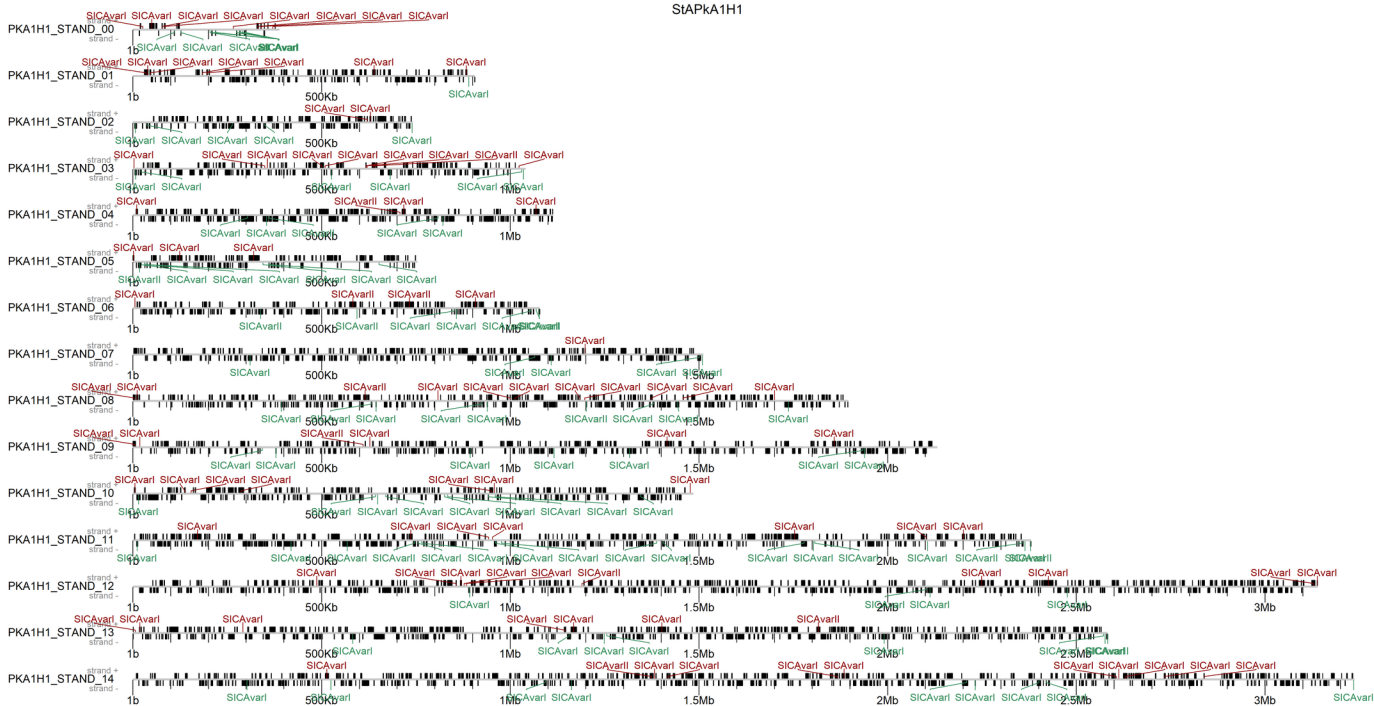


SI Fig 5. Mauve plot of chromosome 08 for StAPkA1H1, sks047, sks048 and the PKNH reference genome. Chromosome 8 of the PKNH reference shows more fragmentation than other chromosomes in the genome which may have influenced the chromosome structure inferred for the draft genomes generated here. Extensive mosaicism has been described in *P. knowlesi* chromosome 8 due to an overrepresentation of genes expressed in the mosquito stage of the parasite's life cycle. Regions of low coverage are still apparent in the draft genomes compared with the PKNH reference genome (red boxes).

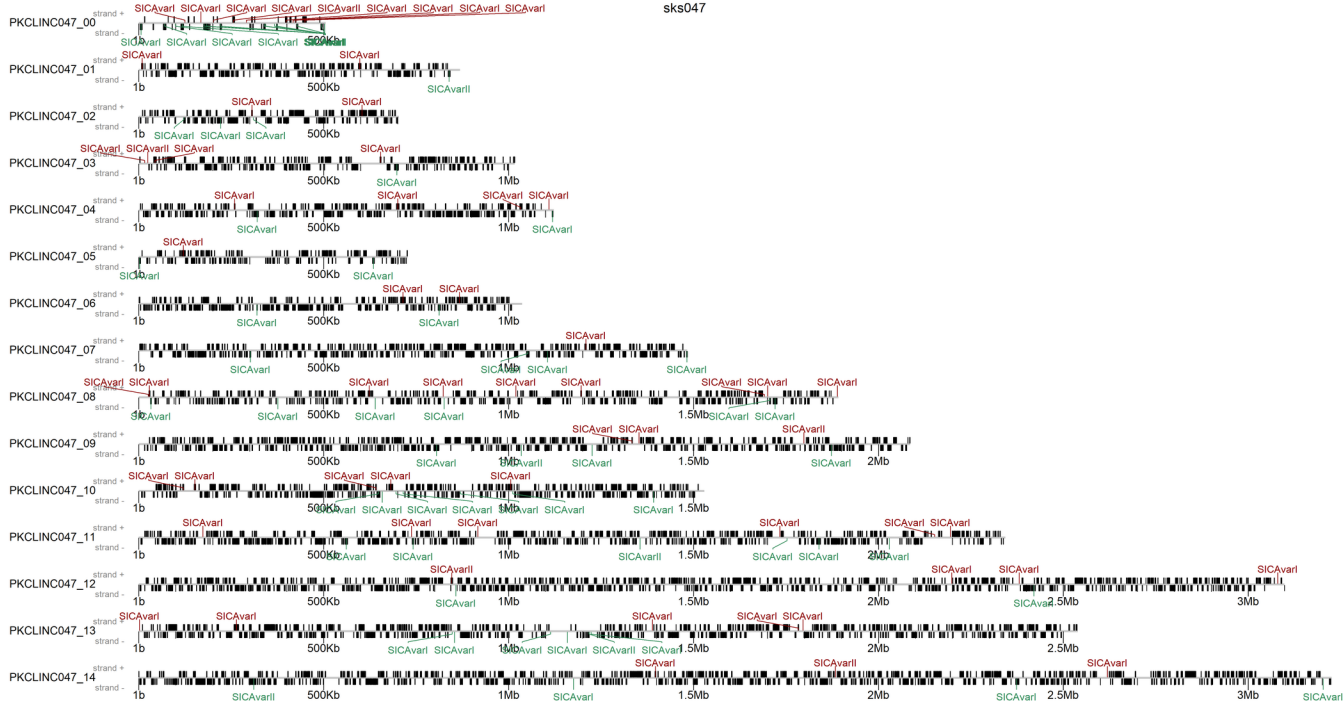
SI Fig 6B. StAPkA1H1 positioning of non-*SICAv*ar multigene family members are shown using karyoploteR. Genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. Identified members of select multigene families are given and colour coded based on being on the positive or negative strand e.g. *TrpRA* on the positive strand is slate blue and coral on the negative strand. The *SICAv*ar gene family members are presented separately in **SI Fig 7**.

SI Fig 6C. *sks047* positioning of non-*SICavar* multigene family members are shown using karyoploteR. Genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. Identified members of select multigene families are given and colour coded based on being on the positive or negative strand e.g. *TrpRA* on the positive strand is slate blue and coral on the negative strand. The *SICavar* gene family members are presented separately in **SI Fig 7**.

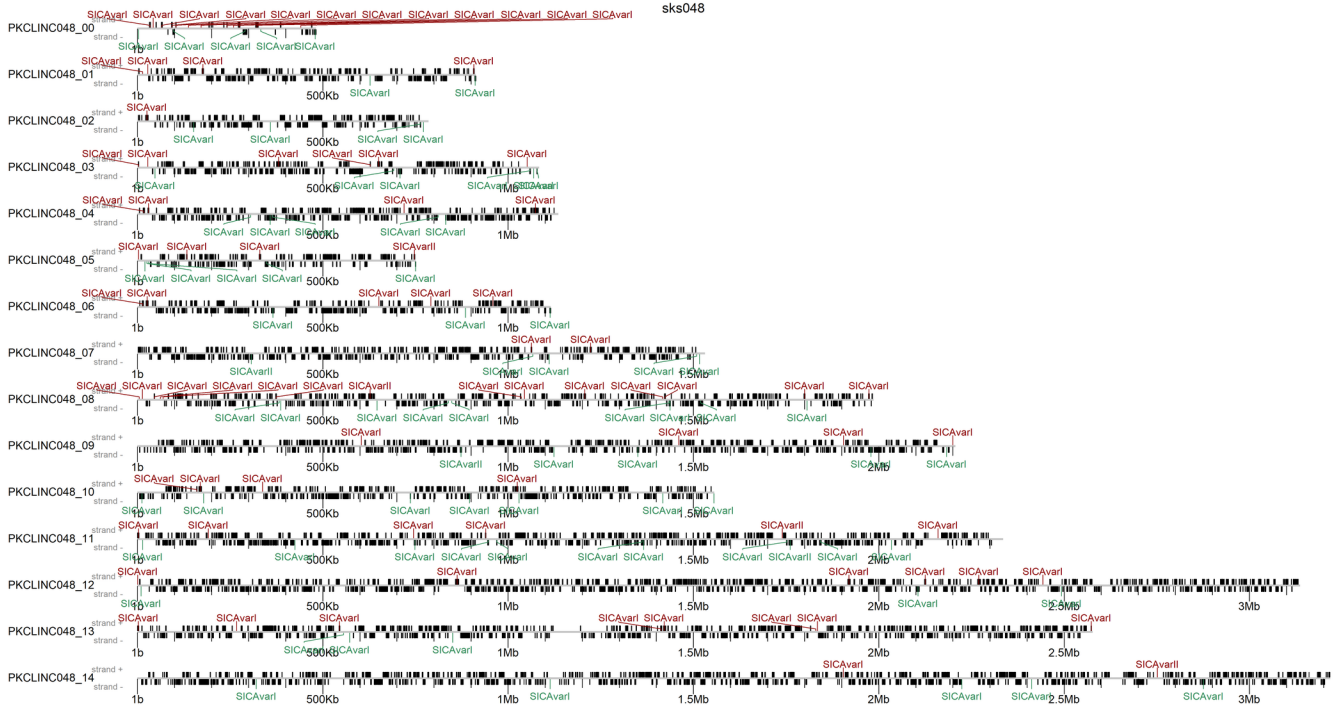
SI Fig 6D. *sks048* positioning of non-*SICAvAr* multigene family members are shown using karyoploteR. Genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. Identified members of select multigene families are given and colour coded based on being on the positive or negative strand e.g. *TrpRA* on the positive strand is slate blue and coral on the negative strand. The *SICAvAr* gene family members are presented separately in **SI Fig 7**.



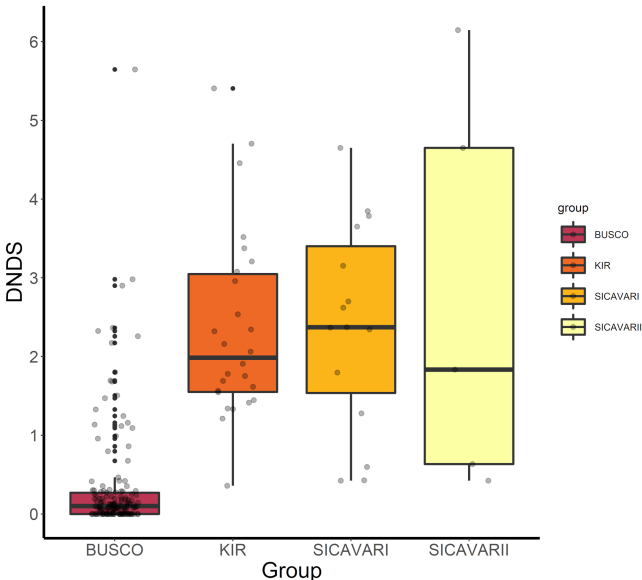
SI Fig 7B. StAPKA1H1 positioning of *SICAvAr* multigene family members are shown using karyoplotR. Annotated genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. *SICAvAr* genes and gene fragments on the positive strand are in red font and on the negative strand below in green font.



SI Fig 7C. sks047 positioning of *SICAvAr* multigene family members are shown using karyoploteR. Annotated genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. *SICAvAr* genes and gene fragments on the positive strand are in red font and on the negative strand below in green font.



SI Fig 7D. sks048 positioning of *SICAvAr* multigene family members are shown using karyoplotR. Annotated genes are shown as black squares marked along the chromosome linear map. Genes on the positive strand appear above the map line and those on the negative strand below. *SICAvAr* genes and gene fragments on the positive strand are in red font and on the negative strand below in green font.



SI Fig 8. Box plot to represent dN/dS ratios for gene-clusters from each gene type: BUSCO; KIR; SICAVar type 1 and SICAVar type 2 in the combined dataset from draft genomes StAPkA1H1, sks047, sks048 and the PKNH reference. There were 154 BUSCO, 27 KIR, 15 SICAVar type 1 and 5 SICAVar type 2 gene clusters with mean number of genes 5, 8.59, 32.2 and 18.2 per cluster respectively. Clusters containing SICAVar type 1, or SICAVar type 2 or KIR genes had a statistically significant greater mean dN/dS value when compared to BUSCO gene clusters (Wilcoxon rank sum test p -value adjustment method Bonferroni: SICAVar type 1, = $4.1\text{e-}08$; SICAVar type 2 = 0.0063 and KIR, $p = 6.7\text{e-}13$) suggesting these gene family members are under selection pressure.