

Quality Control and Genome Assembly

Escherichia coli _MODI - EC4326 (SRR3951549)

In Our Study of an Escherichia Coli sample related to infectious disease, we utilized the BV-BRC genome pipeline on the illumina platform. To ensure high-quality assembly results, we employed the Unicycler v0.4.8 assembler alongside samtools v1.13 (which incorporates htslib 1.13+ds). This combination facilitated the refinement of base calls, thorough quality checks, data trimming, correction of missassemblies, and gap filling within the genome. The findings obtained include:

Total length of good contigs	5,319,720
Min contig_length	300
Largest contig	374,756
GC (Guanine and Cytosine)%	50.28
N50	146,357
N90	24,885
L50	11
L90	42
Assembly_Time	0.78 hours

Escherichia coli _SCHI0016.S.230 (SRR11094155)

For the genome assembly of this second sample, we again used the Unicycler v0.4.8 assembler to optimize outcomes. This process involved correcting bases, conducting detailed quality assessments, and addressing any misassemblies while ensuring gaps were filled appropriately. We evaluated the quality of these assemblies using QUAST (v 5.2.0), which helped in verifying the accuracy and completeness of the results. Additional findings generated include:

Total length of good contigs	4,880,275
Min contig_length	300
Largest contig	779,948
GC (Guanine and Cytosine)%	50.70
N50	359,895
N90	74,690
L50	5
L90	16
Assembly_Time	0.36 hours