RexPo 22
THE 1st INTERNATIONAL CONFERENCE
ON DRUG REPURPOSING
MAASTRICHT 2ND-3RD SEP 2022

DESMOND 2.0: Identification of differentially expressed biclusters for unsupervised patient stratification

Olga Zolotareva[1,2], Olga I. Isaeva[3], Michael Hartung[1], Andreas Maier[1], Fernando M. Delgado-Chaves[1,4], Katharina Kaufmann[1], Alexey Savchik[5], Zoe Chervontseva[5], Niklas Probul[1], Alexandra Abisheva[6], Evgenia Zotova[6], Olga Tsoy[1], David B. Blumenthal[7], Martin Ester[8,9], Jan Baumbach[1,10]

[1] *Institute for Computational Systems Biology, University of Hamburg, Hamburg (Germany)*
[2] *Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University of Munich, Munich (Germany)*
[3] *Division of Molecular Oncology & Immunology and Division of Tumor Biology & Immunology, The Netherlands Cancer Institute, Amsterdam (the Netherlands)*
[4] *Computer Science Division, Higher Polytechnic School, Pablo de Olavide University, Seville (Spain)*
[5] *A.A.Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow (Russia)*
[6] *Altius Institute for Biomedical Sciences, Seattle (USA)*
[7] *Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen (Germany)*
[8] *School of Computing Science, Simon Fraser University, Burnaby (Canada)*
[9] *Vancouver Prostate Centre, Vancouver (Canada)*
[10] *Department of Mathematics and Computer Science, University of Southern Denmark, Odense (Denmark)*

Abstract

Unsupervised patient stratification based on omics data is traditionally approached by clustering methods which may be inefficient for datasets with multiple patterns overlapping in rows and columns. Biclustering methods that are searching for submatrices with a specific pattern in a two-dimensional sample-gene matrix represent a promising alternative to conventional clustering [1]. However, in practice, the existing biclustering methods show a limited ability to robustly recover known PAM50 breast cancer subtypes [2]. This motivated us to develop DESMOND, a novel method for the identification of differentially expressed biclusters that uses interaction networks as constraints to improve the robustness of the biclustering results [3]. We applied DESMOND to two independent breast cancer cohorts (TCGA-BRCA [4] and METABRIC [5]) and confirmed that it indeed identified more robust biclusters than other methods. However, found biclusters were small in terms of genes, possibly due to incompleteness of the input network.

Here we present DESMOND 2.0 (https://github.com/ozolotareva/DESMOND2), an updated version of DESMOND which makes two significant advancements. First, it does not require the network and clusters individual genes instead of gene pairs. Second, it provides the choice of two binarization and three clustering methods for the identification of gene and sample sets. These modifications greatly improve the tool's runtime and help to find larger biclusters in terms of genes and to recover known breast cancer subtypes more precisely than its baselines. Moreover, besides PAM50 subtypes, DESMOND identifies a rare neuroendocrine subtype, which prevalence is around 5% in both cohorts [6].

Keywords

patient stratification, molecular subtypes, phenotype heterogeneity, biclustering, differential expression

References

1. Xie J et al. (2019) It is time to apply biclustering: a comprehensive review of biclustering applications in biological and biomedical data. Brief Bioinform. 20(4):1449-1464.
2. Perou CM et al. (2000) Molecular portraits of human breast tumours. Nature 406(6797):747–75.
3. Zolotareva O et al. (2021) Identification of differentially expressed gene modules in heterogeneous diseases. Bioinformatics 37:1691-8.
4. Liu J et al. (2018) An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell, 173, 400–416.E11.
5. Pereira B et al. (2016) The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nat. Commun., 7, 11908.
6. Sinn HP, and Kreipe H. (2013) A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition. Breast Care (Basel) 8(2):149-54.