

7.8 Pilon Polishing - Round 2 (Script 10)

Script: 10_pilon2

```
#!/bin/bash
#SBATCH --job-name=pilon_SY14
#SBATCH --output=/scratch/grp/msc_appbio/Group2_ABCC/Gene_Assembling/logs/pilon_SY14_%j.out
#SBATCH --error=/scratch/grp/msc_appbio/Group2_ABCC/Gene_Assembling/logs/pilon_SY14_%j.err
#SBATCH --cpus-per-task=4
#SBATCH --mem=120G
#SBATCH --time=8:00:00
#SBATCH --partition=msc_appbio

# Load anaconda module
module load anaconda3/2022.10-gcc-13.2.0

# Activate pilon environment
source activate /scratch/grp/msc_appbio/Group2_ABCC/conda_envs/pilon_env

# Set Java memory
export _JAVA_OPTIONS="-Xmx100G"

# Run Pilon
pilon --genome SY14.contigs.fasta \
    --frags SY14_illumina_mapped.bam \
    --output SY14_polished \
    --changes

echo "Pilon polishing completed!"
echo "Number of contigs:"
grep -c ">" SY14_polished.fasta
echo "Assembly size:"
grep -v ">" SY14_polished.fasta | tr -d '\n' | wc -c
echo "Corrections made:"
wc -l SY14_polished.changes
```

Issue: Java OutOfMemoryError despite allocating 100 GB space. Pilon polishing could not complete due to insufficient memory for processing the genome assembly with high-coverage Illumina data.