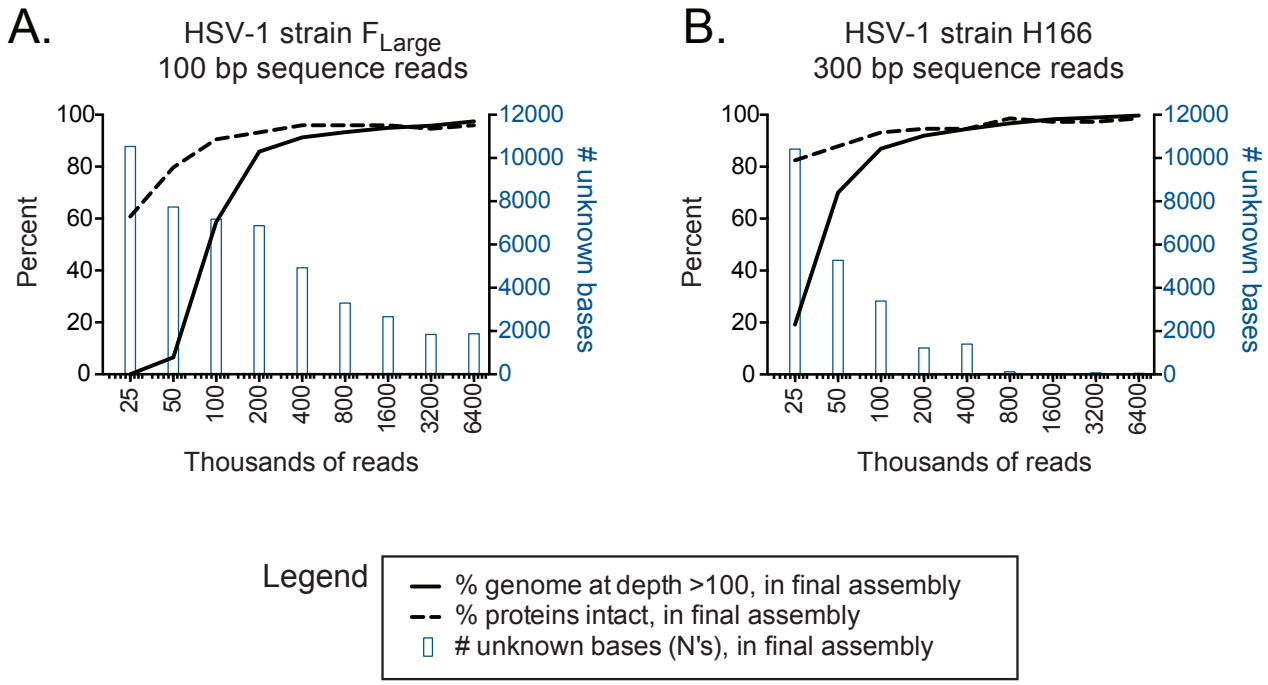
**Supplemental Figure S5. Serial assemblies demonstrate how increasing amounts of input data produce improved viral genome assemblies.**




**Supplemental Figure S5. Serial assemblies demonstrate how increasing amounts of input data produce improved viral genome assemblies.** We used VirGA to generate serial assemblies of both 100 bp paired-end sequence reads **(A)** and 300 bp paired-end sequence reads **(B)**, doubling the number of input reads each time (x-axis). For each assembly, we calculated the percent of the genome with sequencing coverage depth greater than or equal to 100-fold (solid line), the percent of proteins that were assembled intact (gap-free, dashed line), and counted the number of undetermined bases ("N"s in the final assembly, blue histograms). **(A)** For 100 bp, paired-end data, we observed that a minimum of 400,000 reads were needed to produce an HSV-1 assembly with 90% of the proteins intact. **(B)** For 300 bp, paired-end data, we observed that just 200,000 reads were needed to produce an HSV-1 assembly with 90% of the proteins intact. Results were verified with serial assemblies of data from HSV-1 $F_{Syncytial}$ and $H166_{Syncytial}$ as well (data not shown).