Canadian Bioinformatics Workshops

www.bioinformatics.ca
Learning objectives of the course

- Module 1: Introduction to RNA sequencing
- Module 2: RNA-seq alignment and visualization
- Module 3: Expression and Differential Expression
- **Module 4: Isoform discovery and alternative expression**

- Tutorials
  - Provide a working example of an RNA-seq analysis pipeline
  - Run in a ‘reasonable’ amount of time with modest computer resources
  - Self contained, self explanatory, portable
Learning Objectives of Module

• Explore use of Cufflinks in reference annotation based transcript (RABT) assembly mode and ‘de novo’ assembly mode.
  – Both modes require a reference genome sequence...

Review of gene expression
Methods to study splicing by RNA-seq


Useful resources and discussion

• Best approach to predict novel and alternative splicing events from RNA-seq data
  – http://www.biostars.org/p/68966/
  – http://www.biostars.org/p/62728/

• Alternative splicing detection
  – http://www.biostars.org/p/65617/
  – http://www.biostars.org/p/11695/

• Identifying genes that express different isoforms in cancer vs normal RNA-seq data
  – http://www.biostars.org/p/50365/

• Cufflinks / Cuffdiff Output - How are tests different?
  – http://www.biostars.org/p/13525/

• Visualisation of Alternative splicing events using RNA-seq data
  – http://www.biostars.org/p/8979/
Types of alternative expression - part 1

- Simple transcription
- Alternative transcript initiation
- Alternative splicing

Module 4 – Isoform discovery and alternative expression  bioinformatics.ca

Types of alternative expression – part 2

- Alternate 3' splice sites
- Mutually exclusive exons
- Intron retention
- Alternative polyadenylation

Module 4 – Isoform discovery and alternative expression  bioinformatics.ca
Sequencing methods for studying alternative isoforms

Cufflinks alternative splicing tests
Cufflinks alternative splicing tests

Introduction to tutorial

(Module 4)
Module 4 – Isoform discovery and alternative expression

Bowtie/Tophat/Cufflinks/Cuffdiff RNA-seq Pipeline

Sequencing: RNA-seq reads (2 x 100 bp)
Read alignment: Bowtie/Tophat alignment (genome)
Transcript compilation: Cufflinks
Gene identification: Cufflinks (cuffmerge)
Differential expression: Cuffdiff (A:B comparison)

Visualization: CummRbund

Inputs:
- Raw sequence data (.fastq files)
- Reference genome (.fa file)
- Gene annotation (.gtf file)

Bowie/Tophat/Cuffinks/Cuffdiff RNA-seq Pipeline

Module 4 – Rerun Cufflinks in alternative ‘modes’

Sequencing: RNA-seq reads (2 x 100 bp)
Read alignment: Bowtie/Tophat alignment (genome)
Transcript compilation: Cufflinks
Gene identification: Cufflinks (cuffmerge)
Alternative expression: Cuffdiff (A:B comparison)

Visualization: CummRbund

Inputs:
- Raw sequence data (.fastq files)
- Reference genome (.fa file)
- Gene annotation (.gtf file)
What if I don’t have a reference genome for my species?

• Have you considered sequencing the genome of your species?

• If that is not practical or you simply prefer a transcript discovery approach that does not rely on prior knowledge of the genome or transcriptome there are some tools available ...
  – Unfortunately de novo transcriptome assembly is beyond the scope of this workshop
  – The good news is that the skills you learn here will help you figure out how to install and run those tools yourself

Methods to study splicing by RNA-seq

We are on a Coffee Break & Networking Session