GENOME-WIDE MAPPING OF DNA METHYLATION VARIANTS AFFECTING GENE EXPRESSION LEVELS IN GLIOMAS WITH RESPECT TO THEIR GRADE AND IDH GENE MUTATION STATUS

Michal J Dabrowski[1], Agata Dziedzic[1], Rafał Guzik[2], Michal Draminski[1], Bartosz Wojtas[2], Karolina Stępniak[2], Bartłomiej Gielniewski[2], Jacek Koronacki[1], Bozena Kaminska[2]

[1]Institute of Computer Science, Computational Biology Lab, Warsaw, Poland, [2]Nencki Institute of Experimental Biology, Laboratory of Molecular Neurobiology, Warsaw, Poland

Here we report on differentially methylated promoters in gliomas of different histopathological WHO grades and *IDH* gene mutation status. To obtain methylomes for 20 samples we used SeqCap Epi CpGiant Methylation panel and performed bisulphite conversion followed by Illumina NGS sequencing. For these samples RNA-seq was also performed. In order to analyze methylomes we developed CytoMeth tool based on the Roche pipeline. Our tool compiles a set of open source software enabling fast and transparent rough data processing. The sum of unique CpG sites across all samples was above 5 mln, while the common part constituted ~ 300,000. We analysed genomic regions +/- 2kB from TSS that did not differ in the number of cytosines between samples subgroups. For these cytosines in each of over 34,000 promoters we computed: max, mean, SD, median of B-values, as well as number of CpG sites, % of highly methylated, medium and low methylated cytosines. To select genes whose promoter differs in any characteristic between pilocytic astrocytoma (PA) and other glioma subtypes (diffuse astrocytoma and glioblastoma, hereafter HGG), we used two approaches: (i) Wilcoxon test with false discovery rate correction and (ii) MCFS-ID algorithm which performs non-linear range specific feature selection and discovers interdependencies between features. As biologically relevant and not age-specific, we selected genes that differed in expression between PA and HGG and at the same time did not differ between PA and adult normal brain samples (NB) and did differ between HGG and NB. That pipeline was applied for cytosines intersected with promoters with- and without-respect to the strand, which returned two sets: StSp, NotStSp, respectively. For these sets we performed unsupervised text clustering based on gene descriptions from KEGG and REACTOME and verified whether genes were up- or down-regulated.

There were 70% of genes common for StSp and NotStSp, in consequence clusters corresponding to each other were characterized by similar key-words but only in the StSp set there was gene cluster associated with ABC transporters. Homogenous clusters, where all genes were up-regulated in HGG, were described as follows: [1] homeobox, hoxc, morphogenesis; [2] development, embryonic, differentiation; [3] mitotic, cell cycle, repair, G protein, replication; [4] transcription. Moreover, clusters with majority of up-regulated genes in PA were described by: [5] synaptic, adhesion, membrane, receptor, neurotransmitter, protein tyrosine phosphatases; [6] positive/negative regulation, apoptotic process. It is worth to mention that we also discovered a group of antisense genes i.e. *PLCE1-AS1* whose promoter was highly methylated in PA indicating a multistage regulatory mechanism of its protein coding gene expression.

Supported by the NCN grant 2015/16/W/NZ2/00314.