RNA-Seq Analysis

Hadrien Gourlé - SLU

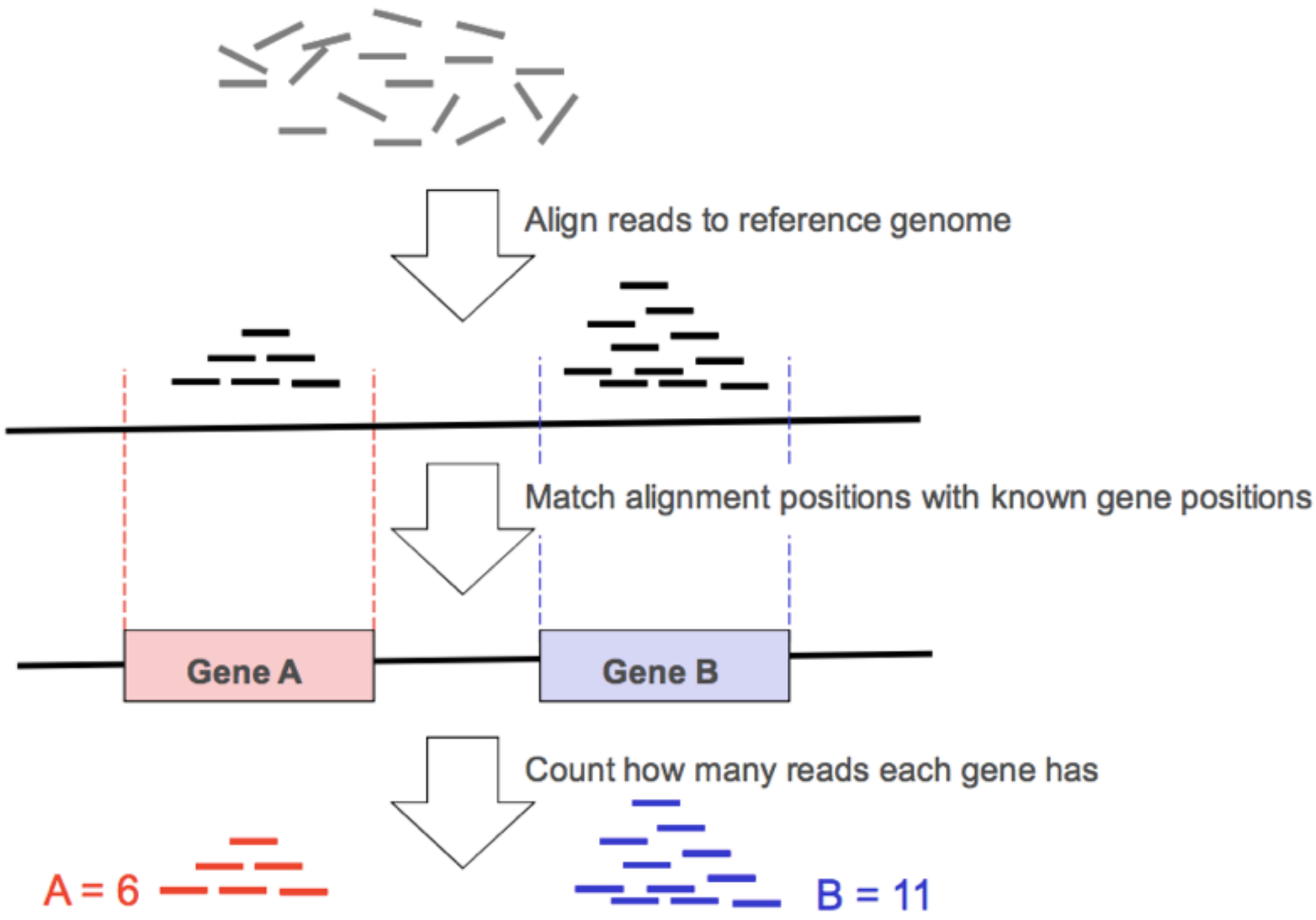
Introduction to RNA-Seq

- What can I do with RNA-Seq?
  - Differential expression
  - Discover new genes and isoforms
  - New Transcriptomes

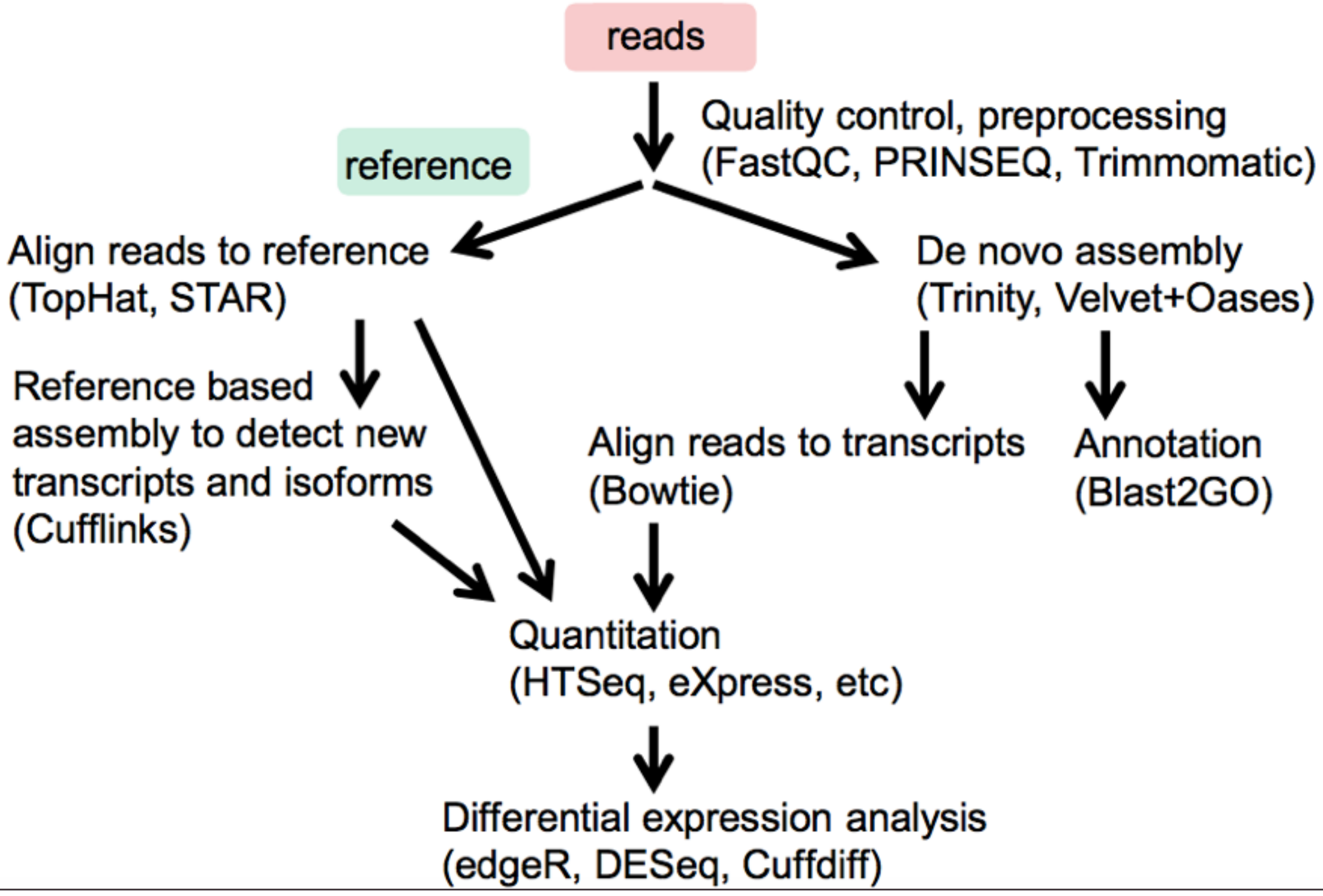
Microarrays?

- Rna-Seq has:
  - a better detection range
  - Can detect new genes and isoforms
  - But the data is voluminous and hard to analyse

RNA-Seq workflow



RNA-Seq workflow



Some things to take into account

- Non uniform coverage
  - Biases in library construction
  - Uniqueness of genomic regions (repeats)
- Longer transcripts give more counts

Alignment to the genome

- Goal is to find where a read originated from
- Mapping to:
  - transcriptome (gene count)
  - genome (finding new transcripts)
- If organisms have introns: Spliced alignment!
  - tens of aligners available. We'll use tophat

Quantitation

- Cufflinks
- HT-Seq

- Given a BAM file and a list of genomic features (e.g. genes), counts how many reads map to each feature.