# ENCODE and BedTools

## 1 Download data from ENCODE

Download .bed files for ATF3 and SRF in HCT116 cells from the ENCODE repository

1.1. Navigate to https://www.encodeproject.org/

1.2. Search for ”ATF3 HCT-116” ->Add to cart

1.3. Search for ”SRF HCT-116” ->Add to cart

1.4. View Cart

   (1.4.1) Output Type: peaks
   (1.4.2) File Type: bed broadPeak
   (1.4.3) Genome Assembly: hg19
   (1.4.4) Lab: Richard Myers, HAIB
   (1.4.5) Status: Released

   ---

   **Total: 4 bed files**

1.5. Download data

```
xargs -L 1 curl -O -L < files.txt
```

1.6. Gunzip files

```
gunzip *.gz
```

1.7. Rename bed files (info in metadata.tsv)

```
mv ENCFF001UDK.bed ATF3.rep1.bed
mv ENCFF001UDL.bed ATF3.rep2.bed
mv ENCFF001UEM.bed SRF.rep1.bed
mv ENCFF001UEN.bed SRF.rep2.bed
```

## 2 Intersect replicates

Intersect ChIP-Seq peak replicates to get high confidence peaks

2.1. Module load BedTools

```
module load bedtools
```

2.2. Intersect ATF3 replicates

```
bedtools intersect -a ATF3.rep1.bed -b ATF3.rep2.bed > ATF3.repinter.bed
```

2.3. Intersect SRF replicates

```
bedtools intersect -a SRF.rep1.bed -b SRF.rep2.bed > SRF.repinter.bed
```

## 3 Intersect high confidence ATF3 and SRF peaks

Intersect ATF3 and SRF peaks that are in both replicates

```
bedtools intersect -a ATF3.repinter.bed -b SRF.repinter.bed > ATF3_SRF.intersect.bed
```

## 4 Visualize on IGV

4.1. Open X2GO

4.2. Start IGV

4.3. Set genome to **hg19**

4.4. File ->Load from ENCODE (2012) ->Load HCT-116 ATF3 and SRF RawSignal

4.5. Import intersected .bed files that we generated

4.6. Scroll through regions using ctrl + F

## 5 Intersect shuffled ATF3 and SRF peaks

To get a qualitative sense of your background expectation, shuffle ATF3 and SRF peaks and intersect them

5.1. Shuffle ATF3 peaks

```
bedtools shuffle -i ATF3.repinter.bed -g hg19.sorted.genome > ATF3.shuffle.bed
```

5.2. Shuffle SRF peaks

```
bedtools shuffle -i SRF.repinter.bed -g hg19.sorted.genome > SRF.shuffle.bed
```

5.3. Intersect shuffled peaks

```
bedtools intersect -a ATF3.shuffle.bed -b SRF.shuffle.bed > ATF3_SRF.shuffled.bed
```

# 6 Quantify intersection statistical significance

6.1. Measure Jaccard Statistic

```
bedtools jaccard -a ATF3.repinter.bed -b SRF.repinter.bed
```

6.2. Perform Fisher's exact test

```
bedtools fisher -a ATF3.repinter.bed -b SRF.repinter.bed -g hg19.sorted.genome
```

# 7 Optional: Merge high confidence ATF3 and SRF peaks

7.1. Method 1: Step-wise with intermediate files

```
cat ATF3.repinter.bed SRF.repinter.bed > ATF3_SRF.cat.bed
bedtools sort -i ATF3_SRF.cat.bed > ATF3_SRF.cat.sorted.bed
bedtools merge -i ATF3_SRF.cat.sorted.bed > ATF3_SRF.cat.sorted.merged.bed
```

7.2. Method 2: Using pipes

```
cat ATF3.repinter.bed SRF.repinter.bed | bedtools sort -i stdin | \
bedtools merge -i stdin > ATF3_SRF.merged.bed
```