Thesis Outline

Chapter 1. Introduction

1. RNA Binding Protein Function
   1. RBPs role in normal function of cell
      1. Splicing
      2. Stability
      3. Translation
      4. Others (less important because not focus of thesis)
   2. RBPs role in disease
      1. RBPs in neurodegenerative diseases
         1. Role in ALS
            1. TAF15
            2. FUS
            3. TDP-43
            4. hnRNPA2B1
   3. RBPs role in development
      1. Role in Stem cell differentiation
         1. (MSI2)
         2. LIN28
         3. RBFOX2
2. Methods to identify RBP binding
   1. Non-transcriptome approaches
      1. EMSA
      2. RNA-compete
      3. RNA Bind-n-seq
   2. CLIP
      1. iCLIP
      2. par-CLIP
      3. ir-CLIP
      4. eCLIP
      5. RIP-seq
3. CLIP Computational Methods
   1. Unique Molecular Identifier detection
      1. Darnel et. al
      2. UMI-tools
      3. Single cell UMI collapsing approaches
   2. Peak Finding Tools
      1. Piranah
      2. Pyicos
      3. Paralyzer
      4. PIPE-CLIP
      5. Others…
   3. Motif calling
      1. Methods
         1. HOMER / MEME / others
         2. Secondary Structure peak identification
            1. List of computational methods
            2. Prospective on shape-seq / pip-seq / other in-vitro computational methods
   4. Pipelines / reproducible science
      1. PIPE-CLIP
      2. Ule’s website (general clip reporting)
      3. There are a few others out there that you should hunt down
4. ENCODE
   1. Review of current accomplishments up till now
      1. This is going to be a lot, don’t know how much detail to go into
   2. Quality Control
5. Quality Control Field
   1. Landet et al paper
   2. Other QC papers I cite in my paper
   3. Also look for even more ChIP-seq papers
   4. Cross-contamination QC
   5. UMI methods
   6. ChIP QC / approaches

Chapter 2. Computational Improvements Bioinformatics Tools (or fold into something else? / cut)

1. Technical Improvements
   1. CLIPper
      1. Initial Optimization
      2. Removal of HTseq
   2. Pybedtools
      1. Identified memory leak and caused it to break far less
   3. Umi-tools / pysam
      1. Identified errors in pysam, and increased speed of UMI tools by ~1000x in some cases
2. Methodological Improvements
   1. Made clipper handle pre-mrna and mRNA distributions in the same run
   2. Need to look back at smoothing spline optimizations that I made / tried to make many years ago

Chapter 3. Guidelines and Best Practices for eCLIP experiments and analysis

1. eCT allows for library complexity estimation
2. Peak Saturation Analysis
3. QC metrics to verify individual and groups of experimental datasets
4. Public pipeline to run data from our lab + data generated by individual
5. Discussion of best ways to normalize and visualize RBP binding

Chapter 4. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses

1. Copy and paste published paper /w context

Chapter 5. Applications of CLIP-seq to biological questions

1. UPF1
2. MSI2

Conclusion / Prospectus