

Genetic Diversity of Toscana Virus

Ximena Collao,^{1,2} Gustavo Palacios,¹
Sara Sanbonmatsu-Gámez, Mercedes Pérez-Ruiz,
Ana I. Negredo, José-María Navarro-Marí,
Marc Grandadam, Ana Maria Aransay,
W. Ian Lipkin, Antonio Tenorio,³
and María-Paz Sánchez-Seco³

Distribution of Toscana virus (TOSV) is evolving with climate change, and pathogenicity may be higher in non-exposed populations outside areas of current prevalence (Mediterranean Basin). To characterize genetic diversity of TOSV, we determined the coding sequences of isolates from Spain and France. TOSV is more diverse than other well-studied phleboviruses (e.g., Rift Valley fever virus).

Toscana virus (TOSV) belongs to the family *Bunyaviridae* and genus *Phlebovirus*. According to the Eighth Report of the International Committee on Taxonomy of Viruses, phleboviruses are classified into 9 serocomplexes (1), 1 of which includes Rift Valley fever virus (RVFV), a deadly pathogen for cattle and humans. TOSV belongs to the sandfly virus serotype Naples serocomplex. Phleboviruses have segmented RNA genomes comprised of 3 units: L (large), M (medium), and S (small). Based on the analysis of the G_N glycoprotein (M segment), 4 lineages of TOSV have been proposed (2). Moreover, phylogenetic analysis has demonstrated that TOSV isolates from Spain differ from the original isolates from Italy, TOSV strain ISS Phl.3 (3).

TOSV is widespread in the Mediterranean Basin, and evidence of human infection has been found in Italy, France, Spain, Portugal, Cyprus, and Turkey. The main clinical manifestation is neurologic dysfunction (4). Although the virus is an important pathogen (4), little genomic information is available for TOSV. We believe that obtaining genome sequence information for viruses with poor representation in public databases is an urgent task for the virologic community, especially for programs of virus surveillance and study of emerging pathogens. The increas-

ing popularity of nucleic acid-based methods of detection underscores the need for a deeper knowledge of sequence variability of wild-type strains to ensure the sensitivity and specificity of the diagnostic assays. We therefore characterized the genetic diversity of TOSV by determining the coding sequences of isolates from Spain and France.

The Study

Viruses were recovered from Vero E6 (ATCC CRL-1586) cell cultures of cerebrospinal fluid samples (CSF) from 2 TOSV-infected patients and from a pool of sandflies. The first patient was a 30-year-old man living in a rural area of Granada, Spain. He had been admitted to hospital with headache, nausea, and fever; he was found to have meningeal and CSF pleocytosis. TOSV ESH 62100 was detected in an acute-phase CSF sample by reverse transcription-PCR (RT-PCR) (5). The second patient was a 57-year-old woman from Augbanc (Marseille area), France, who had fever, malaise, photophobia, neck rigidity, and vomiting. TOSV-specific immunoglobulin M was detected in an acute-phase serum sample; a convalescent-phase sample indicated seroconversion. Her CSF sample was positive for TOSV by PCR (6) and was the source for the isolation of TOSV H/IMTSSA. TOSV EsPhGR40 was isolated from a pool of female sandflies captured in the metropolitan area of Granada (3).

Virus sequences were obtained by using conserved primers (primer sequences available on request). GenBank accession numbers of the new TOSV sequences were EF120631 and FJ153279–FJ153286. To assess the potential for establishing a simple cutoff for classification of TOSV genotypes (e.g., similar to programs created for mumps) (7), we used pairwise sequence comparison to compare these sequences with all published phleboviral sequences. Sequence analysis of the M segment was sufficient to enable determination of genotypes (Figure 1). Two different clusters were clearly distinguished. We propose to name those clusters genotypes A and B. The previously described TOSV lineages that include isolates from Italy all clustered in the proposed TOSV A genotype. Comparison with RVFV, for which 7 lineages (A to G) have been described (8), showed higher divergence in L and M segments and nonstructural (NS) and nucleoprotein (N) genes between the proposed TOSV genotypes than that found among RVFV lineages (Table).

Pairwise comparison demonstrated the following: 1) all TOSV S segments were highly conserved; 2) the L segment demonstrated less conservation than the N gene at the nucleotide or deduced amino acid levels; 3) TOSV M segments were the most divergent; and 4) variation in M segments was higher than that among RVFV strains but less

Author affiliations: Institute of Health "Carlos III," Madrid, Spain (X. Collao, A.I. Negredo, A. Tenorio, M.-P. Sánchez-Seco); Columbia University, New York, New York, USA (G. Palacios, W.I. Lipkin); University Hospital "Virgen de las Nieves," Granada, Spain (S. Sanbonmatsu-Gámez, M. Pérez-Ruiz, J.-M. Navarro-Marí); Institut de Médecine Tropicale du Service de Santé des Armées, Marseille, France (M. Grandadam); and Centro de Investigación Cooperación bioGUNE, Vizcaya, Spain (A.M. Aransay)

DOI: 10.3201/eid1504.081111

^{1,3}These pairs of authors contributed equally to this work.

²Current affiliation: Valparaíso University, Valparaíso, Chile.

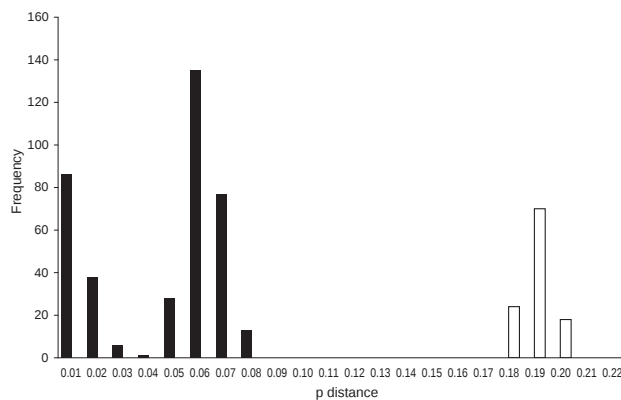


Figure 1. Histogram showing distribution of nucleotide pairwise (p) distances in the medium segment of Toscana virus. p distances are for nucleotides; frequencies are for intervals of 0.01. Validity of this method was confirmed by analysis of variance, comparing the scores of sequence comparisons within genotypes to those between genotypes. Black bars indicate intralineage distribution; white bars indicate interlineage distribution.

pronounced than that within the group of sandfly fever Sicilian viruses or sandfly fever Naples viruses (Table). With regard to the conserved S segments, the NS gene was more variable than the N gene, and despite a nucleotide pairwise difference of 12.1% in the N coding region, the amino acid sequences were completely conserved (100% identity);

Phylogenetic analysis of all phlebovirus sequences was performed by using the maximum-likelihood method available in PAUP* under tree bisection and reconnection branch swapping; the best-fit model of nucleotide substitu-

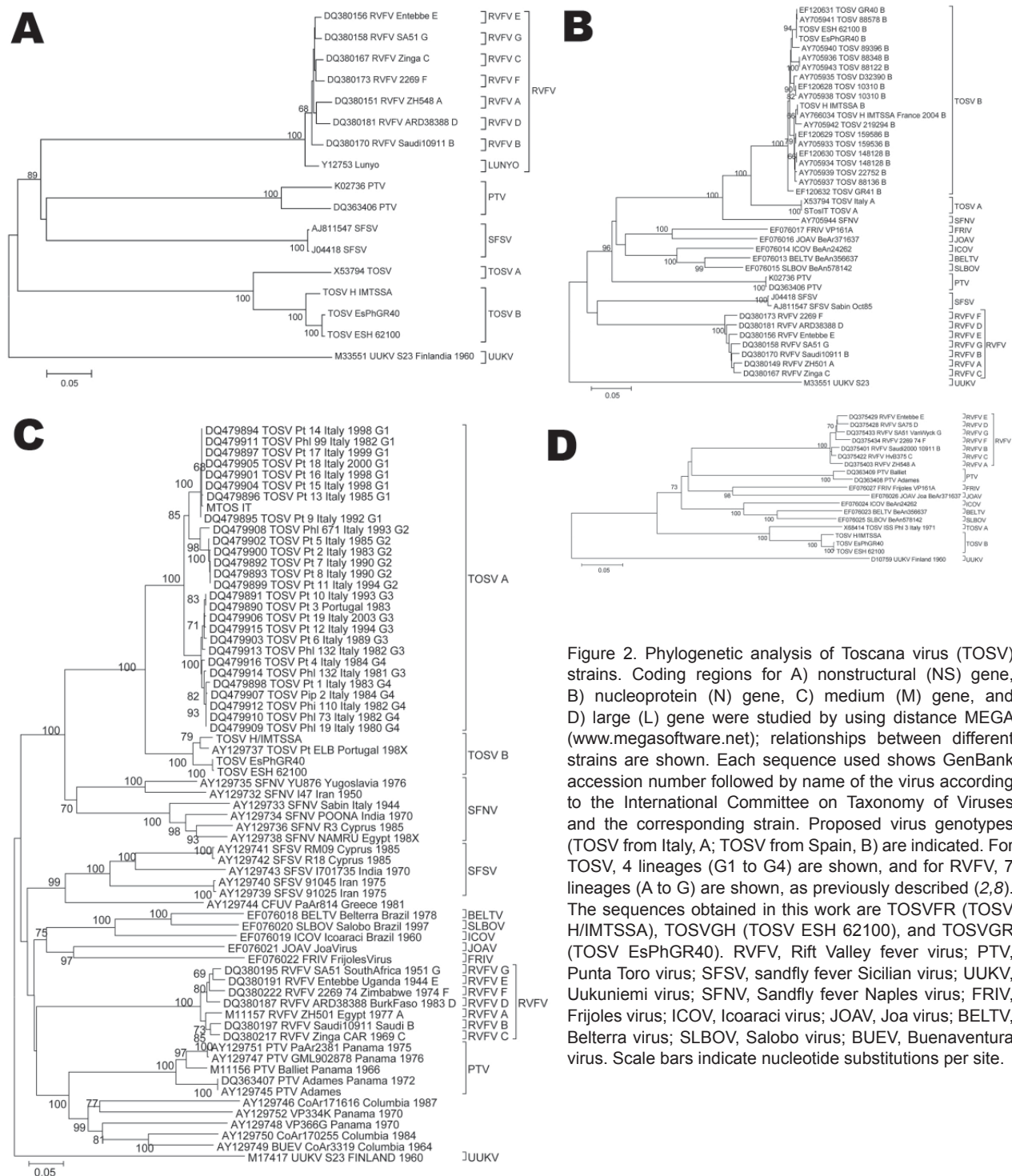
tion in each instance was determined by using MODELTEST (9) (parameter values available upon request) (Figure 2). Time to most recent common ancestor (TMRCA) was estimated for the M segment by using the data from 32 dated samples collected over 35 years and the Bayesian Markov Chain Monte Carlo approach (BEAST package; 10). We applied a relaxed molecular clock with an uncorrelated exponential distribution of rates, a general time reversible + I + F4 model of nucleotide substitution and logistic population growth. Only the M segment had enough sequence representation to perform the analysis. Nonetheless, because glycoproteins are present on the surface of virions, they are the proteins most exposed to the selective pressures of the host.

The estimated mean rate of evolutionary change was 9.1×10^{-5} substitutions/site/year (95% highest probability density [HPD] = 2.5×10^{-4} to 2.7×10^{-6} substitutions/site/year). Under this rate the mean TMRCA was estimated to be 3,265 years, although with wide variance (178–11,000 years). Studies of RVFV, a closely related phlebovirus (11) have estimated the mean RVFV M segment rate to be 2.42×10^{-4} (95% HPD = 1.8×10^{-4} to 3.0×10^{-4} substitutions/site/year) (8). TOSV M segment TMRCA is higher than RVFV M segment TMRCA (117.3 years [variance 95–143 years]) (8). Although the wide variance of the TMRCA calculations might be affected by the scarce sequence information available, the differences observed might also be related to their biological differences. Both are phleboviruses transmitted by the bite of arthropods that could also be implied as amplifier hosts; however, because TOSV is believed not to be amplified on its

Table. Sequence differences among phleboviruses*

Comparison	Similarity of sequences (amino acid similarity), mean % $\pm$ SD			
	L segment	M segment	NS gene	N gene
Within TOSV	91.7 $\pm$ 0.7 (98.8 $\pm$ 0.5)	92.4 $\pm$ 0.6 (95.4 $\pm$ 0.9)	90.7 $\pm$ 0.8 (95.8 $\pm$ 1)	95.9 $\pm$ 0.8 (100)
Within TOSV A	NA	99.3 $\pm$ 0.1 (99.4 $\pm$ 0.5)	NA	100 (100)
Within TOSV B	98.2 $\pm$ 0.2 (99.7 $\pm$ 0.1)	95.5 $\pm$ 0.4 (98.2 $\pm$ 0.5)	97.3 $\pm$ 0.4 (98.5 $\pm$ 0.5)	98.5 $\pm$ 0.3 (100)
Between TOSV A and TOSV B	85.2 $\pm$ 0.4 (97.0 $\pm$ 0.4)	81.8 $\pm$ 0.4 (90.0 $\pm$ 1.1)	83.6 $\pm$ 1.1 (89.6 $\pm$ 1.6)	87.9 $\pm$ 1.4 (100)
Within TOSV A G1 to G4	NA	99.3 $\pm$ 0.1 (99.3 $\pm$ 0.2)	NA	NA
Between TOSV A G1, G2, G3, and G4	NA	95.6 $\pm$ 0.4 (97.3 $\pm$ 0.5)	NA	NA
Within RVFV	97.7 $\pm$ 0.3 (99.8 $\pm$ 0.1)	95.4 $\pm$ 0.5 (98.6 $\pm$ 0.5)	97.2 $\pm$ 0.3 (98.5 $\pm$ 0.4)	97.6 $\pm$ 0.5 (99.8 $\pm$ 0.1)
Within RVFV lineages	98.8 $\pm$ 0.1 (99.5 $\pm$ 0.1)	98.8 $\pm$ 0.1 (99.5 $\pm$ 0.1)	98.8 $\pm$ 0.2 (99.1 $\pm$ 0.3)	99.1 $\pm$ 0.2 (99.9 $\pm$ 0.1)
Between RVFV lineages	96.6 $\pm$ 0.2 (99.2 $\pm$ 0.2)	96.6 $\pm$ 0.2 (99.1 $\pm$ 0.2)	97.0 $\pm$ 0.5 (98.0 $\pm$ 0.6)	97.5 $\pm$ 0.5 (99.6 $\pm$ 0.3)
Within SFSV	NA	76.5 $\pm$ 1.3 (80.2 $\pm$ 2.1)	99.7 $\pm$ 0.2 (98.9 $\pm$ 0.7)	100 (100)
Within SFNV	NA	72.2 $\pm$ 1.2 (72.9 $\pm$ 2.1)	NA	NA
Within PTV	96.5 $\pm$ 0.7 (100)	92.5 $\pm$ 0.8 (97.5 $\pm$ 0.9)	87.6 $\pm$ 1.4 (94.2 $\pm$ 1.6)	100 (100)
Between SFSV and PTV	NA	45.5 $\pm$ 1.8 (32.3 $\pm$ 3.2)	40.2 $\pm$ 1.9 (23.7 $\pm$ 3.1)	55.0 $\pm$ 3.6 (52.5 $\pm$ 6.4)
Between SFNV and TOSV	NA	57.2 $\pm$ 1.7 (53.6 $\pm$ 3.1)	NA	78.8 $\pm$ 2.8 (93.2 $\pm$ 3.3)

*Based on complete genome sequences. To estimate the evolutionary divergence over sequence pairs between groups, the average number of nucleotide differences per site over all sequence pairs between groups and among groups was calculated. Standard error estimates were obtained by a bootstrap procedure (500 replicates). For each pairwise sequence comparison, all positions containing alignment gaps and missing data were eliminated. L, large; M, medium; NS, nonstructural; N, nucleoprotein; TOSV, Toscana virus; NA, not applicable; RVFV, Rift Valley fever virus; SFSV, sandfly fever Sicilian virus; SFNV, sandfly fever Naples virus; PTV, Punta Toro virus.



mammalian hosts, it seems that the RVFV cycle includes mammals. The high conservation of RVFV genome sequences has been interpreted as indicating that the overall tolerance for mutation within the RVFV genome is low or that the viruses in the group have a relatively recent common ancestor (12). More studies are needed to understand

the significance of the high degree of purifying selection observed in TOSV.

Conclusions

Our main goal, to improve the knowledge of sequence information for this neglected genus, was achieved by the

addition of 3 full genome sequences of TOSV. Geographic distribution may be different for each TOSV genotype. Whereas the TOSV A genotype circulates in Italy, France, and Portugal, the TOSV B genotype circulates in Spain, Portugal, and France. Both genotypes have been reported previously in France (13); however, the cocirculation of both genotypes in Portugal is confusing because sequences obtained independently in 2 laboratories from strains allegedly obtained from the same patient clustered in different clades (DQ479890 and AY129737) (14,15). Geographic differences in genotype distribution may relate to differences in vector distribution.

The distribution of phlebotomines is evolving with climate change and has implications for the epidemiology of vector-borne infectious diseases. Preexisting immunity almost certainly plays a role in limiting illness in the Mediterranean Basin; pathogenicity may be higher in naive populations outside areas of current TOSV prevalence. Thus, TOSV might be considered a threat for human health, and research and surveillance programs should aim to prevent its spread to new areas.

Acknowledgments

We thank all members of the Red de Enfermedades Viricas Transmitidas por Artrópodos y Roedores, a multidisciplinary group founded by the Fondo de Investigaciones Sanitarias, Spanish Ministry of Health (G03/059), which has partially funded this work, for their constructive discussions, comments, and advice.

This study was supported in part by National Institutes of Health awards AI062705, AI051292, AI059576, HL083850, and AI57158 (Northeast Biodefense Center–Lipkin).

Dr Collao is a postdoctoral research fellow at the University of Valparaiso, Chile. She obtained her doctoral degree in Spain while investigating TOSV and other phleboviruses. Her current work focuses on the development of arbovirus surveillance programs in Chile.

References

- Nichol ST, Beaty BJ, Elliott RM, Goldbach R, Plyusin A, Schmaljohn C, et al. Family *Bunyaviridae*. London: Elsevier Academic Press; 2005.
- Venturi G, Ciccozzi M, Montieri S, Bartoloni A, Francisci D, Nicoletti L, et al. Genetic variability of the M genome segment of clinical and environmental Toscana virus strains. *J Gen Virol*. 2007;88:1288–94. DOI: 10.1099/vir.0.82330-0
- Sanbonmatsu-Gamez S, Perez-Ruiz M, Collao X, Sanchez-Seco MP, Morillas-Marquez F, de la Rosa-Fraile M, et al. Toscana virus in Spain. *Emerg Infect Dis*. 2005;11:1701–7.
- Charrel RN, Gallian P, Navarro-Mari JM, Nicoletti L, Papa A, Sanchez-Seco MP, et al. Emergence of Toscana virus in Europe. *Emerg Infect Dis*. 2005;11:1657–63.
- Sanchez-Seco MP, Echevarria JM, Hernandez L, Estevez D, Navarro-Mari JM, Tenorio A. Detection and identification of Toscana and other phleboviruses by RT-nested-PCR assays with degenerated primers. *J Med Virol*. 2003;71:140–9. DOI: 10.1002/jmv.10465
- Peyrefitte CN, Devetakov I, Pastorino B, Villeneuve L, Bessaud M, Stolidi P, et al. Toscana virus and acute meningitis, France. *Emerg Infect Dis*. 2005;11:778–80.
- Palacios G, Jabado O, Cisterna D. Molecular identification of mumps virus genotypes from clinical samples: standardized method of analysis. *J Clin Microbiol*. 2005;43:1869–78. DOI: 10.1128/JCM.43.4.1869-1878.2005
- Bird BH, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST. Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. *J Virol*. 2007;81:2805–16. DOI: 10.1128/JVI.02095-06
- Posada D. Using MODELTEST and PAUP* to select a model of nucleotide substitution. *Curr Protocols in Bioinformatics*. Jan 2003: 1–14
- Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214. DOI: 10.1186/1471-2148-7-214
- Jenkins GM, Rambaut A, Pybus OG, Holmes EC. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol*. 2002;54:156–65. DOI: 10.1007/s00239-001-0064-3
- Weaver SC. Evolutionary influences in arboviral disease. *Curr Top Microbiol Immunol*. 2006;299:285–314. DOI: 10.1007/3-540-26397-7_10
- Charrel RN, Izri A, Temmam S, Delaunay P, Toga I, Dumon H, et al. Cocirculation of 2 genotypes of Toscana virus, southeastern France. *Emerg Infect Dis*. 2007;13:465–8.
- Liu DY, Tesh RB, Travassos Da Rosa AP, Peters CJ, Yang Z, Guzman H, et al. Phylogenetic relationships among members of the genus *Phlebovirus* (*Bunyaviridae*) based on partial M segment sequence analyses. *J Gen Virol*. 2003;84:465–73. DOI: 10.1099/vir.0.18765-0
- Venturi G, Madeddu G, Rezza G, Ciccozzi M, Pettinato ML, Cilliano M, et al. Detection of Toscana virus central nervous system infections in Sardinia Island, Italy. *J Clin Virol*. 2007;40:90–1. DOI: 10.1016/j.jcv.2007.06.005

Address for correspondence: María-Paz Sánchez-Seco, Laboratory of Arbovirus and Imported Viral Diseases, National Center of Microbiology, Institute of Health “Carlos III,” Ctra Pozuelo-Majadahonda km 2, 28220 Madrid, Spain; email: paz.sanchez@isciii.es

All material published in *Emerging Infectious Diseases* is in the public domain and may be used and reprinted without special permission; proper citation, however, is required.

Search past issues of EID at www.cdc.gov/eid