

ARH-seq: identification of differential splicing in RNA-seq data.

Rasche A¹, Lienhard M², Yaspo ML², Lehrach H², Herwig R¹.

Author information:

1 Max-Planck-Institute for Molecular Genetics, Department of Vertebrate Genomics, Ihnestrasse 63-73, 14195 Berlin, Germany herwig@molgen.mpg.de.

2 Max-Planck-Institute for Molecular Genetics, Department of Vertebrate Genomics, Ihnestrasse 63-73, 14195 Berlin, Germany.

Abstract

The computational prediction of alternative splicing from high-throughput sequencing data is inherently difficult and necessitates robust statistical measures because the differential splicing signal is overlaid by influencing factors such as gene expression differences and simultaneous expression of multiple isoforms amongst others. In this work we describe ARH-seq, a discovery tool for differential splicing in case-control studies that is based on the information-theoretic concept of entropy. ARH-seq works on high-throughput sequencing data and is an extension of the ARH method that was originally developed for exon microarrays. We show that the method has inherent features, such as independence of transcript exon number and independence of differential expression, what makes it particularly suited for detecting alternative splicing events from sequencing data. In order to test and validate our workflow we challenged it with publicly available sequencing data derived from human tissues and conducted a comparison with eight alternative computational methods. In order to judge the performance of the different methods we constructed a benchmark data set of true positive splicing events across different tissues agglomerated from public databases and show that ARH-seq is an accurate, computationally fast and high-performing method for detecting differential splicing events.

© The Author(s) 2014. Published by Oxford University Press on behalf of Nucleic Acids Research.

[PMID: 24920826](#) [PubMed - indexed for MEDLINE]