

Tophat-Cufflinks-Cuffdiff, ignoring novel transcripts

1. Align reads from each sample to the reference genome using tophat

For Illumina/basespace data:

```
nohup tophat -p 4 -r <mate-inner-distance> -G <gfffile> <bowtie_index_prefix> <R1.fastq> <R2.fastq> &>tophat.log &
```

For ABI SOLiD (colospace) data:

```
nohup tophat --bowtie1 -C -p 4 -r <mate-inner-distance> -G <gfffile> <bowtie_colospace_index_prefix> <R1.fasta> <R2.fasta> <R1.qual> <R2.qual> &>tophat.log &
```

The output file (accepted_hits.bam) is the alignment output which will be used in following steps.

2. Identify differentially expressed transcripts using cuffdiff

If you have more than one replicate for a sample, supply the SAM files for the sample as a single comma-separated list.

```
nohup cuffdiff -o <outputdirectory> <gfffile> <sample1_accepted_hits.bam> <sample2_accepted_hits.bam> <sample3_accepted_hits.bam> &cuffdiff.log &
```

Several output files, consisting of raw and normalized counts for genes, isoforms and transcription start sites are generated. More about the output files at [cufflinks manual](#)