

1. Index the ENCODE reference genome using LongRanger version 2.1.2

```
$ longranger mkref GRCh38_no_alt_analysis_set_GCA_000001405.15.fasta
```

This create a directory of special indexed files which is described here:

<http://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/references>

We used this version, but it turns out their software doesn't actually support using any alternate haplotypes.

2. For each sample, align reads, call and phase variants using LongRanger version 2.0.0

```
$ longranger-2.1.2/longranger-cs/2.1.2/bin/wgs --id=sge_Aug_encodev2_2_local --sex=$SEX --  
fastqs=$folder --sample=mysample --reference=$REF --maxjobs=10 --jobinterval=1000
```

SEX = male or female, as appropriate

FASTQ = directory to the FASTQ files

REF=ENCODE reference from step 1

This creates the phased_variants.vcf file, the phased_posorted_bam.bam file, and the loupe.loupe files that we posted on the website. I would refer to the LongRanger website for all the details: <http://support.10xgenomics.com/genome-exome/software/pipelines/latest/what-is-long-ranger>