Example pipeline

***1. Run STAR (build genome first)***

STAR --genomeDir /path/to/star/genome --runThreadN 4 --readFilesIn paired\_reads\_1.fastq paired\_reads\_2.fastq --outSAMtype BAM SortedByCoordinate --outWigType wiggle --outWigNorm None -–outSAMunmapped Within --outFileNamePrefix aligned\_reads

*This command will generate the following files:*

Contains both mapped and unmapped reads. Unmapped reads need to be extracted as part of the find\_circ pipeline to detect non-canonical backsplicing

# All reads in the form of a BAM file #

\* aligned\_reads\_STARAligned.sortedByCoord.out.bam

# Uniquely mapped reads only #

\* aligned\_reads\_STARSignal.Unique.str1.out.wig # (-) strand coverage

\* aligned\_reads\_STARSignal.Unique.str2.out.wig # (+) strand coverage

# Multi-mapped and uniquely mapped reads #

\* aligned\_reads\_STARSignal.UniqueMultiple.str1.out.wig # (-) strand coverage

\* aligned\_reads\_STARSignal.UniqueMultiple.str2.out.wig # (+) strand coverage

# Splice junction counts #

\* aligned\_reads\_STARSJ\_out.tab # Canonical splice junction counts file needs to

# be converted to bed format appropriate for

# circleVis processing

2. Run star\_sj\_convert on canonical splice junction file

star\_sj\_convert aligned\_reads\_STARSJ\_out.tab

*This command will generate the following file:*

\* aligned\_reads\_STARSJ\_out.tab.canonical.bed # Final canonical splice junction

# file

3. Run find\_circ to generate backsplice junction calls

samtools view -hf 4 aligned\_reads\_STARAligned.sortedByCoord.out.bam | samtools view -Sb - | ./unmapped2anchors.py aligned\_unmapped\_reads.bam > aligned\_unmapped\_reads\_anchors.fastq

mkdir -p circles\_out

STAR --genomeDir /path/to/genome --readFilesIn aligned\_unmapped\_reads\_anchors.fastq |find\_circ.py --genome=/path/to/fasta --prefix=canonical --name=my\_sample --stats=circles\_out/my\_sample\_stats.txt > circles\_out/my\_sample\_splice\_sites.bed

4. Run find\_circ\_convert on find\_circ junction file

find\_circ\_convert my\_sample\_splice\_sites.bed

*This command will generate the following file:*

\* my\_sample\_splice\_sites.bed.circles.bed # Final backsplice junction file

# Uniquely mapped coverage will be used for this example, but it is up to the user to choose unique or unique + multi #

5. Using a GTF file with the same chromosome names as the genome used for your alignement (i.e. if chromosome 1 is labeled ‘chr1’ in the STAR genome, it must also be labeled ‘chr1’ in the GTF file you use. It cannot be ‘chr1’ in one file and ‘1’ in the other.), run circbuild.

circbuild –-gtf Homo\_sapiens.GRCh38.92.gtf --wigneg aligned\_reads\_STARSignal.Unique.str1.out.wig --wigpos aligned\_reads\_STARSignal.Unique.str2.out.wig –-splicejunction aligned\_reads\_STARSJ\_out.tab.canonical.bed --circlejunction my\_sample\_splice\_sites.bed.circles.bed --output my\_sample\_name

*This command will generate the following file:*

\* my\_sample\_name.db

6. Run circplot using your gene of interest

circplot --database my\_sample\_name.db --gene EGFR

*This command will generate the following files:*

#Open either one of them with any web browser

\* EGFR\_ENST00000275493.svg

\* EGFR\_ENST00000275493.html