

# ChIP-Seq Analysis

Jessica Murphy

April 9, 2019

## Overview

The following data is from a transcription factor ChIP-Seq experiment that looked at DNA binding to androgen receptors (AR). ChIP-seq was performed on two biological replicates from a human bladder cancer cell line. A control ChIP(with IgG)-Seq was performed as well.

## Peak Calling

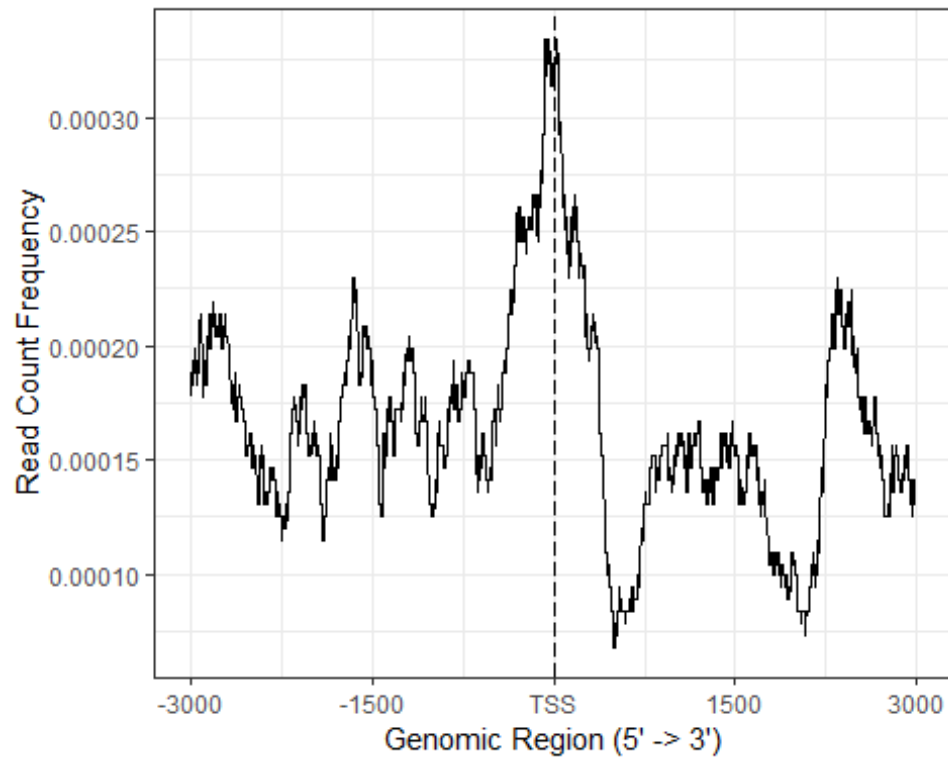
Peaks were called for the two ChIP samples to determine which areas of the genome were interacting/binding with the receptors. MACS2 was used with the following code.

```
macs2 callpeak -t "/BIOS6660/Homework8/data/chip1.markdup.bam" -c  
"/BIOS6660/Homework8/data/input.markdup.bam" -g hs -n chip1_peaks -f bam --  
keep-dup auto --bdg macs2 callpeak -t  
"/BIOS6660/Homework8/data/chip2.markdup.bam" -c  
"/BIOS6660/Homework8/data/input.markdup.bam" -g hs -n chip2_peaks -f bam --  
keep-dup auto --bdg
```

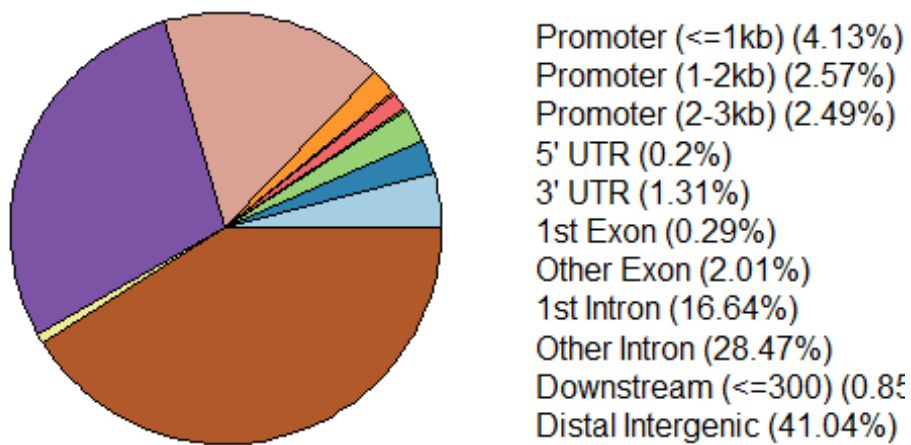
The results are located on Yampa under: /BIOS6660/Homework8/data/

## Peak Annotation

The promoter regions were determined and the distance between the peaks and the promoter regions are shown in the graph below.

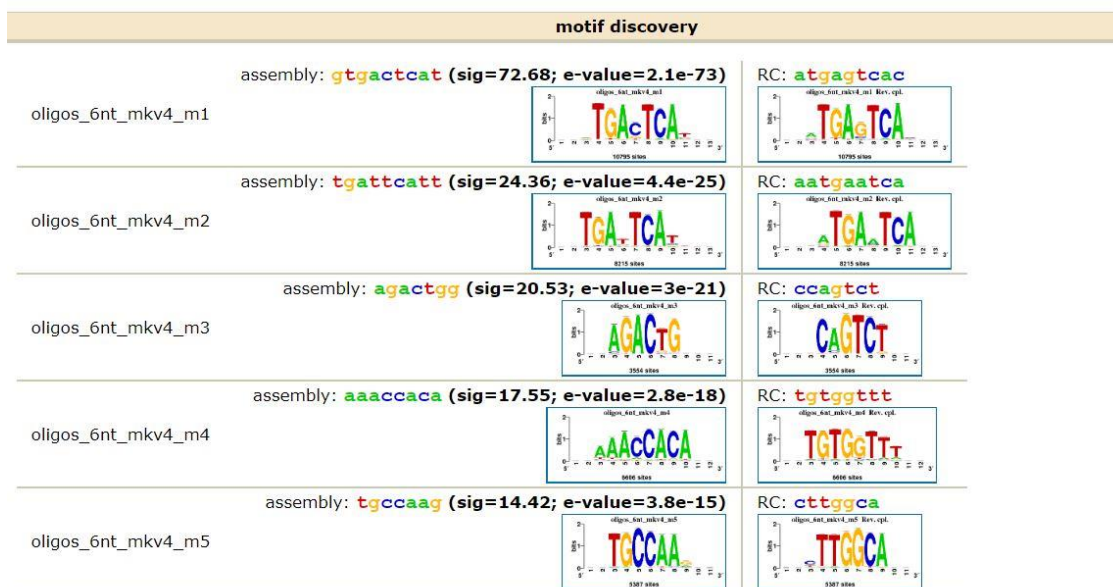


The peaks were then annotated and the results can be seen below.



## Motif Analysis

RSAT was used to identify which motifs were enriched based on these peaks. The results can be found [here](#). The top five motifs are shown below.



## Peak Integration

The resulting peaks were integrated with gene expression candidates by determining which candidates had a peak within 3,000 bp of the transcription start site (TSS). There were 240 differentially expressed candidates and 36 had a peak within 3,000 bp of the TSS.

## Enrichment Analysis

Enrichment analysis was then performed on these 36 genes using enrichR to determine if they are part of any known biological pathway or system. Using the GO Biological Process 2018 database, the top 10 processes with the smallest p-values are listed below.

Process	Overlap	P-value
cellular potassium ion homeostasis ( <a href="#">GO:0030007</a> )	2/13	0.000243
mitotic cell cycle arrest ( <a href="#">GO:0071850</a> )	2/14	0.000283
cellular monovalent inorganic cation homeostasis ( <a href="#">GO:0030004</a> )	2/17	0.000421
negative regulation of lymphocyte activation ( <a href="#">GO:0051250</a> )	2/19	0.000528
potassium ion homeostasis ( <a href="#">GO:0055075</a> )	2/25	0.000921
protein localization to lysosome ( <a href="#">GO:0061462</a> )	2/30	0.001328
regulation of sodium ion transport ( <a href="#">GO:0002028</a> )	2/35	0.001806
negative regulation of intracellular signal transduction ( <a href="#">GO:1902532</a> )	3/161	0.003008
regulation of apoptotic process ( <a href="#">GO:0042981</a> )	6/815	0.003075

negative regulation of mitotic cell cycle ([GO:0045930](#))

2/48 0.003373