Raw sequence reads
297 million reads

After Quality Control, Read trimming, Adapter removal
264 million reads (88.89%)
Trimmomatic (Bolger et al. 2014), fastq-mcf (Aronesty 2011)

Removal of Aiptasia host, algal endosymbiont, PhiX, 28S rRNA sequences
8.6 million reads (2.89%)
BBsplit from BBmap (Bushnell 2016)

Paired-end read pairs based taxonomic classification
38,090 annotated sequences (bacteria, archaea, viruses)
CLARK (Ounit 2015)

Virus community: taxonomic composition and structure
3,293 sequences,
116 distinct viral taxa of 40 families
MetagenomeSeq (Paulson et al. 2013) & R package vegan (Oksanen et al. 2017)