

# **Supplementary materials**

Differential gene expression analysis in mental  
disorders

Network Based Analysis

**Rayan Slatni**

June 2023

Gene	Function	Control	Bipolar	Depression	Schizophrenia
FKBP4	FKBP Prolyl Isomerase 4	4.68	4.47	4.37	4.60
LMTK2	Lemur Tyrosine Kinase 2	1.92	2.04	1.68	1.48
NFIB	Nuclear Factor I B	2.09	2.28	2.32	2.44
MAG	Myelin Associated Glycoprotein	5.23	5.14	4.82	4.27
SOX10	SRY-Box Transcription Factor 10	2.91	2.72	2.62	2.15
FSCN1	Fascin Actin-Bundling Protein 1	5.46	5.27	5.38	5.17
MAP3K12	Mitogen-Activated Protein Kinase	2.09	1.78	1.98	2.20
PLXND1	Plexin D1	2.48	2.19	2.48	2.35
NCAM1*	Neural Cell Adhesion Molecule 1	4.49	4.14	4.18	4.31
STIP1*	Stress Induced Phosphoprotein 1	3.14	2.88	2.93	3.10
ARHGEF1*	Rho Guanine Nucleotide Exchange Factor 1	1.58	1.11	1.50	1.45
TFE3	F Binding To IGHM Enhancer 3	2.15	1.84	2.08	2.04
RPS6KA1	Ribosomal Protein S6 Kinase A1	1.34	0.96	1.17	0.98
DAPK1	Death Associated Protein Kinase 1	0.76	0.78	0.79	1.12
SLC29A1	Solute Carrier Family 29 Member 1	2.42	2.07	2.30	2.37
NEDD4L*	Ubiquitin Protein Ligase	1.88	1.73	1.92	2.10
TPM4	Tropomyosin 4	1.39	1.19	1.48	1.53
PDHB	Pyruvate Dehydrogenase E1 Sub Beta	3.39	3.67	3.20	3.29
DGKA	Diacylglycerol Kinase Alpha	2.51	2.37	2.75	2.67
CACAN1C	Calcium Voltage-Gated Channel	3.50	3.24	3.38	3.32
ARHGEF2	Rho Guanine Nucleotide Exchange Factor 2	4.08	3.72	4.00	3.82

Table 1: Mean of gene expression level for each group of the important genes. The value are calculated on data with log2 transformation and normalization respect the housekeeping gene. The genes with \* have an expression level that is not coherence with the literature or with network enrichment analysis

```
R version 4.2.3 (2023-03-15 ucrt)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 22621)

Matrix products: default

locale:
[1] LC_COLLATE=Italian_Italy.utf8
[2] LC_CTYPE=Italian_Italy.utf8
[3] LC_MONETARY=Italian_Italy.utf8
[4] LC_NUMERIC=C
[5] LC_TIME=Italian_Italy.utf8

attached base packages:
[1] stats4      stats       graphics  grDevices ut
[6] datasets    methods    base

other attached packages:
[1] hgu95av2.db_3.13.0   pathfindR_2.1.0
[3] pathfindR.data_2.0.0 org.Hs.eg.db_3.16.0
[5] AnnotationDbi_1.60.2 IRanges_2.32.0
[7] S4Vectors_0.36.2    KEGGgraph_1.58.3
[9] KEGGREST_1.38.0     BiocManager_1.30.20
[11] gprofiler2_0.2.1    gProfileR_0.7.0
[13] rScudo_1.14.0      colorBlindness_0.1.9
[15] glue_1.6.2          RCy3_2.18.0
[17] gtools_3.9.4        e1071_1.7-13
[19] vctrs_0.6.1         ggrepify_0.4.16
[21] factoextra_1.0.7   useful_1.2.6
[23] GEOquery_2.66.0     lubridate_1.9.2
[25] forcats_1.0.0       stringr_1.5.0
[27] dplyr_1.1.0          purrr_1.0.1
[29] readr_2.1.4          tidyrr_1.3.0
[31] tibble_3.2.1         tidyverse_2.0.0
[33] ROCR_1.0-11         reshape2_1.4.4
[35] RColorBrewer_1.1-3  randomForest_4.7-1.1
[37] pROC_1.18.0          MASS_7.3-58.3
[39] glmnet_4.1-7         Matrix_1.5-4
[41] ggpubr_0.6.0         gplots_3.1.3
[43] genefilter_1.80.3   caret_6.0-94
[45] lattice_0.21-8      ggplot2_3.4.2
[47] Biobase_2.58.0      BiocGenerics_0.44.0
```

Figure 1: R session information

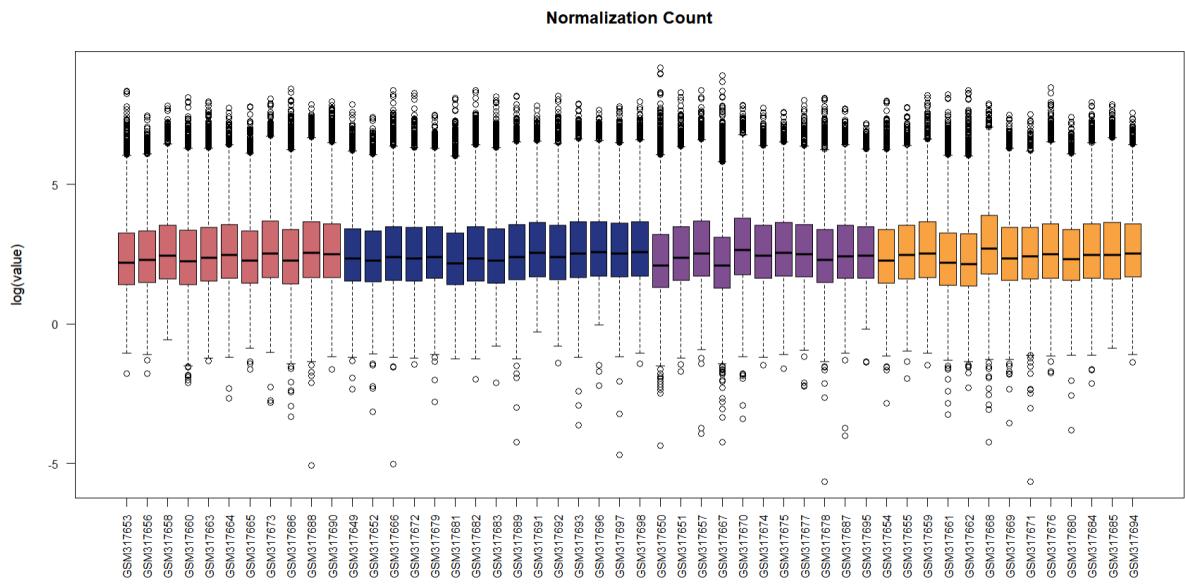


Figure 2: Gene distribution among samples(n=50). Bipolar (n=11 red), control (n=15 blue), depression (n=11, orange), schizophrenia (n=13, purple).

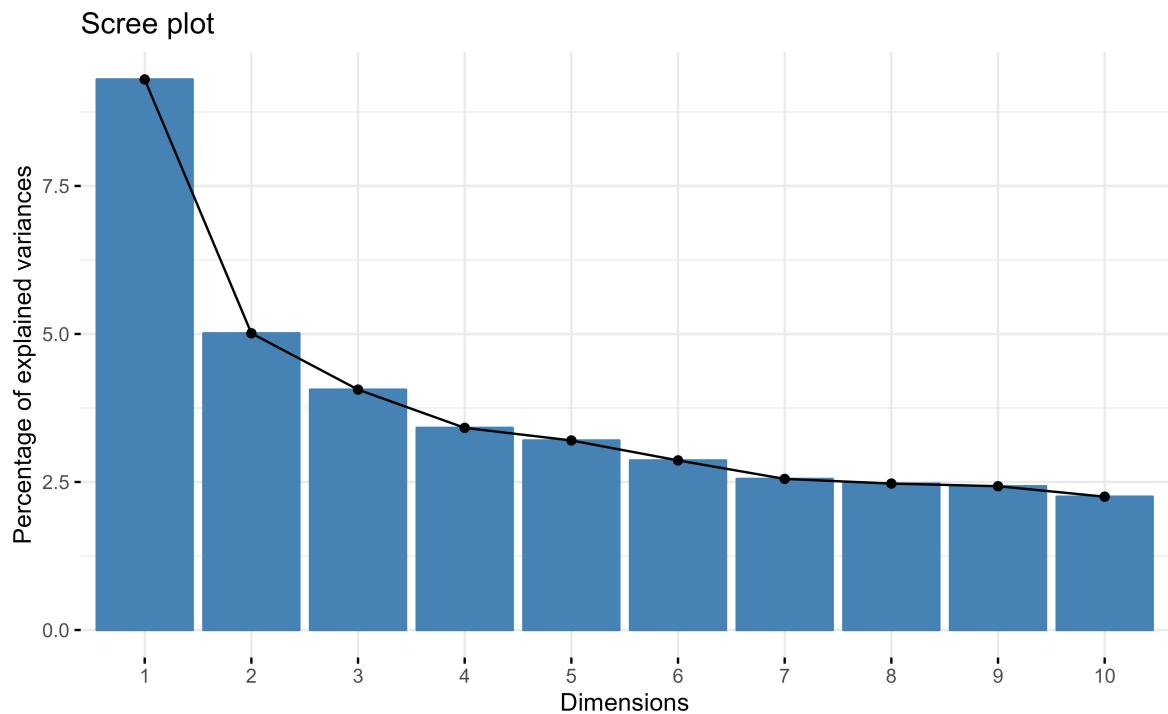


Figure 3: Scree plot of first 10 PC

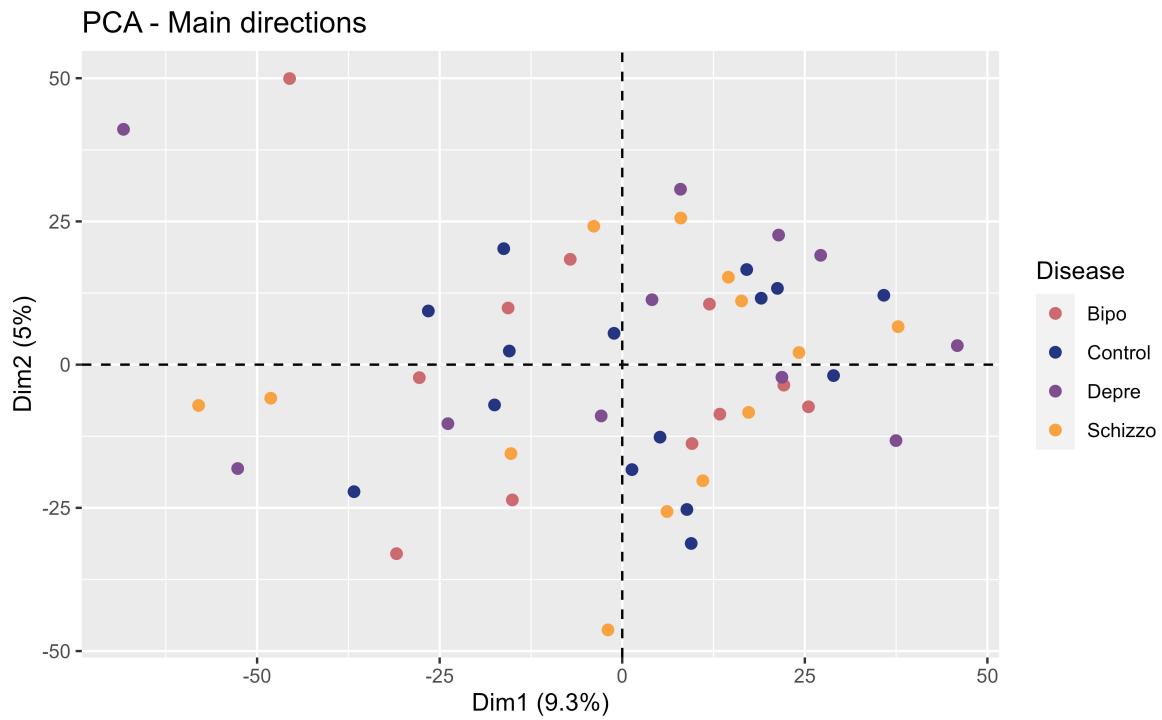


Figure 4: PCA plot

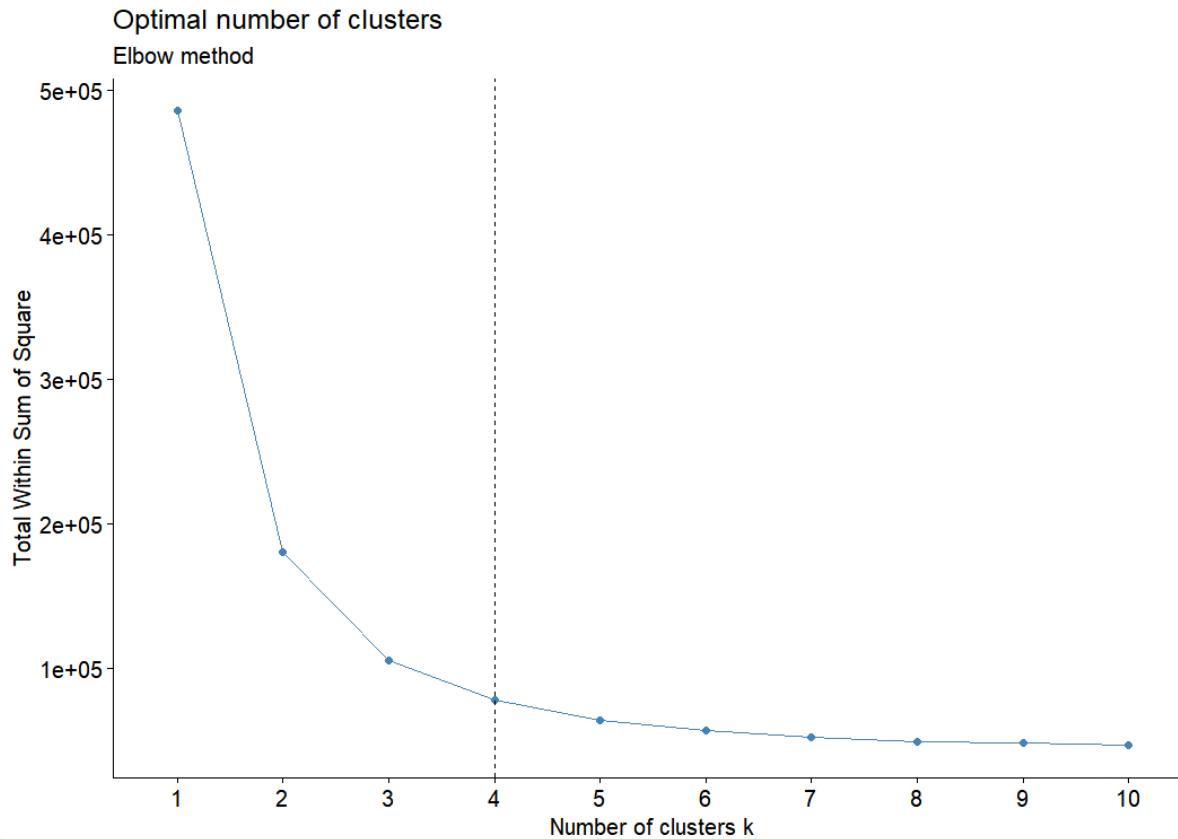


Figure 5: Elbow method for selecting the optimal number of cluster

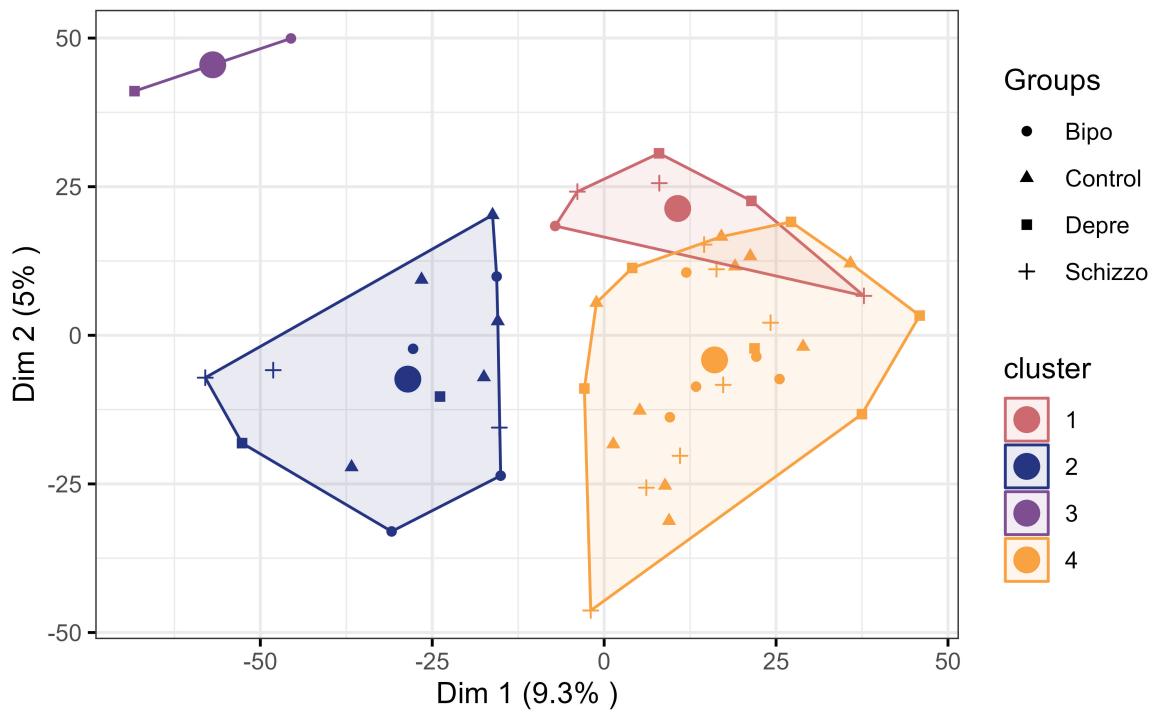


Figure 6: Kmeans plot

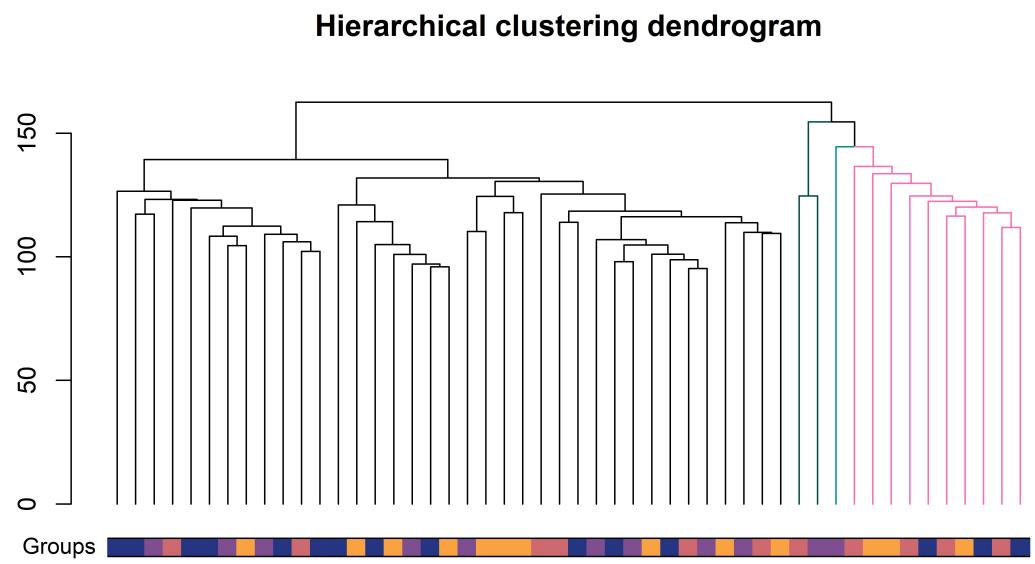


Figure 7: Hierarchical clustering dendrogram: bipolar (red), control (blue), depression (orange), schizophrenia (purple)

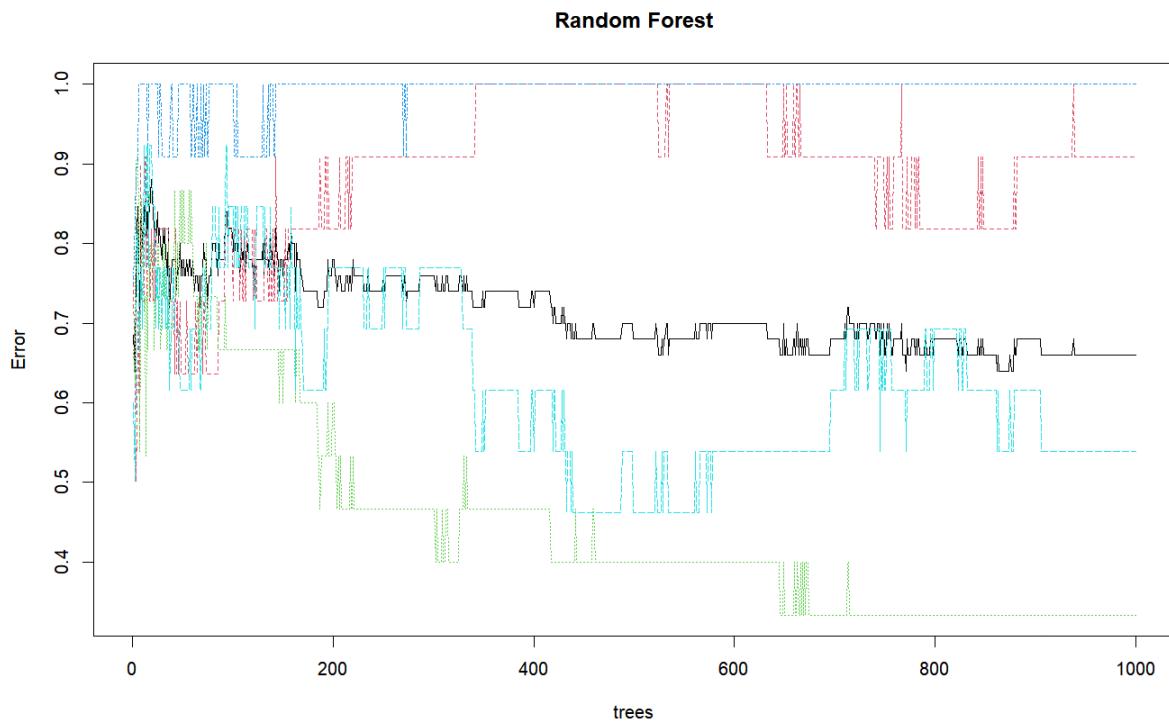


Figure 8: Random Forest error: each color represent a group, black is the mean.

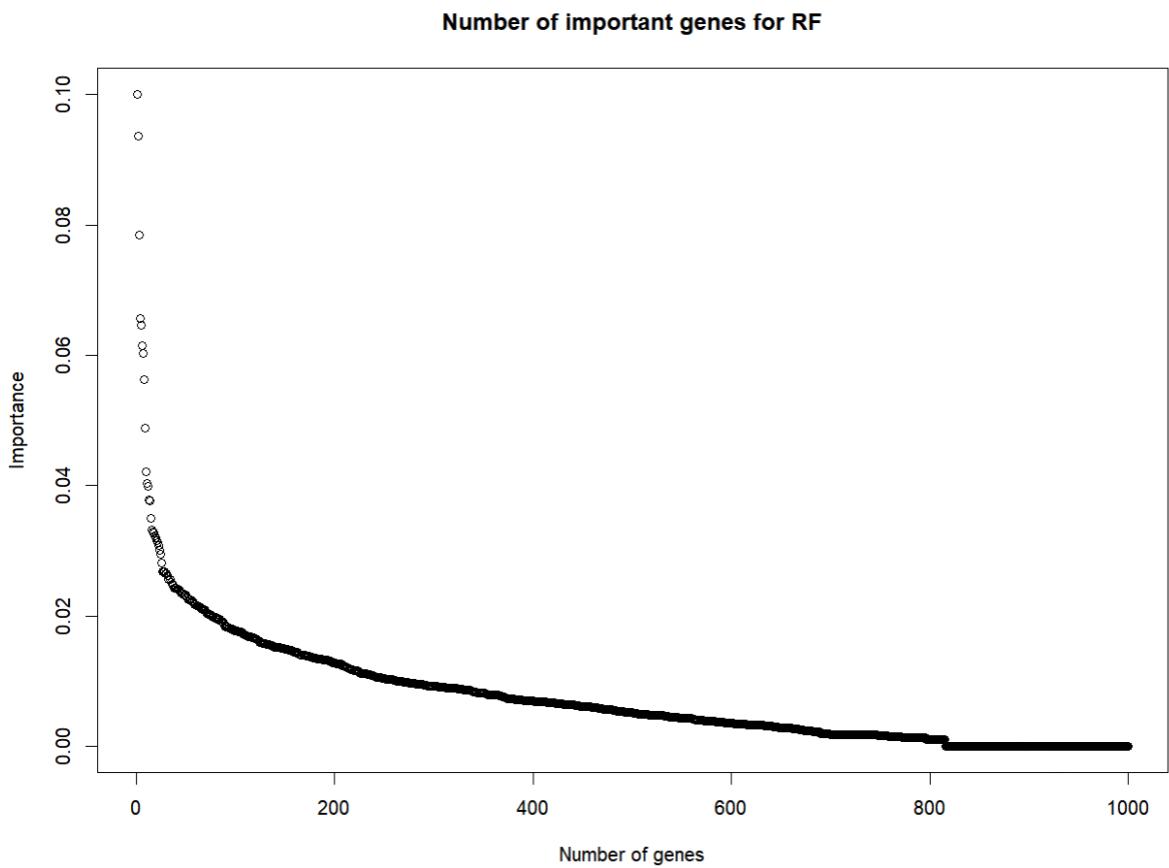


Figure 9: Number of genes important for RF

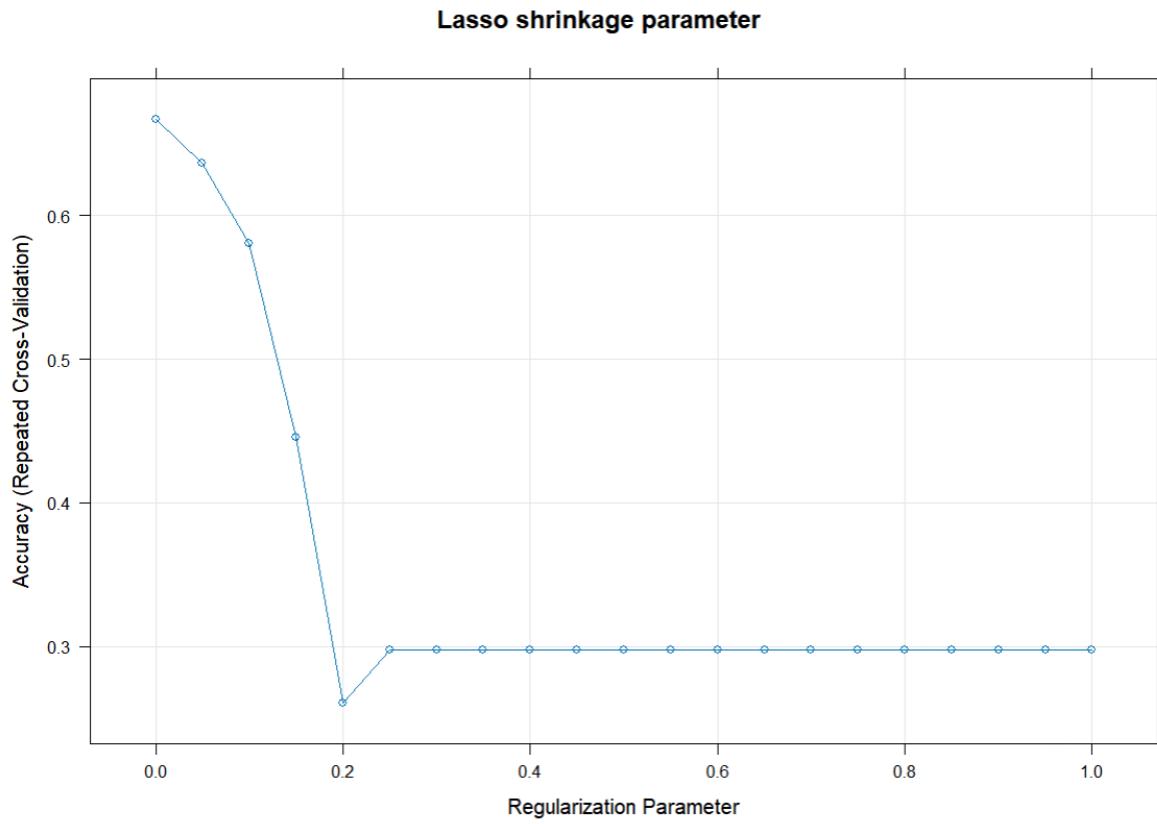


Figure 10: Lasso shrinkage parameter

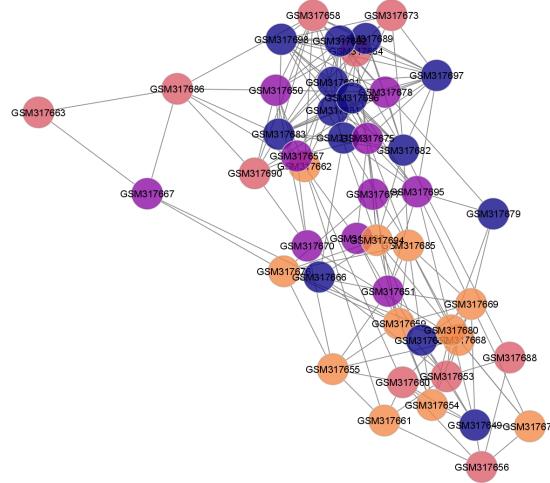


Figure 11: Scudo network with Cytoscape: bipolar (red), control (blue), depression (orange), schizophrenia (purple)

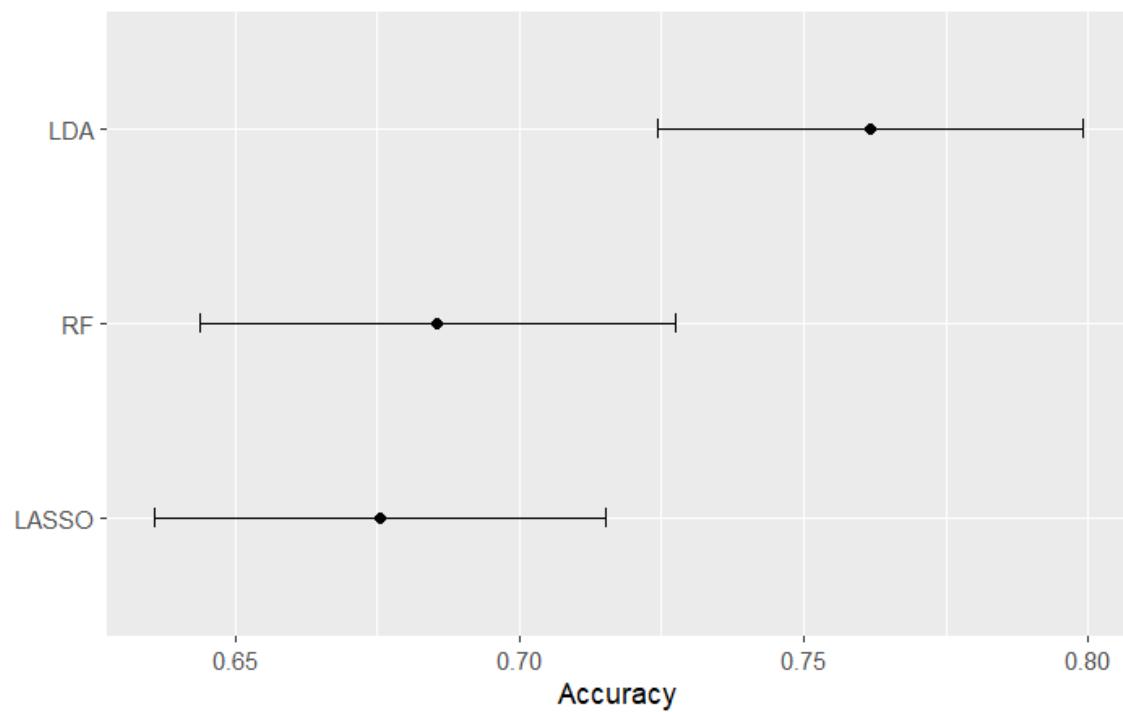


Figure 12: Model accuracy comparing LDA, Lasso and RF using repeated cross-validation

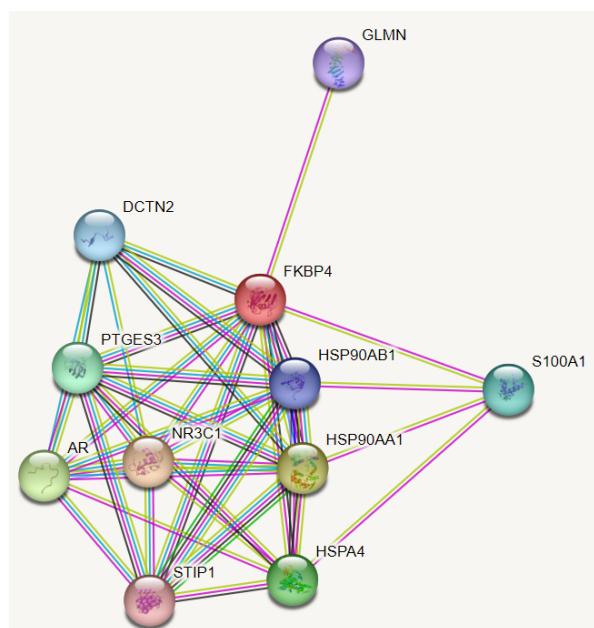


Figure 13: String network linking FKBP4 to STIP1

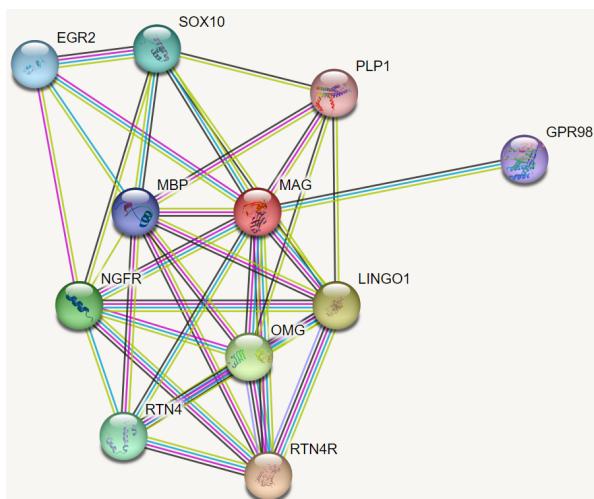


Figure 14: String network linking MAG to SOX10