bit metagenomics workflow

Raw data fastQC
• fastqc/multiqc

→ Quality filter/trim
• bbduk.sh

→ Filtered data fastQC
• fastqc/multiqc

→ Assembly
• Megahit

Binning of contigs
• MetaBat2

← Generate coverage information (mapping)
• bowtie2

Taxonomic classification
• CAT

Quality assessment
• checkm

→ MAG taxonomic classification
• GTDB-tk

Gene prediction
• prodigal

→ Functional annotation
• KOFamScan

MAG functional summarization
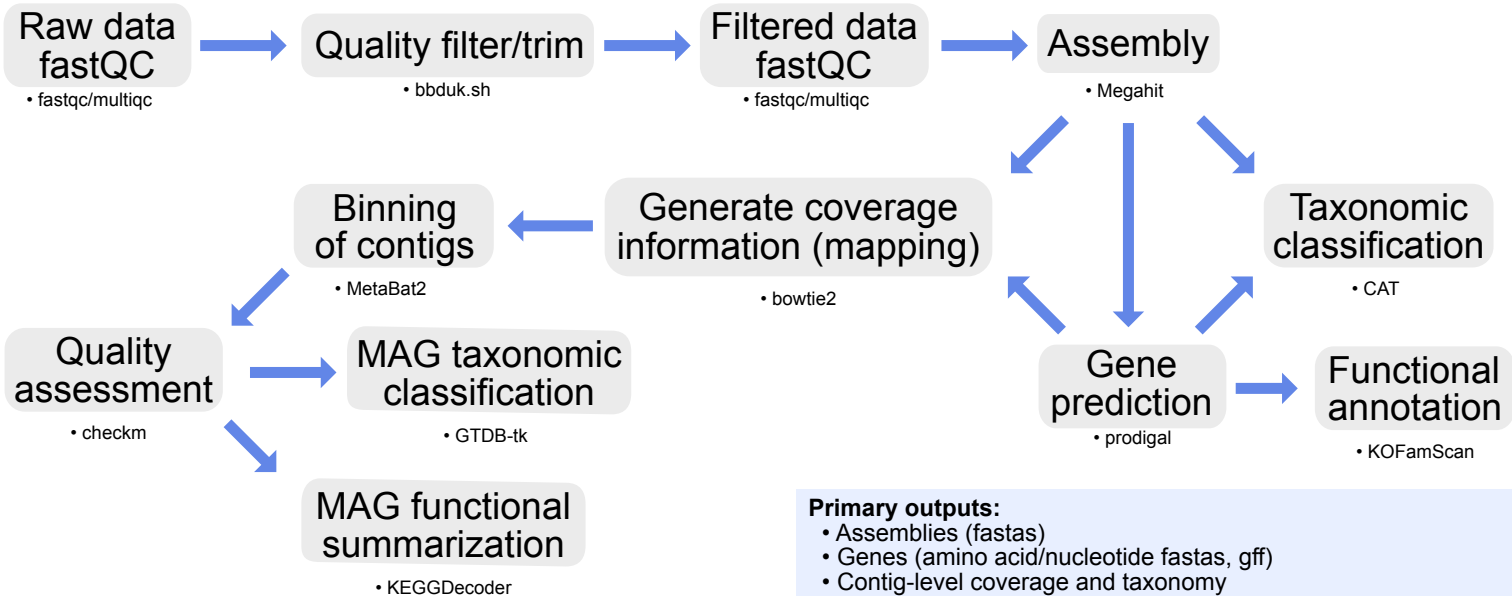• KEGGDecoder

Primary outputs:
- Assemblies (fastas)
- Genes (amino acid/nucleotide fastas, gff)
- Contig-level coverage and taxonomy
- Gene-level functional annotation, coverage, and taxonomy
- KEGG KO all-sample combined table with coverages
- Combined all-sample taxonomy table with coverages
- High-quality **M**etagenome-**A**ssembled **G**enomes (MAGs, fasta)