2.1 Run Duration

Start Time (UTC) {date & time}	Finish Time (UTC) 2023-08-08 10:38:09.882861

2.2 Samples filtered out from the analysis

Note - to pass the FastQC filter samples must have a percentage GC content within a specified range (>= 25 and <= 75%), a minimum average base quality (default 20), a minimum sequence length (default 70nt) and a minimum number of sequences (default 500).

If the table below is blank, all samples have passed filtering.

Sample name	Problem

2.3 Sequence and Contig counts

Pipeline step	Mean (SEM)	Percentage (SEM)
Starting Sequences	110000.0 (10000.00)	100
Sequences remaining after quality filtering (fastp)	108416.9 (9858.01)	98.56
Sequences remaining after human decontamination (Bowtie2)	53126.1 (5332.49)	48.30

2.4 Metatranscriptome coverage

Coverage (%)
39.40%

2.5 Output files

Metatranscriptome and Quality Assessment

thinned_metatranscriptome_rep_seq.fasta: This is the thinned version of the metatranscriptome produced by RNA-Bloom.

short_summary.specific.bacteria_odb10.busco_output.txt: The report from BUSCO provides completeness assessment by employing sets of Benchmarking Universal Single-Copy Orthologs from OrthoDB.

quast_report.tsv: This report will show the number of contigs generated from RNA-Bloom and the N50. N50 statistic defines assembly quality in terms of contiguity. In a set of contigs, the N50 is defined as the sequence length of the shortest contig at 50% of the total assembly length.

Functional Annotation Abundance

isoforms_functional_abundance_EC_TPM.tsv: Functional abundance matrix using TPM counts for EC annotations.

isoforms_functional_abundance_EC_expected_count.tsv: Functional abundance matrix using raw counts for EC annotations.

isoforms_functional_abundance_GOs_TPM.tsv: Functional abundance matrix using TPM counts for GO annotations.