

RESEARCH ARTICLE

# Exploring Genomic, Geographic and Virulence Interactions among Epidemic and Non-Epidemic St. Louis Encephalitis Virus (Flavivirus) Strains

Luis A. Diaz<sup>1,2</sup>\*, Sandra E. Goñi<sup>3</sup>, Javier A. Iserte<sup>4</sup>, Agustín I. Quaglia<sup>1</sup>, Amber Singh<sup>5</sup>, Christopher H. Logue<sup>6</sup>, Ann M. Powers<sup>5</sup>, Marta S. Contigiani<sup>1</sup>



**1** Laboratorio de Arbovirus, Instituto de Virología “Dr. J. M. Vanella”, Facultad de Ciencias Médicas, Universidad Nacional de Córdoba, Córdoba, Argentina, **2** Instituto de Investigaciones Biológicas y Tecnológicas, CONICET—Universidad Nacional de Córdoba, Córdoba, Argentina, **3** Área de Virosis Emergentes y Zoonóticas, Laboratorio de Ingeniería Genética y Biología Celular y Molecular, Instituto de Microbiología Básica y Aplicada, Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Buenos Aires, Argentina, **4** Laboratorio de Bioinformática Estructural, Fundación Instituto Leloir, CABA, Buenos Aires, Argentina, **5** Centers for Diseases Control and Prevention, Fort Collins, Colorado, United States of America, **6** Novel and Dangerous Pathogens Training, Public Health England, Porton Down, United Kingdom

OPEN ACCESS

**Citation:** Diaz LA, Goñi SE, Iserte JA, Quaglia AI, Singh A, Logue CH, et al. (2015) Exploring Genomic, Geographic and Virulence Interactions among Epidemic and Non-Epidemic St. Louis Encephalitis Virus (Flavivirus) Strains. PLoS ONE 10(8): e0136316. doi:10.1371/journal.pone.0136316

**Editor:** Tian Wang, University of Texas Medical Branch, UNITED STATES

**Received:** March 11, 2015

**Accepted:** August 2, 2015

**Published:** August 27, 2015

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This project was supported by Agencia Nacional de Promoción Científica y Tecnológica, Fondo Nacional de Ciencia y Tecnología (ANPCYT-FONCYT) Grant number 2010/38060 and 2013/1779 ([www.agencia.mincyt.gob.ar](http://www.agencia.mincyt.gob.ar)), Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) ([www.conicet.gov.ar](http://www.conicet.gov.ar)), Secretaría de Ciencia y Técnica -Universidad Nacional de Córdoba (SECYT-UNC) (<http://www.unc.edu.ar/investigacion/cienciaytecnologia/novedades-informacion-cyt>),

\* These authors contributed equally to this work.

\* [adrian.diaz@conicet.gov.ar](mailto:adrian.diaz@conicet.gov.ar)

## Abstract

*St. Louis encephalitis virus* (SLEV) is a re-emerging arbovirus in South America. In 2005, an encephalitis outbreak caused by SLEV was reported in Argentina. The reason for the outbreak remains unknown, but may have been related to virological factors, changes in vectors populations, avian amplifying hosts, and/or environmental conditions. The main goal of this study was to characterize the complete genome of epidemic and non-epidemic SLEV strains from Argentina. Seventeen amino acid changes were detected; ten were non-conservative and located in proteins E, NS1, NS3 and NS5. Phylogenetic analysis showed two major clades based on geography: the North America and northern Central America (NANCA) clade and the South America and southern Central America (SAsCA) clade. Interestingly, the presence of SAsCA genotype V SLEV strains in the NANCA clade was reported in California, Florida and Texas, overlapping with known bird migration flyways. This work represents the first step in understanding the molecular mechanisms underlying virulence and biological variation among SLEV strains.

## Introduction

*Saint Louis encephalitis virus* (SLEV) (*Flavivirus*, *Flaviviridae*), a reemerging arbovirus in South America, has been responsible for febrile disease and human encephalitis cases in Argentina in 2002, 2005 and 2010, and in Brazil in 2004 and 2006[1–4]. SLEV first reemerged in the

International Society for Infectious Diseases (<http://www.isid.org/>), UNESCO-IUMS-SGM (United Nations Educational, Scientific and Cultural Organization, International Union Microbiology Societies-Society of General Microbiology) (<http://www.iums.org/>), Bunge y Born Foundation (FBBEI8/10) (<http://www.fundacionbyb.org>), and the US Centers for Disease Control and Prevention ([www.cdc.gov](http://www.cdc.gov)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

central region of Argentina (Córdoba and Santa Fe Provinces) in 2002, when two cases of encephalitis and three fever cases were reported [5]. The first human epidemic of SLEV outside of North America was reported in 2005, when 47 laboratory-confirmed clinical cases of SLE, including nine fatalities, were reported in Córdoba Province (Argentina) [1].

The geographic distribution of SLEV encompasses tropical, sub-tropical and much of the temperate-tropical zones of the Western Hemisphere which includes most of the populated land masses of North and South America [6]. In the USA, SLEV is naturally maintained by mosquito-to-bird transmission involving several *Culex* (*Cx.*) spp. Mosquitoes and a variety of bird species, including the house sparrow (*Passer domesticus*), house finch (*Haemorhous mexicanus*) and mourning doves (*Zenaidamacrourea*) [6].

Most phylogenetic relationships among SLEV strains are based on the complete envelope gene sequence [7]. A total of 8 genotypes are presently recognized (I-VIII) [8]. Genotypes I and II are prevalent in the USA, while the others are traditionally found only in Central and South America [8]. During the 2005 SLEV outbreak in Argentina two genotype III SLEV strains were isolated from *Cx. quinquefasciatus* mosquitoes [9]. Moreover, during the reemergence of SLEV in Buenos Aires province, molecular evidence suggested an association with same genotype III [10]. Although not completely understood, evidence supports a multicausal event for SLEV reemergence. Introduction of a novel more virulent strain, an increase of *Cx. quinquefasciatus* and *Cx. interfor* mosquito vector abundance, and highly susceptible avian host availability are potential explanations [11,12]. A retrospective study shows that no previous activity was documented for this genotype in Cordoba city prior to the 2005 outbreak [11]. Moreover, some biological differences among epidemic (CbaAr-4005, Ep) and non-epidemic (79V-2533, NEp) viral strains were detected. For example, house sparrows inoculated with SLEV CbaAr-4005 Ep developed higher and long lasting viremias [12].

In flaviviruses, several mutations have been associated with phenotype alteration, including those linked to virulence. Unfortunately, the genotypic evidence associated with a virulent SLEV phenotype is lacking. The main objective of this study was to characterize the complete genome of SLEV strains from Argentina, to identify molecular differences among Ep and NEp SLEV strains, and to associate these differences with ecologic, epidemiologic and geographic trends.

## Materials and Methods

### Viral strains

Two SLEV genotype III strains were completely sequenced (79V-2533 and CbaAr-4005). The 79V-2533 strain was isolated from *Culex* (*Culex*) spp. mosquitoes collected in Esperanza city (Santa Fe Province, Argentina) in 1978, during a non-epidemic period. The epidemic strain (CbaAr-4005) was isolated from *Cx. quinquefasciatus* mosquitoes collected during the SLEV human encephalitis outbreak of 2005. Viruses were propagated on VERO cell monolayers inoculated with 100µl of CbaAr-4005 or 79V-2533 viral strains. Virus was harvested on the 6<sup>th</sup> day post-infection (dpi) by centrifuging the supernatant after one freeze/thaw cycle.

### RNA extraction, reverse transcription and PCR amplification

Vero cell supernatant was used for RNA viral extraction employing the commercial QIAamp viral RNA MiniSpin Kit (Qiagen). For reverse transcription and PCR, two commercial kits were used, Titan One Tube RT-PCR System (Roche) and Titan One Tube RT-PCR Kit (Roche, Reaction Mix 2), following the manufacturer's instructions.

## Primer design and genome sequencing

The sequencing strategy and primer design to obtain the complete sequence for CbaAr-4005 and 79V-2533 strains were based on a consensus sequence generated from Kern217 (DQ525916.1 and NC\_007580.2) and Argentine66 (AY632544.1).

For the 5'UTR amplification and sequencing, a commercial kit (Ambion #AM1700 First Choice RLM Race) was used. For the 3'UTR, the A-Plus Poly (A) Polymerase Tailing Kit (Epicentre Biotechnologies) was used. The manufacturer instructions were applied with the following exceptions: 2.5µl RNAase inhibitor (40U/µl), and 0.5µl A-Plus Poly A (4U/µl) were added and the reaction incubated at 37°C for 10min. The Titan One Tube RT-PCR Kit (Roche) was used for genomic amplification. For sequencing of the amplified fragments the same protocol as that for the 5'UTR was used.

The partial fragments generated for each strain were analyzed and individually selected for assembly to generate a consensus genome sequence using the SeqMan II (v. 5.03) program provided within the LaserGene (DNAStar) package. The complete genome sequences for both strains were submitted to GenBank with the accession numbers FJ753286.2 and FJ753287.2.

## Multiple alignments and bioinformatics analyses

Multiple alignments of 29 analyzed sequences were generated by ClustalW [13] with the MEGA 6 program (<http://www.megasoftware.net>). Most of sequences analyzed possessed coding region sequences only, missing the 5' and 3' non coding regions as well as the last portion of NS5 gene. Therefore, for the bioinformatics analysis, all sequences were truncated in order to use the same sequence range.

The coding region was identified using Clone Manager ([http://www.scied.com/pr\\_cmpro.htm](http://www.scied.com/pr_cmpro.htm)) software. The polyprotein cleavage sites were determined following the method described by Rice and Strauss [14]. The SignalP-NN program was employed (<http://www.cbs.dtu.dk/services/>) for proteolytic sites and protease recognition signals. A comparative study among other flaviviruses was carried out based on previously published findings [15]. In the Ep and NEp comparative analyses, conservative and non conservative amino acid substitutions were identified and classified according to the Dayhoff matrix.

The relative homology (RH) profile was calculated for the nucleotide and amino acid sequences of the 29 ORF sequences analyzed. These profiles reflect the conservation grade between different regions of genome and proteins. The relative homology was calculated by a strategy based in overlapping windows, using the consensus symbols obtained in the ClustalW alignments. The media value was calculated using the addition of the total values divided for the window size and plotted (Ghiringhelli, P.D. and Iserte, J.A., unpublished). For homology calculations between protein sequences, the window size was 11 residues and arbitrary values of +1 for identical residues, +0.5 and +0.25 for conservative changes, and -1 for non identical residues were assigned. The non coding regions were analyzed by Sequence Logos tool (<http://weblogo.berkeley.edu/logo.cgi>) and the RNA secondary structure was modeled (MFold, <http://www.bioinfo.rpi.edu/~zukerm>). Briefly, the Sequence Logo is a graphic representation of a multiple alignment of sequences. After the alignments were generated, the sequences were manually curated (using *Dengue 2 virus* (DENV2) strains (NC001474.2: DENV-2/16681 and DQ181806.1: DENV-2/Th2\_0038\_74) and SLEV CbaAr-4005 and 79V-2533 strains), the 5'UTR and 3'UTR were compared, and the different signals and structures were located [16].

Recombination in SLEV polyprotein open reading frame sequences was tested for with SBP and GARD software on Datamonkey web server [17].

## Phylogenetic analyses

An alignment containing 29 SLEV strains (Table 1) and 3 sequences as the outgroup strains (*West Nile virus*, *Japanese encephalitis virus*, and *Murray Valley Encephalitis virus*) was generated using MEGA6 program (<http://www.megasoftware.net/>). The substitution model used in the phylogenetic reconstruction was GTR with a proportion of invariant sites and Gamma distributed rates. The substitution model was selected with jModelTest 2.1.5 [18]. Phylogenetic reconstruction was made by Bayesian inference using BEAST 2.0 [19]. Site model was partitioned by codon position. Substitution models in Beast were tested and compared with the model selected with jModelTest. Monte Carlo Markov chain (MCMC) length was 50 million, sampling every 1000. The Tracer program (<http://tree.bio.ed.ac.uk/software/tracer/>) was used to analyze convergence of MCMC. The tree presented is the maximum clade credibility tree summarized from MCMC data with a burn in of 10%. Phylogenetic reconstruction was also made in MEGA6 software using the Maximum Likelihood method, with 1000 replicates for bootstrap analysis.

## Geographic distribution and genetic diversity

The association between geographic and genetic distance of the 29 SLEV strains examined was explored through symmetric classic Procrustes analysis [20]. Genetic and geographic distance matrices were constructed. Geographic coordinates were obtained by Google Earth application and geodesic distances were calculated. Genetic distances were calculated using Kimura 2-parameter as the evolutionary model under MEGA 6 software [21]. Distance matrices were used as input for ordination analysis. Non Metric Multidimensional Scaling (NMDS) of each distance matrix showed the ordination constellation and potential outlier in pairwise fashion. After removal of the outliers, the genotype distance's NMDS was transposed, rotated and scaled on the geographic distance's NMDS. The selected statistics metrics were Procrustes's correlation ( $t_0$ ), sum of squared errors ( $m^2$ ), scaling and rotation. The significance of  $t_0$  was tested against of the distribution pseudo Procrustes's correlation ( $t$ ) with 10,000 permutations. All analysis were performed in R [22] with the Vegan package [20]. The workflow script to run this analysis on R is available in the supplementary data file provided (please see S1 File).

## Association of mutations with virulence and viremia

We used bibliographic data available in order to determine biological features associated with all studied viral strains including: source of isolation (mosquito, bird, mammals), viremia in birds, and virulence in mice [12,23–25]. SLEV strains were classified as high viremia (HV) or low viremia (LV), and high pathogenicity (HP) or low pathogenicity (LP) (Table 1). An amino acid change survey was carried out for each protein in order to find non conservative mutations through the ORF.

Multivariate analyses were employed to explore patterns between genomic mutations and biological characteristics. Analysis of 200 point mutations observed in a pool of 14 SLEV strains with respect to the multidimensional distribution of all descriptors (biological features) was carried out. As a consequence of the binary nature of the data (i.e: a Chi-square distance between mutation) a Correspondence Analysis was selected [26]. Once CA analyses were carried out, correlations between source of isolation (birds, mammals, mosquitoes), viremogenic capacity in birds (High and Low) and virulence in mice (High and Low) were tested. Spearman's correlation indexes and bootstrap significance were calculated using "corrgram" and "psych" packages in R. After multivariate analysis, selected mutations were tested in contingency tables for virulence and viremia using a Fisher's exact test and ODD ratio. These analyses were carried out with Infostat software package ([www.infostat.com.ar](http://www.infostat.com.ar)).

**Table 1. St. Louis encephalitis virus strains and their biological features.**

SLEV Strain	Geographic Origin	Year of isolation	Host	Rodent Pathogenicity <sup>¥</sup>	Avian Viremia <sup>‡</sup>	Accession Number
79V_2533	Santa Fe, Argentina	1979	<i>Culex</i> spp. (Mosquito)	n.d.	n.d.	FJ753287.1
CbaAr_4005	Córdoba, Argentina	2005	<i>Culex quinquefasciatus</i> (Mosquito)	High <sup>D</sup>	High <sup>C</sup>	FJ753286.1
BeAr23379	Belem, Brasil	1960	<i>Sabethes belisarioi</i> (Mosquito)	Low <sup>A</sup>	Low <sup>B</sup>	EF158048.1
Kern217	California, USA	1989	<i>Culex tarsalis</i> (Mosquito)	n.d.	n.d.	DQ525916.1
904.3	Kentucky, USA	1955	<i>Colaptes auratus</i> (Bird)	High <sup>A</sup>	High <sup>B</sup>	EF158049.1
MSI_7	Mississippi, USA	1975	<i>Passer domesticus</i> (Bird)	High <sup>A</sup>	High <sup>B</sup>	DQ359217.1
GMO94	Guatemala	1969	<i>Culex nigripalpus</i> (Mosquito)	High <sup>A</sup>	High <sup>B</sup>	EF158051.1
V2380_42	Texas, USA	2001	<i>Culex quinquefasciatus</i> (Mosquito)	n.d.	n.d.	EF158052.1
BeAn246262	Belem, Brazil	1973	<i>Didelphis marsupialis</i> (Mammal)	High <sup>A</sup>	High <sup>B</sup>	EF158053.1
75D90	Perú	1975	Mosquito	High <sup>A</sup>	Low <sup>B</sup>	EF158054.1
TBH28	Florida, USA	1962	Human	n.d.	n.d.	EF158055.1
TRVL_9464	Trinidad	1955	<i>Psorophora ferox</i> (Mosquito)	Low <sup>A</sup>	Low <sup>B</sup>	EF158056.1
78_A_28	Guatemala	1978	Mosquito	n.d.	n.d.	EF158057.1
65V310	México	1965	<i>Butorides virescens</i> (Bird)	Low <sup>A</sup>	High <sup>B</sup>	EF158059.1
GML903797	Panamá	1983	Chickens (Bird)	n.d.	n.d.	EF158060.1
69M_1143	Florida, USA	1969	<i>Procyon lotor</i> (Mammal)	Low <sup>A</sup>	Low <sup>B</sup>	EF158061.2
FL79_411	Florida, USA	1979	<i>Culex nigripalpus</i> (Mosquito)	n.d.	n.d.	EF158062.1
CorAn9124	Córdoba, Argentina	1966	<i>Calomys musculinus</i> (Mammal)	Low <sup>A</sup>	Low <sup>B</sup>	EF158063.1
GML902612	Panamá	1973	<i>Haemagogus equines</i> (Mosquito)	n.d.	n.d.	EF158064.1
TNM4_711K	Tennessee, USA	1974	<i>Culex pipiens</i> (Mosquito)	High <sup>A</sup>	High <sup>B</sup>	EF158065.1
GHA 3	Florida, USA	1962	Human	High <sup>A</sup>	n.d.	EF158066.1
BeAn247377	Belem, Brasil	1973	<i>Hylophilax poecilonata</i> (Bird)	n.d.	n.d.	EF158067.1
CorAn9275	Córdoba, Argentina	1967	<i>Mus musculus</i> (Mammal)	Low <sup>A, D</sup>	Low	EF158068.1
72V_4749	Colorado, USA	1972	<i>Culex tarsalis</i> (Mosquito)	High <sup>A</sup>	Low <sup>B</sup>	EF158069.1
Parton	Missouri, USA	1933	Human	n.d.	n.d.	EF158070.1
Hubbard	Missouri, USA	1937	Human	n.d.	n.d.	EU566860.1
IMP115	California, USA	2003	<i>Culex tarsalis</i> (Mosquito)	-	-	JF460774.1
Palenque-A770	Palenque, México	2008	<i>Culex nigripalpus</i> (Mosquito)	-	-	JQ957869.1
Palenque-C475	Palenque, México	2008	<i>Culex nigripalpus</i> (Mosquito)	-	-	JQ957868.1

<sup>¥</sup> Pathogenicity feature observed in swiss albino mice or monkeys according to cited literature.

<sup>‡</sup>Viremia observed in subcutaneously inoculated birds (house sparrows/chicks).

<sup>A</sup>[23];

<sup>B</sup>[24];

<sup>C</sup>[12];

<sup>D</sup>[25].

n.d.: No data available.

doi:10.1371/journal.pone.0136316.t001



## Results

### Epidemic vs Non-Epidemic strains

Full length, consensus genome sequences from both SLEV strains examined (CbaAr-4005 and 79V-2533) were obtained using traditional RT-PCR and capillary sequencing. Each full length genome consisted of a total of 10936 nt. The ORF is composed of 3429 amino acids (10287 nt) while the 5' and 3' untranslated regions (5'UTR and 3'UTR) consisted of 98 nt and 548 nt, respectively. When comparing the Ep vs NEp SLEV strains, a total of 133 changes were observed at the nucleotide level resulting in 17 amino acid substitutions (13% of the total amino acids). Of these 17, 10 non-conservative changes were found (59%) (Fig 1). Only four changes in the structural proteins were identified, with only one being non conservative ( $E_{88}$ ) (Table 1). The remaining amino acid substitutions were identified in nonstructural proteins. Of these changes, 9 were considered non conservative: NS1<sub>70</sub>, NS1<sub>112</sub>, NS1<sub>290</sub>, NS2A<sub>82</sub>, NS3<sub>92</sub>, NS4B<sub>139</sub>, NS5<sub>178</sub>, NS5<sub>189</sub>, NS5<sub>507</sub> (Table 2).

The relative homology profile carried out with all 27 SLEV strains analyzed indicated that amino acid mutations occurred along the entire genome (without inclusion of the Palenque strains). Hypervariable regions ( $HR \leq 0.6$ ) were detected only in NS1 and NS4B (Fig 1), while variable regions were present in the rest of the proteins. Most of the compared sequences showed indices  $\geq 0.8$ , with few regions between 0.6 and 0.8, indicating high conservation among SLEV strains. When Palenque strains were included the appearance of some areas with more variability in NS3 and NS5 proteins were observed.

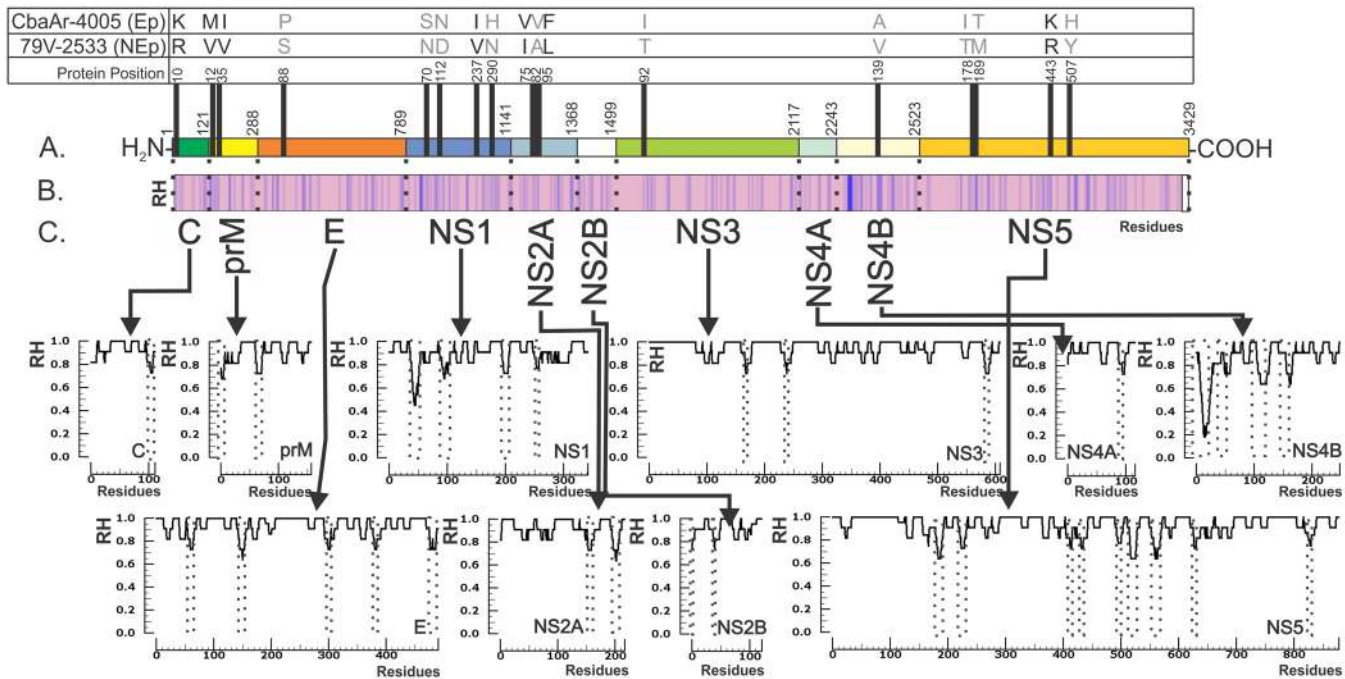
### Protease cleavage sites

The identified cleavage sites remained preserved in most of the SLEV strains analyzed. Exceptions include the BeAn-247377 strain where the standard SWPAS site between NS2A and NS2B was found to be GWPAS, and LALGM between NS3 and NS4A changed to AAGKR in strain I MP115. Differences in amino acids at positions 2, 4, and 5 relative to the cleavage site were also detected. Less conservation can be found in the C protein of the Palenque strains where the conserved motif **PSKKR/GGTRS** changed to **PSKKK/GGSGS**. The protease involved in this cleavage is of viral origin, suggesting this change may be related with changes in the catalytic site of the enzyme by modifying its recognition site. The cleavage sites located between NS4A/2K remain conserved among the flaviviruses analyzed (ROCV, WNV, JEV, MVEV); while most other sites exhibited some degree of variation at amino acids 1 and 2 (Table 3).

### Secondary structure and signals in non-coding regions

Only a few changes were detected between Ep and NEp strains in the non-coding sequences. A total of 26 transitions and 3 transversions (1 for the 5'UTR and 2 for the 3'UTR) were identified. For SLEV, only three strains have been sequenced in these regions: Kern217, Hubbard, and Imperial Valley. When comparing available sequences with those reported here (CbaAr-4005 and 79V-2533), only a few changes were observed; the changes are mainly transitions as only 11.5% of the changes were transversions and only one INDEL was present.

The predicted structure was made using MFold with the sequence data from strain 79V-25533; the results are shown in Fig 2. For both Ep and NEp strains, the same structures were obtained. For the 5'UTR and 3'UTR ends, both stem loop structures and pseudoknots were identified. For the 5'UTR, there are three stem loop structures (SLA, SLB, and capsid hairpin-CHP-), while for the 3'UTR, one large stem loop (SL) and one short stem loop (SSL) were



**Fig 1. Amino acid differences between epidemic (Ep) and non-epidemic (NEp) SLEV strains and the polyprotein conservation profile.** The horizontal bar represents the SLEV polyprotein. (A). Black vertical lines represent residues (in one letter code) that differed between strains studied (CbaAr-4005 (Ep) and 79V-2533 (NEp)). In grey, the non-conservative changes and their respective individual protein positions are highlighted. (B). Polyprotein homology profile for all studied SLEV strains. The light pink areas show the more conserved regions while in the less conserved zones are in blue. (C). Individual protein homology profile for all studied SLEV strains. The Y axis shows the Relative Homology (RH) between 0 and 1 while the X axis shows the residue position. In each protein profile, the zones that show an RH value lower than 0.8 are boxed.

doi:10.1371/journal.pone.0136316.g001

**Table 2. Conservative and non-conservative amino acids mutations detected between SLEV strains CbaAr-4005 and 79V-2533.**

Protein <sub>position</sub>	Viral strain		Dayhoff classification criteria
	CbaAr-4005(Epidemic)	79V-2533(Non-Epidemic)	
C <sub>10</sub>	K	R	Conservative
prM <sub>12</sub>	M	V	Conservative
prM <sub>35</sub>	I	V	Conservative
E <sub>88</sub>	P	S	Non-conservative
NS1 <sub>70</sub>	S <sub>Ep</sub>	N <sub>NEp</sub>	Non-conservative
NS1 <sub>112</sub>	N <sub>Ep</sub>	D <sub>NEp</sub>	Non-conservative
NS1 <sub>237</sub>	I <sub>Ep</sub>	V <sub>NEp</sub>	Conservative
NS1 <sub>290</sub>	H <sub>Ep</sub>	N <sub>NEp</sub>	Non-conservative
NS2A <sub>75</sub>	V <sub>Ep</sub>	I <sub>NEp</sub>	Conservative
NS2A <sub>82</sub>	V <sub>Ep</sub>	A <sub>NEp</sub>	Non-conservative
NS2A <sub>95</sub>	F <sub>Ep</sub>	L <sub>NEp</sub>	Conservative
NS3 <sub>92</sub>	I <sub>Ep</sub>	T <sub>NEp</sub>	Non-conservative
NS4B <sub>139</sub>	A <sub>Ep</sub>	V <sub>NEp</sub>	Non-conservative
NS5 <sub>178</sub>	A <sub>Ep</sub>	V <sub>NEp</sub>	Non-conservative
NS5 <sub>189</sub>	I <sub>Ep</sub>	T <sub>NEp</sub>	Non-conservative
NS5 <sub>443</sub>	K <sub>Ep</sub>	R <sub>NEp</sub>	Conservative
NS5 <sub>507</sub>	H <sub>Ep</sub>	Y <sub>NEp</sub>	Non-conservative

doi:10.1371/journal.pone.0136316.t002

**Table 3. Protease cleavage sites identified in SLEV sequences analyzed.**

Cleavage Site	Protease	SLEV	Changes (Strains)
<b>VirC/AnchC</b>	viral serine protease	<b>PSKKR/GGTRS</b>	<b>PSKKR/GGTGS</b> (ImperialValley, 79V-2533, CbaAr-4005, CorAn9124, CorAn9275, BeAn247377, GML902612, GML903797, TRVL9464, 75D90, BeAn246262, BeAr23379) <b>PSKKR/GGTRP</b> (GMO94)
<b>AnchC/prM</b>	host signalase	<b>GLASS/LQLST</b>	<b>GLASS/LQLSN</b> (69M1143)
<b>Pr/M</b>	furin	<b>RRSRR/SISVQ</b>	<b>KRSRR/SISVQ</b> (TRVL9464)
<b>M/E</b>	host signalase	<b>APAYS/FNCLG</b>	-
<b>E/NS1</b>	host signalase	<b>TSVQA/DSGCA</b>	-
<b>NS1/NS2A</b>	Unknown	<b>SRVTA/GVAGG</b>	<b>SRVTA/GIAGG</b> (69M1143) <b>SRVTA/GVGG</b> (Parton)
<b>NS2A/NS2B</b>	viral serine protease	<b>PNGKR/SWPAS</b>	<b>PNGKR/GWPAS</b> (BeAn247377) <b>PNGKR/SWPAG</b> (72V-4749)
<b>NS2B/NS3</b>	viral serine protease	<b>KHSKR/GGALW</b>	-
<b>NS3/NS4A</b>	viral serine protease	<b>AAGKR/SALGM</b>	<b>AAGKR/LALGM</b> (Imperial Valley)
<b>NS4A/2K</b>	viral serine protease	<b>EPEKQR/SQTDN</b>	-
<b>2K/NS4B</b>	host signalase	<b>GVVAA/NEMGL</b>	<b>GVVAA/NEMDL</b> (Hubbard)
<b>NS4B/NS5</b>	viral serine protease	<b>PKGKR/GGKKG</b>	-

doi:10.1371/journal.pone.0136316.t003

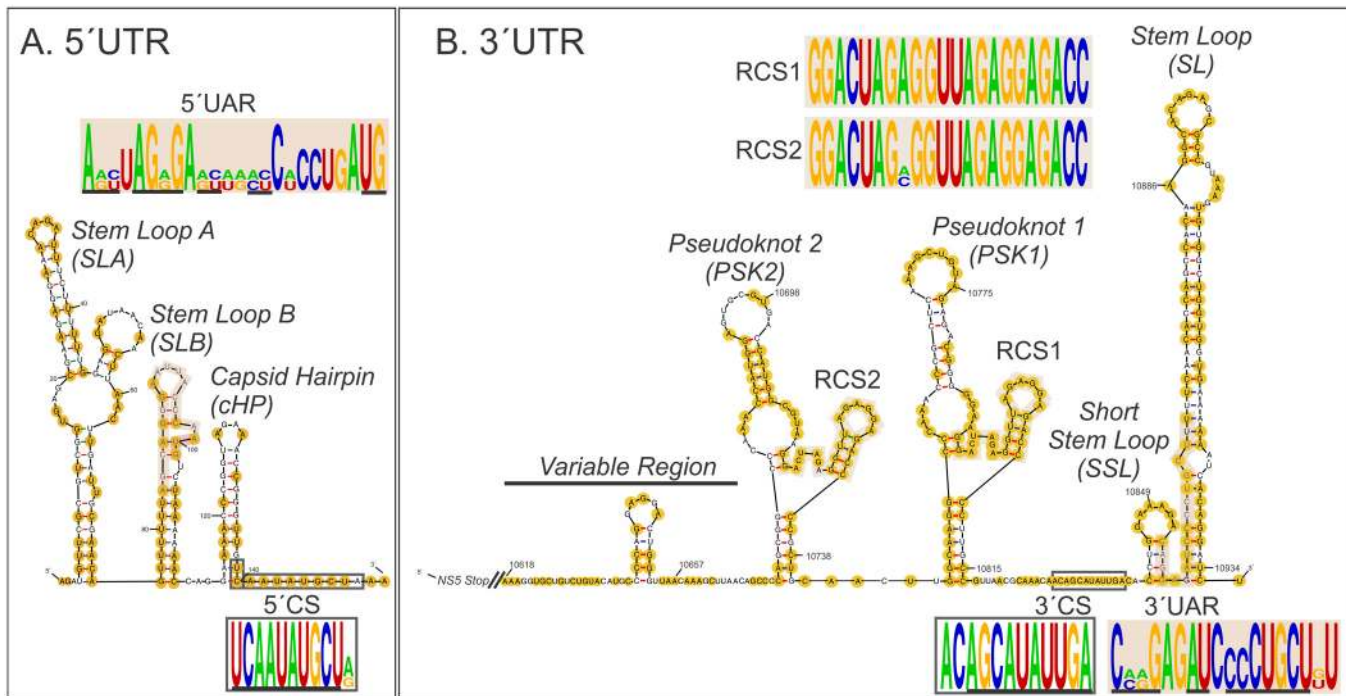
found close to the 3' end followed by two pseudoknot structures (PK1 and PK2) and a third stem loop.

Fig 2 shows the conserved structures identified in non-coding regions for the 2 SLEV strains sequenced and 2 sequences of DENV2 (NC001474.2: DENV-2/16681 and DQ181806.1: DENV-2/Th2\_0038\_74). The 5' upstream UAG region (5'UAR) and 5' conserved sequence (5'CS) in the 5'UTR (Fig 2A), and the 3' upstream UAG region (3'UAR), the 3' conserved sequence (3'CS) and repeated conserved sequence 1 and 2 (RCS1 and RCS2) in the 3'UTR (Fig 2B) were highlighted. The Sequence Logos from alignment of the respective portions are shown.

### Pathogenicity and Viremia phenotypes

A total of 200 non-conservative mutations were mapped across 14 SLEV strains. Their association with biological features was investigated. According to the CA, the first two axes account for 84.8% of the total variation observed in the data set (Fig 3). High correlations between HV-HP and LV-LP mutations were observed. This pattern was confirmed through the Spearman's correlation test, where strains that developed high viremia in birds were also found to be highly pathogenic in mice (see Fig 4 for Spearman's correlation index values). Interestingly positive and significant correlations were detected among strains isolated from Mml/LP and Mml/LV. Strains isolated from mosquitoes showed significant correlations with those HV and HP strains (Fig 4). However, no statistically significant differences were observed between





**Fig 2. Predicted secondary structures for 5'UTR and 3'UTR of a NEp SLEV strain.** For both ends, the conserved sequences (grey shadowed and boxed) and structures being comparison with DENV2 are shown [16]. The conserved residues in the alignment are shadowed. (A). 5'UTR scheme, where structures of Stem Loop A (SLA), Stem Loop B (SLB), capsid hairpin (cHP), and the 5' upstream AUG region (5'UAR), and 5' conserved sequence (5'CS) are shown. (B). 3'UTR scheme, showing the Stem Loop (SL), Short Stem Loop (SSL), the Pseudoknot 1 and 2 (PSK1 and PSK2), Variable Region, the 3' upstream AUG region (3'UAR) sequences, the 3' conserved sequence (3'CS), and the repeated conserved sequences 1 and 2 (RCS1 and RCS2).

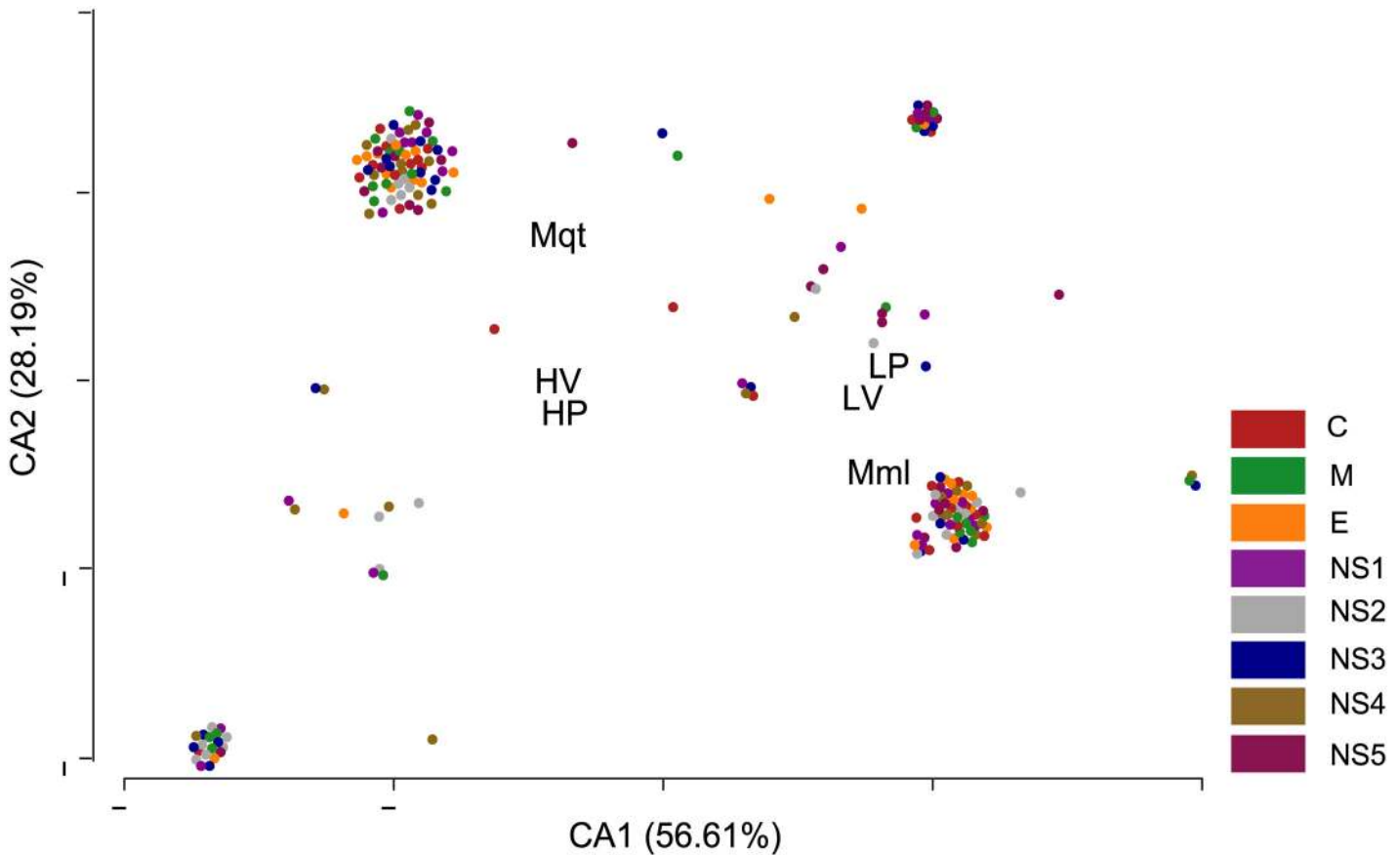
doi:10.1371/journal.pone.0136316.g002

specific point mutations and their presence in high/low viremia strains and high/low pathogenic strains (Exact Fisher's test;  $p > 0.01$ ).

### Phylogenetic analyses

Both Maximum Likelihood (Fig 5A) and Bayesian analyses (Fig 5B) methods showed the same phylogenetic reconstruction. The best fit model for both methods was GTR+I+G (parameters:  $G = 1$ ,  $I = 0.38$ ). High confidence values per node were observed with bootstrap values ranging from 70–100. The tree presented (Fig 5B) is the maximum clade credibility tree summarized from MCMC data with a burn in of 10%.

No pattern was found among phenotype (high/low viremia in birds, high/low pathogenicity in mice, or source of isolation) and phylogenetic classification. Additionally, no patterns based on either the host or year of isolation were identified. However, two main groups reflecting geographic origin could be distinguished: North America/northern-Central America (NAnCA) (Genotypes I and II: USA, MEX and GTM) and South America/southern-Central America (SAsCA) (Genotypes III, IV, V, VI and VII: ARG, BRA, PERU, PAN and TRN) (Fig 5C). Occasional exceptions were found such as one strain from USA (Imperial Valley, CA) that clustered in the SAsCA group. Although CbaAr-4005 and 79V-2533 strains were close to the NAnCA cluster, they constituted a separate lineage (Fig 5). In order to corroborate this result, a Procrustes analysis was carried out. It showed a high correlation among geographic and genetic distance. The goodness of fit in the NMDS was 0.012 and 0.039 for geographic and genetic distances, respectively. The configuration obtained between NMDSs based on 10000 permutations was highly significant ( $t_0 = 0.521$ ,  $m^2 = 0.728$  scale factor = 0.521,  $p$ -value = 0.001).



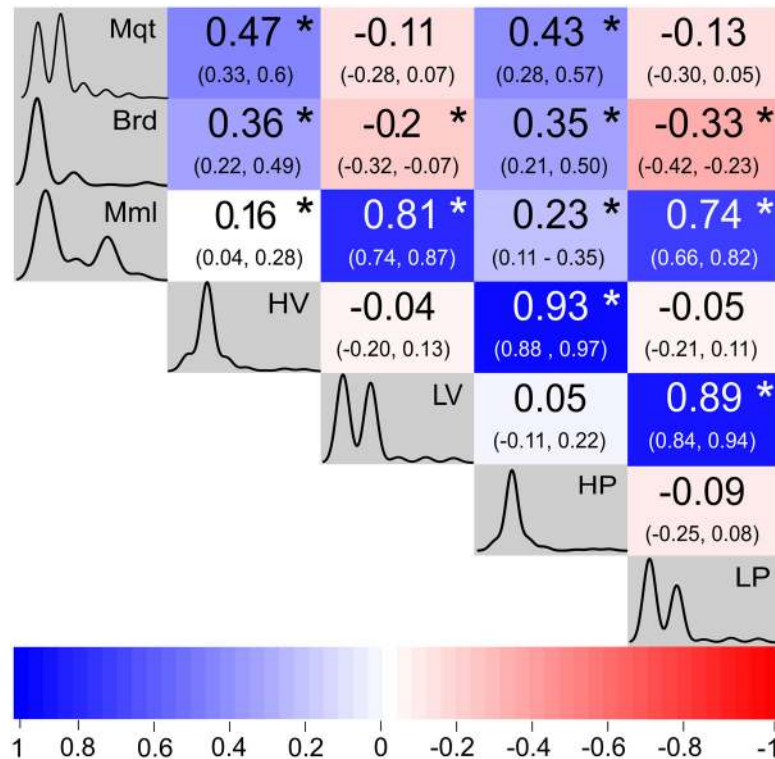
**Fig 3. Correspondence analysis exploring the distribution of 200 non-conservative mutations in 14 SLEV strains through biological features.** Axes CA1 and CA2 account for 84.8% of the total variation observed. HP: high pathogenicity, LP: low pathogenicity, HV: high viremia, LV: low viremia. Source of isolation: Mqt: mosquitoes, Mml: mammals, Brd: birds. Colored dots represent each mutation analyzed distributed by protein.

doi:10.1371/journal.pone.0136316.g003

## Discussion

The occurrence of human encephalitis outbreaks in Central and South America by flaviviruses is historically unusual. The first and largest encephalitis outbreak in human populations reported in South America was caused by *Rocio virus* in coastal region of southern Sao Paulo, Brazil. During that outbreak, a total of 465 cases with 61 deaths was recorded[27]. Beginning in 2002, human encephalitis caused by SLEV became frequent in Argentina and Brazil, with extra heamorrhagic manifestations reported in patients from Brazil [3]. Human outbreaks in populated areas of Argentina were reported in Córdoba (2005), Paraná (2006), Buenos Aires (2010), and San Juan (2012)[1,5,28]. Molecular evidence suggested some association between genotype III SLEV strains and encephalitis[9,10]. In this study, we explored potential molecular markers with epidemiological and biological features through *in silico* sequence comparison and bioinformatic analysis of epidemic and non epidemic SLEV strains isolated in the Americas.

The substitutions identified in the flavivirus structural proteins (C, prM and E), nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) and the 3' and 5' UTRs influence the biological (virulence, neuroinvasiveness, and viremia profiles) and adaptive behavior of the viruses[29,30]. In our study comparing the coding region of Ep and NEp SLEV strains, we detected 17 amino acid changes of which 59% were non-conservative (Fig 1). Some of these



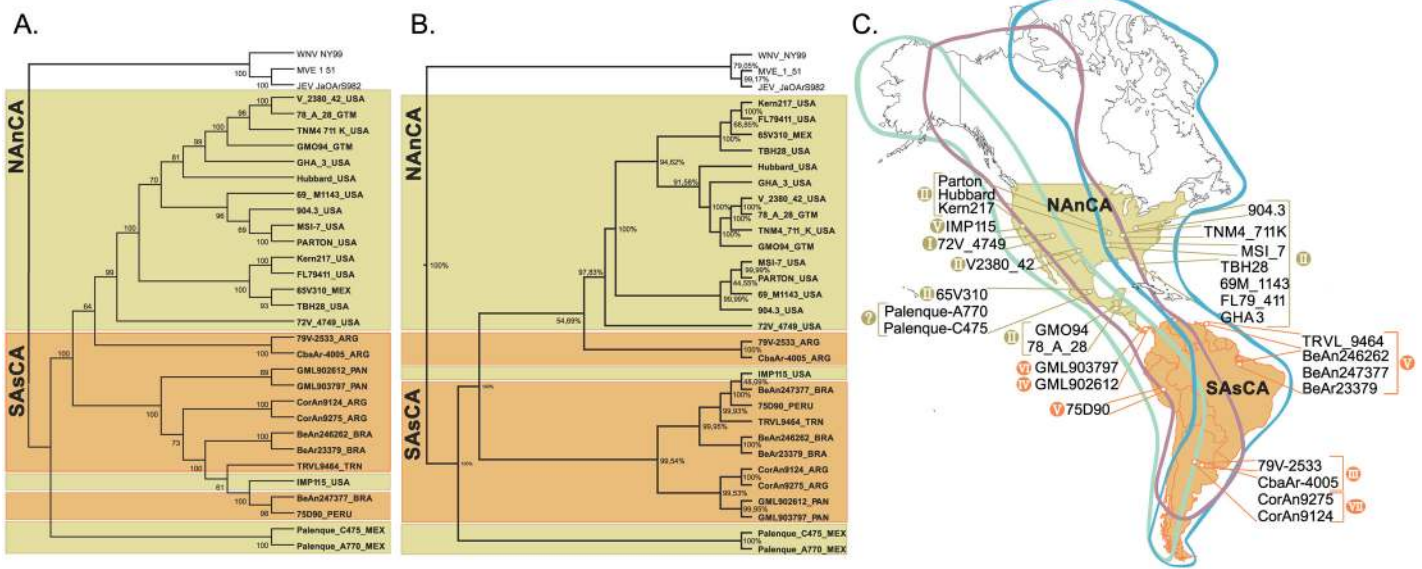
**Fig 4. Spearman's correlation indices and confidence intervals obtained comparing source of isolation (Mml: mammals, Brd: birds, Mqt: mosquito) and biological behavior (viremia in birds and pathogenicity in mice) in 14 SLEV strains.** Gradient scale colors indicate strength of correlation (positive or negative). \*: statistically significant differences.

doi:10.1371/journal.pone.0136316.g004

mutations are responsible for differences in biological behavior observed between Ep and NEp SLEV strains in animal models[12,23–25].

This study explored the association between 200 non-conservative point mutations and biological characteristics. Although multivariate analysis suggested there were associations of pooled mutations with biological features (Fig 3), no statistically significant differences were detected when analyzing mutations individually. This could be explained by the low number of viral strains characterized. Additionally, the difficulty in identifying molecular markers related to particular phenotypic characteristics could be masked or lost by compensatory sequence mutations that modify the viral fitness when the strains are subjected to cell culture for isolation[31,32].

It was also observed that several strains showed changes at cleavage site residues but only when the viral serine protease is involved in the processing at that site (Table 3). In the P1 and P2 positions, a predominance of positively charged amino acid residues were observed, while a glycine residue (G) is present in positions P1' and P2' (site P2P1/P1'P2'), in both DENV and WNV. It was also observed in both viruses, that an Arg or Lys was absolutely necessary for the cleavage at P1 or P2, while the Gly presence was not exclusively required for DENV [33]. Likewise, Arg or Lys residues were present in every strain of SLEV analyzed here. While WNV and DENV had a G residue at P1' and P2', a Lys was commonly found at P2' in the SLEV strains. A reverse genetic system would be necessary to analyze these changes and determine their impact on the phenotype.



**Fig 5. Phylogenetic analyses performed using the SLEV sequences listed in Table 1. (A): Maximum Likelihood analysis. (B): Bayesian inference analysis. (C) Geographic distribution of analyzed St. Louis encephalitis virus strains and migratory bird flyway.** West Nile virus (Acc. Number: DQ211652), Murray Valley Encephalitis virus (Acc. Number: AF161266) and Japanese Encephalitis virus (Acc. Number: M18370.1) were used as outgroups. Abbreviations for the different geographic origin of strains were used as follow: USA (United States); MEX (Mexico); GTM (Guatemala); ARG (Argentina); BRA (Brazil); PERU (Peru); TRN (Trinidad); PAN (Panamá).

doi:10.1371/journal.pone.0136316.g005

Of particular interest is the strong correlation detected between viremia in birds and pathogenicity in mice. It is known that SLEV viral strains are not able to cause neurological damage in birds but they do develop high viremias. This finding is contrary to what is observed during SLEV infection in mice, where viral loads are low but neurological signs are observed. Our findings suggest that mutations important for viral replication in birds are also significant in the neuroinvasion process in mammals. Experimental evidences suggests a role for E and non-structural proteins during this process [34].

Two reports described the conserved UTR secondary structure of SLEV and other flaviviruses [35,36]. Because these structures are highly conserved among flaviviruses, it is possible to infer their functional importance in viral RNA replication [37] and translation [38]. For example, it has been reported that changes in the WNV 3'UTR (along with other mutations in the coding region) generate temperature sensitive phenotypes, reduce plaque production, and result in attenuated neurovirulence in rodents [39]. A high degree of sequence conservation was observed when comparing DENV-2 and SLEV strains (Fig 2). For the potential 5'UAR (5'-AGCAGGGAAUUACCCAAUG-3') found in the 5'UTR, complementary residues (underlined) within the 3'UAR (5'-CAGGAGAUCCCCUGCUUU-3') were identified. These regions likely come together to form the panhandle structure for circularization of the genomic RNA such as is observed for other flaviviruses [40]. The RCS1 and RCS2 are totally conserved between these viruses and are present in all mosquito-borne flaviviruses. They have been shown in DENV to be implicated in RNA synthesis [41]. However, the impact of subtle variations in these regions between strains is unknown and therefore, it is very important to undertake studies assessing the structure and function of these regions in SLEV strains with different phenotypes.

The phylogenetic analysis based on complete ORF data and Procrustes analysis indicates a strong association between SLEV strains with their geographic distribution with two main



clades based on geography (North America /northern Central America and South America/southern Central America) (Fig 5). However, some exceptions are observed indicating that viral flow among the SLEV strains across its geographic range does occur [42]. Because SLEV is maintained by mosquito vectors and avian hosts, the actual distribution of SLEV along its northward and southward dispersion could be explained by bird migration. The same dispersal mechanism was postulated for WNV and its distribution across the Americas [43]. Although not completely understood, birds migrate across the continent through 3 main flyways (Pacific Americas, Mississippi Americas and Atlantic Americas flyways) (Fig 5C). Interestingly, the presence of South American genotype V SLEV strains in North America has been reported in California, Florida and Texas[8] overlapping with migration flyways (Fig 5C). Conversely, North America genotype II was reported in Argentina and Brazil[8,11]. Based on previously available genetic data, Auguste et al.[42] suggested that SLEV may have originated in Brazil. However, recent evidence obtained from the isolation of a new SLEV variant (Palenque strain) indicates a Central America origin[44].

This study represents the first theoretical step in understanding the molecular mechanisms underlying the virulence features and biological variation among SLEV strains. We have identified highly variable sites and point mutations among high pathogenicity/high viremia and low pathogenicity/low viremia strains that may provide the foundation necessary for development of molecular biology tools such as reverse genetic systems for SLEV. These tools will allow us to identify molecular markers for SLEV virulence and to dissect the mechanisms driving these interactions.

## Supporting Information

**S1 File. Procrustes analysis script to run on R.**  
(PDF)

## Acknowledgments

We thank Dr. Daniel Ghiringhelli for his advice in manual alignment corrections. Many thanks to International Union of Microbiology Societies and International Society of Infectious Diseases for supporting the internship of LAD at the CDC, Ft. Collins. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

## Author Contributions

Conceived and designed the experiments: LAD CHL AMP. Performed the experiments: LAD CHL AS. Analyzed the data: LAD SEG JAI AIQ. Contributed reagents/materials/analysis tools: LAD AMP CHL AS MSC JAI AIQ. Wrote the paper: LAD SEG JAI CHL AMP MSC AIQ. Designed one of the software used in the analysis: JAI.

## References

1. Spinsanti LI, Diaz LA, Glatstein N, Arselan S, Morales MA, Farias AA, et al. Human outbreak of St. Louis encephalitis detected in Argentina, 2005. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*. 2008; 42(1):27–33. Epub 2008/02/06. doi: [10.1016/j.jcv.2007.11.022](https://doi.org/10.1016/j.jcv.2007.11.022) PMID: [18249032](https://pubmed.ncbi.nlm.nih.gov/18249032/).
2. Rocco IM, Santos CL, Bisordi I, Petrella SM, Pereira LE, Souza RP, et al. St. Louis encephalitis virus: first isolation from a human in Sao Paulo State, Brazil. *Revista do Instituto de Medicina Tropical de Sao Paulo*. 2005; 47(5):281–5. doi: [/S0036-46652005000500008](https://doi.org/10.10036-46652005000500008). PMID: [16302112](https://pubmed.ncbi.nlm.nih.gov/16302112/).



3. Mondini A, Cardeal IL, Lazaro E, Nunes SH, Moreira CC, Rahal P, et al. Saint Louis encephalitis virus, Brazil. *Emerging infectious diseases*. 2007; 13(1):176–8. PMID: [17370543](#); PubMed Central PMCID: PMC2725838.
4. Rosa R, Costa EA, Marques RE, Oliveira TS, Furtini R, Bomfim MR, et al. Isolation of saint louis encephalitis virus from a horse with neurological disease in Brazil. *PLoS neglected tropical diseases*. 2013; 7(11):e2537. doi: [10.1371/journal.pntd.0002537](#) PMID: [24278489](#); PubMed Central PMCID: PMC3836713.
5. Spinsanti LI, Basquiera AL, Bulacio S, Somale V, Kim SC, Re V, et al. St. Louis encephalitis in Argentina: the first case reported in the last seventeen years. *Emerging infectious diseases*. 2003; 9(2):271–3. Epub 2003/02/27. PMID: [12604006](#); PubMed Central PMCID: PMC2901953.
6. Reisen WK. Epidemiology of St. Louis encephalitis virus. *Advances in virus research*. 2003; 61:139–83. Epub 2004/01/13. PMID: [14714432](#).
7. Kramer LD, Chandler LJ. Phylogenetic analysis of the envelope gene of St. Louis encephalitis virus. *Archives of virology*. 2001; 146(12):2341–55. Epub 2002/01/29. PMID: [11811684](#).
8. Rodrigues SG, Nunes MR, Casseb SM, Prazeres AS, Rodrigues DS, Silva MO, et al. Molecular epidemiology of Saint Louis encephalitis virus in the Brazilian Amazon: genetic divergence and dispersal. *The Journal of general virology*. 2010; 91(Pt 10):2420–7. Epub 2010/07/02. doi: [10.1099/vir.0.019117-0](#) PMID: [20592112](#).
9. Diaz LA, Re V, Almiron WR, Farias A, Vazquez A, Sanchez-Seco MP, et al. Genotype III Saint Louis encephalitis virus outbreak, Argentina, 2005. *Emerging infectious diseases*. 2006; 12(11):1752–4. Epub 2007/02/08. doi: [10.3201/eid1211.060486](#) PMID: [17283629](#); PubMed Central PMCID: PMC3372344.
10. Valinotto LE, Barrero PR, Viegas M, Alvarez Lopez MC, Mistchenko AS. Molecular evidence of St. Louis encephalitis virus infection in patients in Buenos Aires, Argentina. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*. 2012; 54(4):349–51. Epub 2012/05/23. doi: [10.1016/j.jcv.2012.04.005](#) PMID: [22608840](#).
11. Diaz LA, Albrieu Llinas G, Vazquez A, Tenorio A, Contigiani MS. Silent circulation of St. Louis encephalitis virus prior to an encephalitis outbreak in Cordoba, Argentina (2005). *PLoS neglected tropical diseases*. 2012; 6(1):e1489. Epub 2012/02/04. doi: [10.1371/journal.pntd.0001489](#) PMID: [22303490](#); PubMed Central PMCID: PMC3269431.
12. Diaz LA, Nemeth NM, Bowen RA, Almiron WR, Contigiani MS. Comparison of argentinean saint louis encephalitis virus non-epidemic and epidemic strain infections in an avian model. *PLoS neglected tropical diseases*. 2011; 5(5):e1177. Epub 2011/06/02. doi: [10.1371/journal.pntd.0001177](#) PMID: [21629729](#); PubMed Central PMCID: PMC3101189.
13. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*. 1994; 22(22):4673–80. Epub 1994/11/11. PMID: [7984417](#); PubMed Central PMCID: PMC308517.
14. Rice CMS. Production of flavivirus polypeptides by proteolytic processing. *Virology*. 1990; 170:10–20.
15. Medeiros DB, Nunes MR, Vasconcelos PF, Chang GJ, Kuno G. Complete genome characterization of Rocio virus (Flavivirus: Flaviviridae), a Brazilian flavivirus isolated from a fatal case of encephalitis during an epidemic in Sao Paulo state. *The Journal of general virology*. 2007; 88(Pt 8):2237–46. Epub 2007/07/12. doi: [10.1099/vir.0.82883-0](#) PMID: [17622628](#).
16. Paranjape SM, Harris E. Control of dengue virus translation and replication. *Current topics in microbiology and immunology*. 2010; 338:15–34. Epub 2009/10/06. doi: [10.1007/978-3-642-02215-9\\_2](#) PMID: [19802575](#).
17. Delpont W, Poon AF, Frost SD, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010; 26(19):2455–7. Epub 2010/07/31. doi: [10.1093/bioinformatics/btq429](#) PMID: [20671151](#); PubMed Central PMCID: PMC2944195.
18. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*. 2012; 9(8):772. Epub 2012/08/01. doi: [10.1038/nmeth.2109](#) PMID: [22847109](#).
19. Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, et al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*. 2014; 10(4):e1003537. Epub 2014/04/12. doi: [10.1371/journal.pcbi.1003537](#) PMID: [24722319](#); PubMed Central PMCID: PMC3985171.
20. Oksanen JB, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, et al. vegan: Community Ecology Package. R package version 2.0–10 2013. Available: <http://CRAN.R-project.org/package=vegan>.

21. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution*. 2013; 30(12):2725–9. Epub 2013/10/18. doi: [10.1093/molbev/mst1197](https://doi.org/10.1093/molbev/mst1197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/); PubMed Central PMCID: PMC3840312.
22. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014.
23. Monath TP, Cropp CB, Bowen GS, Kemp GE, Mitchell CJ, Gardner JJ. Variation in virulence for mice and rhesus monkeys among St. Louis encephalitis virus strains of different origin. *The American journal of tropical medicine and hygiene*. 1980; 29(5):948–62. Epub 1980/09/01. PMID: [6254384](https://pubmed.ncbi.nlm.nih.gov/6254384/).
24. Bowen GS, Monath TP, Kemp GE, Kerschner JH, Kirk LJ. Geographic variation among St. Louis encephalitis virus strains in the viremic responses of avian hosts. *The American journal of tropical medicine and hygiene*. 1980; 29(6):1411–9. Epub 1980/11/01. PMID: [7446828](https://pubmed.ncbi.nlm.nih.gov/7446828/).
25. Rivarola ME, Tauro LB, Llinas GA, Contigiani MS. Virulence variation among epidemic and non-epidemic strains of Saint Louis encephalitis virus circulating in Argentina. *Memorias do Instituto Oswaldo Cruz*. 2014; 109(2):197–201. Epub 2014/05/09. PMID: [24810175](https://pubmed.ncbi.nlm.nih.gov/24810175/); PubMed Central PMCID: PMC4015247.
26. Legendre P, Legendre LFJ. *Numerical Ecology*. Elsevier, editor: Elsevier; 1998. 852 p.
27. de Souza Lopes O, de Abreu Sacchetta L, Coimbra TL, Pinto GH, Glasser CM. Emergence of a new arbovirus disease in Brazil. II. Epidemiologic studies on 1975 epidemic. *American journal of epidemiology*. 1978; 108(5):394–401. PMID: [727209](https://pubmed.ncbi.nlm.nih.gov/727209/).
28. Seijo A, Morales A, Poustis G, Romer Y, Efron E, Vilora G, et al. [Outbreak of St. Louis encephalitis in the Metropolitan Buenos Aires Area]. *Medicina*. 2011; 71(3):211–7. Epub 2011/07/13. PMID: [21745768](https://pubmed.ncbi.nlm.nih.gov/21745768/).
29. Beasley DW, Davis CT, Whiteman M, Granwehr B, Kinney RM, Barrett AD. Molecular determinants of virulence of West Nile virus in North America. *Archives of virology Supplementum*. 2004;(18):35–41. Epub 2004/05/04. PMID: [15119761](https://pubmed.ncbi.nlm.nih.gov/15119761/).
30. Arroyo J, Guirakhoo F, Fenner S, Zhang ZX, Monath TP, Chambers TJ. Molecular basis for attenuation of neurovirulence of a yellow fever Virus/Japanese encephalitis virus chimera vaccine (ChimeriVax-JE). *Journal of virology*. 2001; 75(2):934–42. Epub 2001/01/03. doi: [10.1128/JVI.75.2.934-942.2001](https://doi.org/10.1128/JVI.75.2.934-942.2001) PMID: [11134306](https://pubmed.ncbi.nlm.nih.gov/11134306/); PubMed Central PMCID: PMC113989.
31. Ciota AT, Ehrbar DJ, Matakchiero AC, Van Slyke GA, Kramer LD. The evolution of virulence of West Nile virus in a mosquito vector: implications for arbovirus adaptation and evolution. *BMC evolutionary biology*. 2013; 13:71. Epub 2013/03/22. doi: [10.1186/1471-2148-13-71](https://doi.org/10.1186/1471-2148-13-71) PMID: [23514328](https://pubmed.ncbi.nlm.nih.gov/23514328/); PubMed Central PMCID: PMC3626576.
32. Parks CL, Lerch RA, Walpita P, Wang HP, Sidhu MS, Udem SA. Comparison of predicted amino acid sequences of measles virus strains in the Edmonston vaccine lineage. *Journal of virology*. 2001; 75(2):910–20. Epub 2001/01/03. doi: [10.1128/JVI.75.2.910-920.2001](https://doi.org/10.1128/JVI.75.2.910-920.2001) PMID: [11134304](https://pubmed.ncbi.nlm.nih.gov/11134304/); PubMed Central PMCID: PMC113987.
33. Shiryayev SA, Strongin AY. Structural and functional parameters of the flaviviral protease: a promising antiviral drug target. *Future virology*. 2010; 5(5):593–606. doi: [10.2217/fvl.10.39](https://doi.org/10.2217/fvl.10.39) PMID: [21076642](https://pubmed.ncbi.nlm.nih.gov/21076642/); PubMed Central PMCID: PMC2976050.
34. Neal JW. Flaviviruses are neurotropic, but how do they invade the CNS? *The Journal of infection*. 2014; 69(3):203–15. Epub 2014/06/01. doi: [10.1016/j.jinf.2014.05.010](https://doi.org/10.1016/j.jinf.2014.05.010) PMID: [24880028](https://pubmed.ncbi.nlm.nih.gov/24880028/).
35. Brinton MA, Dispoto JH. Sequence and secondary structure analysis of the 5'-terminal region of flavivirus genome RNA. *Virology*. 1988; 162(2):290–9. Epub 1988/02/01. PMID: [2829420](https://pubmed.ncbi.nlm.nih.gov/2829420/).
36. Brinton MA, Fernandez AV, Dispoto JH. The 3'-nucleotides of flavivirus genomic RNA form a conserved secondary structure. *Virology*. 1986; 153(1):113–21. Epub 1986/08/01. PMID: [3016981](https://pubmed.ncbi.nlm.nih.gov/3016981/).
37. Filomatori CV, Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik AV. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes & development*. 2006; 20(16):2238–49. Epub 2006/08/03. doi: [10.1101/gad.1444206](https://doi.org/10.1101/gad.1444206) PMID: [16882970](https://pubmed.ncbi.nlm.nih.gov/16882970/); PubMed Central PMCID: PMC1553207.
38. Chiu WW, Kinney RM, Dreher TW. Control of translation by the 5'- and 3'-terminal regions of the dengue virus genome. *Journal of virology*. 2005; 79(13):8303–15. Epub 2005/06/16. doi: [10.1128/JVI.79.13.8303-8315.2005](https://doi.org/10.1128/JVI.79.13.8303-8315.2005) PMID: [15956576](https://pubmed.ncbi.nlm.nih.gov/15956576/); PubMed Central PMCID: PMC1143759.
39. Davis CT, Galbraith SE, Zhang S, Whiteman MC, Li L, Kinney RM, et al. A combination of naturally occurring mutations in North American West Nile virus nonstructural protein genes and in the 3' untranslated region alters virus phenotype. *Journal of virology*. 2007; 81(11):6111–6. Epub 2007/03/23. doi: [10.1128/JVI.02387-06](https://doi.org/10.1128/JVI.02387-06) PMID: [17376926](https://pubmed.ncbi.nlm.nih.gov/17376926/); PubMed Central PMCID: PMC1900292.
40. Gritsun TS, Gould EA. Origin and evolution of flavivirus 5'UTRs and panhandles: trans-terminal duplications? *Virology*. 2007; 366(1):8–15. Epub 2007/07/31. doi: [10.1016/j.virol.2007.04.011](https://doi.org/10.1016/j.virol.2007.04.011) PMID: [17658577](https://pubmed.ncbi.nlm.nih.gov/17658577/).

41. Gebhard LG, Filomatori CV, Gamarnik AV. Functional RNA elements in the dengue virus genome. *Viruses*. 2011; 3(9):1739–56. Epub 2011/10/14. doi: [10.3390/v3091739](https://doi.org/10.3390/v3091739) PMID: [21994804](https://pubmed.ncbi.nlm.nih.gov/21994804/); PubMed Central PMCID: PMC3187688.
42. Auguste AJ, Pybus OG, Carrington CV. Evolution and dispersal of St. Louis encephalitis virus in the Americas. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2009; 9(4):709–15. Epub 2008/08/19. doi: [10.1016/j.meegid.2008.07.006](https://doi.org/10.1016/j.meegid.2008.07.006) PMID: [18708161](https://pubmed.ncbi.nlm.nih.gov/18708161/).
43. Diaz LA, Komar N, Visintin A, Dantur Juri MJ, Stein M, Lobo Allende R, et al. West Nile Virus in Birds, Argentina. *Emerg Infect Dis*. 2008; 14(4):3. doi: [10.3201/eid1404.071257](https://doi.org/10.3201/eid1404.071257)
44. Kopp A, Gillespie TR, Hobelsberger D, Estrada A, Harper JM, Miller RA, et al. Provenance and geographic spread of St. Louis encephalitis virus. *mBio*. 2013; 4(3):e00322–13. Epub 2013/06/14. doi: [10.1128/mBio.00322-13](https://doi.org/10.1128/mBio.00322-13) PMID: [23760463](https://pubmed.ncbi.nlm.nih.gov/23760463/); PubMed Central PMCID: PMC3685209.