# Computing Coverage with igvtools

In this exercise we will compute total and strand-specific coverage from RNA-Seq alignments. The coverage files will be examined visually in IGV, and strand information compare with known annotations.

## Data

Data for this tutorial was provided by Brian Haas can be found in the "data/rna" directory. This dataset was derived from S. Pombe samples and includes RNA-Seq strand-specific paired-end reads, a genome fasta reference file, annotations, and files from Tophat and Cufflinks analysis.

## Install igvtools

To install the latest version of igvtools from www.broadinstitute.org/igv/download and save it to your home directory. (Note: its not necessary to install igvtools here, but the remainder of the instructions assume this location)

**"cd" to your home directory (just type cd <enter>)**

Unzip the archive.

**unzip igvtools_2.3.37.zip**

Verify the installation by executing the following

**~/IGVTools/igvtools**

You should see a version and usage string followed by a summary of commands.

Now add igvtools to your path environment variable

**set path = ( $path ~/IGVTools )**

## Compute total and strand-specific coverage

Compute a total coverage file in "tdf" format as follows.

### "cd" to the data/rna directory

execute the following

**igvtools count Sp_ds.bam Sp_ds.bam.tdf genome.fa**

This should produce a file called "Sp_ds.bam.tdf" containing coverage data for the entire bam.

Now we'll compute a strand-specific coverage file. This is possible because the RNA library is strand preserving. Specifically, for this library, the strand of the first read of each pair corresponds to the strand of the fragment sequenced. We'll use the –strands arugment to separate the coverage count by strand

**igvtools count --strands first Sp_ds.bam Sp_ds.bam.strands.tdf genome.fa**

## View the coverage files in IGV

Launch igv

Load the custom genome and annotations

- From the menu "Genomes > Load Genome from File"… select "genome.fa" in the data/rna folder.
- From the menu "File > Load from File…" select "genes.bed" in the data/rna folder

Load the bam file "Sp_ds.bam"

Note: the total coverage "tdf" file is loaded automatically.

Load the strand specific coverage file "Sp_ds.bam.strands.tdf"

Zoom in on a peak until alignments are shown

Color alignments by strand by right-clicking in the alignment track, then selecting “Color alignments by > first-of-pair strand”.

Verify that coverage tracks and alignments correspond with the known strands of the gene annotations. This verifies that the rna-seq data is indeed strand specific.