

Phylogenetic analysis of the 2020 West Nile virus (WNV) outbreak in Andalusia (Spain)

Carlos S. Casimiro-Soriguer^{1,2,+}, Javier Perez-Florido^{1,2,+}, Jose L. Fernandez-Rueda¹, Irene Pedrosa-Corral^{3,4}, Vicente Guillot-Sulay^{3,4}, Nicola Lorusso⁵, Luis Javier Martinez-Gonzalez⁶, Jose M. Navarro-Mari^{3,4}, Joaquin Dopazo^{1,2,7,8,*}, Sara Sanbonmatsu-Gómez^{3,4*}

- ¹ Clinical Bioinformatics Area. Fundación Progreso y Salud (FPS). Hospital Virgen del Rocio. 41013. Sevilla. Spain;1; e-mail@e-mail.com
- ² Computational Systems Medicine, Institute of Biomedicine of Seville (IBIS), Hospital Virgen del Rocio. 41013. Sevilla. Spain ;
- ³ Laboratorio de Referencia de Virus de Andalucía. Servicio de Microbiología. Hospital Virgen de las Nieves, Granada. Spain
- ⁴ Instituto de investigación biosanitaria, ibs.GRANADA. Granada.Spain
- ⁵ Dirección General de Salud Pública y Ordenación Farmacéutica. Consejería de Salud y Familias. Junta de Andalucía;
- ⁶ GENYO. Centre for Genomics and Oncological Research: Pfizer - University of Granada - Andalusian Regional Government. Granada.Spain;
- ⁷ Bioinformatics in Rare Diseases (BiER). Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER). FPS. Hospital Virgen del Rocio. 41013. Sevilla. Spain;
- ⁸ Functional Genomics Node (INB). FPS. Hospital Virgen del Rocio. 41013 Sevilla. Spain.
- + Equal contributions
- * Correspondence: JD: joaquin.dopazo@juntadeandalucia.es, SB-G: saral.sanbonmatsu.sspa@juntadeandalucia.es

Abstract: During the last decades West Nile Virus (WNV) outbreaks have continuously occurred in the Mediterranean area. In August 2020 a new WNV outbreak affected 71 people with meningoencephalitis in Andalusia and 6 more cases in Extremadura (south-west of Spain), causing a total of eight deaths. The whole genomes of four viral isolates were obtained and phylogenetically analyzed in the context of recent outbreaks. The Andalusian viral samples belonged to the lineage 1 and were relatively similar to previous outbreaks occurred in the Mediterranean region. Here we present a detailed analysis of the outbreak, including an extensive phylogenetic study.

Keywords: West Nile Virus; outbreak; meningoencephalitis; epidemiology; phylogeny; whole genome sequencing

1. Introduction

West Nile virus (WNV) is an flavivirus with an enzootic cycle involving birds and mosquitoes [1], which can ultimately be transmitted to horses and/or humans causing disease outbreaks [2]. Currently, the virus is considered a recurrent zoonosis with a wide geographic distribution [3]. WNV has circulated in the Mediterranean region during the last years, being the last outbreaks in Spain in 2008, 2010 and 2016 [4, 5]. Other outbreaks have been reported in the region in Italy by 2011, 2012 [6], and 2013, including WNV from lineages 1 and 2 [7], and also in Cyprus by 2016 [8].

In August 2020, five West Nile Fever (WNF) human cases with unknown lymphocytic meningoencephalitis were first identified in the province of Seville (Andalusia) in two neighboring municipalities in front of the Guadalquivir marsh (Puebla del Río and Coria del Río). A month later, another group of human cases was identified in the province of Cádiz. The whole outbreak, as reported by Andalusian Epidemiological Surveillance System (SVEA), comprised 71 human cases of WNV meningoencephalitis, with eight deaths. Later, six more cases occurred in Badajoz [9], in the autonomous community of Extremadura, bordering Andalusia. Here we provide a detailed molecular characterization of the last WNV outbreak in Andalusia (Spain). The five cases confirmed by PCR (4 from Sevilla and 1 from Cádiz) belonged to the lineage 1 [9].

2. Materials and Methods

2.1. Molecular diagnosis and culture isolation

According to the Protocol for West Nile fever surveillance and alert, clinical samples from suspected cases of human WNV neurological disease were referred to our Andalusian Virus Reference Laboratory for laboratory diagnosis [10]. The following diagnostic tests were performed in order to confirm WNV infection [11, 12]: detection of virus-specific IgM and IgG antibodies in serum and/or cerebrospinal fluid (CSF) by ELISA testing (Euroimmun, Lübeck, Germany) and detection of specific viral nucleic acids in CSF, serum and/or urine. Nucleic acid extraction from clinical samples was performed by using QI-Asymphony DSP virus/pathogen mini kit (Qiagen, Hilden, Germany). Real time reverse transcription polymerase chain reaction (qRT-PCR) targeting a conserved sequence of the 3'-untranslated (3'UTR) region of WNV genome was used to confirm the presence of specific viral RNA [13] in 7 urine samples and 1 CSF from 7 patients. For virus isolation, all of the procedures were performed within certified biosafety cabinets under biosafety level 3 (BSL3) containment. WNV RNA-positive samples were inoculated onto confluent monolayers of Vero cells. Passage to fresh Vero cell tubes was performed after 10 days of incubation or when cytopathic effect was observed [14, 15]. Viral growth was confirmed by qRT-PCR [13] of the cell culture supernatant. WNV was isolated from two urine samples, one case from Seville and the other one from Cadiz. Purified RNA from qRT-PCR positive samples with ct values <30 along with cell culture supernatant from one WNV isolate was used for sequencing.

2.2. Viral sequencing

RNA was quantified by NanoDrop (Thermo Scientific) and verified through bioanalyzer 2100 using RNA 6000 Nano Kit (Agilent Technologies). SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific) was used for cDNA synthesis with 11 µl of RNA and 1 µl of random hexamer primers. To perform multiplex PCR, Q5® Hot Start High-Fidelity (New England Biolabs) protocol was followed by adding primers previously described [16]. After the PCR performed with a set of 41 primer pairs, the amplified regions were purified with Agencourt AMPure XP beads (Beckman Coulter). Library preparation was performed through Illumina DNA Prep kit following manufacturer's recommendation. Samples were pooled in equal concentrations after quantification by Qubit 4 fluorometer (Invitrogen). Sequencing was carried out on a Miseq system using a Nano v2 kit (Illumina).

2.3. Viral genomic data processing

Sequencing data (150bp x 2) were analyzed using in-house scripts and the nf-core/viralrecon pipeline software [17]. Briefly, after read quality filtering, sequences for each sample were aligned to the NY99 lineage 1 WNV genome (NC_009942.1) using *bowtie 2* algorithm [18], followed by primer sequence removal and duplicate read marking using *iVar* [19] and *picard* [20] tools respectively. Genomic variants are identified through *iVar* software, using a minimum allele frequency threshold of 0.25 for calling variants and a filtering step to keep variants with a minimum allele frequency threshold of 0.75. Using the set of high confidence variants and the NY99 genome, a consensus genome per sample is finally built using *iVar*.

The sequences of the four WNV isolates are available in the European Nucleotide Archive (ENA) database under the project identifier PRJEB43037.

2.4. Phylogenetic analysis

A phylogenetic analysis was performed on the obtained consensus genomes in the context of a representative set of world-wide WNV isolates (See Table S1) using the Augur application [21]. For the multiple alignment, the strain NC_009942.1 was used as a reference. The results can be viewed in the Nextstrain [22] local server of the Andalusian Genomic Epidemiology System (SIEGA) [23].

3. Results

The outbreak occurred between August and mid-September and affected the Andalusian provinces of Seville, with 56 cases and 5 deaths, and Cádiz with 15 cases and 3 deaths (a 11.3% mortality) followed by new cases in the neighborhood region of Extremadura [9]. Figure 1 summarizes the number of probable and confirmed cases [10] of WNV meningoencephalitis [10], by province of exposure in Andalusia.

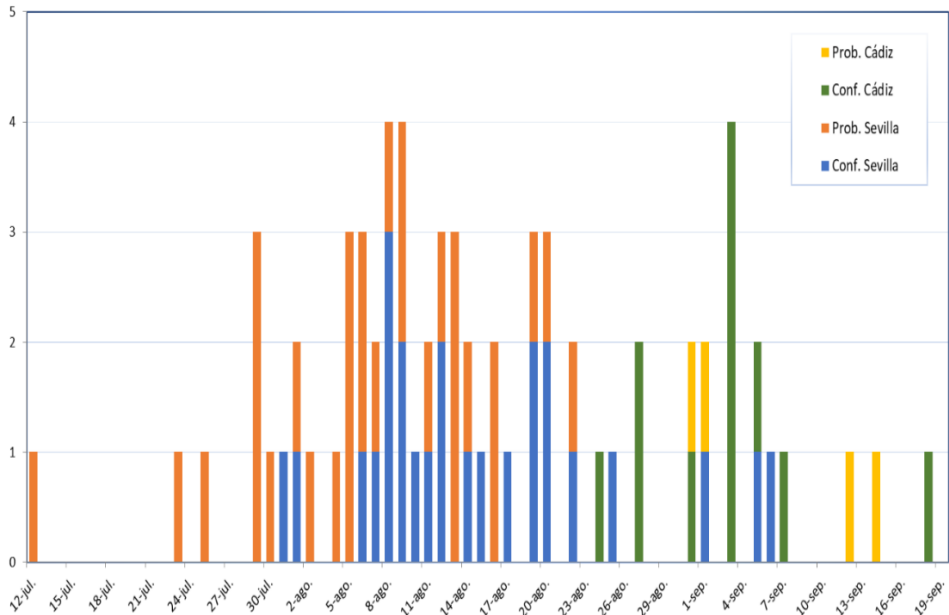


Figure 1. Epidemic curve confirmed and probable cases of WNV meningoencephalitis, by province of exposure, Andalusia 2020.

Table 1. This is a table. Tables should be placed in the main text near to the first time they are cited.

Sample	Mean depth	High confidence variants	Genome coverage
2_44013532	1190x	461	96.1%
4_44013537	1043x	420	79.7%
5_44013538	1205x	453	94.5%
6_44025400	1145x	451	94.7%

The SIEGA Nextrain server [23] displays four WNV sequences (listed in Table 1) belonging to the Spanish 2020 outbreak in the context of the rest of viral isolates worldwide that have been sequenced during the last years (see Table S1). Table S2 contains the genetic distances among all the samples that allowed the reconstruction of the whole WNV phylogeny. Figure 2 depicts a detailed view of the genetically closest isolates that include the Spanish 2010 outbreak (JF719069) and the 2008 [4] and 2011 [24] Italian outbreaks, all of them located in the Mediterranean region.

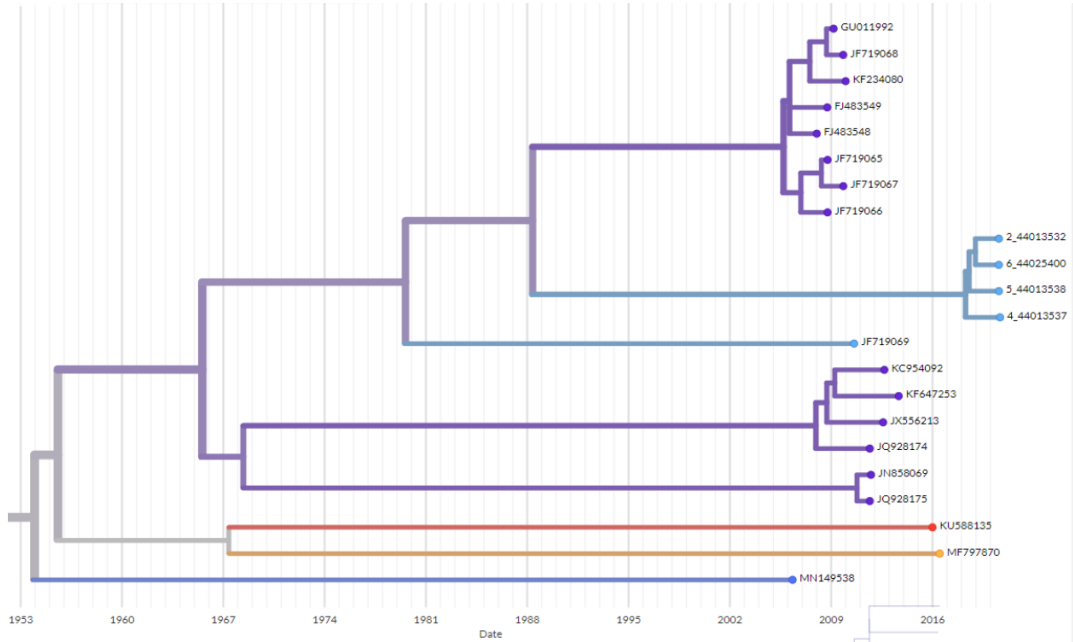


Figure 2. Sequences of the Spanish 2020 WNV outbreak and the closest relatives from previous outbreaks in Italy (darkest blue) and the Spanish 2010 outbreak (JF719069).

The phylogenetic analysis confirms the initial assignment to lineage 1 by PCR in [9]. Table 2 provides the estimated genetic distances among the WNV isolates shown in Figure 2. The genetic diversity in the current outbreak (between 4 and 8 nucleotide differences) is slightly lower, probably because its short duration, but in the range to that observed in previous outbreaks, such as the 2008 (4 to 23, with a median of 10.5, nucleotide differences) and 2012-2013 [6, 7] (9 to 32 differences) Italian outbreaks.

Table 2. This is a table. Tables should be placed in the main text near to the first time they are cited.

	2_44013532	4_44013537	5_44013538	6_44025400	JF719069	KU588135	MF797870	MN149538	IT08 ¹	IT12-13 ²	IT11 ³
2_44013532	0	8	5	5	182	330	293	252	136	238,5	248
4_44013537	8	0	8	7	165	302	265	234	127	219,5	224
5_44013538	5	8	0	4	178	320	289	250	133	232,5	244
6_44025400	5	7	4	0	178	322	289	250	133	233,5	242
JF719069	182	165	178	178	0	298	250	219	137	197	202
KU588135	330	302	320	322	298	0	246	258	295	277,5	298
MF797870	293	265	289	289	250	246	0	210	258,5	233,5	260
MN149538	252	234	250	250	219	258	210	0	216	197,5	220,5
IT08 ¹	136	127	133	133	137	295	258,5	216	10,5	198	206,25
IT12-13	238,5	219,5	232,5	233,5	197	277,5	233,5	197,5	198	9,75	192,25
IT11 ³	248	224	244	242	202	298	260	220,5	206,25	192,25	2

¹ Sequences GU011992, JF719068, KF234080, FJ483549, FJ483548, JF719065, JF719067 and JF719066, from the 2008-2009 Italian outbreak [25]
² Sequences KC954092, KF647253, JX556213 and JQ928174 from the 2012 and 2013 Italian outbreaks [6, 7].
³ Sequences JN858069 and JQ928175 from the 2011 Italian outbreak [24]

4. Discussion

The West Nile Encephalitis Surveillance Plan in Spain of the Ministry of Agriculture, Fisheries and Food [26], which contemplates entomological, ornithological and equine

surveillance during the vector activity season (April to November), discovered an especially high vector activity in the affected area, with an abundant presence of *Culex perexiguus* in rice-growing areas and *Culex pipiens* in urban areas. Actually, a total of 125 equine foci were declared by the Plan (58 in the province of Seville, 49 in Cádiz, and the rest in other provinces) since August 10, 2020 [9], reinforcing the importance of surveillance plans and supporting the establishment of a One Health, genomics-informed, real-time, global pathogen surveillance approach in the control of zoonosis [27] such as the SIEGA initiative [28].

Table 2 provides the estimated genetic distances among the WNV shown in Figure 2. The genetic diversity in the current outbreak (between 4 and 8 nucleotide differences) is slightly lower, but in the range to that observed in previous outbreaks, such as the 2008 (4 to 23 nucleotide differences) and 2011 (9 to 32 differences) Italian outbreaks. This lower nucleotide divergence is probably due to the comparatively shorter duration of the outbreak. This could be achieved by the fast implementation of Public Health measures that allowed the control of the outbreak. These measures included the reinforcement of the active epidemiological surveillance of lymphocytic meningoencephalitis human cases and the passive and active surveillance activities in horses, birds and *Culex* mosquitoes in the areas defined as risk zones. Prevention campaigns were carried out in the affected areas by informing of the recommendations for individual protection against mosquito bites. Precautionary measures were taken for blood and tissue donations. Vector population control actions for the area of influence of the outbreak consisted of a Phase 1 to give an urgent response to the elimination of the majority of the adult population of mosquitoes that could carry the virus; and a Phase 2 or larvicidal directed to the larvae of the vector.

In addition, a Rapid Assessment of the Risk of West Nile Virus Meningoencephalitis was carried out in Spain, in collaboration with the Spanish Ministry of Health [26], which considers the risk of transmission in Andalusia moderate, with a high impact due to severe neuroinvasive disease and deaths associated [9].

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: List of WNV whole genomes available in GenBank, with their IDs, isolation dates and places, and publication references when available. Table S 2: Matrix of genetic distances (as number of nucleotide changes between pairs of sequences) estimated among all the WNV whole genomes in Table S1.

Author Contributions: Conceptualization, JD, JMNM. and SSG.; methodology, JPF, CSCS, LJMG; software, JLFR; formal analysis, CSCS, JPF.; investigation, SSG, IPC, VGS.; resources, JMNM; data curation, CSCS.; writing—original draft preparation, JD, SSG, NL, LJMG, JPF; writing—review and editing, , JD, SSG, NL, LJMG, JPF; visualization, CSCS, JLFR.; supervision, JD, SSG.; funding acquisition, JD. All authors have read and agreed to the published version of the manuscript.

Funding: Please add: This research was funded by grants PT17/0009/0006 from the ISCIII, co-funded with European Regional Development Funds (ERDF) as well as H2020 “ELIXIR-EXCELERATE fast-track ELIXIR implementation and drive early user exploitation across the life sciences” (GA 676559).

Data Availability Statement: The sequences of the four WNV isolates presented here are available in the European Nucleotide Archive (ENA) database under the project identifier PRJEB43037.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Mackenzie JS, Gubler DJ, Petersen LR: **Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses.** *Nature medicine* 2004, **10**(12):S98-S109.

2. Kramer LD, Styer LM, Ebel GD: **A global perspective on the epidemiology of West Nile virus.** *Annu Rev Entomol* 2008, **53**:61-81.
3. Hubálek Z, Halouzka J: **West Nile fever--a reemerging mosquito-borne viral disease in Europe.** *Emerging infectious diseases* 1999, **5**(5):643.
4. Sotelo E, Fernandez-Pinero J, Llorente F, Vazquez A, Moreno A, Agüero M, Cordioli P, Tenorio A, Jimenez-Clavero MA: **Phylogenetic relationships of Western Mediterranean West Nile virus strains (1996–2010) using full-length genome sequences: single or multiple introductions?** *Journal of general Virology* 2011, **92**(11):2512-2522.
5. López-Ruiz N, del Carmen Montaña-Remacha M, Durán-Pla E, Pérez-Ruiz M, Navarro-Marí JM, Salamanca-Rivera C, Miranda B, Oyonarte-Gómez S, Ruiz-Fernández J: **West Nile virus outbreak in humans and epidemiological surveillance, west Andalusia, Spain, 2016.** *Eurosurveillance* 2018, **23**(14):17-00261.
6. Barzon L, Pacenti M, Franchin E, Pagni S, Lavezzo E, Squarzon L, Martello T, Russo F, Nicoletti L, Rezza G: **Large human outbreak of West Nile virus infection in north-eastern Italy in 2012.** *Viruses* 2013, **5**(11):2825-2839.
7. Barzon L, Pacenti M, Franchin E, Lavezzo E, Masi G, Squarzon L, Pagni S, Toppo S, Russo F, Cattai M *et al*: **Whole genome sequencing and phylogenetic analysis of West Nile virus lineage 1 and lineage 2 from human cases of infection, Italy, August 2013.** *Eurosurveillance* 2013, **18**(38):20591.
8. Richter J, Tryfonos C, Tourvas A, Floridou D, Paphitou NI, Christodoulou C: **Complete Genome sequence of West Nile virus (WNV) from the first human case of neuroinvasive WNV infection in Cyprus.** *Genome announcements* 2017, **5**(43).
9. **Meningoencefalitis por virus del Nilo occidental en España (2ª actualización).** Evaluación rápida de riesgo [https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/docs/20201203_ERR_Nilo_Occidental.pdf]
10. **Protocolo de vigilancia y alerta de Fiebre del Nilo Occidental.** [Protocol for West Nile fever surveillance and alert]. [https://juntadeandalucia.es/export/drupaljda/salud_5af95879cc803_fiebre_nilo250817.pdf]
11. Martínez Yoldi MJ, Ruiz M, Sánchez-Seco Fariñas MP, Vázquez González A: **Diagnóstico microbiológico de las principales arbovirosis importadas y autóctonas.** In: *Procedimientos en Microbiología Clínica*. Edited by Mansilla Cercenado E, Moreno Cantón R: Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica (SEIMC); 2020.
12. Antón AIN, de Ory Manchón F, Fariñas MPS-S, Narváez LF, Cámara MIG, Mari JMN, Matanzo AT: **Diagnóstico microbiológico de arbovirosis y robovirosis emergentes.** *Enfermedades infecciosas y microbiología clinica* 2015, **33**(3):197-205.
13. Vázquez A, Herrero L, Negredo A, Hernández L, Sánchez-Seco MP, Tenorio A: **Real time PCR assay for detection of all known lineages of West Nile virus.** *Journal of Virological methods* 2016, **236**:266-270.
14. Monath TP: **The Arboviruses:: Epidemiology and Ecology**, vol. 1: CRC Press; 2020.
15. Barzon L, Pacenti M, Franchin E, Squarzon L, Sinigaglia A, Ulbert S, Cusinato R, Palù G: **Isolation of West Nile virus from urine samples of patients with acute infection.** *Journal of clinical microbiology* 2014, **52**(9):3411-3413.
16. Hepp CM, Cocking JH, Valentine M, Young SJ, Damian D, Samuels-Crow KE, Sheridan K, Fofanov VY, Furstenau TN, Busch JD: **Phylogenetic analysis of West Nile Virus in Maricopa County, Arizona: Evidence for dynamic behavior of strains in two major lineages in the American Southwest.** *PloS one* 2018, **13**(11):e0205801.
17. **nf-core/viralrecon: nf-core/viralrecon v1.1.0 - Steel Pangolin (Version 1.1.0).** [<https://zenodo.org/record/3905178#.YBumlOhKi71>]
18. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Methods* 2012, **9**(4):357-359.
19. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, Tan AL, Paul LM, Brackney DE, Grewal S: **An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar.** *Genome biology* 2019, **20**(1):1-19.
20. **Picard. A set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF.** [<http://broadinstitute.github.io/picard/>]
21. Huddleston J, Hadfield J, Sibley TR, Lee J, Fay K, Ilcisin M, Harkins E, Bedford T, Neher RA, Hodcroft EB: **Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens.** *Journal of Open Source Software* 2021, **6**(57):2906.

22. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA: **Nextstrain: real-time tracking of pathogen evolution**. *Bioinformatics* 2018, **34**(23):4121-4123.
23. **West Nile Virus epidemiological local server - SIEGA** [<http://nextstrain.clinbioinfoispa.es/wnv>]
24. Barzon L, Pacenti M, Franchin E, Squarzon L, Lavezzo E, Toppo S, Martello T, Cattai M, Cusinato R, Palù G: **Novel West Nile virus lineage 1a full genome sequences from human cases of infection in north-eastern Italy, 2011**. *Clinical Microbiology and Infection* 2012, **18**(12):E541-E544.
25. Barzon L, Squarzon L, Cattai M, Franchin E, Pagni S, Cusinato R, Palu G: **West Nile virus infection in Veneto region, Italy, 2008-2009**. *Euro Surveill* 2009, **14**(31).
26. **Surveillance program for the West Nile Virus 2021 [Programa de Vigilancia de fiebre del Nilo occidental 2021]** [https://www.mapa.gob.es/es/ganaderia/temas/sanidad-animal-higiene-ganadera/programafiebreelnilooccidental2021_tcm30-437515.pdf]
27. Gardy JL, Loman NJ: **Towards a genomics-informed, real-time, global pathogen surveillance system**. *Nature Reviews Genetics* 2018, **19**(1):9.
28. **Integrated System for Genomic Epidemiology of Andalusia [Sistema Integrado de Epidemiologia Genomica de Andalusia - SIEGA]** [<http://clinbioinfoispa.es/projects/siega/>]