Diagnosis and Genomic Characterization of the Largest Western Equine Encephalitis Virus Outbreak in Uruguay During 2023–2024

Gonzalo Tomás
Universidad de la República
Ana Marandino
Universidad de la República
Sirley Rodríguez
Ministerio de Ganadería Agricultura y Pesca
Gabriel Luz Wallau
Instituto Aggeu Magalhães (IAM)- 20 Fundação Oswaldo Cruz (FIOCRUZ)
Filipe Zimmer Dezordi
Instituto Aggeu Magalhães (IAM)- 20 Fundação Oswaldo Cruz (FIOCRUZ)
André Luiz Sá de Oliveira
Instituto Aggeu Magalhães (IAM)- 20 Fundação Oswaldo Cruz (FIOCRUZ)
Claudia Techera
Universidad de la República
Lucía Calleros
Universidad de la República
Sofía Grecco
Universidad de la República
Joaquín Williman
Universidad de la República
Ramiro Pérez
Ministerio de Ganadería Agricultura y Pesca
Lucía Bassetti
Ministerio de Ganadería Agricultura y Pesca
Raúl Negro
Ministerio de Ganadería Agricultura y Pesca
Lucía Moreira Marrero
Universidad de la República
Adriana Delfraro
Universidad de la República
Roberto Vidal
Ministerio de Ganadería Agricultura y Pesca
Yanina Panzera
Universidad de la República
Ruben Pérez
rperez@fcien.edu.uy

Universidad de la República

Article

Keywords: arbovirus, alphavirus, genome evolution, diagnosis, qPCR, PCR enrichment, next-generation sequencing

Posted Date: July 25th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4547844/v1

Additional Declarations: No competing interests reported.

Abstract

The Western Equine Encephalitis Virus (WEEV) is transmitted between mosquitoes and wild birds. Mosquitoes can spread the virus to horses and human populations, causing severe encephalitis and death. The most recent large outbreak of WEEV occurred in the Southern cone of South America from November 2023 to April 2024, affecting many equines and humans in Argentina and Uruguay. We identified and genetically characterized WEEV strains during this outbreak to understand their evolutionary trends and rapid expansion in the country and at international borders. The virus affected 1,086 horses and caused 388 deaths in all regions of Uruguay. We obtained genomes from 15 strains using a novel multiplex PCR assay combined with next-generation Illumina sequencing. The phylogenetic analysis revealed that samples from Uruguay and Brazil collected during 2023–2024 and an Argentine strain from 1958 share a common evolutionary origin and are distinct from North American strains. Phylogenetic and epidemiological data on the outbreak suggest that it originated in Argentina and spread to Uruguay and Brazil, likely by movements of infected birds. Genomic analysis also revealed mispairing in real-time PCR primers and probes that may affect official diagnostic protocols, highlighting the need for assay updates. Our research emphasizes the need to map the genetic diversity of WEEV in South America to understand their epidemiology and develop effective control approaches.

INTRODUCTION

Arthropod-borne viruses (arboviruses) are high-burden priority pathogens of global concern [1]. Among these, the *Alphavirus* genus (Togaviridae family) inflicts considerable morbidity and mortality in animals and humans [2]. This genus is broadly split into Old World and New World alphaviruses, each with distinct characteristics and impact on health [3–5]. The Old World alphaviruses comprise Semliki Forest, O'Nyong-Nyong, Ross River, Chikungunya, and Sindbis viruses. The New World alphaviruses include the Venezuelan equine encephalitis virus, the Eastern Equine Encephalitis virus, and the Western Equine Encephalitis Virus [6].

The Western Equine Encephalitis Virus (WEEV) is transmitted between mosquitoes (*Culex tarsalis* in North America and *Aedes albifasciatus* in South America), having wild birds as primary enzootic cycle reservoirs [7,8]. The virus might spill over into horses and human populations through mosquitoes that opportunistically feed on mammals and start epizootic cycles. Equine and human infections, though typically asymptomatic, can lead to severe encephalitis and death, with case-fatality rates ranging from 3–4% in humans and 15–20% in equines [9]. Horses and humans are dead-end hosts and do not contribute to further epizootic transmissions [7]. There are no effective treatments for active WEEV infection, but vaccines are available for the prophylactic therapy of equines [10].

Most data available from WEEV epizootic outbreaks was registered from passive surveillance and outbreak events between the 40s and the 90s in North America. The first large-scale outbreak was recorded in the western United States around 1940, with about half a million equine cases and many thousands of human infections [11], followed by another significant outbreak in the 1980s in the United States and Canada [12]. In Mexico, epizootics involving more than 41 horses were reported in 2019 [13]. Over the past decades, infections in the United States and Canada have dropped significantly [14–16].

WEEV diagnosis and classification were initially conducted using serological methods and, more recently, through genomic phylogenetic and phylodynamic analysis [27]. The WEEV genome comprises a positive-sense, single-stranded RNA molecule of approximately 11.7 Kb in length. Non-coding regions flank it at the 5' and 3' ends and are involved in virus replication and translation [28]. The genomic RNA encodes four nonstructural proteins (nsP1 to nsP4). A subgenomic RNA encodes five structural proteins: a nucleocapsid protein (C), two envelope glycoproteins (E3 and E2), and two smaller accessory peptides (6K and E1). Envelope glycoprotein heterodimer (E1 + E2) contains virus-specific neutralizing epitopes; nucleocapsid protein has broadly cross-reactive epitopes with other alphavirus.

Studies of WEEV were limited in recent years due to the pathogen's submergence [14]. Detailed phylogenetic analysis of WEEV has been mainly restricted to North American strains [14,15,29]. The A, B1, B2, B3, and C1 groups were proposed to characterize the WEEV variability. Groups A and B are composed mainly of North American strains, but a recently described group C was proposed for one ancient strain from Argentina and three strains collected in 2024 from Brazilian municipalities bordering Uruguay [30]. However, remarkably few genomes from South America were available, making it difficult to determine their phylogenetic relationships with the remaining sequences.

The most recent large WEEV outbreak unfolded in the Southern cone of South America from November 2023 to April 2024, affecting many equines and humans in Argentina and Uruguay [27]. In Brazil, only three infected equines have been identified by retrospective analyses of deceased horses suspected of rabies infection [30].

Due to the large scale of Argentina and Uruguay outbreaks, the Pan American Health Organization/World Health Organization issued an epidemiological alert on the risk to human health associated with WEEV infection in December 2023, emphasizing the importance of strengthening epidemiological surveillance, diagnosis, intersectoral coordination, and vector control. As part of a coordinated response to this unfolding emergency, we present the epidemiological and genomic characterization of the largest WEEV outbreak in Uruguay from 2023 to 2024. This was possible by developing and testing a new next-generation sequencing protocol for WEEV diagnostic and characterization. Our findings reveal the evolutionary trends and rapid expansion of WEEV in the country and at international borders. We also identified genomic variability that affects WEEV qPCR primers and probes and the binding effectivity of referenced protocols for diagnostics [27], highlighting the need for diagnostic assay reevaluation.

MATERIALS AND METHODS

Dataset and sampling collection

The MGAP collected all the samples and data on the horses affected by the outbreak, including date, location, number of infected animals, clinical signs, and disease outcome. To display this data, we set up a Microreact instance [31], which harbors the whole dataset from positive cases in Uruguay.

Samples for genetic studies corresponded to brain tissue and cerebrospinal fluid collected from deceased horses exhibiting clinical signs of encephalitis. These samples from the outbreak's peak were transported under refrigerated conditions for molecular diagnosis (Table 1).

WEEV diagnosis

Total RNA extraction was performed using the taco™ Automatic Nucleic Acid Extraction System (GeneReach Biotechnology, Taiwan, China). The virus was initially identified using a generic nested PCR to detect and classify alphaviruses [32]. The nested PCR targeted the nsP4 (RNA-dependent RNA polymerase) region and produced a 481 bp-long amplicon in the first PCR round and a 195 bp-long amplicon in the second one. Complementary DNA (cDNA) and the first PCR round were performed using AgPath-IDTM One-Step RT-PCR Reagents (Thermo Fisher Scientific Inc, USA). The second round used Platinum™ Taq DNA Polymerase (Thermo Fisher Scientific Inc, USA).

WEEV was confirmed by the RT-qPCR described by Lambert et al. [33] and suggested by the WOAH Terrestrial Manual [34]. This hydrolysis probe-based assay amplifies a 67 bp fragment in the E1 coding region. An alternative WEEV-specific RT-qPCR method was also applied for comparison purposes [35]. This assay also uses a hydrolysis probe and amplifies an 80 bp fragment within the E2 coding region. Amplification reactions were conducted using a 7500 Real-Time PCR System (Applied Biosystems, Waltham, MA, USA) and the MIC-4 Thermal Cycler Real-time qPCR (Bio Molecular Systems BMS. Australia).

Amplicon-based next-generation sequencing (NGS)

Primer design

All available full-length WEEV genomes were retrieved from the NCBI database to generate a comprehensive dataset. Multiple alignments were performed using MAFFT v7.490 in Geneious Prime® 2023.1.2 [36]. To enrich the viral genome before massive sequencing, we designed primer sets generating overlapping amplicons covering the entire genome. The suitability of each primer was evaluated using the OligoAnalyzer™ tool from IDT (https://www.idtdna.com/pages/tools/oligoanalyzer) to ensure optimal Tm and avoidance of hairpin and dimer structures.

Multiplex PCR-NGS reaction

The Illumina Microbial Amplicon Prep (IMAP) kit was used to construct amplicon-based libraries, following the manufacturer's protocol. Each sample underwent an RT reaction with 8.5 µL of total RNA, followed by two multiplex-PCR amplification reactions (pool 1 and pool 2) using 5 µL of cDNA. Pool 1 and pool 2 multiplex PCR reactions generated non-overlapping amplicons. Each primer was used at a final concentration of 0.4 µM. The PCR cycling protocol was the same for both pools. It consisted of an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of denaturation at 95°C for 30 seconds, and an annealing/extension step at 60°C for 2 minutes. A final extension step at 72°C for 5 minutes completed the amplification process. Generation of the expected amplicons was verified by 0.8% agarose gel electrophoresis. Subsequently, 10 µL from each PCR reaction (pools 1 and 2) was pooled and subjected to tagmentation and indexing. The library's quality and fragment length distribution was assessed using the Agilent high-sensitivity DNA kit (Agilent, Santa Clara, CA, USA) on a Fragment Analyzer™ System (Advanced Analytical Technologies Inc., Heidelberg, Germany). Genome sequencing was performed on an Illumina MiSeq platform (Illumina, USA) using the MiSeq Reagent Kit v2 (300 cycles).

Genome assembly

The raw data was processed using the Geneious BBDuk v38.84 tool. Illumina adapters and primer sequences were removed, and low-quality reads (Phred quality scores < 30) were filtered out. The remaining reads were aligned to MN477208 as the reference genome using Minimap2 v2.17 (Li, 2018). Assemblies were visually and manually curated to generate majority consensus sequences; annotations were transferred from reference strains.

Phylogenetic analysis

We obtained all WEEV genomes larger than 7kb from the BV-BRC database on 25.02.2024 for a detailed phylogenetic analysis. This dataset included six recently submitted Uruguayan strains (PP620641–PP620646) obtained through a probe-capture enrichment NGS assay.

Recombination was assessed with the RDP program in the dataset utilized for phylogenetic analysis [37]. Maximum-likelihood trees were inferred from IQTREE v2.1.2, and FastTree v2.1.11 was implemented within Geneious [38]. We also investigated the temporal signal of the current 2023 – 2024 outbreak samples. The root-to-tip divergence was low (R ~ 0.45); therefore, we performed a phylogenetic reconstruction with no temporal scale. The resulting tree was midpoint rooted as it was unclear which sister alphavirus should be used as an outgroup.

Amino acidic markers

Using the same genomic dataset as in the phylogenetic analysis, we characterized nucleotide and amino acid mutations in the coding sequence of each viral protein using TreeTime [39] and augur-auspice (https://auspice.us/) via Nextstrain [40]. Mutations in each tree branch of the major lineages and clades described in this study were compared with the reference basal genome (GQ287646) and further visually inspected in the original alignment.

RESULTS

Virus identification and diagnosis

WEEV was initially characterized as the causative agent of the outbreak by sequencing a generic RT-nested PCR followed by Sanger sequencing of a 195 bp fragment of the coding region of the nsP4 protein [32]. Once identified, the virus was diagnosed using the qPCR methods described by Lambert et al. and Brault et al. [33,35]. The Ct values obtained with the Lambert et al. method were relatively high, with an average Ct value of 36.5 (Table 1). When the Brault et al. method was used in the same samples, lower Ct values were obtained in all cases, with an average Ct value of 27.5 (Table 1).

Epidemiological characterization of the outbreak

Equine cases were tracked from November 28, 2023, to April 10, 2024. Mortality and morbidity were determined over an estimated equine population of 405,644 animals. The number of equines exposed to the WEEV infection was 16,863 from 605 establishments (farms, ranches, racecourses, and backyards) (https://microreact.org/project/gofP6cvtZmUv2ZMaGExUYC-weevuruguayrs). The virus affected 1,086 horses, including 80 confirmed cases with molecular diagnostics, and produced 388 deaths across all 19 Uruguayan Departments (https://microreact.org/project/nvwRMyT8EVHq4yqbpUiM6v-weevuruguayrspositive) (Fig. 1a and b).

Clinical signs in horses include fever, anorexia, and depression. In severe cases, progression leads to hyperexcitability, blindness, ataxia, severe neurologic depression, recumbency, convulsions, and death.

The number of infected horses showing clinical symptoms and death increased from November 2023, reaching its peak in the first weeks of December 2023 (Fig. 1a). These 2023 cases were primarily concentrated in the Uruguayan departments that border Argentina (Fig. 1c). This was followed by a decrease between the second half of December and the first week of January 2024 and a new increase in cases in the second week of January 2024. This second peak in 2024 was more concentrated in the central and east departments of Uruguay (Fig. 1d).

Developing a new NGS protocol for obtaining WEEV genomes

Nine primer sets were designed within conserved genome regions, incorporating degenerate bases as necessary (Table 2). Amplicons range from 1031 to 1736 bp long, with overlapping areas averaging 100 bp (Supplementary Fig. 1). Most samples yielded the expected numbers and lengths of amplicons, producing five amplicons for pool 1 and four for pool 2 (Supplementary Fig. 1).

The nine amplicons were included in IMAP Illumina libraries and sequenced. Filtered reads were aligned with the reference genome, resulting in 15 full-length and 4 partial genome sequences, exhibiting breadth coverage ranging from 53.6–94.2%. Three strains previously sequenced by probe-capture NGS sequencing (PP620641, PP620642, PP620646) were re-sequenced in the current study (Table 1).

NGS of complete WEEV genomes produced an average of 481,750 reads per sample, ranging from 268,709 to 1,489,6367 reads. The mean coverage depth of CDS regions ranged from 765 to 17,322×.

Genome characterization and variability

The 15 full-length consensus sequences generated (11,216 nt) have 11,159 nt corresponding to coding sequences (CDS). Overall, these WEEV genomes were highly conserved, with a nucleotide similarity ranging from 99.75 to 100% and an amino acid similarity of 99.78 to 100%. Nonstructural protein genes (nsP1 to nsP4) showed similar levels of heterogeneity as structural protein genes (C, E3, E2, 6K, and E1), with an average nucleotide similarity ranging from 98.99% (nsP3) to 99.37% (C) and an average amino acid similarity ranging from 99.34% (nsP3) to 99.92% (E3).

Phylogenetic analysis

There was no significant evidence of recombination in the entire genomic dataset analyzed (72 sequences). The maximum-likelihood (ML) phylogeny of the genome data set shows that most WEEV sequences belong to either a North American (NA) or South America (SA) lineage (Fig. 2). These NA and SA lineages share a common ancestor and have three basal South American strains (two from Argentina and one from Guyana). The NA lineage is divided into sublineages A and B. The

A sublineage includes mosquitoes, humans, and horse samples collected from 1930 to 1971 in the United States, Canada, Russia, and Cuba. The B sublineage includes samples from mosquitoes and horses from 1946 to 2005. Within sublineage B are three genogroups: B1, B2, and B3 [15]; B1 and B2 are paraphyletic, and B3 is monophyletic.

The SA lineage encompasses a monophyletic clade with the synchronic horse samples from Uruguay and Brazil collected during 2023–2024 and a basal Argentine strain from 1958 (KT844543). The Brazilian and Uruguayan 2023–2024 clade is divided into three subclades: I, II, and III (aLRT/UFBOOTS > 98) (Fig. 3). Strains from Brazil and Uruguay are clustered in subclades I, while subclades II and III are composed of genomes from Uruguay only (Fig. 3a). The subclade strains are widely distributed across various Uruguayan Departments (Figs. 3b, c, and d) and exhibit a clear polytomy, evidencing low within-clade genetic variability and fast strain spreading.

Amino acid markers

Different amino acid markers defined the monophyletic groups described in this study. The 2023–2024 clade depicted two amino acid markers, and their three subclades contained one or three unique substitutions (subclade I: nsP2-H297Y, nsP2-V460I, and E3-P15L, subclade II: nsP2-K418E, nsP3-A44T, and E1-I213L, and subclade III: nsP2- I641T). One residue in subclade I (E2-413V) was also shared by the B3 monophyletic genogroup.

Variability in target qPCR sequences

Our analysis of complete WEEV genome sequences revealed a single mismatch in a primer delineated by Brault et al.[35], along with several mismatches in the probes and primers outlined by Lambert et al. [33](Fig. 4).

In Brault's assay, we identified a mismatch in the reverse primer (19 nt) exhibited in the 2023–2024 clade but not in the basal strain within the SA lineage.

The differences were notably more pronounced in Lambert's assay, where mismatches were observed in both NA and SA lineages. Within the NA lineage (B3 genogroup), strains exhibited one to two mismatches in the 23-nt long probe, while the basal strain and strains from the 2023–2024 clade in the SA lineage showed two to three mismatches. Moreover, certain strains within the 2023–2024 clade of the SA lineage displayed a mismatch with the forward primer (20 nt). The reverse primer (20 nt) exhibited one to three differences across variants from both NA and SA lineages (Fig. 4).

DISCUSSION

The WEEV was responsible for large human and equine outbreaks in western North America throughout the early to mid-20th century, leading to thousands of equine and human deaths due to encephalitis. However, during the late 20th century, a marked reduction in WEEV transmission was recorded, with the last human case documented in 1998 and a decrease in infected mosquito pools identified in longitudinal surveillance programs [14–16].

A similar trend has occurred in South America in recent decades, with only sporadic cases occurring after the 1980s [22]. This period of epidemiological silence contrasts with the present outbreak from 2023 to 2024, one of the largest and deadliest in South America's history.

Although the clinical data suggested an equine encephalitis virus, the outbreak was unequivocally characterized by sequencing a conserved coding region of the RNA-dependent RNA polymerase (nsP4) that distinguishes WEEV from other related alphaviruses [32]. The WEEV identification allowed the deployment of more cost and time-effective qPCR methods, including the PAHO/WHO recommended assay [27,33]. However, the Lambert et al. methodology [33] performed poorly in Uruguayan strains, likely due to the mismatches in probes and PCR primers (Fig. 4). In comparison, the alternative method described by Brault et al. [35] has greater homology in the target region of the 2023–2024 outbreak strains, resulting in lower Ct values and improved sensitivity (Fig. 4).

The 2023–2024 outbreak affected 1,419 equine and 58 human cases in Argentina. In Uruguay, it impacted 1086 equines, with approximately 35% of them dying from severe neurological symptoms. Considering that Argentina has eight times more horses than Uruguay [41], the outbreak has been particularly severe in Uruguay. However, it is important to note that differences in reporting may exist. The estimated mortality (horse deaths/population size = 388/405,644) was 0.09%, and the morbidity (affected horses/population size = 1,086/405,644) was 0.26%. During the outbreak, vaccination was recommended for the affected ranches and equine gathering events (Resolution Nº 282, MGAP). The Ministry of Public Health reported five human WEEV cases during the same period. These findings underscore the threat of WEEV to equine and human health and highlight the need for ongoing surveillance.

The multiplex-NGS methodology applied to the 2023–2024 South American outbreak yielded 15 complete genomes directly from deceased horse samples. These new WEEV sequence data significantly increased the number of publicly available genomes in GenBank and allowed us to perform the first large-scale genomic study of this arbovirus in South America. Our findings support the existence of two main lineages of WEEV. The North American lineage contains strains from various hosts (mosquitoes, birds, horses, and humans) collected from 1930 to 2005. These strains were relatively genetically stable, with only a maximum of 3.7% nucleotide sequence divergence [15]. This lineage was extensively described and divided into groups A and B. Group B comprised three genogroups (B1 to B3) that are not necessarily monophyletic, bearing specific amino acid markers and circulation periods [14,51]. Group A includes the original California strain (KJ554965.1), obtained from a horse in 1930, the human isolate McMillan (GQ287640.1) from Canada (1941), the Fleming strains (MN477208.1) from the USA, and strains from Russia (1962) and Cuba (1971). In the United States, group A might have become extinct in the 1940s and was displaced by Group B, which became predominant. Strains from ancestral groups A and B1 are generally more virulent than recent strains from groups B2 and B3. When all three group B sublineages were circulating, there was a concurrent increase in estimated viral population size between 1965 and the late 1980s. However, after the late 1980s, a reduction in estimated population size occurred when the group B3 viruses became predominant in North America [14,15].

South America has a more limited strain diversification than North America, possibly due to the fewer available strains from historical outbreaks. The South American lineage comprises horse strains from 2023 and 2024 (Uruguay and Brazil) and the CBA87 strain collected in 1958 from Argentina (Córdoba).

The two American lineages share a common ancestor with basal strains from Argentina (horse and mosquito) and Guyana (horse). The phylogenetic clustering suggests that WEEV in America shared a common origin and diversified into North and South lineages.

The 2023–2024 outbreak started in Argentina and was later detected in Uruguay and Brazil. Strains from Uruguay and Brazil are closely related and bear several clade-specific mutations. They are here denoted as a clade because they compose a monophyletic group, share a recent common ancestor, and are synchronous. This clade comprises the Uruguayan and Brazilian subclade I and the Uruguayan unique monophyletic subclades II and III (Fig. 3). The Argentine CBA87 strain collected in 1958 from a horse is basal for the 2023–2024 outbreak clade. This old Argentine strain might represent an ancestor of the outbreak's source that persisted in an enzootic transmission cycle.

This study identified distinct nucleotide synonymous changes and amino acids that define WEEV phylogenetic clustering. We analyzed the 2023–2024 clade to determine the presence of neurovirulence and transmission markers previously described. All the 2023–2024 clade strains contained the E2-214R residue previously associated with low neurovirulence in a murine model [52]. Residue E2-214R was also identified as necessary for the efficient infectivity of the mosquito vector *Culex tarsalis* [52]. Additionally, the 2023–2024 clade lacks the six amino acid mutations previously described to increase fitness in avian and mosquito hosts (nsP3-52I, nsP4-602S, C-89R, C-250W, E2-23T, and E1-374S) [14]. The clade contains all six residues resembling the ancestral, less fit state (nsP3-52T, nsP4-602N, C-89K, C-250K, E2-23A, and E1-374T).

Our analysis revealed no marker linked to distinct WEEV phenotypic changes. All the unique amino acid markers detected in our analysis that have not been previously studied for their virulence or transmission characteristics should be considered subclade markers. Whether these markers have functional implications or are influenced by the mosquito or avian species involved remains unanswered, highlighting the need for further research.

Interestingly, the 2023–2024 outbreak spread rapidly in Uruguay, beginning in the region bordering Argentina and reaching the central and eastern areas. This spread pattern aligns with WEEV emerging in Argentina and spreading to Uruguay following a west-to-east expansion (Fig. 1). Similar behavior may have occurred in Brazil, where municipalities in the West detected the first cases, followed by a city from the East part of Rio Grande do Sul state.

Historically, equine encephalitis activity follows multiyear cycles, with epidemics occurring after periods of excessive rainfall starting during the preceding year [53,54]. In November and December 2023, the weather in neighboring Argentina, Brazil, and Uruguay regions showed high temperatures and rain, leading to mosquito proliferation, particularly those of the *Aedes albifasciatus* species. The virus may have emerged in Argentina from genetically close strains that invaded Uruguay and Brazil. Genetic polymorphisms in the original Argentine viral population and different crossing-border events would explain the different subclades observed in Uruguay and Brazil (Fig. 3a).

Mosquitoes have a narrow flight range of less than a kilometer, and horse transportation between regions does not impact the transmission because the viremia in these animals is insufficient to infect new mosquitoes. Birds and, to a lesser extent, some mammals, such as lagomorphs and rodents, are the amplification hosts that transmit the virus to the local mosquito population [22]. House sparrows are a natural enzootic amplification/reservoir host for WEEV in North America, developing a high-titer, short-lived viremia that peaks on day 1 post-infection without detectable morbidity [55]. Therefore, wild birds with fast displacement across the country are the more likely amplification host associated with transmission. The outbreak coincided with migratory and indigenous birds arriving in areas with abundant water sources, including wells, lakes, ponds, streams, or rivers. Migratory birds remain the most plausible means of transporting the virus long distances from southern locations.

A limitation of this study is the absence of molecular diagnostics in asymptomatic and recovered horses. The lack of Argentine WEEV sequences for comparison and the unavailability of sequences from mosquitoes, amplification hosts, and humans hinders a more comprehensive description of the outbreak's epidemiological scenario.

CONCLUSION

Genomic epidemiology studies remain a priority for a better understanding of viruses' biology and epidemiology and for developing effective treatments, vector control approaches, and vaccines. In this sense, we expect other researchers to adopt the method for NGS-sequencing WEEV genomes and adjust the qPCR methodology by considering the mismatches found in probes and primers. Our research emphasizes the importance of mapping the genetic diversity of WEEV in South America, which may uncover additional enzootic lineages beyond the three lineages identified in this study. Genomic surveillance for certain viruses in this region has been neglected, which hinders our ability to monitor, control, and prevent future outbreaks. The scientific community needs additional resources to improve our understanding of emerging and re-emerging pathogens that may affect animals and humans within a One Health approach.

Declarations

Acknowledge

We want to express our gratitude to Valeria Gayo from DILAVE for effectively coordinating the research activities in this study. We thank Biko S.A. (Elena Fernández and Joaquín Lozano) for technical support (Illumina sequencing).

Financial support: this study did not receive any financial support.

Conflicts of Interest: The authors declare no conflict of interest.

Institutional Review Board Statement: The study does not require ethical approval because it was conducted by the Ministry of Livestock, Agriculture, and Fisheries of Uruguay (MGAP) as part of public health surveillance by official veterinary authorities (DILAVE) in response to equine encephalitis outbreaks.

Data Availability statement

The assembled WEEV sequences generated in this study are available in the GenBank database under accession numbers PP747340 to PP747354. The raw sequencing dataset used for phylogenetic inferences is available as Supplementary Dataset 1. No custom code or scripts were used to generate or analyze this dataset.

Author Contribution Statement

GT: conceptualization, methodology (sequencing and qPCR), and validation. **AM** and **SR**: conceptualization, methodology (sequencing and qPCR), and data interpretation. **GW**, **FD**, **AO**: bioinformatic analysis and data interpretation. **CT**, **LC**, **SG**, **JW**: methodology (sequencing and qPCR). **RaP**, **LB**, **RN**, **LM**, and **AD**: molecular diagnosis (PCR and qPCR). **RV**: geolocalization and analysis of outbreak data. **YP**: conceptualization, methodology (sequencing). **RP**: conceptualization, bioinformatic analysis and data interpretation, writing (original draft), overall organization. All authors revised and approved the final manuscript.

References

1. Girard, M., Nelson, C. B., Picot, V. & Gubler, D. J. Arboviruses: A global public health threat. Vaccine 38, 3989–3994 (2020).
2. Azar, S. R., Campos, R. K., Bergren, N. A., Camargos, V. N. & Rossi, S. L. Epidemic Alphaviruses: Ecology, Emergence, and Outbreaks. Microorganisms 8, 1167 (2020).
3. Levinson, R. S., Strauss, J. H. & Strauss, E. G. Complete sequence of the genomic RNA of O'nyong-nyong virus and its use in the construction of alphavirus phylogenetic trees. Virology 175, 110–123 (1990).
4. Strauss, J. H. & Strauss, E. G. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev 58, 491–562 (1994).
5. Weaver, S. C. *et al.* A comparison of the nucleotide sequences of eastern and western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses. Virology 197, 375–390 (1993).
6. Gould, E. A. *et al.* Understanding the alphaviruses: Recent research on important emerging pathogens and progress towards their control. Antiviral Research 87, 111–124 (2010).
7. Aréchiga-Ceballos, N. & Aguilar-Setién, A. Alphaviral equine encephalomyelitis (Eastern, Western and Venezuelan). Rev Sci Tech 34, 491–501 (2015).
8. Avilés, G., Sabattini, M. S. & Mitchell, C. J. Transmission of Western Equine Encephalomyelitis Virus by Argentine Aedes albifasciatus (Diptera: Culicidae). Journal of Medical Entomology 29, 850–853 (1992).
9. Lecollinet, S. *et al.* Viral Equine Encephalitis, a Growing Threat to the Horse Population in Europe? *Viruses* 12, 23 (2019).
10. Stromberg, Z. R., Fischer, W., Bradfute, S. B., Kubicek-Sutherland, J. Z. & Hraber, P. Vaccine Advances against Venezuelan, Eastern, and Western Equine Encephalitis Viruses. Vaccines 8, 273 (2020).
11. Howitt, B. F. Viruses of Equine and of St. Louis Encephalitis in Relationship to Human Infections in California, 1937–1938. Am J Public Health Nations Health 29, 1083–1097 (1939).
12. Centers for Disease Control. Western equine encephalitis–United States and Canada, 1987. MMWR Morb Mortal Wkly Rep 36, 655–659 (1987).
13. Cobos-Marín, L., Rodríguez-Monterde, A. & Valdés-Vázquez, L. M. Encefalitis equina del oeste. Vet Méx OA 6, (2019).
14. Bergren, N. A. *et al.* "Submergence" of Western equine encephalitis virus: Evidence of positive selection argues against genetic drift and fitness reductions. PLoS Pathog 16, e1008102 (2020).

15. Bergren, N. A. *et al.* Western Equine Encephalitis Virus: Evolutionary Analysis of a Declining Alphavirus Based on Complete Genome Sequences. J Virol 88, 9260–9267 (2014).
16. Robb, L. L. *et al.* Continued Evidence of Decline in the Enzootic Activity of Western Equine Encephalitis Virus in Colorado. J Med Entomol 56, 584–588 (2019).
17. Bruno-Lobo, G., Bruno-Lobo, M., Travassos, J., Pinheiro, F. & Pazin, I. Estudos sobre arbovirus III. Isolamento de um vírus sorologicamente relacionado ao subgrupo Western-Sindbis de um caso de encefalomielite eqüina ocorrido no Rio de Janeiro. Anais de Microbiologia do Rio de Janeiro vol. 9 183–195 (1961).
18. Acha, P. N. & Szyfres, B. *Zoonosis y enfermedades transmisibles comunes al hombre y a los animales*. (Organización Panamericana de la Salud, Washington, D.C., 2003).
19. Mitchell, C. J. *et al.* Arbovirus Isolations from Mosquitoes Collected during and after the 1982–1983 Epizootic of Western Equine Encephalitis in Argentina. The American Journal of Tropical Medicine and Hygiene 36, 107–113 (1987).
20. Somma, M. *et al.* Arbovirus en el Uruguay. Archivos de Pediatría de Uruguay vol. 41 359–363 (1970).
21. Delfraro, A. *et al.* Fatal human case of Western equine encephalitis, Uruguay. Emerg Infect Dis 17, 952–954 (2011).
22. González Ayala, S. E., Morales, M. A. & Enría, D. A. Reemergencia de la encefalitis equina del oeste (WEEV) en la Argentina: una revisión de aspectos epidemiológicos, virológicos y clínicos de relevancia. *ASEI* (2024) doi:10.52226/revista.v32i114.315.
23. Silva, M. L. C. R. *et al.* Outbreaks of Eastern equine encephalitis in northeastern Brazil. J VET Diagn Invest 23, 570–575 (2011).
24. Araújo, F. A. A. *et al.* Anticorpos antialfavírus detectados em equinos durante diferentes epizootias de encefalite equina, Paraíba, 2009. *RBCV* 19, 80–85 (2012).
25. Burgueño, A. *et al.* Genomic Characterization and Seroprevalence Studies on Alphaviruses in Uruguay. Am J Trop Med Hyg 98, 1811–1818 (2018).
26. Diniz, D. D. M. *et al.* Seroprevalence and risk factors associated with seropositivity for equine encephalomyelitis virus in horses in Rio Grande do Norte, Brazil. Arq. Inst. Biol. 89, e00462020 (2022).
27. PAHO/WHO. *Laboratory Guidelines for the Detection and Diagnosis of Western Equine Encephalitis Virus Human Infection*. www.paho.org (2024).
28. Netolitzky, D. J. *et al.* Complete genomic RNA sequence of western equine encephalitis virus and expression of the structural genes. J Gen Virol 81, 151–159 (2000).
29. Fallah, H. M. & Kramer, L. D. Genetic variation among isolates of western equine encephalomyelitis virus from California. The American Journal of Tropical Medicine and Hygiene 60, 708–713 (1999).
30. Campos, A. S. *et al.* Molecular epidemiology of Western equine encephalitis virus in Brazil. Preprint at https://doi.org/10.1101/2024.04.15.24305848 (2024).
31. Argimón, S. *et al.* Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microbial Genomics 2, (2016).
32. Sánchez-Seco, M. P., Rosario, D., Quiroz, E., Guzmán, G. & Tenorio, A. A generic nested-RT-PCR followed by sequencing for detection and identification of members of the alphavirus genus. Journal of Virological Methods 95, 153–161 (2001).
33. Lambert, A. J., Martin, D. A. & Lanciotti, R. S. Detection of North American eastern and western equine encephalitis viruses by nucleic acid amplification assays. J Clin Microbiol 41, 379–385 (2003).
34. OIE. Equine encephalomyelitis (Eastern, Western and Venezuelan). in *OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* vol. Chapter 3.6.5. 16 (World Organization for Animal Health, Paris, 2021).
35. Brault, A. C., Fang, Y. & Reisen, W. K. Multiplex qRT-PCR for the Detection of Western Equine Encephalomyelitis, St. Louis Encephalitis, and West Nile Viral RNA in Mosquito Pools (Diptera: Culicidae). Journal of Medical Entomology 52, 491–499 (2015).

36. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30, 772–780 (2013).
37. Martin, D. P. *et al.* RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol 7, veaa087 (2021).
38. Price, M. N. *et al.* FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Molecular Biology and Evolution 26, 1641–1650 (2009).
39. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. Virus Evolution 4, (2018).
40. Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34, 4121–4123 (2018).
41. HSI. Horsemeat production in South America. (2014).
42. Djordjevic, S. P. *et al.* Genomic surveillance for antimicrobial resistance - a One Health perspective. Nature reviews. Genetics 25, 142–157 (2024).
43. Kalinich, C. C. *et al.* Real-time public health communication of local SARS-CoV-2 genomic epidemiology. PLoS Biology 18, (2020).
44. Oude Munnink, B. B. *et al.* The next phase of SARS-CoV-2 surveillance: real-time molecular epidemiology. Nature medicine 27, 1518–1524 (2021).
45. Gu, W., Miller, S. & Chiu, C. Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. Annual Review of Pathology: Mechanisms of Disease 14, 319–338 (2019).
46. Tinning, M. & Genome, A. Next-Generation Sequencing: an overview of technologies and applications. ARC Centre of Excellence in Bioinformatics (2013).
47. Suminda, G. G. D. *et al.* High-throughput sequencing technologies in the detection of livestock pathogens, diagnosis, and zoonotic surveillance. Comput Struct Biotechnol J 20, 5378–5392 (2022).
48. Hall, R. J. *et al.* Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. Journal of Virological Methods 195, 194–204 (2014).
49. Kozarewa, I., Armisen, J., Gardner, A. F., Slatko, B. E. & Hendrickson, C. L. Overview of Target Enrichment Strategies. *Curr Protoc Mol Biol* 112, 7.21.1–7.21.23 (2015).
50. Burke, C. W. *et al.* Complete Coding Sequence of Western Equine Encephalitis Virus Strain Fleming, Isolated from a Human Case. Microbiol Resour Announc 9, e01223-19 (2020).
51. Weaver, S. C. *et al.* Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. J Virol 71, 613–623 (1997).
52. Mossel, E. C. *et al.* Molecular Determinants of Mouse Neurovirulence and Mosquito Infection for Western Equine Encephalitis Virus. PLoS ONE 8, e60427 (2013).
53. Grady, G. F. *et al.* Eastern equine encephalitis in Massachusetts, 1957–1976. A prospective study centered upon analyses of mosquitoes. Am J Epidemiol 107, 170–178 (1978).
54. Mermel, L. A. Association of Human Eastern Equine Encephalitis With Precipitation Levels in Massachusetts. JAMA Netw Open 3, e1920261 (2020).
55. *The Arboviruses. Volume 1*. (CRC Press, Taylor & Francis Group, Boca Raton London New York, 2019).

Tables

Table 1. Samples sequenced in this study by the multiplex PCR-NGS assay.

Sample	Sampling Date*[1]	Location	qPCR (Ct)*[2]	Accession number*[3]
51	11/30/2023	Durazno	39.0/28.1	PP747340
70*	12/01/2023	Paysandú	34.0/27.0	PP747341
95	12/04/2023	San José	40.0/26.0	PP747342
107	12/06/2023	Río Negro	37.3/29.8	PP747343
158*	12/08/2023	San José	35.6/25.8	PP747344
184	12/11/2023	Paysandú	37.9/31.0	PP747345
191	12/12/2023	Río Negro	38.7/27.3	PP747346
197	12/13/2023	Río Negro	39.5/28.4	PP747347
198	12/13/2023	San José	34.9/24.9	PP747348
199	12/13/2023	Artigas	40.6/31.1	PP747349
227	12/27/2023	Lavalleja	41.1/28.1	PP747350
228	12/27/2023	Lavalleja	38.7/28.5	PP747351
230	12/27/2023	Treinta y Tres	38.5/27.2	PP747352
248	01/15/2024	San José	37.0/28.5	PP747353
250	01/16/2024	Río Negro	31/22.3	PP747354

* Sequences from samples 70, 158, and 255 were previously available in GenBank (PP620641, PP620642, PP620646).

*[1] All samples are from equine hosts and obtained from brain tissue or cerebral cortex, except for sample 186, which is from cerebrospinal fluid.

*[2] The Ct values obtained using the Lambert et al. [33] and Brault et al. [35] methodologies are indicated and separated by a bar.

*[3] Accession numbers of complete coding sequences.

Table 2. Primers for amplifying the complete WEEV genome (PCR enrichment). The position of the primer is relative to the MN477208 WEEV strain.

Amplicon	Pool	Forward sequence 5′–3′	Genome position	Reverse sequence 5′–3′	Genome position
1	1	TAGAGGAACCTACCCTACAAAC	14–35	GTCAGTTGCCAGAATCCCT	1089–1071
2	2	GTTTCTTTTGCTGTGTGTACG	1018–1038	TTCTTCGCACTCACAACCAG	2279–2260
3	1	ATYGGAGTCTATGGAGTGCC	2191–2210	ATGGACTCTACTTTCTTCCC	3599–3580
4	2	TATCTAAGATGAAGGGCAAATCTG	3524–3547	CTGTTCTTTCTTCGCAGAGC	4878–4859
5	1	GCCCAGTAGAGGAGTCAGAG	4766–4785	GAAAGGTATCCCARTACTCATCATTG	6352–6327
6	2	GTTACCCAAATGCGAGAAYTAC	6253–6274	TTTTAGGCTGAGTAGGTTTTGG	7743–7722
7	1	AAACAACGAKCACCTAATCC	7659–7678	CCTGTAATCCATTCTGCTGTTG	9394–9373
8	2	TCTTATTAGGCACACAGACCAC	9182–9203	CGTATGCCTCACTCAGTTGT	10212–10193
9	1	CGTTTTTGGCGGTGTGTA	10127–10144	GAGTCTCATAAAAGTGATATAGTGTTG	11403–11377

Figures

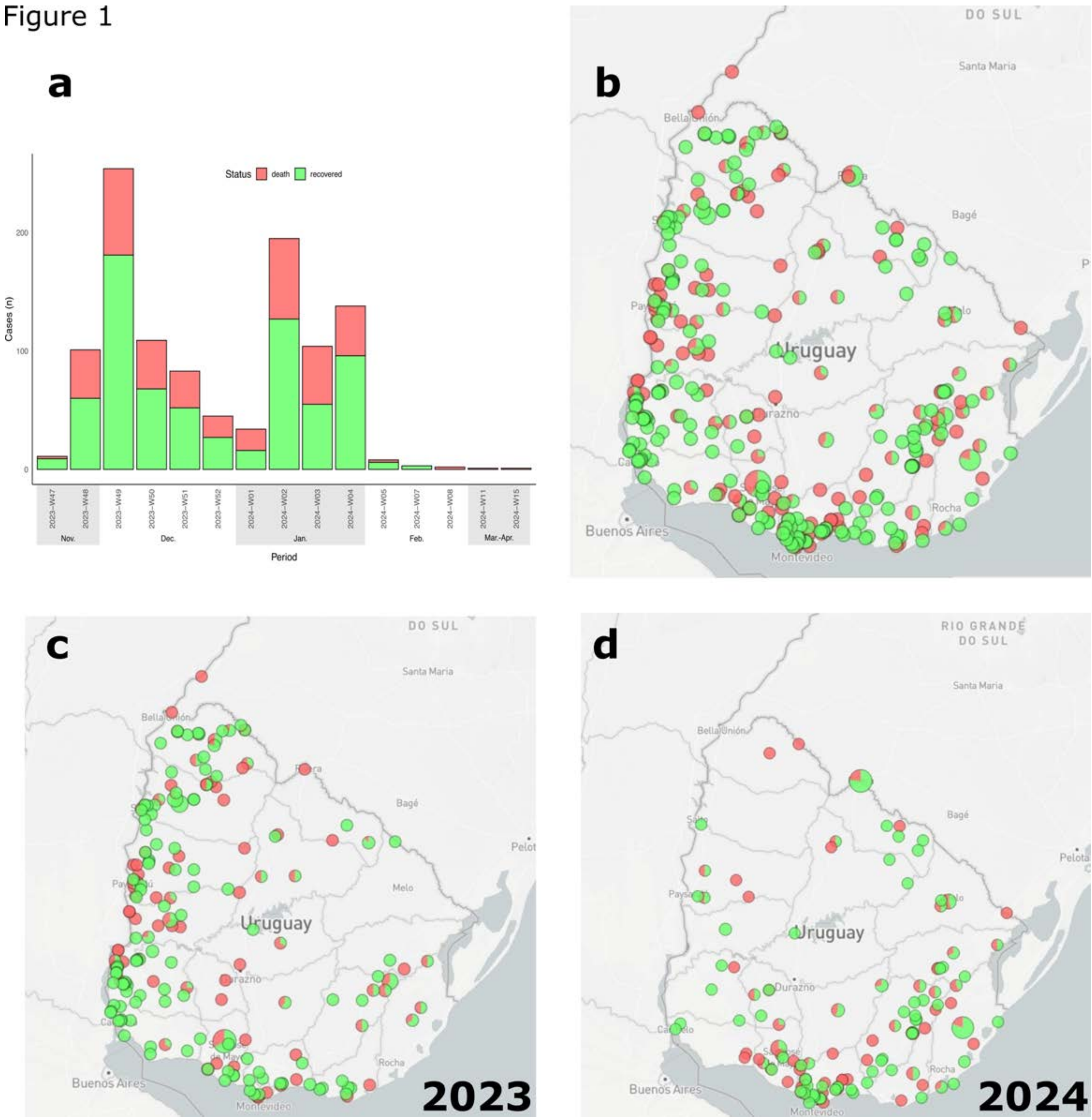

Figure 1

Distribution of WEEV positive cases in Uruguayan equines. **a** – Histogram plot illustrating the distribution of WEEV-positive cases in Uruguay during 2023–2024. **b** – Geographic distribution of WEEV-positive cases during 2023–2024. **c**– WEEV-positive cases during 2023 are located mainly in regions bordering Argentina. **d** – WEEV-positive cases during 2024 are primarily located in central and eastern regions.

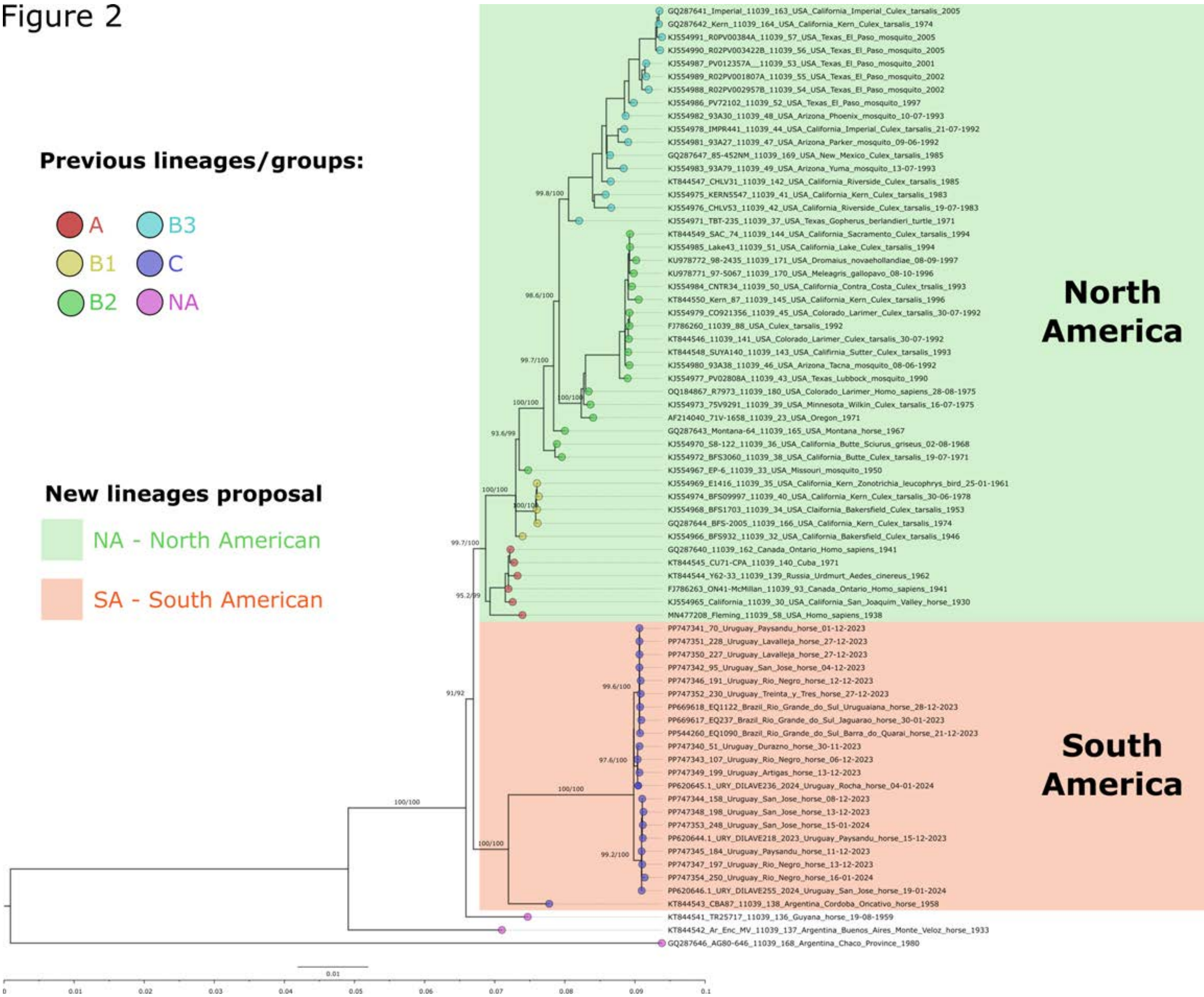

Figure 2

Phylogenetic analysis of complete genomes. WEEV genomes, including the 15 genomes from the current 2023-24 outbreak. Tip color follows groups and lineages previously reported in the literature. The new proposed North American (green) and South American (orange) lineages are depicted. Branch supports depicted above each branch are aLRT and ultrafast bootstrap.

Figure 3

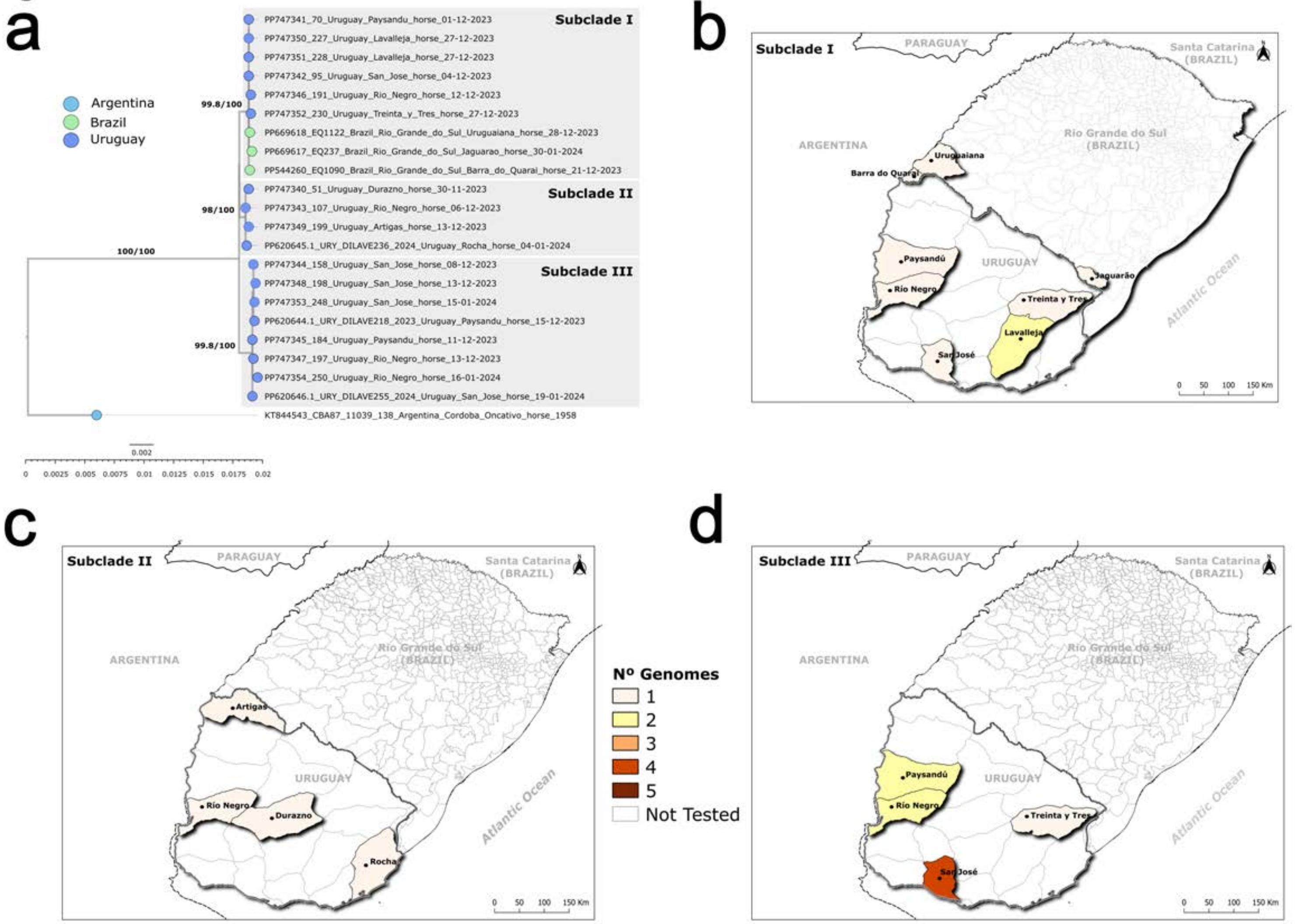

Figure 3

Geographic distribution of the I, II, and III subclades of the 2023-24 outbreak clade. **a** – Zoom of the clade, including an early Argentina strain (1958) and the 23 strains from the 2023-24 outbreak. Tip colors are countries of origin. Branch supports are aLRT and ultrafast bootstrap values. **b** – Geographic distribution of subclade I samples from Uruguay and Brazil. **c** – Geographic distribution of Uruguayan subclade II samples. **d** – Geographic distribution of Uruguayan subclade III samples. Heat map colors indicate the number of genomes per department.

Figure 4

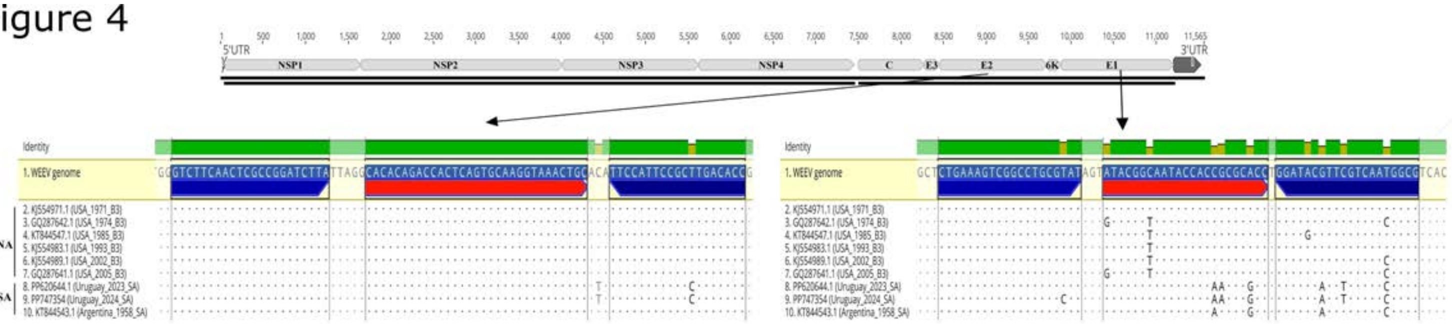

Figure 4

Mismatches in probes and primers of two qPCR assays used for WEEV diagnosis. **Top**: representation of the WEEV genome showing qPCR target regions. **Right bottom**: Alignment of the probe and primer regions of the official qPCR diagnostic protocol targeting the E1 coding region [33]. Several SNPs may affect the binding effectivity of primers and probes to South American strains. **Left bottom**: Alignment of the probe and primer regions of an alternative qPCR assay targeting the E2 coding region [35]. These regions are more conserved among WEEV strains.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryDataset1.fasta
- SupplementaryFigure1.pdf