

## Whalen Symposium Abstract

All life requires regulation of gene expression. Gene expression is the process by which genetic information encoded in a gene, flows from DNA to protein. DNA is copied into a single-stranded RNA which carries the coding instructions of the gene, known as messenger-RNA, or mRNA. This process is called transcription. Following transcription is the process of translation, in which the RNA molecule is decoded to produce a protein that is expressed and functions in the cell. RNA polymerase is the enzyme that copies the genetic information in DNA to RNA. In higher organisms, the RNAs are first copied to precursor molecules known as pre-mRNAs, that are highly processed before they can be properly translated. For example, splicing is the process by which non-coding regions of the gene are removed from the pre-mRNA molecule. This leaves the mRNA with only the coding regions of the gene, known as exons. The exons are thus spliced together into a mature mRNA molecule to be translated into a functional protein. Alternative splicing is the process in which the exons can be arranged in different combinations as the introns are removed. The variety of how the exons can be arranged gives rise to many different proteins and potential physical characteristics.

RNA polymerase has an unusual C-terminal or tail domain (CTD), which has ~30-50 seven amino acid repeats. This domain is known to recruit factors that process the RNA such as splicing factors during transcription. Interestingly, the number of heptad repeats in the CTD are thought to correlate with the complexity of the organism, suggesting that the role of the CTD is important for regulation. We are using the single-celled eukaryotic model organism, *Schizosaccharomyces pombe*, more commonly known as fission yeast, to study mechanisms of gene expression. What we learn in *S. pombe* will give us insights into mechanisms of higher organisms such as humans. The results of our previous study have shown that large scale

mutational changes to the CTD have led to increased gene expression for large regions at the ends of the chromosomes (Inada et al., 2016). Currently, we are generating an array of single amino acid mutations at each position in the CTD in *S. pombe* in order to examine their genome-wide effects on gene expression and splicing. This mutagenesis will be conducted en masse such that each amino acid in the CTD is changed to each of the possible 20 amino acids. We are currently cloning the RNA polymerase II gene, *rpb1*, into a vector by amplifying *rpb1* DNA sequence and ligating it into a plasmid. The ligated plasmid will then be replaced with the prepared library of single amino acid mutations oligos, transformed into *S. pombe*, and assayed to measure and examine their genome-wide effects on gene expression. Our goal is to identify the role of each amino acid in the CTD on splicing in gene expression. This data will help us begin to understand the mechanisms of gene expression via RNA splicing and apply our insights in more complex organisms, like humans.

## Reference

Inada, Maki, Robert J. Nichols, Jahan-Yar Parsa, Christina M. Homer, Ruby A. Benn, Reyal S. Hoxie, Hiten D. Madhani, Stewart Shuman, Beate Schwer, and Jeffrey A. Pleiss. "Phospho-Site Mutants of the RNA Polymerase II C-Terminal Domain Alter Subtelomeric Gene Expression and Chromatin Modification State in Fission Yeast." *Nucleic Acids Research*, July 8, 2016, gkw603. <https://doi.org/10.1093/nar/gkw603>.