

Untargeted metagenomic sequencing identifies Toscana virus in patients with idiopathic meningitis, southern Spain, 2015 to 2019

Fabiana Gámbaro¹, Ana Belén Pérez^{2,3,4}, Matthieu Prot¹, Eduardo Agüera^{2,3,5}, Artem Baidaliuk¹, María Paz Sánchez-Seco^{4,6}, Luis Martínez-Martínez^{2,3,4,5}, Ana Vázquez^{6,7}, María Dolores Fernandez-García^{3,6,7,8}, Etienne Simon-Loriere^{1,8}

1. G5 Evolutionary Genomics of RNA Viruses, Institut Pasteur, Université Paris Cité, Paris, France
2. Hospital Universitario Reina Sofía, Córdoba, Spain
3. Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Córdoba, Spain
4. CIBER de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain
5. Universidad de Córdoba, Córdoba, Spain
6. National Centre for Microbiology, Instituto de Salud Carlos III, Madrid, Spain
7. CIBER de Epidemiología y Salud Pública (CIBERESP), Instituto de Salud Carlos III (ISCIII), Madrid, Spain
8. These authors contributed equally to this work and share last authorship and correspondence

Correspondence: Etienne Simon-Loriere (etienne.simon-loriere@pasteur.fr), María Dolores Fernandez-García (mdfernandez@isciii.es)

Citation style for this article:

Gámbaro Fabiana, Pérez Ana Belén, Prot Matthieu, Agüera Eduardo, Baidaliuk Artem, Sánchez-Seco María Paz, Martínez-Martínez Luis, Vázquez Ana, Fernandez-García María Dolores, Simon-Loriere Etienne. Untargeted metagenomic sequencing identifies Toscana virus in patients with idiopathic meningitis, southern Spain, 2015 to 2019. *Euro Surveill.* 2023;28(45):pii=2200913. <https://doi.org/10.2807/1560-7917.ES.2023.28.45.2200913>

Article submitted on 21 Nov 2022 / accepted on 15 Apr 2023 / published on 09 Nov 2023

Background: Various pathogens, including bacteria, fungi, parasites, and viruses can lead to meningitis. Among viruses causing meningitis, Toscana virus (TOSV), a phlebovirus, is transmitted through sandfly bites. TOSV infection may be suspected if patients with enterovirus- and herpesvirus-negative aseptic (non-bacterial) meningitis recall recent insect bites. Other epidemiological factors (season, rural area) may be considered. The broad range of possible meningitis aetiologies poses considerable diagnosis challenges. Untargeted metagenomic next-generation sequencing (mNGS) can potentially identify pathogens, which are not considered or detected in routine diagnostic panels. **Aim:** In this retrospective, single-centre observational study, we investigated mNGS usefulness to understand the cause of meningitis when conventional approaches fail. **Methods:** Cerebrospinal fluid (CSF) samples from patients hospitalised in southern Spain in 2015–2019 with aseptic meningitis and no aetiology found by conventional testing, were subjected to mNGS. Patients' demographic characteristics had been recorded and physicians had asked them about recent insect bites. Obtained viral genome sequences were phylogenetically analysed. **Results:** Among 23 idiopathic cases, TOSV was identified in eight (all male; median age: 39 years, range: 15–78 years). Five cases lived in an urban setting, three occurred in autumn and only one recalled insect bites. Phylogenetic analysis of TOSV segment sequences supported one intra-genotype reassortment event. **Conclusions:** Our study highlights the usefulness of mNGS for identifying viral pathogens directly in CSF. In southern Spain, TOSV should be considered regardless of recalling of insect

bites or other epidemiological criteria. Detection of a disease-associated reassortant TOSV emphasises the importance of monitoring the spread and evolution of phleboviruses in Mediterranean countries.

Introduction

Neurological infections are a considerable cause of morbidity and mortality worldwide, with 10.6 million viral or bacterial meningitis cases reported for 2017 alone [1]. Various pathogens, including viruses, bacteria, fungi, and parasites, can lead to meningitis and encephalitis [2]. Since the mid-2000s, molecular diagnostic tests have improved rapidly and multiplex PCR approaches now allow with high sensitivity and specificity to detect a wide variety of pathogens that can occur in cerebrospinal fluid (CSF). Despite the high global prevalence of central nervous system (CNS) infections and improved differential diagnostic methods, an investigation in the United Kingdom in 2011–2014 found that the reason for meningitis remained unknown in a large proportion (42%) of cases [3]. Moreover, a study in Houston, Texas, between 2005 and 2010, estimated that for 81% of patients with meningitis the cause was undetermined [4].

Diagnosing meningitis can indeed entail challenges. One of these is the broad range of potential aetiological agents, which cannot be singled out or deduced from the clinical presentation of patients alone. For finding the cause, several selective tests may be required, potentially raising considerations on their availability and costs. Another challenge is that pathogen-specific diagnostic assays (molecular or antigenic) may fail to

KEY PUBLIC HEALTH MESSAGE

What did you want to address in this study?

In a high proportion of meningitis cases, the causative pathogen remains unknown despite extensive diagnostic testing. Unlike conventional testing, which targets known or suspected pathogens, metagenomic next-generation sequencing (mNGS) can potentially capture genomic data from any pathogen in a patient sample. We investigated by mNGS samples of patients hospitalised in southern Spain in 2015–2019 with undetermined cause of aseptic meningitis.

What have we learnt from this study?

Toscana virus (TOSV), a virus transmitted by sandflies, was identified in eight of 23 samples from patients with meningitis of unknown origin. Of the eight TOSV-positive patients, three had occurred in autumn and five lived in urban areas. Only one recalled recent insect bites, an event that may raise suspicion of TOSV infection. By analysing the genetic information of TOSV strains affecting the patients, insights into how TOSV evolves and spreads were gained.

What are the implications of your findings for public health?

The study reveals the utility of mNGS and the need to consider including TOSV in the differential diagnosis of aseptic meningitis in southern Spain regardless of the recall of insect bites or environmental conditions such as summer-season or rural residence. Panels for laboratory routine diagnostic of neurological infections in this region's hospitals might also target TOSV, since timely diagnosis is crucial for public health risk management strategies.

detect a given pathogen because of genetic divergence. Furthermore, diagnostic tests may not be designed to detect pathogens, which are not suspected or commonly known to cause meningitis, such as emerging or re-emerging ones that have newly acquired this ability. An example of a virus for which the clinical picture resulting from infection appeared to change is Zika virus, which despite being known since 1947, was only associated with a risk of neurological afflictions in 2014–2016, upon its extensive spread in the Pacific and South America [5]. Last, as the amount of pathogen in CSF may vary from the moment of infection onwards, the timing of a patient seeking medical care or the timing of sampling may influence the chance of reliable detection (due to clearance of the pathogen or its amounts being below the test detection threshold).

Metagenomic next-generation sequencing (mNGS) is a promising approach for diagnosing infectious diseases since it can overcome some of the challenges faced by conventional diagnostic techniques such as PCR. For instance, mNGS has the potential to capture genomic information of any pathogen present in a biological sample, thereby saving the need for a broad panel of pathogen-specific tests. Due to its unbiased nature, mNGS can also identify novel (to a specific region or population), or highly divergent pathogens from those previously known, including new recombinant forms, or forms presenting atypical neurological manifestations [6]. In addition, mNGS provides pathogen genomic data, a resource increasingly leveraged in Public Health.

In this study, we used mNGS to retrospectively analyse leftover CSF samples from patients with aseptic meningitis in southern Spain (Andalusia region) for whom no cause had been found after routine clinical testing.

Methods

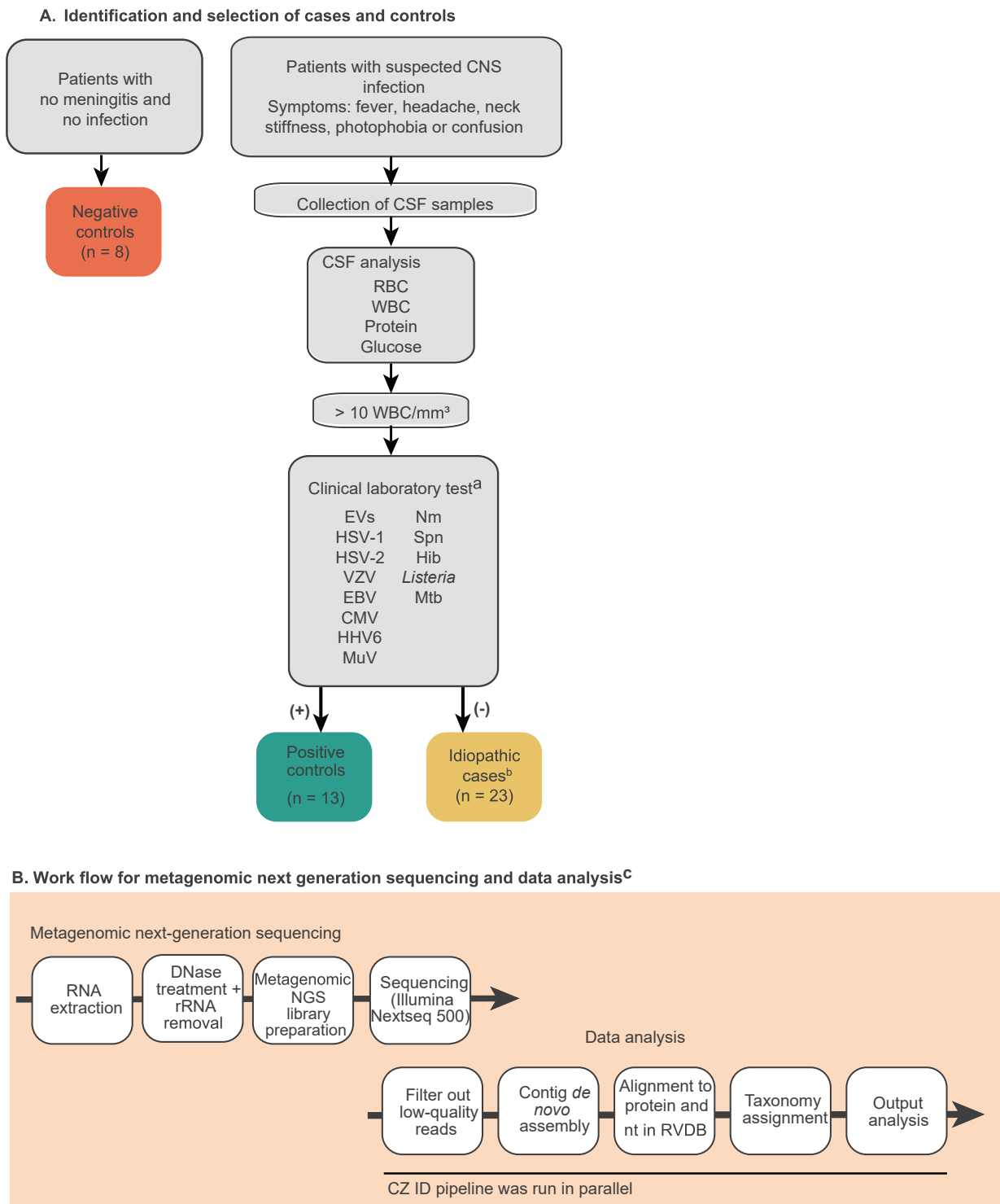
Sample selection and laboratory methods

Clinical CSF specimens used in this study were collected for routine clinical care from patients with suspected viral CNS infection in the Department of Neurology in the University Hospital Reina Sofia (Cordoba, Spain) between 2015 and 2019.

The Department of Neurology is responsible for the treatment of patients aged 14 years and older. It is the reference unit for patients with suspected CNS infections in the Córdoba province, where the number of inhabitants during the study period ranged from 795,611 in 2015 to 782,979 in 2019 [7]. CSF samples were obtained at patient admission according to standard procedures. For these CSF samples, protein and glucose levels were quantified, and white and red blood cells (WBC and RBC, respectively) were counted. Testing for a range of pathogenic agents according to the clinical suspicion of the attending physician was performed in the Unit of Clinical Microbiology. Included tested pathogens were enteroviruses (EVs), herpesviruses (cytomegalovirus (CMV), Epstein–Barr virus (EBV), herpes simplex viruses type 1 and type 2, human herpes virus type 6 and varicella zoster virus (VZV)), as well as bacterial pathogens (*Haemophilus influenzae type b*, *Listeria*, *Mycobacterium tuberculosis*,

FIGURE 1

(A) Criteria to classify patients as negative or positive controls, or as idiopathic cases, and (B) workflow for metagenomic sequencing and analysis, Andalusia, Spain, 2015–2019 (n = 44 samples analysed)



CNS: central nervous system; CSF: cerebrospinal fluid; CZ-ID: Chan Zuckerberg ID; NGS: next generation sequencing; RBC: red blood cells; rRNA: ribosomal ribonucleic acid; RVDB: reference virus database; RT-qPCR: reverse transcription-quantitative PCR; TOSV: Toscana virus; WBC: white blood cells.

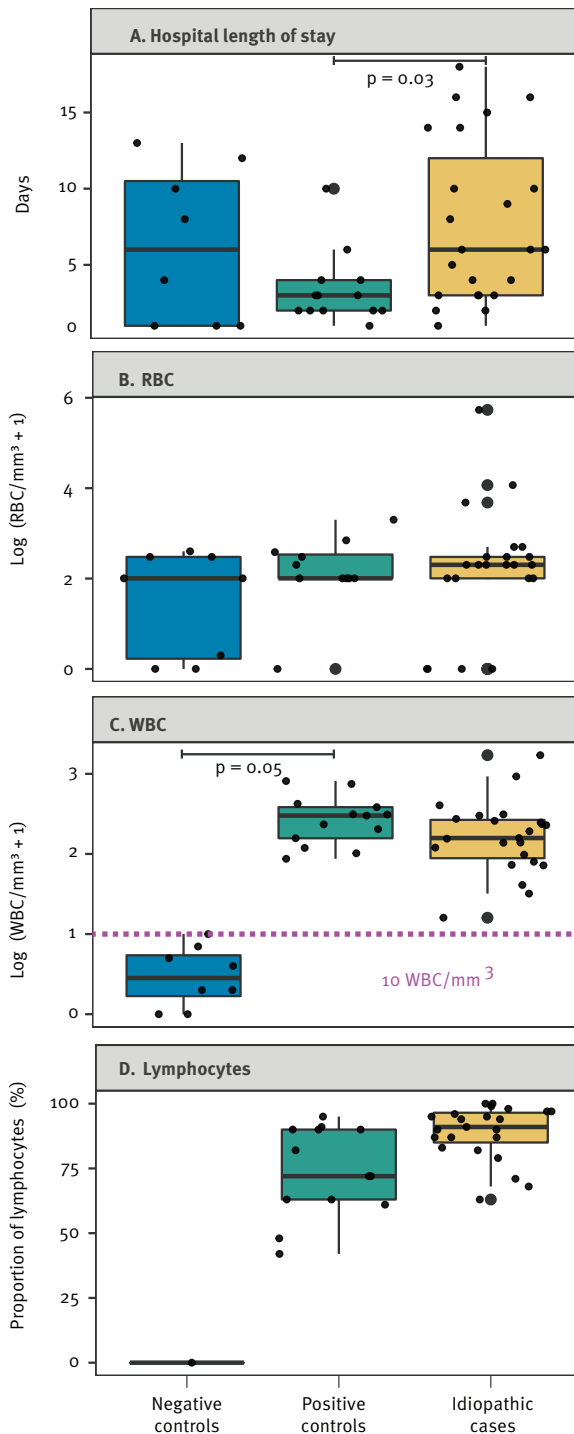
^a CMV: cytomegalovirus; EBV: Epstein–Barr virus; EVs: enteroviruses; HH6: human herpesvirus type 6; Hib: *Haemophilus influenzae* type b; HSV-1: herpes simplex virus 1; HSV-2: herpes simplex virus 2; Mtb: *Mycobacterium tuberculosis*; MuV: mumps virus; Nm: *Neisseria meningitidis*; Spn: *Streptococcus pneumoniae*; VZV: varicella zoster virus.

^b Among the idiopathic cases, one sample (LCR_1152) had been prior tested for TOSV by RT-qPCR, but this had yielded a negative result. This sample was then shown by metagenomic sequencing to be positive for TOSV.

^c Data were analysed using custom scripts and tools on the CZID platform (<https://czid.org/>).

FIGURE 2

Respective comparison of (A) hospital length of stay and (B–C) blood cell abundance in cerebrospinal fluid samples from negative controls, positive controls and idiopathic aseptic meningitis cases, Andalusia, Spain, 2015–2019 (n = 44 samples analysed)



ANOVA: analysis of variance; CSF: cerebrospinal fluid; RBC: red blood cells; RTqPCR: reverse-transcription quantitative PCR; TOSV: Toscana virus; WBC: white blood cells.

Comparison of characteristics of clinical samples among the cases (yellow) and two types of controls (positive (green) and negative (blue)) are shown in the Figure. Significantly different samples tested with one-way ANOVA with Tukey's post-hoc test are indicated with the pvalue, p. The lymphocyte count was only available for one of the eight negative controls; no statistical test was performed. Norm RBC in CSF adults: 0 cells/mm³; norm WBC in CSF adults: <5 cells/mm³ of which lymphocytes 60–70%.

Neisseria meningitidis and *Streptococcus pneumoniae*). For some patients, serum samples were also obtained. In this case, these were screened for *Borrelia*, *Brucella*, *Coxiella*, *Cryptococcus*, *Rickettsia*, hepatitis viruses, HIV, *Mycoplasma pneumoniae*, parvovirus and *Treponema pallidum*. For patients with clinical suspicion for mumps virus (MuV), Toscana virus (TOSV) or West Nile virus (WNV), CSF was also tested for these viruses in the Virus Reference Laboratory of Andalusia at the University Hospital Virgen de las Nieves (Granada, Spain).

Of note, in Spain as in the EU, infections by TOSV are not notifiable and there is no case definition for TOSV. Testing is considered in patients with enterovirus- and herpesvirus-negative aseptic meningitis recalling recent insect bites. Other epidemiological factors such as summer season or rural residence may also be considered.

Sex, age, location of residence (rural or urban), month of sample collection, reporting of insect bites, length of hospital stay and clinical parameters were recorded. Information on routine laboratory testing performed in all selected patients (in CSF and serum samples) is provided in Supplementary Table S1. The laboratory methods used are summarised in Supplementary Table S2.

Residual CSFs were stored at -45°C and those with sufficient volume (1 mL) were selected. Samples included in this study were: (i) positive controls, which were randomly chosen CSF samples from patients with meningitis for whom the causative pathogen was previously identified by quantitative (q)PCR test; (ii) negative controls, which were randomly chosen CSF samples from patients with no meningitis and no infection but in whom another diagnosis was made (epilepsy or cognitive impairment); and (iii) idiopathic samples which were randomly chosen CSF samples from patients diagnosed with aseptic meningitis with unknown aetiological agent after conventional routine laboratory testing in the hospital (Figure 1). Meningitis-suspected cases were identified as described previously [8].

Metagenomic next generation sequencing

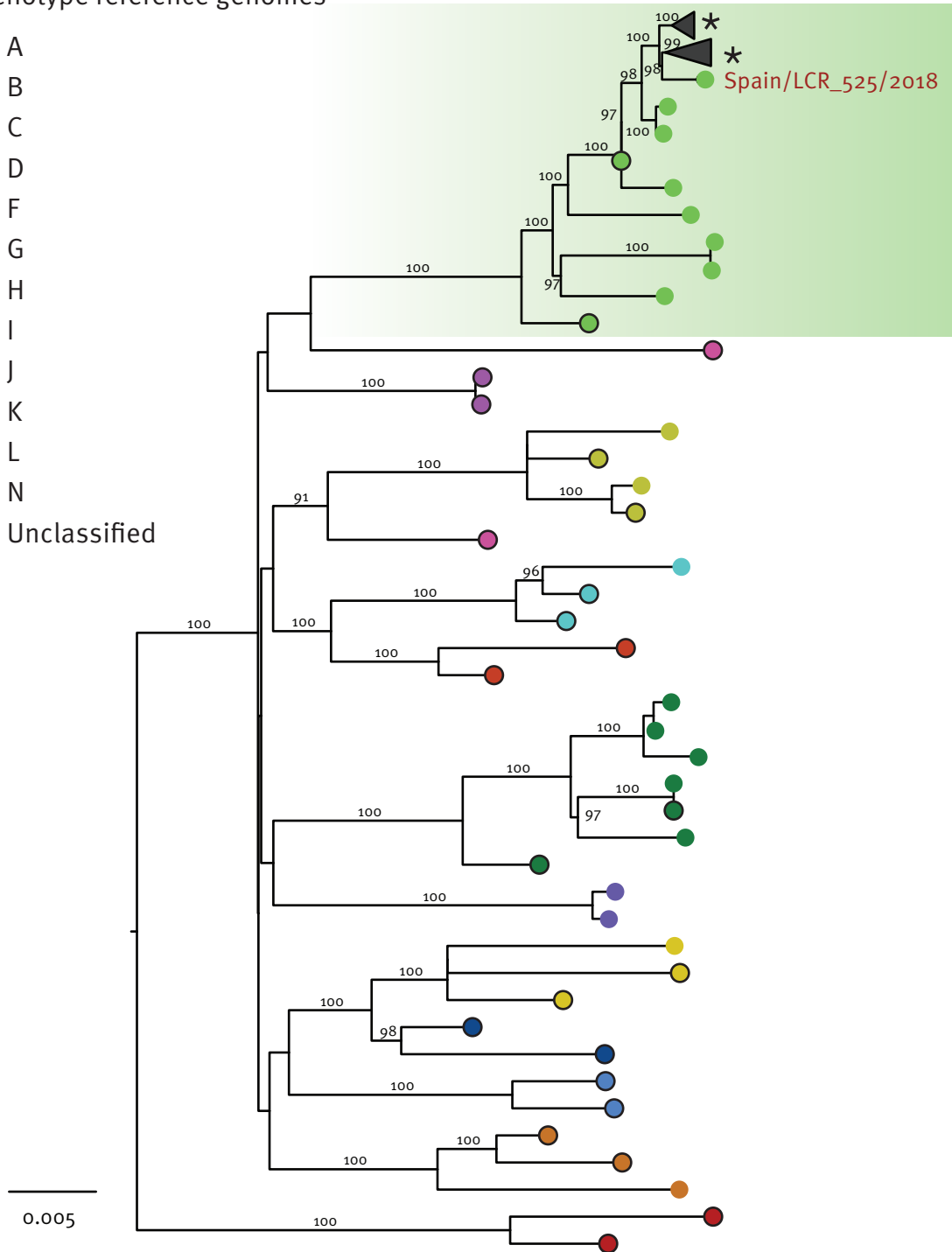
The sample processing for mNGS and subsequent data analysis are described in the Supplementary Material. Briefly, the samples for mNGS underwent a ribosomal RNA depletion step prior to randomly primed reverse transcription and library preparation (Figure 1). Samples were sequenced on an Illumina NextSeq500, with a median depth of 11.7 million reads/sample (interquartile range (IQR): 9.9–16.5 million). After filtering out human and low-quality reads, the resulting FASTQ files were processed using an in-house pipeline (Supplementary Material), as well as on the CZID platform (<https://czid.org/>), a cloud-based open-source tool for pathogen detection from metagenomic data [9].

FIGURE 3

Maximum likelihood phylogenetic analysis of a full-length mumps virus genome sequence with other mumps virus sequences including those of the WHO reference strains for the different genotypes, Andalusia, Spain, 2018 (n = 211 sequences)

Genotype reference genomes

- A
- B
- C
- D
- F
- G
- H
- I
- J
- K
- L
- N
- Unclassified



MuV: mumps virus; WHO: World Health Organization.

The tree leaves have circles filled with colour according to the MuV genotype. A black circle around the coloured tip denotes a WHO reference strain. The light green box highlights the MuV genotype G, the genotype to which the genome generated in this study (MuV/Spain/LCR_525/2018) belongs. The collapsed clusters with the asterisks correspond to sequences from the United States (2015–2017) and Canada (2017). Ultrafast bootstrap values above 80% are indicated above the branches leading to corresponding nodes. The scale bar indicates nucleotide substitutions per site. The tree is midpoint-rooted for visualisation. A complete expanded tree is shown in Supplementary Figure S2.

Phylogenetic analysis

To build the TOSV dataset for phylogenetic reconstruction, the set of sequences generated during this study were combined with a set of partial and full-length sequences available in Virus Pathogen Resource (ViPR) [10] and GenBank [11] up to June 2022; the latter sequence set included representatives of the different TOSV genotypes. Partial sequences of less than 190 bp in length were excluded. We put together a total of 112, 88, and 80 sequences for the small (S), medium (M), and large (L) segments of the TOSV genome, respectively. The resulting dataset was aligned using Multiple Alignment using Fast Fourier Transform (MAFFT) v7.467 [12] for each segment separately and alignments were visually inspected in Geneious Prime 2020.2. We inferred maximum-likelihood (ML) phylogenies using IQ-TREE v2.0.6 [13]. Tree reconstruction was performed using the default settings and the best-fitting model provided by ModelFinder [14], followed by Ultrafast Bootstrap Approximation with 1,000 replicates [15] as implemented in IQ-TREE 2.

We similarly retrieved from ViPR all full-length MuV genome sequences available from 2008 up to October 2021 [10] and used these as well as the different World Health Organization (WHO) MuV genotype reference strain sequences to analyse the complete genome generated by mNGS in this study. This dataset containing a total of 211 sequences was employed to infer the ML phylogeny.

Recombination analysis

Recombination analyses were carried out on concatenated genomes using the three segments of TOSV. A total of 29 genomes, including the seven generated in this study, were included in the analysis as described in Supplementary Table S3. We studied recombination using a combination of six methods (RDP, GENECONV, MaxChi, Bootscan, SisScan and 3SEQ) implemented in Recombination Detection Program (RDP)₄ [16], and we considered recombination signals detected by more than three methods for breakpoint identification. Similarity plot and bootscanning analysis were performed using SimPlot v3.5.1 [17], with a 400-nt window moving in 40-nt steps and using a Kimura 2-parameter method with a transitions-transversions ratio of two with 1,000 resampling.

Results

Patient characteristics

The sample cohort consisted of positive controls (enteroviral meningitis, $n=12$ and mumps meningitis, $n=1$); negative controls ($n=8$); and idiopathic meningitis samples ($n=23$). All patients originated from Cordoba province and the majority (26/44) lived in large urban centres, while the remaining patients lived in villages (populations between 900 and 9,000 inhabitants). The age of the patients ($n=44$, 26 male, 18 female) ranged from 14 to 95 years (mean: 41 years). Among the 23 idiopathic cases (17/23 male), 16 resided in urban areas.

The age of the 23 cases ranged from 14 to 83 years (mean: 42 years), nine reported no insect bites when asked, and two explicitly mentioned insect bites, and for the 12 remainder this information was not available on the clinical record.

The WBC was higher in idiopathic (median: 157.0 cells/mm³; IQR: 88.0–265.5 cells/mm³; norm adults: <5 cells/mm³ of which lymphocytes 60–70%) and positive control (median: 300.0 cells/mm³; IQR: 156.0–383.0 cells/mm³) cases than in negative control cases (median: 2.0 cells/mm³; IQR: 0.8–5 cells/mm³). However, this difference was not statistically different (Figure 2). Lymphocyte predominance was observed in idiopathic (median: 91.0%; IQR: 85.0–96.5%) and positive control (median: 73.8%; IQR: 63.0–90.0%) CSF samples. Epidemiological and clinical characteristics recorded with the CSF samples and routine laboratory testing performed are provided in Supplementary Tables S1 and S2.

Metagenomic sequencing

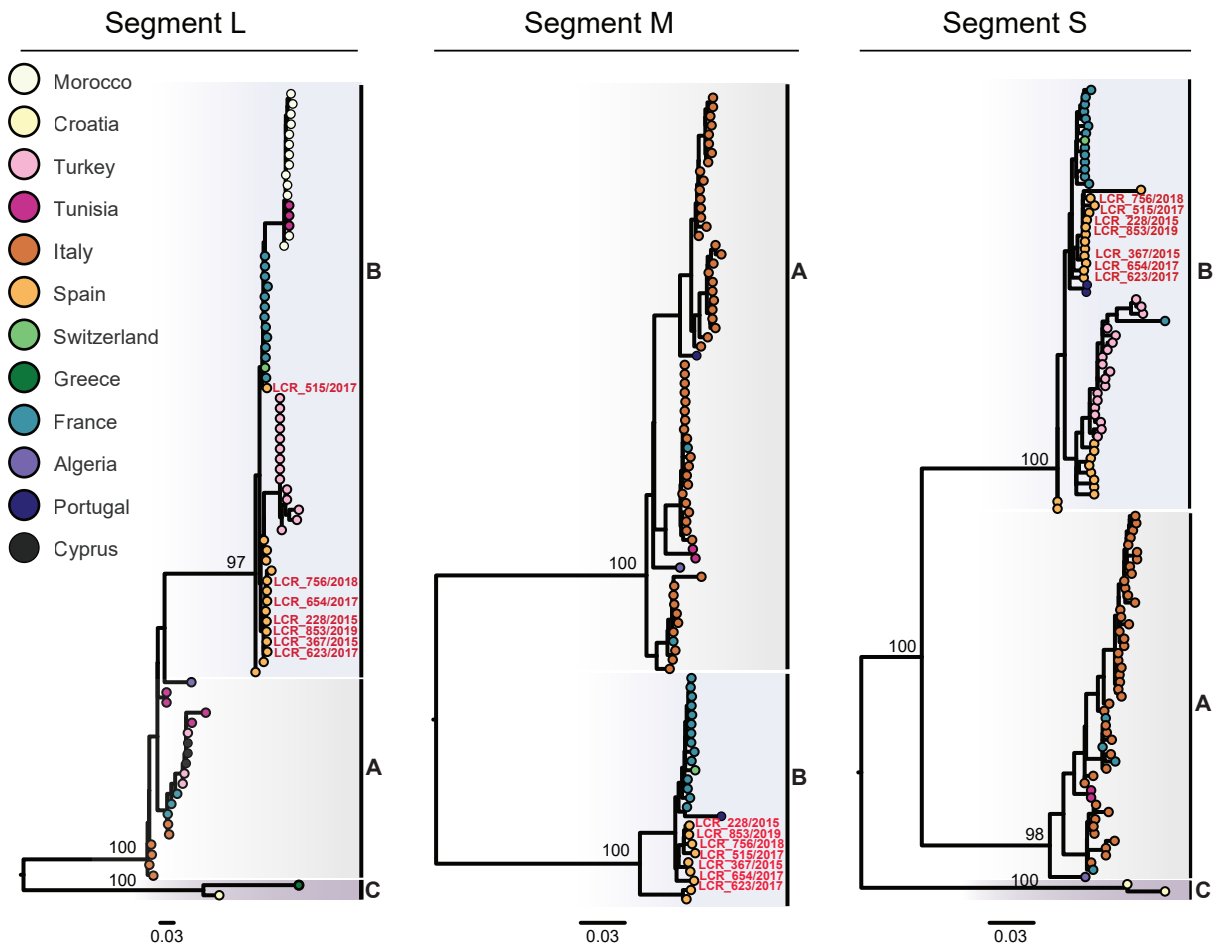
In positive-control samples, we identified complete EV ($n=12$) [8] and MuV ($n=1$) genomes by mNGS, in accordance with prior diagnosis. In the negative control samples, neither approach revealed potential pathogens in the sequencing data. Surprisingly, among the 23 idiopathic cases with unknown aetiology, we detected TOSV genomic material in eight cases. TOSV had only been prior suspected for one (sample LCR_1152) of these TOSV-positive cases, however, the TOSV reverse-transcription (RT)-qPCR ordered during hospitalisation for this sample was negative. To rule out possible cross-contamination with another TOSV-positive sample, this sample was processed independently twice from the RNA extraction step. We reconstructed partial to almost complete segments (between 40–97% of the genome covered) for seven TOSV-positive samples, and only a partial fragment of the S and L segments for one sample as shown in Supplementary Table S4 and Figure S1. No pathogen could be identified from the rest of the idiopathic samples ($n=15$).

Characteristics of the mumps virus positive sample and phylogenetic analysis

The positive control CSF sample found with MuV (LCR_525) was used to verify the capacity of mNGS, in our setting, to obtain exploitable genomic data. The sample had been obtained from a woman admitted to the hospital with parotitis, severe headache, and photophobia. The CSF specimen contained 747 WBC/mm³ and had tested positive for MuV by an in-house RT-qPCR [18]. mNGS corroborated this diagnosis and allowed to reconstruct a near-complete MuV genome sequence. This genome sequence was subsequently phylogenetically analysed using a dataset of 210 publicly available full-length MuV genome sequences, including those of the WHO reference genomes for the different MuV genotypes (A–N). The MuV/Spain/LCR_525/2018 genome derived from the LCR_525

FIGURE 4

Maximum likelihood phylogenetic analysis of segment L, segment M and segment S of Toscana virus sequences generated in this study, Andalusia, Spain, 2015–2019 (n=7 novel samples analysed)



The maximum likelihood phylogenetic trees were inferred using IQ-TREE 2. Genotype A sequences are highlighted in grey, B in blue, and C in purple. Novel genomes reported here are noted in red. The circles at the branch tips are coloured according to country. Branch support values for the main lineages are ultrafast bootstrap percentages, and the scale bars represent the number of nucleotide substitutions per site. Trees with full taxa names are provided in Supplementary Figures S4–6.

sample clustered within genotype G, together with sequences collected from the United States (US) during 2015–2017 and Canada in 2017. The two sequences immediately basal to MuV/Spain/LCR_525/2018 are genomes collected in the US in 2011 (GenBank accession number: KY969482) and the Netherlands in 2010 (GenBank accession number: MW261742) (Figure 3). This is also illustrated in Supplementary Figure S2.

Toscana virus amplicon-based sequencing

As mNGS only allowed to reconstruct partial genome sequences from some idiopathic cases, we built a PCR panel to improve coverage, as follows. Using the information obtained from mNGS, we designed a short PCR amplicon panel based on Quick et al. [19] targeting the known diversity of TOSV genotype B using Primal Scheme (<https://primalscheme.com/>). To test and optimise the panel, we used TOSV genomic RNA obtained from the European Virus Archive–Global (<https://www.european-virus-archive.com/>, Supplementary material). The optimised PCR

panel can be found at <https://github.com/Simon-LorriereLab/TOSV> and in Supplementary Table S5. Using the targeted approach, we were able to generate almost complete TOSV genome sequences for seven of eight samples as shown in Supplementary Table S6 and Figure S3.

Characteristics of patients testing positive for Toscana virus

All TOSV-infected patients were male. The age of patients ranged from 15 to 78 years (median: 39 years). The length of the hospital stay ranged from 2 to 16 days (median: 7 days). All these cases were detected between July and November in different years, compatible with the seasonal activity of sandflies in Mediterranean countries [20]. These patients were admitted to the hospital with severe headache and fever ($\geq 38^\circ\text{C}$), and their CSF contained high levels of WBC (median: 194.5 cells/mm³; IQR: 127.8–252.8 cells/mm³). Some of them also experienced vomiting, sensitivity to light and neck stiffness (Supplementary Table

S1). All TOSV-positive patients reported no insect bites to the physician, except for one patient (sample 654) who reported insect bites during a trip to Portugal. Finally, the majority (5/8) lived in Cordoba city, a large urban centre of over 300,000 inhabitants.

Phylogenetic analysis of Toscana virus

In combination with a set of publicly available partial and complete TOSV sequences corresponding to the three TOSV genotypes (A, B and C), we generated ML phylogenies for each genome segment separately. The estimated phylogenies placed the novel sequences within genotype B with high support for the three segments (node support of 100, 100 and 97% for segments S, M and L, respectively) (Figure 4). Genotype B comprises sequences collected from France (2004–2015), Morocco (2008–2011, partial genomes), Portugal (1983), Spain (1998–2004), Switzerland (2018), Tunisia (2012, partial genomes) and Turkey (2013–2018, partial genomes), as can be viewed in Supplementary Figures S4–6.

Phylogenies inferred using only complete or almost complete genomes showed similar topologies, where genotype B may be consistently subdivided into two lineages, B.1 and B.2 (Figure 5 and Supplementary Figures S4, S5 and S6). While lineage B.1 contained sequences collected from France, Morocco, Switzerland or Tunisia, lineage B.2 included sequences from Spain and Turkey suggesting that the circulation of the distinct viral lineages may be segregated between these two sets of countries, with some countries appearing to have different lineages despite their geographic proximity. In a phylogenetic tree based on complete genome information (Figure 5), two sequences collected from Spain in 1993–1999 (strain ESH_62100) and 2003–2004 (strain EsPhGR40) appeared basal to the two groups forming B.1 and B.2 lineages [21].

Interestingly, the phylogenetic placement of sample TOSV/Spain/LCR_515/2017 from one of the study cases did not seem consistent between segment L and the other segments analyses, suggesting intra-genotype reassortment. Indeed, the L segment of sample TOSV/Spain/LCR_515/2017 appears within the lineage B.1 cluster, whereas both segments M and S figure within the lineage B.2 cluster. To further explore this incongruence, we performed a recombination analysis using concatenated genomes of the TOSV sequences and a combination of models implemented in RDP4. The concatenated TOSV/Spain/LCR_515/2017 genome presented strong signal of recombination in five different methods including 3SEQ. Similarly, similarity plot and bootscanning analysis showed higher similarity of TOSV/Spain/LCR_515/2017 to B.2 lineage across the S and M segment region and higher similarity to B.1 sequences across the L segment, further supporting an intra-genotype reassortment event; this is illustrated in Supplementary Figure S7.

Discussion

In this study, we used mNGS on CSF samples from aseptic meningitis cases in southern Spain with unknown aetiology. To validate our methodology, we included positive and negative control samples. Our work highlights the value of mNGS as a potent diagnostic tool as we confirmed viruses that had been previously identified by conventional PCR in all positive controls (EV, $n=12$ and MuV, $n=1$ genomes, two RNA viruses well-known to cause meningitis). In addition, our approach allowed to reconstruct complete genomes from these positive controls, thus providing useful molecular epidemiology data such as the first whole-genome sequence of MuV in Spain or the identification of a novel Echovirus 13 recombinant form [8]. In this study we detected TOSV genetic material in eight of 23 idiopathic aseptic meningitis cases, a rate of identification of an aetiology similar to the one reported in the study from Saha et al. (40%) in Bangladesh using a similar dataset (idiopathic meningitis; $n=25$) and a similar RNA mNGS approach on CSF [22].

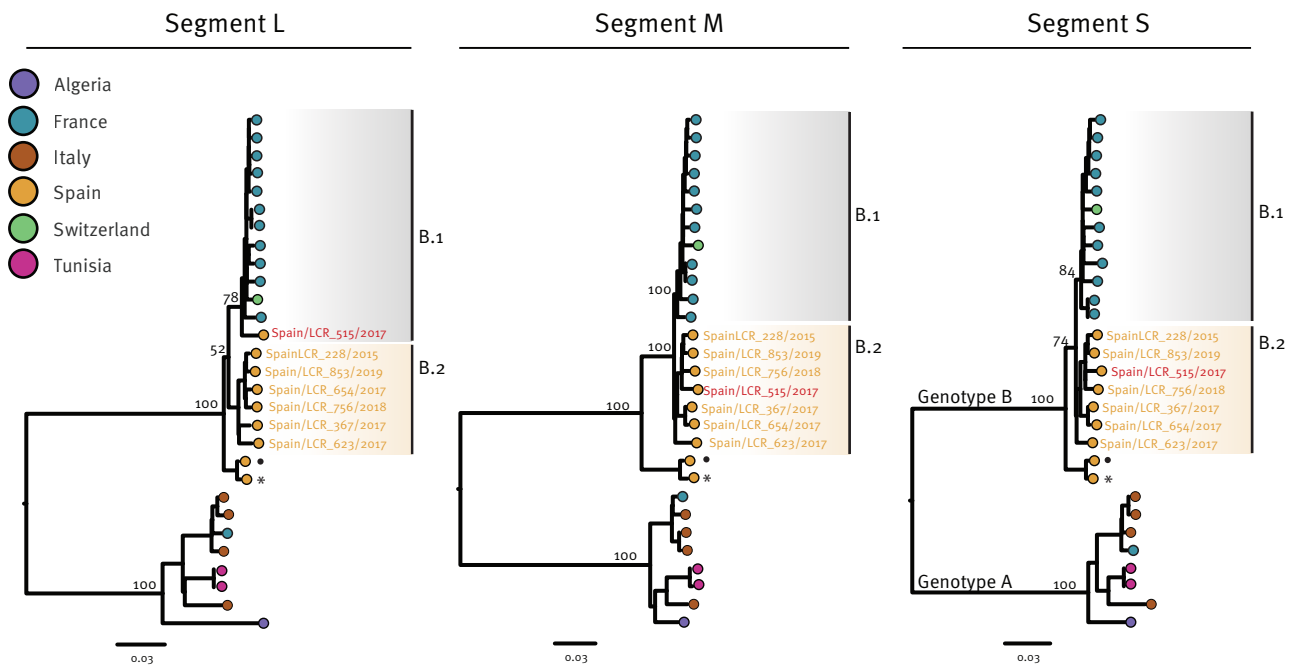
Interestingly, although based on smaller numbers of patients, the frequency of TOSV detection appeared notably higher than observed in previous retrospective CSF studies in the region of Madrid and southern Spain between 1998 and 2013, where 58 of 1,393 idiopathic aseptic meningitis cases (4.2%) were identified as TOSV cases by PCR [23]. Such results could be attributed to the enhanced viral detection capability of mNGS over conventional testing alone as reported in previous studies [24,25], or higher incidence of TOSV in the region during the study period.

As the mNGS failed to generate full-length TOSV genomes for a fraction of samples (4/8 samples, Supplementary Table S4), we developed and optimised a highly multiplexed PCR amplicon panel, using an approach known to help with challenging samples (low viral load or degraded RNA) [19]. Using this amplicon panel, we obtained almost complete sequences from all but one identified sample (Spain/LCR_1152/2019). This sample was the only one that had been prior tested for TOSV by RT-qPCR, with a negative result. Although none of the CSF samples contained high amounts of TOSV RNAs, to address the risk of cross-contamination, this sample was processed a second time by a different experimenter in different facilities starting from the RNA extraction and the result did not change. A possible explanation for the discrepancy between mNGS and RT-qPCR could be that TOSV RNAs were too degraded for RT-qPCR detection. Indeed, the RT-qPCR target region – among others – was not recovered by the sequencing of this sample. This result highlights the strength of mNGS to detect pathogens in clinical specimens even with degraded samples.

TOSV is a phlebovirus transmitted by sandflies of the genus *Phlebotomus* [26], and it has been detected in many countries surrounding the Mediterranean Sea. Although seroprevalence studies suggest that TOSV

FIGURE 5

Phylogenetic analysis of complete Toscana virus genome sequences, including those generated in this study, highlighting the incongruent position of the TOSV/Spain/LCR_515/2017 strain sequence, Andalusia, Spain, 2015–2019 TOSV (n = 29 sequences)



TOSV: Toscana virus.

All available complete TOSV genomes (n = 22) were combined with the genomes generated in this study (n = 7) to infer maximum likelihood trees using IQ-TREE 2. Proposed lineages B.1 and B.2 within genotype B are highlighted. Strains ESH_62100 and EsPhGR40, basal to B.1 and B.2 in this analysis are noted with an asterisk and a dot symbol, respectively. The numbers at the nodes indicate bootstrap support values in per cent. The scale bar represents nucleotide substitutions per site. Taxa names and metadata of each genome are provided in Supplementary Table S3. Phylogenies can be visualised interactively at <https://github.com/Simon-LorriereLab/TOSV>.

causes asymptomatic infections in most cases, some individuals experience severe disease with fever, intense headache, vomiting, or neurological diseases such as meningitis or encephalitis. Indeed, a recent study showed that TOSV is the most common cause of summer viral meningitis in central Italy, with meningitis cases by TOSV outnumbering those by EV [27]. In Spain, sporadic meningitis cases and series of TOSV infections have been documented in the Andalusia region, the Mediterranean coast and Madrid [23,28]. In Andalusia, where Cordoba study province is located, 107 TOSV positive cases have been diagnosed from 1988 to 2020 and the region is considered to be endemic for TOSV [29]. Despite its recognised clinical importance, TOSV remains a neglected virus [26]. In Spain, cases are often identified retrospectively in research studies indicating a lack of clinical suspicion for TOSV in patients with meningitis [23]. Only one patient in the current study cohort (LCR_1152) had been tested for TOSV by clinicians. This patient did not report recent insect bites, but had an outdoor occupation associated with a higher risk for TOSV infection [29]. The lack of suspicion for TOSV may lie in the rare report of insect bites (only one TOSV positive patient mentioned insect bites to the doctor when asked). The difficulty in remembering insect bites could be explained

by the long incubation period for TOSV neuroinvasive infection, estimated at 12 days (95% confidence interval (CI):10.2–14.4) [30]. It is also noteworthy that the majority (5/8) of TOSV positive cases identified here indicated living in large urban centres which differed from previous studies reporting most TOSV cases in rural areas in Spain [23]. This difference may reflect an evolution of the geographic distribution of sandflies, impacting on the epidemiology of associated human and animal pathogens such as TOSV [31]. Additional studies, including detailed patient information (e.g. employment, outdoor activities, out-of-town secondary residences) would help to evaluate the factors associated with higher risk of TOSV infection. Finally, some TOSV positive cases (3/8) corresponded to infections that occurred during autumn months (September–November). Although detection of TOSV in Spain outside of summertime has been sporadically reported [23], it is important to raise awareness among clinicians on this potential expansion of temporality. Overall, our study calls for considering the inclusion of TOSV in the differential diagnosis of patients presenting with enterovirus- and herpesvirus-negative aseptic meningitis regardless of history of insect bites mentioned by the patient, summer season or rural residence.

Another reason contributing to the neglected nature of this virus is the lack of laboratory testing for TOSV in some Mediterranean countries. In Andalusia, most hospital laboratories do not offer testing, and TOSV-suspected samples are sent to the Regional Reference Virology Laboratory in Granada, possibly delaying diagnosis. Including TOSV in the routine diagnosis of aseptic meningitis cases at the hospitals in southern Spain could result in shortened length of hospital stay (median 7 days in this study for TOSV-positive patients), reduced associated costs [3], and improved patient care. The recurrent detection of TOSV across the Mediterranean region [32] and the detection of TOSV and its vector in European countries outside this region, such as Germany (2006–2016, n=4) [33], calls for increased TOSV awareness and vigilance in Europe. Importantly, the timely identification of a vector-borne infection is crucial for public health risk management strategies such as local vector control programmes.

Analysis of the complete TOSV genomes provides evidence of intra-genotype reassortment involving the L segment. The mechanism of reassortment between different viruses within the *Phlebovirus* genus has been widely described [34], consisting in all cases of the replacement of the M segment [35,36]. This TOSV reassortment has multiple implications. Indeed, while the genomic data previously available for the B.1 and B.2 lineages suggest a strong geographical clustering within genotype B, this genomic exchange indicates that both lineages must co-circulate or have co-circulated in recent years, in Spain or the south of Europe. This also suggests that TOSV prevalence – and genetic diversity – in its reservoir or vector populations is underestimated. We hope that the targeted approach developed here, together with the ongoing increase in pathogen genomic sequencing efforts, will help enrich the databases (including completing partial genomes), allowing high-resolution phylodynamic analysis to better understand the epidemiology and evolution of this virus.

Our work has several limitations. The storage conditions of the samples used in this study were not optimal: the samples were collected between 2015 and 2019, stored at -45°C until processed in 2019, and subjected to multiple freeze-thaw cycles. This fact could have affected the quality of the RNA, contributing to the proportion of cases that remained unsolved. In addition, our metagenomic approach focuses on RNA so, in terms of using this method for DNA pathogens, while the mRNA of such pathogens can be potentially detected [37], if these replicate at a low level, they might go undetected.

In the Andalusia Region, aseptic meningitis is a mandatorily and urgently notifiable disease. From 2016 to 2017, a total of 567 cases of viral meningitis were reported in Andalusia of which 275 (48.5%) remained with an undiagnosed aetiology [29]. In conclusion, to reduce these numbers, our work calls for (i) testing

patients presenting with enterovirus- and herpesvirus-negative aseptic meningitis in Andalusia for TOSV regardless of their recalling of insect bites, summer season or rural living and (ii) including TOSV as a target in neurological panels in the routine laboratory-testing at hospitals in southern Spain. Finally, this work demonstrates the utility of mNGS to identify viral pathogens directly in CSF and provides a methodological improvement to facilitate TOSV full-genome sequencing [32]. The detection of a novel, disease-associated, reassortant TOSV emphasises the importance of full-genome surveillance to monitor the spread and evolution of phleboviruses.

Ethical statement

Ethical approval for the experimental protocol was given by the Ethical Review Board of Cordoba (Comité de Ética de Investigación de Córdoba) under protocol 201999903552445. Written informed consent was obtained from all patients/legal guardians. The procedures were carried out in accordance with approved guidelines, regulations and the principles of the Declaration of Helsinki.

Funding statement

ESL laboratory received funding from Institut Pasteur, from the INCEPTION programme (Investissements d'Avenir grant ANR-16-CONV-0005), from the NIH PICREID program (Award Number U01AI151758) and from the Labex IBEID (ANR-10-LABX-62-IBEID). MDFG acknowledges funding from the Junta de Andalucía (PI-0216-2019) and Instituto de Salud Carlos III (Acción Estratégica en Salud Intramural PI20CIII/00005).

Data availability

Raw sequencing reads generated in this study were deposited on the European Nucleotide Archive (<https://www.ebi.ac.uk/ena/browser/home>) with study number PRJEB52962. The GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) accession numbers of the assembled virus genomes are OP554790–OP554810.

Acknowledgements

We would like to thank all the healthcare workers who took care of the patients involved in this study. We acknowledge the authors, originating and submitting laboratories of the sequences from GenBank used in this work. We acknowledge the European virus archive–Global (EVAg) for providing control TOSV RNAs. This work used the computational and storage services provided by the IT department at Institut Pasteur, Paris.

Conflict of interest

None declared.

Authors' contributions

Conceptualisation and design of experiments: MDFG, ESL; Sample collection: EA; Clinical data collection: EA, ABP,

LMM; Project administration: ABP; Sequencing: MDFG, MP, FG; Phylogenetic analysis: MDFG, FG, ESL; Writing – original draft: MDFG, FG, ESL. Writing – critical reviewing and editing: ABP, AB, AV, ESL, FG, LMM, MDFG, MPSS; Study supervision (contributed equally): MDFG, ESL. All authors read and approved the final version of the manuscript.

References

- James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. , GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789–858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7) PMID: 30496104
- John CC, Carabin H, Montano SM, Bangirana P, Zunt JR, Peterson PK. Global research priorities for infections that affect the nervous system. *Nature*. 2015;527(7578):S178–86. <https://doi.org/10.1038/nature16033> PMID: 26580325
- McGill F, Griffiths MJ, Bonnett LJ, Geretti AM, Michael BD, Beeching NJ, et al. , UK Meningitis Study Investigators. Incidence, aetiology, and sequelae of viral meningitis in UK adults: a multicentre prospective observational cohort study. *Lancet Infect Dis*. 2018;18(9):992–1003. [https://doi.org/10.1016/S1473-3099\(18\)30245-7](https://doi.org/10.1016/S1473-3099(18)30245-7) PMID: 30153934
- Shukla B, Aguilera EA, Salazar L, Wootton SH, Kaewpoowat Q, Hasbun R. Aseptic meningitis in adults and children: Diagnostic and management challenges. *J Clin Virol*. 2017;94:110–4. <https://doi.org/10.1016/j.jcv.2017.07.016> PMID: 28806629
- Faria NR, Azevedo RDS, Kraemer MUG, Souza R, Cunha MS, Hill SC, et al. Zika virus in the Americas: Early epidemiological and genetic findings. *Science*. 2016;352(6283):345–9. <https://doi.org/10.1126/science.aaf5036> PMID: 27013429
- Ramachandran PS, Wilson MR. Metagenomics for neurological infections - expanding our imagination. *Nat Rev Neurol*. 2020;16(10):547–56. <https://doi.org/10.1038/s41582-020-0374-y> PMID: 32661342
- Instituto Nacional de Estadística (INE). Córdoba: Población por municipios y sexo. [Population by municipalities and sex]. Madrid: INE. Spanish. Available from: <https://www.ine.es/jaxiT3/Datos.htm?t=2901>
- Gámbaro F, Pérez AB, Agüera E, Prot M, Martínez-Martínez L, Cabrerizo M, et al. Genomic surveillance of enterovirus associated with aseptic meningitis cases in southern Spain, 2015–2018. *Sci Rep*. 2021;11(1):21523. <https://doi.org/10.1038/s41598-021-01053-4> PMID: 34728763
- Kalantar KL, Carvalho T, de Bourcy CFA, Dimitrov B, Dingle G, Egger R, et al. IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience*. 2020;9(10):giaa111. <https://doi.org/10.1093/gigascience/giaa111> PMID: 33057676
- Pickett BE, Greer DS, Zhang Y, Stewart L, Zhou L, Sun G, et al. Virus pathogen database and analysis resource (ViPR): a comprehensive bioinformatics database and analysis resource for the coronavirus research community. *Viruses*. 2012;4(11):3209–26. <https://doi.org/10.3390/v4113209> PMID: 23202522
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. *Nucleic Acids Res*. 2013;41(Database issue) D1:D36–42. PMID: 23193287
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80. <https://doi.org/10.1093/molbev/mst010> PMID: 23329690
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268–74. <https://doi.org/10.1093/molbev/msu300> PMID: 25371430
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14(6):587–9. <https://doi.org/10.1038/nmeth.4285> PMID: 28481363
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*. 2018;35(2):518–22. <https://doi.org/10.1093/molbev/msx281> PMID: 29077904
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol*. 2015;1(1):vev003. <https://doi.org/10.1093/ve/vev003> PMID: 27774277
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999;73(1):152–60. <https://doi.org/10.1128/JVI.73.1.152-160.1999> PMID: 9847317
- Uchida K, Shinohara M, Shimada S, Segawa Y, Doi R, Gotoh A, et al. Rapid and sensitive detection of mumps virus RNA directly from clinical samples by real-time PCR. *J Med Virol*. 2005;75(3):470–4. <https://doi.org/10.1002/jmv.20291> PMID: 15648065
- Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*. 2017;12(6):1261–76. <https://doi.org/10.1038/nprot.2017.066> PMID: 28538739
- Alten B, Maia C, Afonso MO, Campino L, Jiménez M, González E, et al. Seasonal Dynamics of Phlebotomine Sand Fly Species Proven Vectors of Mediterranean Leishmaniasis Caused by *Leishmania infantum*. *PLoS Negl Trop Dis*. 2016;10(2):e0004458. <https://doi.org/10.1371/journal.pntd.0004458> PMID: 26900688
- Collao X, Palacios G, Sanbonmatsu-Gámez S, Pérez-Ruiz M, Negro Al, Navarro-Marí JM, et al. Genetic diversity of Toscana virus. *Emerg Infect Dis*. 2009;15(4):574–7. <https://doi.org/10.3201/eid1504.081111> PMID: 19331735
- Saha S, Ramesh A, Kalantar K, Malaker R, Hasanuzzaman M, Khan LM, et al. Unbiased Metagenomic Sequencing for Pediatric Meningitis in Bangladesh Reveals Neuroinvasive Chikungunya Virus Outbreak and Other Unrealized Pathogens. *MBio*. 2019;10(6):e02877–19. <https://doi.org/10.1128/mBio.02877-19> PMID: 31848287
- Ministerio De Sanidad. Informe de situación y evaluación del riesgo de enfermedad por flebovirus transmitidos por flebotomos en España. [Situation report and risk assessment of sandfly-borne phlebovirus disease in Spain]. Madrid: Centro de Coordinación de Alertas y Emergencias Sanitarias (CCAES); 2019. Spanish. Available from: https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/analisisituacion/doc/ER_Flebovirus.pdf.
- Piantadosi A, Mukerji SS, Ye S, Leone MJ, Freimark LM, Park D, et al. Enhanced Virus Detection and Metagenomic Sequencing in Patients with Meningitis and Encephalitis. *MBio*. 2021;12(4):e0114321. <https://doi.org/10.1128/mBio.01143-21> PMID: 34465023
- Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, et al. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl J Med*. 2019;380(24):2327–40. <https://doi.org/10.1056/NEJMoa1803396> PMID: 31189036
- Ayhan N, Charrel RN. An update on Toscana virus distribution, genetics, medical and diagnostic aspects. *Clin Microbiol Infect*. 2020;26(8):1017–23. <https://doi.org/10.1016/j.cmi.2019.12.015> PMID: 31904562
- Gori Savellini G, Gandolfo C, Cusi MG. Epidemiology of Toscana virus in South Tuscany over the years 2011–2019. *J Clin Virol*. 2020;128:104452. <https://doi.org/10.1016/j.jcv.2020.104452> PMID: 32474372
- Sanbonmatsu-Gámez S, Pedrosa-Corral I, Navarro-Marí JM, Pérez-Ruiz M. Update in Diagnostics of Toscana Virus Infection in a Hyperendemic Region (Southern Spain). *Viruses*. 2021;13(8):1438. <https://doi.org/10.3390/v13081438> PMID: 34452304
- Informe Semanal. Sistema Vigilancia Epidemiológica de Andalucía (SVEA). Meningitis Vírica 2016–2017. [Viral meningitis in Andalusia 2016–2017]. Spanish. 2019;24(3):BES 0319. Available from: https://www.repositoriosalud.es/bitstream/10668/3049/1/SVEAIS_24_03_2019.pdf
- Laroche L, Jourdain F, Ayhan N, Bañuls AL, Charrel R, Prudhomme J. Incubation Period for Neuroinvasive Toscana Virus Infections. *Emerg Infect Dis*. 2021;27(12):3147–50. <https://doi.org/10.3201/eid2712.203172> PMID: 34808074
- Alkan C, Bichaud L, de Lamballerie X, Alten B, Gould EA, Charrel RN. Sandfly-borne phleboviruses of Eurasia and Africa: epidemiology, genetic diversity, geographic range, control measures. *Antiviral Res*. 2013;100(1):54–74. <https://doi.org/10.1016/j.antiviral.2013.07.005> PMID: 23872312
- Ayhan N, Charrel RN. An update on Toscana virus distribution, genetics, medical and diagnostic aspects. *Clin Microbiol Infect*. 2020;26(8):1017–23. <https://doi.org/10.1016/j.cmi.2019.12.015> PMID: 31904562
- Dersch R, Sophocleous A, Cadar D, Emmerich P, Schmidt-Chanasit J, Rauer S. Toscana virus encephalitis in Southwest Germany: a retrospective study. *BMC Neurol*. 2021;21(1):495. <https://doi.org/10.1186/s12883-021-02528-7> PMID: 34937553
- Briese T, Calisher CH, Higgs S. Viruses of the family Bunyaviridae: are all available isolates reassortants?

- Virology. 2013;446(1-2):207-16. <https://doi.org/10.1016/j.virol.2013.07.030> PMID: 24074583
35. Collao X, Palacios G, de Ory F, Sanbonmatsu S, Pérez-Ruiz M, Navarro JM, et al. Granada virus: a natural phlebovirus reassortant of the sandfly fever Naples serocomplex with low seroprevalence in humans. *Am J Trop Med Hyg.* 2010;83(4):760-5. <https://doi.org/10.4269/ajtmh.2010.09-0697> PMID: 20889862
 36. Palacios G, Tesh R, Travassos da Rosa A, Savji N, Sze W, Jain K, et al. Characterization of the Candiru antigenic complex (Bunyaviridae: Phlebovirus), a highly diverse and reassorting group of viruses affecting humans in tropical America. *J Virol.* 2011;85(8):3811-20. <https://doi.org/10.1128/JVI.02275-10> PMID: 21289119
 37. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J Virol Methods.* 2014;195:194-204. <https://doi.org/10.1016/j.jviromet.2013.08.035> PMID: 24036074

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2023.