Infection, Genetics and Evolution 41 (2016) 93–99

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

# Phylogeographic analysis of the 2000–2002 foot-and-mouth disease epidemic in Argentina

Barbara Brito [a,*], Guido König [b], Sebastian Cabanne [b], Claudia Perez Beascoechea [c,1], Luis Rodriguez [a], Andres Perez [d]

[a] USDA/ARS Foreign Animal Disease Research Unit, Plum Island Animal Disease Center, P.O. Box 848, Greenport, NY 11944, USA
[b] Instituto de Biotecnología, INTA, Buenos Aires, Argentina
[c] FMD Virology Department, OIE FMD Reference Laboratory, DLA, Servicio Nacional de Sanidad y Calidad Agroalimentaria (SENASA), Dirección de Laboratorio Animal, Argentina
[d] University of Minnesota, Department of Veterinary Population Medicine, College of Veterinary Medicine, Saint Paul, MN, USA.

ARTICLE INFO

*Article history:*
Received 5 November 2015
Received in revised form 18 March 2016
Accepted 25 March 2016
Available online xxxx

*Keywords:*
Foot-and-mouth disease
Viral transmission
Molecular epidemiology
Phylogeography
Argentina

ABSTRACT

Foot-and-mouth disease (FMD) is a highly transmissible disease of hooved livestock. Although FMD has been eradicated from many countries, economic and social consequences of FMD reintroductions are devastating. After achieving disease eradication, Argentina was affected by a major epidemic in 2000–2002, and within few months, FMD virus spread throughout most of the country and affected >2500 herds. Available records and viral strains allowed us to assess the origins, spread and progression of this FMD epidemic, which remained uncertain. We used whole genome viral sequences and a continuous phylogeographic diffusion approach, which revealed that the viruses that caused the outbreaks spread fast in different directions from a central area in Argentina. The analysis also suggests that the virus that caused the outbreaks in the year 2000 was different from those found during the 2001 epidemic. To estimate if the approximate overall genetic diversity of the virus was related to disease transmission, we reconstructed the viral demographic variation in time using Bayesian Skygrid approach and compared it with the epidemic curve and the within-herd transmission rate and showed that the genetic temporal diversity of the virus was associated with the increasing number of outbreaks in the exponential phase of the epidemic. Results here provide new evidence of how the disease entered and spread throughout the country. We further demonstrate that genetic data collected during a FMD epidemic can be informative indicators of the progression of an ongoing epidemic.



## 1. Introduction

Foot-and-mouth disease (FMD) is arguably one of the most important diseases of hooved livestock worldwide and it is endemic in many African and Asian countries (OIE, 2009). In South America, official disease status and effectiveness of control programs are disparate between countries, resulting in sporadic epidemics in disease-free areas. One of the most dramatic examples of this situation took place in 2000–2002, when a devastating FMD epidemic affected Argentina, few months after the country discontinued vaccination and was recognized as free-without vaccination by the International Organisation of Animal Health (OIE). There was an initial outbreak in August 2000 that was controlled using stamping out and movement control policy (Perez et al., 2004a). Unfortunately early in 2001, another introduction FMD virus (FMDV) serotype A, occurred, resulting in a total of 2519 herds affected throughout the country. This epidemic was finally controlled implementing massive vaccination, animal movement control, and active surveillance strategies (Perez et al., 2004b). However, there was a need to reformulate the vaccination strategy twice to incorporate the strains from 2000 (A/Arg/00) and 2001 (A/Arg/01) viruses (Mattion et al., 2004; Perez et al., 2004a). Since then, the country conducts mandatory vaccination campaigns twice a year nationwide, except in southern areas of Argentina, the only region in the country that has remained free from FMD. This large epidemic in Argentina was well documented and previous studies of this epidemic have been aimed at describing the epidemiological features (Perez et al., 2004b), molecular epidemiology (Konig et al., 2007; Perez et al., 2008), vaccine matching (Mattion et al., 2009), and within-herd transmission (Brito et al., 2011).

During the last decade, methods for analyzing sequences and using them to reveal epidemiologic aspects of infectious diseases based on genetic data of the pathogens, a field known as phylodynamics, have been rapidly emerging and evolving (Grenfell et al., 2004; Volz et al., 2013). Within this field, the reconstruction of the phylogeny and the geographical location of virus' ancestors, referred to as phylogeography has

* Corresponding author at: Universidad de Chile, Facultad de Ciencias Veterinarias y Pecuarias, Departamento de Medicina Preventiva, Av. Santa Rosa 11735, La Pintana, Santiago, Chile.
*E-mail address:* barbara.brito.r@gmail.com (B. Brito).
[1] Deceased 10/8/2015.

http://dx.doi.org/10.1016/j.meegid.2016.03.026
1567-1348/

shown to have an important application to understand viral spread. Additionally, important information contained in the estimated phylogeny can be used to infer temporal changes in the viral genetic diversity (estimated by the effective viral population size). Previous studies have suggested that the increase of the viral population size or genetic diversity is related to an increase of disease incidence, and transmission rate of the virus due to the association between the coalescent rate and rate of transmission (Frost & Volz, 2010; Holmes et al., 1995). However, estimates of infectious diseases transmission rates normally rely on partial data obtained from a proportion of individuals reporting the disease, so the real incidence is approximated using different epidemiological tools. In the case of some animal diseases, reporting is mandatory, and clinical disease is evident in most cases, so animal health authority records allow having accurate information about disease occurrence. Such is the case of FMDV when introduced into a disease-free country.

The objective of the study here was to estimate the spread pattern of FMDV A/Arg/01 in one of the largest epidemics reported and documented in a disease-free country. We used Bayesian methods to reconstruct the phylogeny, and a random walk diffusion model to account for the spatial spread of the virus. We further determined the viral population size and compared it with the epidemic curve and the within-herd transmission rate of the FMD epidemic in Argentina in 2001. Results from the estimated viral population size obtained from sequence analysis suggest that this approach can be applied to detect the duration of the exponential growth of cases over the course of an epidemic.

## 2. Methods

### 2.1. Study population and sampling

Veterinarians of the National Argentine Veterinary Services (SENASA), visited all affected herds and collected samples from clinical lesions, as well as all data pertaining to herd demographics and diseased animals. Collected samples were stored at −70 °C. All records of infected herds were categorized into 3 different periods of the epidemic (exponential, saturation, and declining phase), and geographically in 3 different regions, as previously described (Perez et al., 2004b). Finally, thirty specimens from different time (epidemic periods) and geographical categories previously described (the number of sequences included from each epidemic period and geographic category was proportional to the number of outbreak events occurred within those categories), and which had a relatively high viral load were sequenced. All samples were collected in cattle, which was, almost exclusively, the only susceptible species affected during the epidemic.

Information of the computed intra-herd transmission coefficient (the rate at which a susceptible individual acquired the infection from an infected one) for every herd was available from our prior study (Brito et al., 2011).

### 2.2. RNA extraction, RT-PCR, and virus sequencing

Whole genome sequencing was performed at the National Institute of Agricultural Technology. Viral RNA was inactivated and extracted directly from the disrupted tissue (field samples) using Tri-Reagent® (MRC) according to the manufacturer's protocol. RNA was reverse transcribed into cDNA using 50 ng of random hexamer primers (Invitrogen), 2 pmol of specific reverse primer, and 400 U of MMLV-RT Superscript III (Invitrgoen) following the manufacturer's indications except for cDNA treatment with RNAse H (not performed). cDNA was used as template for the amplification of 7 overlapping fragments using Taq polymerase HiFi (Invitrogen). PCR tests were performed in order to amplify the complete genome in seven 0.4–2.0 kb-long PCR fragments. An automated Sanger sequencer was used for sequencing, obtaining a mean coverage of 3.75×. Finally, 30 complete genome sequences were obtained. All primers and conditions used to sequence this virus are available in (Appendix A). GenBank accession number and related information of the sequences are shown in Table 1.

**Table 1**
Location and collection date of sequences used in the study.

| Sequence name | Collection date | Latitude | Longitude | GenBank accesion number |
|---|---|---|---|---|
| 38950_For_Pat | 12/31/00 | 24.083 | 58.375 | KX002194 |
| 39064_SL_Ped | 2/10/01 | 34.250 | 65.375 | KX002195 |
| 39111_Cor_Gpaz | 2/18/01 | 27.917 | 57.875 | KX002196 |
| AGLopez01 | 3/7/01 | 33.750 | 62.125 | KX002204 |
| AAreco01 | 3/13/01 | 34.417 | 59.625 | KX002203 |
| 129_LPam_Capital | 3/21/01 | 36.750 | 64.125 | KX002178 |
| ATLauquen01 | 3/27/01 | 35.917 | 62.875 | KX002205 |
| 258_ER_Diamante | 3/29/01 | 32.083 | 60.375 | KX002193 |
| 423_Cor_Union | 4/17/01 | 32.083 | 62.375 | KX002197 |
| 467_SFe_Irio | 4/20/01 | 32.917 | 61.375 | KX002198 |
| 548_LP_Guatrech | 4/26/01 | 37.583 | 63.375 | KX002199 |
| 820_Cor_Groca | 5/12/01 | 34.583 | 63.375 | KX002200 |
| 915_SEst_Riv | 5/15/01 | 29.917 | 62.125 | KX002201 |
| 973_SL_Pede | 5/15/01 | 33.583 | 65.375 | KX002202 |
| 1155_BsAs_MCh | 5/25/01 | 37.417 | 57.875 | KX002176 |
| 1293_BsAs_Indio | 5/28/01 | 35.417 | 57.625 | KX002179 |
| 1185_ER_Par | 5/28/01 | 31.583 | 59.875 | KX002177 |
| 1300_BS_As_Gguido | 6/3/01 | 36.583 | 57.875 | KX002180 |
| 1441_BsAs_Lflores | 6/16/01 | 35.583 | 59.125 | KX002181 |
| 1452_longMCH | 6/17/01 | 37.417 | 57.875 | KX002182 |
| 1514_LPam_HU | 6/17/01 | 38.083 | 64.125 | KX002183 |
| 1893_longMCH | 7/10/01 | 37.417 | 57.625 | KX002184 |
| 2027_Cor_Col | 7/20/01 | 31.083 | 63.875 | KX002186 |
| 2026_Tuc_Tuc | 8/7/01 | 26.9165 | 65.125 | KX002185 |
| 2071_LPam_Chali | 8/8/01 | 36.750 | 66.625 | KX002189 |
| 2063_Cor_Ccuat | 8/15/01 | 29.583 | 58.875 | KX002188 |
| 2059_BsAs_Hirig | 8/21/01 | 36.083 | 61.625 | KX002187 |
| 2084_StEst_OjoAgua | 9/1/01 | 29.583 | 63.125 | KX002190 |
| 2107_BsAs_Merced | 10/11/01 | 34.583 | 59.375 | KX002191 |
| 2118_LPam_Lmahu | 11/15/01 | 36.917 | 66.875 | KX002192 |

### 2.3. Phylogenetic analysis

Sequences of the FMDV A/Arg/2001 whole genome were aligned using MUSCLE (Edgar, 2004) software. Previous to estimation of the phylogeny, we assessed virus recombination using the Recombinant Identification Program (Siepel et al., 1995). Recombination analysis is used to estimate the most appropriate phylogenetic analysis, and to understand the potential mechanisms that the virus uses to evolve (in this case to estimate whether this virus that caused a massive outbreak, was a recombinant). We used whole genome of FMDV A strains A/Arg/00 (previous strain that caused a smaller outbreak in Argentina in 2000, GenBank: AY593782), as well as A24 and A25 (vaccine viruses used in previous years, GenBank: AY593767, AY593769) to compare similarities and to establish potential recombination. We also assessed recombination among all the sequenced viruses in this study. This method searches at different sizes of regions within sequences, identifies similarities between these regions and test for potential recombination.

To obtain the topology of the phylogenetic tree and evolutionary parameters, we used Bayesian Evolutionary Analysis by Sampling trees (BEAST v1.8) (Drummond and Rambaut, 2007). We initially estimated independently the phylogeny of each of the 12 protein-coding segments. For each protein-coding segment alignment, we selected the best codon partition scheme and substitution model using PartitionFinder software (Lanfear et al., 2012), based on the Bayesian information criterion. Using the selected substitution model and partition scheme we reconstructed the phylogeny using two different clock models for each protein coding segment: the strict clock and the uncorrelated lognormal relaxed clock (ucln) model. The (mean) clock rate for each protein was used to determine if different proteins differed in their evolution rate. The rates estimated were later used as informative initial values to set the multilocus analysis for the whole genome. For

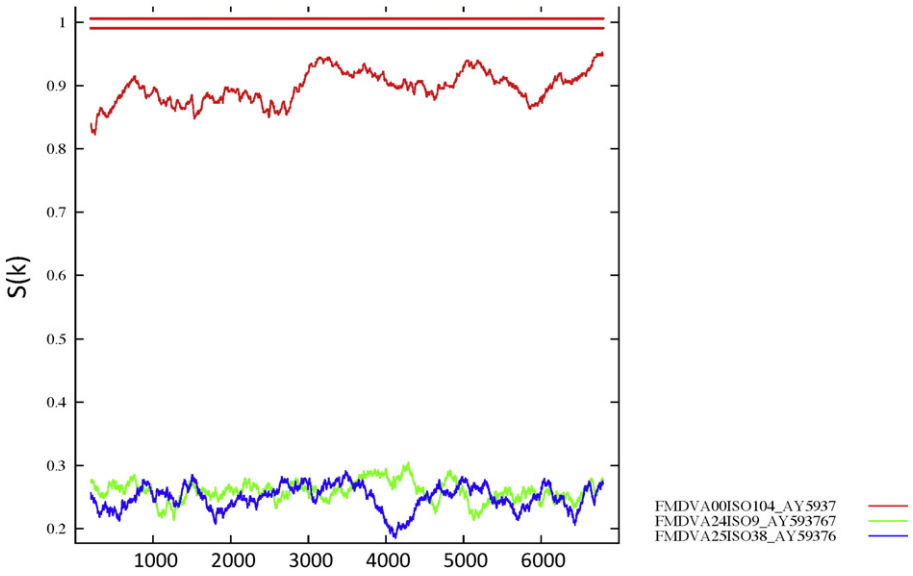


**Fig 1.** Recombination analysis. The x-axis represents the nucleotide position. The y-axis shows similarities of sequences A/Arg/01 (red), A24 (green), and A25 (blue) with the background sequence A/Arg/01. There was no recombination detected by this analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

independent protein coding regions analysis we used the coalescent Bayesian skyline tree prior because of its flexibility assumption on the population growth. We set the chain length at 10(Frost & Volz, 2013) iterations in BEAST and visualized the results of the iterations using Tracer v1.6 (Rambaut and Drummond, 2007).

After obtaining separate estimates for each protein-coding segment, we constructed a model for the whole genome. We constructed 8 different phylogenetic models to determine the best fit for the data. We used the substitution model partition scheme for each protein in all models. The first phylogenetic model consisted in linked clock model of proteins 2A, 2B and 2C, and for 3A, 3B and 3C, and made separate clock estimates for protein Leader pro, VP1, VP2, VP3 and VP4. The second model consisted in unlinking clock models of all protein coding regions (Leader pro, VP1, VP2, VP3, VP4, 2B, 2C, 3A, 3B, 3C and 3D). In the third phylogenetic model, clock estimates were made separately for all proteins except for 2B and 2C, which were linked. The fourth model consisted in linking all proteins partitions into one clock model. Strict clock was used for all the first 4 models. The remaining 4 models consisted in the same scheme described above, but using ucln instead of strict clock models. Because there was no evidence of recombination, all protein coding segments were linked to one tree model; the Bayesian Skygrid tree prior, which allows for the estimation of the effective viral population size for multilocus sequence data (Gill et al., 2013). Each model was run for 10 (Frost & Volz, 2013) iterations or until all parameters reached >200 effective sample size (ESS). Mixing and convergence were assessed using Tracer v1.6. To compare models we used stepping-stone and path sampling methods (Baele et al., 2012). The trees generated before reaching convergence of the Markov Chain Monte Carlo chains were burned out. The maximum clade credibility tree (MCC) was annotated and visualized in Figtree (Rambaut, 2009). Computational resources available in CIPRES Science Gateway were used to run all analyses (Miller et al., 2010).

### 2.4. Phylogeographic analysis

Once the best model for the multilocus data was selected, we estimated the phylogeny using a spatial diffusion relaxed random walk model to reconstruct continuous traits. Latitude and longitude in decimal coordinates of each of the sequenced viruses were used a bivariate continuous traits (Lemey et al., 2010). This relaxed random walk approach allows different rates of spatial diffusion among branches. The estimation of spatial spread of the virus can be achieved using the known location of the sequences in the tips of the phylogeny and reconstructing the ancestors' location. The maximum clade credibility (MCC) tree was annotated and visualized in Figtree (Rambault, 2006–2009). Convergence and mixing of the simulations was assessed using Tracer (Rambaut and Drummond, 2007). To visualize the estimated spatial and temporal diffusion of the virus, the animated online platform Phylowood was used (Landis and Bedford, 2014).

### 2.5. Effective viral population size, epidemic curve and within-herd transmission rate

To approximate genetic diversity in time, we reconstructed the demographic history of the virus by plotting the effective viral population size in time, estimated with the Bayesian Skygrid model for multilocus sequence analysis, using Tracer v1.6. We also reconstructed the demographic history using only VP1 protein coding region with a Bayesian Skyline plot.

We plotted the viral population size, the epidemic curve and within herd transmission coefficient obtained from previous works (Brito et al., 2011; Perez et al., 2004b).

The objective of this analysis is to determine if, during the course of an epidemic, relevant information about the epidemic and prediction of near-future outcomes can be obtained by analyzing viral sequences, specifically, changes in viral population size.

We compared the relationship between epidemiological variables (incidence and within herd transmission rates) estimated by the viral genetic diversity inferred by VP1 and by whole genome analysis.

First we plot the viral population size and the number of new herds infected in time (epidemic curve) in time, to assess potential relationship between these two variables. We also wanted to determine if viral diversity was related to within-herd transmission, which directly influence the number of new herds infected (the probability of transmission between herds will depend, in part, on how many animals become infected within a herd). Within-herd transmission is a result of the virus interacting with susceptible hosts by acquiring properties that allow the virus to become more or less transmissible, and that may also be reflected in overall viral genetic diversity.

We compared the relationship between these epidemic parameters (new herds infected and within-herd transmission rates) and viral genetic diversity inferred by VP1 coding region and by whole genome

analysis. The importance of this comparison relies on the fact that sequencing of VP1 coding region is the most frequently used technique in reference diagnostic laboratories. We wanted to determine if using whole genome analysis might provide more accurate information to assess the epidemic in time.

## 3. Results

### 3.1. Phylogenetic analysis

Whole genome of FMDV A/Arg/01 sequences was 8142 nucleotide-long including UTR and polyC regions, and 6999 nucleotide-long open-reading frame. There was no evidence of recombination among sequences included in this study or with previous A/Arg/00 or vaccine strains (Fig 1).

The best partition scheme for each protein is shown in Appendix B Table 1. Best partition schemes of most proteins suggested that codon position 3 should be estimated as a partition different to that of codon positions 1 and 2. Protein segment 2A from sequenced viruses only had heterogeneity in only 3 nucleotide-sites, and the amino acid sequences were identical. Because this protein was so conserved, it was not included in further phylogenetic analysis.

The clock rate estimated for each protein-coding segment is shown in Appendix B Table 2. In general clock rates estimated were similar using ucln and strict clock. There were differences between the estimated rates of the proteins: VP1, VP4, 3A and 3C had higher substitution rates, while VP2 and 2B had the lowest rates.

Model comparison of whole genome sequences, using path sampling and stepping stone (Appendix B Table 3) determined that the best phylogenetic model was the approach that linked all protein coding regions into one clock estimate using an UCLN. The MCC tree with obtained the node's 95% highest posterior density (HPD) credibility intervals (bars), and the nucleotide substitution rates (nucleotide substitution/site/year) represented by a color gradient are shown in Fig 2. Posterior values of nodes showed that this topology has well supported estimates.

The MCC tree shows three main clades (named as clades A, B and C in Fig. 2) that evolved during the epidemic. The time most recent common ancestor (tMRCA) for all A/Arg/01 viruses was estimated at May-2000 (95% HPD Dec. 1999–Aug. 2000), while tMRCA of all A/Arg/01 and the A/Arg/00 virus (GenBank: AY593782) was estimated in May 1994 (95% HPD Feb. 1993–May 1995).

The overall number of substitutions per site per day of FMDV A/Arg/01 was 8.35E-3. The number of substitutions per site per day was: clade A 8.295E-3 (95%HPD 6.07E-3, 1.03E-2), clade B 8.29E-3 (95%HPD 6.29E-3, 1.04E-2), and clade C 8.94E-3 (95%HPD 6.14E-3, 1.23E-2).

### 3.2. Phylogeographic analysis

The animated diffusion of FMDV is shown in (Video 1). The location of all sequence's ancestor was located ~300 km northwest to Buenos Aires province, where a high number of bovine herds are concentrated. The virus spread in all directions from this initial ancestor. Clade B (which in phylowood tree is the upper clade) spread mostly southeastern and southwestern directions, while clade A (the second clade from top to bottom) tend to spread in all directions and had the largest geographic coverage. Clade C tends to spread in northeastern and northwestern directions (Fig. 4).

### 3.3. Effective viral population size, epidemic curve and transmission rate

The effective viral population size, an approximation of viral genetic diversity, estimated using Bayesian Skyline tree prior (VP1) and Bayesian Skygrid (whole genome) is shown in Fig 3. Day 0 was considered as the first day in which lineage A/Arg/01 was sampled from an outbreak (31-Dec.-2000). Initial high within transmission rate observed before day 100, was associated with increasing genetic viral diversity. The number of outbreaks per week had an initial growth between days 50 and 100 after the beginning of the epidemic, that coincided with the

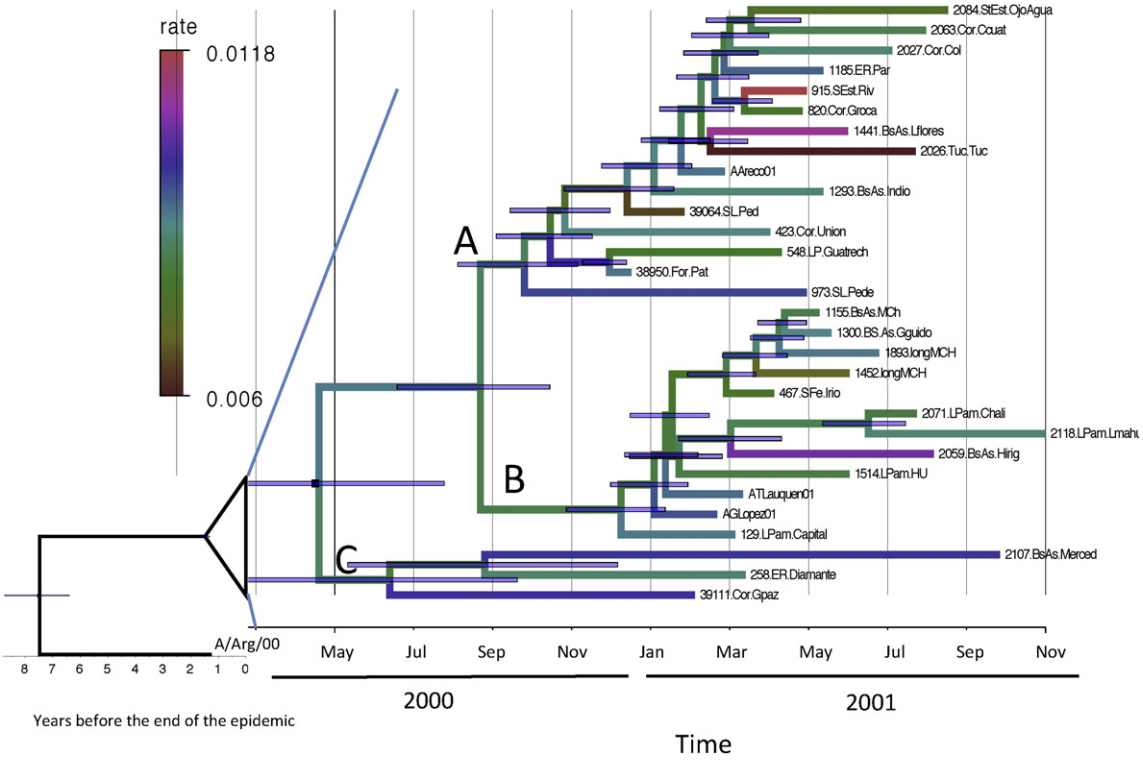


**Fig 2.** Phylogenetic tree with 95% credibility interval. Color of branches represent the substitution rate per site per day. The horizontal axis represents time. The three main genetic clades are indicated by letters A, B, and C. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

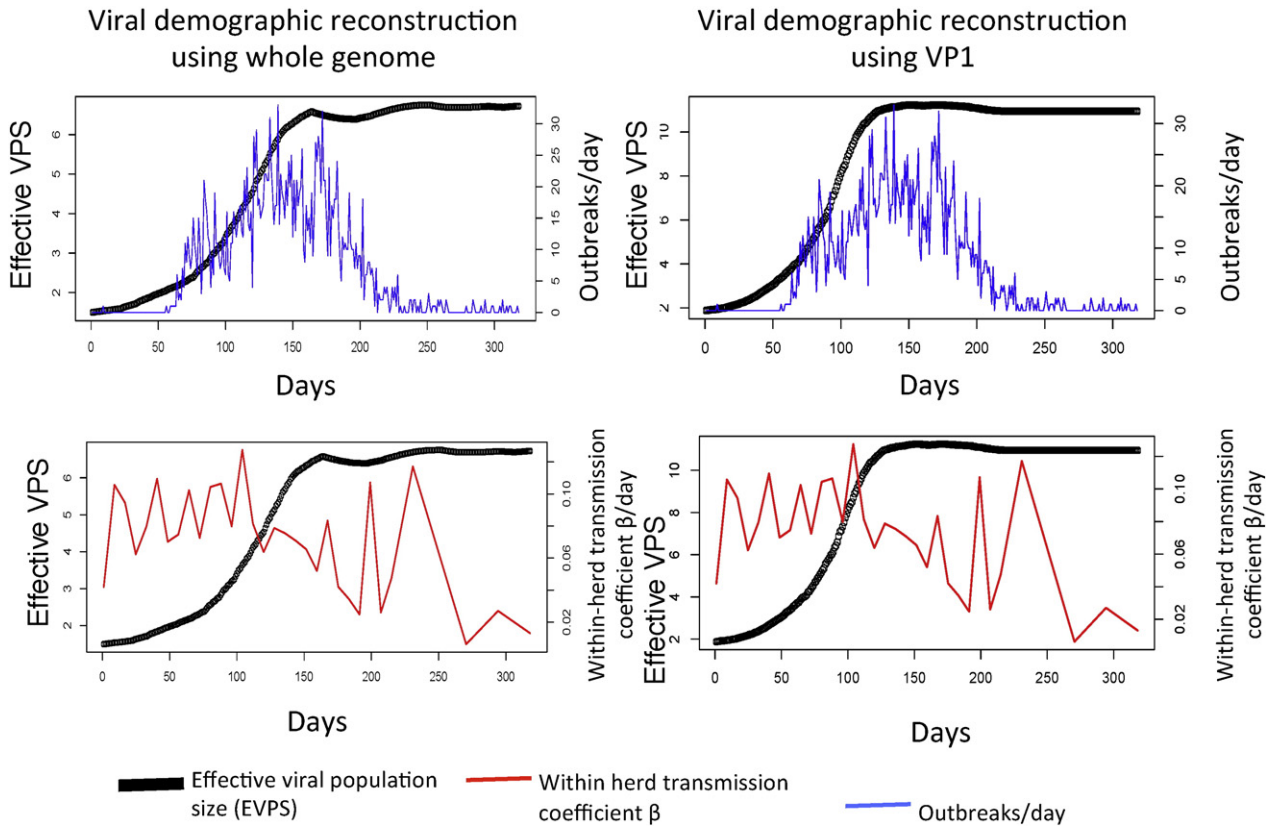


**Fig 3.** Mean viral population size, umber of outbreaks per day, and intra-herd transmission coefficient. Panels in the left area of the figure correspond to the analysis made using whole genome for demographic reconstruction (Bayesian Skygrid), while panels in the right refer to demographic reconstruction using only VP1-coding sequence (Bayesian skyline). X-axes refer to time (day 0 correspond to the first A/Arg/01 isolate collection date, 31-Dec.-2000). Left y-axes in all graphs indicate the scale of the effective viral popularion size (black line). Right y-axes on the graph indicate the number of outbreaks per day (blue lines-top panels) or the within-herd transmission coefficient (red lines-bottom panels). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

viral population size estimated by whole genome analysis, but was preceded by the increase genetic diversity estimated by VP1 analysis.

The peak of the epidemic was ~day 120. The peak of viral genetic diversity estimated by VP1concided with the peak of cases in the epidemic curve, and was later reached by the genetic diversity estimated using whole genome analysis. The viral population size reached a plateau and was maintained high until the end of the epidemic, in contrast with the number of cases indicated by the epidemic curve, which decreased after day ~170.

## 4. Discussion

Our phylogeographic analysis estimated that the FMDV ancestor originating the epidemic in 2001 diverged into 3 genetic clades during the epidemic. The diffusion model used to determine the spread of FMDV A/Arg/01, showed an initial ancestor virus in a central area of the country, from where it spread simultaneously into multiple directions.

The most recent common ancestor for all FMDV A/Arg/01 viruses was predicted to have emerged almost 7 years before the beginning of the 2001 epidemic (Fig. 2). Although Argentina suffered two consecutive epidemics in 2000 and 2001, before this analysis there was still a debate regarding the possibility that the two epidemics were caused by the same virus. However, our analyses strongly suggest that the FMDVs that caused the first outbreaks were not recently related.

The mechanism by which the 2001 virus was introduced in the country remains unknown, but an illegal animal movement from the north is a likely scenario. However, the number of cases was maintained low, until it reached a region with high intensive livestock farming. In 2001, the provinces with the highest cattle population, and the highest number of intensive production farm were Buenos Aires, Cordoba and Santa Fe. Our analysis suggests that the virus did not begin to spread widely until it reached the predicted ancestor within a more central location in Santa Fe province (Video 1), where there is frequent animal movement between intensive production systems allowed a high number of transmission events to occur (Perez et al., 2004b).

Recombination of FMDV has been previously reported, mostly studied in endemic places where different subtypes and lineages cocirculate in a geographic area. FMDV recombination is not frequently found in capsid protein regions, but frequently found in conserved non-structural genes, so detection of recombinants requires whole genome sequencing (Haydon et al., 2004; Jackson et al., 2007; Jamal et al., 2011). We did not find evidence of recombination of FMDV A/Arg/01 with FMDV A24 strain, contemporarily used as vaccine, or the A/Arg/00. Although recombination has been described for FMDV serotype viruses, there was no evidence that FMDV A/Arg/01 acquire new antigenic or pathogenic characteristics through this mechanism. Previous studies have attributed the pathogenic characteristics of the A/Arg/01 virus to a few amino-acid changes within the capsid-coding region (Garcia-Nunez et al., 2010).

It has been previously shown that changes in viral population size determined using a coalescent model, reflect the number of lineages/strains present in time and is associated with the number of new infections (i.e., incidence) (Frost and Volz, 2010; Holmes et al., 1995). We observed a high association of the viral population size and the epidemic curve during the phase of exponential growth of the epidemic (Fig 3). However, after reaching a plateau, the viral population size remained high while the number of new cases decreased. Using VP1 reached faster a plateau (compared to using whole genome) following the exponential phase of the epidemic. Interestingly, when the number of cases

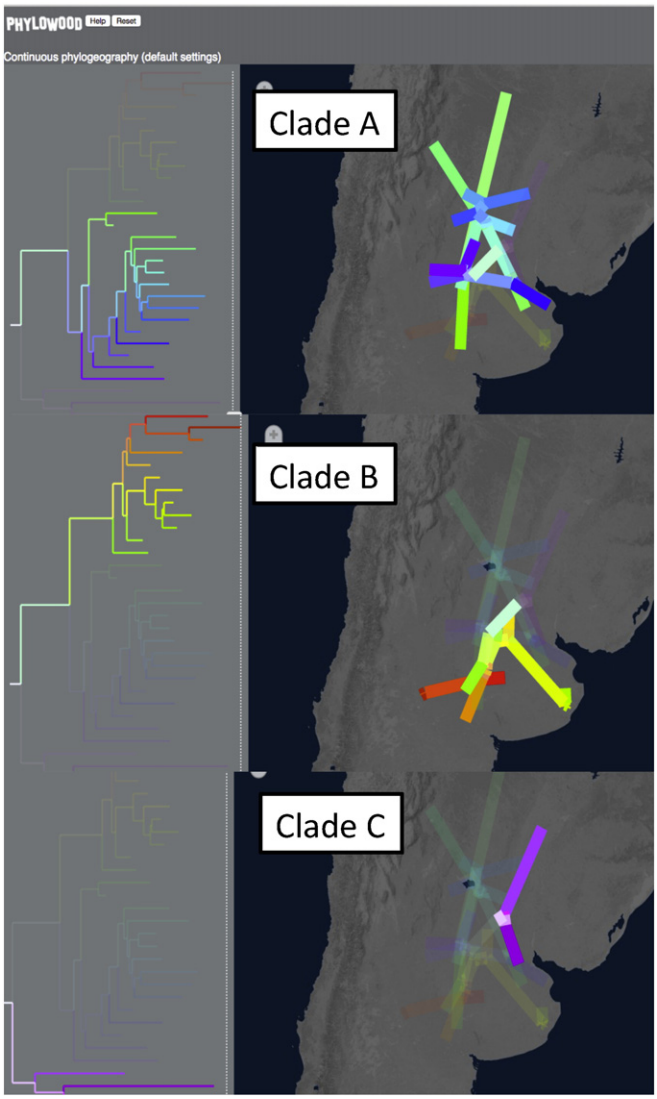


**Fig. 4.** Geographic spread of the three major genetic clades of A/Arg/01 virus during the epidemic inferred by continuous diffusion phylogeographic models.

declined, the viral population size remained constant after reaching the plateau, which may be attributed to a sudden bottleneck, changes of generation time during the epidemic, or due to low number of viruses obtained by the end of the epidemic, when few infected herds remained (de Silva et al., 2012; Frost & Volz, 2013). A similar observation has been documented in a recent study of FMDV O/Cathay strain through a 10-year period, where the viral population size decreased only few years after the number of reported outbreaks decreased (Di Nardo et al., 2014), however, outbreaks records used were only those reported to the OIE.

Assessing the growth of the epidemic using molecular tools could be extremely helpful during an epidemic. By reaching a plateau measuring viral population size, decisions makers can approximate that the exponential phase of the disease is coming to an end, and therefore predict that the number of cases will be decreasing in the future, so control measures are taken. This real-time assessment can help guiding health authorities through the course of an outbreak. However, further validation of these methods should be made with different viral strains, and with data from other epidemic scenarios.

During phylogenetic analysis, we found that 2A protein was extremely conserved among all sequences sampled, with an identical amino acid sequence in all viruses sampled. Negative selection may have affected this protein because of advantageous characteristics that it may have conferred to the virus. However, this finding needs to be further investigated.

Estimation of epidemic growth and genetic diversity was previously used in a study to assess the evolution of pandemic influenza. In contrast to our study, they found that the viral population size decreased before the peak of the epidemic and discussed that this finding may be attributed to bias in sampling and reporting (de Silva et al., 2012). In contrast, incidence during the 2001 FMD outbreak in Argentina was well documented throughout the entire epidemic, because it was mandatory to report cases as soon as they occurred, hence, providing an acceptable scenario to assess virus phylodynamics in animal populations. The outbreak only declined by reducing the number of susceptible hosts using vaccination and by restricting animal movement (Perez et al., 2004a). The relevance of this finding is that epidemiologic features can be inferred using genetic data. In this case, the deceleration in the growth of the viral population size coincided with the peak of the epidemic. This type of analysis could be applied to future epidemics with incomplete epidemic data, to infer the phase of the epidemic using viral population size and evaluate the effectiveness of control measures (i.e., vaccination).

## 5. Conclusions

In conclusion phylogeographic and genetic analysis of FMDV A/Arg/01 using whole genome data, revealed important aspects of the 2001 Argentine epidemic. The virus was widely spread from a central area of the country, where surveillance should be strengthened in future surveillance. We also observed a high association between the growing phase of the epidemic curve and the viral population size. This observation might be a useful tool to predict effectiveness of control measures early in the course of viral epidemics.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2016.03.026.

## Author contributions

Conceived and designed the study: AP, GK, BB and LR. Viral sequencing: GK and SC. Data compilation: SC, CPB and BB. Sequence analysis: BB, GK and SC. Wrote the paper: BB, AP, GK, SC, LR, and CPB.

## Acknowledgments

Financial support: This paper has been funded in part by grants from the USDA-ARS Project “Developing predictive tools for prevention of Foot-and-Mouth Disease” (58-1940-7-124F), CEEZAD, and the MnDrive.

## References

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167.

Brito, B.P., Perez, A.M., Cosentino, B., Rodriguez, L.L., Konig, G.A., 2011. Factors associated with within-herd transmission of serotype A foot-and-mouth disease virus in cattle, during the 2001 outbreak in Argentina: a protective effect of vaccination. Transbound. Emerg. Dis. 58, 387–393.

de Silva, E., Ferguson, N.M., Fraser, C., 2012. Inferring pandemic growth rates from sequence data. Journal of the Royal Society, Interface/the Royal Society 9, 1797–1808.

Di Nardo, A., Knowles, N.J., Wadsworth, J., Haydon, D.T., King, D.P., 2014. Phylodynamic reconstruction of O CATHAY topotype foot-and-mouth disease virus epidemics in the Philippines. Vet. Res. 45, 90.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Frost, S.D., Volz, E.M., 2010. Viral phylodynamics and the search for an ‘effective number of infections’. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 365, 1879–1890.

Frost, S.D., Volz, E.M., 2013. Modelling tree shape and structure in viral phylodynamics. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 368, 20120208.

Garcia-Nunez, S., Konig, G., Berinstein, A., Carrillo, E., 2010. Differences in the virulence of two strains of foot-and-mouth disease virus serotype A with the same spatiotemporal distribution. Virus Res. 147, 149–152.
Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713–724.
Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303, 327–332.
Haydon, D.T., Bastos, A.D., Awadalla, P., 2004. Low linkage disequilibrium indicative of recombination in foot-and-mouth disease virus gene sequence alignments. J Gen Virol 85, 1095–1100.
Holmes, E.C., Nee, S., Rambaut, A., Garnett, G.P., Harvey, P.H., 1995. Revealing the history of infectious disease epidemics through phylogenetic trees. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 349, 33–40.
Jackson, A.L., O'Neill, H., Maree, F., Blignaut, B., Carrillo, C., Rodriguez, L., Haydon, D.T., 2007. Mosaic structure of foot-and-mouth disease virus genomes. J. Gen. Virol. 88, 487–492.
Jamal, S.M., Ferrari, G., Ahmed, S., Normann, P., Belsham, G.J., 2011. Molecular characterization of serotype Asia-1 foot-and-mouth disease viruses in Pakistan and Afghanistan; emergence of a new genetic group and evidence for a novel recombinant virus. Infect. Genet. Evol. 11, 2049–2062.
Konig, G.A., Palma, E.L., Maradei, E., Piccone, M.E., 2007. Molecular epidemiology of foot-and-mouth disease virus types A and O isolated in Argentina during the 2000–2002 epizootic. Vet. Microbiol. 124, 1–15.
Landis, M.J., Bedford, T., 2014. Phylowood: interactive web-based animations of biogeographic and phylogeographic histories. Bioinformatics 30, 123–124.
Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Lemey, P., Rambaut, A., Welch, J.J., Suchard, M.A., 2010. Phylogeography takes a relaxed random walk in continuous space and time. Mol. Biol. Evol. 27, 1877–1885.
Mattion, N., Goris, N., Willems, T., Robiolo, B., Maradei, E., Beascoechea, C.P., Perez, A., Smitsaart, E., Fondevila, N., Palma, E., De Clercq, K., La Torre, J., 2009. Some guidelines for determining foot-and-mouth disease vaccine strain matching by serology. Vaccine 27, 741–747.
Mattion, N., Konig, G., Seki, C., Smitsaart, E., Maradei, E., Robiolo, B., Duffy, S., Leon, E., Piccone, M., Sadir, A., Bottini, R., Cosentino, B., Falczuk, A., Maresca, R., Periolo, O., Bellinzoni, R., Espinoza, A., Torre, J.L., Palma, E.L., 2004. Reintroduction of foot-and-mouth disease in Argentina: characterisation of the isolates and development of tools for the control and eradication of the disease. Vaccine 22, 4149–4162.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE). New Orleans, LA, USA, pp. 1–8.
OIE, World Organization for Animal Health, 2009. Foot and Mouth Disease — Technical Disease Card. Available at: http://www.oie.int/en/animal-health-in-the-world/technical-disease-cards/.
Perez, A.M., Konig, G., Spath, E., Thurmond, M.C., 2008. Variation in the VP1 gene of foot-and-mouth disease virus serotype A associated with epidemiological characteristics of outbreaks in the 2001 epizootic in Argentina. J. Vet. Diagn. Investig. 20, 433–439.
Perez, A.M., Ward, M.P., Carpenter, T.E., 2004a. Control of a foot-and-mouth disease epidemic in Argentina. Prev. Vet. Med. 65, 217–226.
Perez, A.M., Ward, M.P., Carpenter, T.E., 2004b. Epidemiological investigations of the 2001 foot-and-mouth disease outbreak in Argentina. Vet. Rec. 154, 777–782.
Rambault, A., 2006–2009. Fig tree. Tree Figure Drawing Tool, Version 1.3.1.
Rambaut, A., 2009. Molecular evolution, phylogenetics and epidemiology: FigTree. Available at: http://tree.bio.ed.ac.uk/software/figtree/.
Rambaut, A., Drummond, A., 2007. Tracer Version 1.4. Computer Program and Documentation Distributed by the Author. Available at: http://tree.bio.ed.ac.uk/software/tracer/.
Siepel, A.C., Halpern, A.L., Macken, C., Korber, B.T., 1995. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. AIDS Res. Hum. Retrovir. 11, 1413–1416.
Volz, E.M., Koelle, K., Bedford, T., 2013. Viral phylodynamics. PLoS Comput. Biol. 9, e1002947.