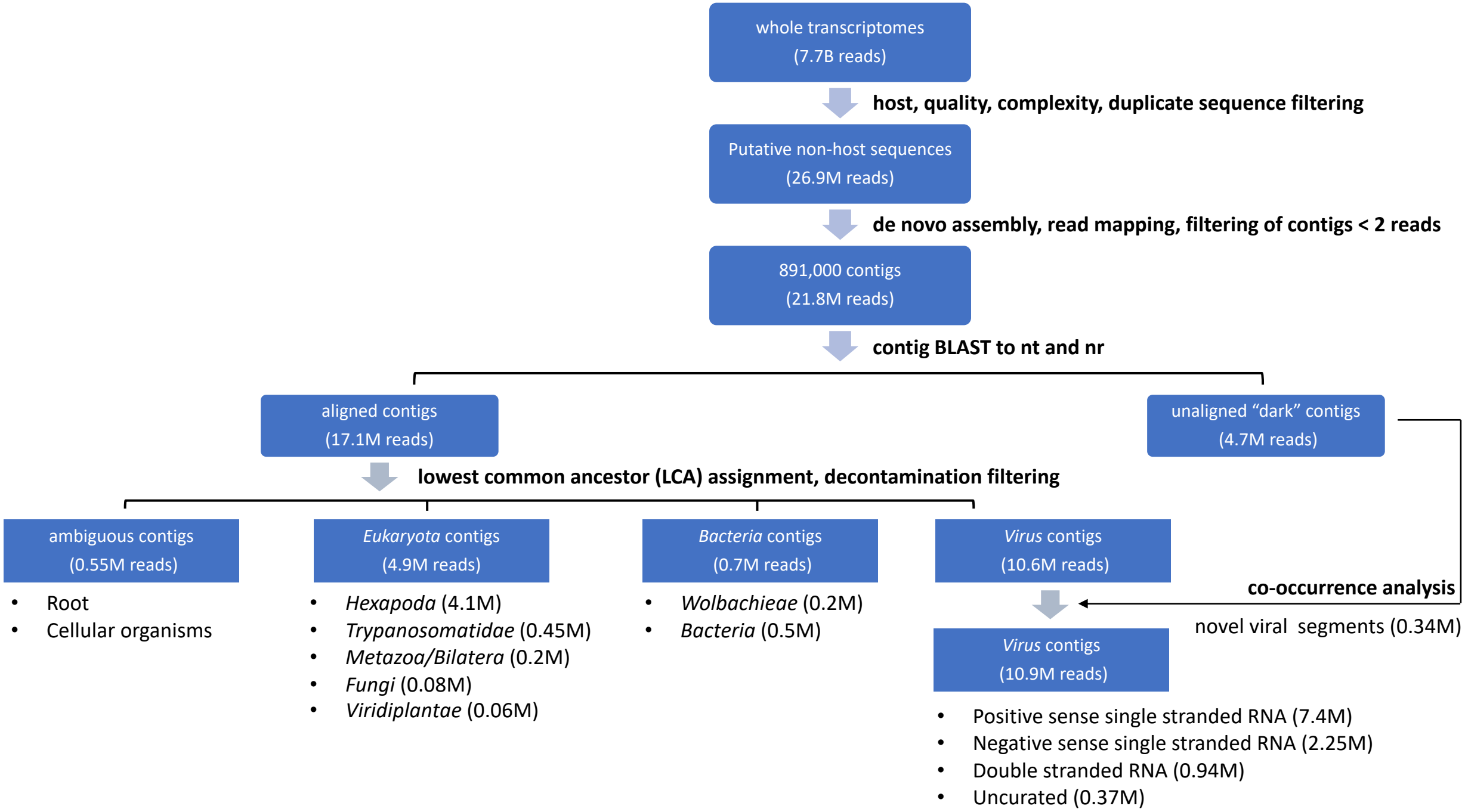
whole transcriptomes
(7.7B reads)
host, quality, complexity, duplicate sequence filtering
Putative non-host sequences
(26.9M reads)
de novo assembly, read mapping, filtering of contigs < 2 reads
891,000 contigs
(21.8M reads)
contig BLAST to nt and nr
aligned contigs
(17.1M reads)
unaligned "dark" contigs
(4.7M reads)
lowest common ancestor (LCA) assignment, decontamination filtering
ambiguous contigs
(0.55M reads)
• Root
• Cellular organisms
Eukaryota contigs
(4.9M reads)
• Hexapoda (4.1M)
• Trypanosomatidae (0.45M)
• Metazoa/Bilatera (0.2M)
• Fungi (0.08M)
• Viridiplantae (0.06M)
Bacteria contigs
(0.7M reads)
• Wolbachieae (0.2M)
• Bacteria (0.5M)
Virus contigs
(10.6M reads)
co-occurrence analysis
novel viral segments (0.34M)
Virus contigs
(10.9M reads)
• Positive sense single stranded RNA (7.4M)
• Negative sense single stranded RNA (2.25M)
• Double stranded RNA (0.94M)
• Uncurated (0.37M)