Software needed:
bowtie2
samtools

1. Build database for reference genome

You only need to do this once per reference genome. I put all my assemblies into a folder called “references” so I can easily access them with minimal change to the script

```
bowtie2-build YWAR_min1000_sm.fasta YWARv0
```

2. Map duplicate filtered samples to genome assembly

We convert the files directly to bam and sort them with sam tools in order to save space. While mapping, we also add read groups that identify each sample as well as the library and lane. This information is used by GATK later.

```
for sample in `ls *.1.1.fq.gz | cut -f1 -d'.'`
do
     bowtie2 -x References/YWARv0/YWARv0 \
     --threads 8 -1 $sample.1.1.fq.gz -2 $sample.2.2.fq.gz \
     --rg-id $sample --rg SM:$sample --rg ID:$sample \
     --rg LB:YWAR1 --rg PU:YWAR1 --rg PL:illumina | \
     samtools view -bhS - | \
     samtools sort - $sample
done
```

Alignment rates will be output as stdout. It’s good to pay attention to these at this step, especially if you are using a genome from another species. You may want to fine tune the parameters in bowtie a bit.

NOTE: My script for this is RADmap.sh