**Bacterial Genome Assembly Using Short and Long Reads**

**Retrieving the sequencing reads**

# Download the long reads

wget -q https://zenodo.org/record/10669812/files/DRR187567.fastq.bz2 && bzcat

DRR187567.fastq.bz2 > long\_reads.fastq

# Download the paired end short reads

wget -q https://zenodo.org/record/10669812/files/DRR187559\_1.fastqsanger.bz2 && bzcat DRR187559\_1.fastqsanger.bz2 > illumina\_1.fastq

wget -q https://zenodo.org/record/10669812/files/DRR187559\_2.fastqsanger.bz2 && bzcat DRR187559\_2.fastqsanger.bz2 > illumina\_2.fastq

#**Quality control of raw reads and trimming**

fastqc -t 3 \*.fastq 2> fastqc.log && multiqc .

**#Trim the short reads**

fastp -i illumina\_1.fastq -I illumina\_2.fastq -f 15 -F 15 -t 2 -T 2 -w 4 -o illumina\_1.trimmed.fastq -O illumina\_2.trimmed.fastq

**#Trim the long reads**

filtlong --min\_length 1000 -1 illumina\_1.trimmed.fastq -2 illumina\_2.trimmed.fastq long\_reads.fastq > long\_reads.filtered.fastq

#**Assembly of long reads**

flye --nano-corr long\_reads.filtered.fastq -t 4 -o flye\_Saureus

**#Visualize the assembly graph**

**Bandage image flye\_Saureus/assembly\_graph.gfa flye\_Saureus/assembly\_graph.png --width 800 --height 800**

**#Polishing the assembly with short reads**

**BWA-MEM2 Indexing**

bwa-mem2 index flye\_Saureus/assembly.fasta 2> polishing\_indexing.log

**BWA-MEM2 Mapping**

bwa-mem2 mem -t 4 -o Saureus\_polishing\_map.sam flye\_Saureus/assembly.fasta illumina\_1.trimmed.fastq illumina\_2.trimmed.fastq 2> polishing\_mapping.log

**Polypolish Polishing**

polypolish polish flye\_Saureus/assembly.fasta Saureus\_polishing\_map.sam > Saureus\_polished\_assembly.fasta

**#Evaluating the assembly**

**busco -i Saureus\_polished\_assembly.fasta -l bacteria -o Saureus\_BUSCO -m geno -c 4 -f > busco.log**

**#Annotation of the assembled genome**

**prokka Saureus\_polished\_assembly.fasta --genus Staphylococcus --species aureus --cpus 4 --outdir Saureus\_PROKKA --force 2> prokka.log**