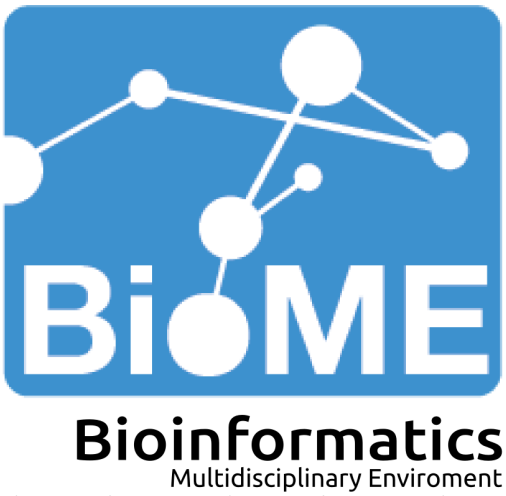


RNASeq Analysis Report

Copyright © 2016 BioME - Bioinformatics Multidisciplinary Environment
HTTP://BIOINFO.IMD.UFRN.BR/
Report compiled in December, 2016

1. Report

1.1 Pre-Mapping Analysis

1.1.1 Contamination Test

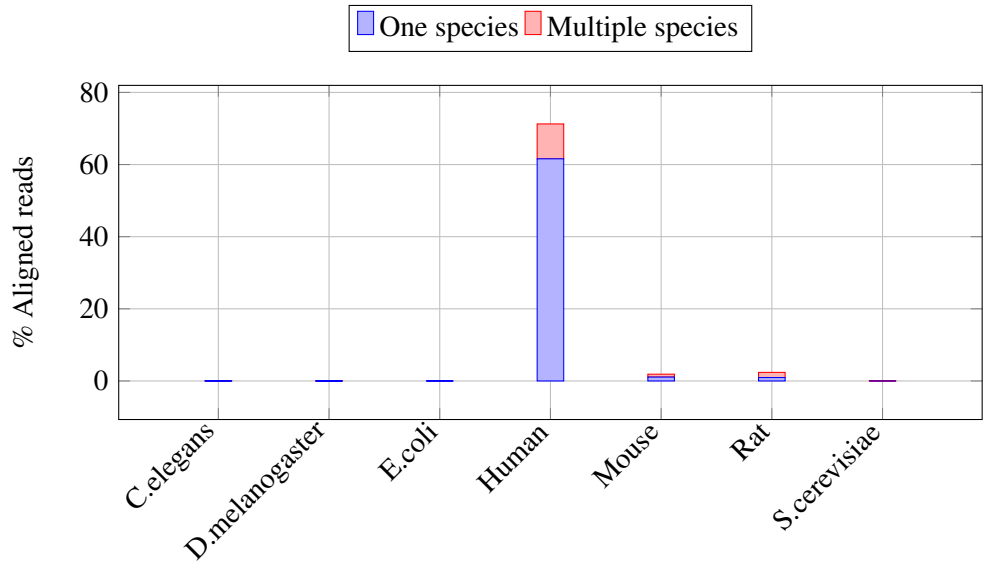


Figure 1.1: Contamination Test For Sample 1 - Cerebellum. All reads were aligned to the genomes of the species shown in the X axis.

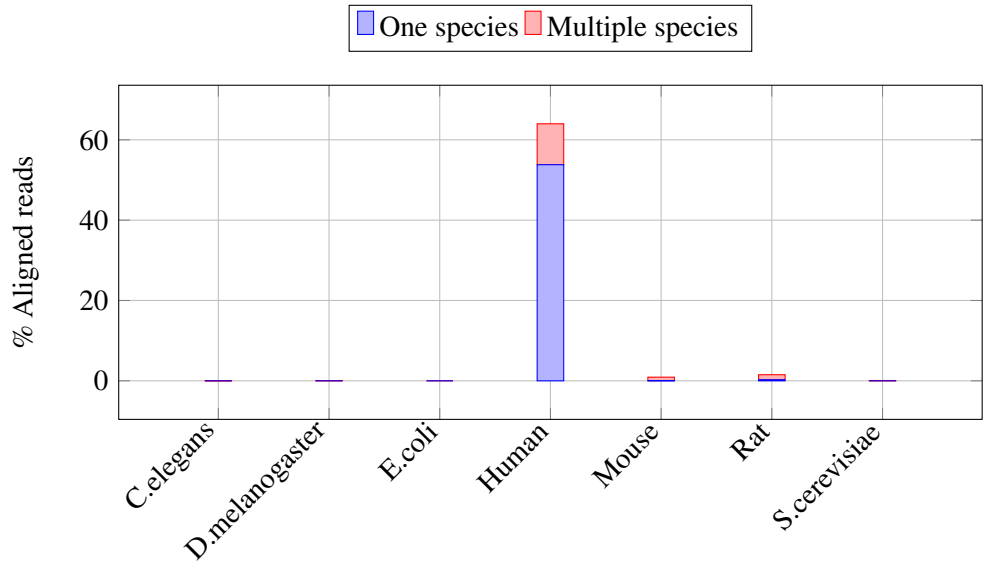


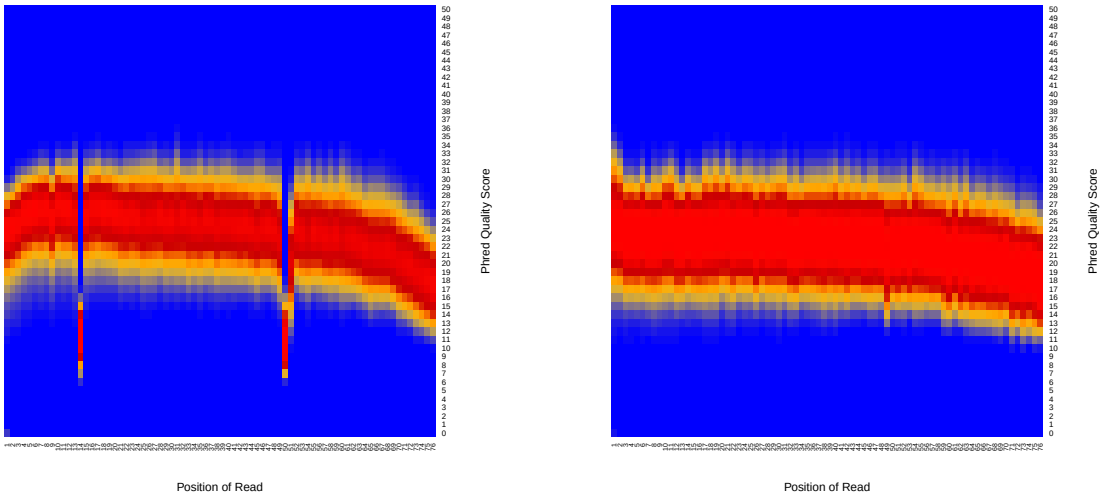
Figure 1.2: Contamination Test For Sample 2 - Prefrontal Cortex. All reads were aligned to the genomes of the species shown in the X axis.

1.2 Post-Mapping Analysis

1.2.1 Mapping and Read Quality

Sample ID	Sequenced reads	Aligned reads	MAPQ < 30
Sample 1 - Cerebellum	32.698.558	21.185.378	1.510.419 (7,129% of aligned reads)
Sample 2 - Prefrontal Cortex	24.325.223	14.814.722	804.050 (5,43% of aligned reads)

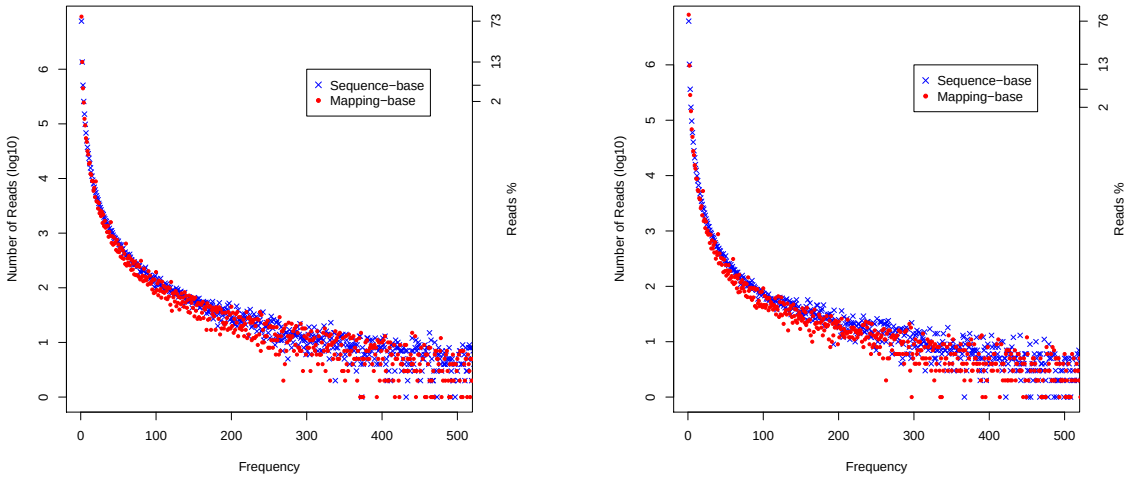
Table 1.1: Mapping overview.



(a) Sample 1 - Cerebellum

(b) Sample 2 - Prefrontal Cortex

Figure 1.3: Heatmap of read quality. X-axis corresponds to the nucleotide position within the read. Y-axis corresponds to the Phred score of the corresponding nucleotide.



(a) Sample 1 - Cerebellum

(b) Sample 2 - Prefrontal Cortex

Figure 1.4: Duplication rate of reads. Reads starting and ending at the exact same position can be products of artefacts at the amplification step. This plot reports the number of such cases.

1.2.2 Expression Distribution

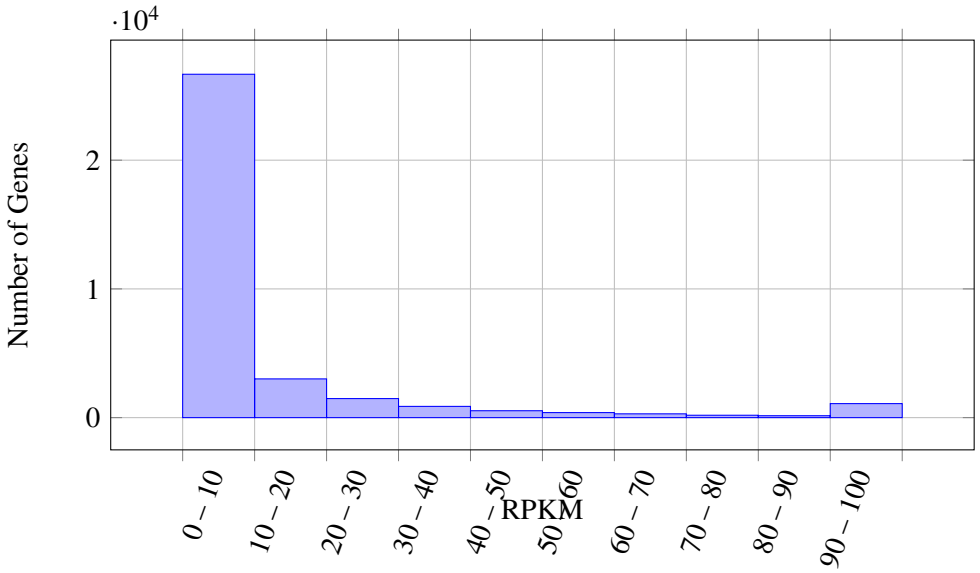


Figure 1.5: Expression Distribution For Sample 1 - Cerebellum.

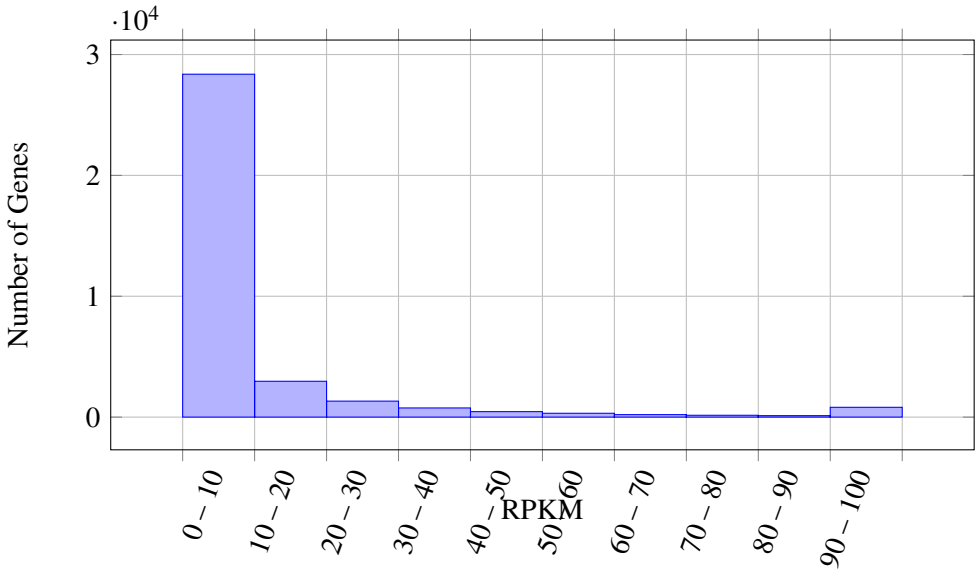
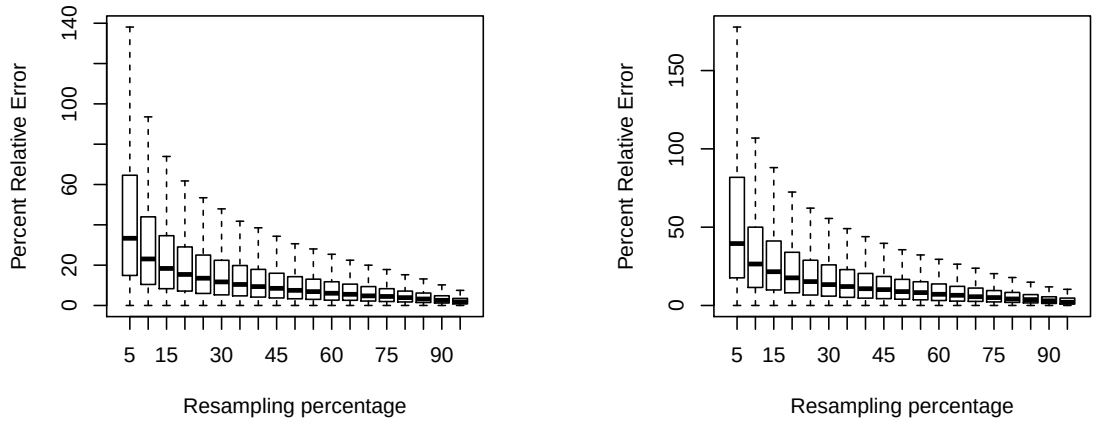


Figure 1.6: Expression Distribution For Sample 2 - Prefrontal Cortex.

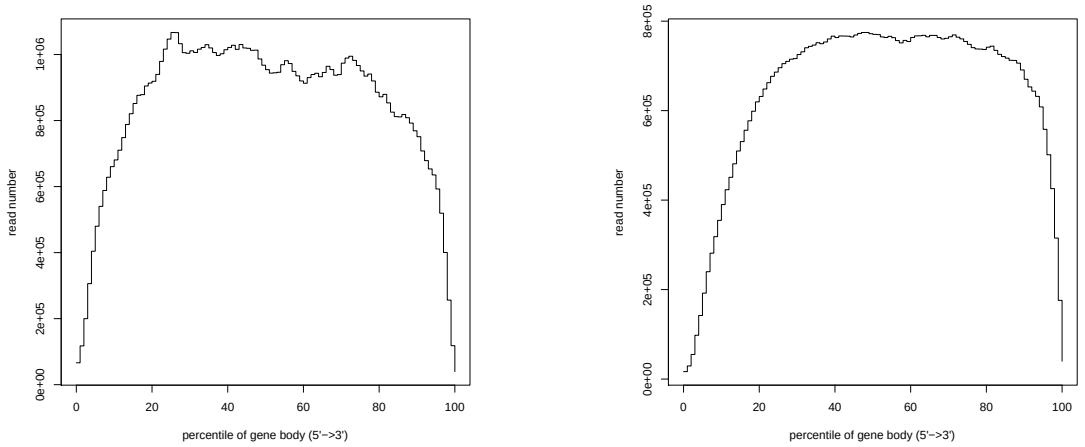
1.2.3 Coverage Quality



(a) Sample 1 - Cerebellum

(b) Sample 2 - Prefrontal Cortex

Figure 1.7: RPKM saturation analysis for the 25% lowest expressed genes



(a) Sample 1 - Cerebellum

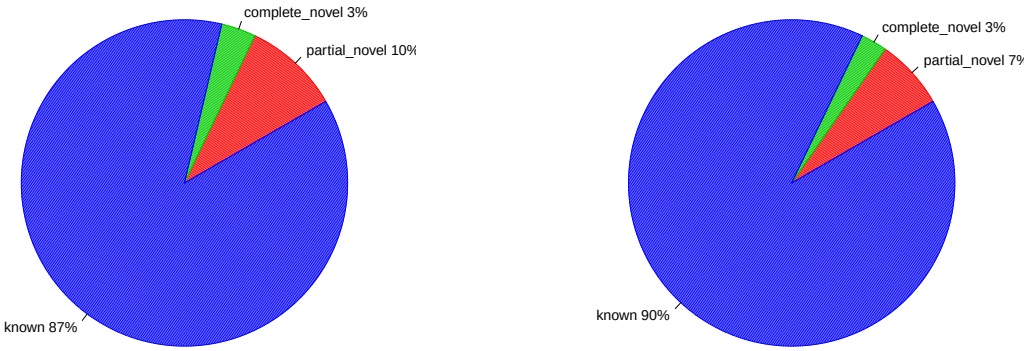
(b) Sample 2 - Prefrontal Cortex

Figure 1.8: Overall transcript coverage. X-axis is a representation of a transcript where 0 and 100 correspond to the 5' and 3' ends, respectively.

1.2.4 Splicing Overview

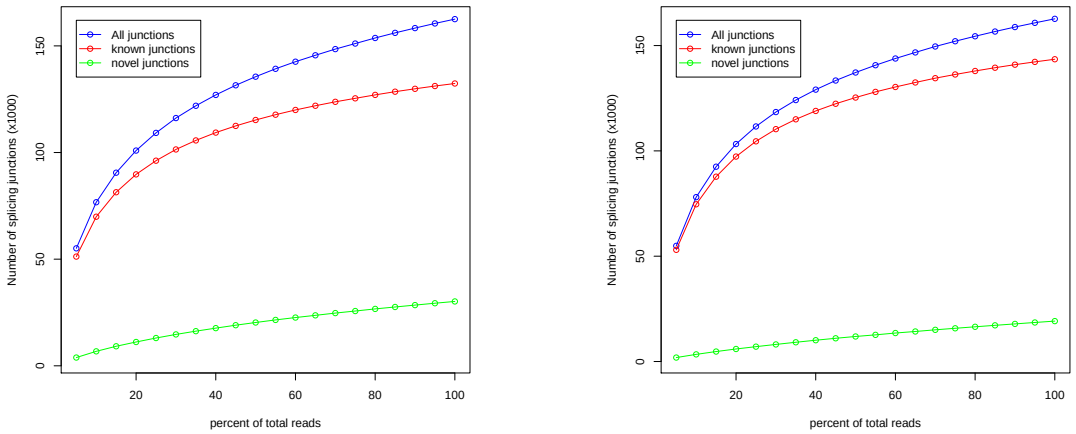
Sample ID	Identified Junctions	Known Junctions	Unknown Junctions
Sample 1 - Cerebellum	162.548	141.428	21.120
Sample 2 - Prefrontal Cortex	162.696	147.152	15.544

Table 1.2: Splicing junctions statistics. Known junctions are based on the respective set of Reference Sequences.



(a) Sample 1 - Cerebellum (b) Sample 2 - Prefrontal Cortex

Figure 1.9: Junctions overview



(a) Sample 1 - Cerebellum (b) Sample 2 -Prefrontal Cortex

Figure 1.10: Saturation analysis for splicing junctions

1.3 Differential Expression Analysis

Sample ID	Total Genes	DEG (p <= 0.01)
Prefrontal Cortex vs Cerebellum	55.766	236

Table 1.3: Statistics on differentially expressed genes

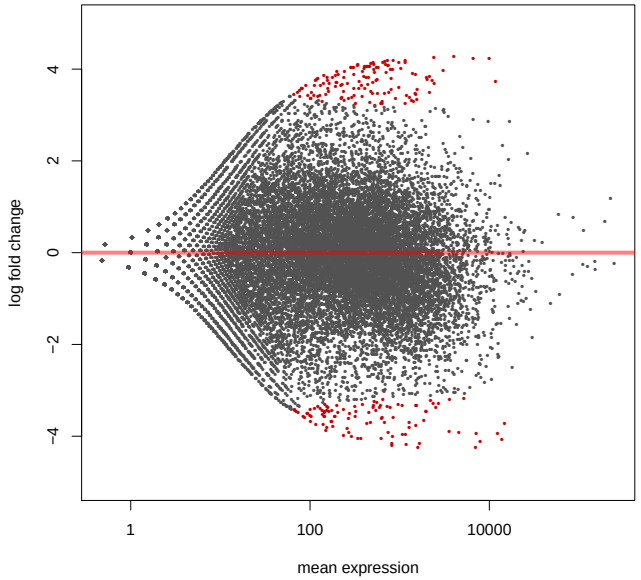


Figure 1.11: Overview of differentially expressed genes. Cerebellum is the reference.

1.4 Functional Analysis

1.4.1 Gene Ontology - Biological Process

Figure 1.12: Frequency of genes by GO group. All differentially expressed genes were used.

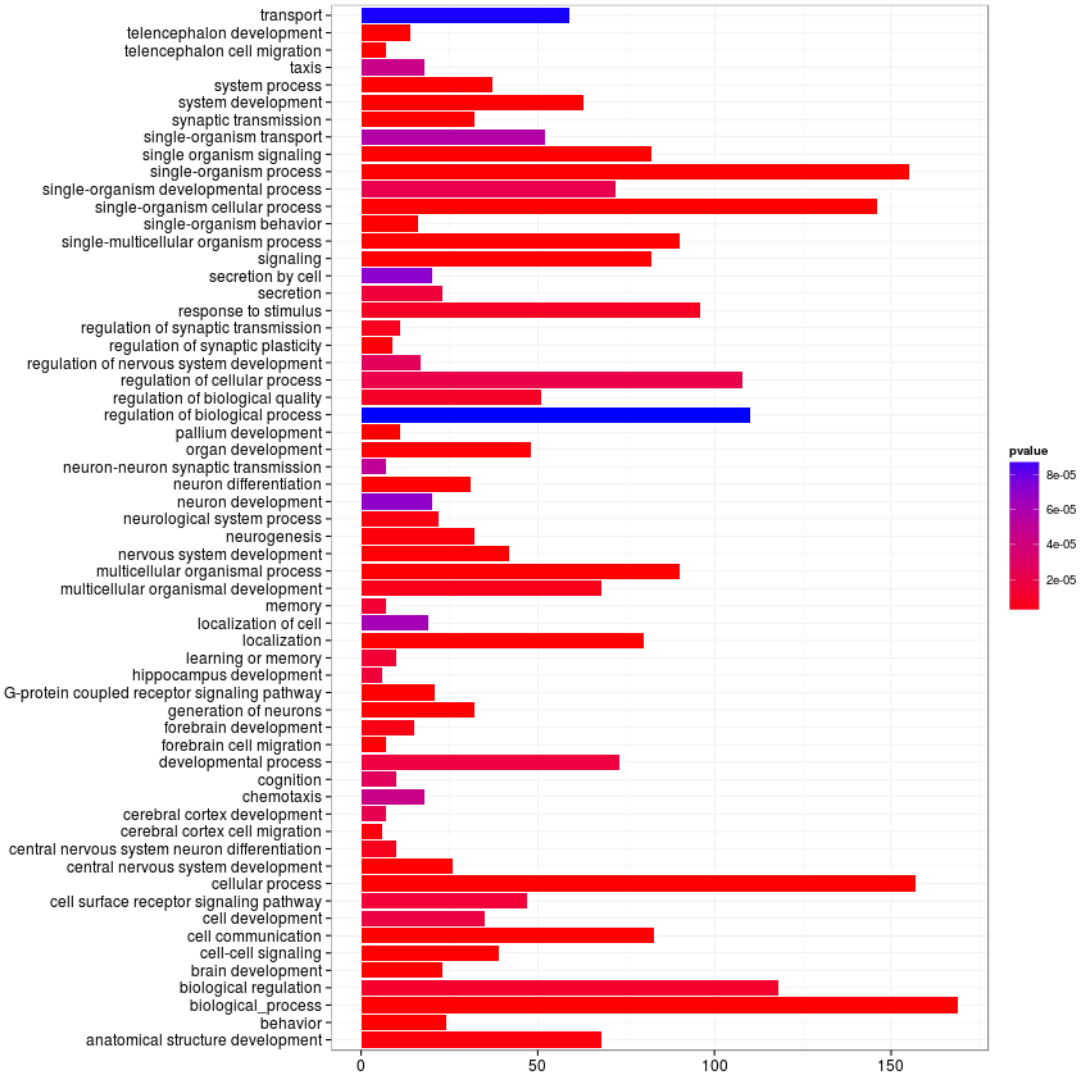


Figure 1.13: Most Enriched GO groups.

1.4.2 Disease Ontology

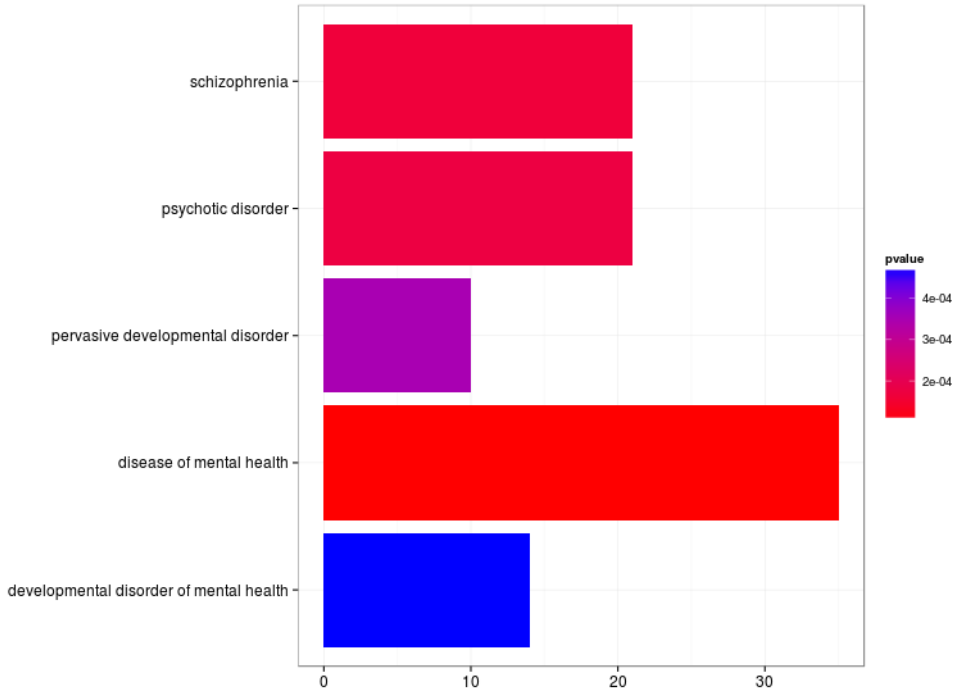


Figure 1.14: Most Enriched DO groups.

1.4.3 KEGG Pathways

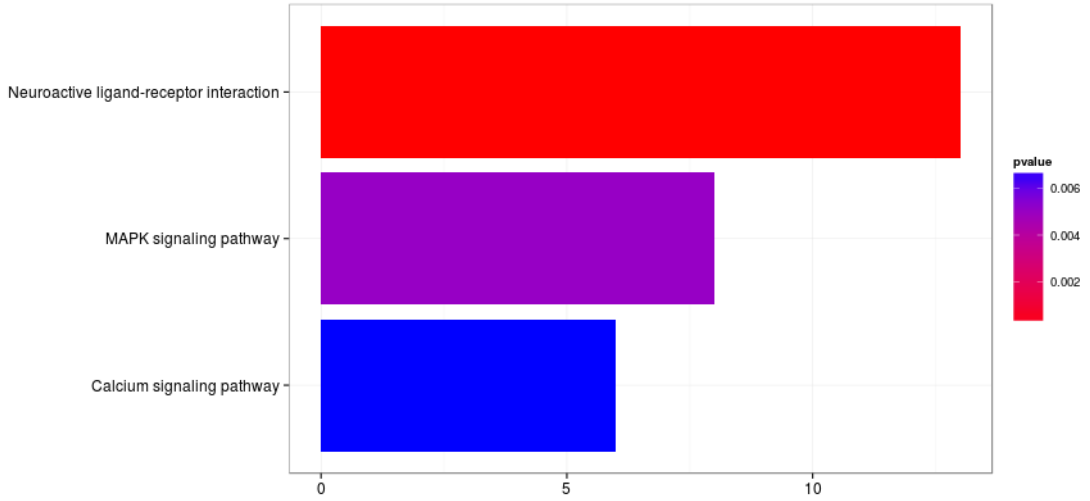


Figure 1.15: Most enriched KEGG pathways. All differentially expressed genes were used.

Neuroactive Ligand-Receptor Interaction

ILLUSTRATIVE FIGURE! FINAL REPORT WILL CONTAIN COMPLETE SET OF DATA.

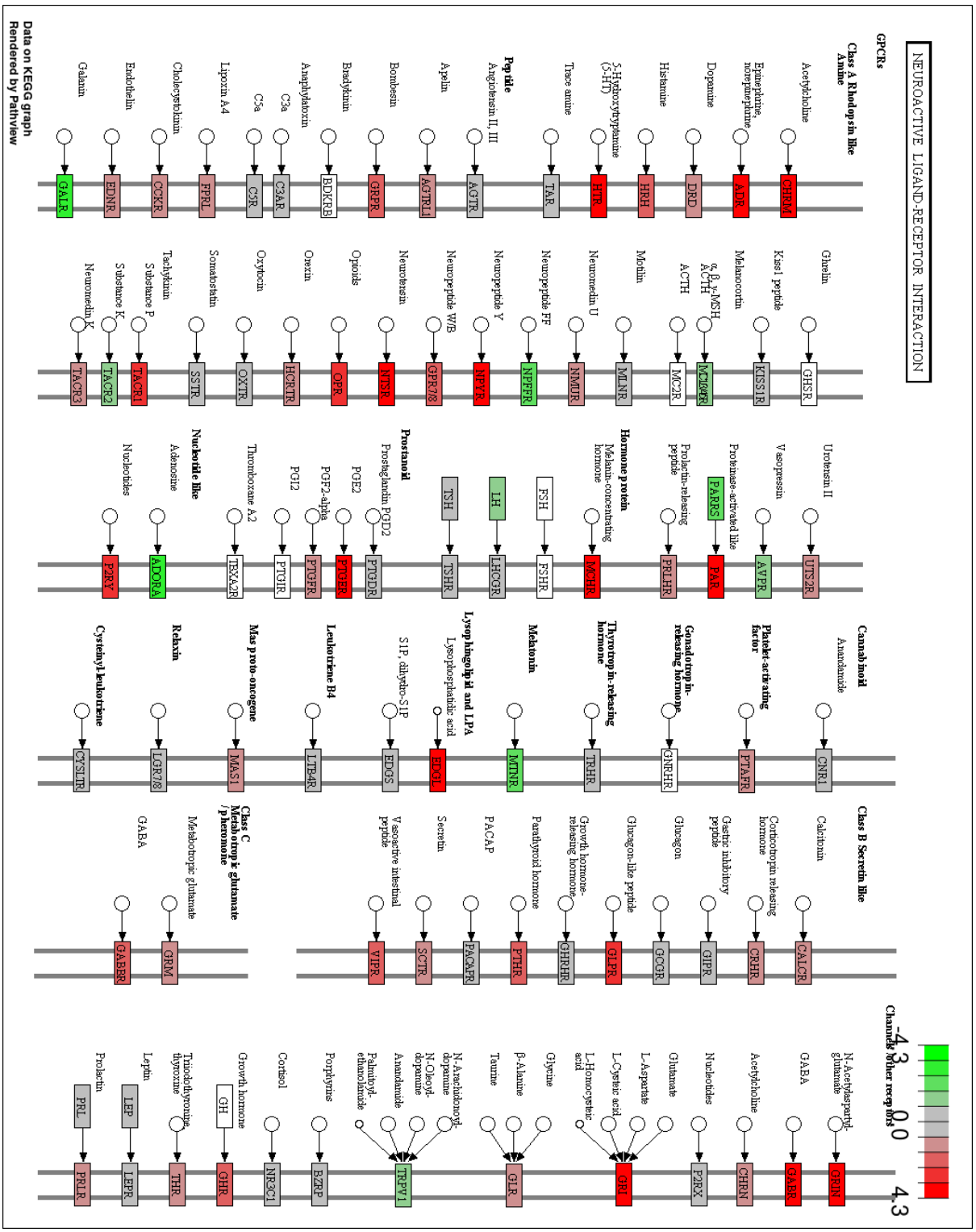


Figure 1.16: Representation of an enriched KEGG pathway.

MAPK Signalaing Pathway

ILLUSTRATIVE FIGURE! FINAL REPORT WILL CONTAIN COMPLETE SET OF DATA.

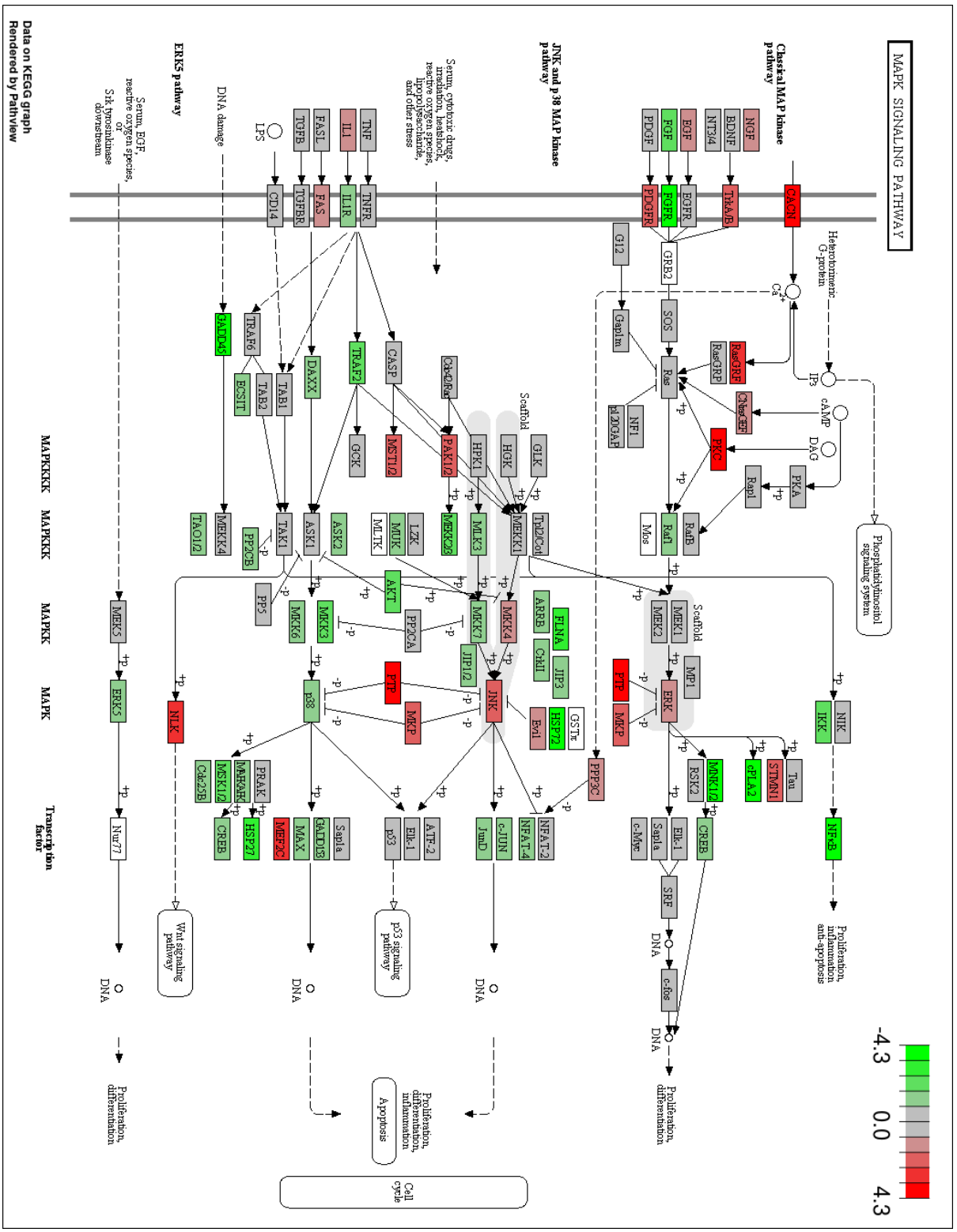


Figure 1.17: Representation of an enriched KEGG pathway.

Calcium Signaling Pathway

ILLUSTRATIVE FIGURE! FINAL REPORT WILL CONTAIN COMPLETE SET OF DATA.

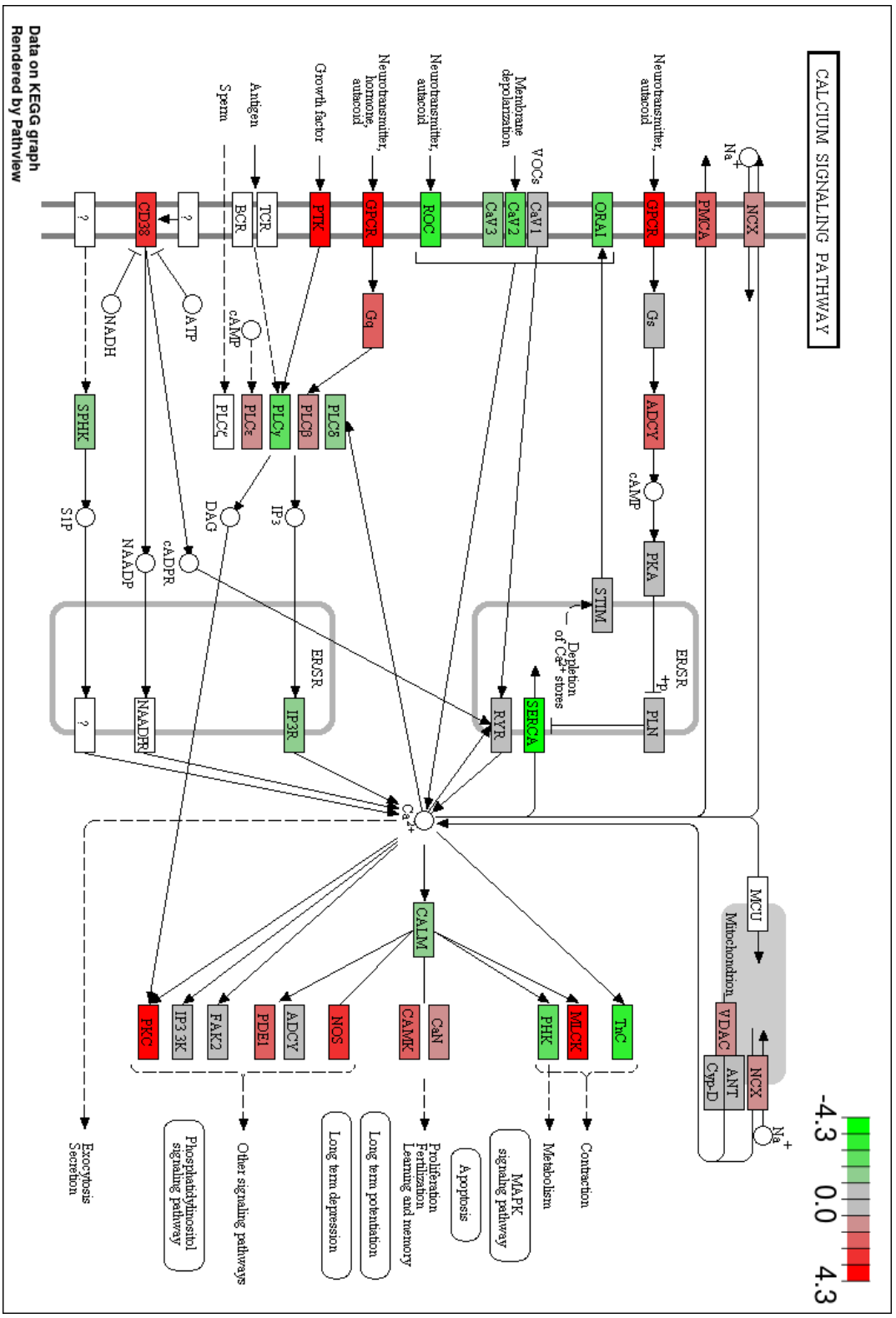


Figure 1.18: Representation of an enriched KEGG pathway.