doi 10.18699/bgrs2024-5.1-14

A principal components analysis and functional annotation of differentially expressed genes in brain regions of gray rats selected for tame or aggressive behavior

Chadaeva I.[1]*, Kozhemyakina R.[1], Shikhevich S[1], Bogomolov A.[1, 2], Kondratyuk E.[1, 3, 4], Oshchepkov D.[1, 2], Orlov Yu. L.[5, 6]*, Markel A.L.[1, 2]

[1] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia;*

[2] *Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia*

[3] *Siberian Federal Scientific Centre of Agro-BioTechnologies, Krasnoobsk, Russia*

[4] *Research Institute of Clinical and Experimental Lymphology, Novosibirsk, Russia*

[5] *Institute of Biodesign and Complex Systems Modeling, Moscow, Russia*

[6] *Agrarian and Technological Institute, Peoples' Friendship University of Russia, Moscow, Russia*

** ichadaeva@bionet.nsc.ru*

Key words: animal domestication; artificial selection; behavioral genetics; brain; differentially expressed gene; aggressive and tame rats; RNA-seq; principal component analysis; functional annotation of genes; gene network

Motivation and Aim: The aim of this work was to analyze and carry out functional annotation of differentially expressed genes (DEG) in four brain regions (the hypothalamus, the hippocampus, the periaqueductal gray matter (PAG), and the tegmental region of the midbrain (MTg) of gray rats selected for tame or aggressive behavior.

Methods and Algorithms: To analyze the DEGs, we used our previously obtained high-throughput RNA sequencing data from four brain regions of gray rats of the two strains: the hypothalamus [1], the hippocampus [2], the PAG [3], and the MTg [4]; these data have been deposited in the NCBI SRA database (ID = PRJNA668014). The total list of DEGs was subjected to principal component analysis: we performed this analysis using the PAST4.04 software package. Functional annotation of DEGs was carried out with the help of Web service DAVID Bioinformatics Resources. The construction of gene networks based on interactions between the DEGs under study was performed by means of web service STRING.

Results: On the basis of previously published transcriptomes of the hypothalamus, hippocampus, PAG, and MTg of tame and aggressive rats, we compiled a total list of DEGs, which consisted of 112 genes. The data on differential gene expression were then subjected to principal component analysis (PCA). It revealed a first principal component (here and further: PC1), which explained the largest proportion of variance of gene expression in the brain regions between the two strains (29 %) and a second principal component (here and further: PC2; 17 % of the variance). As for the PC1, the rats were subdivided into two groups on the basis of the 26 genes (out of 112 DEGs) that showed a significant correlation with this principal component (Fig. 1).

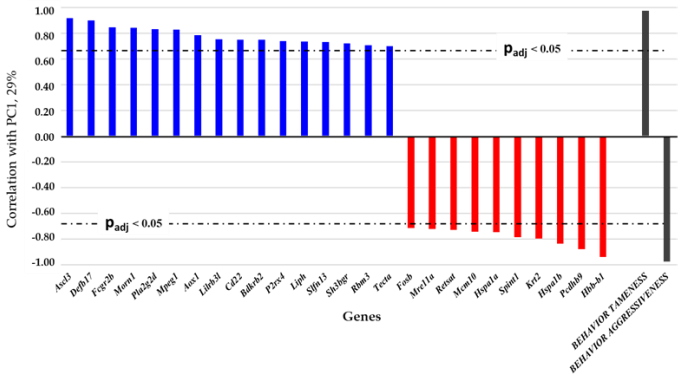

Fig. 1. A statistically significant linear correlation between expression levels of DEGs and the PC1. Genes whose expression levels are higher in tame rats are highlighted in blue; the genes whose expression is higher in aggressive rats are highlighted in red

The PC1 reflects gene expression changes that are caused by artificial selection based on the behavioral response of rats to humans. Accordingly, this component's positive coefficient of correlation with 16 genes points to a connection with the tame behavioral phenotype, and the negative correlation with 10 genes indicates an association with the rat phenotype of the aggressive response. The contribution of the PC2 accounted for only 17 % of the total variance of gene expression under our experimental conditions. The PC2 correlates with the expression levels of five genes in the hippocampus and reflects hippocampus-specific changes of gene expression in tame and aggressive rats.
Functional annotation of the DEG list was carried out using DAVID (Fig. 2).

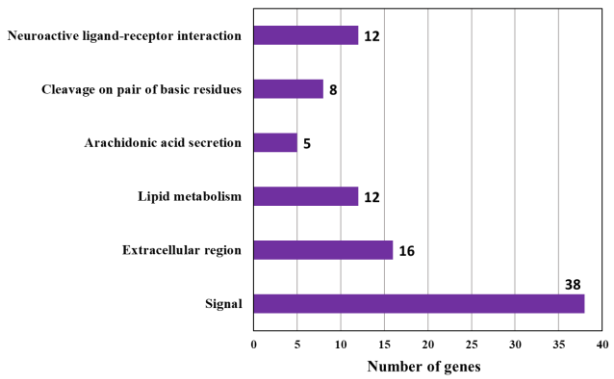

Fig. 2. The distribution of all DEGs among GO and KEGG terms in the brain regions of tame and aggressive rats according to DAVID. Biological processes are presented that are significantly enriched within the total list of DEGs

The DAVID analysis made it possible to determine which biological processes could lead to changes in tame and aggressive rats: "Neuroactive ligand-receptor interaction", "Cleavage on pair of basic residues", "Arachidonic acid secretion", "Lipid metabolism", "Extracellular region", and "Signal."
To identify possible relations between some DEGs a gene association network was built using STRING (Fig. 3). The gene network contains one large cluster, which includes such relevant genes as *Pcp2*, *Nmb*, *Fosb*, *Nr4a3*, *Shox2*, *Hspa1a*, *Hspa1b* and others: a total of 77 DEGs out of 112. In this total list, 33 DEGs were found to not participate in any interactions.

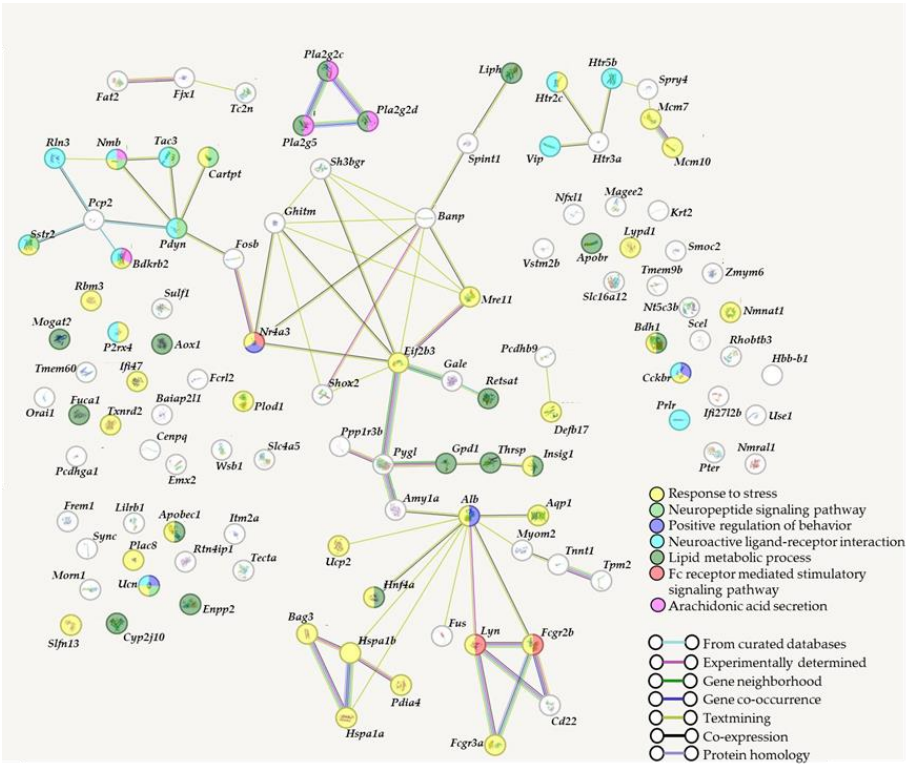

Fig. 3. The gene association network built by means of web service STRING on the basis of protein–protein interactions corresponding to some DEGs of tame and aggressive rats. Genes that are related to one of seven biological processes are highlighted in a color(s), and interactions between genes are indicated by lines

Thus, the DAVID and STRING functional enrichment analysis results complement each other and identify overrepresented terms that drive behavioral responses. All this together determines a multifactor system controlling the two different types of behavior towards human for which selection is carried out.

Conclusion: We studied the transcriptomes of two strains of gray rats selected for behavior towards human. We identified and analyzed genes that change their expression in rat brain regions as a result of the artificial selection for tame or aggressive behavior. The results obtained confirm the presence of a multifactorial system for controlling the behaviors studied.

Funding: The study is supported by Russian government projects FWNR-2022-0015.

References

1. Chadaeva I. et al. Domestication explains two-thirds of differential gene expression variance between domestic and wild animals; the remaining one-third reflects intraspecific and interspecific variation. *Animals.* 2021;11:2667. doi 10.3390/ani11092667
2. Oshchepkov D. et al. Stress reactivity, susceptibility to hypertension, and differential expression of genes in hypertensive compared to normotensive patients. *Int J Mol Sci.* 2022;23(5):2835. doi 10.3390/ijms23052835
3. Shikhevich S. et al. Differentially expressed genes and molecular susceptibility to human age-related diseases. *Int J Mol Sci* 2023;24(4):3996. doi 10.3390/ijms24043996
4. Oshchepkov D. et al. Transcription factors as important regulators of changes in behavior through domestication of gray rats: quantitative data from RNA sequencing. *Int J Mol Sci.* 2022;23(20):12269. doi 10.3390/ijms232012269