A single-subject method to detect pathways enriched with alternatively spliced genes

A. Grant Schissler
Department of Mathematics & Statistics
University of Nevada, Reno (United States)
aschissler@unr.edu

Abstract

RNA-Sequencing data offers an opportunity to enable precision medicine, but most methods rely on gene expression alone. To date, no methodology exists to identify and interpret alternative splicing patterns within a patient. This study develops methodology and conducts computational experiments to test the hypothesis that pathway aggregation of subject-specific alternatively spliced genes can inform upon disease mechanisms and predict survival. We propose the N-of-1-*pathways*[1, 2, 3] Alternatively Spliced (N1PAS) method that takes an individual patient's paired-sample RNA-Seq isoform expression data (e.g., tumor vs non-tumor, before-treatment vs during-therapy) and genetic pathway annotations as inputs. N1PAS quantifies the degree of alternative splicing via Hellinger[4] distances followed by two-stage clustering to determine pathway enrichment. We provide a clinically relevant 'odds ratio' along with statistical significance to quantify pathway enrichment. We validate our method in clinical samples and find that our method selects relevant pathways. Importantly, our studies also unveil highly heterogeneous single-subject alternative splicing patterns that cohort-based approaches may overlook. Finally, we aggregate our patient-specific results to predict cancer survival, translating transcriptome data into clinically actionable information.

Keywords

Single-subject design • Hellinger distance • RNA-sequencing
Bioinformatics • High-dimensional data • Mixture modeling • Clustering

References

[1] Vincent Gardeux, Ikbel Achour, Jianrong Li, Mark Maienschein-Cline, Haiquan Li, Lorenzo Pesce, Gurunadh Parinandi, Neil Bahroos, Robert Winn, Ian Foster, Joe G N Garcia, and Yves a Lussier. 'N-of-1-pathways' unveils personal deregulated mechanisms from a single pair of RNA-Seq samples: towards precision medicine. *Journal of the American Medical Informatics Association*, 21(6):1015–1025, 2014.

[2] A. Grant Schissler, Vincent Gardeux, Qike Li, Ikbel Achour, Haiquan Li, Walter W. Piegorsch, and Yves A. Lussier. Dynamic changes of RNA-sequencing expression for precision medicine: N-of-1-pathways Mahalanobis distance within pathways of single subjects predicts breast cancer survival. *Bioinformatics*, 31(12):i293–i302, 2015.

[3] A Grant Schissler, Walter W Piegorsch, and Yves A Lussier. Testing for differentially expressed genetic pathways with single-subject N-of-1 data in the presence of inter-gene correlation. *Statistical Methods in Medical Research*, 27(12):3797–3813, 2018.

[4] Marla Johnson and Elizabeth Purdom. Clustering of mRNA-Seq data for detection of alternative splicing patterns. *Biostatistics*, 18(2):295–307, 2017.