**Supplementary Files**

# First isolation of the Sindbis Virus in mosquitoes from Southwestern Spain reveals a new recent introduction from Africa

Rafael Gutiérrez-López1,2*; Mª José Ruiz-López3,4; Juan Ledesma1; Sergio Magallanes3,4; Cristina Nieto1; Santiago Ruiz6, Carolina Sanchez-Peña7; Ulises Ameyugo7; Juan Camacho1; Sarai Varona5,8, Isabel Cuesta5; Isabel Jado-García1; Mª Paz Sanchez-Seco1,2; Jordi Figuerola3,4; Ana Vázquez1,4

## Methods; Viral sequences analysis

The sequences were analysed for viral consensus genome reconstruction using viralrecon pipeline v2.6.0 (https://github.com/nf-core/viralrecon) (accessed on 23rd November 2023) [1] written in Nextflow (https://www.nextflow.io/) in collaboration between the nf-core community (https://nf-co.re/) and the Bioinformatics Unit of the Institute of Health Carlos III (BU-ISCIII) (https://github.com/BU-ISCIII) (further details in Supplementary material). In this pipeline, fastq files containing raw reads were first analysed for quality using FastQC v0.11.9. [2] Raw reads were trimmed using fastp v.0.23.2, [3] where a sliding window quality filtering approach was performed, scanning the reads with a 4-base wide sliding window, cutting 3′ and 5′ end bases when the average quality per base drops below a Qphred33 of 20. Reads shorter than 50 nucleotides and reads with more than 10% of read quality under Qphred 20 were removed. Additionally, poly-X sequences were removed from read ends. Trimmed reads were mapped against the reference SINV genome (GenBank accession number: OK644705) with bowtie2

v.2.4.4, [4] Picard v.3.0.0 (https://github.com/broadinstitute/picard) and SAMtools v.1.16.1 [5] were used to generate viral genome mapping stats.

Variant calling was carried out using ivar variants v.1.4. [6] This program calls low and high frequency variants and includes in the consensus genome sequences as ambiguous nucleotides those variants with an allele frequency higher than 20%. Although only high-frequency variants are kept in the consensus, all variants were annotated using SnpEff v.5.0.e [7] and, SnpSift v.4.3. [8] Finally, BEDtools v2.30.0 [9] was used to obtain the viral genome consensus with filtered variants and masking genomic regions with coverage values lower than 10×. Final summary reports were created using MultiQC v.1.14. [9]

**References**

1. H. Patel, S. Varona, S. Monzón, J. Espinosa-Carrasco, M.L. Heuer, A. Underwood. **nf-core/viralrecon: Nf-core/viralrecon v2.4.1—Plastered Magnesium Marmoset. Zenodo**. 2022.

2. S. Andrews. **FastQC: A Quality Control Tool for High Throughput Sequence Data**. 2010. Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (23rd November 2023).

3. S. Chen, Y. Zhou, Y. Chen, J. Gu. **Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor**. Bioinformatics. 34(2018), i884–i890.

4. B. Langmead, S.L. Salzberg. **Fast Gapped-Read Alignment with Bowtie 2**. Nat. Methods 9(2012), 357–359.

5. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G Marth, G. Abecasis, R. Durbin. **1000 Genome Project Data Processing Subgroup The Sequence Alignment/Map Format and SAMtools**. Bioinformatics. 25(2009), 2078–2079.

6. N.D. Grubaugh, K. Gangavarapu, J. Quick, N.L. Matteson, J.G. De Jesus, B.J. Main, A. L. Tan, L. M. Paul, D. E. Brackney, S. Grewal, N. Gurfield, K. K. A. Van Rompay, S. Isern, S. F. Michael, L. L. Coffey, N. J. Loman, K. G. Andersen **An Amplicon-Based Sequencing Framework for Accurately Measuring Intrahost Virus Diversity Using PrimalSeq and IVar**. Genome Biol. 20(2019), 8.

7. P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, D.M. Ruden. **A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila Melanogaster* Strain w 1118; Iso-2; Iso-3**. Fly. 6(2012), 80–92.

8. P. Cingolani, V.M. Patel, M. Coon, T. Nguyen, S.J. Land, D.M. Ruden, X. Lu. **Using *Drosophila Melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift**. Front. Gene. 3(2012), 35.

9. P. Ewels, M. Magnusson, S. Lundin, M. Käller. **MultiQC: Summarize analysis results for multiple tools and samples in a single report.** Bioinformatics. 32(2016), 3047–3048.

# Supplementary tables

**Table S1**. Sindbis virus strains included in this study.

| Gen Bank accession number | Reference | Isolate | Year of Isolation | Location | Isolation source | Used in Philogenetic tree | Used in molecular tree |
|---|---|---|---|---|---|---|---|
| OK644705 | Ayhan et al., 2022 | P29 Algeria | 2017 | Algeria | *Culex perexiguus* | X | X |
| KY616987 | Sigei et al., 2018 | BONI_566 | 2013 | Kenya | *Aedes ochraceus* | X | X |
| OL943984 | Meno,K. Unpublished data | KYA14MP133/SA | 2014 | South Africa | *Culex pipiens* | X | X |
| OL943985 | Meno,K. Unpublished data | MAR14MP222/SA | 2014 | South Africa | *Culex univittatus* | X | X |
| OL943986 | Meno,K. Unpublished data | KYA14MP134/SA | 2014 | South Africa | *Culex univittatus* | X | X |
| OL943983 | Meno,K. Unpublished data | GAU14MP070/SA | 2014 | South Africa | *Culex univittatus* | X | X |
| KY616985 | Sigei et al., 2018 | BONI_584 | 2013 | Kenya | *Aedes tricholabis* | X | X |
| MK045258 | Ling et al., 2019; Lundström & Pfeffer 2010. | SA80_394 | 1980 | Saudi Arabia | *Culex univittatus* | X | X |
| MK045257 | Ling et al., 2019; Lundström & Pfeffer 2010. | SA80_480 | 1980 | Saudi Arabia | *Culex univittatus* | X | X |
| MK045256 | Ling et al., 2019; Norder et al., 1996 | 1038 | 1964 | Israel | *Streptopeli turtur* | X | X |
| MK045253 | Ling et al., 2019; Norder et al., 1997 | GREIS | 1975 | Italy | *Hyalomma marginatum* | X | X |
| MK045252 | Ling et al., 2019; Norder et al., 1998 | AZ_16 | 1977 | Azerbaijan | *Nycticorax nycticorax* | X | X |
| NC001547 | Strauss et al., 1984 | Unknow | Unknow | Unknow | Unknow | X | |
| MT121982 | Mikryukova,T.P., Ternovoy,V.A., Protopopova,E.V., Konovalova,S.N., Shvalov,A.N. and Loktev,V.B. Unpublished data | AR-339 | 1952 | Egypt | *Culex pipiens/Culex univittatus* | X | X |
| OM179803 | Johnson et al., 1977 | K512 | 1971 | Kenya | *Mansonia uniformis* | X | X |
| MG679378 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | LEIV-Ast03-1-844 | 2003 | Russia | Unknow | X | X |
| MG679380 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | Tatarstan | Unknow | Russia | Unknow | X | |
| MG679373 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | Altay | Unknow | Russia | Unknow | X | |
| MG679376 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | F-720 | 1990 | Armenia | *Bubulcus ibis* | X | X |
| MG679377 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | LEIV-Ast03-1-839 | 2003 | Russia | Unknow | X | X |
| MF459683 | Kutchko et al., 2018 | Girdwood | Unknow | Unknow | Unknow | X | |
| HM147984 | Forrester et al., 2012 | Unknow | Unknow | Unknow | Unknow | X | |
| MH212167 | Ouilibona et al., 2018 | ArB7761 | 1977 | Central African Republic | *Culex cireneus* | X | X |
| MF409178 | Tricou et al., 2019 | HB741570 | 1974 | Central African Republic | *Homo sapiens* | X | X |
| MG779535 | Scheuch et al., 2018 | Z158523 | 2016 | Germany | *Culex pipiens/Culex torrentium* | X | X |
| JX570540 | Eiden et al., 2014 | Berlin-2010A | 2012 | Germany | *Corvus corone cornix* | X | X |
| MT270145 | Korhonen et al., 2020 | FIN_2018_H_05 | 2018 | Finland | *Homo sapiens* | X | X |
| JQ771793 | Sane et al., 2012 | Ilomantsi-2005M | 2005 | Finland | Mosquito | X | X |
| MF543016 | Ziegler et al., 2019 | Giessen_2016-A | 2016 | Germany | *Columba palumbus* | X | X |
| MF589985 | Jansen,S., Luehken,R., Augsten,X., Oerther,S., Joest,H., Schmidt-Chanasit,J. and Cadar,D. Unpublished data | BNI-10865 | 2016 | Germany | *Culex pipiens/Culex torrentium* | X | X |
| MG779533 | Scheuch et al., 2018 | 15Z03121 | 2015 | Germany | *Culex pipiens/Culex torrentium* | X | X |
| GU361118 | Jost et al., 2010 | 28.9 | 2009 | Germany | *Culex torrentium* | X | X |
| GU361116 | Jost et al., 2010 | 5.3 | 2009 | Germany | *Anopheles maculipennis* | X | X |
| MT270144 | Korhonen et al., 2020 | FIN_2018_H_25 | 2018 | Finland | *Homo sapiens* | X | X |
| JQ771796 | Sane et al., 2012 | Ilomantsi-2002C | 2002 | Finland | *Homo sapiens* | X | X |
| KF737350 | Bergqvist et al., 2015 | Lovanger | 2013 | Sweden | *Culiseta morsitans* | X | X |
| MG495620 | Jaworski,L., Luehken,R., Schmidt-Chanasit,J. and Cadar,D. Unpublished data | BNI-CuliMo543 | 2016 | Germany | *Culex pipiens* | X | X |
| MG679381 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | LEIV-9298 | Unknow | Russia | Unknow | X | |
| MG679379 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | SV-1383 | Unknow | Russia | *Somateria mollissima* | X | |
| JQ771799 | Sane et al., 2012 | LEIV-9298 | 1983 | Russia | Mosquito | X | X |
| JQ771795 | Sane et al., 2012 | Ilomantsi-2002B | 2002 | Finland | *Homo sapiens* | X | X |
| JQ771794 | Sane et al., 2012 | Ilomantsi-2002A | 2002 | Finland | *Homo sapiens* | X | X |
| JQ771797 | Sane et al., 2012 | Johannes-2002 | 2002 | Finland | *Homo sapiens* | X | X |
| JQ771798 | Sane et al., 2012 | Kiihtelysvaara-2002 | 2002 | Finland | *Homo sapiens* | X | X |
| MF409177 | Tricou,V., Nakoune,E., Selekon,B., Kazanji,M. and Berthet,N.Unpublished data | ArB489 | 1985 | Central African Republic | Arhropod | X | |
| MG679375 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | Stavropol | Unknow | Russia | Unknow | X | |
| MG679374 | Belyaev,A.M., Uryvaev,L.V., Ionova,K.S., Parasjuk,N.A., Lvov,D.K. and Gromashevsky,V.L. Unpublished data | Kyzylagach LEIV-65A | 1963 | Azerbaijan | Unknow | X | |
| MH229928 | Wang,M. Unpublished data | YN_222 | 2013 | China | Unknow | X | |

**Table S2.** Infection rates in the six mosquito species analysed (95% confident intervals are included).

| Mosquito species | Total mosquitoes | Range mosquitoes per pool | Pools analyzed (N) | Positive pools (N) | Infection rate | CILow 95% | CIHigh 95% |
|---|---|---|---|---|---|---|---|
| *Aedes caspius* | 2365 | [1-51] | 65 | 3 | 1.3 | 0.3 | 3.3 |
| *Aedes detritus* | 200 | [1-50] | 25 | 0 | 0 | 0 | 10 |
| *Culex laticintus* | 225 | [1-50] | 33 | 1 | 4.5 | 0.3 | 20 |
| *Culex modestus* | 1037 | [1-51] | 63 | 6 | 6.5 | 3 | 10 |
| *Culex perexiguus* | 20807 | [1-51] | 535 | 122 | 6.9 | 6 | 8 |
| *Culex pipiens* | 7286 | [1-51] | 428 | 5 | 0.7 | 0.2 | 2 |

**Table S3.** Differences in the infection rate estimates for SINV between the mosquito species and the provinces analysed.

| **Mosquito species** | *Aedes caspius* | *Culex laticintus* | *Culex modestus* | *Culex perexiguus* |
|---|---|---|---|---|
| *Culex laticintus* | 0.31 | | | |
| *Culex modestus* | 0.07 | 0.74 | | |
| *Culex perexiguus* | **0.048** | 0.72 | 0.99 | |
| *Culex pipiens* | 0.6 | 0.14 | **<0.01** | **<0.01** |

| **Province** | Cádiz | Huelva | Málaga |
|---|---|---|---|
| Huelva | 0.42 | | |
| Málaga | 0.25 | 0.47 | |
| Sevilla | **0.016** | 0.84 | 0.56 |

**Table S4.** Province, study site, mosquito species and pool size of the samples in which 16 Sindbis virus strains were isolated in 2022 in Southwestern Spain.

| Sequence ID | Study site | Province | Mosquito species | Pool size | Ct-value | GenBank accesion number |
|---|---|---|---|---|---|---|
| CA_JA220236 | Jerez de la Frontera | Cadiz | *Cx. perexiguus* | 50 | 24 | PP879145 |
| CA_JA220241 | Vejer de la Frontera | Cadiz | *Cx. perexiguus* | 50 | 30 | PP879146 |
| CA_JA220432 | Barbate | Cadiz | *Cx. perexiguus* | 50 | 29 | PP879147 |
| CA_JA220545 | Barbate | Cadiz | *Cx. perexiguus* | 50 | 27 | PP879148 |
| HU_22C1470 | Gibraleón | Huelva | *Cx. perexiguus* | 2 | 24 | PP879149 |
| HU_HP0597 | Calatilla | Huelva | *Cx. perexiguus* | 8 | 30 | PP879150 |
| SE_22C1621 | La Puebla del Río | Seville | *Cx. perexiguus* | 50 | 30 | PP879151 |
| SE_22C1639 | Coría del Río | Seville | *Cx. perexiguus* | 50 | 26 | PP879152 |
| SE_22C1668 | Palomares del Río | Seville | *Cx. perexiguus* | 50 | 25 | PP879153 |
| SE_22C1689 | La Cañada de los Pajaros | Seville | *Cx. perexiguus* | 50 | 30 | PP879154 |
| SE_22C1706 | Dehesa de Abajo | Seville | *Cx. perexiguus* | 50 | 30 | PP879155 |
| SE_22C1774 | Pinares de Colina | Seville | *Cx. perexiguus* | 50 | 30 | PP879156 |
| SE_JA220058 | Almensilla | Seville | *Cx. perexiguus* | 50 | 29 | PP879157 |
| SE_JA220061 | Bollulos de la Mitación | Seville | *Cx. perexiguus* | 50 | 25 | PP879158 |
| SE_JA220120 | Los Palacios y Villafranca | Seville | *Cx. perexiguus* | 50 | 30 | PP879159 |
| SE_JA220877 | Los Palacios y Villafranca | Seville | *Cx. perexiguus* | 50 | 30 | PP879160 |

**Table S5**. Pairwise-distances of the complete genome between the 16 SINV strains from mosquitoes collected in 2022 in Southwestern Spain, the prototype SINV strain EgAR339 (Genbank accession number NC_001547), and SINV-I P29 Argelia strain (GenBank accession number OK644705).

| GenBank accession number | NC001547 | OK644705 | PP879145 | PP879146 | PP879147 | PP879148 | PP879149 | PP879150 | PP879151 | PP879152 | PP879153 | PP879154 | PP879155 | PP879156 | PP879157 | PP879158 | PP879159 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC001547 | | | | | | | | | | | | | | | | | |
| OK644705 | 0.038 | | | | | | | | | | | | | | | | |
| PP879145 | 0.038 | 0.009 | | | | | | | | | | | | | | | |
| PP879146 | 0.038 | 0.009 | 0.000 | | | | | | | | | | | | | | |
| PP879147 | 0.038 | 0.009 | 0.000 | 0.000 | | | | | | | | | | | | | |
| PP879148 | 0.038 | 0.01 | 0.001 | 0.001 | 0.001 | | | | | | | | | | | | |
| PP879149 | 0.038 | 0.009 | 0.000 | 0.001 | 0.001 | 0.001 | | | | | | | | | | | |
| PP879150 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | | | | | | | | | | |
| PP879151 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | | | | | | | | | |
| PP879152 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | | | | | | | | |
| PP879153 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | | | | | | | |
| PP879154 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | | | | | | |
| PP879155 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | | | | | |
| PP879156 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | | | | |
| PP879157 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | | | |
| PP879158 | 0.038 | 0.01 | 0.000 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.000 | 0.001 | 0.000 | 0.000 | 0.001 | 0.000 | 0.001 | | |
| PP879159 | 0.038 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | |
| PP879160 | 0.038 | 0.09 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 |