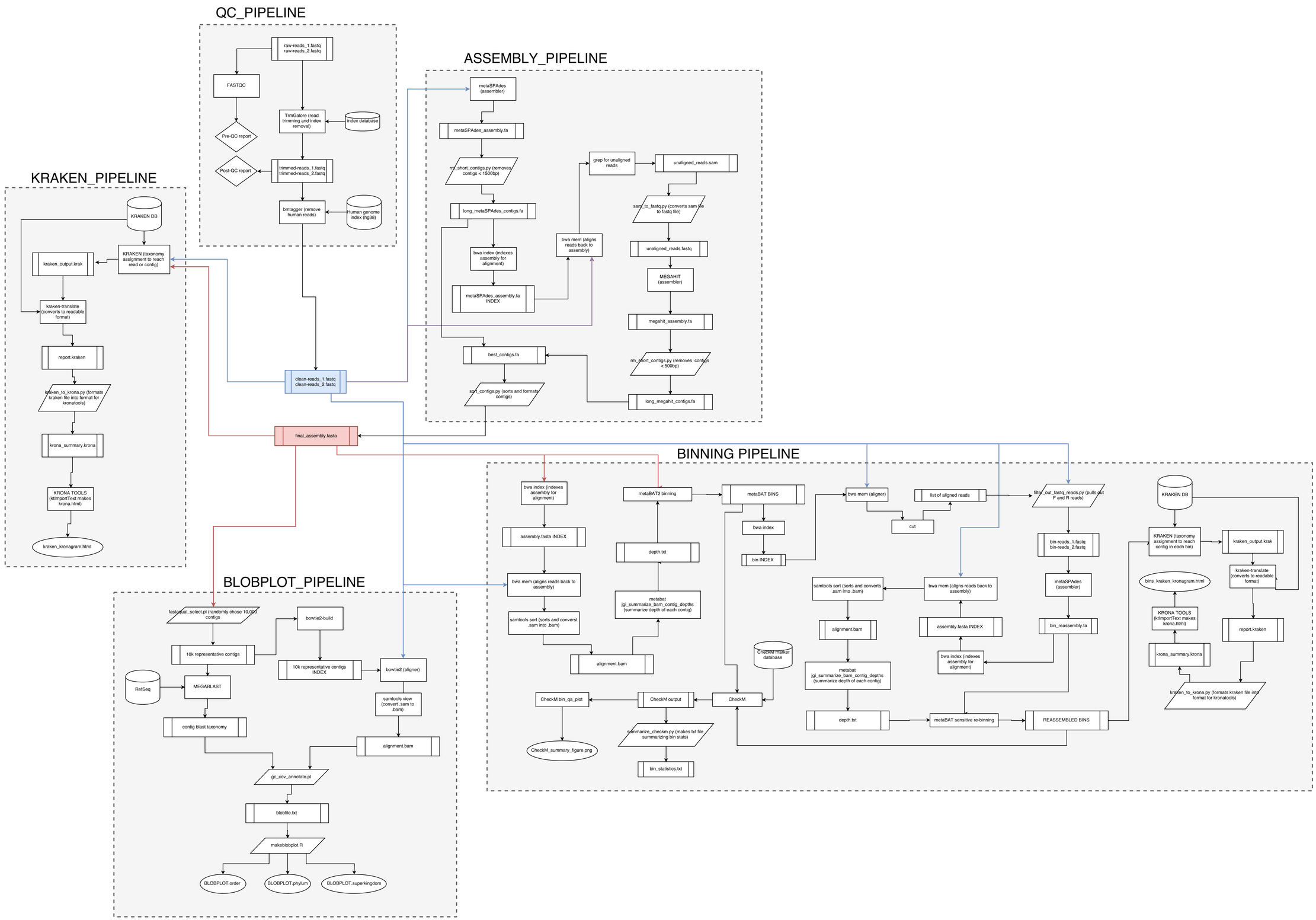
QC_PIPELINE
raw-reads_1.fastq
raw-reads_2.fastq
FASTQC
TrmGalore (read trimming and index removal)
index database
Pre-QC report
Post-QC report
trimmed-reads_1.fastq
trimmed-reads_2.fastq
bmtagger (remove human reads)
Human genome index (hg38)
ASSEMBLY_PIPELINE
metaSPAdes (assembler)
metaSPAdes_assembly.fa
rm_short_contigs.py (removes contigs < 1500bp)
long_metaSPAdes_contigs.fa
bwa index (indexes assembly for alignment)
metaSPAdes_assembly.fa INDEX
bwa mem (aligns reads back to assembly)
grep for unaligned reads
unaligned_reads.sam
sam_to_fastq.py (converts sam file to fastq file)
unaligned_reads.fastq
MEGAHIT (assembler)
megahit_assembly.fa
rm_short_contigs.py (removes contigs < 500bp)
long_megahit_contigs.fa
best_contigs.fa
sort_contigs.py (sorts and formats contigs)
KRAKEN_PIPELINE
KRAKEN DB
KRAKEN (taxonomy assignment to reach read or contig)
kraken_output.krak
kraken-translate (converts to readable format)
report.kraken
kraken_to_krona.py (formats kraken file into format for kronatools)
krona_summary.krona
KRONA TOOLS (ktImportText makes krona.html)
kraken_kronagram.html
clean-reads_1.fastq
clean-reads_2.fastq
final_assembly.fasta
BINNING PIPELINE
bwa index (indexes assembly for alignment)
assembly.fasta INDEX
bwa mem (aligns reads back to assembly)
samtools sort (sorts and converst .sam into .bam)
alignment.bam
metaBAT2 binning
depth.txt
metabat jgi_summarize_bam_contig_depths (summarize depth of each contig)
metaBAT BINS
bwa index
bin INDEX
bwa mem (aligner)
cut
list of aligned reads
filter_out_fastq_reads.py (pulls out F and R reads)
bin-reads_1.fastq
bin-reads_2.fastq
metaSPAdes (assembler)
bin_reassembly.fa
assembly.fasta INDEX
bwa index (indexes assembly for alignment)
bwa mem (aligns reads back to assembly)
samtools sort (sorts and converts .sam into .bam)
alignment.bam
metabat jgi_summarize_bam_contig_depths (summarize depth of each contig)
depth.txt
metaBAT sensitive re-binning
REASSEMBLED BINS
CheckM marker database
CheckM
CheckM output
CheckM bin_qa_plot
CheckM_summary_figure.png
summarize_checkm.py (makes txt file summarizing bin stats)
bin_statistics.txt
KRAKEN DB
KRAKEN (taxonomy assignment to reach contig in each bin)
kraken_output.krak
kraken-translate (converts to readable format)
report.kraken
kraken_to_krona.py (formats kraken file into format for kronatools)
krona_summary.krona
KRONA TOOLS (ktImportText makes krona.html)
bins_kraken_kronagram.html
BLOBPLOT_PIPELINE
fastaqual_select.pl (randomly chose 10,000 contigs)
10k representative contigs
bowtie2-build
10k representative contigs INDEX
bowtie2 (aligner)
samtools view (convert .sam to .bam)
alignment.bam
RefSeq
MEGABLAST
contig blast taxonomy
gc_cov_annotate.pl
blobfile.txt
makeblobplot.R
BLOBPLOT.order
BLOBPLOT.phylum
BLOBPLOT.superkingdom