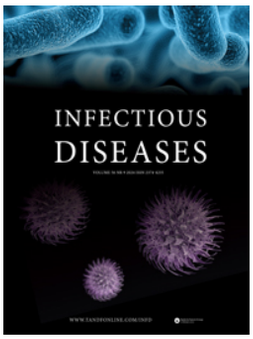

Infectious Diseases

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/infd20

The key role of Spain in the traffic of West Nile virus lineage 1 strains between Europe and Africa

Pilar Aguilera-Sepúlveda, Cristina Cano-Gómez, Rubén Villalba, Vítor Borges, Montserrat Agüero, Daniel Bravo-Barriga, Eva Frontera, Miguel Ángel Jiménez-Clavero & Jovita Fernández-Pinero

To cite this article: Pilar Aguilera-Sepúlveda, Cristina Cano-Gómez, Rubén Villalba, Vítor Borges, Montserrat Agüero, Daniel Bravo-Barriga, Eva Frontera, Miguel Ángel Jiménez-Clavero & Jovita Fernández-Pinero (2024) The key role of Spain in the traffic of West Nile virus lineage 1 strains between Europe and Africa, Infectious Diseases, 56:9, 743-758, DOI: 10.1080/23744235.2024.2348633

To link to this article: https://doi.org/10.1080/23744235.2024.2348633

Taylor & Francis
Taylor & Francis Group

INFECTIOUS DISEASES,
2024; VOL. 56,
NO. 9, 743–758

https://doi.org/10.1080/23744235.2024.2348633

RESEARCH ARTICLE

OPEN ACCESS Check for updates

The key role of Spain in the traffic of West Nile virus lineage 1 strains between Europe and Africa

Pilar Aguilera-Sepúlveda[a], Cristina Cano-Gómez[a], Rubén Villalba[b], Vítor Borges[c], Montserrat Agüero[b], Daniel Bravo-Barriga[d], Eva Frontera[d], Miguel Ángel Jiménez-Clavero[a,e] and Jovita Fernández-Pinero[a]

[a]Centro de Investigación en Sanidad Animal (CISA-INIA), CSIC, Valdeolmos, Spain; [b]Laboratorio Central de Veterinaria (LCV), Ministry of Agriculture, Fisheries and Food (MAPA), Algete, Spain; [c]Genomics and Bioinformatics Unit, Department of Infectious Diseases, National Institute of Health Doutor Ricardo Jorge (INSA), Lisbon, Portugal; [d]Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad de Extremadura, Cáceres, Spain; [e]CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain

ABSTRACT

Background: West Nile Virus (WNV) is a zoonotic arbovirus worldwide spread. Seasonal WNV outbreaks occur in the Mediterranean basin since the late 1990's with ever-increasing incidence. In Southern Spain WNV is endemic, as disease foci - caused by WNV lineage 1 (WNV-L1) strains - occur every year. On the contrary, WNV-L2 is the dominant lineage in Europe, so most European WNV sequences available belong to this lineage, WNV-L1 sequences being still scarce.

Methods: To fill this gap, this study reports the genetic characterisation of 27 newly described WNV-L1 strains, involved in outbreaks affecting wild birds and horses during the last decade in South-Western Spain.

Results: All strains except one belong to the Western Mediterranean-1 sub-cluster (WMed-1), related phylogenetically to Italian, French, Portuguese, Moroccan and, remarkably, Senegalese strains. This sub-cluster persisted, spread and evolved into three distinguishable WMed-1 phylogenetic groups that co-circulated, notably, in the same province (Cádiz). They displayed different behaviours: from long-term persistence and rapid spread to neighbouring regions within Spain, to long-distance spread to different countries, including transcontinental spread to Africa. Among the different introductions of WNV in Spain revealed in this study, some of them succeeded to get established, some extinguished from the territory shortly afterwards. Furthermore, Spain's southernmost province, Cádiz, constitutes a hotspot for virus incursion.

Conclusion: Southern Spain seems a likely scenario for emergence of exotic pathogens of African origin. Therefore, circulation of diverse WNV-L1 variants in Spain prompts for an extensive surveillance under a One Health approach.

KEYWORDS
West Nile virus
Full genome sequencing
PrM-E sequencing
Phylogenetic analysis
Wild birds
Horses
Mediterranean

ARTICLE HISTORY
Received 2 February 2024
Revised 20 April 2024
Accepted 22 April 2024

CONTACT
Jovita Fernández-Pinero
fpinero@inia.csic.es
Miguel Ángel Jiménez-Clavero
majimenez@inia.csic.es
Centro de Investigación en Sanidad Animal (CISA-INIA), CSIC, Valdeolmos, Spain

Present addresses: Cristina Cano-Gómez, Laboratorio Central de Veterinaria (LCV), Ministry of Agriculture, Fisheries and Food (MAPA), Algete, Spain; Daniel Bravo-Barriga, Departamento de Sanidad Animal, Grupo de Investigación en Zoonosis y Sanidad Animal (GISAZ), Universidad de Córdoba, Córdoba, Spain.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23744235.2024.2348633.

Introduction

West Nile virus (WNV; family *Flaviviridae*, genus *Orthoflavivirus*) is a mosquito-borne zoonotic, epizootic and epornitic virus belonging to the Japanese encephalitis serocomplex. Its genome consists of a single-stranded positive-sense RNA of approximately 11,000 nt encoding a single polyprotein, which is excised into three structural proteins: capsid (C), pre m/membrane (PrM/M) and envelope (E), and seven non-structural (NS) proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 [1]. Under natural conditions, WNV circulates in an enzootic cycle between mosquitoes, mainly *Culex* genus (vectors), and birds (reservoir hosts) while mammals (particularly humans and horses) are considered incidental, dead-end hosts [1] as they do not develop enough viremia titre to infect vectors.

WNV is characterised by a remarkable phenotypic variation and a high genetic diversity; currently, there are at least 8 lineages described [2], being lineages 1 (WNV-L1) and 2 (WNV-L2) the most widely distributed and the ones associated with disease in animals and humans [3].

First described in Uganda in 1937, WNV was recognised as a significant emerging pathogen around the 1990s, when the frequency, severity, and geographic range of outbreaks increased dramatically through European and Mediterranean countries and eventually spread through the Americas [1]. Nowadays, WNV is one of the most widespread arboviruses in the world, reported in all continents except Antarctica [4,5]. This makes WNV one of the most relevant examples of emerging infectious diseases, with considerable impact on public and animal health [6].

The first description of WNV in the Western Mediterranean dates back to 1962 in the Camargue, France [7]. It reappeared in the region in the mid-1990s, with successive sporadic outbreaks reported by neighbouring countries including Algeria, France, Italy, Morocco, Portugal, Spain and Tunisia [8]. The situation became more pressing since 2008, when outbreaks in Italy started to be more frequent and severe [9], a trend that has continued in Europe and Mediterranean countries since then [10,11]. WNV outbreaks in Europe were due to lineage 1 strains until the introduction of WNV-L2, first described in 2004 in Hungary [12], which quickly dispersed across the continent [11,13–15], changing radically the WNV landscape in this region. Nowadays, WNV-L2 is the predominant lineage circulating in Europe [4,11], while WNV-L1 circulation is currently restricted to France [2], Southern Spain [16–18], Italy [9,19], Cyprus [20], and neighbouring areas, North-Western African countries [21–23], Turkey [24] and Israel [25].

In Spain, WNV-L1 was first isolated in 2007 from golden eagles [26]. Since then, it has been detected multiple times in South-Western regions of the country, either in mosquitos or in vertebrate hosts (birds, horses, humans) [16–18,27–33]. The most intense and severe transmission season in Spain occurred in 2020, with 77 human cases (including 8 deaths) and 139 outbreaks among equids reported, all of them due to WNV-L1 [16]. Meanwhile, an incursion of WNV-L2 in the farthest part of the country (North-East) was first detected in 2017, affecting wild birds (raptors), which continued in 2020, 2021 and 2022, as reported in wild birds, mosquitoes, and humans, respectively [34–36]. Despite the key position of Spain, close to Africa and crossed by key bird migration routes between this continent and Europe, little is known regarding WNV dynamics in this country and its role in WNV epidemiology in the context of the Mediterranean region. This gap is particularly notable when it comes to WNV-L1, which has become a minor variant in Europe. The knowledge on when, how and from where WNV-L1 is introduced into Europe, and how it remains and progresses in this continent, is still very limited today [37], as these studies are bound to the availability of suitable samples for viral characterisation, which are scarce. Under these premises, this study provides 13 complete and 14 partial genome sequences of WNV-L1, obtained from WNV foci occurred in wild birds and horses between 2010 and 2020, in diverse Spanish provinces. This notable contribution made possible a powerful phylogenetic analysis, enabling to better depict the relatedness among WNV-L1 strains circulating in South-Western Spain and their relationships with other strains from Mediterranean and Western African origins, shedding light into their evolutionary history and dynamics across this vast region. Moreover, polyprotein profiles of the different WNV-L1 variants described were analysed. In addition, a simple and affordable characterisation method was developed and here described, based on a single conventional RT-PCR followed by conventional sequencing, which enables the study and classification of closely related WNV-L1 strains.

Materials and methods

Sample collection

Samples ($n = 27$) were collected from horses ($n = 11$; cerebrospinal fluid, organs) and wild birds ($n = 16$;

organs, feathers) with clinical suspicion of WNV infection, that was later confirmed by laboratory diagnostic techniques. The collection encompasses a period of 11 years (2010 to 2020) in five provinces belonging to three Spanish Autonomous Communities (AC): Cádiz, Seville and Huelva (AC: Andalusia), Toledo (AC: Castile-La Mancha), and Cáceres (AC: Extremadura) (Figure 1). The detailed information of the field samples used in this study (and other Spanish sequences previously published) is described in Table 1. Part of the samples employed in this study ($n = 20$) were initially sent to the Spanish National Reference Laboratory for WNV (LCV, Algete, Spain) for laboratory confirmation of WNV detection. The remaining samples ($n = 7$) were directly sent to CISA, where further molecular and virological characterisations were performed on all samples. All samples were shipped frozen to CISA and preserved at −80 °C until further analysis. During the analytical phase, samples were processed as previously described [38].

Viral isolation

Nineteen suitable WNV-L1 positive samples were subjected to virus isolation by inoculation of Vero cells as previously described [26]. Positive results (cytopathic effect observed) were confirmed by RT-PCR. Samples were considered negative for virus isolation if, after 3 cell culture passages, cytopathic effect was absent.

Viral genome detection assays

Viral RNA was extracted from 100 µl of sample–homogenates and swabs using the BioSprint 15 workstation (Qiagen) and the BioSprint DNA blood kit (Qiagen), as previously described [36].

Samples initially sent to the LCV for confirmation of WNV detection, were analysed by a real-time RT-PCR [39]. The rest of the samples were examined at CISA by a real–time multiplex RT–PCR that distinguishes between WNV-L1, WNV-L2 and Usutu virus (USUV) [40]. For both PCR-based methods, samples with Ct value > 40 were considered negative. Ct values for WNV detection are shown in Table 1.

WNV genome sequencing

Nucleic acid amplification of the WNV-L1 whole genome was carried out by conventional RT-PCRs, using primer sets previously described [41]. Additionally, 26 new primers were specifically designed for this study (Table S1) to cover the complete viral genomes.

Figure 1. Map of Spain and its provinces. The provinces where WNV-L1 positive samples have been collected during this study are coloured in brown.

Table 1. Information of all the collected samples in this study as well as all previously Spanish WNV-L1 published sequences.

Year	Province	Species	ID	ID (abbrev.)	Sample (F/I) Field sample (F) or Isolate (I)	Source of the sample/isolate	WNV RRT-PCR (Ct value)	Genome sequencing. Field sample (F)/Isolate (I)	GenBank accession no.	Reference of the sequence
2007	Toledo	Golden eagle	GE-1b/B	GE-1b/B	Brain (I)	N/A	N/A	Complete (I)	FJ766331	Sotelo et al. 2009
2007	Toledo	Golden eagle	GE-2o/V	GE-2o/V	Oropharyngeal swab (I)	N/A	N/A	Complete (I)	FJ766332	Sotelo et al. 2009
2008	Seville	*Culex perexiguus*	HU6365/08	HU6365/08	Pool mosquitoes (F)	N/A	N/A	Complete (F)	JF707789	Sotelo et al. 2011
2010	Cádiz	Horse	Spain/2010/H-1b	Spain/2010/H-1b	Brain (F)	N/A	N/A	Complete (F)	JF719069	Sotelo et al. 2011
2010	Cádiz	Horse	5617_Spain/2010/H-b	5617	Brain (F)	LCV	30,37[1]	Partial (F)	OM302333	This study
2011	Cádiz	Cape Barren goose	4797_Spain/2011/CBG-b	4797	Brain (F)	LCV	30,86[1]	Partial (F)	OM302331	This study
2011	Cádiz	Red-legged partridge	3011_Spain/2011/RLP-b	3011	Brain (F)	LCV	32,64[1]	Partial (F)	OM302334	This study
2012	Cádiz	Horse	2712_Spain/2012/H-b	2712	Brain (F)	LCV	33,07[1]	Partial (F)	OM302332	This study
2012	Cádiz	Horse	SPA12-01_Spain/2012/H-b/2V	SPA12-01	Brain (I)	LCV/CISA	27,82[2]	Complete (I)	OM302323	Ruiz-López et al. 2023b
2012	Cádiz	Horse	SPA12-02_Spain/2012/H-cf/2V	SPA12-02	Cerebrospinal fluid (I)	LCV/CISA	28,58[2]	Complete (I)	OM302322	Ruiz-López et al. 2023b
2013	Seville	Red-legged partridge	3179_Spain/2013/RLP-os	3179	Oral swab (F)	LCV	33,73[1]	Partial (F)	OM302330	This study
2013	Huelva	Horse	3442_Spain/2013/H-b	3442	Brain (F)	LCV	29,37[1]	Partial (F)	OM302326	This study
2013	Huelva	Horse	3589_Spain/2013/H-b	3589	Brain (F)	LCV	33,99[1]	Partial (F)	OM302327	This study
2015	Seville	Red-legged partridge	2496_Spain/2015/RLP-cs	2496	Cloacal swab (F)	LCV	33,22[1]	Partial (F)	OM302338	This study
2015	Cádiz	Horse	2315_Spain/2015/H-b	2315	Brain (F)	LCV	30,98[1]	Partial (F)	OM302325	This study
2015	Cádiz	Horse	2232_Spain/2015/H-b	2232	Brain (F)	LCV	35,42[1]	Partial (F)	OM302329	This study
2016	Seville	Duck	2439_Spain/2016/D-f	2439	Feathers (F)	LCV	22,10[1]	Complete (F)	OM302324	This study
2016	Cáceres	Cinereous vulture (16/959)	2442-3_Spain/2016/CV-pf	2442-3	Pectoral feathers (F)	LCV	26,50[1]	Partial (F)	OM302328	This study
2016	Cáceres	Cinereous vulture (16/953)	2442-4_Spain/2016/CV-pf	2442-4	Pectoral feathers (F and I)	LCV/CISA	23,70[1] (F) and 17,70[2] (I)	Complete (I)	OM302315	This study
2016	Cáceres	Cinereous vulture (16/953)	2442-5_Spain/2016/CV-b	2442-5	Brain (F)	LCV	14,30[1]	Complete (F)	OM302316	This study
2016	Cáceres	Cinereous vulture (16/953)	2442-5_Spain/2016/CV-b/3B	2442-5	Brain (I)	LCV/CISA	12,50[1]	Complete (I)	OM302314	This study
2016	Seville	Horse	2500_Spain/2016/H-b/2V	2500	Brain (I)	LCV/CISA	21,49[2]	Complete (I)	OM302320	This study
2016	Seville	Horse	2500_Spain/2016/H-b	2500	Brain (F)	LCV	19,80[2]	Complete (F)	OM302321	This study
2017	Toledo	Golden eagle	3963_Spain/2017/GE-b/2V	3963	Brain (I)	CERI/CISA	13,34[2]	Complete (I)	OM302319	This study
2017	Cáceres	Griffon vulture	B1_Spain/2017/GV-s	B1	Spleen (F)	UEX	34,47[2]	Partial (F)	OM302335	This study
2017	Málaga	Northern goshawk	Spain/2017/NG-b	Spain/2017/NG-b	Brain (I)	N/A	16,60[2]	Complete (I)	MW915462	Aguilera-Sepúlveda et al. 2021
2018	Cáceres	Horse	C79B_Spain/2018/H-b	C79B	Brain (F)	UEX	37,66[2]	Almost complete (F)	OM302313	This study
2018	Cáceres	Little owl	a2153_Spain/2018/LO-l	a2153	Lung (F)	UEX	21,89[2]	Complete (F)	OM302317	This study
2018	Cáceres	Little owl	101279_Spain/2018/LO-l	101279	Lung (F)	UEX	21,54[2]	Partial (F)	OM302336	This study
2018	Cáceres	Little owl	101279_Spain/2018/LO-h	101279	Heart (F)	UEX	26,09[2]	Partial (F)	OM302337	This study
2018	Cáceres	Little owl	101279_Spain/2018/LO-h/2V	101279	Heart (I)	UEX/CISA	15,54[2]	Complete (I)	OM302318	This study
2020	Seville	Human	2_44013532	44013532	Urine (F)	N/A	N/A	Almost complete (F)	HG994353	Pérez-Florido et al. 2021
2020	Seville	Human	4_44013537	44013537	Urine (F)	N/A	N/A	Almost complete (F)	HG994351	Pérez-Florido et al. 2021
2020	Seville	Human	5_44013538	44013538	Urine (F)	N/A	N/A	Almost complete (F)	HG994352	Pérez-Florido et al. 2021
2020	Cádiz	Human	6_44025400	44025400	Urine (I)	N/A	N/A	Almost complete (F)	HG994354	Pérez-Florido et al. 2021
2020	Cádiz	Cinereous vulture	SPA20-01_Spain/2020-01/CV-b/3V	SPA20-01	Brain (I)	LCV/CISA	18,78[2]	Complete (I)	OP713603	This study
2020	Seville	Horse	SPA20-02_Spain/2020-02/H-b/3V	SPA20-02	Brain (I)	LCV/CISA	17,84[2]	Complete (I)	OP713602	This study
2020	Cádiz	Horse	SPA20-05_Spain/2020-05/H-b/4V	SPA20-05	Brain (I)	LCV/CISA	18,19[2]	Complete (I)	OP713601	This study
2020	Seville	*Culex perexiguus*	20C124	20C124	Pool mosquitoes (F)	N/A	N/A	Complete (F)	OP643863	Ruiz-López et al. 2023a
2021	Seville	*Culex perexiguus*	21C560	21C560	Pool mosquitoes (F)	N/A	N/A	Complete (F)	OP643864	Ruiz-López et al. 2023a
2022	Cádiz	*Culex perexiguus*	JA22 423	JA22 423	Pool mosquitoes (F)	N/A	N/A	Complete (F)	OQ357821	Ruiz-López et al. 2023b
2022	Cádiz	*Culex perexiguus*	JA22 544	JA22 544	Pool mosquitoes (F)	N/A	N/A	Complete (F)	OQ357822	Ruiz-López et al. 2023b
2022	Seville	*Culex perexiguus*	22C2392	22C2392	Pool mosquitoes (I)	N/A	N/A	Complete (I)	OQ357819	Ruiz-López et al. 2023b
2022	Seville	*Culex perexiguus*	22C2285	22C2285	Pool mosquitoes (I)	N/A	N/A	Complete (I)	OQ357820	Ruiz-López et al. 2023b

[1]Jiménez-Clavero et al. [39].
[2]Del Amo et al. [40].
LCV = Laboratorio Central de Veterinaria, National Reference Laboratory, Algete, Spain. CISA = Centro de Investigación en Sanidad Animal (CISA-INIA, CSIC), Valdeolmos, Spain. UEX = Veterinary Faculty, University of Extremadura, Cáceres, Spain. CERI = Centro de Estudios de Rapaces Ibéricas, Toledo, Spain.

Due to the limited amount of some samples, a simplified method was designed to characterise the WNV present and to expand the further molecular studies. For that, specific primers were designed for analysis of a 621 nt region of the viral genome. This fragment encodes the entire M protein and the N-terminal region of the E glycoprotein (here called 'PrM-E'), corresponding to amino acids 193-399 of the whole polyprotein. The targeted region holds 3% of nucleotide divergence between two Spanish representative sequences of WMed-1 (SPA12-02, Cádiz 2012) and WMed-2 (GE 2o/V, Toledo 2007), being the most variable and informative genome portion between both sub-clusters.

The designed primer sequences delimiting a fragment of 666 nt were: WNV-639F (5′-CCAGAAGACATCGACTG TTGGTG-3′) and WNV-1305R (5′-CGCATGTGTCAATGCTTC CTTTG-3′), corresponding respectively to the 639 and 1305 nucleotide positions of the SPA12-02 isolate (GenBank accession no. OM302322). One-step RT-PCR kit (Qiagen) was used to run conventional RT-PCR assay according to the manufacturer's instructions. Amplifications were performed following this thermocycling conditions: 30 min at 50 °C, 15 min at 95 °C, 40× (30 s at 94 °C, 45 s at 58 °C, and 1 min at 72 °C), and 7 min at 72 °C.

Amplified PCR products were purified using ExoSAP-IT kit (GE Healthcare) and then bidirectionally sequenced by Sanger's method (Big Dye Terminator Cycle Sequencing kit (version 3.1), in an ABI 3730 XL DNA Analyser (Applied Biosystems).

Sequence and phylogenetic analyses

Analysis and assembly of the full linear genomes were performed using SeqMan software (DNASTAR. Madison, WI). PrM-E genome sequences were also edited by SeqMan. The consensus sequences obtained were compared with other 45 full or partial genome sequences of WNV-L1, respectively, belonging to the Mediterranean and Middle East regions, and Senegal, available in GenBank. Multiple sequence alignments were performed by ClustalW tool (available in MEGA7). Phylogenetic trees were generated by the maximum likelihood algorithm, using TN93 + G as the optimal nucleotide substitution model provided by MEGA7 for both whole and PrM-E genome sequences analysis. Bootstrap values were inferred for 1000 replicates.

Further genetic and amino acidic homology studies were carried out with the support of MEGA7 software by estimating the evolutionary divergence between sequences and between the established WMed groups. These analyses were conducted using the maximum composite likelihood method.

In order to get an advanced analysis, visualisation and exploration of phylogenetic and genomic data together with geographic and temporal data (or other relevant metadata available), the 58 WNV complete genome sequences employed in this study were analysed through the Nextstrain [42] generic build available at the online bioinformatics platform INSaFLU-TELEVIR (https://insaflu.insa.pt/; v.2.0.0; https://github.com/INSaFLU/nextstrain_builds) [43]. Mutation annotation refers to the GenBank reference genome AF260968.1 (strain Eg101, Egypt 1951). The interactive Nextstrain JSON file (including phylogenetic, geotemporal and mutation data) and metadata table are available as Dataset S1, for navigation in https://auspice.us/.

Finally, the algn2pheno (https://github.com/insapathogenomics/algn2pheno; v1.1.5) tool was applied to extract the whole repertoire of amino acid changes of the 58 WNV complete genome sequences, taking as input protein-specific alignments retrieved from a full genome alignment using nextalign (https://github.com/neherlab/nextalign; v.2.11.9). The algn2pheno report with amino acid changes found *per* sample *per* protein, in comparison to the reference genome AF260968.1 (strain Eg101, Egypt 1951), is available as Table S2.

Results

Virus isolation

Eight WNV-L1 isolates were obtained out of the nineteen samples that were subjected to virus isolation in Vero cells: 2442-4, 2442-5, 2500, 3963, 101279, SPA20-01, SPA20-02 and SPA20-05 (Table 1).

Spanish WNV-L1 sequences

From the 27 WNV-L1 positive samples subjected to sequencing, 13 full genome linear sequences of 10,220 (id. C79B), 10,926 (id. 2442-5), 10,931 (id. SPA20-02 and SPA20-05), 10,932 (id. 2439, 2442-4, 101279, a2153, 2500 and 3963), 10,948 (id. 2442-5), 10,963 (id. SPA20-01) and 10,979 (id. 2500) nucleotides were obtained. This panel represents 50% of all Spanish full-length WNV-L1 sequences described up to date ($n = 26$). In addition, 14 samples were partially sequenced obtaining a fragment of 621 nt of the PrM-E coding region. Table 1 compiles all Spanish WNV-L1 sequences obtained from 2007 to 2022 from this and previous studies, indicating species,

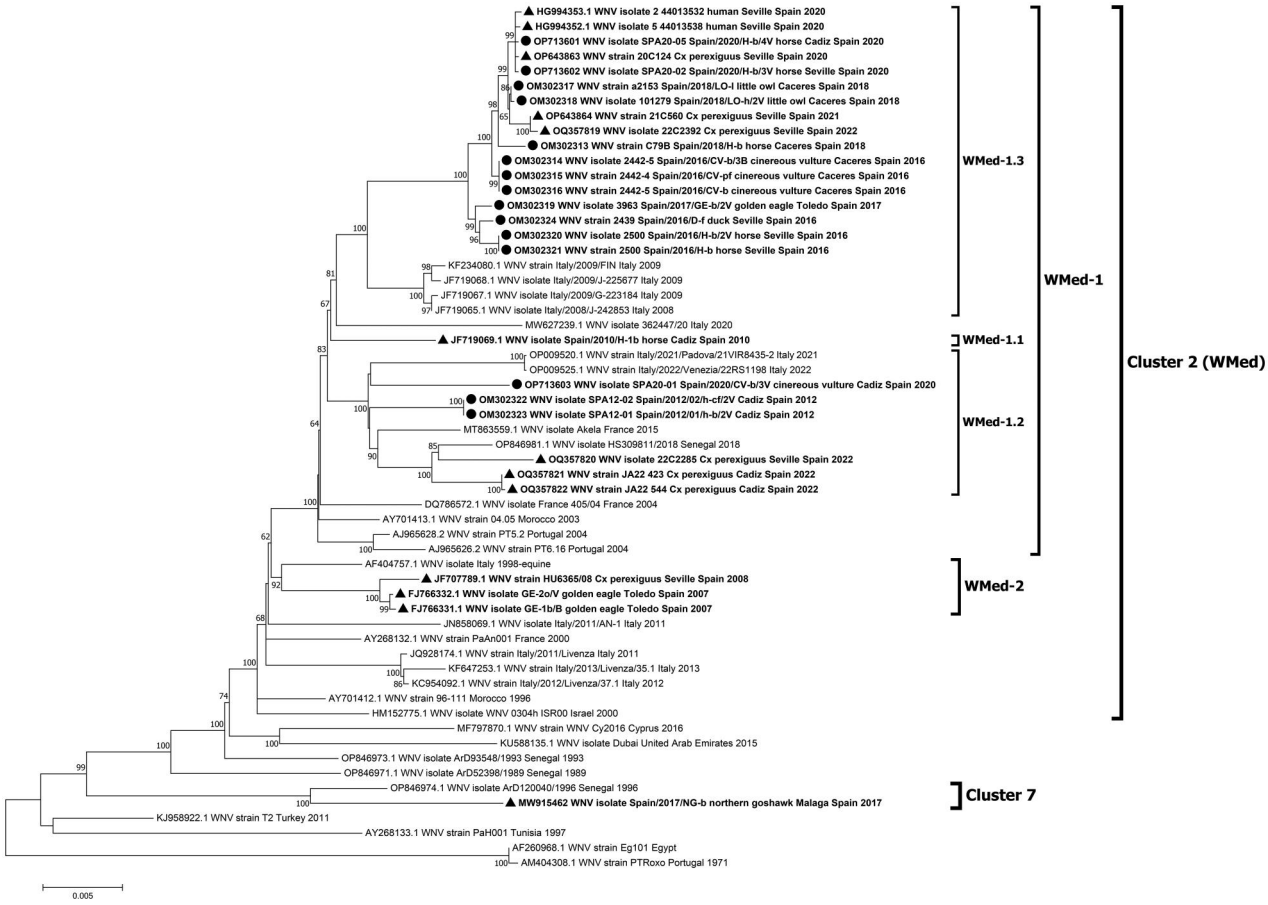

Figure 2. Phylogenetic analysis of complete genome nucleotide sequences of WNV-L1 from Western Mediterranean region. Viral sequences are identified by GenBank accession number, sample code, species, Spanish province, country, and year of isolation. Sequences emphasised in bold and with a black circle (●) were generated during this study. Sequences emphasised in bold and with a black triangle (▲) are previously published Spanish sequences. Percentages of successful bootstrap replicates over 60% are indicated at tree nodes.

type of specimen and province from where they were collected. Figure 1 shows the provinces where WNV-L1 positive samples have been collected during this study.

The nucleotide sequences obtained in this study ($n = 27$) have been lodged in the GenBank sequence database under accession nos. OM302313-OM302321, OM302324-OM302338, OP713601-OP713603.

Phylogenetic analysis of the Spanish WNV complete genome sequences

Phylogenetic relationships between the obtained full genome sequences and other WNV-L1 complete sequences available in GenBank were established. The new 13 sequences were compared with 45 WNV-L1 isolates, including previously available Spanish isolates ($n = 15$), and sequences from Senegal and other Mediterranean and Middle East countries ($n = 30$) (Figure 2). Overall, the sequences included in the analysis encompass a time period covering from 1989 to 2022, with two exceptions (Egypt, 1951 and Portugal, 1971) used as outgroup.

Phylogenetic analyses placed the new complete Spanish sequences in the defined cluster 2 of WNV clade 1a, also known as Western Mediterranean (WMed) cluster, and in particular within the Western Mediterranean-1 (WMed-1) sub-cluster [8]. This monophyletic group also included other related Spanish variants previously described, collected in 2010, 2012, 2020, 2021 and 2022, as well as French (2015) and Senegalese (2018) representatives, and ancient (2008-2009) and recent (2020-2022) Italian strains. Remarkably, the recent Spanish isolates differ from the earliest ones (2007-2008), which group in the Western Mediterranean-2 (WMed-2) subcluster [8]. Besides, all of them widely differ from the Spanish isolate detected in Málaga in 2017, which belongs to cluster 7, within WNV-L1 clade 1a [23,31].

A closer examination of the phylogenetic tree shown in Figure 2 reveals that WMed-1 diverged into at least three main genetic groups, here denoted WMed-1.1, WMed-1.2 and WMed-1.3, based on chronology (earliest

isolate in each group). Notably, WMed-1.2 showed a highly diverse composition, both temporally and spatially, comprising sequences of Spain, from Cádiz and Seville, (2012, 2020, 2022), France (2015), Italy (2021-2022) and Senegal (2018). Meanwhile, WMed-1.3 formed a well-defined clade of closely related Spanish sequences from most recent cases (2016-2022), originating in four provinces: Seville, Cádiz, Cáceres and Toledo. These included isolates causing disease in wild birds, horses and humans, and isolates from *Culex perexiguus* pools obtained in Seville in the last seasons (2020-2022) [17]. Interestingly, earlier Italian isolates (2008-2009) clustered within this group.

Comparative genome analysis

Pairwise comparison of new Spanish full-length sequences collected between 2007 and 2022 revealed nucleotide divergences ranging from 0.01% to 4.72%. Isolates 101279 and a2153 (both from Cáceres, 2018) were the most closely related, showing 99.99% nucleotide identity. On the other hand, the isolate Spain/2017/NG-b (Málaga, 2017) and C79B (Cáceres, 2018) exhibited the highest divergence, with 95.28% nucleotide identity.

The mean nucleotide identity between WMed-1 and WMed-2 was estimated to be 97.81%, while dropped to 95.46% or 95.85% between the outgrouped Spain/2017/NG-b isolate and WMed-1 or WMed-2, respectively. Within the different Spanish WMed-1 groups, the mean nucleotide distance was estimated at 0.0172 between WMed-1.1 and WMed-1.2, 0.0175 between WMed-1.1 and WMed-1.3, and 0.0209 between WMed-1.2 and WMed-1.3 groups; in all these cases, the overall identity surpassed 97.90%.

Analysis of partial genome sequences (PrM-E)

In this study, a practical sequencing method targeted at a region encoding 621 nt of the PrM-E proteins was developed, which was successfully applied to samples where full genome sequencing attempts failed. This strategy enabled to add 14 new samples to the phylogenetic analysis, resulting in a total of 27 linear sequences of 621 nt (nucleotide positions 662 to 1283 of SPA12-02), covering the period from 2010 to 2020 (Table 1).

Overall, the phylogenetic tree obtained with these 27 Spanish PrM-E sequences, along with those from the 45 WNV-L1 isolates previously employed reproduced the same topology as in the phylogenetic analysis based on complete WNV-L1 sequences (Figure 3). Phylogenetic analysis placed all except one of the 14 new PrM-E sequences within the defined WMed-1 sub-cluster.

A closer examination of the PrM-E phylogenetic analysis revealed further details of WMed-1. First, WMed-1.1 group comprised only two equine samples collected in 2010 (Spain/2010/H-1b) and 2012 (id. 2712) in the same province (Cádiz). In turn, WMed-1.2 revealed a more complex history, starting with equine and avian cases in Cádiz in 2010-2012, then spotting in France (2015) and Senegal (2018), and later re-emerging in 2020 in Spain (again in Cádiz) where it remained later in circulation, as seen in mosquito pools from Cádiz and Seville in 2022. Moreover, WMed-1.2 variants were also identified in Italy (2021-2022) for the first time, causing significant human outbreaks [44]. Regarding WMed-1.3, this analysis added up one more Spanish province (Huelva) to the already wide territory in Spain from which the abundant sequences of this group were found. In addition, the PrM-E study includes sequences from equine and avian cases occurring in 2013 in Seville and Huelva, indicating that this variant may have arisen in 2013 or earlier.

This phylogenetic analysis also displayed two outlier sequences, assigned neither to WMed-1 nor to WMed-2. The first one corresponds to a red-legged partridge collected in Cádiz in 2011 (sample id. 3011). This sequence fitted in the WMed cluster, but clearly differed from any of the monophyletic groups from the Mediterranean Basin. The second exception, Spain/2017/NG-b isolate (Málaga, 2017), also revealed by full genome analysis, which is phylogenetically distant to any other known Mediterranean isolate [31].

Analysis of the viral polyproteins

A study of the viral polyproteins (Table 2) was performed comparing all the available Spanish WNV-L1 representatives ($n = 21$, 2007-2022), excluding only those from the 2020 human outbreak ($n = 4$), due to incomplete polyprotein sequences in GenBank. This study revealed amino acidic divergence ranging from 0.03% among the closest isolates (representing a difference in one amino acidic position) to 0.88% (observed between the Spain/2017/NG-b isolate and HU6365/08), corresponding to a divergence in thirty amino acidic positions.

Comparing the earliest Spanish isolates (2007-2008), the unique belonging to the WMed-2 sub-cluster, with the rest of Spanish sequences, a specific signature was identified in these isolates, which is characterised by

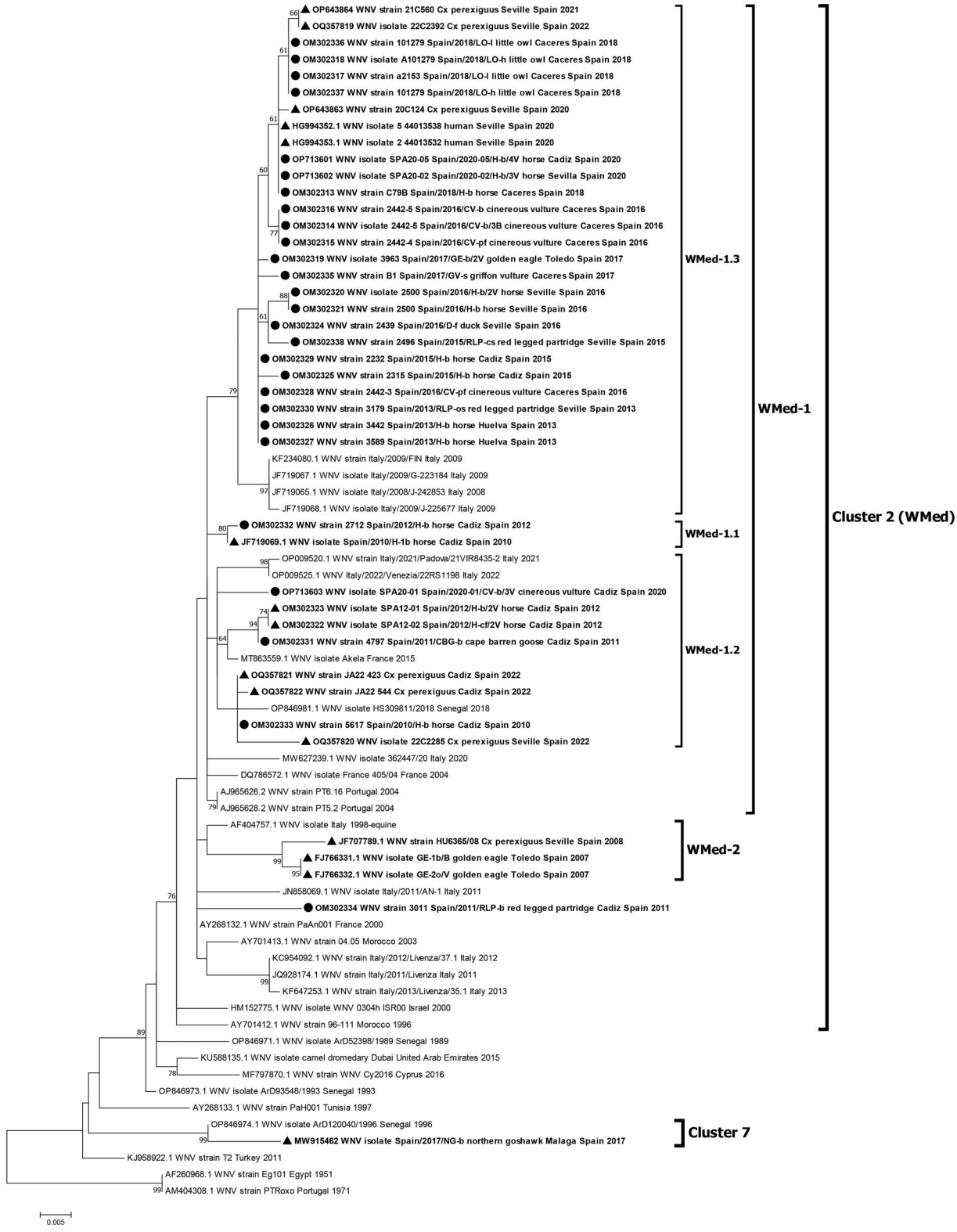

Figure 3. Phylogenetic analysis of partial genome (621 nt) nucleotide sequences of WNV-L1 from Western Mediterranean region. Viral sequences are identified by GenBank accession number, sample code, species, Spanish province, country, and year of isolation. Sequences emphasised in bold and with a black circle (●) were generated during this study. Sequences emphasised in bold and with a black triangle (▲) are previously published Spanish sequences. Percentages of successful bootstrap replicates over 60% are indicated at tree nodes.

eight unique amino acid positions: C-M34V, E-A51T, E-P88S, NS1-Y35H, NS1-D208H, NS1-E289G, NS4B-A115T and NS5-M436I (numbering of aa changes refers to the AF260968.1, Egypt 1951 reference genome annotation) (Dataset S1 and Table S2). Furthermore, these eight exclusive positions were not found in the rest of the

Table 2. Comparison of the amino acid substitutions between all the complete genomes of Spanish WNV-L1 isolates published up to date.

		Cluster 2 (WMed)																					Cluster 7
		WMed-2		WMed-1																			
		Initial Reference		WMed-1.1	WMed-1.2					WMed-1.3													
		FJ766331																					
		Toledo	JF707789	JF719069	OM302323	OP713603	OQ357822	OQ357821	OQ357820	OM302324	OM302320	OM302314	OM302319	OM302313	OM302317	OM302318	OP713602	OP713601	OP643863	OP643864	OQ357819	MW915462
		2007	Seville	Cádiz	Cádiz	Cádiz	Cádiz	Cádiz	Seville	Seville	Seville	Cáceres	Toledo	Cáceres	Cáceres	Cáceres	Seville	Cádiz	Seville	Seville	Seville	Málaga
	Amino	GE-1b/B	2008	2010	2012	2020	2022	2022	2022	2016	2016	2016	2017	2018	2018	2018	2020	2020	2020	2021	2022	2017
Viral	acid	(Golden	HU6365/08	Spain/2010/H-1b	SPA12-01	SPA20-01	JA22 544	JA22 423	22C2285	2439	2500	2442-5	3963	C79B	a2153	101279	SPA20-02	SPA20-05	20C124	21C560	22C2392	Spain/2017/NG-b
protein	position	eagle)	(*Cx. perexiguus*)	(Horse)	(Horse)	(Cinereous vulture)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(Duck)	(Horse)	(C. Vulture)	(Golden eagle)	(Horse)	(Little owl)	(Little owl)	(Horse)	(Horse)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(N. Goshawk)
C	34	V	.	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
	52	A	.	.	.	.	.	.	.	.	.	T	.	T	T	T	T	T	T	T	T	.
	95	I	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.	.	.	.
	109	T	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	110	G	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	115	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V
	117	L	.	.	.	.	.	.	.	M	M	M	M	M	M	M	M	M	M	M	M	.
PrM	20	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	43	Y	.	.	.	.	.	.	.	.	.	.	.	H	.	.	.	.	.	.	.	.
	94	L	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	156	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
E	15	V	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
	24	V	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	50	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.
	51	T	.	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
	84	K	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	88	S	.	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
	93	R	.	.	K	K	K	K	K	.	.	.	.	.	.	.	.	.	.	.	.	.
	293	E	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
	481	I	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NS1	35	H	.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	70	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
	83	G	.	.	.	.	D	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	88	V	.	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	170	K	.	.	.	.	.	.	.	.	.	R	.	R	R	R	R	R	R	R	R	.
	175	N	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.
	208	H	.	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
	289	G	.	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
	338	Q	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	339	R	.	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.
NS2A	32	I	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.
	119	H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y
	142	I	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	158	L	.	V	.	.	M	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	187	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V
	211	G	.	.	.	.	.	.	.	R	R	R	R	R	R	R	R	R	R	R	R	.
NS2B	84	N	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.
	92	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	L
	102	I	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	V
	119	I	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.	.	.	.	.
NS3	131	P	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	143	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S
	177	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	V	.
	249	P	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

(*continued*)

Table 2. Continued.

		Cluster 2 (WMed)																						Cluster 7
		WMed-2		WMed-1																				
		Initial Reference FJ766331		WMed-1.1	WMed-1.2					WMed-1.3														
		Toledo	JF707789	JF719069	OM302323	OP713603	OQ357822	OQ357821	OQ357820	OM302324	OM302320	OM302314	OM302319	OM302313	OM302317	OM302318	OP713602	OP713601	OP643863	OP643864	OQ357819	MW915462
		2007	2008	2010	2012	2020	2022	2022	2022	2016	2016	2016	2017	2018	2018	2018	2020	2020	2020	2021	2022	2017
	Amino	GE-1b/B	HU6365/08	Cádiz	Cádiz	Cádiz	Cádiz	Cádiz	Seville	Seville	Seville	Cáceres	Toledo	Cáceres	Cáceres	Cáceres	Seville	Cádiz	Seville	Seville	Seville	Málaga
Viral	acid	(Golden		Spain/2010/H-1b	SPA12-01	SPA20-01	JA22 544	JA22 423	22C2285	2439	2500	2442-5	3963	C79B	a2153	101279	SPA20-02	SPA20-05	20C124	21C560	22C2392	Spain/2017/NG-b
protein	position	eagle)	(*Cx. perexiguus*)	(Horse)	(Horse)	(Cinereous vulture)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(Duck)	(Horse)	(C. Vulture)	(Golden eagle)	(Horse)	(Little owl)	(Little owl)	(Horse)	(Horse)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(*Cx. perexiguus*)	(N. Goshawk)
	253	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S
	258	V	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	338	T	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	436	T	.	.	.	.	.	.	.	A	A	A	A	A	A	A	A	A	A	A	A	A
	453	A	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	465	N	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	583	I	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NS4A	5	L	.	.	.	.	.	.	.	.	.	.	.	F	.	.	.	.	.	.	.	.
	6	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V
NS4B	11	S	.	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	13	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	L	.
	21	I	.	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	23	V	.	A	.	.	.	.	.	.	.	I	.	.	.	.	.	.	.	.	.	.
	50	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.
	70	T	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
	115	T	.	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
	241	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I
	245	I	.	.	.	.	.	.	.	M	M	.	.	.	.	V	.	.	.	.	.	.
	249	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	E
NS5	33	I	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	41	K	.	.	.	.	.	.	.	R	R	R	R	R	R	R	R	R	R	R	R	.
	88	C	.	S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	167	M	.	.	.	.	.	.	.	.	.	.	.	I	.	.	.	.	.	.	.	.
	194	K	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	280	K	.	.	.	.	R	R	R	.	.	.	.	.	.	.	.	.	.	.	.	.
	422	R	.	.	.	.	.	.	.	K	K	K	K	K	K	K	K	K	K	K	K	.
	426	A	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
	436	I	.	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
	450	H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y
	503	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q

The first Spanish isolate from 2007 was used as the reference sequence. Black dot indicates the same amino acid as the reference sequence. Amino acids in Spanish WNV-L1 polyproteins that are different from the initial reference are highlighted in bold.

Colours indicate the amino acid category: yellow = hydrophobic; green = polar; pink = acidic; blue = basic.

sequences analysed in this study, suggesting that this signature could be characteristic for these three Spanish WMed-2 isolates. Meaningfully, Spanish sequences belonging to WMed-1 sub-cluster and the dissimilar Spain/2017/NG-b share the same amino acid residues in all these eight positions (Table 2).

On the other hand, Spanish isolates from WMed-1.3 presented three exclusive amino acid changes: C-L117M, NS2A-G211R and NS5-K41R. Moreover, NS5-R422K seems to be a position that defines WMed-1.3 group, as both Spanish and Italian representatives of this group share this amino acidic signature. Notably, WMed-1.3 isolates from Seville (2016) presented a distinctive amino acid change: NS4B-I245M. Similarly, WMed-1.2 group seems to be defined by the single amino acid change E-R93K, even though this change is also observed in the Israel 2000 isolate. Remarkably, WMed-1.2 mosquito sequences from Cádiz (2022) presented three particular amino acid changes: NS1-G83D, NS2A-L158M and NS5-I33T. Finally, the dissimilar Spain/2017/NG-b isolate from Málaga is characterised by seven unique amino acids, which differ from the rest of the 58 isolates analysed (C-I115V, NS2A-I187V, NS2B-I92L, NS3-N143S, NS3-N253S, NS4A-I6V and NS4B-T241I) (Table 2, Dataset S1 and Table S2).

Discussion

Although WNV-L1 has been reported in Spain since the early 2000s [28], it was not until the next decade when its impact grew as it spread to increasingly wider territories and outbreaks became more and more frequent, a trend that was also observed in other Mediterranean countries affected. In an early study, it was found that all WNV-L1 strains circulating in the Western Mediterranean countries up to 2010 constituted a single monophyletic cluster (WMed), reflecting the ability of the virus to establish and remain circulating in the area, evolving into at least two different groups, named WMed-1 and WMed-2 [8]. The present study takes advantage of ten years of WNV surveillance in animals in Spain to gather the most complete collection of WNV-L1 sequences from the Western Mediterranean area as of today. Overall, this study provides 27 new WNV-L1 sequences of which 13 are full genomes. The other 14 sequences are partial genomes that were resolved using a new sequencing method, targeting PrM-E coding region, intended to provide high-resolution information about closely related WMed variants, which reproduced the topology observed using complete genomes, thereby validating this approach. This method was developed as a rapid WMed WNV-L1 genotyping assay for samples reluctant to full-length sequencing. Other studies have used similar approaches, to characterise and classify close related WNV-L1 viruses, in diverse territories such as the USA [45], India [46], Australia [47] or Israel [25,48]. Hence, in this section both full and partial genome analyses will be discussed together.

Our phylogenetic analyses confirm the topology previously observed [49], gathering almost all Spanish sequences in cluster 2 (WMed subtype) within clade 1a. The single exception was Spain/2017/NG-b (Málaga 2017), a dissimilar isolate classified in cluster 7 [31]. A closer examination on the tree topologies suggest that WNV-L1 strains are recurrently introduced into Spain. The first introduction is represented by two sequences from Toledo in 2007 [26]. This strain remained circulating the following season and spread to other territories, since a closely related sequence was identified in Seville in 2008. These sequences belong to the WMed-2 sub-cluster and are closely connected with an Italian WNV strain from 1998 [26]. Interestingly, WNV from this sub-cluster seems to have extinguished, as no related sequences have appeared after 2008, despite many other variants have been identified. The second incursion debuts with an isolate obtained from the first horse outbreak occurred in 2010 in Cádiz, classified in the WMed-1 sub-cluster [8]. Remarkably, this sub-cluster showed to be highly successful at establishing and spreading; it encompasses sequences widely distributed in all the Spanish provinces examined, not only spanning the whole decade (2010-2020) but also including sequences from preceding outbreaks in Italy in 2008 and 2009, as well as sequences from other neighbouring (Portugal, France, Morocco) and non-neighbouring (Senegal) countries. The inclusion of partial PrM-E sequences allowed further discrimination within this sub-cluster, which divided WMed-1 into three subgroups, WMed-1.1, WMed-1.2, and WMed-1.3, each of which probably arose as an independent introduction from other Western Mediterranean areas. Once in Spain, they followed different fates. WMed-1.1 seems to have been restricted to horse outbreaks that occurred in Cádiz between 2010 and 2012. Meanwhile, WMed-1.2 might have appeared also in Cádiz causing outbreaks in horses in 2010, 2011 and 2012, and, after 8 years undetected, it reappeared in 2020, again in Cádiz. After that, this variant was detected in Cádiz and Seville in 2022. In turn, WMed-1.3 might have reached the South of Spain around 2013, and then spread through at least

five provinces (Cádiz, Huelva, Seville, Toledo and Cáceres), causing widespread outbreaks in avian, equine and human hosts, including the most severe outbreak occurring in humans in Spain in 2020 [16]. Interestingly, Spanish sequences in this WMed-1.3 group are closely connected to the 2008-2009 Italian isolates, indicating a possible origin.

Two more introductions may also be hypothesised based on two singular sequences; (1) the divergent WMed sample id. 3011 (Cádiz 2011) likely a distinct WNV-L1 introduction from an unknown Western Mediterranean region around 2011, and (2) the isolate obtained in 2017 in Málaga [31], which roots with an ancient isolate from Senegal (1996), both belonging to cluster 7 within clade 1a, apparently extra-Mediterranean [23,31].

Overall, these results allowed to set five basic principles that contribute to explain the WNV-L1 dynamics observed in Spain as a part of the Western Mediterranean, namely: (1) introduction; (2) endemization, spread and evolution; (3) co-circulation; (4) extinction; (5) re-emergence after silent circulation and alternation.

(1) *Introduction*. WNV-L1 is able to reach occasionally the South of Spain and settle down. Interestingly, 4 out of the 6 proposed introductions are supported by sequences from Cádiz, the southernmost province of Spain and the closest point of continental Europe to Africa. Indeed, the earliest representative in Spain from each variant of WMed-1 was found in Cádiz with the only exception of WMed-1.3, suggesting this province as a hotspot for the introduction of WNV in Spain. In fact, other mosquito-borne epornitic flaviviruses such as Bagaza [50] or Usutu [30] have emerged or have been detected in Cádiz, respectively. In consequence, it is advisable to reinforce flavivirus surveillance in the area in order to prevent future outbreaks.

The introduction of WNV into Europe through migratory birds from their African wintering grounds is a plausible hypothesis [4,22,23,51]. Indeed, recently described close-related Tunisian (2016) and Italian (2016-2017) isolates belonging to WMed-1 support this idea [22]. In fact, billions of birds move yearly between Europe and Africa [23], so near countries from both continents might share common or close-related viral strains. In this regard, published Senegalese sequences [5] were incorporated to the phylogenetic analyses showing a close relationship with some Spanish representatives, supporting the hypothesis above mentioned. This situation emphasises the need of coordinated surveillance actions among European and African countries in order to early detect circulating strains and to prevent national and international spread of WNV, as has also been recently proposed [23].

On the other hand, this study shows a close relationship between Italian and Spanish WNV-L1 isolates from the two main WMed sub-clusters. Similarly, WNV-L2 strains that are currently circulating in North-Eastern Spain (2017-2020) are related to an Italian cluster which also spread to France [36]. Both events strongly support the idea that WNV circulating in the Western Mediterranean can persist for long periods and move back and forth through this extensive area. Furthermore, the strong connection between Italian and Spanish strains could also be associated with common migratory bird routes [2,23,52]. However, the role of migratory birds in WNV dispersal needs to be carefully weighted with the evidence available. Indeed, alternative hypothesis such as animal trade (legal or illegal) should not be disregarded.

(2) *Endemization, spread and evolution*. Besides the ability of WNV-L1 strains to recurrently reach, establish and develop endemic cycles in Spain, some of them were shown to persist for long periods and spread to neighbouring provinces (e.g. WMed-1.3) or even countries (e.g. WMed-1.2), while others vanish shortly after being detected (e.g. WMed-1.1). As for their evolution, variants from the same territory and year normally cluster together and diverge from those circulating elsewhere. In fact, the specific signatures found in the polyprotein, e.g. within WMed-1.3, when comparing strains from distant provinces (Seville, Toledo, Cáceres) suggest adaptive evolution to different scenarios in the same way as they reflect the temporal evolution in a same territory. Similarly, the polyproteins of the Spanish isolates belonging to the WMed-1.3 variant share a specific amino acid signature that differs from the Italian representatives belonging the same WMed-1.3 group. Likewise, Spanish WNV-L2 isolates (2017-2020) share some exclusive amino acid positions within their phylogenetic Lombardy cluster, which contains isolates from this Italian region (2013-2018) and France (2018) [36]. All the same, more studies will be necessary in order to go deeper in the evolutionary history of this complex virus.

(3) *Co-circulation*. As a consequence of divergence and spread, different variants can co-circulate in the same area, which is also supported in this study. Specifically, strains belonging to WMed-1.1 and WMed-1.2 co-circulated at least during 2010 and 2012 in Cádiz province, while the same occurred for WMed-1.2 and WMed-1.3 in Cádiz in 2020 and in Seville in 2022. On the other hand, more distant variants, namely Spain/

2017/NG-b (Málaga 2017) or 3011 (Cádiz 2011), co-circulated during the same seasons as WMed-1.3 (Toledo and Cáceres 2017) or WMed-1.2 (Cádiz 2011), respectively. In fact, these distant newly introduced strains were apparently not successful at establishing enzootic cycles, which may be due to competition with the locally-adapted variants, perhaps more fitted for the circulation in these ecological settings.

(4) *Extinction*. Evidence for extinction of specific WNV strains has also been provided e.g. by the WMed-2 sub-cluster, first detected in Italy in 1998 and then arising a decade later (2007-2008) in Spain. The molecular analysis suggests that this sub-cluster, causing initial outbreaks in both countries, disappeared after a while, as no new related isolates have been reported in the last 15 years, despite intense surveillance. Additionally, the study of the polyprotein also revealed eight exclusive amino acid positions that characterise only WMed-2 Spanish isolates. These exclusive positions may represent an unsuccessful trial for adaptation to local conditions, leading to the extinction of this WMed-2 sub-cluster. In fact, other variants included in the study (belonging to different groups, sub-clusters and even clusters) share the same amino acidic substitutions in those specific positions, suggesting that their conservation is important for adaptation and survival in the conditions locally prevailing. Furthermore, as remarked above, other strains such as Spain/2017/NG-b (Málaga 2017) or 3011 (Cádiz 2011), as well as those within the WMed-1.1 group, detected transiently in Spain, may have become extinct too, possibly due to similar constrains in their particular amino acidic substitutions along their polyprotein.

(5) *Re-emergence after silent circulation and alternation*. Evidence is provided that after a period of latency, a particular variant can re-emerge in the same territory. This might be explained by silent circulation during a period of time that may span several years, or, alternatively, by re-introduction of a closely similar variant in the territory from another geographical location. Interestingly, WMed-1.2 variant that circulated in Cádiz during three seasons (2010-2012), remained unnoticed during eight years, after which it re-emerged in the same province. Notably, this variant was detected in France (2015), Senegal (2018) and Italy (2021-2022) for the first time [2,5,9,23], evidencing a strong ability to spread as well as pointing out a remarkable network of geographical connections for this type of WNV-L1 strains, going back and forth between Africa and Europe.

This exhaustive phylogenetic study on WNV-L1 from Spain provides a wide perspective of the epidemiology, evolution and ecology of this particular WNV lineage, involving a wider territory including Western Europe and Northern Africa-Sahel. By focusing sampling on diseased animals, the study may have biased the detection of strains towards those most pathogenic, therefore missing less pathogenic variants. In fact the study reveals that some strains may be circulating silently for years, remaining unnoticed until a spill-over occurs under certain scenarios [53]. In this regard, mosquito, human and animal surveillance complement each other, and therefore, biases affecting the pathogenicity of the strains are expected to be avoided.

To conclude, the analysis of this important collection of Spanish sequences (2010-2020) from a wide range of vertebrate hosts demonstrates that certainly more than one genetic variant of WNV-L1 is still present and actively circulating in South-Western Spain. These WNV-L1 strains are shared over a wide territory involving the Western Mediterranean and the Sahel areas, with likely contribution of wild birds in the dispersal. Overall, WNV-L1 strains pose an increasing risk of spill-over into the human populations living in these territories. The high seroprevalence levels found in birds and horses in the same regions where the studied sequences originated corroborates this view [29,33,54–56]. Indeed, human cases may appear after detection in animals [57], as was recently observed in 2020 in the country [16]. These spill-overs may recur in future transmission seasons, so it is of utmost importance to reinforce animal and mosquito (i.e. One Health) surveillance and to refine diagnostic tools, especially in those areas with an increased risk of WNV circulation.

Acknowledgments

We thank Amalia Villalba, Ana Robles and María del Carmen Barbero for their technical support at CISA facilities and Encarnación Madueño for sequencing the samples at CISA. We thank Centro de Recuperación de Fauna y Educación Ambiental 'Los Hornos' for providing some of the samples from Cáceres. In addition, we thank Natalia Romeral-Requena (Centro de Estudios de Rapaces Ibéricas, CERI) for providing the sample from Toledo. Finally, we thank Daniel Sobral (National Institute of Health Doutor Ricardo Jorge, INSA) for implementing Nextstrain pipeline in INSaFLU. This work was supported by the Spanish INIA-MAPA agreements EG17-141 and AEG21-198, by the INIA grant E-RTA2015-00002-C02-00, AEI grant PID2020-116768RR-C21 and by Extremadura regional project IB16135. Pilar Aguilera-Sepúlveda was funded by a FPI predoctoral fellowship from INIA. Vítor Borges work was co-financed through the DURABLE project. The

DURABLE project has been co-funded by the European Union, under the EU4Health Programme (EU4H), Project no. 101102733. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This study was supported by Ministerio de Agricultura, Pesca y Alimentación (EG17-141 and AEG21-198); Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (E-RTA2015-00002-C02-00); Agencia Estatal de Investigación (PID2020-116768RR-C21); Junta de Extremadura (IB16135); European Union-EU4 Health Programme (101102733).

ORCID

Pilar Aguilera-Sepúlveda http://orcid.org/0000-0002-2587-528X
Cristina Cano-Gómez http://orcid.org/0000-0003-4979-983X
Vítor Borges http://orcid.org/0000-0003-3767-2209
Daniel Bravo-Barriga http://orcid.org/0000-0003-3481-7310
Eva Frontera http://orcid.org/0000-0003-0222-7850
Miguel Ángel Jiménez-Clavero http://orcid.org/0000-0003-2125-9743
Jovita Fernández-Pinero http://orcid.org/0000-0001-9919-0112

References

[1] Chancey C, Grinev A, Volkova E, et al. The global ecology and epidemiology of West Nile virus. Biomed Res Int. 2015; 2015:376230. doi: 10.1155/2015/376230.

[2] Beck C, Goffart IL, Franke F, et al. Contrasted epidemiological patterns of West Nile virus lineages 1 and 2 infections in France from 2015 to 2019. Pathogens. 2020;9(11): 908. doi: 10.3390/pathogens9110908.

[3] Pérez-Ramírez E, Llorente F, Del Amo J, et al. Pathogenicity evaluation of twelve West Nile virus strains belonging to four lineages from five continents in a mouse model: discrimination between three pathogenicity categories. J Gen Virol. 2017;98(4):662–670. doi: 10.1099/jgv.0.000743.

[4] Mencattelli G, Ndione MHD, Rosà R, et al. Epidemiology of West Nile virus in Africa: an underestimated threat. PLoS Negl Trop Dis. 2022;16(1):e0010075. doi: 10.1371/journal.pntd.0010075.

[5] Ndione MHD, Ndiaye EH, Faye M, et al. Re-Introduction of West Nile virus lineage 1 in Senegal from Europe and subsequent circulation in human and mosquito populations between 2012 and 2021. Viruses. 2022;14(12):2720. doi: 10.3390/v14122720.

[6] Beck C, Jimenez-Clavero MA, Leblond A, et al. Flaviviruses in Europe : complex circulation patterns and their consequences for the diagnosis and control of West Nile disease. Int J Environ Res Public Health. 2013;10(11):6049–6083. doi: 10.3390/ijerph10116049.

[7] Murgue B, Murri S, Triki H, et al. West Nile in the Mediterranean basin: 1950-2000. Ann N Y Acad Sci. 2006; 951(1):117–126. doi: 10.1111/j.1749-6632.2001.tb02690.x.

[8] Sotelo E, Fernández-Pinero J, Llorente F, et al. Phylogenetic relationships of Western Mediterranean West Nile virus strains (1996-2010) using full-length genome sequences: single or multiple introductions? J Gen Virol. 2011;92(Pt 11):2512–2522. doi: 10.1099/vir.0.033829-0.

[9] Mencattelli G, Iapaolo F, Fusco G, et al. West Nile virus lineage 1 in Italy : newly introduced or a re-occurrence of a previously circulating strain ? Viruses. 2022;14(1):64. doi: 10.3390/v14010064.

[10] Barrett ADT. West Nile in Europe: an increasing public health problem. J Travel Med. 2018;25(1):30289526. doi: 10.1093/jtm/tay096.

[11] Vilibic-Cavlek T, Savic V, Petrovic T, et al. Emerging trends in the epidemiology of West Nile and Usutu virus infections in Southern Europe. Front Vet Sci. 2019;6:437. doi: 10.3389/fvets.2019.00437.

[12] Bakonyi T, Ivanics É, Erdélyi K, et al. Lineage 1 and 2 strains of encephalitic West Nile virus, Central Europe. Emerg Infect Dis. 2006;12(4):618–623. doi: 10.3201/eid1204.051379.

[13] Hernández-Triana LM, Jeffries CL, Mansfield KL, et al. Emergence of West Nile virus lineage 2 in Europe: a review on the introduction and spread of a mosquito-borne disease. Front Public Health. 2014;2:271. doi: 10.3389/fpubh.2014.00271

[14] Popescu CP, Florescu SA, Cotar AI, et al. Re-emergence of severe West Nile virus neuroinvasive disease in humans in Romania, 2012 to 2017–implications for travel medicine. Travel Med Infect Dis. 2018;22:30–35. doi: 10.1016/j.tmaid.2018.03.001.

[15] Christova I, Papa A, Trifonova I, et al. West Nile virus lineage 2 in humans and mosquitoes in Bulgaria, 2018–2019. J Clin Virol. 2020;127:104365. doi: 10.1016/j.jcv.2020.104365.

[16] García San Miguel Rodríguez-Alarcón L, Fernández-Martínez B, Sierra Moros MJ, et al. Unprecedented increase of West Nile virus neuroinvasive disease, Spain, summer 2020. Eurosurveillance. 2021;26(19):2002010. doi: 10.2807/1560-7917.ES.2021.26.19.2002010.

[17] Ruiz-López MJ, Muñoz-Chimeno M, Figuerola J, et al. Genomic analysis of West Nile virus lineage 1 detected in mosquitoes during the 2020–2021 outbreaks in Andalusia, Spain. Viruses. 2023;15(2):266. doi: 10.3390/v15020266.

[18] Ruiz-López MJ, Aguilera-Sepúlveda P, Cebrián-Camisón S, et al. Re-Emergence of a West Nile virus (WNV) variant in South Spain with rapid spread capacity. Viruses. 2023; 15(12):2372. doi: 10.3390/v15122372.

[19] Barzon L, Pacenti M, Montarsi F, et al. Rapid spread of a new West Nile virus lineage 1 associated with increased risk of neuroinvasive disease during a large outbreak in northern Italy, 2022: One Health analysis. J Travel Med. 2022;taac125. doi: 10.1093/jtm/taac125.

[20] Richter J, Tryfonos C, Tourvas A, et al. Complete genome sequence of West Nile virus (WNV) from the first human case of neuroinvasive WNV infection in Cyprus. Genome Announc. 2017;5(43):e01110-17. doi: 10.1128/genomeA.01110-17.

[21] Selim A, Radwan A, Arnaout F. Seroprevalence and molecular characterization of West Nile virus in Egypt. Comp Immunol Microbiol Infect Dis. 2020;71:101473. doi: 10.1016/j.cimid.2020.101473.

[22] Amdouni J, Monaco F, Portanti O, et al. Detection of enzootic circulation of a new strain of West Nile virus lineage 1 in sentinel chickens in the North of Tunisia. Acta Trop. 2020;202:105223. doi: 10.1016/j.actatropica.2019.105223.

[23] Mencattelli G, Ndione MHD, Silverj A, et al. Spatial and temporal dynamics of West Nile virus between Africa and Europe. Nat Commun. 2023;14(1):6440. doi: 10.1038/s41467-023-42185-7.

[24] Akıner MM, Öztürk M, Başer AB, et al. Arboviral screening of invasive *Aedes* species in northeastern Turkey: West Nile virus circulation and detection of insect-only viruses. PLoS Negl Trop Dis. 2019;13(5):e0007334. doi: 10.1371/journal.pntd.0007334.

[25] Schvartz G, Farnoushi Y, Berkowitz A, et al. Molecular characterization of the re-emerging West Nile virus in avian species and equids in Israel, 2018, and pathological description of the disease. Parasit Vectors. 2020;13(1):528. doi: 10.1186/s13071-020-04399-2.

[26] Jiménez-Clavero MA, Sotelo E, Fernández-Pinero J, et al. West Nile virus in golden eagles, Spain, 2007. Emerg Infect Dis. 2008;14(9):1489–1491. doi: 10.3201/eid1409.080190.

[27] Sotelo E, Fernández-Pinero J, Jiménez-Clavero MÁ. La fiebre/encefalitis por virus West Nile: reemergencia en Europa y situación en España. Enferm Infecc Microbiol Clin. 2012;30(2):75–83. doi: 10.1016/j.eimc.2011.09.002.

[28] Figuerola J, Soriguer R, Rojo G, et al. Seroconversion in wild birds and local circulation of West Nile virus, Spain. Emerg Infect Dis. 2007;13(12):1915–1917. doi: 10.3201/eid1312.070343.

[29] Bravo-Barriga D, Aguilera-Sepúlveda P, Guerrero-Carvajal F, et al. West Nile and Usutu virus infections in wild birds admitted to rehabilitation centres in Extremadura, Western Spain, 2017-2019. Vet Microbiol. 2021;255:109020. doi: 10.1016/j.vetmic.2021.109020.

[30] Vázquez A, Ruiz S, Herrero L, et al. Short report : West Nile and Usutu viruses in mosquitoes in Spain, 2008 – 2009. Am J Trop Med Hyg. 2011;85(1):178–181. doi: 10.4269/ajtmh.2011.11-0042.

[31] Aguilera-Sepúlveda P, Gómez-Martín B, Agüero M, et al. A new cluster of West Nile virus lineage 1 isolated from a northern goshawk in Spain. Transbound Emerg Dis. 2022;69(5):3121–3127. doi: 10.1111/tbed.14399.

[32] García-Bocanegra I, Belkhiria J, Napp S, et al. Epidemiology and spatio-temporal analysis of West Nile virus in horses in Spain between 2010 and 2016. Transbound Emerg Dis. 2018;65(2):567–577. doi: 10.1111/tbed.12742.

[33] Guerrero-Carvajal F, Bravo-Barriga D, Martín-Cuervo M, et al. Serological evidence of co-circulation of West Nile and Usutu viruses in equids from Western Spain. Transbound Emerg Dis. 2020;68(3):1432–1444. doi: 10.1111/tbed.13810.

[34] Busquets N, Laranjo-González M, Soler M, et al. Detection of West Nile virus lineage 2 in North-Eastern Spain (Catalonia). Transbound Emerg Dis. 2019;66(2):617–621. doi: 10.1111/tbed.13086.

[35] Ministerio de Agricultura Pesca y Alimentación M. 2021. Actualización de la situación epidemiológica de la fiebre del Nilo occidental (West Nile fever).

[36] Aguilera-Sepúlveda P, Napp S, Llorente F, et al. West Nile virus lineage 2 spreads westwards in Europe and overwinters in North-Eastern Spain (2017–2020). Viruses. 2022;14(3):569. doi: 10.3390/v14030569.

[37] Rizzoli A, Jiménez-Clavero MA, Barzon L, et al. The challenge of West Nile virus in Europe : knowledge gaps and research priorities. Eurosurveillance. 2015;20(20):21135. doi: 10.2807/1560-7917.ES2015.20.20.21135.

[38] Aguilera-Sepúlveda P, Llorente F, Rosenstierne MW, et al. Detection of a new avian bornavirus in barn owl (*Tyto alba*) by pan-viral microarray. Vet Microbiol. 2024;289:109959. doi: 10.1016/j.vetmic.2023.109959.

[39] Jiménez-Clavero MA, Agüero M, Rojo G, et al. A new fluorogenic real-Time RT-PCR assay for detection of lineage 1 and lineage 2 West Nile viruses. J Vet Diagn Invest. 2006;18(5):459–462. doi: 10.1177/104063870601800505.

[40] Del Amo J, Sotelo E, Fernández-Pinero J, et al. A novel quantitative multiplex real-time RT-PCR for the simultaneous detection and differentiation of West Nile virus lineages 1 and 2, and of Usutu virus. J Virol Methods. 2013;189(2):321–327. doi: 10.1016/j.jviromet.2013.02.019.

[41] Sotelo E, Fernandez-Pinero J, Llorente F, et al. Characterization of West Nile virus isolates from Spain: new insights into the distinct West Nile virus eco-epidemiology in the Western Mediterranean. Virology. 2009;395(2):289–297. doi: 10.1016/j.virol.2009.09.013.

[42] Hadfield J, Megill C, Bell SM, et al. NextStrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34(23):4121–4123. doi: 10.1093/bioinformatics/bty407.

[43] Borges V, Pinheiro M, Pechirra P, et al. INSaFLU: an automated open web-based bioinformatics suite "from-reads" for influenza whole-genome-sequencing-based surveillance. Genome Med. 2018;10(1):46. doi: 10.1186/s13073-018-0555-0.

[44] Barzon L, Montarsi F, Quaranta E, et al. Early start of seasonal transmission and co-circulation of West Nile virus lineage 2 and a newly introduced lineage 1 strain, Northern Italy, June 2022. Eurosurveillance. 2022;27(29):2200548. doi: 10.2807/1560-7917.ES.2022.27.29.2200548.

[45] Anderson JF, Vossbrinck CR, Andreadis TG, et al. A phylogenetic approach to following West Nile virus in Connecticut. Proc Natl Acad Sci U S A. 2001;98(23):12885–12889. doi: 10.1073/pnas.241472398.

[46] Bondre VP, Jadi RS, Mishra AC, et al. West Nile virus isolates from India: evidence for a distinct genetic lineage. J Gen Virol. 2007;88(Pt 3):875–884. doi: 10.1099/vir.0.82403-0.

[47] Prow NA, Edmonds JH, Williams DT, et al. Virulence and evolution of West Nile virus, Australia, 1960-2012. Emerg Infect Dis. 2016;22(8):1353–1362. doi: 10.3201/eid2208.151719.

[48] Lustig Y, Hindiyeh M, Orshan L, et al. Mosquito surveillance for 15 years reveals high genetic diversity among West Nile viruses in Israel. J Infect Dis. 2016;213(7):1107–1114. doi: 10.1093/infdis/jiv556.

[49] May FJ, Davis CT, Tesh RB, et al. Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. J Virol. 2011;85(6):2964–2974. doi: 10.1128/JVI.01963-10.

[50] Agüero M, Fernández-Pinero J, Buitrago D, et al. Bagaza virus in partridges and pheasants, Spain, 2010. Emerg Infect Dis. 2011;17:1498–1501. doi: 10.3201/eid1708.110077.

[51] M'ghirbi Y, Mousson L, Moutailler S, et al. West Nile, Sindbis and Usutu viruses: evidence of circulation in mosquitoes and horses in Tunisia. Pathogens. 2023;12(3):360. doi: 10.3390/pathogens12030360.

[52] García-Carrasco JM, Muñoz AR, Olivero J, et al. Gone (and spread) with the birds: can chorotype analysis highlight the spread of West Nile virus within the Afro-Palaearctic flyway? One Health. 2023;17:100585. doi: 10.1016/j.onehlt.2023.100585.

[53] Figuerola J, Jiménez-Clavero MÁ, Ruíz-López MJ, et al. A One Health view of the West Nile virus outbreak in Andalusia (Spain) in 2020. Emerg Microbes Infect. 2022;11:2570–2578. doi: 10.1080/22221751.2022.2134055

[54] Ferraguti M, La Puente JMD, Soriguer R, et al. West Nile virus-neutralizing antibodies in wild birds from Southern Spain. Epidemiol Infect. 2016;144(9):1907–1911. doi: 10.1017/S0950268816000133.

[55] Abad-Cobo A, Llorente F, Barbero M del C, et al. Serosurvey reveals exposure to West Nile virus in asymptomatic horse populations in Central Spain prior to recent disease foci. Transbound Emerg Dis. 2017;64(5):1387–1392. doi: 10.1111/tbed.12510.

[56] Magallanes S, Llorente F, Ruiz-López MJ, et al. Long-term serological surveillance for West Nile and Usutu virus in horses in South-west Spain. One Health. 2023;17:100578. doi: 10.1016/j.onehlt.2023.100578.

[57] Young JJ, Haussig JM, Aberle SW, et al. Epidemiology of human West Nile virus infections in the european union and european union enlargement countries, 2010 to 2018. Eurosurveillance. 2021;26(19):2001095. doi: 10.2807/1560-7917.ES.2021.26.19.2001095.