

Tutorial - Start diploid mapping for Day 2

Prepare for Day 2

To prepare for Day 2, try to start either BWA or Bowtie2 mapping of these two fastq files:

The data

```
$BI/ngs_course/human_variation/allseqs_R1.fastq
$BI/ngs_course/human_variation/allseqs_R2.fastq
```

against this reference:

The reference

```
$BI/ngs_course/human_variation/ref/hs37d5.fa
```

followed by converting the output to bam format, sorting the bam file, and indexing the output.

For either BWA or Bowtie, you need an indexed reference generated by either `bwa index` or `bowtie-build`. Perhaps one already exists? You then call the alignment program(s) `bwa aln/bwa sampe` or `bowtie`. Conversion, sorting, and indexing of the output are all done via `samtools`.



Submit to the TACC queue

Move into your scratch directory and create a new directory:

```
cds
mkdir day2
cd day2
```

Then create a `commands` file with the commands below and use ***launcher_creator.py*** followed by ***qsub*** to run them.

1. Mapping

a. For BWA, the commands are:

ONE LINE command for ALL bwa operations

```
module load bwa; bwa aln $BI/ngs_course/human_variation/ref/hs37d5.fa $BI/ngs_course
/human_variation/allseqs_R1.fastq > r1.sai && bwa aln $BI/ngs_course/human_variation/ref/hs37d5.fa
$BI/ngs_course/human_variation/allseqs_R2.fastq > r2.sai && bwa sampe $BI/ngs_course
/human_variation/ref/hs37d5.fa r1.sai r2.sai $BI/ngs_course/human_variation/allseqs_R1.fastq $BI
/ngs_course/human_variation/allseqs_R1.fastq > test.sam
```

b. For bowtie2:

One-line command for bowtie

```
module load bowtie/2.0.0b6; bowtie2 -t -x $BI/ngs_course/human_variation/ref/hs37d5_bowtie2 -l $BI
/ngs_course/human_variation/allseqs_R1.fastq -2 $BI/ngs_course/human_variation/allseqs_R2.fastq -S
test.sam
```

2. Convert, sort, and index output:

One-line command for samtools

```
samtools view -S -b test.sam > test.bam && samtools sort test.bam test.sorted && samtools index test.sorted.bam
```

For a bit of extra challenge, you could make two separate `commands` files, one for each mapper, qsub them. Watch out - you'll need to make sure they both write output to different files, or run in different directories!