

Genome Indexing for aligning HiC files

Bowtie2 v2.5.1 was used for this script. We used two files BY4742_SRR6823436.contigs.fasta and SY14.contigs.fasta as the reference genomes for BY4742 and SY14. Both of these files were made during the Genome Assembling process.

```
#!/bin/bash -l
#SBATCH --job-name=hic_build_indices
#SBATCH --output=/scratch/grp/msc_appbio/Group2_ABCC/HiC/logs/%x_%j.out
#SBATCH --error=/scratch/grp/msc_appbio/Group2_ABCC/HiC/logs/%x_%j.err
#SBATCH --time=02:00:00
#SBATCH --mem=8G
#SBATCH --cpus-per-task=4
#SBATCH -p msc_appbio
# Building indices to align efficiently, otherwise mapping would be too slow or inaccurate. Without
indices the mapping step cannot run.
#loading module
module load bowtie2
BASE=/scratch/grp/msc_appbio/Group2_ABCC/HiC
REF=$BASE/references
IDX=$BASE/results/indices
cd $REF
#BY4742_SRR6823436.contigs.fasta (We chose one of the two assembled runs for the BY4742 assembled
reference genome.
echo "Building index for BY4742_SRR6823436.contigs.fasta"
bowtie2-build --threads 4 \
BY4742_SRR6823436.contigs.fasta \
${IDX}/BY4742_SRR6823436
#Using the renamed SY14.contigs.fasta for the SY14 assembled reference genome.
echo "Building index for SY14.fa"
bowtie2-build --threads 4 \
SY14.fa \
${IDX}/SY14
echo "Done building indices."
```