

arXiv.org > stat > arXiv:1106.3211

Search or Article-id

All papers 🚽 Go!

(Help | Advanced search)

Download:

- PDF
- PostScript
- Other formats

Current browse context: stat.ME

< prev | next >

new | recent | 1106

Change to browse by:

q-bio q-bio.GN stat

References & Citations

NASA ADS

Bookmark(what is this?)
Image: Image:

Statistics > Methodology

Statistical Modeling of RNA-Seq Data

Julia Salzman, Hui Jiang, Wing Hung Wong

(Submitted on 16 Jun 2011)

Recently, ultra high-throughput sequencing of RNA (RNA-Seq) has been developed as an approach for analysis of gene expression. By obtaining tens or even hundreds of millions of reads of transcribed sequences, an RNA-Seq experiment can offer a comprehensive survey of the population of genes (transcripts) in any sample of interest. This paper introduces a statistical model for estimating isoform abundance from RNA-Seq data and is flexible enough to accommodate both single end and paired end RNA-Seq data and sampling bias along the length of the transcript. Based on the derivation of minimal sufficient statistics for the model, a computationally feasible implementation of the maximum likelihood estimator of the model is provided. Further, it is shown that using paired end RNA-Seq provides more accurate isoform abundance estimates than single end sequencing at fixed sequencing depth. Simulation studies are also given.

Comments:	Published in at this http URL the Statistical Science (this http URL) by the Institute of Mathematical Statistics (this http URL)
Subjects:	Methodology (stat.ME); Genomics (q-bio.GN)
Journal reference:	Statistical Science 2011, Vol. 26, No. 1, 62-83
DOI:	10.1214/10-STS343
Report number:	IMS-STS-STS343
Cite as:	arXiv:1106.3211 [stat.ME]
	(or arXiv:1106.3211v1 [stat.ME] for this version)

Submission history

From: Julia Salzman [view email] [v1] Thu, 16 Jun 2011 12:17:05 GMT (330kb)

Which authors of this paper are endorsers?