Differential alternative splicing by RNA-seq data in brain areas of laboratory rats with aggressive behavior

Anatoly O. Bragin, Vladimir N. Babenko, Yuriy L. Orlov

Institute of Cytology and Genetics SB RAS, Lavrentyeva 10, 630090, Novosibirsk, Russia,
`{ibragim,bob,orlov}@bionet.nsc.ru`

Arthur I. Dergilev, Irina V. Chadaeva, Arcady L. Markel

Novosibirsk State University, Novosibirsk, Russia,
`arturd1993@yandex.ru, {ichadaeva,markel}@bionet.nsc.ru`

Current work considers molecular mechanisms of complex behavior studies based on omics data using laboratory rat as a model. To study genetic component of aggressive behavior we used RNA-seq data related to aggressive behavior in rat (*Rattus norvegicus*) [1]. Two rat lines were selected over more than 30 years at the Institute of Cytology and Genetics SB RAS, Novosibirsk for studying of genetics factors determining aggression behavior. First rat line was selected for tame behavior towards human. Other rat line was selected for aggressiveness. We have analyzed three rat brain areas: hypothalamus; tegmental area and midbrain raphe nuclei (periaqueduct) [2]. These brain areas were selected due to previous studies indicating their role in behavior: ventral tegmental area contains the dopaminergic and opioidergic neurons and responsible for the positive reinforcement, midbrain raphe nuclei contain serotonergic neurons responsible for inhibitory control of aggression; hypothalamus is involved in the regulation of stress responsiveness. Using the software for gene splicing analysis we found the differentially spliced gene isoforms frequency profiles in our RNA-Seq data in aggressive and tame rats. We discuss role of the alternative splicing for neurospecific genes in behaviour patterns.

RNA-seq sequencing of rat brain areas samples was done using Illumina HiSeq 1500. RNA quality was tested on Bioanalyzer 2100 (Agilent) following the Illumina protocol. Sequencing depth consisted of at least 10 mol reads for each sample. The files obtained in "fastq" format were used for the reads mapping onto the RGSC Rnor_5.0\rn5 reference genome using Tophat2 aligner. Input reads were filtered and trimmed by Trimmomatic software [3]. Trimmomatic omits technical sequences and sequences with low quality from

input data. We use Tophat2 program [4] for mapping reads on the rodent reference genomes. Tophat2 is built on ultrafast, memory-efficient program Bowtie [5] and can identify splice junctions. Files with genome annotations were downloaded from UCSC genome browser [6]. The bam-files - received from Tophat output - were used for detection of differentially expressed genes and alternative splicing analysis in the samples. We have analysed rat and mouse gene expression by Cufflinks v2.0.2 programs. Cufflinks use bam files - with mapped reads – and gff files with reference genome annotation for calculation gene expression. Cufflinks also provide information about differently expressed genes between the samples. Gene expression levels were estimated in FPKM (fragments per kilobase per million reads) values. The detection of splice variants and differential splicing was performed using rMATS software [7].

Recent studies of gene expression in brain on animal models used similar models of laboratory animals [8, 9] and were discussed at previous MCCMB conference. Here we updated the study using series of control experiments of chromatography and RT-PCR. Overall we extended the analysis of difference in alternative splicing profiles across three brain areas and between aggressive and tame rats. To analyze alternative splicing events in the transcriptomics data MATS (multivariate analysis of transcript splicing) and rMATS (replicate MATS) shown to be effective tools [10]. We implemented a Python-based application to tackle rMATs annotation. The rMATs output lacks the information on the short isoforms spectra of particular exon. We implemented the add-on that, based on 3 subsequent exons locations used in alternative exons identification, retrieves from RefGene database short isoforms IDs lacking the exon. While it's not possible to relate exon skipping event to the particular full length isoform due to the short reads length, we can still speculate on tissue specificity of the exon skipping events.

Our study continues recent work on group comparison of RNA-seq profiles in tame and aggressive laboratory animals [1,2]. The alignment of differentially expressed genes with a behavioral phenotype can be further examined using a variety of secondary analyses, e.g., examining if the genes cluster within known gene ontology categories or are part of a known protein-protein interaction network [11]. After constructing differentially expressed gene lists

we used set of tools for coexpression analysis to found characteristic features of gene network related to aggressive behavior. We used such tools as ANDVisio and ANDSystem [12] that allow reconstruction of gene networks controlling biochemical, molecular-genetic, and physiological processes. Note that the scientific literature on genetics of aggressive behavior suggests multi-loci determination of aggressive behavior without any major gene reported.

Next, deeper level of molecular mechanisms controlling brain function is landscape of gene splicing events. A study underlines Ca2+ intrusion into cells via NMDA under leading to chemokine receptor CXCR4 mediated neurotoxicity [13]. The blockade of l-type Ca2+ channels and NMDAR both prevent CXCR4-mediated toxicity of CXCL12 in cerebrocortical neurons. Thus, Grin1 gene is relevant to the stress induced response in neurons. The precise homeostasis of glutamate concentrations necessary for the proper neurotransmission activity may underlie the chronic stress states such as aggression and depression [14].

The 8 isoforms of Grin1 are generated by 3 alternative exons: 5, 21 and 22 [15]. We performed PC (Principal Components) decomposition based on isoforms profile frequencies in 12 samples [2]. Low glutamate levels activate adaptive stress responses that include proteins that protect neurons against more severe stress. Conversely, abnormally high levels of glutamate, resulting from increased release and/or decreased removal, cause neuronal atrophy and depression. The deregulation of the glutamatergic transmission in depression could be underlined by several factors including a decreased inhibition (γ-aminobutyric acid or serotonin) or an increased excitation (primarily within the glutamatergic system). Experimental evidence shows that the activation of N-methyl-D-aspartate receptor (NMDA) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors (AMPAR) can exert two opposite effects on neurogenesis and neuron survival depending on the synaptic or extrasynaptic concentration.

Importantly, the shortest isoform, NM_001270610, features aggressive hypothalamus brain regions, and differs statistically significantly in frequency in tame rats (FDR<8E-6). The longest isoform NM_017010, also proved to be significantly deviated (FDR<1E-2) in is characteristic of tame rat hypothalamus [16]. From previous studies [15] it was elucidated the exon 21 skipping mediates neuron depolarization. It could be speculated that synapse activity

is higher in aggressive rats. The differential alternative splicing revealed that many synapse related genes have statistically significant deviation depending on brain region and aggressive/tame status in rat. It is a novel phenomenon of the transcriptome data related to aggressive behavior [16].

Previous studies reported quite valuable information on the gene expression alteration both in the inherited and the 'induced/acquired' models of aggression [17- Warren et al., 2013]. We reconfirmed the brain specific deviations of gene expression in our data, in particular in synapse specific genes. The systemic research of gene expression in the brain cells using comprehensive experimental approaches is required for interdisciplinary neurobiological studies, and future research has to be conducted using new transcriptome data and software applications for gene expression analysis [1,2].

References

1. N.N. Kudryavtseva, A.L. Markel, Y.L. Orlov (2014) Aggressive behavior: genetic and physiological mechanisms, *Vavilov journal of genetics and breeding,* **18**(4/3):1133–1155. (In Russian)

2. V.N. Babenko et al. (2016) Analysis of differential gene expression by RNA-seq data in brain areas of laboratory animals, *Journal of Integrative bioinformatics*, **13**(4):292.

3. M. Bolger, M. Lohse, B. Usadel (2014) Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, **30**(15):2114-20.

4. C. Trapnell et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, *Nat Protoc.*, **7**(3):562-78.

5. B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, *Genome biology*, **10**(3):R25.

6. D. Karolchik et al. (2003) The UCSC genome browser database, *Nucleic acids research*, **31**(1):51-54.

7. S. Shen et al. (2014) rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data, *Proc. Nat. Acad. Sci. USA,* **111**(51): E5593-601.

8. R. Hitzemann et al. (2013) Genes, behavior and next-generation RNA sequencing, *Genes Brain Behav.*, **12**(1):1-12.

9. H.O. Heyne et al. (2014) Genetic influences on brain gene expression in rats selected for tameness and aggression, *Genetics,* **198**(3):1277-90.

10. S. Shen et al. (2012) MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data, *Nucleic Acids Res*., **40**(8):e61.

11. G. Feng et al. (2012) Using the bioconductor GeneAnswers package to interpret gene lists, *Methods Mol Biol.,* **802**:101-112.

12. V.A. Ivanisenko et al. (2015) ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology, *BMC Syst Biol*., **9**(Suppl 2):S2.

13. A.B. Sanchez et al. (2016) CXCL12-induced neurotoxicity critically depends on NMDA receptor-gated and L-type Ca(2+) channels upstream of p38 MAPK, *J Neuroinflammation*, **13**(1):252.

14. A. Rubio-Casillas, A. Fernández-Guasti (2016) The dose makes the poison: from glutamate-mediated neurogenesis to neuronal atrophy and depression, *Rev Neurosci.*, **27**(6):599-622.

15. J.A. Lee et al. (2007) Depolarization and CaM kinase IV modulate NMDA receptor splicing through two essential RNA elements, *PLoS Biol*. **5**(2):e40.

16. V.N. Babenko et al. (2017) Differential alternative splicing in brain tissues of rats selected by aggressive behavior, *Mol Biol (Mosk)*, 2017. (In press) (In Russian).

17. B.L. Warren et al. (2013) Neurobiological sequelae of witnessing stressful events in adult mice, *Biol Psychiatry,* **73**(1):7-14.