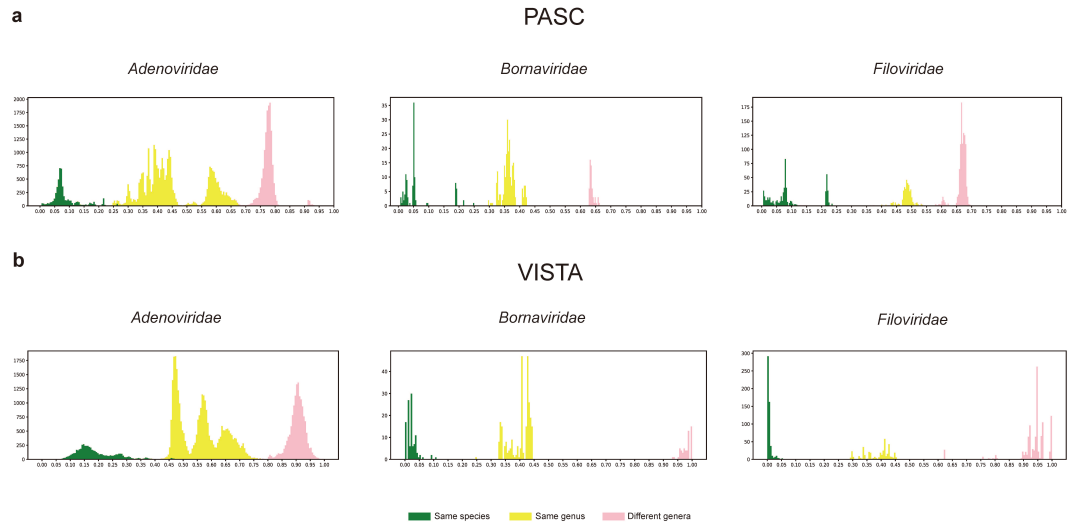
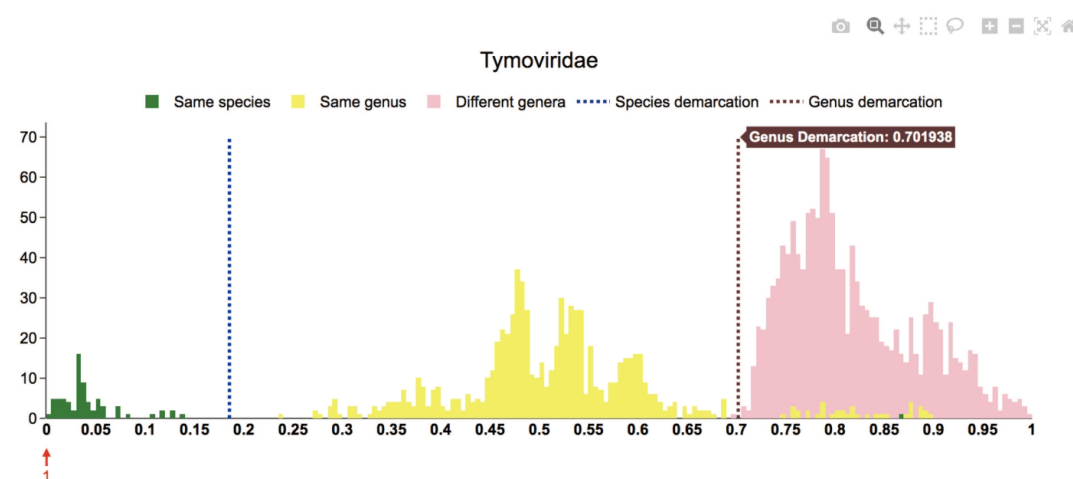




Supplementary Figure S1. Comparison of distance distributions depicted by PASC and VISTA within the families *Adenoviridae*, *Bornaviridae*, and *Filoviridae*. **a, b** Frequency distributions of pairwise distances from all complete genome sequences within three virus families (*Adenoviridae*, *Bornaviridae*, *Filoviridae*) generated by PASC and VISTA. The bars colored in green represent that both genomes belong to the same species; in yellow represent that the pair of genomes belong to different species but the same genus; and in pink represent that they belong to different genera.



Analysis Results:

[Download All Results](#)

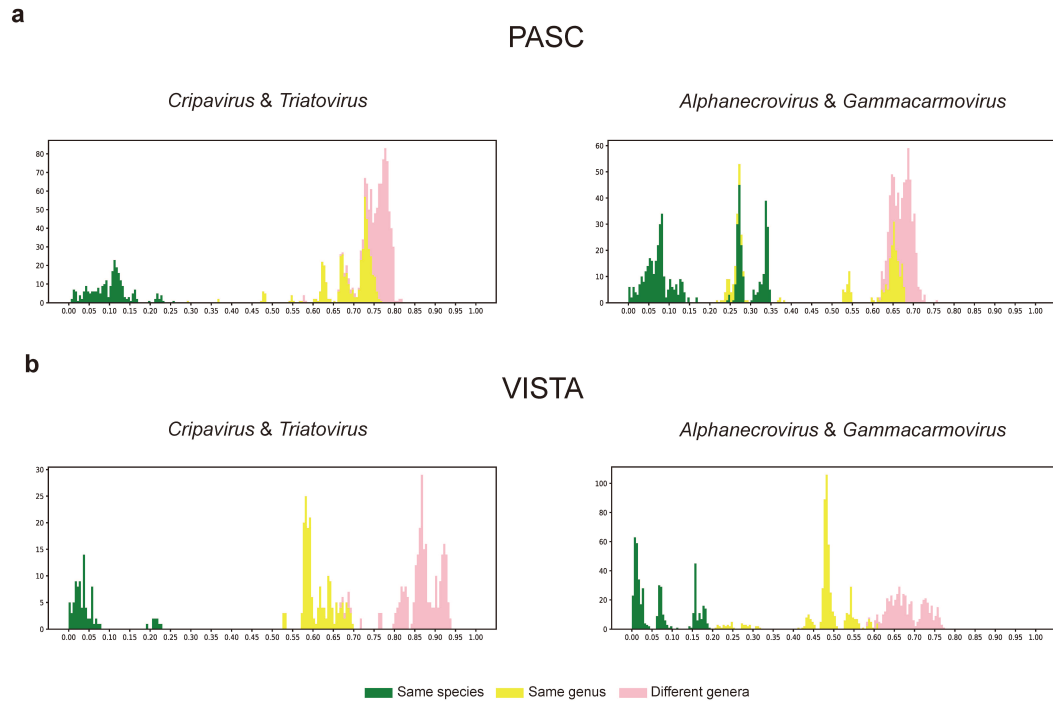
Query ID	Query Name	Closest Match	Closest Match Genus	Closest Match Species	Distance	Assignment	Download
1	gil407868434 ref NC_018703.1 Culex originated Tymoviridae-like virus, complete genome	407868434	Maculavirus	Culex originated Tymoviridae-like virus	0.0	Known Species	

Top 10 Matches:

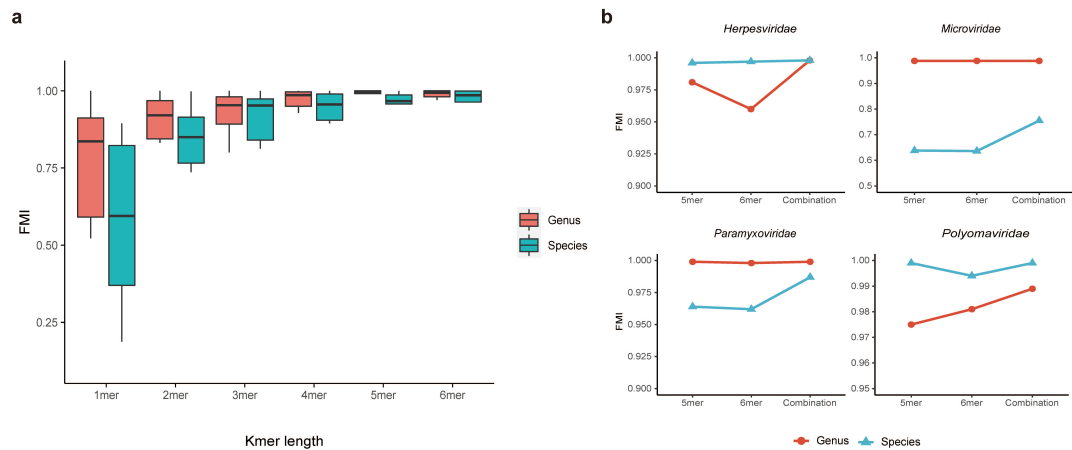
Top Matches	Reference	Reference Genus	Reference Species	Distance	Assignment
1	407868434	Maculavirus	Culex originated Tymoviridae-like virus	0.0	Known Species
2	18138525	Maculavirus	Grapevine fleck virus	0.7839	Novel Genus
3	9629255	Marafivirus	Oat blue dwarf virus	0.8122	Novel Genus
4	612507560	Maculavirus	Bombyx mori Macula-like virus	0.8174	Novel Genus
5	334285101	Maculavirus	Bombyx mori Macula-like virus	0.8187	Novel Genus
6	1045861390	Maculavirus	Grapevine Red Globe virus	0.8239	Novel Genus
7	289629717	Marafivirus	Oat blue dwarf virus	0.8263	Novel Genus
8	1189354694	Marafivirus	Citrus sudden death-associated virus	0.828	Novel Genus
9	76577724	Marafivirus	Citrus sudden death-associated virus	0.8284	Novel Genus
10	1189354696	Marafivirus	Citrus sudden death-associated virus	0.8294	Novel Genus

Showing 1 to 1 of 1 rows

Supplementary Figure S2. Screenshot of the classification result for the sequence NC_018703 generated by the VISTA web server.



Supplementary Figure S3. Comparison of distance distributions depicted by PASC and VISTA between members of different genera. a, b Analysis of two dicistrovirus genera (left) and two tombusvirus genera (right) by PASC (a) and VISTA (b).



Supplementary Figure S4. The impact of k-mer length and k-mer combination on the performance of VISTA. a, The boxplots show how clustering performance improves with increasing the k-mer length for 10 representative virus families (*Poxviridae*, *Mimiviridae*, *Papillomaviridae*, *Totiviridae*, *Filoviridae*, *Alphaflexiviridae*, *Phycodnaviridae*, *Coronaviridae*, *Hepadnaviridae*, *Inoviridae*). **b,** Comparisons of clustering performance between using the k-mer combination and using one fixed-length k-mer.