

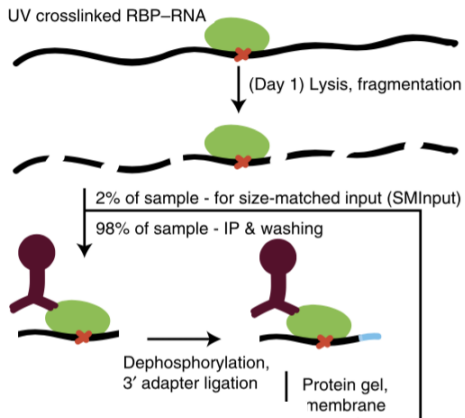
CLIP-Explorer: CLIP-Seq data analysis in Galaxy

Florian Heyl

Albert-Ludwigs-Universität Freiburg – Institute for Computer Science
MeInBio - BioInMe Research Training Group
34th TBI Winterseminar in Bled

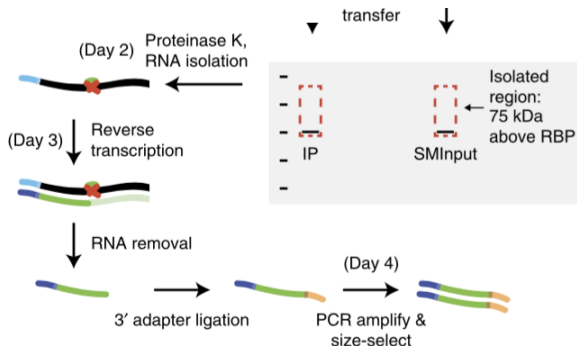
February 15, 2019

1.1 CLIP-Seq in a Nutshell

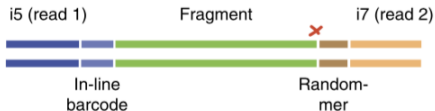


[Nostrand *et al.* 2016]

1.1 CLIP-Seq in a Nutshell



Final library structure:



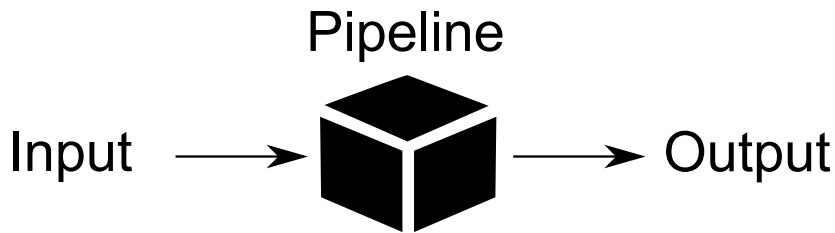
[Nostrand *et al.* 2016]

1.2 Problems of a Data Analysis Pipeline

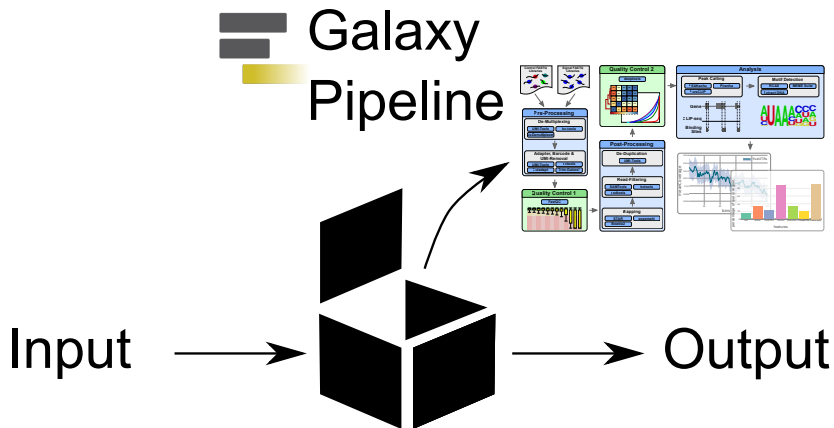


[<https://www.wikihow.com/Play-Jenga>]

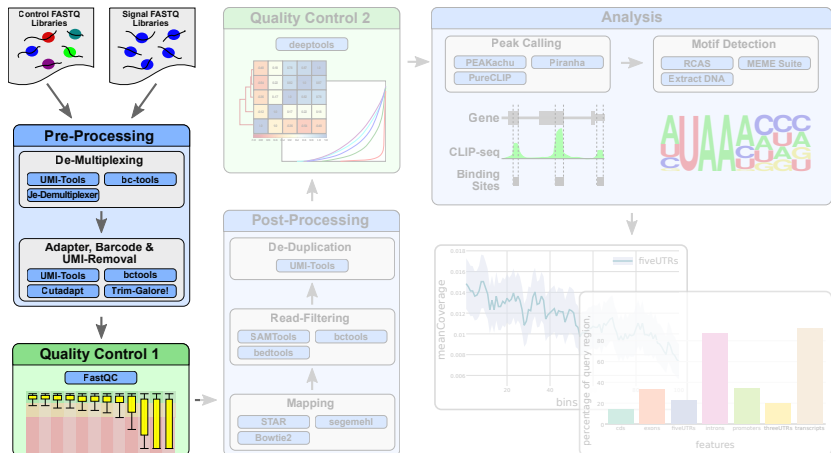
1.2 Problems of a Data Analysis Pipeline



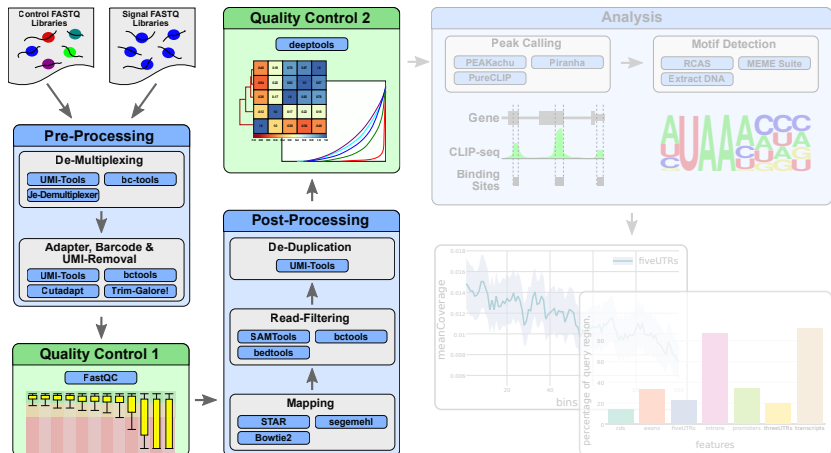
1.2 Problems of a Data Analysis Pipeline



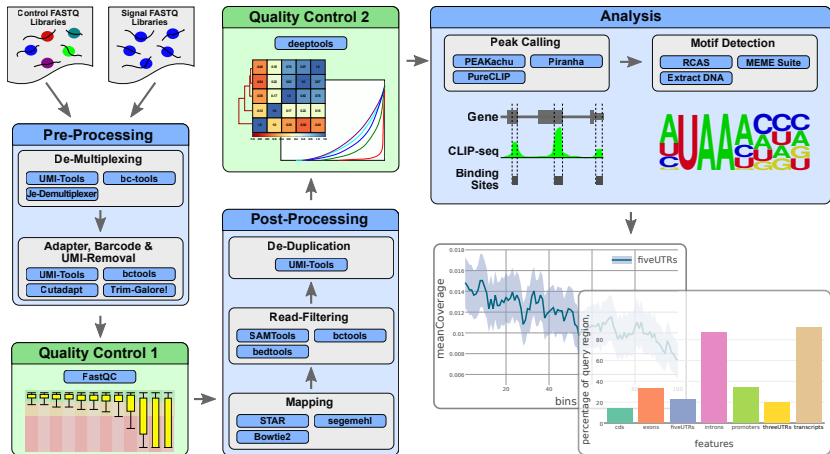
1.3 CLIP-Explorer



1.3 CLIP-Explorer



1.3 CLIP-Explorer



1.4 CLIP-Explorer Available Under

History Options

Send results to a new history

Yes No

1: Background

No data dataset collection available.

2: Paired-end reads collection

No data dataset collection available.

3: Annotation Reference File for RCAS

No GTF dataset available.

<https://clipseq.usegalaxy.eu/>

1.5 Galaxy Training Material for CLIP-Explorer



Galaxy Training!

Transcriptomics

Introduction slides

Galaxy Instances

Input Dataset

Literature EN

Help

Edit

CLIP-Seq data analysis from pre-processing to motif detection

Overview

Questions

- How is raw CLIP-Seq data processed and analysed?
- How do I find binding motifs and targets for a protein (e.g., RBFOX2)?

Objectives

- Remove Adapters, Barcodes and Unique Molecular Identifiers (UMIs) from the reads
- Align trimmed reads with STAR
- De-duplicate the read library
- Inspect the read mapping and de-duplication quality
- Perform peak calling with PEAKachu
- Analyse the peaks and find potential binding motifs and targets
- Check the quality of the peak calling

Requirements

- [Introduction to Galaxy Analyses](#)
- [Sequence analysis](#)
 - Quality Control: [slides](#) - [hands-on](#)
 - Mapping: [slides](#) - [hands-on](#)

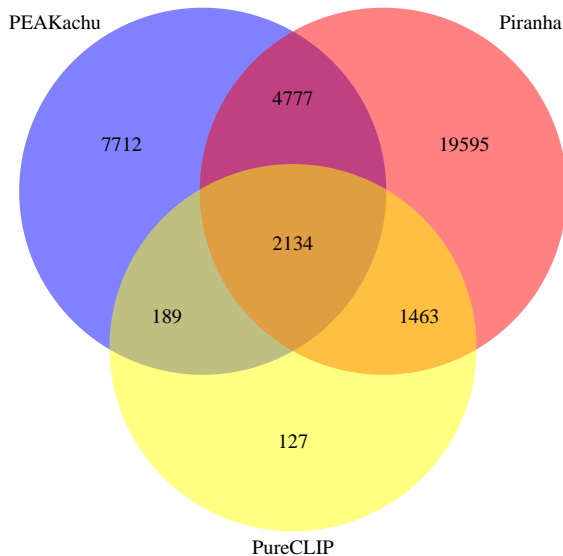
Time estimation: 6 hours

<https://galaxyproject.github.io/training-material/topics/transcriptomics/tutorials/clipseq>

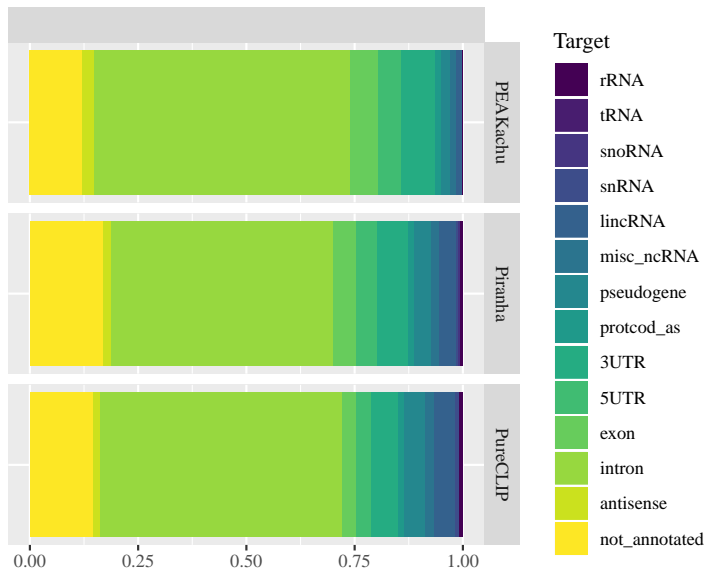
1.6 Analysis of RBFOX2

- ▶ Mainly binds to introns adjacent to differentially spliced exons
- ▶ Development and tissue-specific splicing factor
- ▶ Main motif: TGCATG

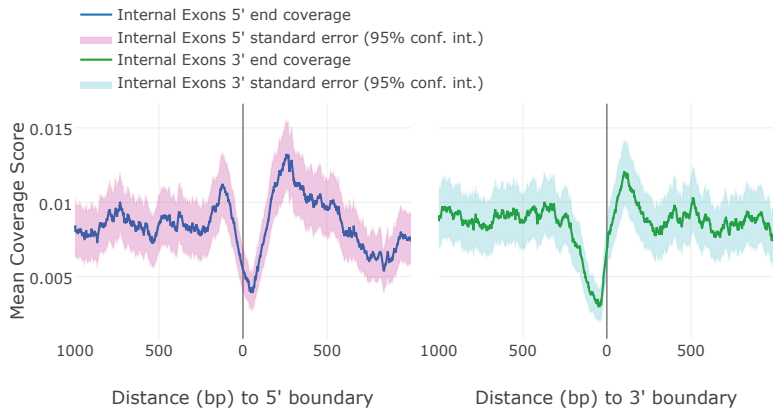
1.6 Analysis of RBFOX2



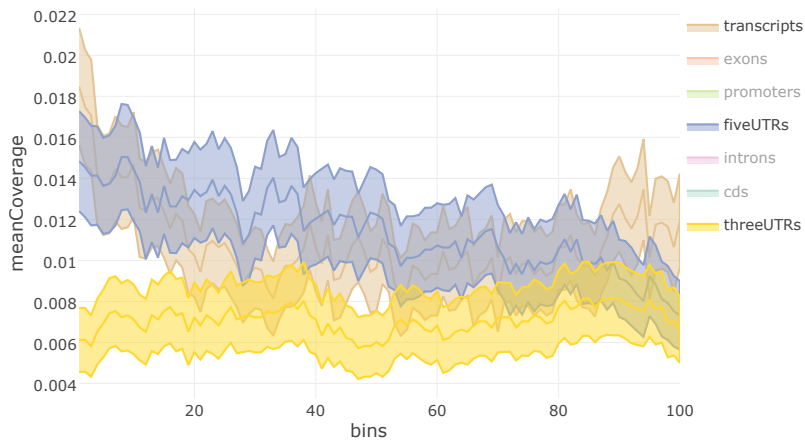
1.6 Analysis of RBFOX2



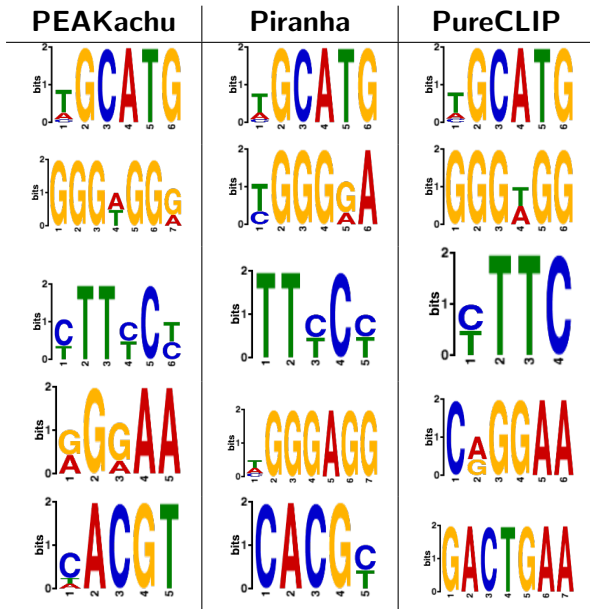
1.6 Analysis of RBFOX2



1.6 Analysis of RBFOX2



1.6 Analysis of RBFOX2



2. THANKS

Thanks to my group:

- ▶ Rolf Backofen
- ▶ Daniel Maticzka := Florian Eggenhofer
- ▶ Galaxy team
- ▶ Everybody Else

Thanks to:

