FOLDGO - THE SOFTWARE PACKAGE FOR FOLD-CHANGE-SPECIFIC GO TERMS IDENTIFICATION IN TRANSCRIPTOME DATA

Wiebe D.S., Mukhin A.M., Omelyanchuk N.A., Mironova V.V.

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia Novosibirsk State University, Novosibirsk, Russia e-mail: daniil.wiebe@gmail.com

Gene expression profiling technologies have revolutionized biological and medical research. Scientists extensively generate transcriptome data to study genome response to external factors or internal conditions. Thousands of datasets are available online, e.g. in GEO (Edgar et al., 2002). The typical scenario of transcriptome data analysis is the identification of differentially expressed genes (DEGs), followed by a functional enrichment analysis of the DEGs set using Gene Ontology (GO, Ashburner et al., 2000). As a result, biological processes, molecular functions, and cellular components that are significantly enriched in up- or down-regulated genes are identified. However, in this analysis, one reduces the distribution of gene expression changes to a binary variable (regulated and not regulated by a factor under investigation). Recently, we showed that functional annotation methods can be extended to identify the functional gene groups that respond to an external factor or an internal condition not only unidirectionally, but also with similar fold-changes (fold-change-specific groups). Fold-change-specific functional gene groups have been discovered in the transcriptome induced by the phytohormone auxin in *Arabidopsis thaliana* roots (Omelyanchuk et al., 2017). We extended this approach to human transcriptome using data from the experiment on AR-v7 expression in the LNCaP cell line (Cottard et al., 2017). As a result, we found functional gene groups responsible for processes crucial for cancer development which were overlooked by classical approaches such as Singular Enrichment Analysis (SEA). Identification of these fold-change-specific functional gene groups may help to clarify the coordination principles in complex molecular-genetic processes, but no ready-to-use tools exist for this task. Here we present the developed method as a FoldGO R package (https://bioconductor.org/packages/release/bioc/html/FoldGO.html) and web-server (https://webfsgor.sysbio.cytogen.ru/), which allow the user to apply fold-change-specific functional enrichment analysis to a transcriptomic dataset of interest. The work was supported by the RFBR, project № 18-34-00871.

References:

1. Ashburner, M., et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics, 2000, 25(1), 25–29.

2. Cottard F, et al. Dual effects of constitutively active androgen receptor and full-length androgen receptor for N-cadherin regulation in prostate cancer. Oncotarget. 2017; 8:72008-72020.

3. Edgar, R., Domrachev, M., & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research, 2002, 30(1), 207–210.

4. Omelyanchuk N.A., et al. Auxin regulates functional gene groups in a fold-specific manner in Arabidopsis root // Nat Sci Rep – 2017 - N7(1) - p2489