Filter out Red Alder Sequencing Reads

In order to focus on sequencing reads from the microbes in the nodule, we will filter out reads that align to the red alder genome as follows:
1) Align the fastq-formatted reads to the red alder genome using minimap2.
2) Extract reads that do not align to red alder and sort them using samtools.
3) Create a fastq file with only the unaligned reads using samtools bam2fastq.
4) Compress the fastq file using gzip.

Activate the environment that contains minimap2 and samtools

```
conda activate seqtools
```

Make a directory and go into it

```
mkdir ~/microbe_fastq
cd ~/microbe_fastq
```

Link to the merged minion reads

```
ln -s /home/cjb/minion/2022/data/minit4/data/AlderNodule3469-3/3469-3/20220701_2144_MC-
113445_FAS21661_134c02ac/fastq_pass/all.fastq 3469-3.all.fastq
```

Run Minimap2 to align the MinION reads to the red alder genome

The -x map-out parameter (allows ~10% error + divergence)

```
minimap2 -x asm20 -L -t 8 -a \
    /home/cjb/minion/2019/indexes/red_alder_genome/consensus.fasta \
    3469-3.all.fastq > 3469-3-minionxredalder.mm2.sam
```

Convert the unmapped reads in the alignment file (sam) to a fastq file

The -f4 includes only reads with the 4 flag (unmapped)

```
samtools fastq -f4 3469-3-minionxredalder.mm2.sam > 3469-3.microbe.fq
```

Compress the new fastq file (note that it will automatically add the extension .gz)

```
gzip 3469-3.microbe.fq
```

Now run these steps with 4956-3 (minit3)

Reads are here: /home/cjb/minion/2022/data/minit3/data/aldernodule4956-3/4956-3/20220701_2154_MC-113286_FAS37509_f3119554/fastq_pass/all.fastq