

Stony Brook University The Graduate School

Doctoral Defense Announcement

Abstract

Probabilistic methods for predicting local rates of transcription elongation
from nascent RNA sequencing data

By

Lingjie Liu

Across all branches of life, transcription elongation is a crucial, regulated phase in gene expression. Many recent studies in eukaryotes have focused on the regulation of promoter-proximal pausing of RNA Polymerase II (Pol II), but rates of productive elongation also vary substantially throughout the gene body, both within and across genes. Here, we introduce a probabilistic model for systematically evaluating potential determinants of the local elongation rate based on nascent RNA sequencing (NRS) data. Our model is derived from a unified model for both the kinetics of Pol II movement along the DNA template and the generation of NRS read counts at steady state. It allows for a continuously variable elongation rate along the gene body, with the rate at each nucleotide defined by a generalized linear relationship with nearby genomic and epigenomic features. We validate this model with simulations and apply it to public PRO-seq and epigenomic data for four cell types, finding that reductions in local elongation rate are associated with cytosine nucleotides, DNA methylation, splice sites, RNA stem-loops, CTCF binding sites, and several histone marks, including H3K36me3 and H4K20me1. By contrast, increases in local elongation rate are associated with thymines, A+T-rich and low-complexity sequences, and H3K79me2 marks. We then introduce a convolutional neural network that improves our local rate predictions. Overall, our analysis is the first to permit genome-wide predictions of relative nucleotide-specific elongation rates based on complex sets of genomic and epigenomic covariates. We have made predictions available for the K562, CD14+, MCF-7, and HeLa-S3 cell types in a UCSC Genome Browser track.

Date: May 2, 2024

Time: 1:00 pm

Place: Cold Spring Harbor Laboratory, Beckman Building, Plimpton Room

To attend virtually, contact the Program Director at martha.furie@stonybrook.edu

Program: Genetics

Dissertation Advisor: Adam Siepel