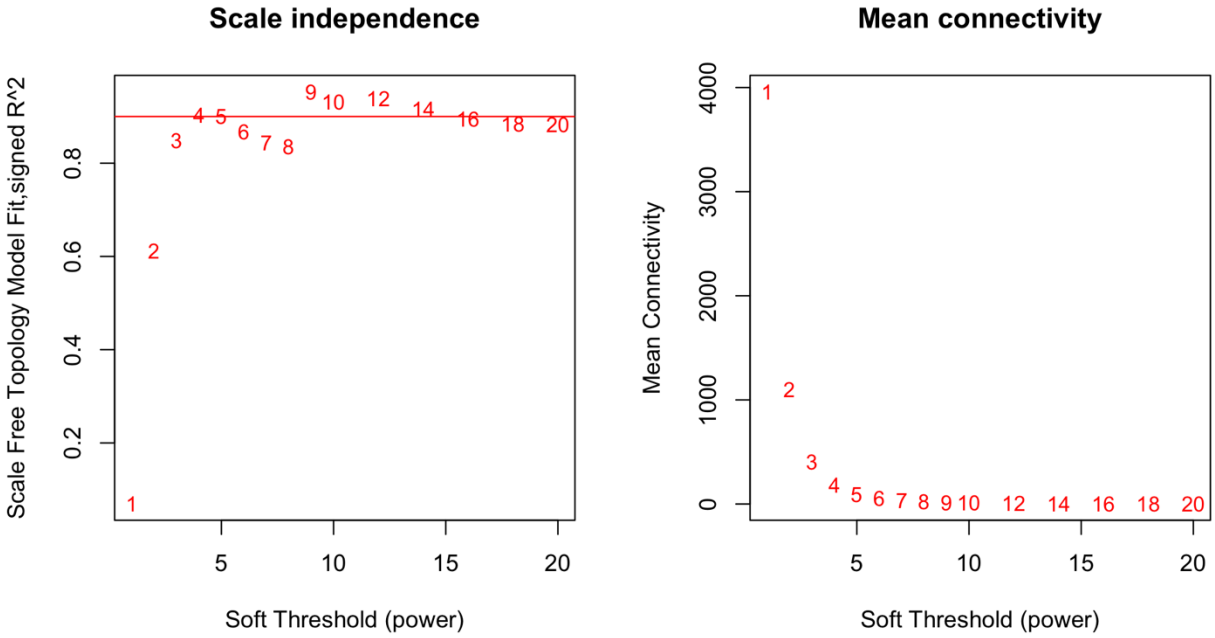
Systems-level network modeling deciphers the master regulators of phenotypic plasticity and heterogeneity in melanoma.

Firstly, after cleaning the dataset and mapping the genes, I proceeded with using WGCNA.

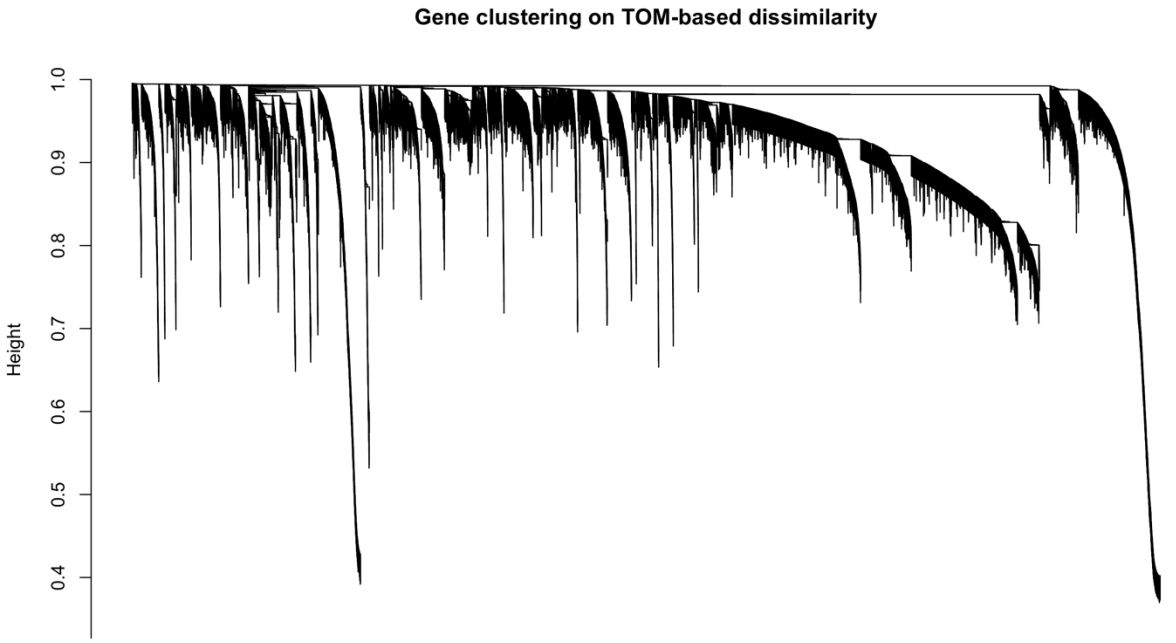
For WGCNA, I had to perform step by step network construction and module detection.

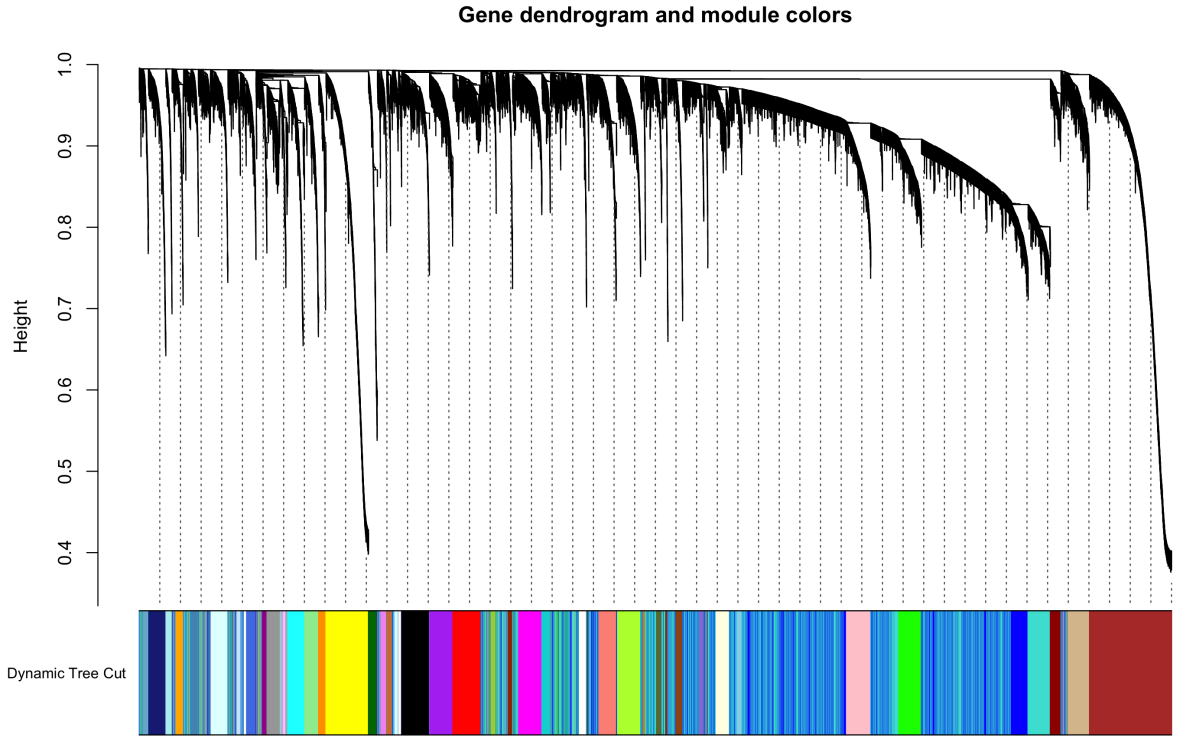


Power from this was identified to be **4.**

Next was the time for co-expression similarity & adjacency. Following this to minimize effects of noise and spurious associations, I transformed the adjacency into Topological Overlap Matrix, and calculate the corresponding dissimilarity.

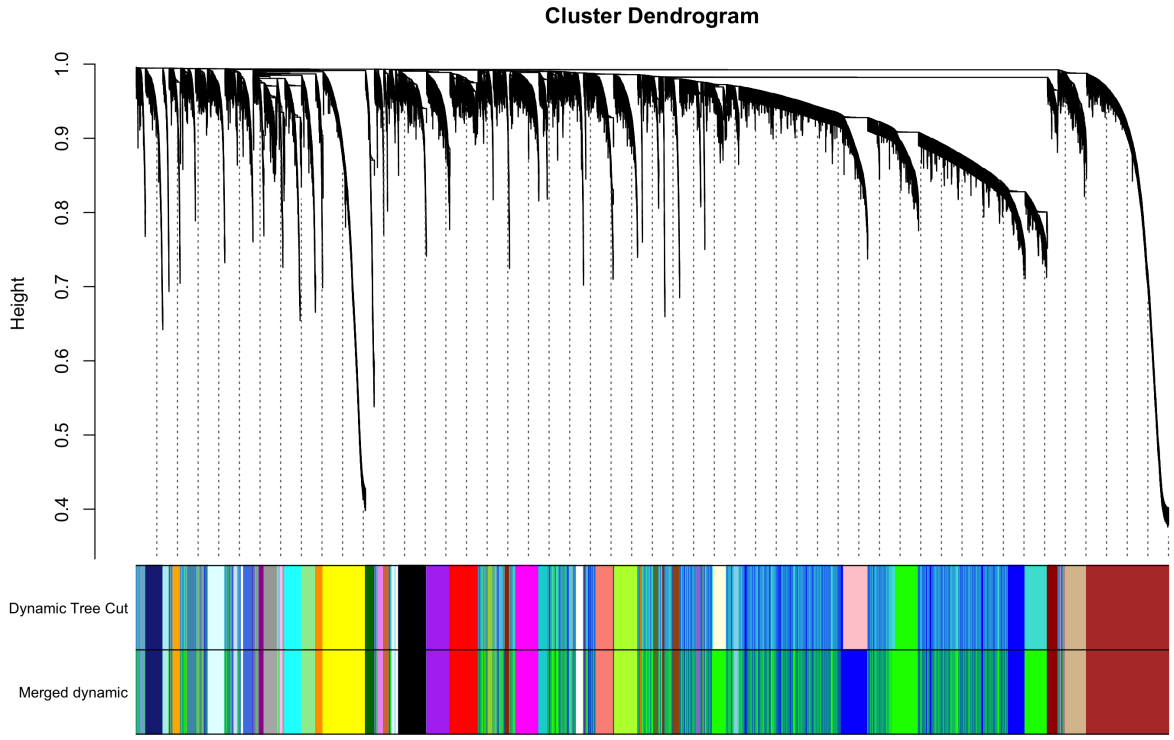
Following this I used hierarchical clustering to produce a hierarchical clustering tree (dendrogram) of genes.

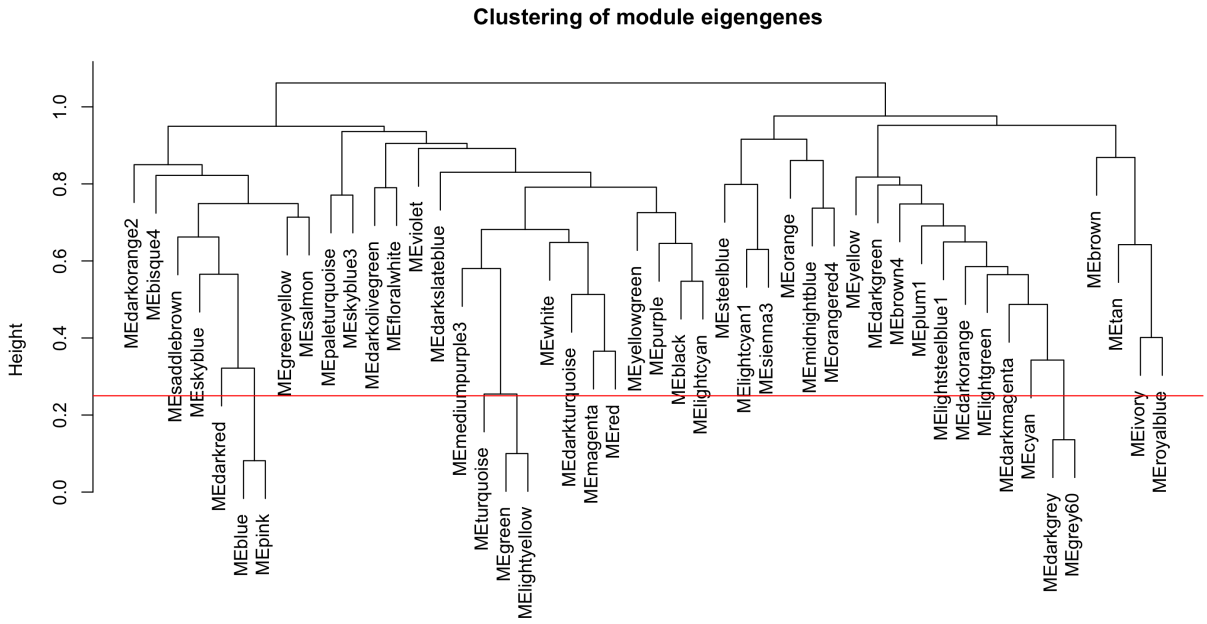




Following this, I merged modules whose expression profiles are very similar. It may be prudent to merge such modules since their genes are highly co-expressed. To quantify co-expression

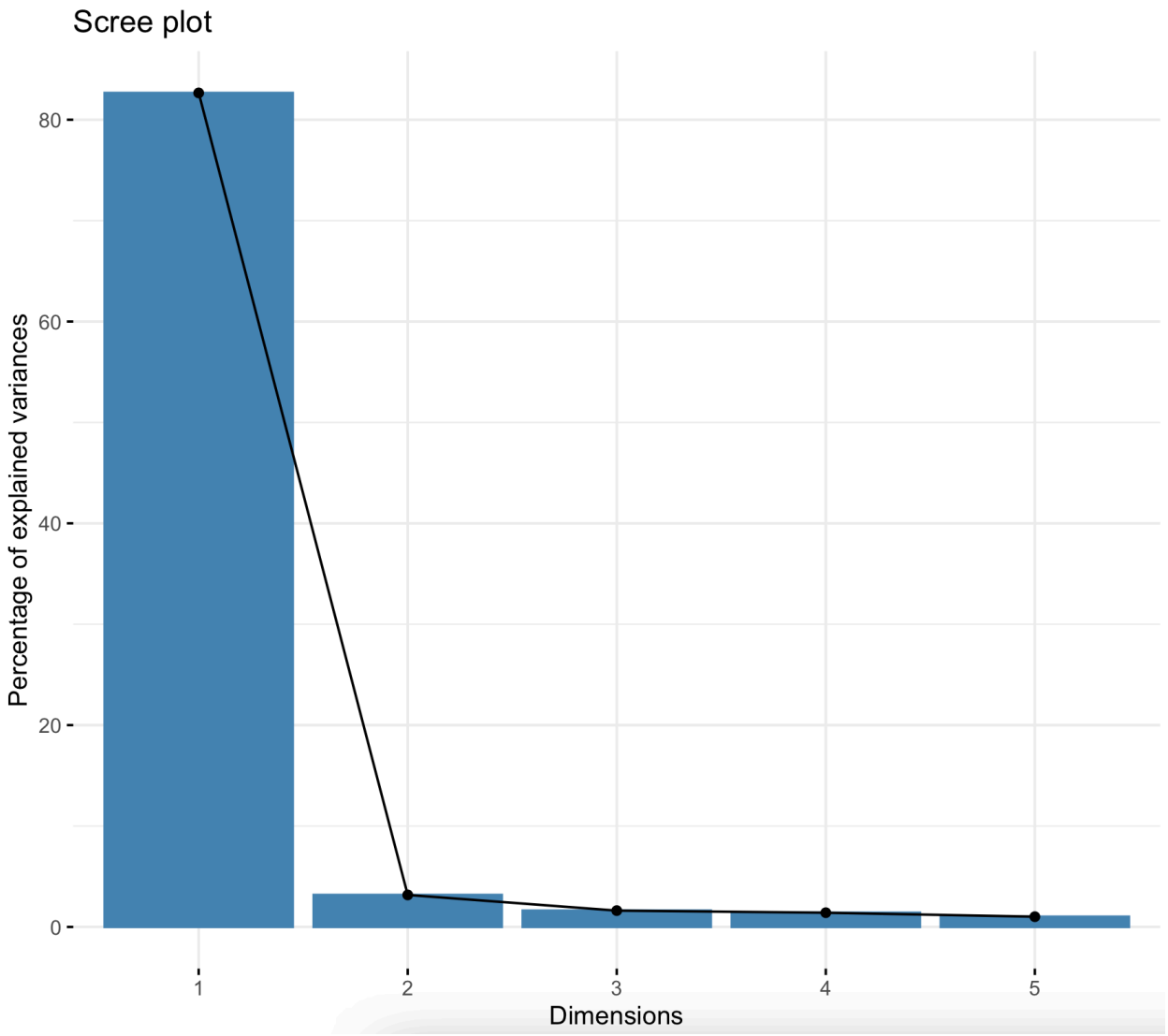
similarity of entire modules, I calculated their eigengenes and clustered them on their correlation.
To see what the merging did to the module colours, I plotted the gene dendrogram again, with the original and merged module colours underneath.





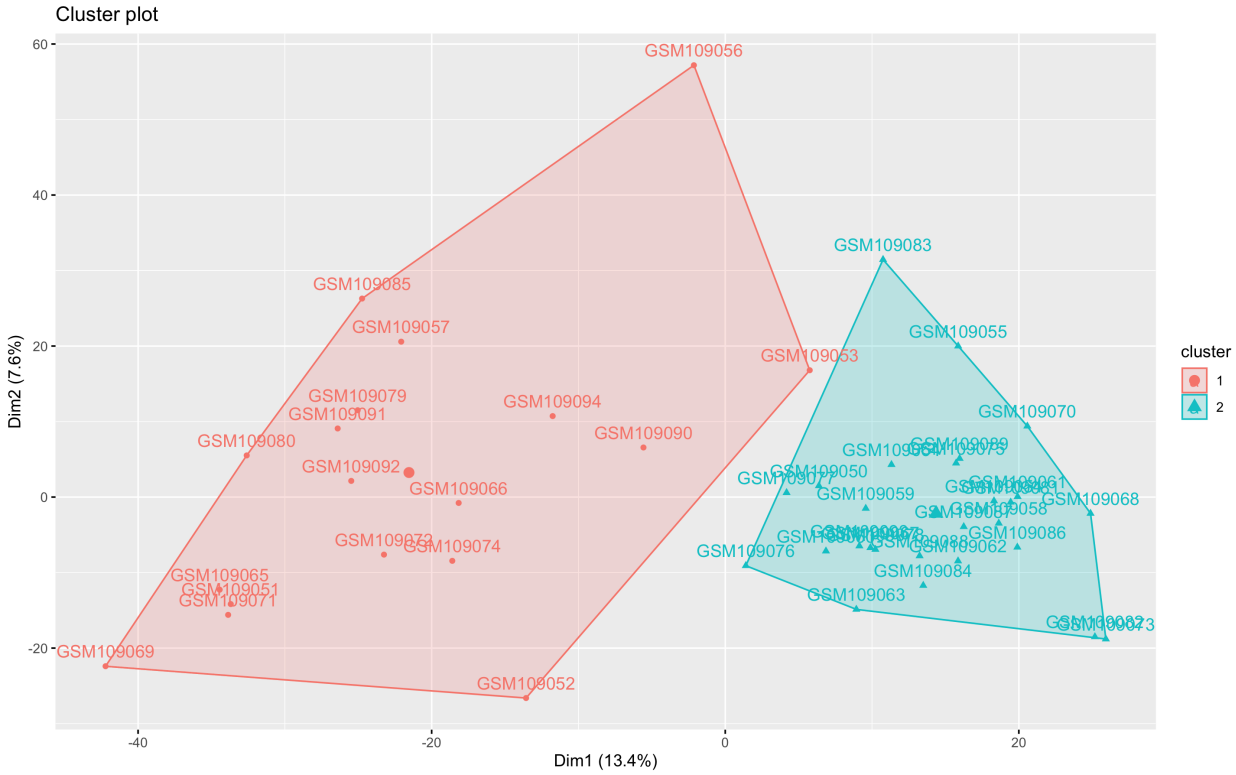
Row-wise variance for each gene was calculated. Then the column containing the variance was arranged in increasing order of values and the lowest 3000 genes were kept, and discarding the rest of them. This gave us the top 3000 genes with maximum variance. Following this, Principal component analysis was carried out.

Scree-Plot which gives an idea about the number of number of factors to retain in an exploratory factor analysis (FA) or principal components to keep in a principal component analysis (PCA).

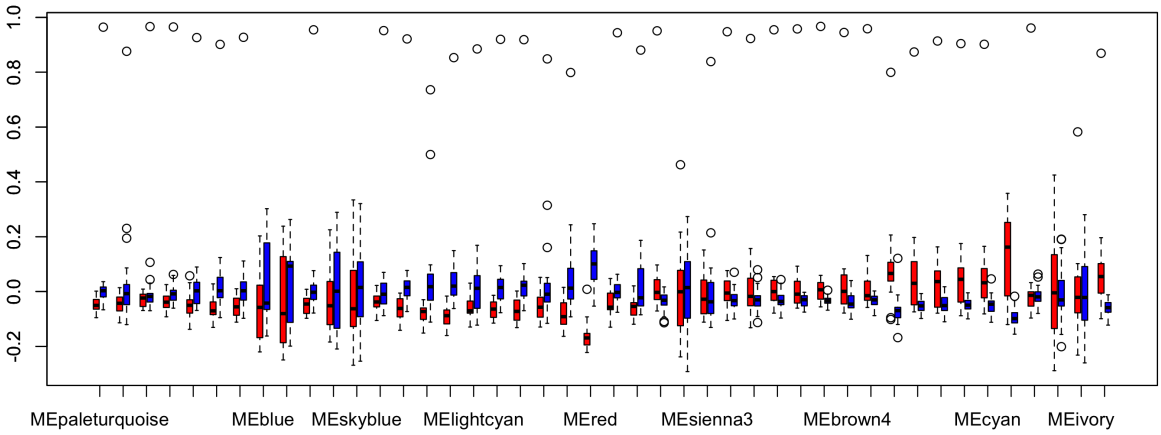


As this graph clearly points that 2 components should be taken into account, thus the rest were discarded. Following this the plot was created for k-means on the optimal number of centres which k-means for our data should contain.

Since the order of Genes might be different for the data used in the paper, and used by me, thus the graph obtained might be different.



Module eigengene values are shown for each sample and the corresponding expression level heatmap of genes in each module. This was performed on the data available on the top 3000 Genes with highest variance.


1.0
0.8
0.6
0.4
0.2
0.0
-0.2
MEpaleturquoise
MEblue
MEskyblue
MElightcyan
MEred
MEsienna3
MEbrown4
MEcyan
MEivory