



GeneDMRs: an R package for Gene-based Differentially Methylated Regions analysis

Wang, Xiao; Kadarmideen, Haja

Published in:

Proceedings of 27th Conference on Intelligent Systems for Molecular Biology and 18th European Conference on Computational Biology

Publication date:

2019

Document Version

Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Wang, X., & Kadarmideen, H. (2019). GeneDMRs: an R package for Gene-based Differentially Methylated Regions analysis. In *Proceedings of 27th Conference on Intelligent Systems for Molecular Biology and 18th European Conference on Computational Biology*

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

GeneDMRs: an R package for Gene-based Differentially Methylated Regions analysis

Xiao Wang, Haja N. Kadarmideen *

Quantitative Genomics, Bioinformatics and Computational Biology Group, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark

Calculating methylation levels using bisulfite sequencing datasets like RRBS within a single gene could help investigate and identify the Gene-based Differentially Methylated Regions (GeneDMRs). Such GeneDMRs are better than single differentially methylated cytosines (DMCs), as it provides an overall gene methylation profiles than single nucleotide level profiles, particularly when the testing known genes prone to epigenetics.

The mean methylation of a gene of one treatment group is defined as:

$$\sum_1^n \frac{\sum_1^m MR_{ij}}{\sum_1^m TR_{ij}} * W_{ij} \text{ and } W_{ij} = \frac{\sum_1^m TR_{ij}}{\sum_1^n \sum_1^m TR_{ij}},$$

where MR_{ij} and TR_{ij} are methylated and total reads number of the involved CpG/DMC j at a given gene of individual i , n is the total individual number of one treatment group, m is total number of CpG/DMC involved in this gene and W_{ij} is the weight of reads of the involved CpG/DMC j of individual i .

GeneDMRs is defined by the comparisons across different treatment groups following logistic regression model:

$$\ln \left(\frac{\pi_i}{1-\pi_i} \right) = u + \beta T_i + e,$$

where π_i is the mean methylation at a given gene, u is the intercept, T_i is the treatment indicator and e is the residual.

The GeneDMRs is a user-friendly R package that can easily output the required results and figures, such as mean methylation levels of all genes and CpG islands or a specific gene or promoter/exon/intron regions with the realized boxplot, heat map and correlation matrix, as well GO terms/pathways in hyper-/hypo-methylated categories by combining with upregulated/downregulated differentially expressed genes. These developed algorithms in the GeneDMRs package have been applied to the real porcine DNA methylation (RRBS) data and example results are given in the Figures 1 and Figure 2.

As more features for GeneDMRs are being updated, the current offline GeneDMRs package is available from xiwa@dtu.dk and hajak@dtu.dk.

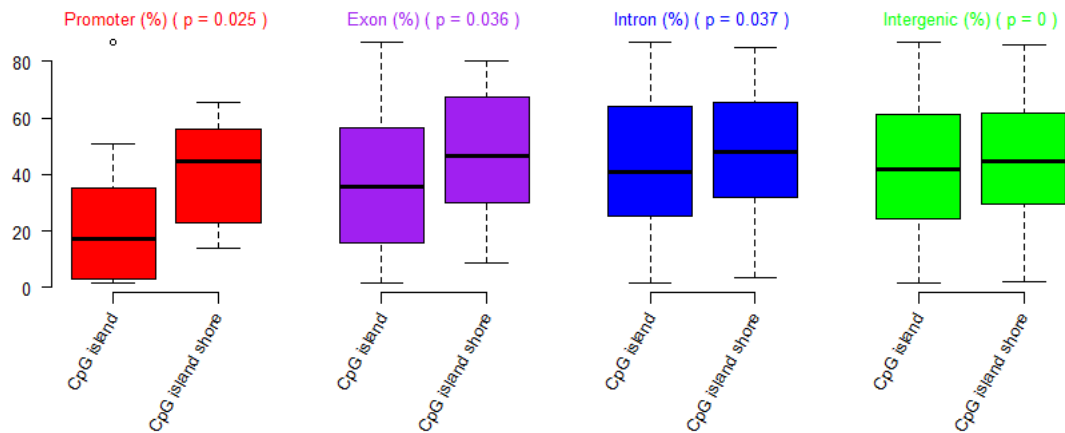


Figure 1. Comparison of methylation levels between CpG islands and CpG island shores at different genic feature with Student's *t*-test.

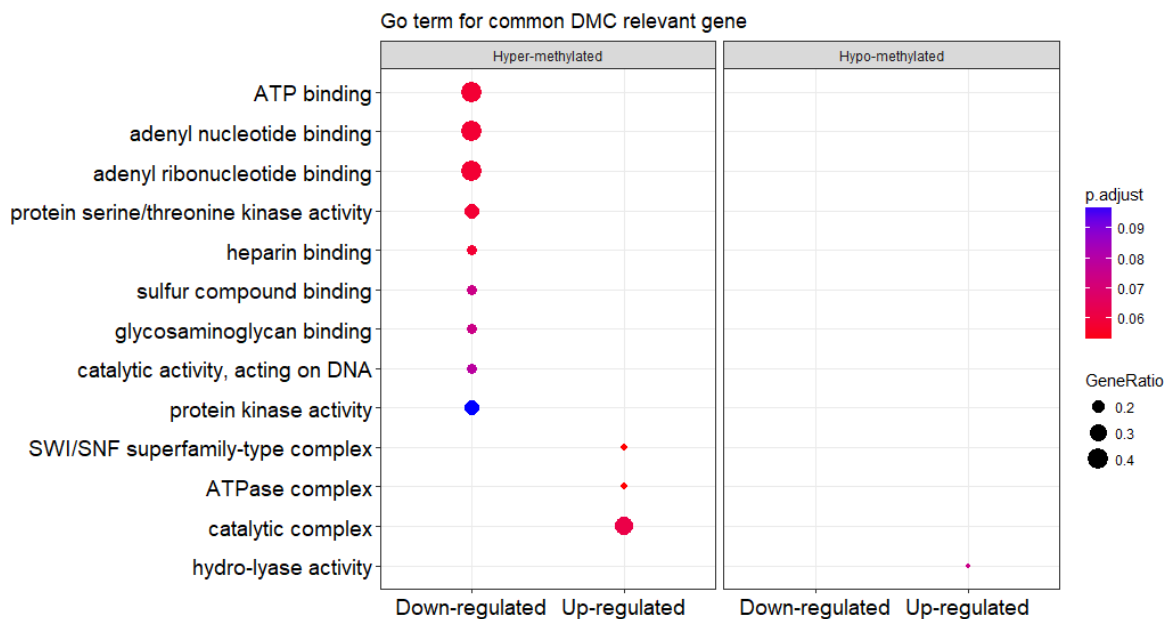


Figure 2. Significant GO terms ($p.adjust < 0.1$) of the up-regulated and down-regulated differentially expressed genes in the hyper-methylated and hypo-methylated categories by combining with R package *clusterProfiler* (Yu *et al.* 2012).

Yu G., Wang L.G., Han Y. & He Q.Y. (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16**, 284-7.