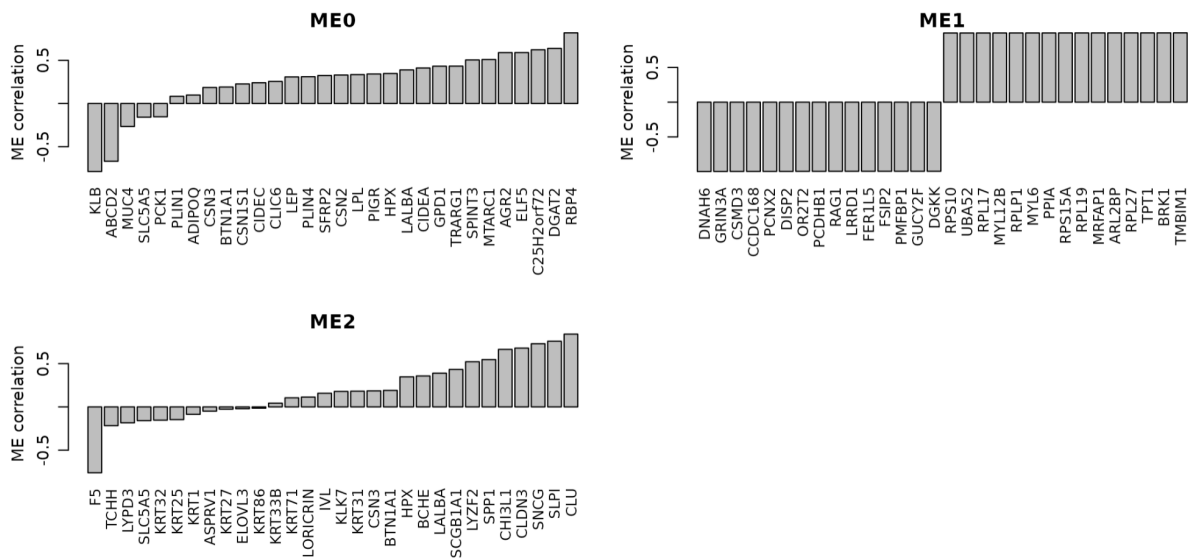
Results:

Module-Trait Relationships: analysis that identifies modules that are significantly associated with the measured traits (small vs. medium dogs), quantifying the association as the correlation of the eigengenes with external traits.

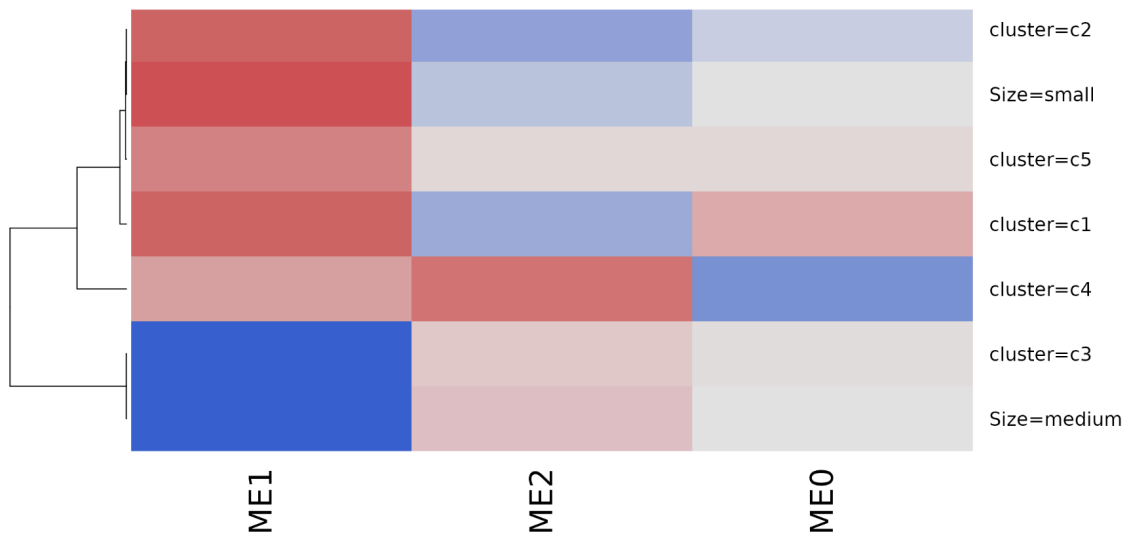
This heatmap shows how gene modules ME1, ME2, and ME0 relate to size characteristics, such as small and medium. Red coloring indicates a positive correlation with higher numbers indicating stronger linkage. Green coloring indicates a negative correlation with lower numbers showing a stronger negative link. White and pale colors indicate weak or no statistically significant linkage. Cluster 3 (c3) shows a correlation of -0.99 with ME1, this showcases a very strong negative correlation, indicating that the genes in ME1 are highly inactive in samples from cluster 3. Cluster 4 (c4) has a correlation of 0.86 with ME2, this shows that there is a very strong positive correlation, meaning genes from cluster 4 are highly active in ME2.



Module membership (eigengene correlation)



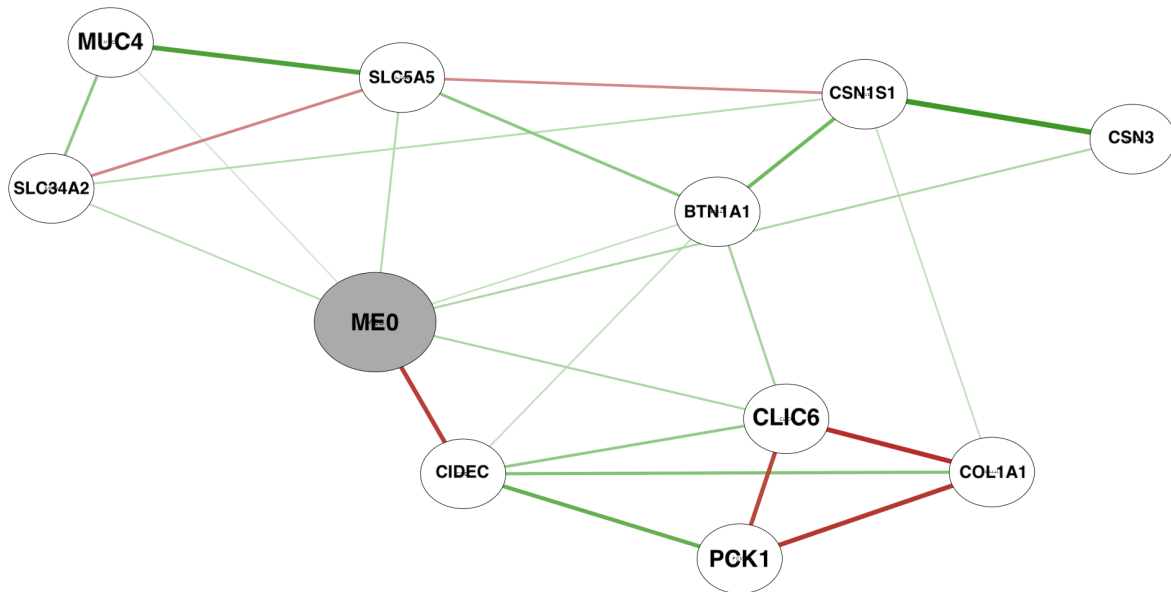
Membership-trait heatmap



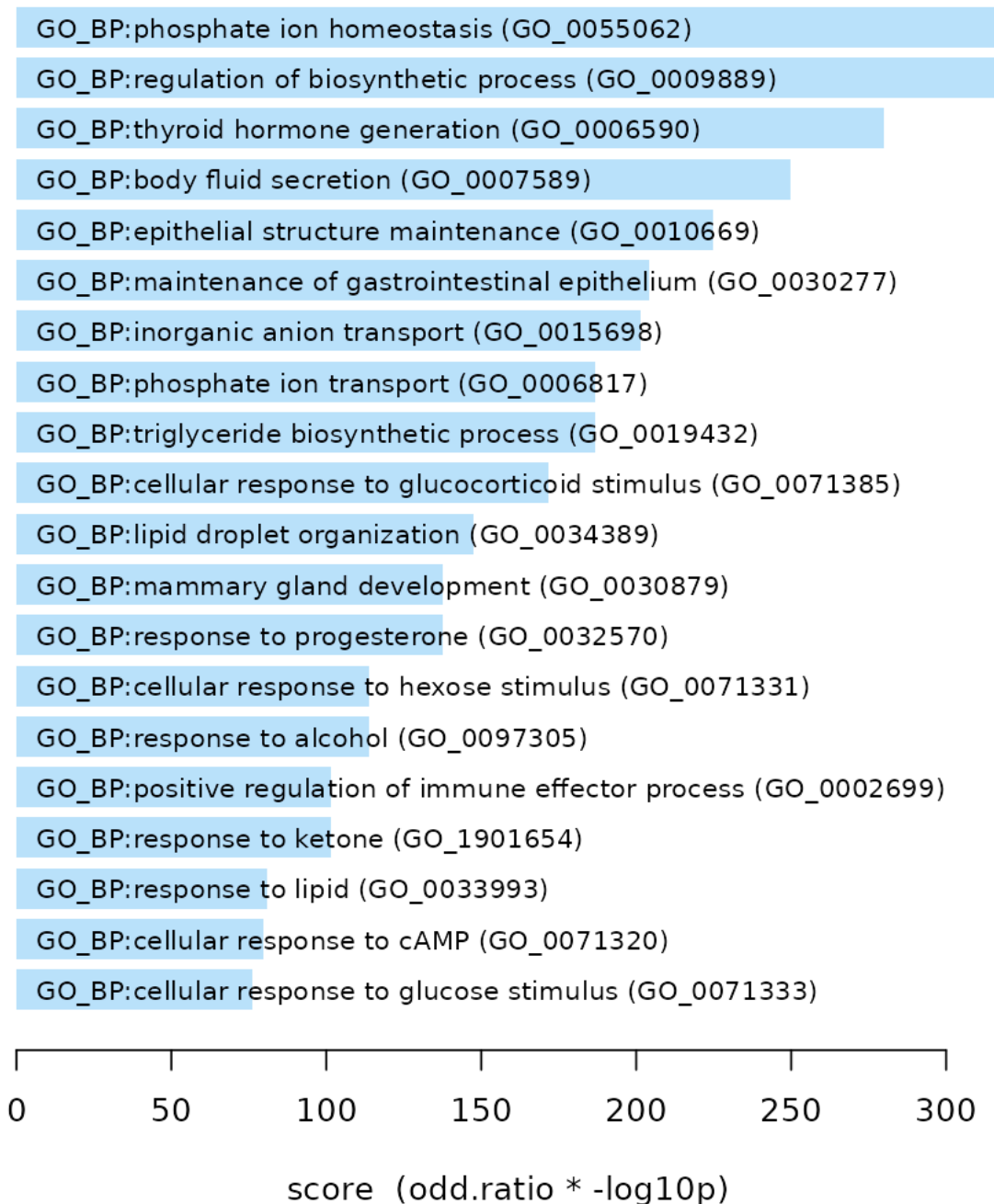
Membership-trait heatmap. We quantify associations of individual genes with our trait of interest (weight) by defining Gene Significance GS as (the absolute value of) the correlation between the gene and the trait. For each module, we also define a quantitative measure of module membership MM as the correlation of the module eigengene and the gene expression profile. Using the GS and MM measures, we can identify genes that have a high significance for weight as well as high module membership in interesting modules.

ME 0:

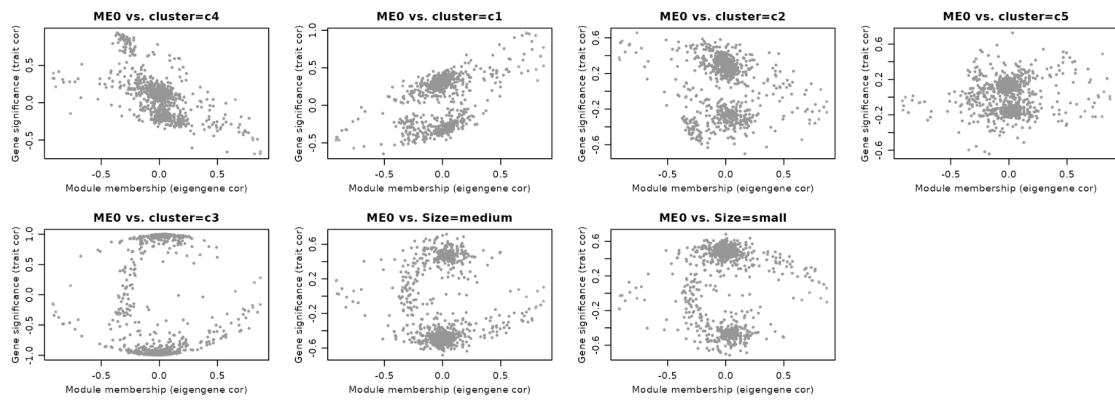
Partial correlation graph centered on module eigen-gene with top most correlated features. Green edges correspond to positive (partial) correlation, red edges to negative (partial) correlation. Width of the edges is proportional to the correlation strength of the gene pair. The regularized partial correlation matrix is computed using the 'graphical lasso' (Glasso) with BIC model selection.



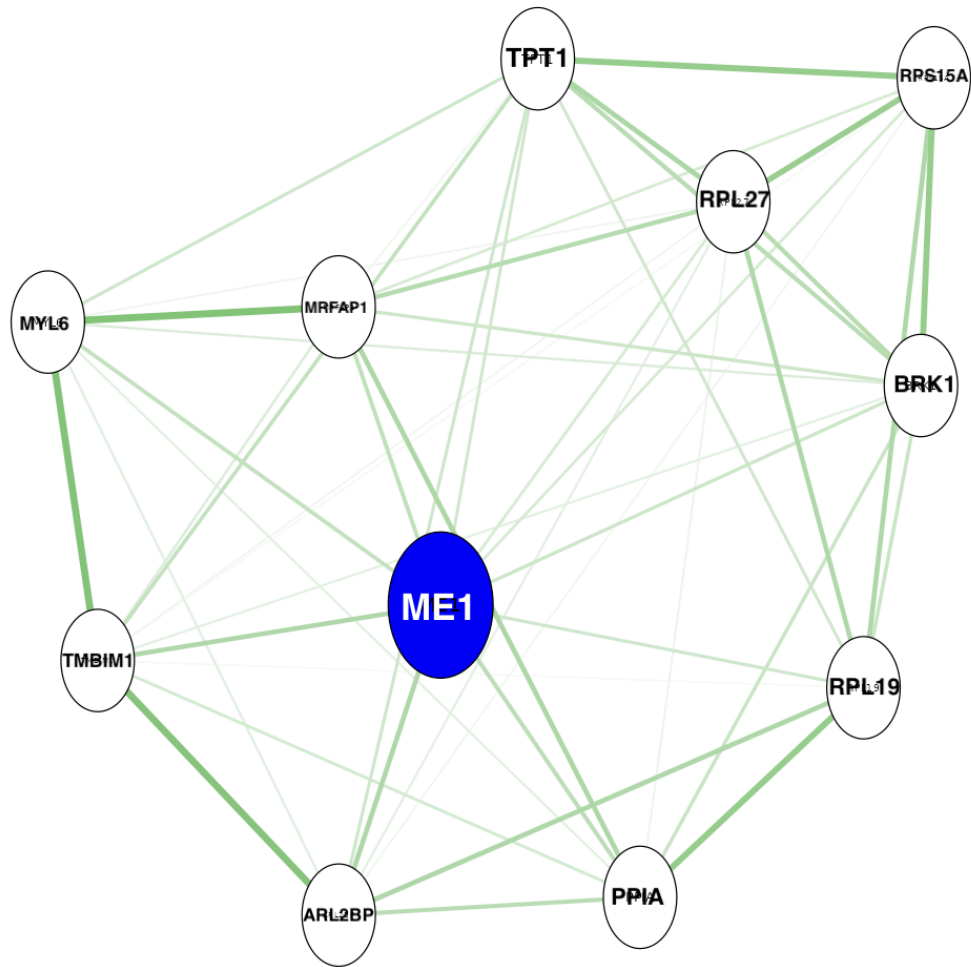
Module Enrichment Plot: Functional enrichment of the selected module.



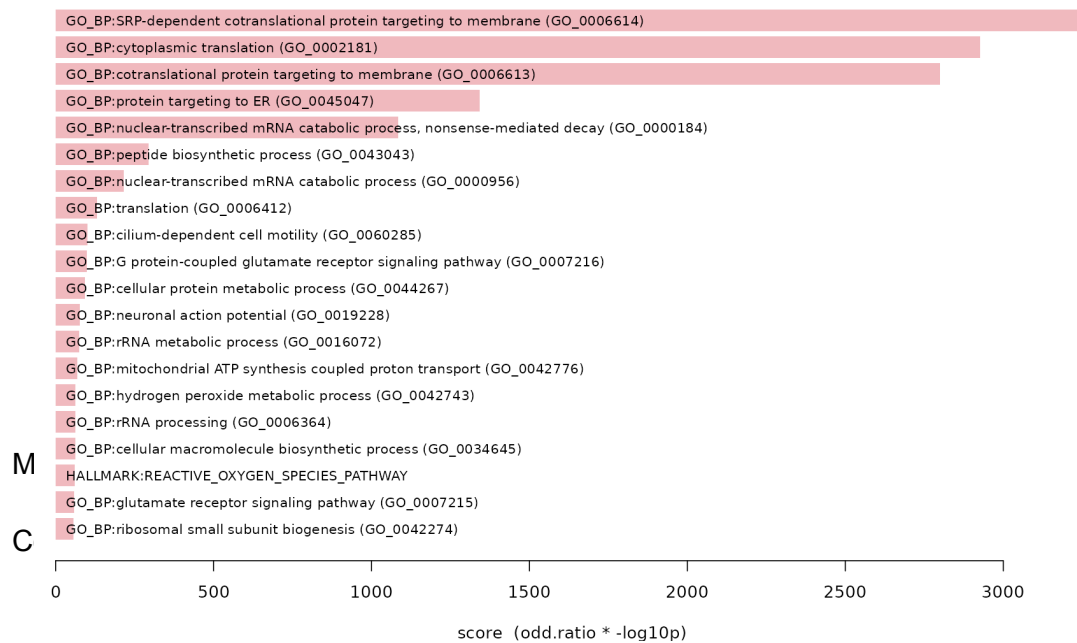
Membership vs. trait correlation



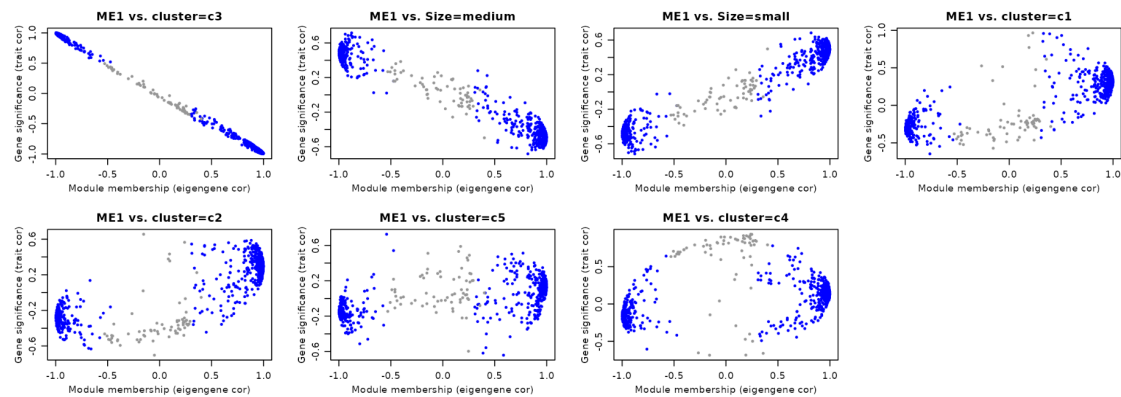
ME1:
Correlation Network



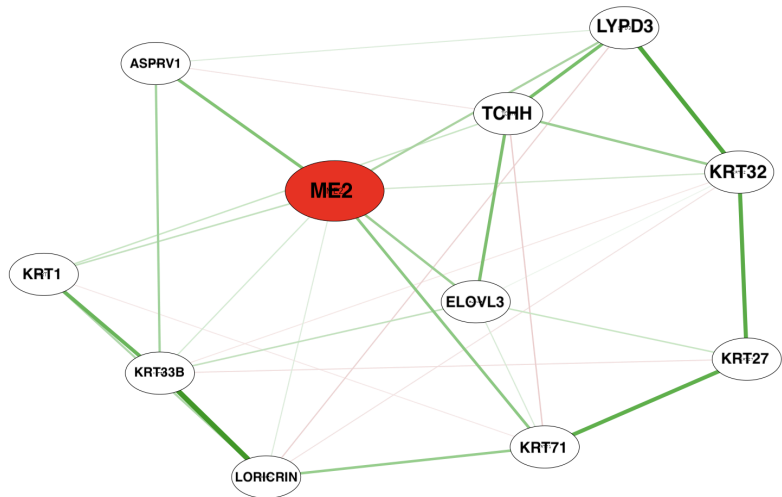
Module Enrichment Plot:



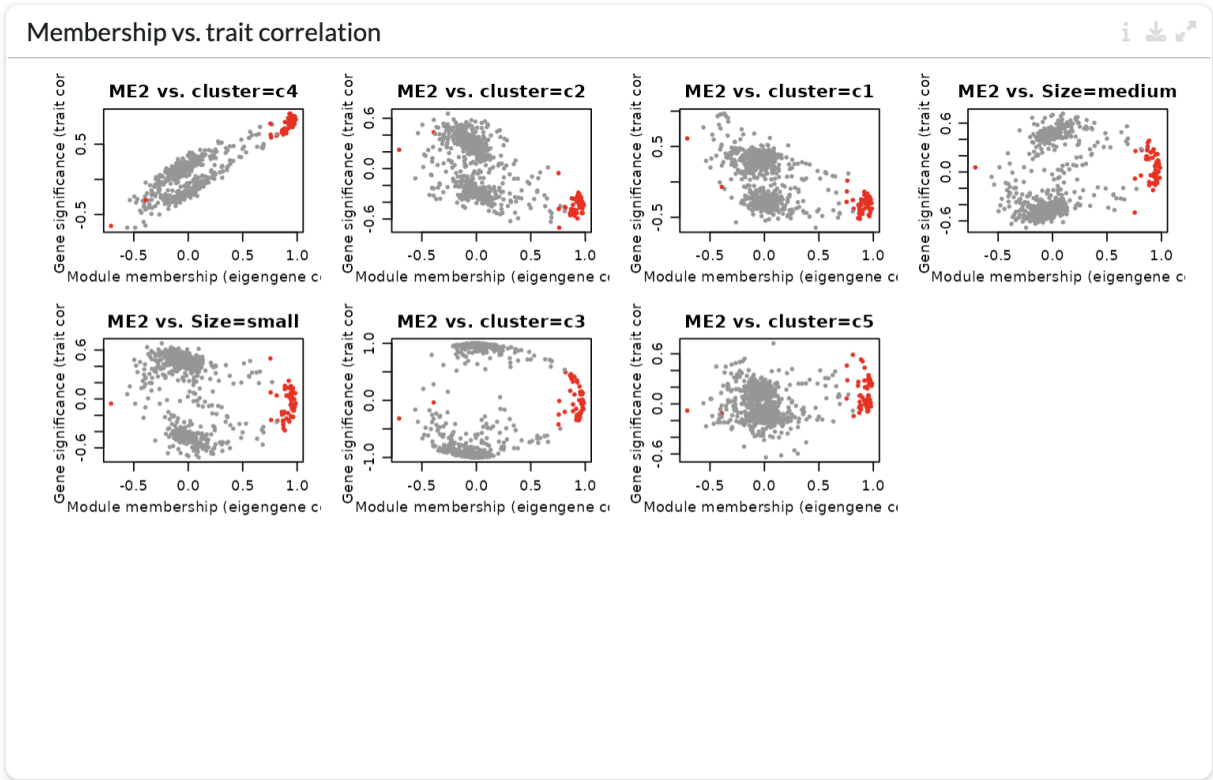
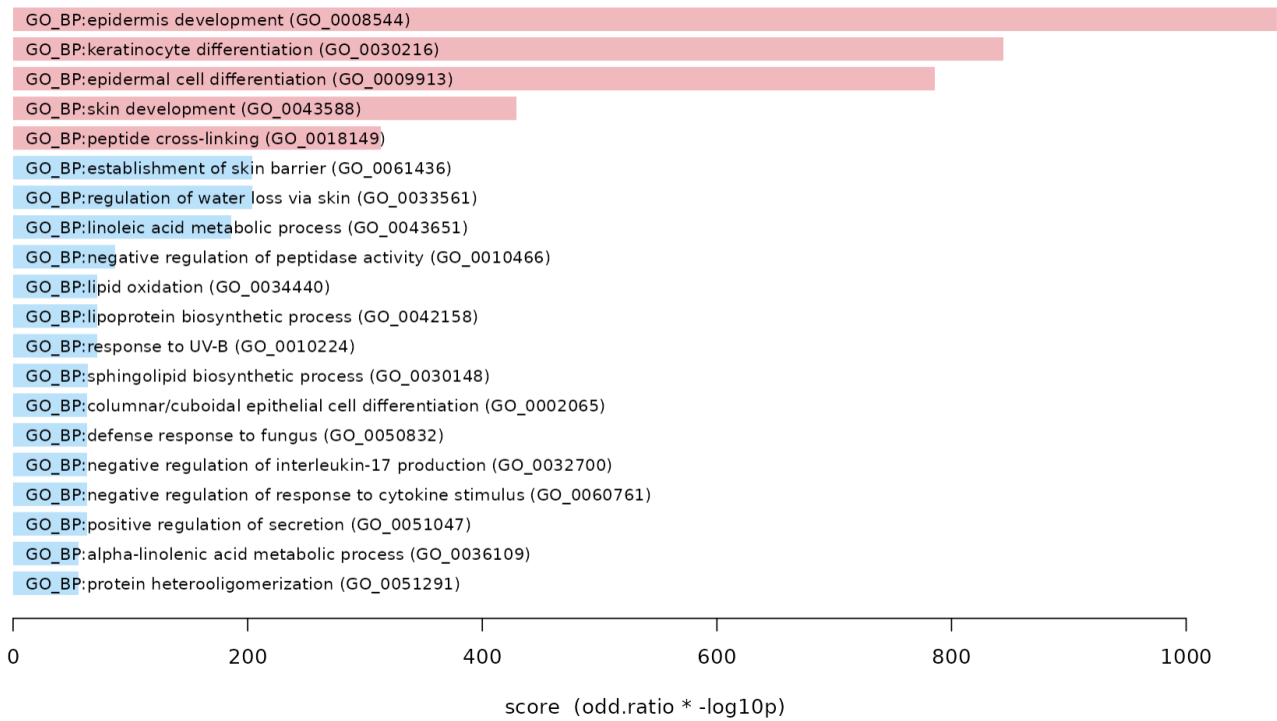
Membership vs. trait correlation



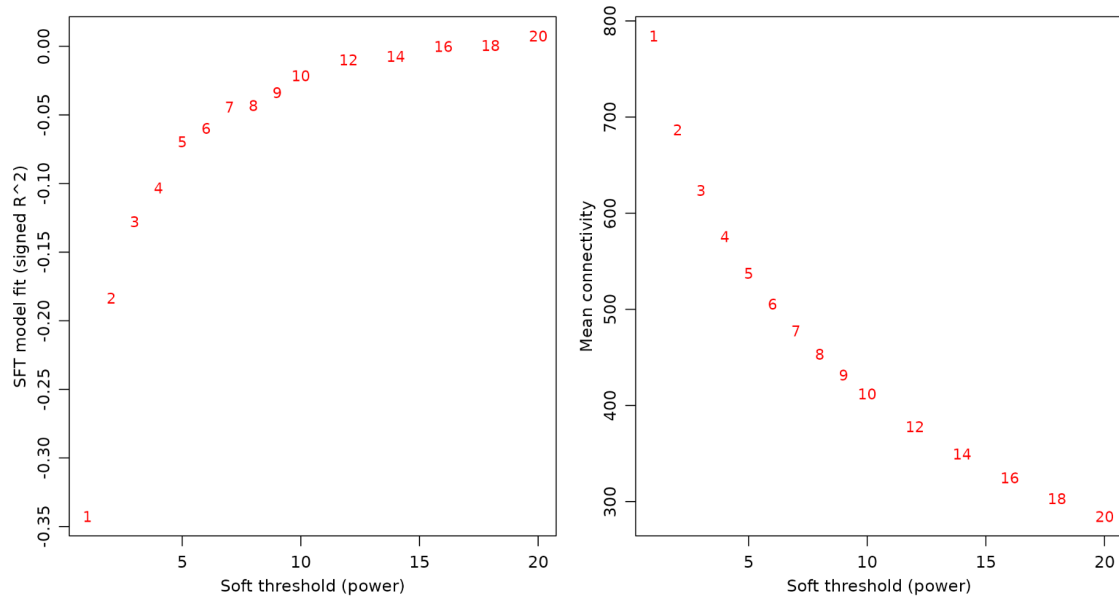
ME2:



Module Enrichment Plot:

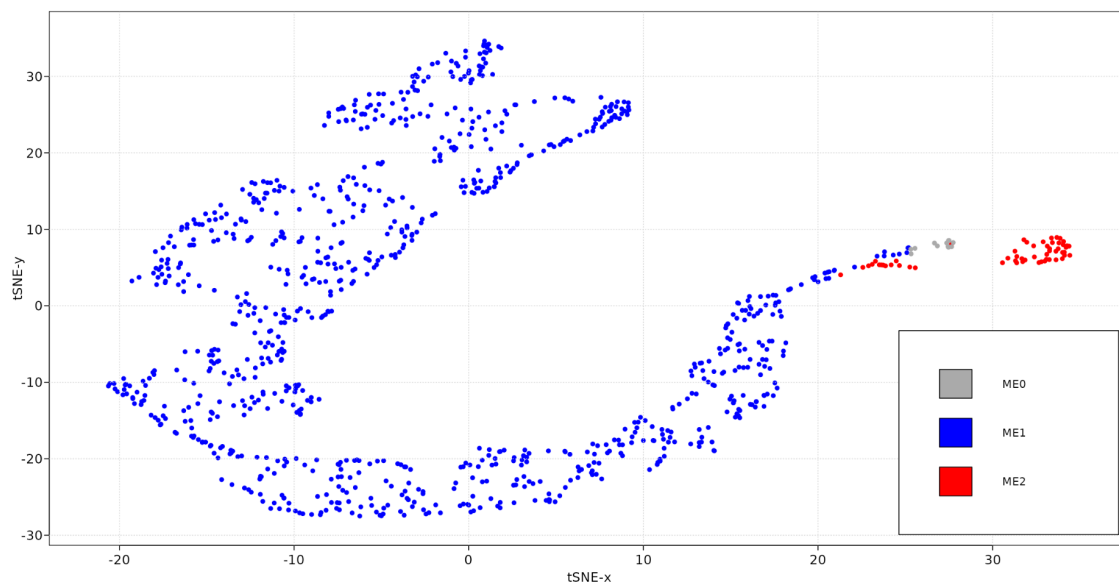


Scale independence and mean connectivity



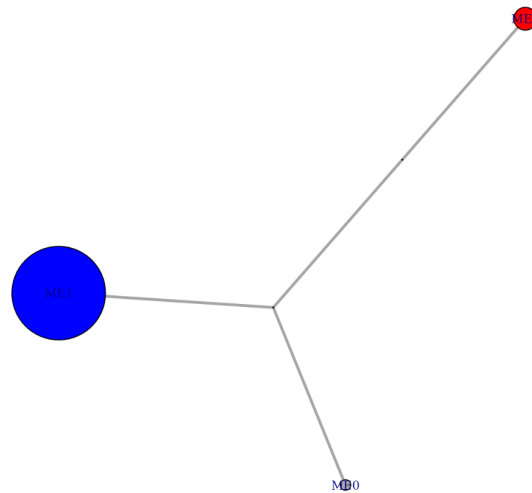
Scale independence and mean connectivity. Scale independence and mean connectivity plots to determine the soft threshold.

Gene clustering



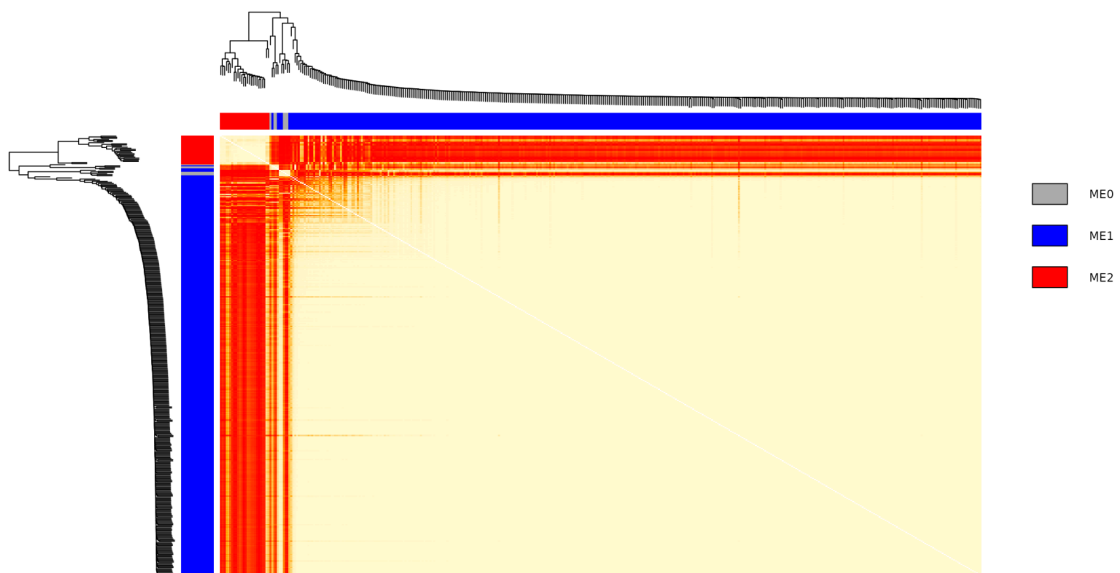
Gene clustering. Dimensionality reduction maps colored by WGCNA module.

Module graph



Module graph. Graph network of WGCNA modules.

TOM heatmap



TOM heatmap. Topological overlap matrix visualized as heatmap.