



Cold  
Spring  
Harbor  
Laboratory

# Advanced Sequencing Technologies & Applications

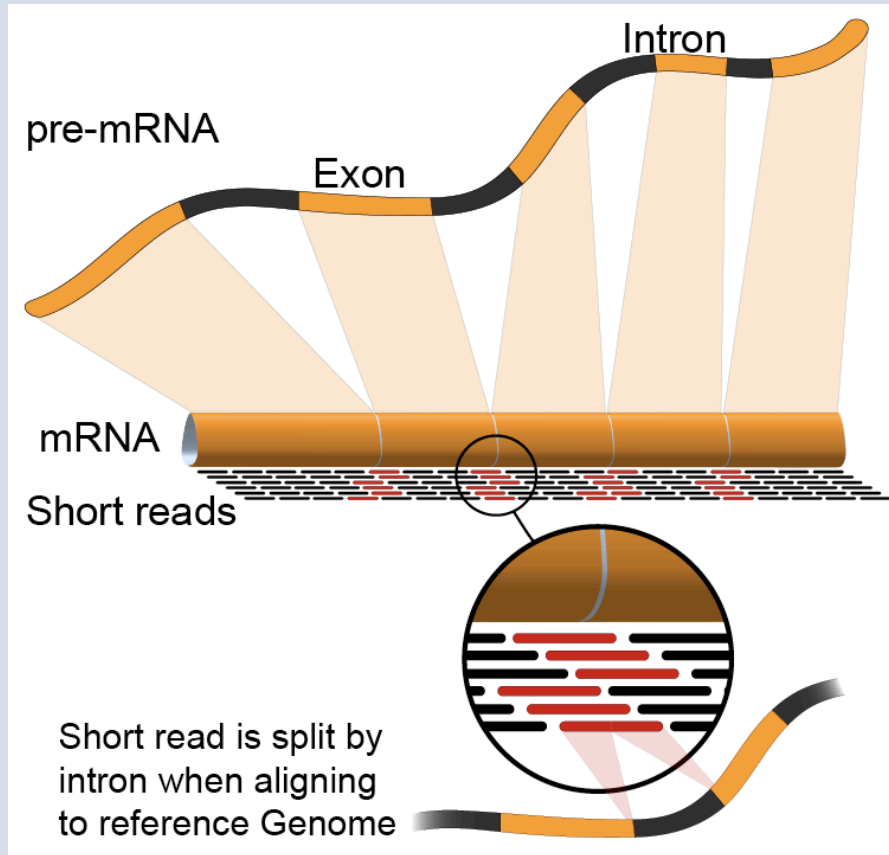
<http://meetings.cshl.edu/courses.html>



Cold  
Spring  
Harbor  
Laboratory

# Module 5 Isoform discovery and alternative expression (lecture)

Malachi Griffith, Obi Griffith, Jason Walker, Ben Ainscough  
Advanced Sequencing Technologies & Applications  
November 11-23, 2014



# Learning objectives of the course

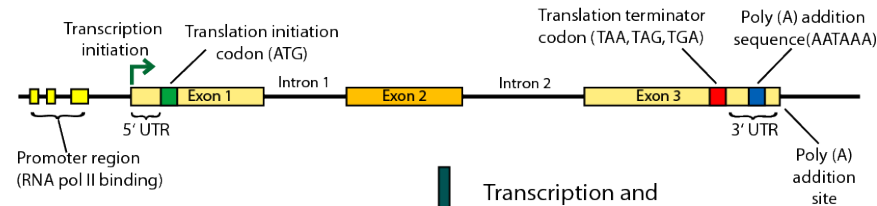
- Module 1: Introduction to cloud computing
- Module 2: Introduction to RNA sequencing
- Module 3: RNA-seq alignment and visualization
- Module 4: Expression and Differential Expression
- **Module 5: Isoform discovery and alternative expression**
  
- Tutorials
  - Provide a working example of an RNA-seq 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...

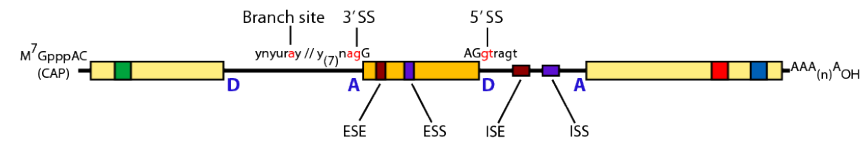
# Review of gene expression

## Double-stranded genomic DNA template



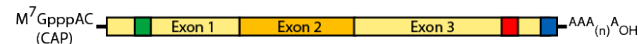
Transcription and polyadenylation

## Single-stranded pre-mRNA (nuclear RNA)



RNA processing

## Mature mRNA

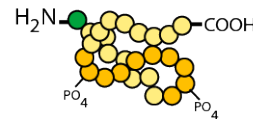


Export to cytoplasm and translation

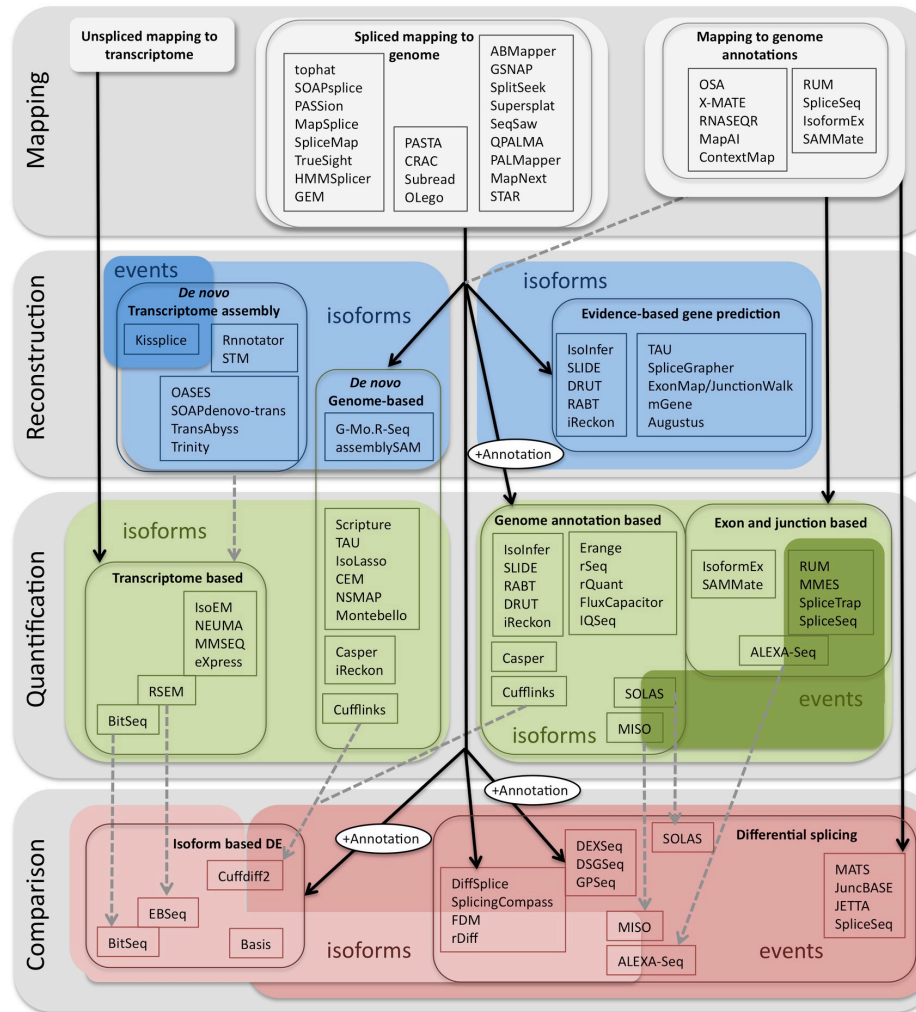
## Protein (amino acid sequence)



Folding, posttranslational modification, subcellular localization, etc.



# Methods to study splicing by RNA-seq



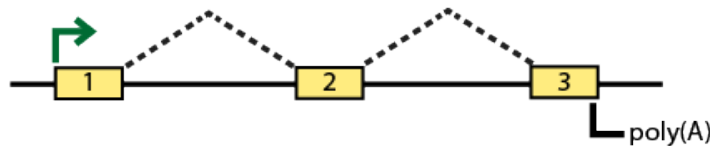
<http://www.rna-seqblog.com/data-analysis/splicing-junction/methods-to-study-splicing-from-rna-seq/>  
<http://arxiv.org/ftp/arxiv/papers/1304/1304.5952.pdf>

# Useful resources and discussion

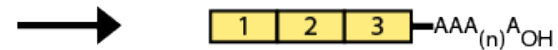
- Best approach to predict novel and alternative splicing events from RNA-seq data
  - <http://www.biostars.org/p/68966/>
  - <http://www.biostars.org/p/62728/>
- Alternative splicing detection
  - <http://www.biostars.org/p/65617/>
  - <http://www.biostars.org/p/11695/>
- Identifying genes that express 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

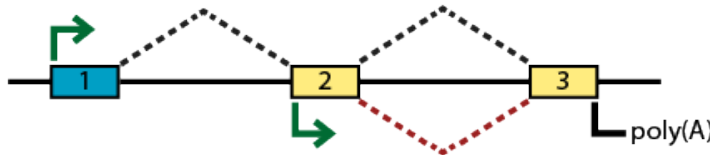
## Simple transcription



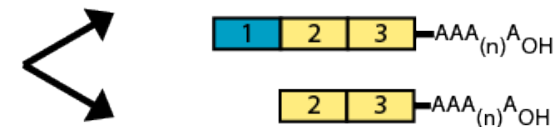
Canonical isoform



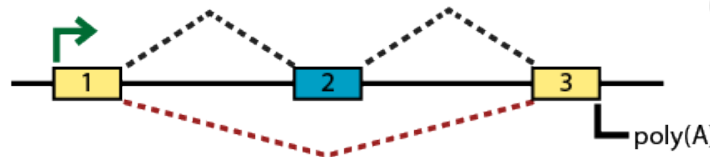
## Alternative transcript initiation



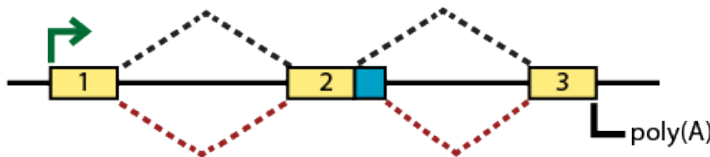
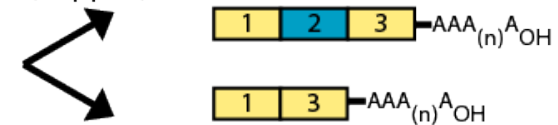
Alternate start sites



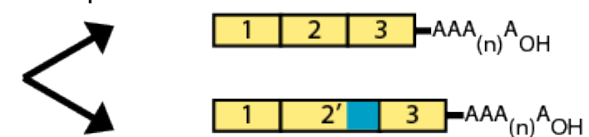
## Alternative splicing



Cassette (skipped) exon

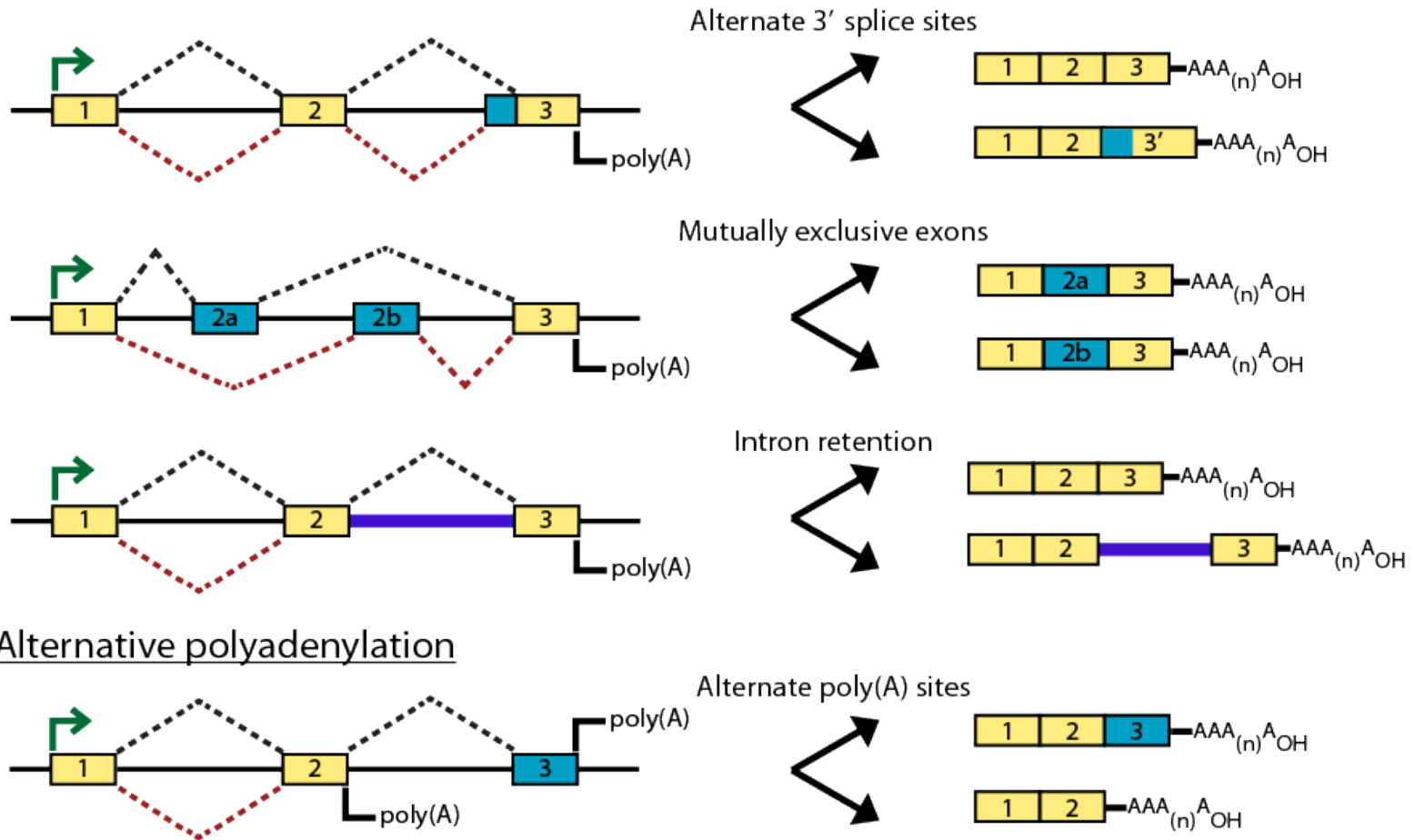


Alternate 5' splice sites

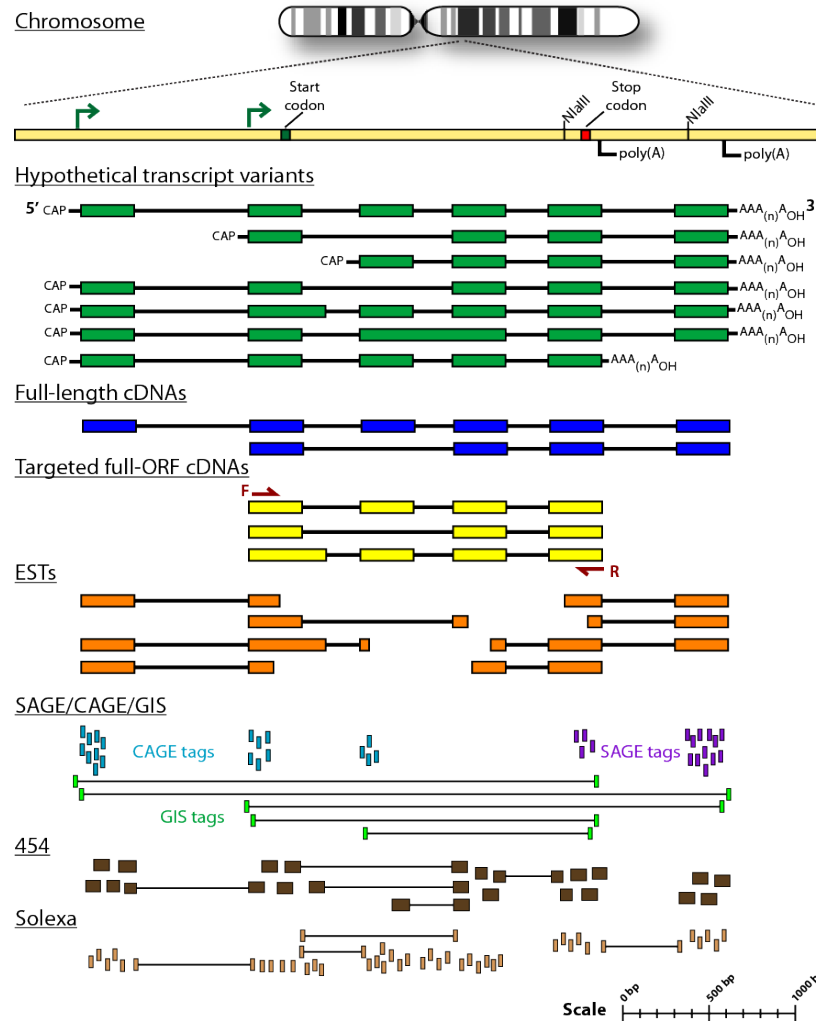




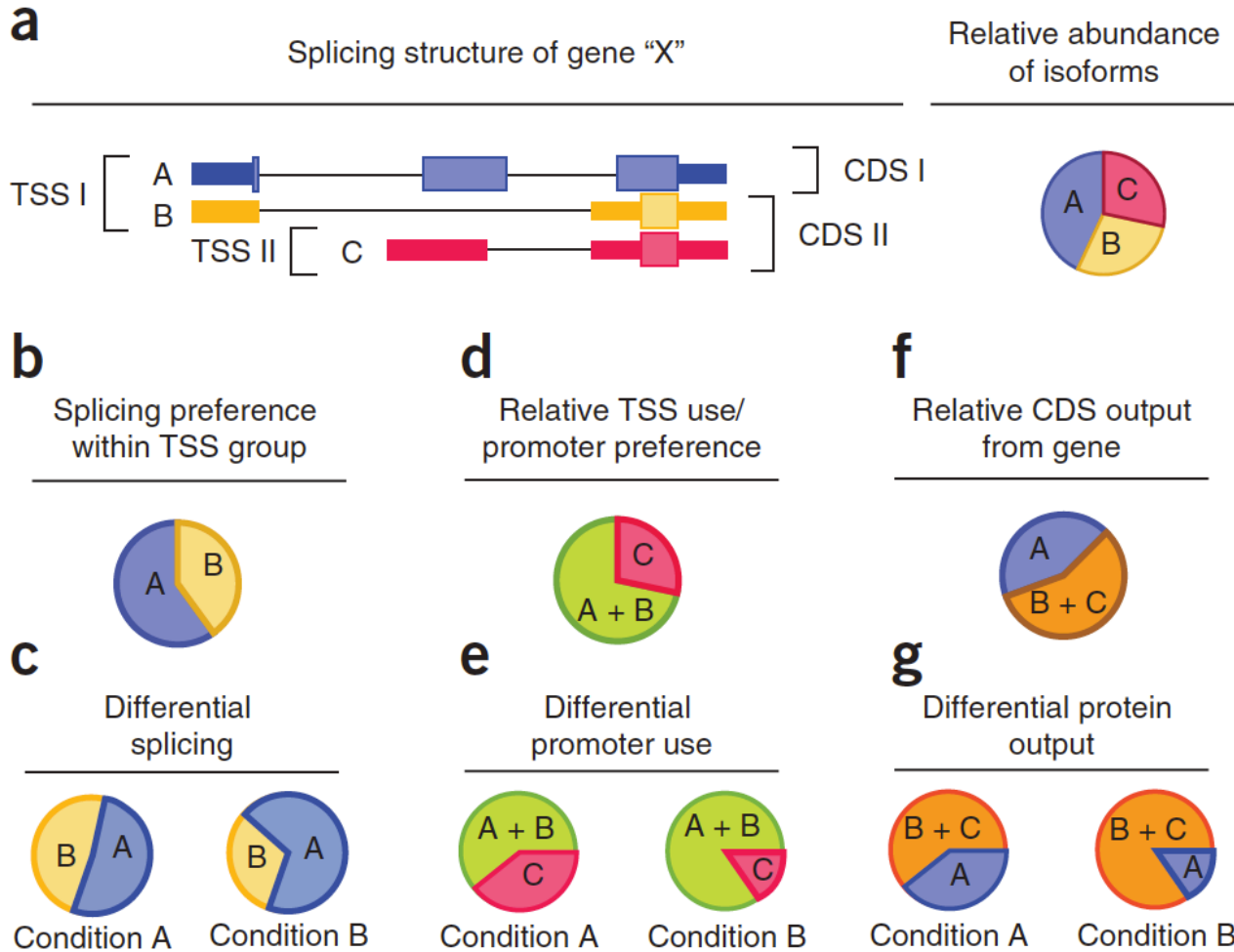
# Types of alternative expression – part 2



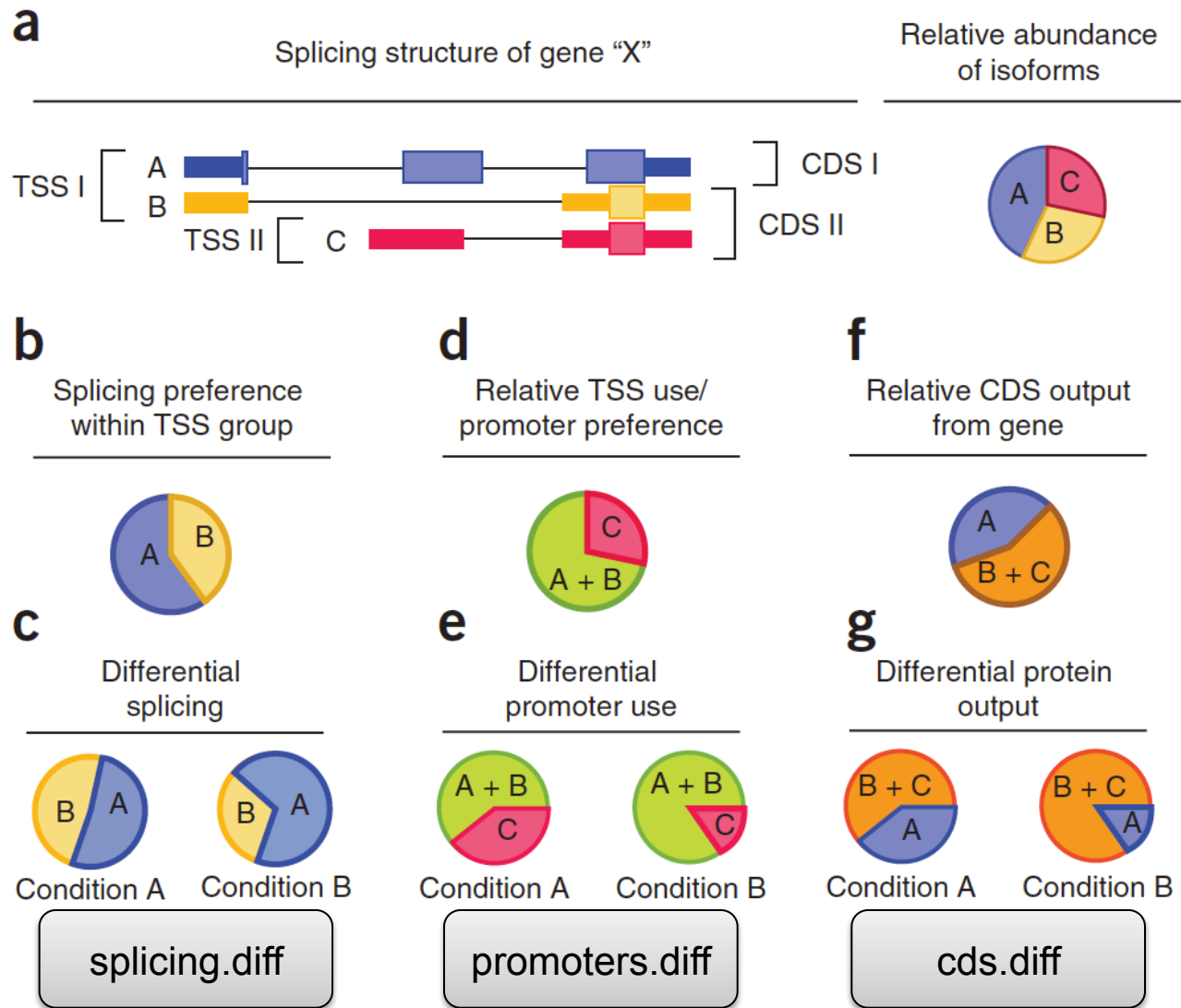
# Sequencing methods for studying alternative isoforms



# Cufflinks alternative splicing tests

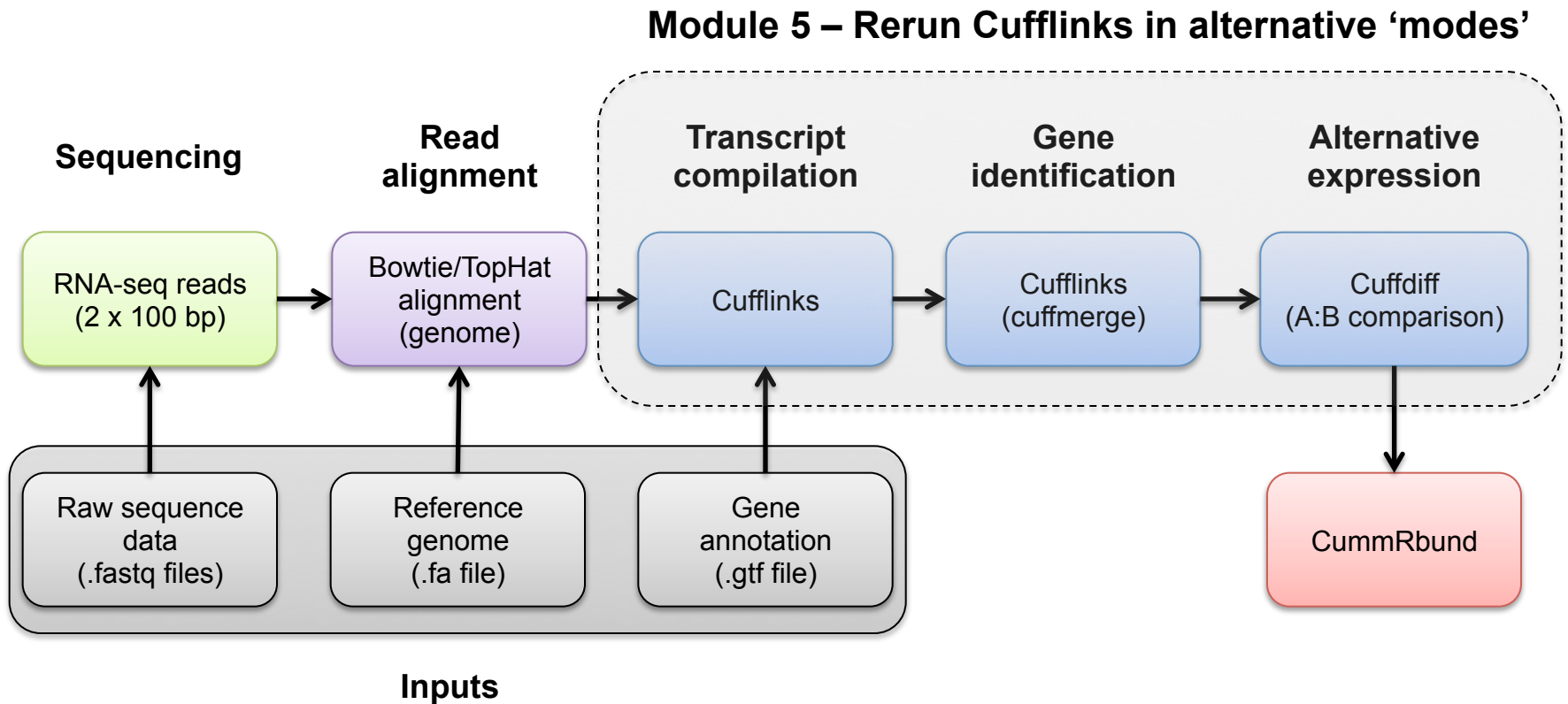


# Cufflinks alternative splicing tests



# **Introduction to tutorial (Module 5)**

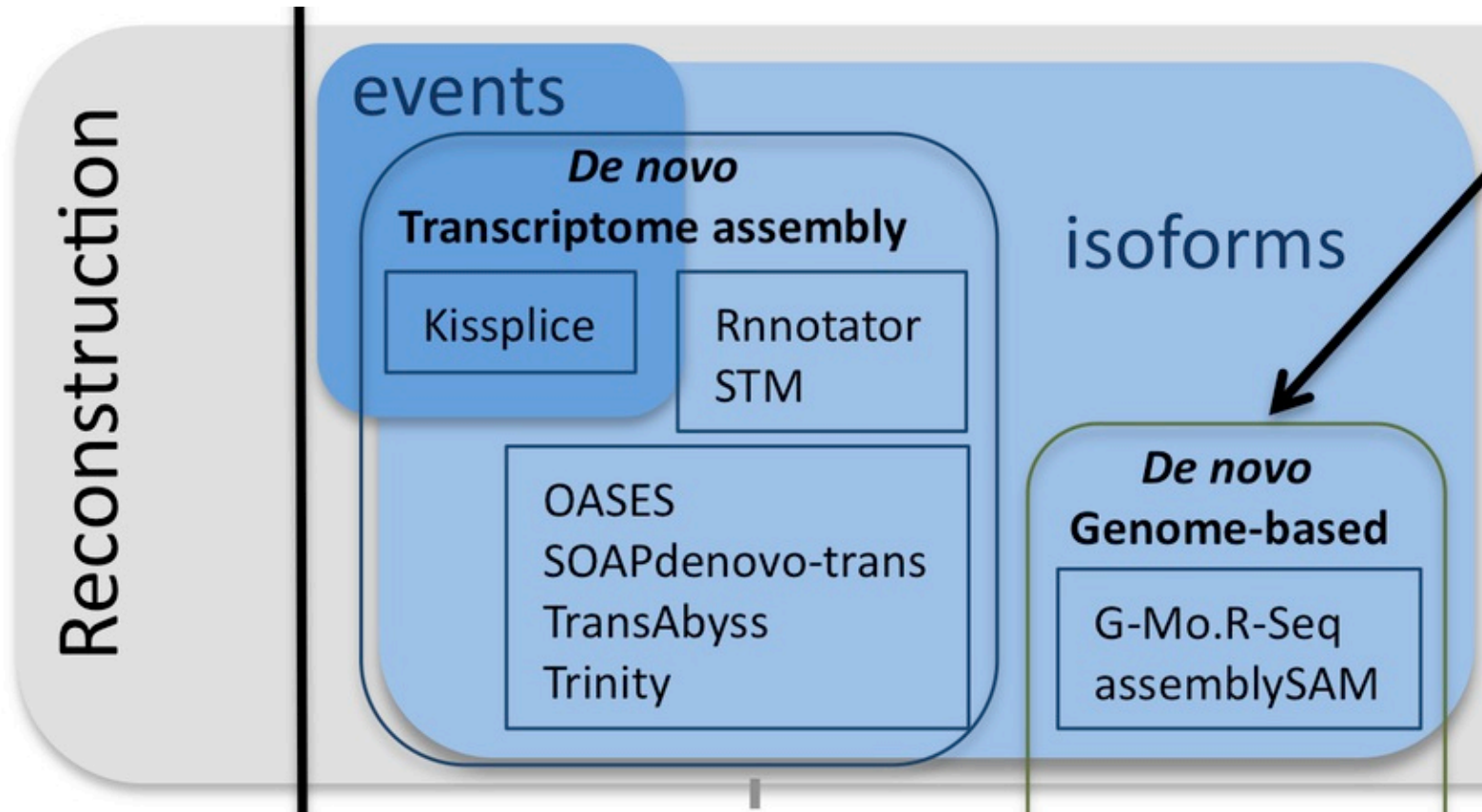
# Bowtie/TopHat/Cufflinks/Cuffdiff RNA-seq Pipeline



# 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>



Break