bioinformatics.ca

Canadian Bioinformatics Workshops

www.bioinformatics.ca

Creative Commons

This page is available in the following languages:
Afrikaans български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE) Euskara Suomeksi français français (CA) Galego עברית hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски srpski (latinica) Sotho svenska 中文 華語 (台灣) isiZulu

creative commons

Attribution-Share Alike 2.5 Canada

You are free:

to Share — to copy, distribute and transmit the work

to Remix — to adapt the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

Disclaimer

Your fair dealing and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full licence) available in the following languages:
English French

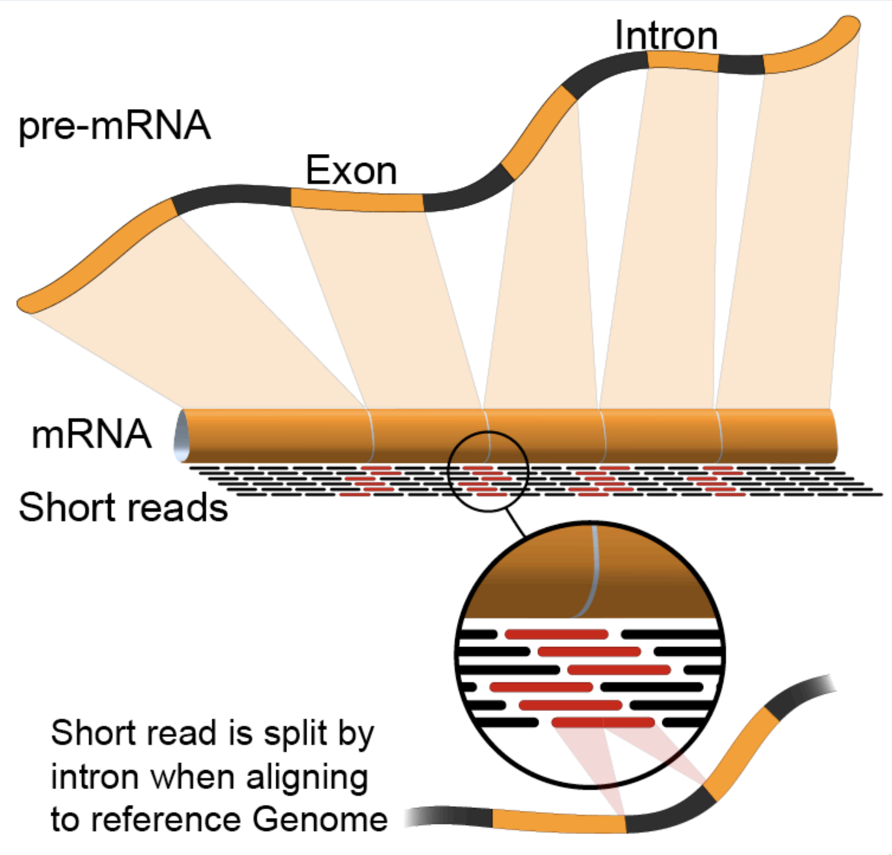
Learn how to distribute your work using this licence

Module 4
Isoform discovery and alternative expression (lecture)

Malachi Griffith & Obi Griffith

Informatics for RNA-seq Analysis

June 8-9, 2015



bioinformatics.ca

Learning objectives of the course

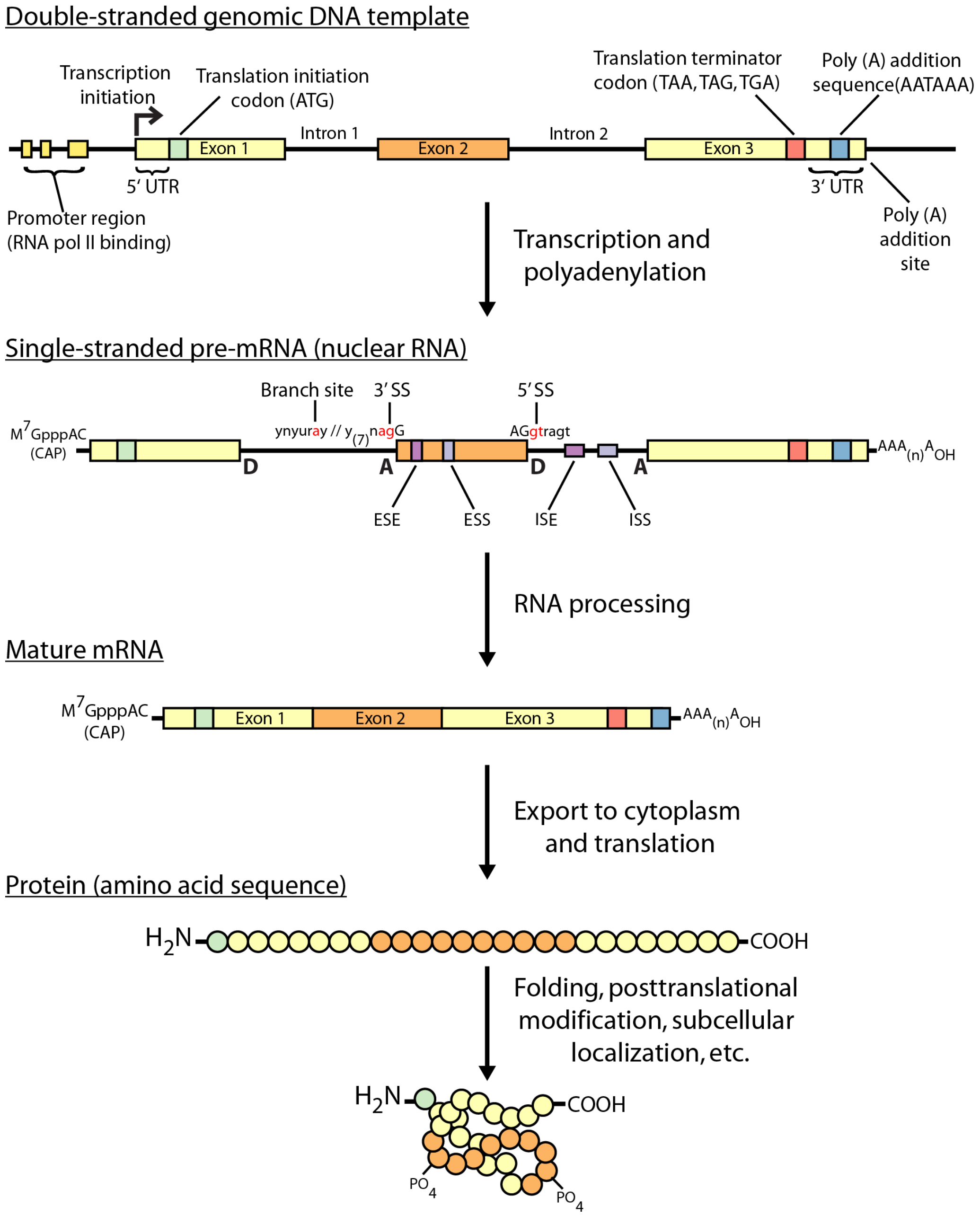
- Module 0: Introduction to cloud computing
- Module 1: Introduction to RNA sequencing
- Module 2: RNA-seq alignment and visualization
- Module 3: Expression and Differential Expression
- **Module 4: Isoform discovery and alternative expression**

- 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

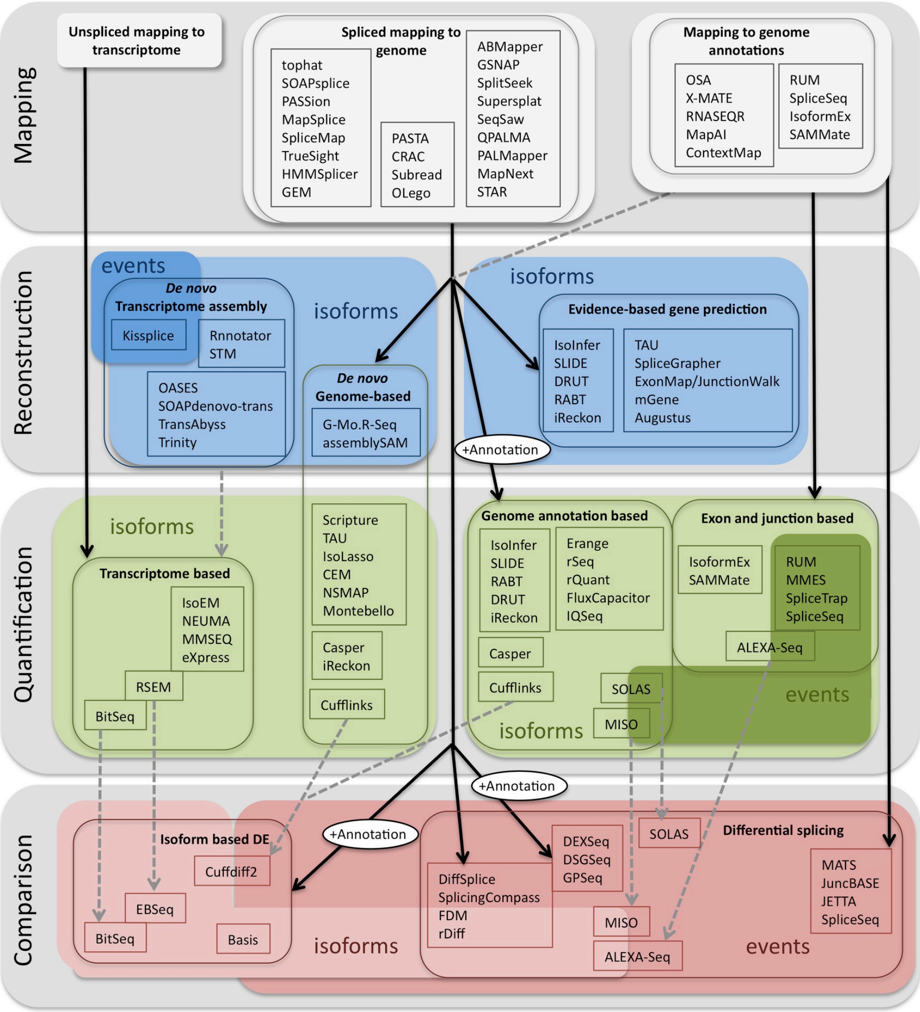
Learning Objectives of Module

- 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



Mapping

Unspliced mapping to transcriptome

Spliced mapping to genome
tophat
SOAPsplice
PASSion
MapSplice
SpliceMap
TrueSight
HMMSplicer
GEM

PASTA
CRAC
Subread
OLego

ABMapper
GSNAP
SplitSeek
Supersplat
SeqSaw
QPALMA
PALMapper
MapNext
STAR

Mapping to genome annotations
OSA
X-MATE
RNASEQR
MapAl
ContextMap

RUM
SpliceSeq
IsoformEx
SAMMate

Reconstruction

events

***De novo* Transcriptome assembly**
Kissplice
Rnnotator
STM
OASES
SOAPdenovo-trans
TransAbyss
Trinity

isoforms

***De novo* Genome-based**
G-Mo.R-Seq
assemblySAM

isoforms

Evidence-based gene prediction
IsoInfer
SLIDE
DRUT
RABT
iReckon

TAU
SpliceGrapher
ExonMap/JunctionWalk
mGene
Augustus

+Annotation

Quantification

isoforms

Transcriptome based
IsoEM
NEUMA
MMSEQ
eXpress
RSEM
BitSeq

Scripture
TAU
IsoLasso
CEM
NSMAP
Montebello

Casper
iReckon

Cufflinks

Genome annotation based
IsoInfer
SLIDE
RABT
DRUT
iReckon

Erange
rSeq
rQuant
FluxCapacitor
IQSeq

Casper
Cufflinks
SOLAS
MISO

isoforms

Exon and junction based
IsoformEx
SAMMate

RUM
MMES
SpliceTrap
SpliceSeq

ALEXA-Seq

events

+Annotation

Comparison

Isoform based DE
Cuffdiff2
EBSeq
BitSeq
Basis

+Annotation

isoforms

DiffSplice
SplicingCompass
FDM
rDiff

DEXSeq
DSGSeq
GPSeq

SOLAS
MISO
ALEXA-Seq

Differential splicing
MATS
JuncBASE
JETTA
SpliceSeq

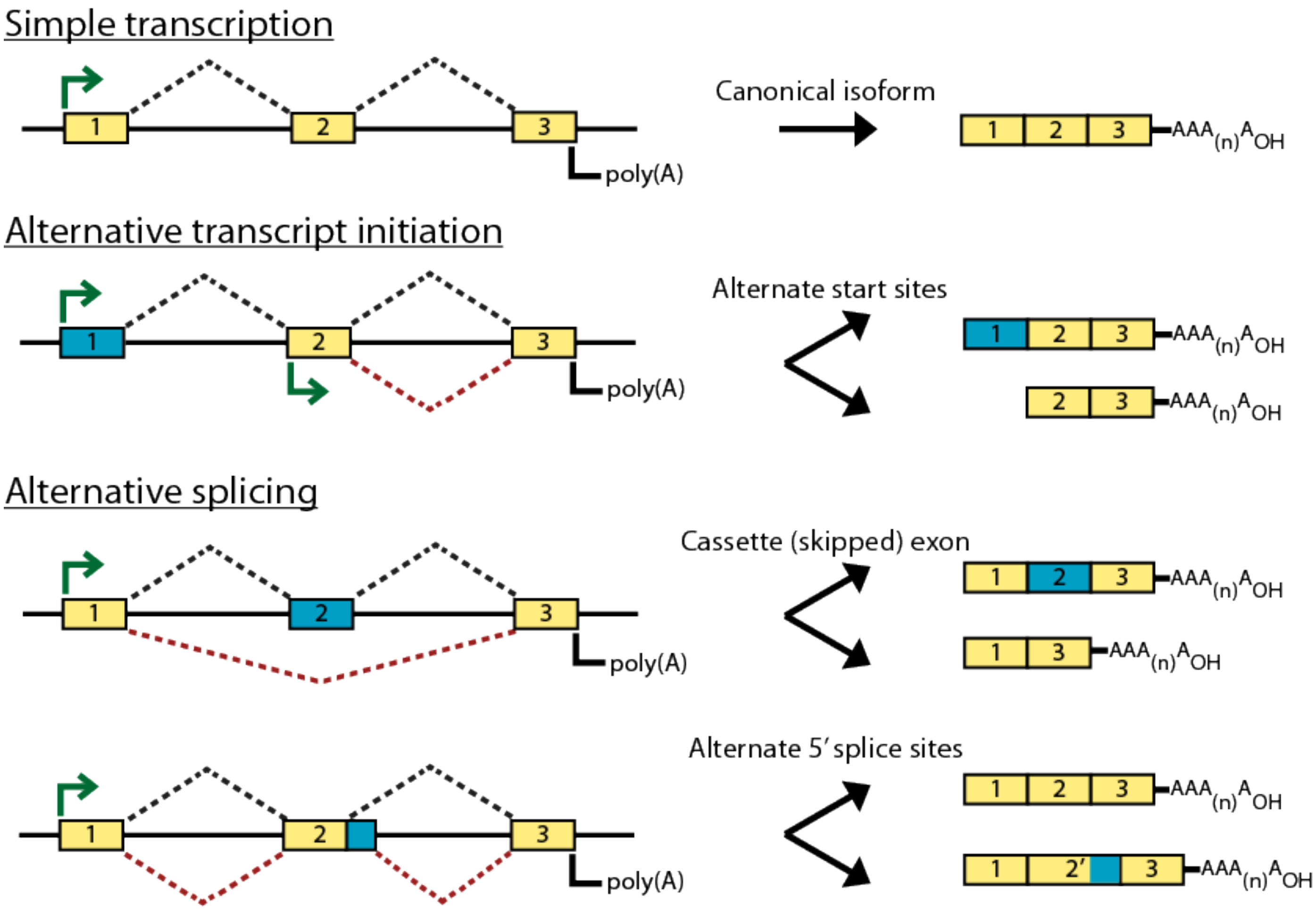
events

http://www.rna-seqblog.com/data-analysis/splicing-junction/methods-to-study-splicing-from-rna-seq/
http://arxiv.org/ftp/arxiv/papers/1304/1304.5952.pdf

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/
- Visualization of alternative splicing events using RNA-seq data
  - http://www.biostars.org/p/8979/

Types of alternative expression - part 1



Simple transcription

Canonical isoform

Alternative transcript initiation

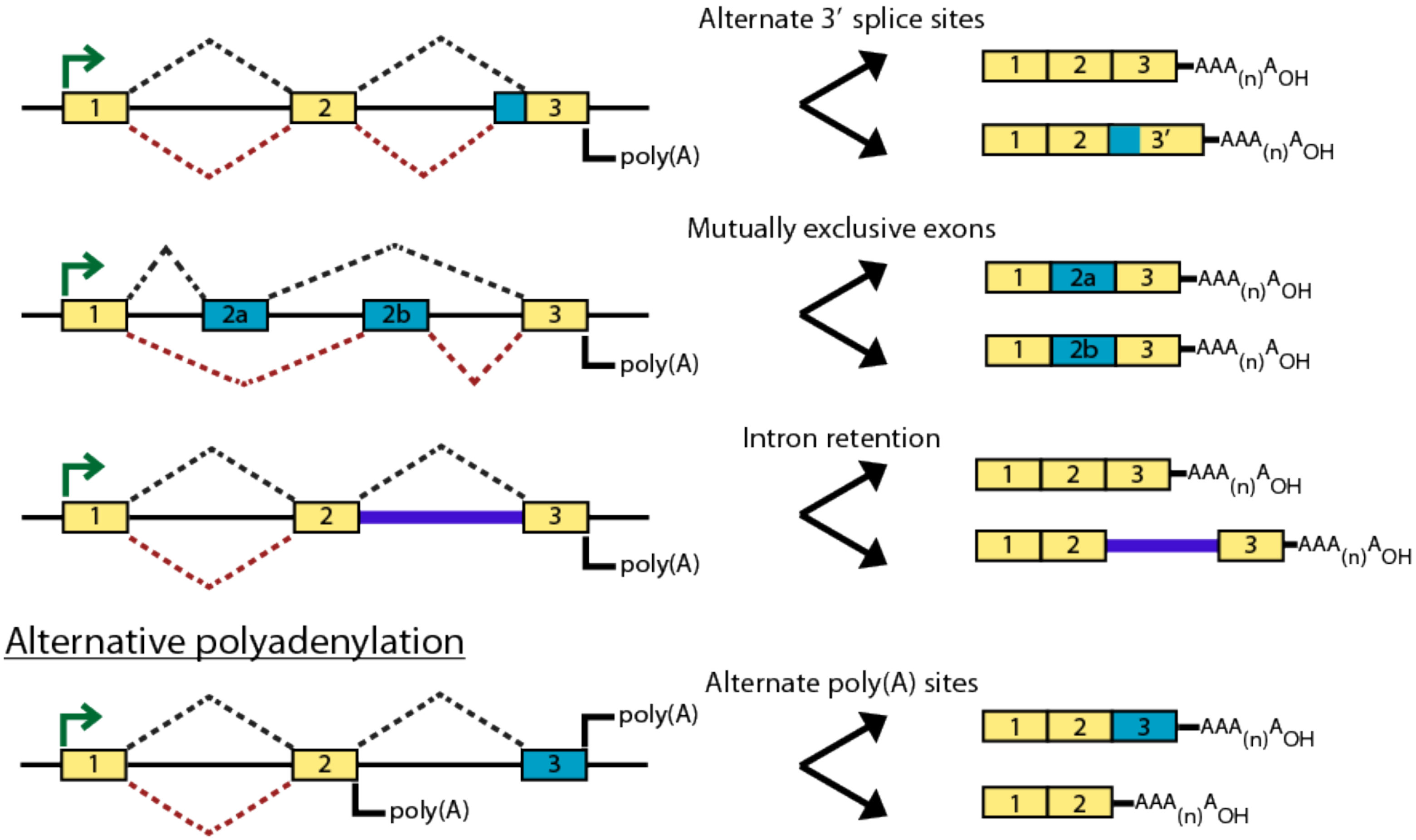
Alternate start sites

Alternative splicing

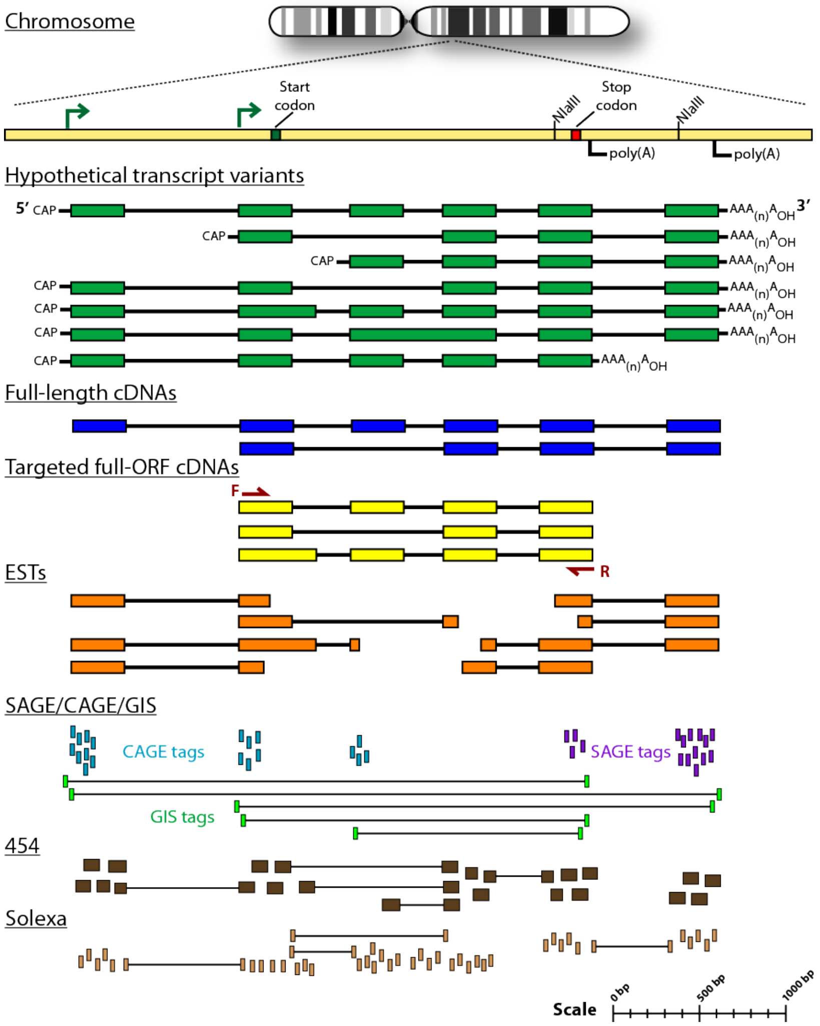
Cassette (skipped) exon

Alternate 5′ splice sites

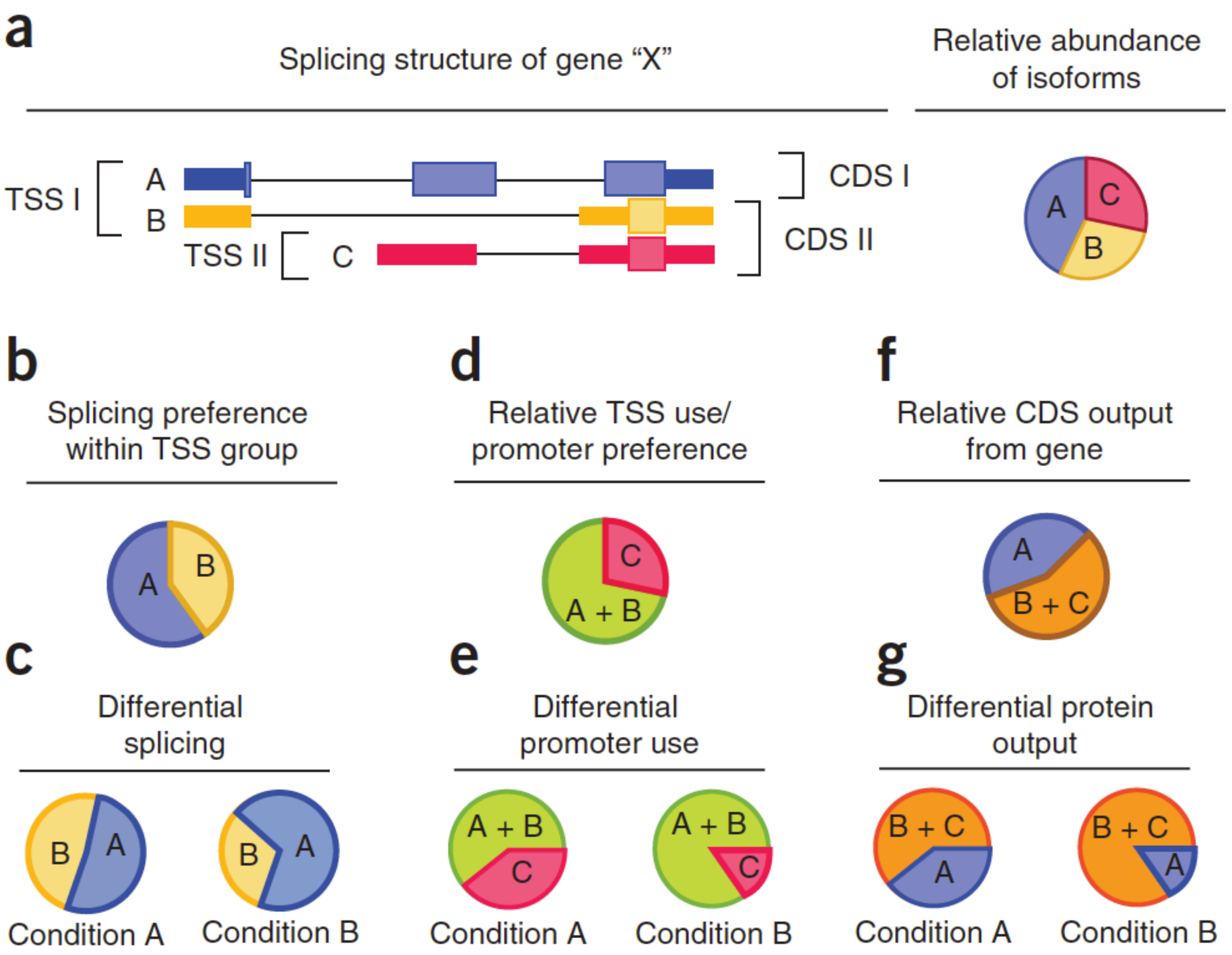
Types of alternative expression – part 2



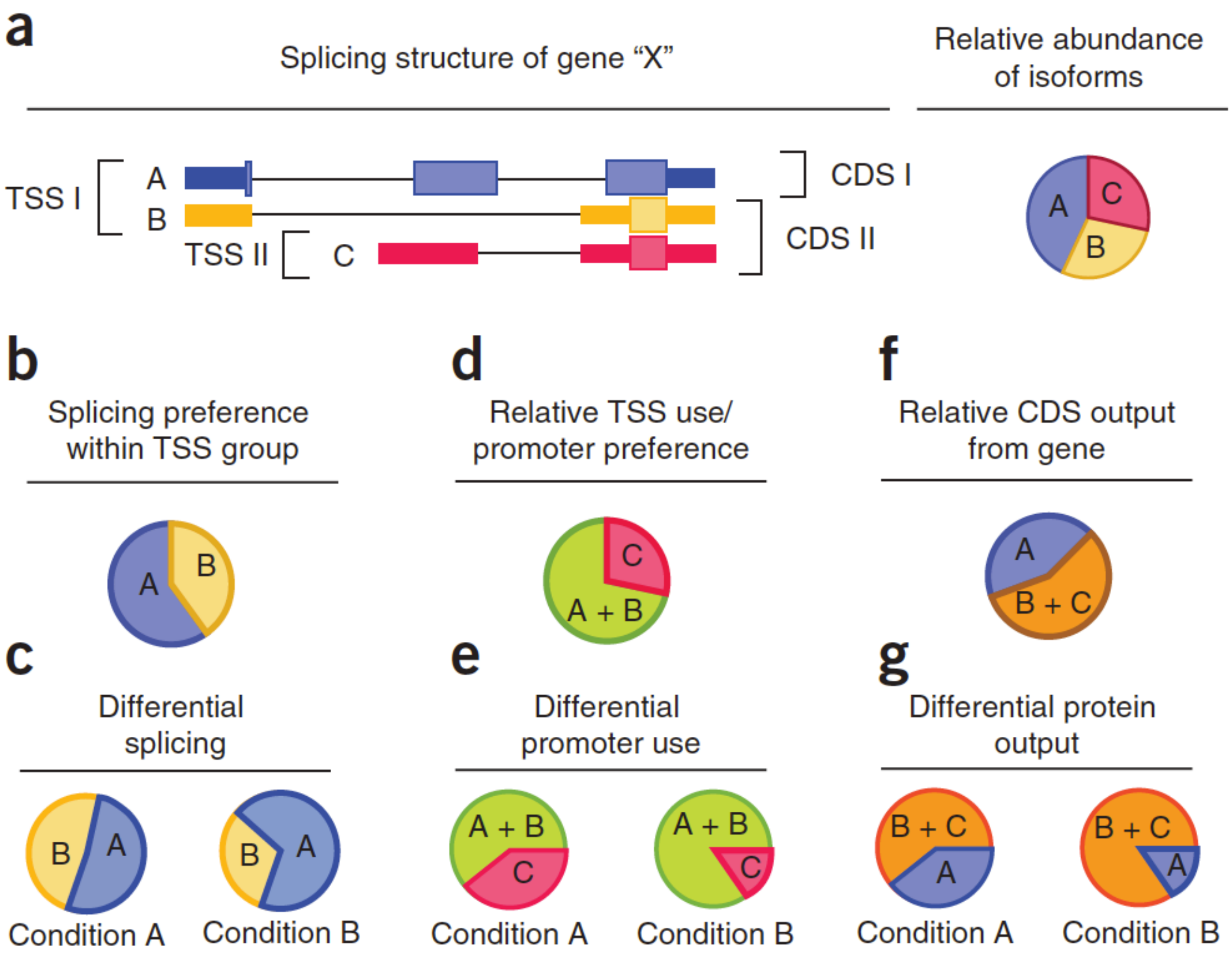
Sequencing methods for studying alternative isoforms



Cufflinks alternative splicing tests



Cufflinks alternative splicing tests



Relevant Cuffdiff output file

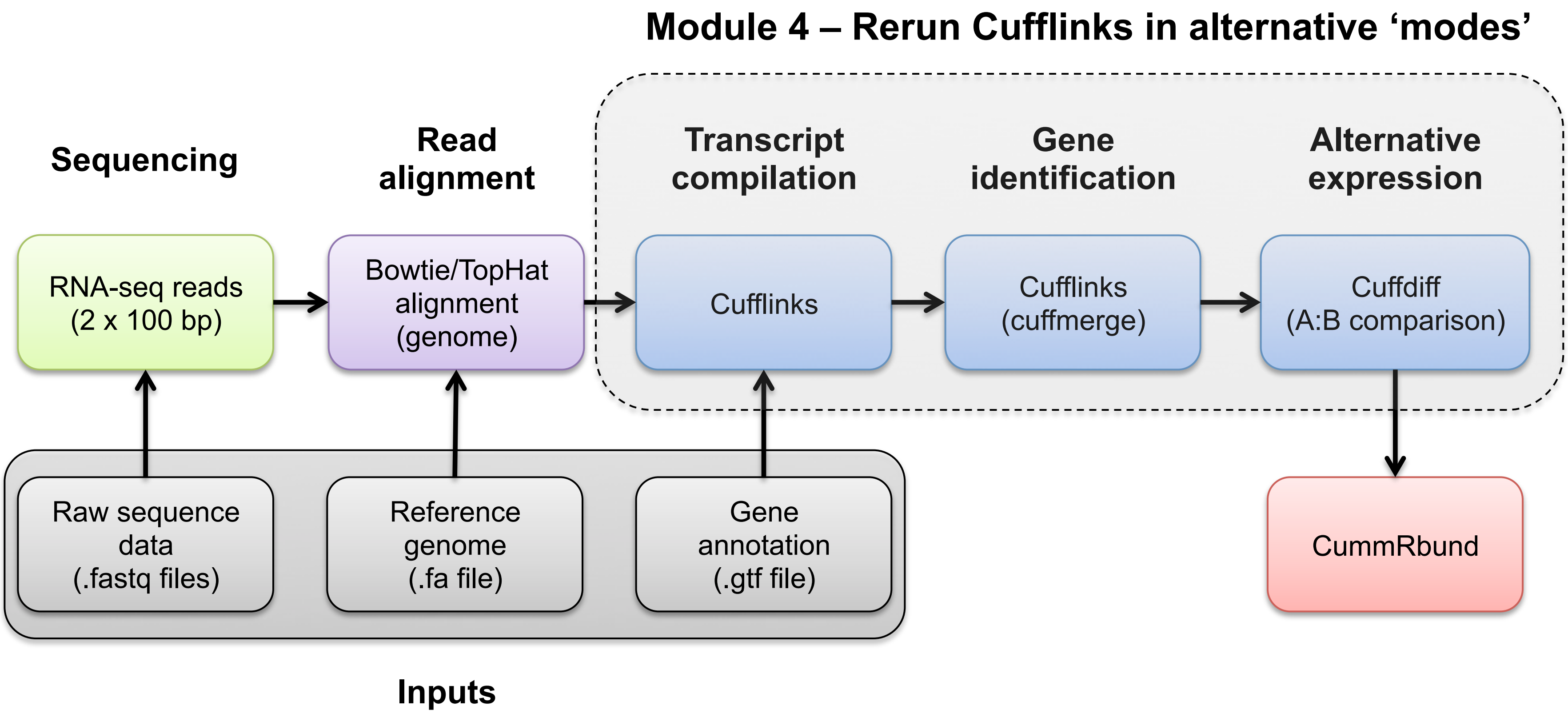
splicing.diff

promoters.diff

cds.diff

Introduction to tutorial
(Module 4)

Bowtie/Tophat/Cufflinks/Cuffdiff RNA-seq Pipeline



Module 4 – Rerun Cufflinks in alternative 'modes'

Sequencing

Read alignment

Transcript compilation

Gene identification

Alternative expression

RNA-seq reads (2 x 100 bp)

Bowtie/TopHat alignment (genome)

Cufflinks

Cufflinks (cuffmerge)

Cuffdiff (A:B comparison)

Raw sequence data (.fastq files)

Reference genome (.fa file)

Gene annotation (.gtf file)

CummRbund

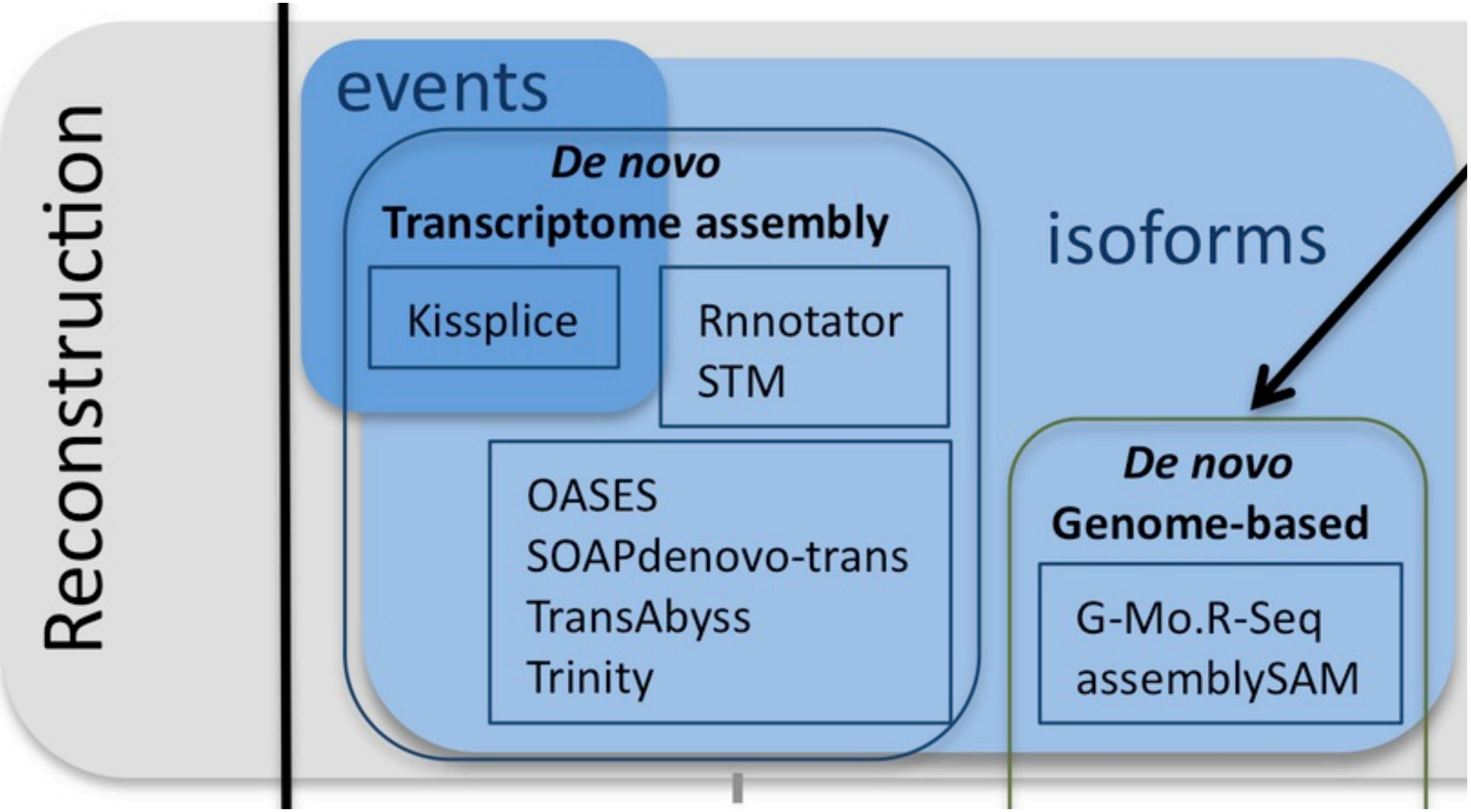
Inputs

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



http://www.rna-seqblog.com/data-analysis/splicing-junction/methods-to-study-splicing-from-rna-seq/
http://arxiv.org/ftp/arxiv/papers/1304/1304.5952.pdf

We are on a Coffee Break & Networking Session

bioinformatics.ca