satuRn: Scalable analysis of differential transcript usage for bulk and single-cell RNA-sequencing applications

This manuscript (permalink) was automatically generated from jgilis/satuRn_md@ad08f20 on August 30, 2022.

Authors

- **Jeroen Gilis**
  0000-0001-8415-0943 · jgilis · GilisJeroen
  Department of Applied Mathematics, Computer science and Statistics, Ghent University · Funded by FWO

- **Kristoffer Vitting-Seerup**
  XXXX-XXXX-XXXX-XXXX · kvittingseerup
  Department of Something, University of Whatever; Department of Whatever, University of Something

- **Koen Van den Berge**
  XXXX-XXXX-XXXX-XXXX · koenvandenberge
  Department of Something, University of Whatever; Department of Whatever, University of Something

- **Lieven Clement**
  XXXX-XXXX-XXXX-XXXX · statOmics
  Department of Something, University of Whatever; Department of Whatever, University of Something

✉ — Correspondence possible via GitHub Issues

Abstract

Alternative splicing produces multiple functional transcripts from a single gene. Dysregulation of splicing is known to be associated with disease and as a hallmark of cancer. Existing tools for differential transcript usage (DTU) analysis either lack in performance, cannot account for complex experimental designs or do not scale to massive single-cell transcriptome sequencing (scRNA-seq) datasets. We introduce satuRn, a fast and flexible quasi-binomial generalized linear modelling framework that is on par with the best performing DTU methods from the bulk RNA-seq realm, while providing good false discovery rate control, addressing complex experimental designs, and scaling to scRNA-seq applications.

Introduction

Studying differential expression (DE) is one of the key tasks in the downstream analysis of RNA-seq data. Typically, DE analyses identify expression changes on the gene level. However, the widespread adoption of expression quantification through pseudo-alignment,1,2 which enables fast and accurate quantification of expression at the transcript level, has effectively paved the way for transcript-level analyses. Here, we specifically address differential transcript usage (DTU) analysis, one type of transcript-level analysis that studies the change in relative usage of transcripts/isoforms within the same gene. DTU analysis holds great potential: previous research has shown that most multi-exon human genes are subject to alternative splicing and can thus produce a variety of functionally different isoforms from the same genomic locus.3–5 The dysregulation of this splicing process has been reported extensively as a cause for disease,6–9 including several neurological diseases such as frontotemporal dementia, Parkinsonism and spinal muscular atrophy, and is a well- known hallmark of cancer.10

References