Employing co-expression networks to identify transcriptome diversification in maize lines

**Author:** Jan Izquierdo i Ramos  
**Scientific director:** Dr. Georg Haberer

Plant Genome and Systems Biology, Environmental Health Center, Helmholtz Munich.

Address: Ingolstädter Landstraße 1, 85764 Neuherberg, Munich

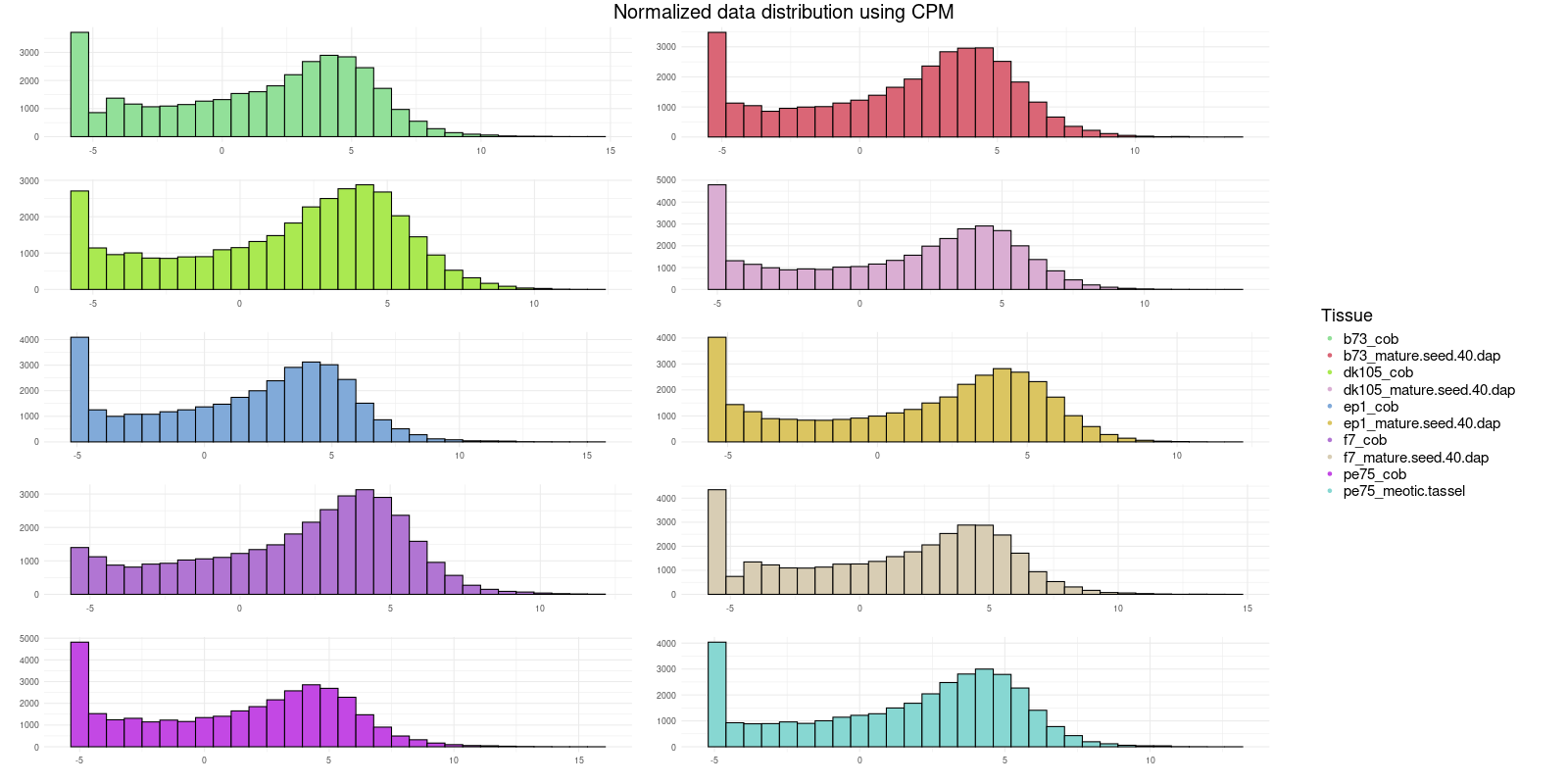
*Bachelor’s Degree in Bioinformatics (UPF-UPC-UB)*

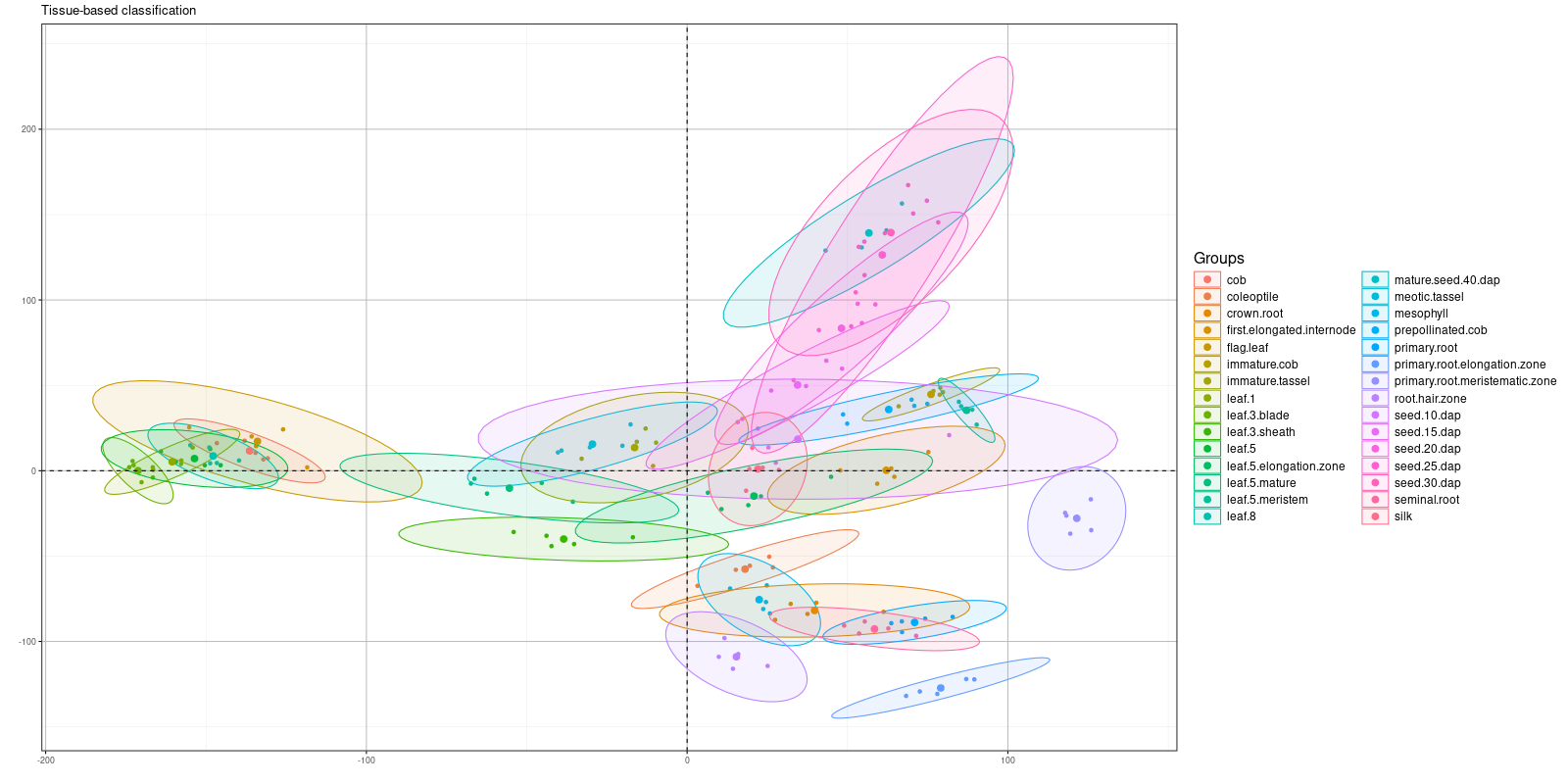
*Final Grade Project*

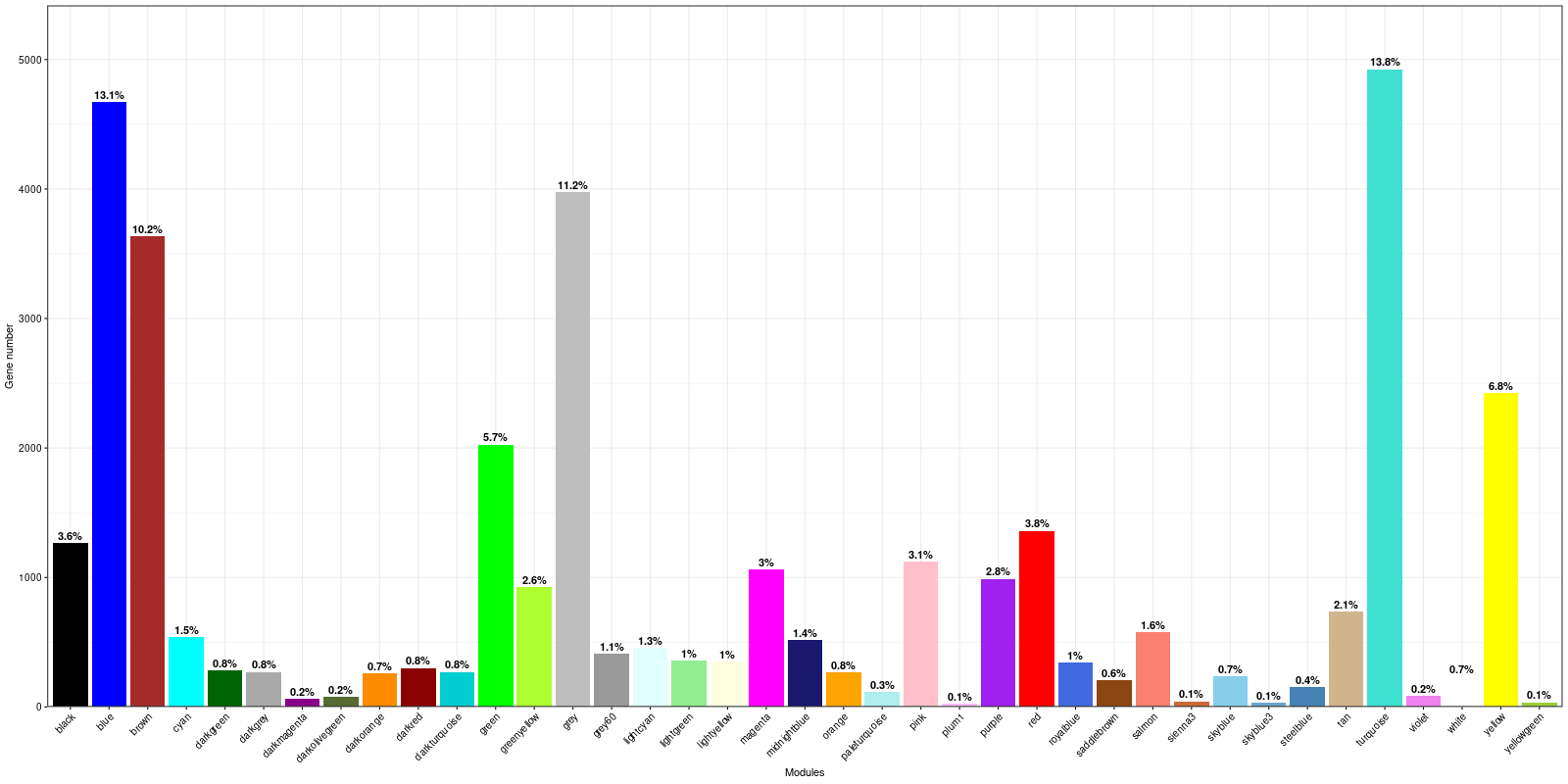
|  |
| --- |
| Employing co-expression networks to identify transcriptome diversification in maize lines  Jan Izquierdo i Ramos1  Scientific director: Dr. Georg Haberer1  1Plant Genome and Systems Biology, Environmental Health Center, Helmholtz Munich, Ingolstädter Landstraße 1, 85764 Neuherberg, Munich  Abstract  **Motivation:** The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog.  **Results:** The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog. The quick brown fox jumps over the lazy dog.  **Supplementary information:** Supplementary material and code available at GitHub: <https://github.com/Janek21/BDBI_TFG_MaizeCoexpression> |

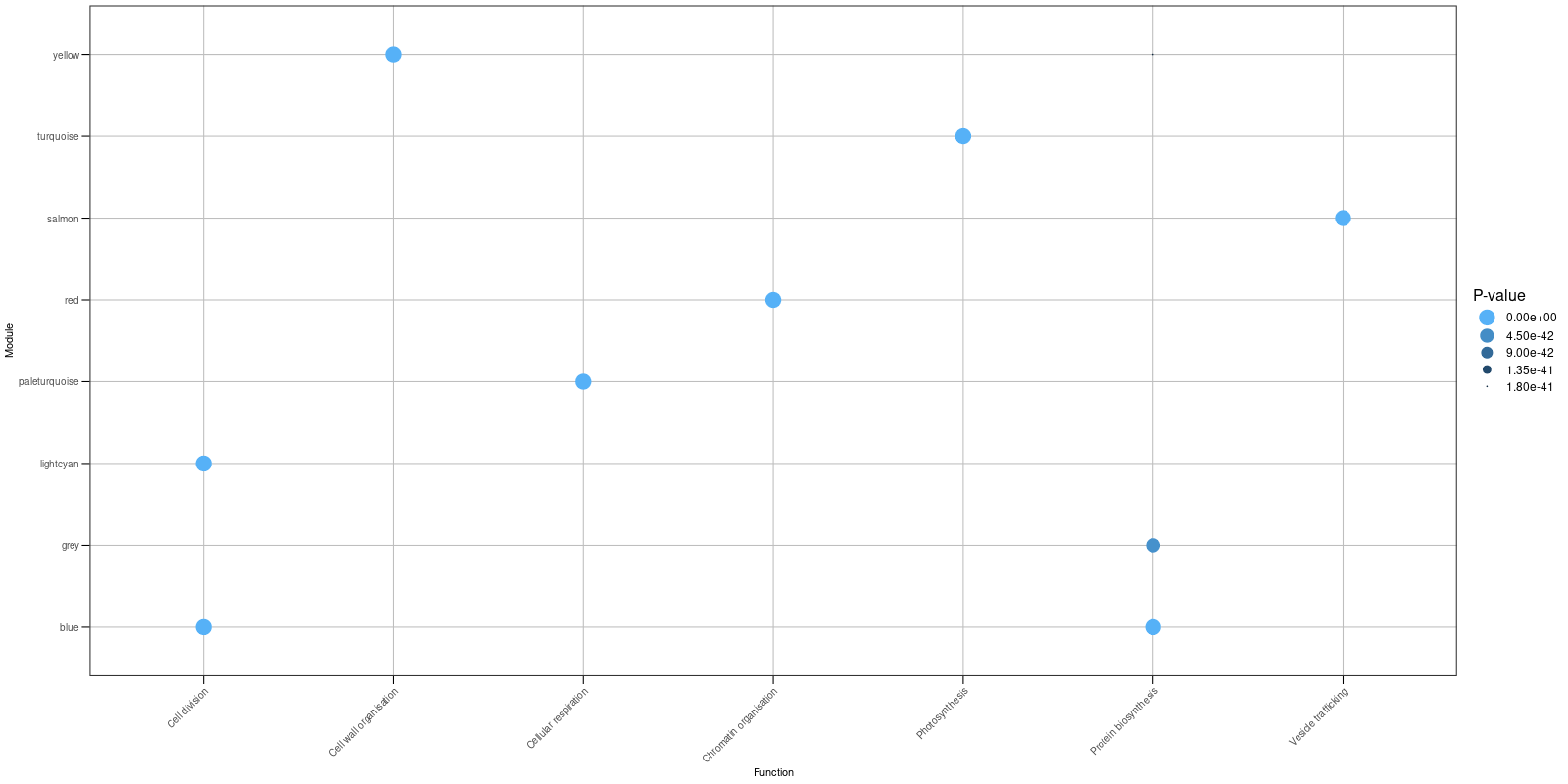
Supplementary material

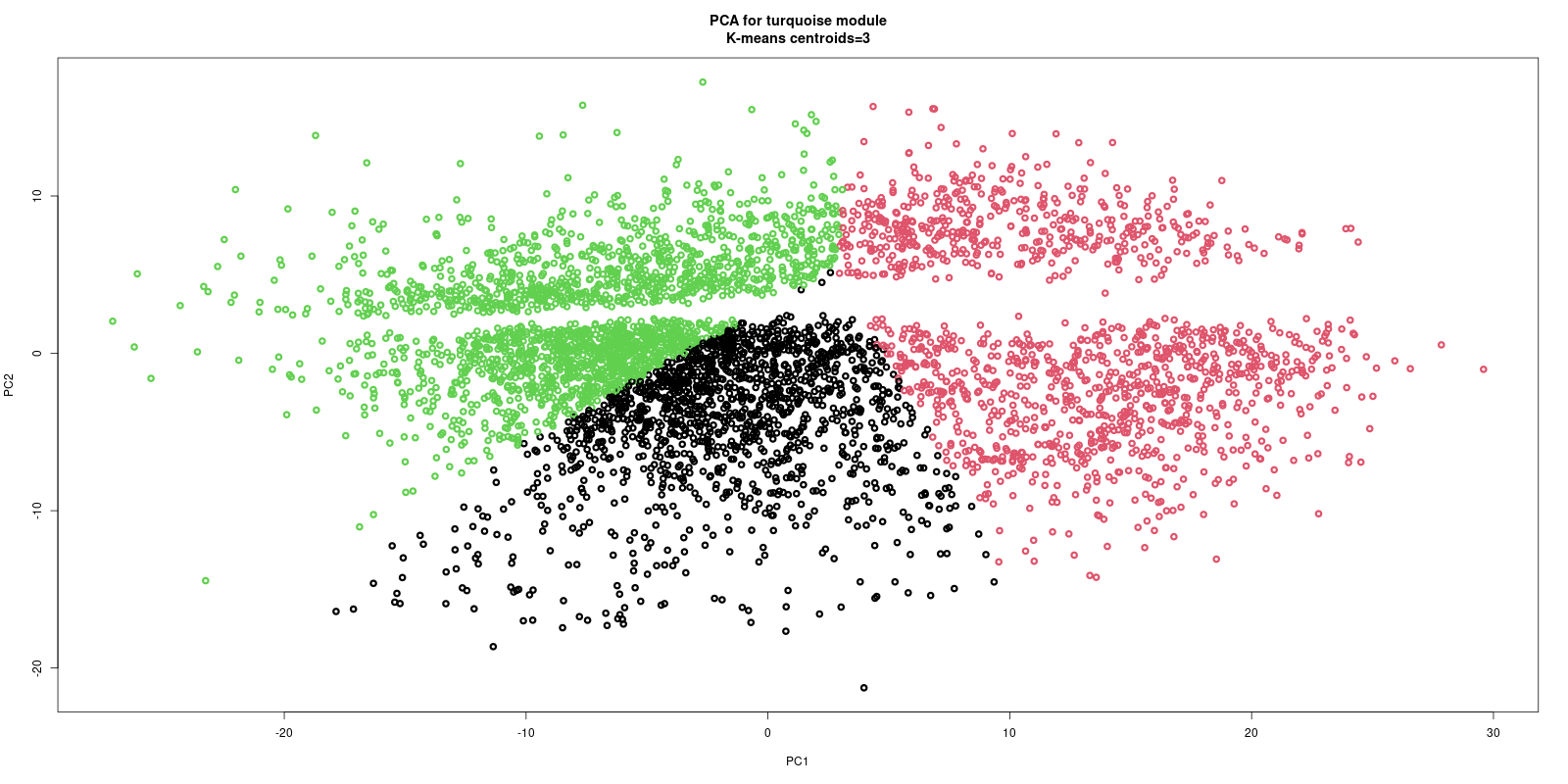
GitHub material: <https://github.com/Janek21/BDBI_TFG_MaizeCoexpression/tree/main/all_figures>

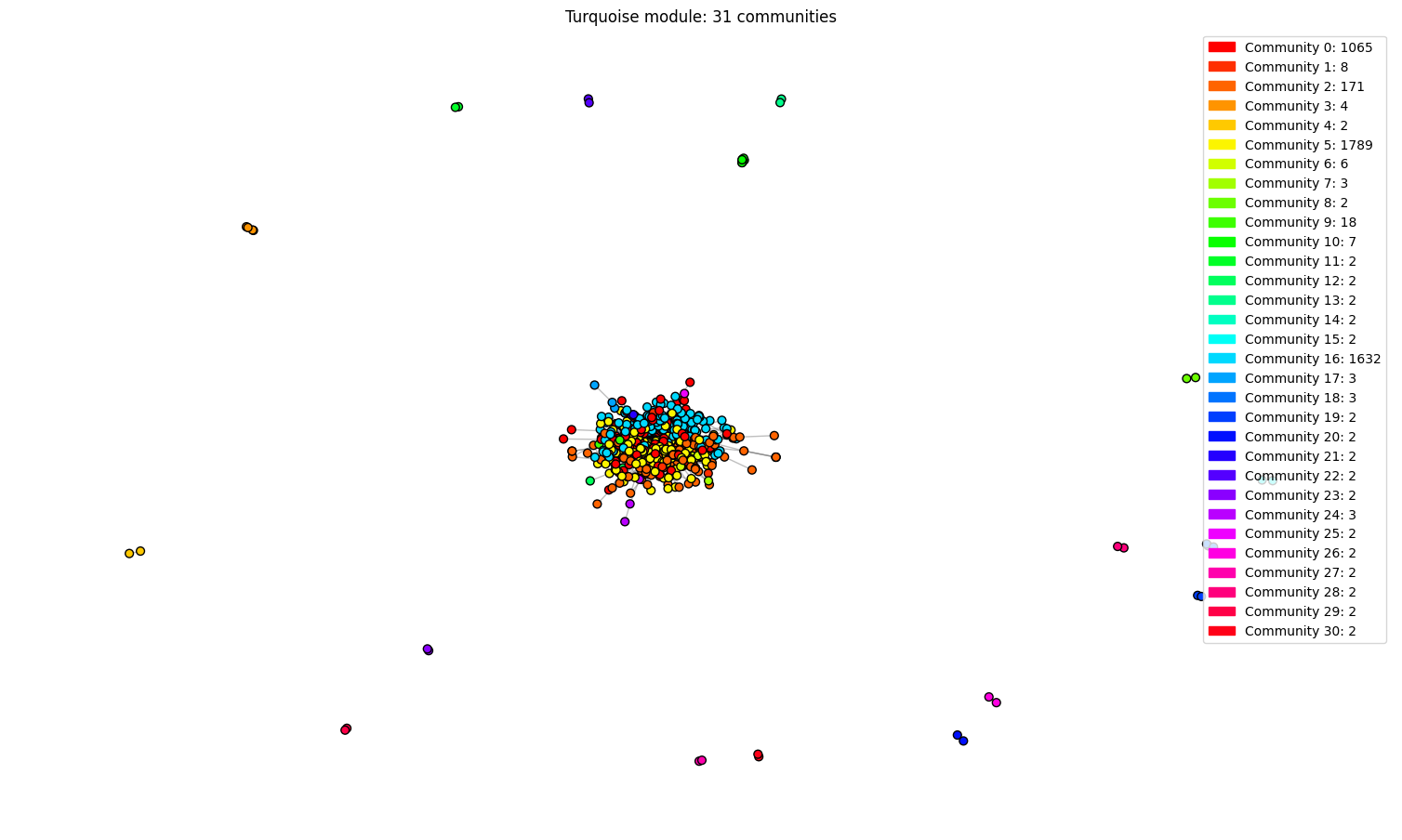
Figure 1: CPM distribution for 2 tissues of each line, the distribution shows overrepresentation in lowly expressed genes, even after filtering.

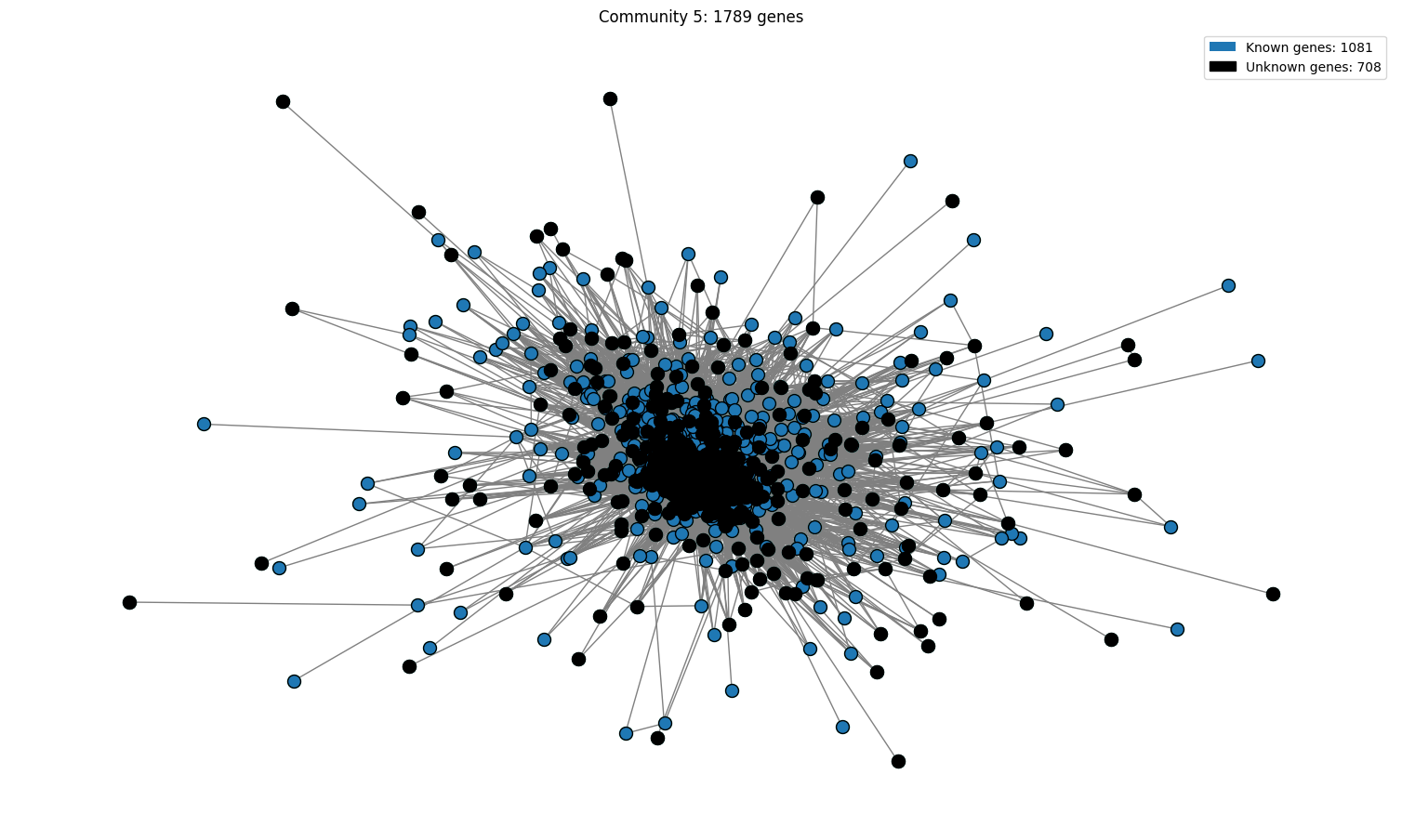
Figure 2: Representations of gene expression colored according to tissue. The representation is a PCA where clusters by expression similarity can be observed.

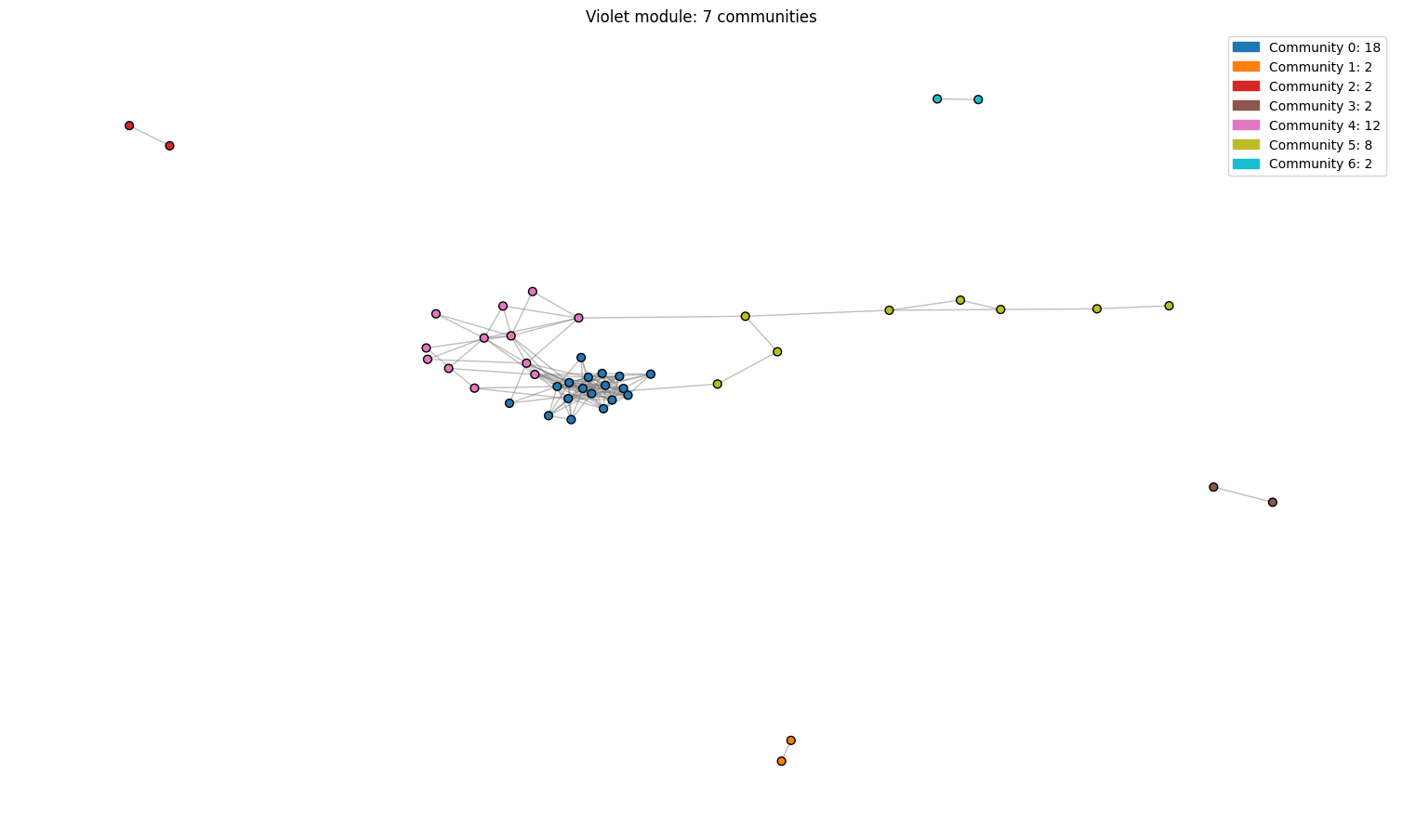
Figure 3: Classification of genes per module according to the network construction. The size of a module can also reflect its significance, since highly correlated genes tend to be involved in similar biological processes. This suggests that the function they perform is essential, as it is carried out by many different genes. (Kitano, 2004)

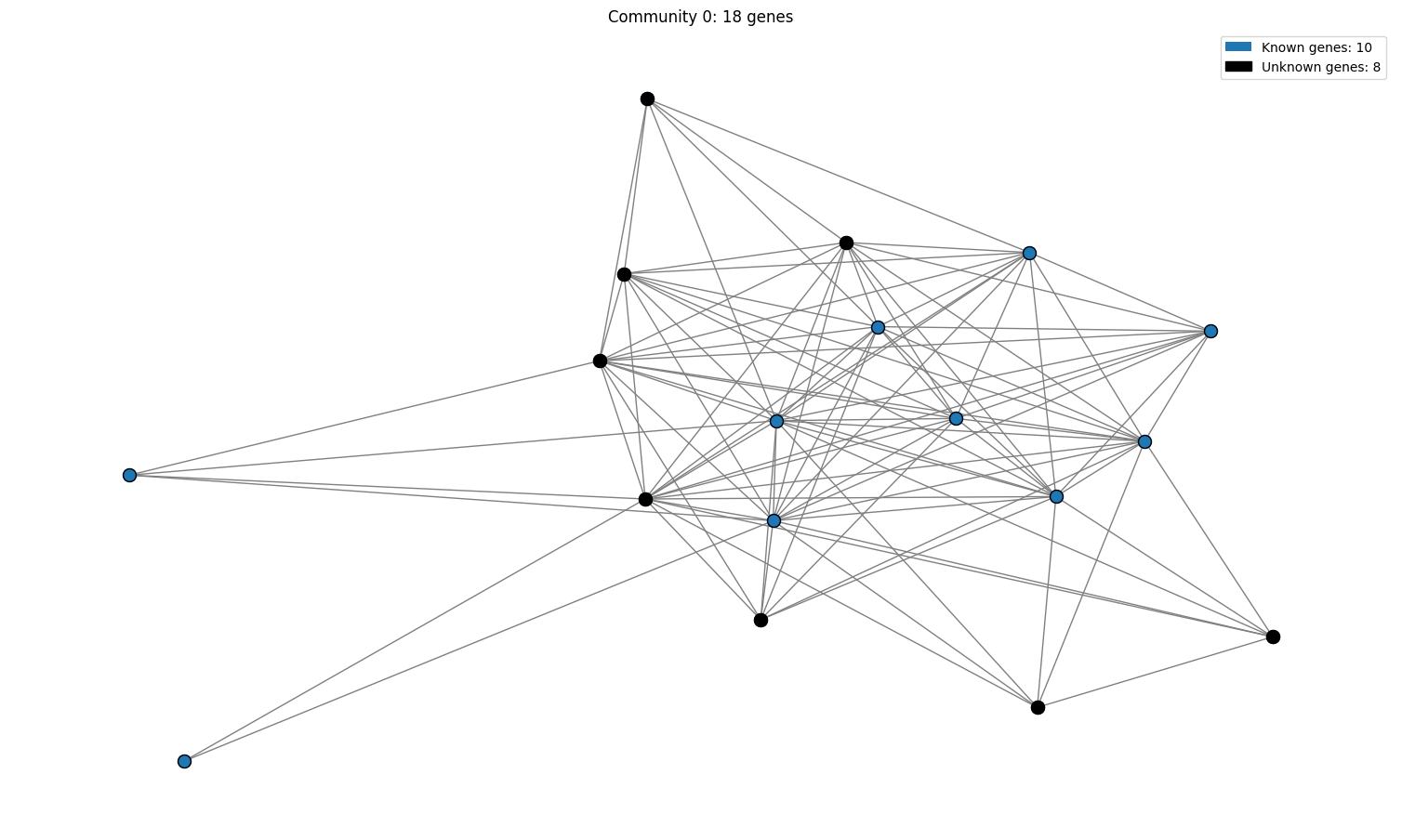
Figure 4: 10 most significant functions for all modules.

Figure 5: PCA of the gene co-expression in the turquoise module, colored by the 3 clusters determined by K-means

Figure 6: Complete visualization of the turquoise module colored according to communities.

Figure 7: This figure represents a community of the turquoise module, the unknown genes are marked in black while the known genes are blue. It can be observed that there is a very high proportion of unknown genes, which translates into a low confidence level in the significant functions.

Figure 8: Complete visualization of the violet module colored according to communities.

Figure 9: Visualization of a violet community.

|  |  |  |
| --- | --- | --- |
| *Community* | *Known* | *Unknown* |
| C0 | 537 | 528 |
| C1 | 952 | 714 |
| C2 | 94 | 98 |
| C3 | 2 | 4 |
| C5 | 1081 | 708 |
| C6 | 5 | 1 |
| C7 | 3 | 0 |
| C9 | 0 | 18 |
| C16 | 939 | 693 |
| C24 | 3 | 0 |

Table 1: Communities in the turquoise module, with the known and unknown genes for each