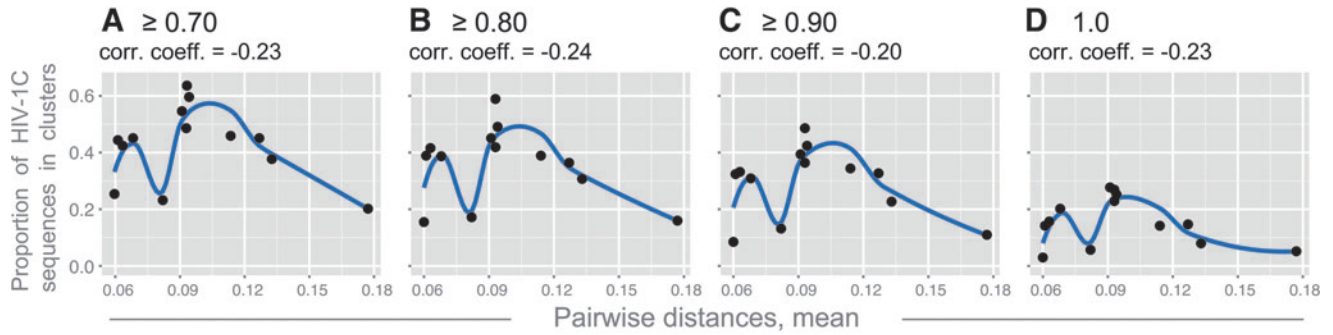
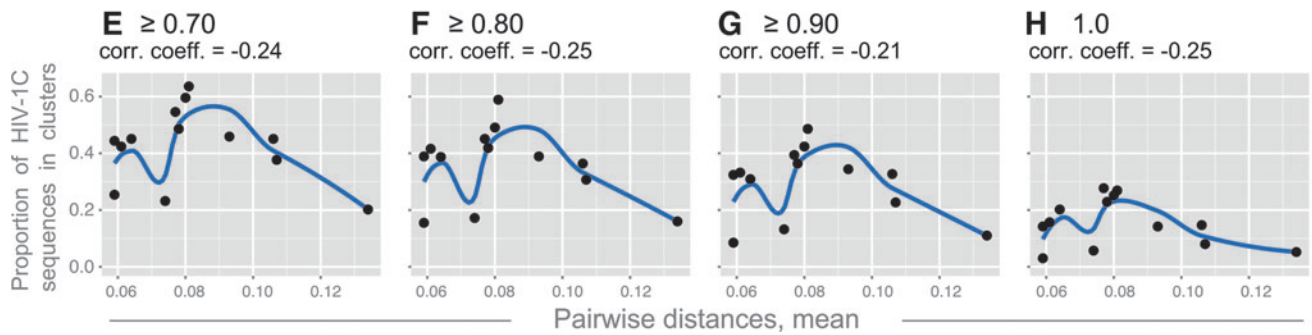


### Pairwise deletion of gaps:



### Complete deletion of gaps:



**SUPPLEMENTARY FIG. S4.** Pairwise distances and HIV clustering. Individual graphs show the relationships between HIV clustering and mean pairwise distances in the sets of sequences used. Axis y shows the proportion of HIV-1C sequences in clusters. Axis x shows pairwise distances. The calculation of pairwise distances differed by treating the gaps and missing data during computation: (**A-D**) Graphs represent *pairwise* deletion of gaps while (**E-H**) show *complete* deletion of gaps in the analyzed multiple sequence alignments.