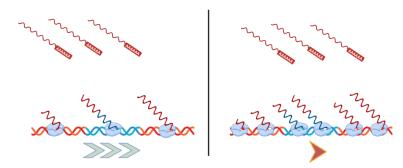
Measuring RNA elongation and splicing speed changes from total RNA-seq

Jakub Koubele (Beyer lab)

Introduction

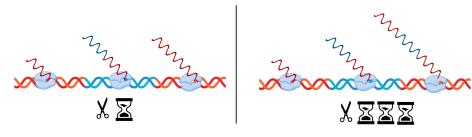
- Changes in RNA elongation and splicing speed influence the abundance and composition of nascent RNA content.
- Total RNA-seq (Ribo-minus) captures both the mature and nascent RNA seq. Elongation and splicing speed changes influence the abundance of intronic reads and their location (coverage), as introns are predominantly present in the nascent RNA.
- I am developing a computational method that estimates the log fold changes (LFC) of RNA Pol II elongation speed and the speed of splicing. The code (work in progress) can be found here: github.com/jkoubele/dev-pol-ii-analysis

Elongation speed changes



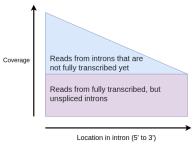
• Slower speed \implies higher occupancy of DNA by RNA polymerases \implies more nascent RNA \implies higher fraction of intronic reads.

Splicing speed changes



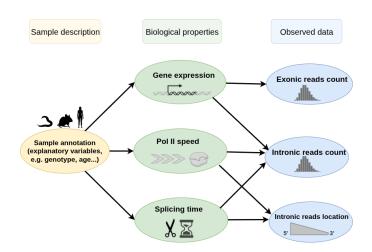
• Slower splicing \implies more intronic reads.

Location of intronic reads



- Intronic reads come from two sources: 1) from introns that are just being transcribed by Pol II, and 2) from introns that were already transcribed, but not spliced (and degraded) yet.
- The amount of Pol II's transcribing intron is inversely proportional to the elongation speed. The amount of fully transcribed but unspliced introns is inversely proportional to the splicing speed.
- Thus, the shape of intron coverage depends on the ratio of elongation and splicing speed.

Full model



- The abundance of exonic and intronic reads, and the location of intronic reads, depend on the LFC of overall gene expression, elongation speed, and splicing speed.
- We therefore have a probability model parametrized by these LFC. I fit the model by maximum-likelihood estimation (MLE)
 - Statistical significance (p-values) of LFC can be obtained as well, via Wald test, likelihood-ratio test, and/or parametric bootstraping.