How to analysis methylation sequencing data using DMRFusion?

Maryam Yassi, Ehsan Shams Davodly, Afsaneh Mojtabanezhad Shariatpanahi, Mehdi Heidari, MohammadHossein Moattar, Mahdieh Dayyani, Mohammad Amin Kerachian
Cancer Genetics Research Unit, Reza Radiotherapy and Oncology Center, Mashhad, Iran

Abstract: DNA methylation is an important epigenetic modification involved in many number of biological processes and diseases. It affects the gene expression pattern and causes changes in the epigenetic function. Computational analysis of the methylation sequencing for detecting differentially methylated regions (DMRs) can assist the researchers explore the underlying reasons of methylation. Herein, DMRFusion is presented as a useful tool for comprehensive DNA methylation analysis of DMRs on methylation sequencing data. It is designed base on the integration of several ranking methods; *Information gain*, *Between versus within Class scatter ratio*, *Fisher ratio*, *Z-score* and *Welch's t-test*. These filtering methods rank CpG sites in hyper- and hypomethylation DMRs through several measurements, such as distance and information theory. Hence, DMRFusion method could identify the single CpG nucleotide with the maximum difference in prominent DMRs. The experimental of the proposed approach on reduced representation bisulfite sequencing (RRBS) data in chronic lymphocytic leukemia cancer displayed 30 nominated regions and CpG sites with a maximum methylation difference detected in the hypermethylation DMRs. We realized that DMRFusion is able to process methylation sequencing data in an efficient and accurate manner and to provide annotation and visualization for DMRs with high fold difference score (*p* value and FDR less than 0.05 and type 1 error: 0.04). Moreover, we applied unsupervised clustering with performance more than 97% for distinguishing between cancer and control groups based on candidate CpG sites

Keywords: DNA methylation*;* Epigenetic; Differentially methylated regions; Reduced representation bisulfite sequencing; Filter method