

## Supplementary Material

### Agua Salud alphavirus defines a novel lineage of insect-specific alphaviruses discovered in the New World

Kyra Hermanns<sup>a</sup>, Marco Marklewitz<sup>a,b,c</sup>, Florian Zirkel<sup>d\*</sup>, Gijs J. Overheul<sup>e</sup>, Rachel A. Page<sup>c</sup>, Jose R. Loaiza<sup>c,e,f</sup>, Christian Drosten<sup>a,b</sup>, Ronald P. van Rij<sup>g</sup>, Sandra Junglen<sup>a,b#</sup>

<sup>a</sup>Institute of Virology, Charité – Universitätsmedizin Berlin, corporate member of Free University Berlin, Humboldt-University Berlin, and Berlin Institute of Health, Berlin, Germany;

<sup>b</sup>German Center for Infection Research (DZIF), Germany;

<sup>c</sup>Smithsonian Tropical Research Institute, Panama City, Republic of Panama;

<sup>d</sup>Institute of Virology, University of Bonn Medical Center, Bonn, Germany;

<sup>e</sup>Centro de Biodiversidad y Descubrimiento de Drogas, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT-AIP), Panama City, Republic of Panama;

<sup>f</sup>Programa Centroamericano de Maestría en Entomología, Vicerrectoría de Investigación y Postgrado, Universidad de Panamá, Republic of Panama;

<sup>g</sup>Department of Medical Microbiology, Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, Nijmegen, The Netherlands

**Supplementary file 1: Amino acid identities of single proteins between ASALV and related alphaviruses.** MAFFT-E alignments were generated for all proteins and identities [%] were calculated in Geneious. For GenBank accession numbers of the alphaviruses see figure 3. Abbreviations: EILV - Eilat virus, TALV - Tai Forest alphavirus, MWAV - Mwinilunga alphavirus, AURAV - Aura virus, SINV - Sindbis virus, WHATV - Whataroa virus, WEEV - western equine encephalitis virus, TROV - Trocara virus, EEEV - eastern equine encephalitis virus, VEEV - Venezuelan equine encephalitis virus.

	EILV	TALV	MWAV	AURAV	SINV	WHATV	WEEV	TROV	EEEV	VEEV
nsP1	68.69	66.48	69.80	69.76	67.78	69.26	55.19	63.82	55.56	56.48
nsP2	63.29	63.54	63.16	64.19	68.28	69.14	53.33	62.95	52.59	54.32
nsP3	36.06	34.32	33.96	43.17	43.73	46.88	31.26	36.91	31.74	31.05
nsP4	74.67	75.61	75.49	78.07	78.72	75.61	70.28	73.85	69.79	71.22
C	49.46	48.91	50.36	52.69	51.60	51.42	46.91	55.11	46.21	45.58
E3	49.21	49.25	54.69	50.82	50.82	45.90	59.02	50.82	43.75	60.00
E2	39.76	41.18	40.47	46.23	47.75	48.70	47.75	43.87	39.29	41.55
6K	36.36	40.00	38.18	40.00	38.18	45.45	45.45	41.82	44.07	48.28
E1	47.72	48.86	48.86	50.23	55.48	56.62	53.20	55.02	49.66	50.00

**Supplementary file 2: Phylogeny based on the NSP-ORF.** The phylogenetic tree was calculated based on a MAFFT-E translational alignment of the nucleotide sequences of the nonstructural protein ORF (nsP1-nsP4) excluding parts of nsP3 of all established alphavirus species, TALV, MWAV and ASALV strain PA-2013-MP416. An optimized maximum-likelihood phylogenetic tree with the GTR substitution model and 1000 bootstrap replicates was calculated using PhyML. The tree was rooted to the midpoint. GenBank accession numbers are shown next to the virus names.

