

Supplementary Material Braun et al. December 2020

Table S1 List of the 49 flaviviruses mapped by DGraph in Fig.1 and 2.

ID	Flavi track ID	Genbank	Viral name	Length of Polyproteins
1	>CFAX92USxXmX	AIM49245	Cell fusing agent virus (CFAV)	3341 aa
2	>KRx99KExNmX_r	NP_891560	Kamiti River virus	3357 aa
3	>CX_NC_008604	YP_899469	Culex flavivirus	3362 aa
4	>BSQx04USmXxX	NC_009026	Aroa virus	3429 aa
5	>IGUx04USmXxX	NC_009026.2	Aroa virus (Iguape)	3416 aa
6	>KOKx04USmXxX	AAV34157	Kokobera virus	3410 aa
7	>ALFx66AUmXxX	AAX82481	Alfuy virus	3434 aa
8	>MVEe99AUmXxX	NP_051124	Murray Valley encephalitic virus	3434 aa
9	>JBEe82JPmNmX_r	NP_059434	Japanese encephalitis virus	3432 aa
10	>USUx01ATmNiX_r	YP_164264	Usutu virus	3434 aa
11	>KUNx60AUmXmX	AAP78942	Kunjin virus	3433 aa
12	>WNe93DEmXxX_r	NP_041724	West Nile virus (WNV)	3430 aa
13	>SLEe04USmXxX_r	YP_001008348	Saint Louis encephalitis virus	3430 aa
14	>ILHx04USmXxX	YP_001040006	Ilheus virus	3424 aa
15	>ROCx04USmXxX	ATG32103	Rocio virus	3425 aa
16	>BAGx04USmXxX	AAV34161	Bagaza virus	3426 aa
17	>KEDx04USmXxX	AAV34156	Kedougou virus	3408 aa
18	>ZIKAx04USmXxX	YP_002790881	Zika virus	3419 aa
19	>DV1v74AUmNxX_r	NP_059433	Dengue virus 1	3392 aa
20	>DV3x56PHmXxX_r	YP_001621843	Dengue virus 3	3390 aa
21	>DV2v88USmXxX_r	NP_056776	Dengue virus 2	3391 aa
22	>DV4v00USmXxX_r	NP_073286	Dengue virus 4	3387 aa
23	>ENTx06UGxXrX	AKP24039	Entebbe bat virus	3411 aa
24	>YOKx71JPxNbX_r	NP_872627	Yokose virus	3425 aa
25	>SEPx06PGxXrX	YP_950478	Sepik virus	3405 aa
26	>YFx27GHmXhX	NP_041726	Yellow fever virus	3411 aa
27	>AHFh95SAthX_r	AFV78154	Alkhumra hemorrhagic fever virus	3416 aa
28	>KFDx03FRtXxX	YP_009513189	Kyasanur Forest disease virus	3416 aa
29	>GGE_DQ235153	ABB90677	Greek goat encephalitis virus	3414 aa
30	>TSE_DQ235151	ABB90675	Turkish sheep encephalitic virus	3414 aa

31	>Llx96GBtNtX_r	AYG86395	Louping ill virus	3414 aa
33	>SSE_DQ235152	ABB90676	Spanish sheep encephalitis virus	3414 aa
33	>TBEe95ATtXrX	NP_043135	Tick-borne encephalitis virus	3414 aa
34	OMSKh02UStNtX	NP_878909	Omsk hemorrhagic fever virus	3414 aa
35	>LGTx56MYtXtX_r	NP_620108	Langat virus	3414 aa
36	>DTx95CTAtXxX	AAL32169	Deer tick virus	3415 aa
37	>POWx58CAtXxX_r	NP_620099	Powassan virus	3415 aa
38	>GGY_DQ235145	YP_009345034	Gadgets Gully virus	3416 aa
39	>KSIx04UZtXxX_r	ABE73208	Karshi virus	3416 aa
40	>RF_DQ235149	ABB90673	Royal Farm virus	3417 aa
41	>KAD_DQ235146	ABB90670	Kadam virus	3404 aa
42	>MEA_DQ235144	ABB90668	Meaban virus	3421 aa
43	>SRE_DQ235150	ABB90674	Saumarez Reef virus	3422 aa
44	>TYU_DQ235148	ABB90672	Tyuleniy virus	3422 aa
45	>MMLe58USxNbX	NP_689391	Montana myotis leukoenc. virus	3374 aa
46	>RBx99FRxXxX_r	NP_620044	Rio Bravo virus	3379 aa
47	>MODx58USxNrX	NP_619758	Modoc virus	3374 aa
48	>APOIX99FRxNrX	NP_620045	Apoi virus	3371 aa
49	>TABx73TTxXbX	NP_658908	Tamana bat virus	3350 aa

Table S2: Enterovirus (EV) sequence fragments, used for Figure 1 i-g and Figure 3, listed according to their number in the plots. The fragments include the VPg (viral protein linked to the genome, “3B”, underlined (1)) and amino acids of the following area from “3C”. VPg is cleaved from the 3rd protein fragment between the terminal Q and G residues marked by /. The sequences, selected for their diversity (2, 3), were not aligned before uploading to DGRAPH. Abbreviations: CVA, CVB: coxsackie virus A or B; HPV: human poliovirus; AFP: acute flaccid paralysis; CNS: central nervous system infection.

#	Strain	Sequence	Characteristics/Disease Phenotype
1	SimN125	GAYTGLPFTKPKVPTIRQAKVQ/GPNFEFAV AMMKRNSCI AKTGQGEFTMLGIWDKWAVLP	Simian
2	EV103	GAYTGLPFNPKVPTIRQAKVQ/GPSFEFAV AMMKRNSAIIKTGQ/GEFTMLGVWDKWAVVP	Primate (with distinct differences to previously known chimpanzee and monkey viruses; (4)
3	CVA20	GAYTGLPNKKPNIPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNVAVLP	Common cold, infant diarrhea
4	CVA24	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNILTATTSKGEFTMLGVHDNVAI LP	“
5	CVA11	GAYTGLPNKRPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNVAVLP	“
6	SV6	GAYTGLPFNPKVPTIRQAKVQ/GPSFEFAV AMMKRNSAIIKTGQ/GEFTMLGIWDKWAVVP	Simian

7	HEV102	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVLP	Isolate similar to CVA20 strains (5);
8	CVA15	GAYTGLPNKRPNVPTIRTAQVQ/GPGFDYAV AMAKRNILTATTSKGEFTMLGVHDNAVILP	Common cold, infant diarrhea
9	CVA13	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVLP	Common cold, infant diarrhea
10	HEV87	GAYTGLPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNASTAKTEYGEFTMLGIYDRWAVLP	AFP (6);
11	HPV2	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVILP	Poliomyelitis Serotype 2
12	HPV3	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVILP	Poliomyelitis Serotype 3
13	PV1 Mahoney	GAYTGLPNKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVILP	Poliomyelitis Serotype 1
14	CVA18	GPYTGLPSKKPNIPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVLP	Oncolytic; isolated from throat swabs of soldiers displaying no detectable illness (7)
15	HEV99	GPYTGLPTRKPNVPTIRTAQVQ/GPGFDYAV AMAKRNILTATTTKGEFTMLGVHDNAVLP	AFP and other
16	CVA17	GAYTGLPNKKPNIPTIRTAQVQ/GPGFDYAV AMAKRNIVTATTSKGEFTMLGVHDNAVLP	Common cold, infant diarrhea
17	SimPic18	GAYSGMPQTKPKVPTIRQAKVQ/GPNFEFAV AMMKRNSAIIKTGQGEFTMLGVWDKWAVVP	Simian
18	HEV75	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNASTVKTEYGEFTMLGIYDRWAVLP	CNS, AFP (8)
19	EV6	GAYTGMPNQPKVPTLRQARVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema
20	CVA21	GAYTGLPSKKPNVPTIRIAQVQ/GPGFDYAV AMAKRNIVTATTNKGEFTMLGVHDNAVLP	Common cold, infant diarrhea
21	CVB3W	GAYTGINQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	CVB3, Woodruff strain, myocardiotropic
22	HEVB	GAYTGLPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNASTVKTEYGEFTMLGIYDRWAVLP	Enterovirus B, see EV87, EV100
23	EC05N	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Fever, rash, aseptic meningitis (sporadic) (9)
24	EV100	GAYTGLPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNASTVKTEYGEFTMLGIYDRWAVLP	EVB; see also HEV87
25	EV4	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema
26	EV96	GAYTGLPSKKPNVPTIRTAQVQ/GPGFDYAV AMAKRNILTATTSKGEFTMLGVHDNAVILP	Hand foot and mouth disease (10) ;
27	EV3	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema
28	EV11	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNASTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema
29	EV7	GAYTGLPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema
30	CVB1	GAYTGMPNQPKVPTLRQAKVQ/GPVFEFAV AMMKRNSSTVKTEYGEFTMLGIYDRWAVLP	Myocarditis, encephalitis, pleurodynia, exanthema
31	CVB6	GAYTGMPNQPKVPTLRQAKVQ/GPAFEFAV AMMKRNSSTVKTEYGEFTMLGVYDRWAVLP	Myocarditis, encephalitis, pleurodynia, exanthema
32	EV30	GAYSGMPNQKSKVPTLRQAKVQ/GPAFEFAV AMMKRNASTVKTEYGEFTMLGIYDRWAVLP	Meningitis, encephalitis, pleurodynia, exanthema

33	CVA22	GAYTGLPNAKPKVPTIRAAKVQ/GPVFDYAV AMAKKNILTATTEKGEFTMLGVYDRVAVLP	Gastroenteritis; herpangina; aseptic meningitis (11);
----	-------	---	--

Fig. S1: PD graph of the E domain 3 of sequences representing WNV, the four DENV serotypes, and Zika virus. The 7P8 sequence is a PCP-consensus derived from DENV strains that represents all four serotypes of DENV. Note that Zika has only slightly higher identity to WNV than it does to the DENV consensus.

Input file: Testfile7P8withFVstrains.txt

"Islands" Removed by 14.0 distance threshold:

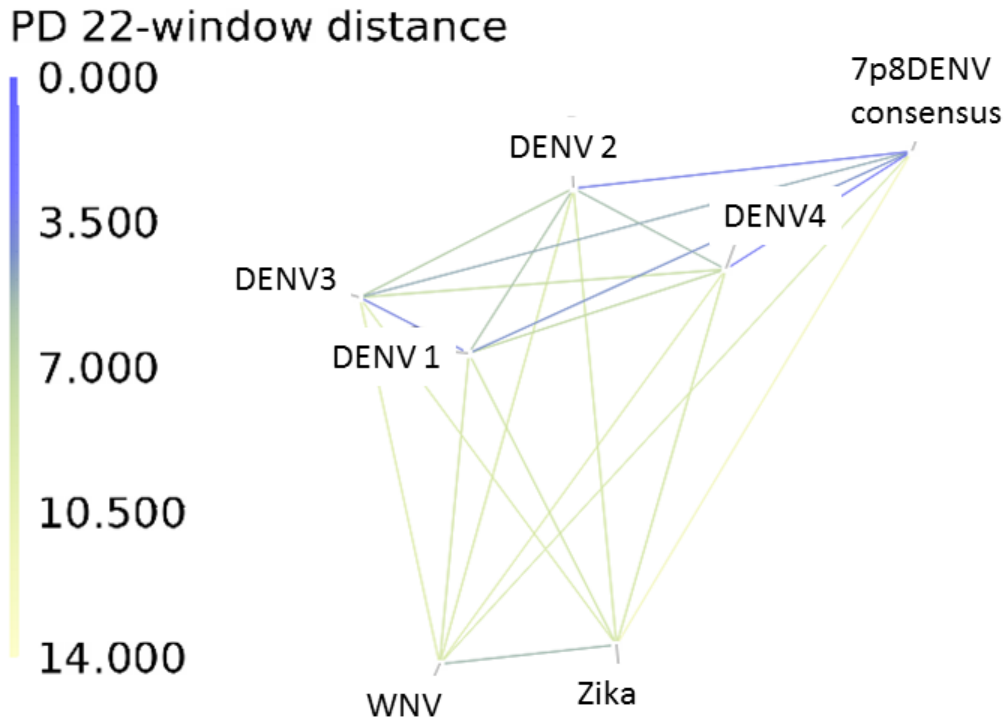
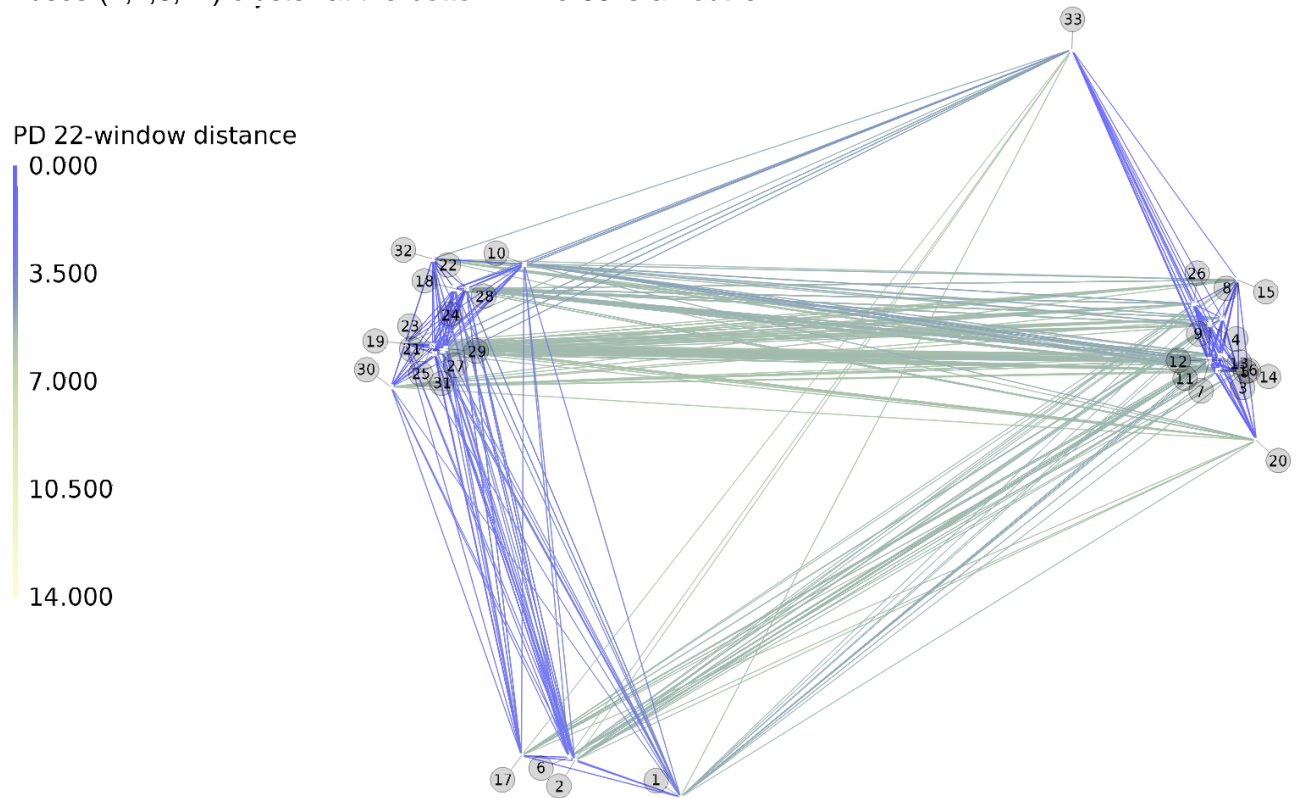


Figure S2: Enterovirus plot with numbering according to Table S2. Nodes in the left cluster are 10,18,19,21-25,27-32, those in the right cluster are 3-5,7-9,11-16,20. The four simian viruses (1,2,6,17) cluster at the bottom while 33 is an outlier.



Using the program: First time users should follow directions in the “read me” file. Input files should be ASCII text files. The headers of each FASTA sequence file can directly be displayed on the resulting graph by pushing “y” to stop the program and “b” to show sequence headers.

The input files for the figures in this paper can be obtained from the GitHub website.

To open the program, click on the executable .jar file. The program will automatically open and supply a list of subroutines for formatting or inputting data. In default mode, where the user supplies a list of sequences in FASTA format, the program will automatically open the selected text file, calculate the inter-sequence PD values as described below, and begin the simulation to determine the best 2D-presentation of the data. For very diverse data, it is recommended that

the user scale the viewing field by pushing 9 to zoom out (this may require pushing 9 several times to get the data into view). After the viewing field is established, the points can be rescattered by pushing 'e', at which point the simulation will start again. The user can stop the simulation or restart it at any time by pushing 'y'. Once the user is satisfied with the graph, it can be saved as a .pdf and .png file by typing s.

References

- 1 Schein CH, Oezguen N, van der Heden van Noort GJ, et al. NMR solution structure of poliovirus uridylyated peptide linked to the genome (VPgpU). *Peptides*. 2010 Aug;**31**(8):1441-8.
- 2 Schein CH, Ye M, Paul AV, et al. Sequence specificity for uridylylation of the viral peptide linked to the genome (VPg) of enteroviruses. *Virology*. 2015 Oct;**484**:80-5.
- 3 Schein CH, Rowold D, Choi KH. Allosteric inhibitors of Coxsackie virus A24 RNA polymerase. *Bioorganic & medicinal chemistry*. 2016 Feb 15;**24**(4):570-7.
- 4 Nix WA, Jiang B, Maher K, Strobert E, Oberste MS. Identification of enteroviruses in naturally infected captive primates. *Journal of clinical microbiology*. 2008 Sep;**46**(9):2874-8.
- 5 Brown BA, Maher K, Flemister MR, et al. Resolving ambiguities in genetic typing of human enterovirus species C clinical isolates and identification of enterovirus 96, 99 and 102. *The Journal of general virology*. 2009 Jul;**90**(Pt 7):1713-23.
- 6 Tao Z, Cui N, Xu A, et al. Complete genome sequence of enterovirus 87 isolated from a child with acute flaccid paralysis in China in 2000. *Virus genes*. 2013 Aug;**47**(1):156-9.
- 7 Au GG, Beagley LG, Haley ES, Barry RD, Shafren DR. Oncolysis of malignant human melanoma tumors by Coxsackieviruses A13, A15 and A18. *Virology journal*. 2011 Jan 18;**8**:22.
- 8 Oberste MS, Michele SM, Maher K, et al. Molecular identification and characterization of two proposed new enterovirus serotypes, EV74 and EV75. *The Journal of general virology*. 2004 Nov;**85**(Pt 11):3205-12.
- 9 Israelsson S, Gullberg M, Jonsson N, Roivainen M, Edman K, Lindberg AM. Studies of Echovirus 5 interactions with the cell surface: heparan sulfate mediates attachment to the host cell. *Virus research*. 2010 Aug;**151**(2):170-6.
- 10 Xu Y, Sun Y, Ma J, et al. A novel Enterovirus 96 circulating in China causes hand, foot, and mouth disease. *Virus genes*. 2017 Jun;**53**(3):352-6.
- 11 Tokarz R, Haq S, Sameroff S, Howie SRC, Lipkin WI. Genomic analysis of coxsackieviruses A1, A19, A22, enteroviruses 113 and 104: viruses representing two clades with distinct tropism within enterovirus C. *The Journal of general virology*. 2013 Sep;**94**(Pt 9):1995-2004.