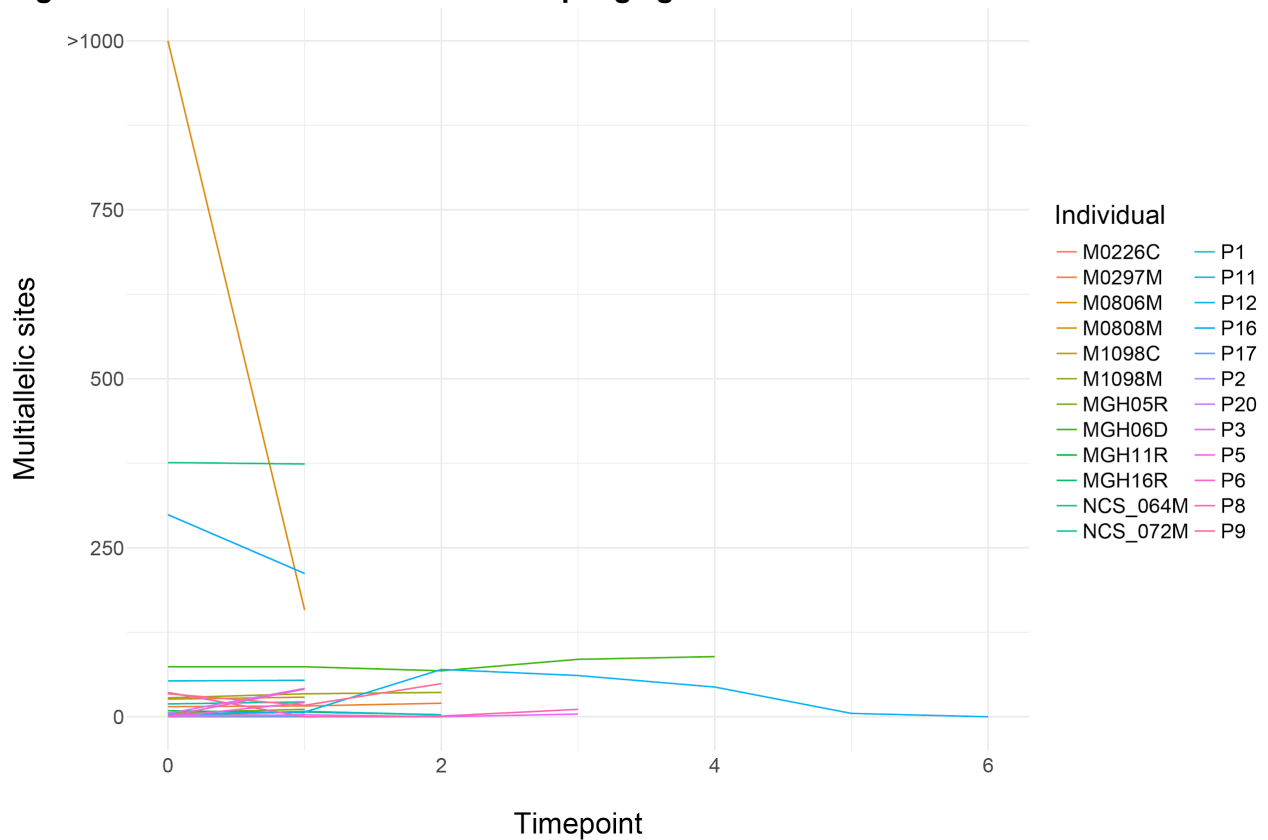


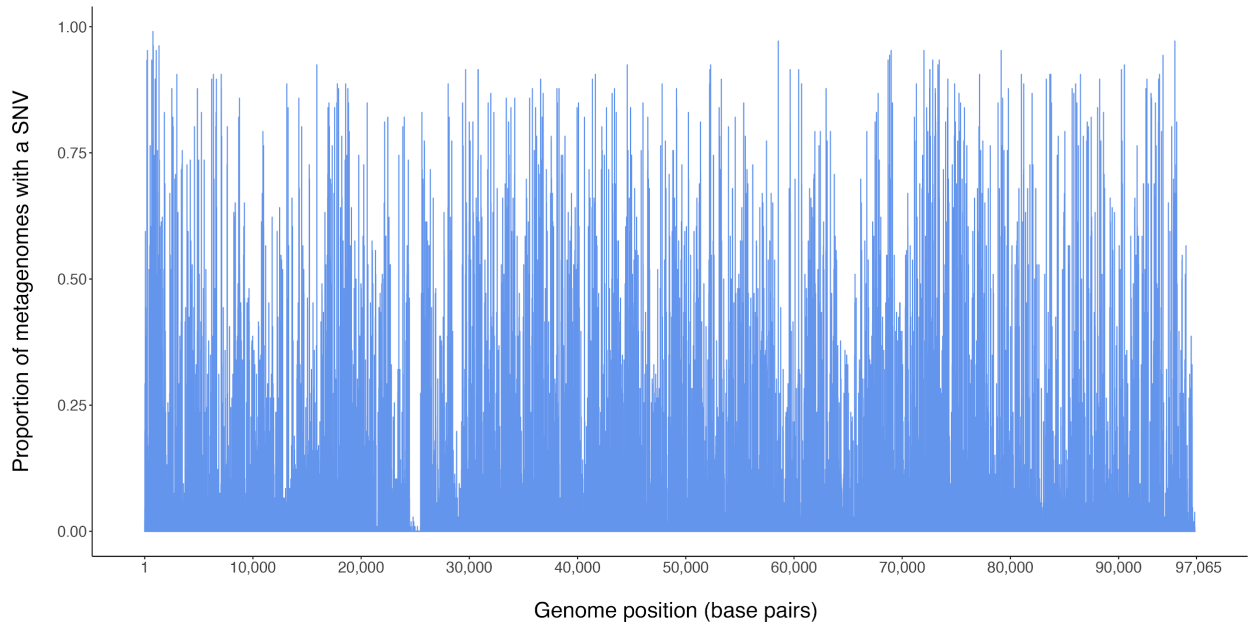
1 **Figure S1: Multiallelic sites in the crAssphage genome**



2
 3 Line plot of the number of multiallelic sites in alignments of metagenomic data to the
 4 crAssphage reference genome from individuals over serial timepoints. Only individuals with
 5 more than one time point are included. Multiallelic sites are defined here as positions with a
 6 non-reference base between 0.1 and 0.9 frequency. The number of multiallelic sites stays
 7 relatively constant across timepoints for most individuals, with the notable exception of
 8 M0806M, who appears to have a high number of multiallelic sites initially which decreases in
 9 subsequent sampling. This indicates that most individuals have a stable single crAssphage
 10 strain or strain population over time.

11
 12
 13

14 **Figure S2. Variant distribution across the genome of crAssphage**



15
16 Bar plot showing the proportion of metagenomes (n=106) with a variant (SNV or indel) at each
17 position in the crAssphage reference genome. SNVs are relatively evenly distributed throughout
18 the genome. The gap in variation at 24460-25494 base pairs corresponds to a gene sequence
19 that is present in the crAssphage reference genome but absent in crAssphage genomes in our
20 sample collection.

21

22