



# Insights into the origin and diversification of bovine viral diarrhea virus 1 subtypes

Matheus N. Weber<sup>1</sup> · Jonas M. Wolf<sup>2,3</sup> · Mariana S. da Silva<sup>3</sup> · Ana Cristina S. Mosena<sup>3</sup> · Renata F. Budaszewski<sup>3</sup> · Vagner R. Lunge<sup>2</sup> · Cláudio W. Canal<sup>3</sup>

Received: 1 June 2020 / Accepted: 30 October 2020 / Published online: 3 January 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, AT part of Springer Nature 2021

## Abstract

In this study, we performed phylogenetic and evolutionary analysis on bovine viral diarrhea virus 1 (BVDV-1) sequences to investigate the origin and temporal diversification of different BVDV-1 subtypes. Dated phylogenies using the complete polyprotein sequence were reconstructed, and the time of the most recent common ancestor (tMRCA) was estimated. The results demonstrated that BVDV-1 subtypes clustered into two phylogenetic clades, where the predominant subtypes worldwide grouped together. In the temporal analysis, the tMRCA of BVDV-1 was 1336, and the diversification into different subtypes appears to have occurred around 363 years ago. The present results help to elucidate the origins of BVDV-1 subtypes and the dynamics of ruminant pestiviruses.

Bovine viral diarrhea (BVD) is a globally distributed cattle disease caused by ruminant pestiviruses that is associated with high economic losses [1]. The genus *Pestivirus* within the family *Flaviviridae* was recently reclassified into 11 species, and bovine viral diarrhea virus 1 (BVDV-1) was assigned to the species *Pestivirus A* [2]. Pestiviruses have a positive-sense single-stranded RNA genome that contains one open reading frame (ORF) flanked by untranslated regions (UTRs) at the 5' and 3' ends [3]. The ORF encodes a polyprotein that is processed into 12 polypeptides [3]. The classification of BVDV-1 into 21 subtypes (a through u) is primarily based on phylogenetic analysis of

partial N<sup>pro</sup> sequences, and the N<sup>pro</sup> tree topology is reflected in trees constructed using the complete genome sequences [4–6]. The current knowledge of the circulating subtypes is important for epidemiological information, development of appropriate diagnostic tools, and introduction of control programs [7].

The high mutation rate is a key driver of quasispecies distribution among pestiviruses [8, 9]. The resulting swarm of genetically unique viruses increases the ability of pestiviruses to adapt to the changing environmental conditions within a single host [10]. This extensive diversity is important for control methods, such as vaccine development, and it is useful for molecular epidemiology and phylodynamics studies.

The rapid diversification of RNA genomes over relatively short periods of time enables precise inference of phylogenetic relationships over time [10–13]. Moreover, in the absence of archeological samples, the age of a virus may be calculated using the concept of a molecular clock [14]. The analysis of the time of the most recent common ancestor (tMRCA) has been used to infer viral epidemiology, with the determination of the historical fluctuations in population dynamics and spatial dispersion of ancestral viral strains to their contemporary locations in a geographical area [10, 12, 15]. Major geographical movements of the virus could be indicative of key migratory events or transport routes followed by the host species.

Handling Editor: Patricia Aguilar.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00705-020-04913-y>.

✉ Matheus N. Weber  
matheusweber@gmail.com

<sup>1</sup> Laboratório de Microbiologia Molecular, Universidade Feevale, Novo Hamburgo, Rio Grande do Sul, Brazil

<sup>2</sup> Laboratório de Diagnóstico Molecular, Universidade Luterana do Brasil (ULBRA), Canoas, Rio Grande do Sul, Brazil

<sup>3</sup> Laboratório de Virologia, Faculdade de Veterinária, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil

BVDV-1 is the most widely dispersed ruminant pestivirus worldwide, where subtypes 1a and 1b are the most frequent and best-studied subtypes [1]. However, some uncommon BVDV-1 subtypes are also epidemiologically important for understanding the evolution of this virus. Here, we report the phylogenetic and evolutionary analysis of complete BVDV-1 polyprotein sequences to elucidate the origin and diversification of this pestivirus.

Complete polyprotein sequences of 43 BVDV-1 strains were downloaded from the GenBank database (Supplement 1). Genome sequences previously reported to have undergone homologous recombination events [16] were intentionally excluded from the analysis. Moreover, insertions observed in cytopathic strains were also removed. The dataset was aligned using MAFFT [17]. Putative recombination events were verified using the Recombination Detection Program version 4 (RDP4) software [18] with the default settings using the algorithms RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, 3Seq, and LARD. Putative recombinant events were considered significant when a *P*-value of 0.01 or less was obtained for the same event using four or more algorithms. Sequences with evidence of recombination events were excluded from the tMRCA analysis.

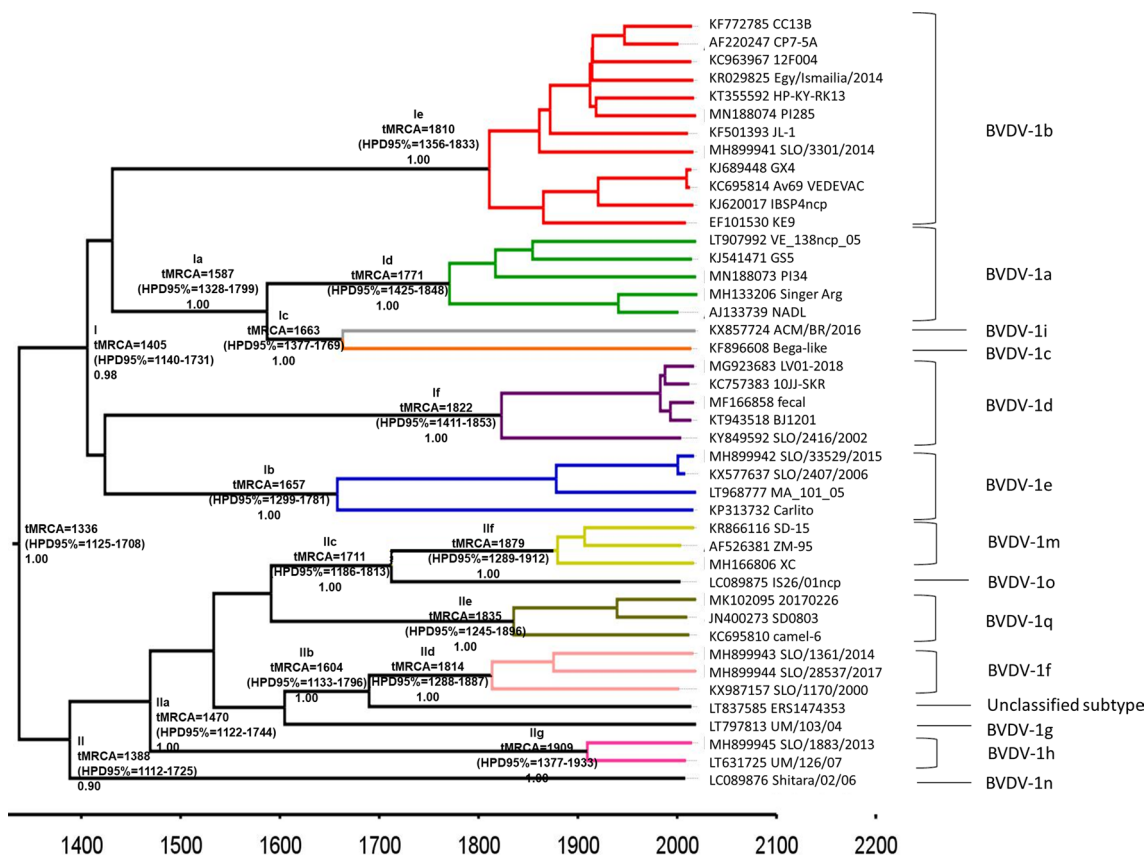
The best-fitting nucleotide substitution (GTR+F+I+G4) model was selected using the hierarchical likelihood ratio, Akaike information criterion, and Bayesian information criterion tests with Model Finder on the IQ-TRE webserver [19]. A maximum-likelihood (ML) phylogenetic tree of the BVDV-1 polyprotein gene sequences was inferred using the best-fitting model using the IQ-TRE web server (<http://iqtree.e.cibiv.univie.ac.at/>). Statistical support for internal branches in the phylogeny was assessed by bootstrapping (1000 replicates) and the approximate likelihood ratio test (aLRT) [20]. We used this tree to obtain root-to-tip regressions in TempEst v1.5 [21] by selecting the root position that maximized the correlation coefficient.

To estimate the tMRCA of each BVDV-1 subtype, dated phylogenies using the complete polyprotein region were reconstructed using the Bayesian Monte Carlo Markov chain (MCMC) method implemented in BEAST v1.10 [22]. All tips were dated with their year of collection. Implementing strict and relaxed molecular clocks, MCMC analysis was performed, and the maximum likelihood estimation (MLE) of the resulting trees were compared using a Bayes factor (BF) to select the best model and parameter values. For each run of  $5.0 \times 10^8$  of MCMCs, the marginal likelihood was estimated using path sampling (PS) and stepping stone (SS) methods, and the resulting BF (ratio of marginal likelihoods) was used to select the best-fitting clock/demographic model (Clocks: Strict, Uncorrelated Lognormal, Random Local, and Fixed Local; Priors: Costant Size, Exponential Growth, Logistic Growth, Expansion Growth, Bayesian Skygrid, GMFR Bayesian

Skyride, Bayesian Skyline, Extended Bayesian Skyline Plot, and Yule Procces, Calibrated Yule). Both SS and PS estimators indicated the relaxed clock (BF = 21.8) to be the best-fitted model to the datasets under analysis. Also, BF analysis showed that the relaxed clock fit the data significantly better than the strict clock (2lnBF between the strict and relaxed clock was 691.39 in favor of the second). Under the relaxed clock, the BF analysis showed that the Bayesian skyline plot was better than other models (2lnBF > 100). As recommended, the Bayesian analysis assumed an uncorrelated lognormal + SDR06 + skyline plot (Supplement 2). The mean evolutionary rate for the BVDV-1 polyprotein was calculated using the alignments applied in the present study. The MCMC chain was run for  $5.0 \times 10^8$  chain steps, and the convergence was evaluated in TRACER v1.5, excluding an initial 10% for burn-in. Maximum clade credibility trees (MCC) were summarized using TreeAnnotator v1.8.3, and the resulting tree was visualized with FigTree v1.4.3. Uncertainty in parameter estimates is reflected in the 95% highest probability density (HPD) values.

A root-to-tip vs. divergence plot of the full dataset showed a correlation between sampling time and genetic distance to the root of the ML tree of the available sequences (correlation coefficient = 0.22; R squared = 4,9084E-2) (Supplement 3), suggesting a moderate temporal signal and the possibility of calibrating a reliable molecular clock despite the limited variability in the year of collection of sequences. The low R-squared value may be explained by the low amplitude of the dates in available BVDV-1 sequences. However, the positive R-squared value allowed the present analysis to be applied in the evolutionary history of BVDV-1. Future studies when more BVDV-1 sequences are available may provide more precise information about the evolutionary origin of BVDV-1.

A Bayesian phylogenetic tree was constructed using complete BVDV-1 polyprotein sequences (Fig. 1). The phylogenetic tree demonstrated the existence of two main clades, BVDV-1 I and BVDV-1 II, supported by posterior probability values of 0.98 and 0.90, respectively. BVDV-1 I included subtypes 1a, 1b, 1c, 1d, 1e, and 1i, while BVDV-1 II included subtypes 1f, 1g, 1h, 1m, 1n, 1o, 1q and one unclassified subtype detected in Italy. The phylogenetic tree showed the same general topology and BVDV-1 subtype clustering presented in previous reports [4–6]. Notably, all genomes of the same subtype clustered together as expected, including the sequence with unclassified subtype detected in Italy. Despite clustering in the BVDV-1f and -1g branch, they did not exhibit a high nucleotide sequence similarity to these other genomes. BVDV-1f and -1g are frequently reported in Italy [1], and these sequences of an unclassified subtype appear to have the same common ancestor as BVDV-1f and -1g.



**Fig. 1** Time-scaled maximum clade credibility tree constructed by Bayesian analysis of BVDV-1 polyprotein gene sequences. The time of the most recent common ancestor (tMRCA), 95% HPD (highest

posterior probability), and significant posterior probability are shown at the nodes. The scale bar indicates the timeline.

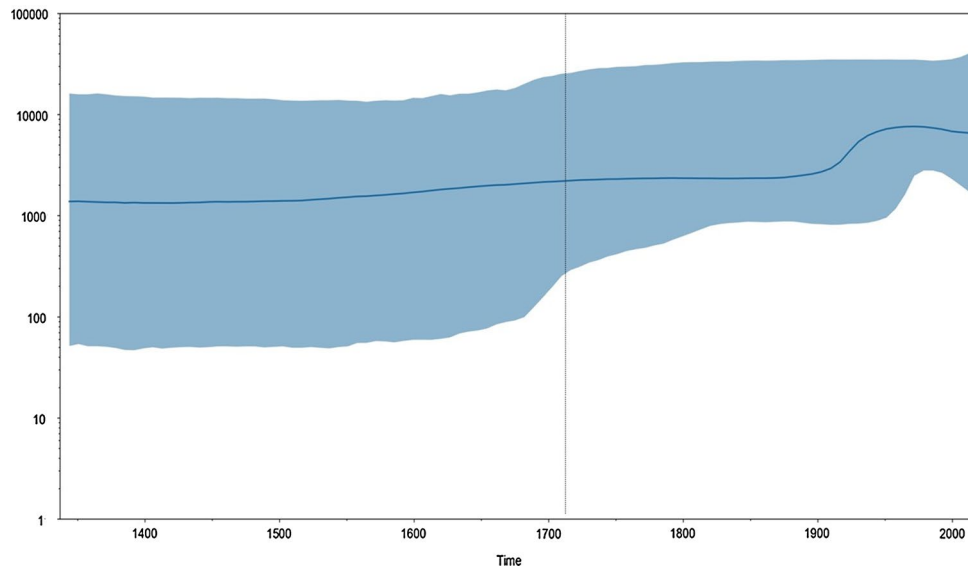
The BVDV-1 substitution rate was  $2.9 \times 10^{-4}$  (95% HPD:  $1.52E-5$ – $7.11E-4$ ) substitutions per site per year. An additional analysis of sequences that grouped in clade I was performed (BVDV-1a, -1b, -1c, -1d, -1e and -1i) and the tMRCAs were similar to those shown in Fig. 1, reinforcing the data (Supplement 4).

Temporal analysis showed that all BVDV-1 strains seem to have one common ancestor dating back to about 1336 (95% HPD: 1125–1708) (Fig. 1), a much earlier date than 1802 (95% HPD: 1522–1939), as reported previously [11]. It was also estimated that all of the BVDV-1 subtypes in circulation worldwide originated approximately 363 years ago (Fig. 1). The complete genome analysis reported here probably increased the statistical confidence to generate more robust data for this analysis, as highlighted by the posterior probability values observed. Partial sequences comprising  $N^{pro}$  and glycoprotein E2 were also analyzed using this methodology, but despite their similar topology, their confidence intervals had larger amplitudes, and they were therefore not included in the results.

The two most common BVDV-1 subtypes worldwide are BVDV-1a and -1b [1]. In our analysis, their common

ancestors emerged earlier than the common ancestor of the other BVDV-1 subtypes, except for BVDV-1e (Fig. 1). Our results suggested the emergence of the BVDV-1a common ancestor around 1771 (95% HPD: 1425–1848) and of BVDV-1b around 1822 (95% HPD: 1411–1853). It is important to highlight that a previous study performed in Canada indicated a recent emergence of BVDV-1a and -1b (1930 and 1955, respectively), but with a limited number of samples from only one country [15], whereas in the present work, we analyzed a larger number of BVDV-1 sequences from different regions of the world. The earlier emergence of these two BVDV-1 subtypes could be attributed to their spread worldwide in cattle accompanying human emigrations to the New World in the sixteenth and seventeenth centuries [23]. The same inference may be made for BVDV-1d, which emerged at nearly the same time as BVDV-1a and -1b (tMRCA: 1810; 95% HPD: 1356–1833) and is also present in the Americas, Europe and Africa [1].

BVDV-1e putatively emerged around 1657 (95% HPD: 1299–1781), temporally close to BVDV-1a and -1b. However, this subtype is mainly restricted to European countries, especially France and Switzerland, where it is the most



**Fig. 2** Bayesian skyline plot (BSP) of BVDV-1 polyprotein gene sequences obtained from GenBank. The effective number of infections is shown on the y-axis. The timeline is shown on the x-axis. The

colored area corresponds to the 95% credibility intervals of highest probability density (95% HPD). The vertical line indicates the lower 95% HPD (dotted) of the tree root.

prevalent BVDV-1 subtype. There have been a few reports of this subtype outside of Europe, but only one case has been reported in Brazil [1].

Fewer sequences were available for the other subtypes analyzed, making similar inferences difficult. Information about these subtypes is restricted to a few reports from Europe (subtypes 1f, 1g and 1h) and Asia (subtypes 1m, 1n, 1o and 1q). Other examples can be observed in Switzerland and Austria, where BVDV-1h is the subtype most frequently observed in cattle [5, 24], which has been maintained by a lack of cattle imports and trade barriers within the country [25]. Additionally, these BVDV-1 subtypes may have evolved independently in these regions, resulting in considerable genetic and antigenic differences due their unique virus-host interactions.

Phylogenetic analysis showed that the overall effective population size was constant between 1336 and 1900, but it grew in the first half of the 20th century. Thereafter, the overall BVDV-1 population seems to have remained constant (Fig. 2). This might be related to an increase in cattle herd sizes since the nineteenth century, which enabled virus to have more host contact, resulting in an increase in phylodynamics.

The absence of virus samples that date back hundreds of years makes it difficult to precisely calibrate molecular clocks, which decreases the precision of the technique [26]. Moreover, in most instances, some molecular clocks have to be calculated using virus samples obtained over a relatively short period. Therefore, in the present work,

we chose to focus on BVDV-1 because there are more data available than for other cattle pestiviruses. Moreover, BVDV-1 is more variable, as indicated by the large number of subtypes compared with BVDV-2 [27, 28]. Some BVDV-1 sequences are older than those of BVDV-2 and Hobi-like virus sequences, which enhances the precision of dating the emergence of BVDV-1 subtypes. It has also been observed that an earlier tMRCA tends to have a larger confidence interval (95% HPD) as the Bayesian method becomes less robust. This seems to be a characteristic of BVDV-1, as reported previously by Li et al. [11], who performed an analysis of the concatenated 5' UTR, N<sup>pro</sup> and E2 genes. However, even with the large confidence interval, the posterior probabilities conferred statistical confidence, which was observed in our analysis using the complete polyprotein (Fig. 1).

We used complete polyprotein sequences to date the origin of BVDV-1 genetic variants and to investigate the relationship between closely related subtypes and their geographical distribution. The ancestral BVDV-1 strain initially diverged into distinct subtypes, including the common subtypes 1a, 1b, 1c, 1d, 1e and 1i and the uncommon subtypes 1f, 1g, 1h, 1m, 1n, 1o and 1q. We inferred that subtypes 1a and 1b, in particular, were spread in cattle that accompanied human emigration during the sixteenth and seventeenth centuries. Thus, the present study may help to elucidate the origins of BVDV-1 subtypes and the molecular epidemiology and dynamics of ruminant pestivirus.

**Funding** Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and Propeq/UFRGS supported this study.

## Compliance with ethical standards

**Conflict of interest** None to declare

## References

- Yeşilbağ K, Alpay G, Becher P (2017) Variability and global distribution of subgenotypes of bovine viral diarrhoea virus. *Viruses* 9:128. <https://doi.org/10.3390/v9060128>
- Smith DB, Meyers G, Bukh J et al (2017) Proposed revision to the taxonomy of the genus Pestivirus, family Flaviviridae. *J Gen Virol* 98:2106–2112. <https://doi.org/10.1099/jgv.0.000873>
- Simmonds P, Becher P, Bukh J et al (2017) ICTV virus taxonomy profile: flaviviridae. *J Gen Virol* 98:2–3. <https://doi.org/10.1099/jgv.0.000672>
- Deng M, Ji S, Fei W et al (2015) Prevalence study and genetic typing of bovine viral diarrhoea virus (BVDV) in four bovine species in China. *PLoS ONE* 10:e0134777. <https://doi.org/10.1371/journal.pone.0121718>
- Vilcek S, Paton DJ, Durkovic B et al (2001) Bovine viral diarrhoea virus genotype 1 can be separated into at least eleven genetic groups. *Arch Virol* 146:99–115
- Weber MN, Silveira S, Machado G et al (2014) High frequency of bovine viral diarrhoea virus type 2 in Southern Brazil. *Virus Res* 191:117–124. <https://doi.org/10.1016/j.virusres.2014.07.035>
- Bachofen C, Stalder H, Braun U et al (2008) Co-existence of genetically and antigenically diverse bovine viral diarrhoea viruses in an endemic situation. *Vet Microbiol* 131:93–102. <https://doi.org/10.1016/j.vetmic.2008.02.023>
- Weber MN, Bauermann FV, Canal CW et al (2017) Temporal dynamics of 'HoBi'-like pestivirus quasispecies in persistently infected calves generated under experimental conditions. *Virus Res* 227:23–33. <https://doi.org/10.1016/j.virusres.2016.09.018>
- Ridpath JF, Bayles DO, Neill JD et al (2015) Comparison of the breadth and complexity of bovine viral diarrhoea (BVDV) populations circulating in 34 persistently infected cattle generated in one outbreak. *Virology* 485:297–304. <https://doi.org/10.1016/j.virol.2015.07.022>
- Chernick A, Godson DL, van der Meer F (2014) Metadata beyond the sequence enables the phylodynamic inference of bovine viral diarrhoea virus type 1a isolates from Western Canada. *Infect Genet Evol* 28:367–374. <https://doi.org/10.1016/j.meegid.2014.01.003>
- Liu L, Xia H, Wahlberg N et al (2009) Phylogeny, classification and evolutionary insights into pestiviruses. *Virology* 385:351–357. <https://doi.org/10.1016/j.virol.2008.12.004>
- Luzzago C, Ebranati E, Sasserà D et al (2012) Spatial and temporal reconstruction of bovine viral diarrhoea virus genotype 1 dispersion in Italy. *Infect Genet Evol* 12:324–331. <https://doi.org/10.1016/j.meegid.2011.12.007>
- Ebranati E, Lauzi S, Cerutti F et al (2018) Highlighting priority areas for bovine viral diarrhoea control in Italy: a phylogeographic approach. *Infect Genet Evol* 58:258–268. <https://doi.org/10.1016/j.meegid.2018.01.006>
- Chao L (1997) Evolution of sex and the molecular clock in RNA viruses. *Gene* 205:301–308. [https://doi.org/10.1016/S0378-1119\(97\)00405-8](https://doi.org/10.1016/S0378-1119(97)00405-8)
- Chernick A, van der Meer F (2017) Evolution of Bovine viral diarrhoea virus in Canada from 1997 to 2013. *Virology* 509:232–238. <https://doi.org/10.1016/j.virol.2017.06.024>
- Weber MN, Streck AF, Silveira S et al (2015) Homologous recombination in pestiviruses: identification of three putative novel events between different subtypes/genogroups. *Infect Genet Evol* 30:219–224. <https://doi.org/10.1016/j.meegid.2014.12.032>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Martin DP, Murrell B, Golden M et al (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:3. <https://doi.org/10.1093/ve/vev003>
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>
- Minh BQ, Nguyen MAT, Von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 30:1188–1195. <https://doi.org/10.1093/molbev/mst024>
- Rambaut A, Drummond AJ, Xie D et al (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>
- Suchard MA, Lemey P, Baele G et al (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4:16. <https://doi.org/10.1093/ve/vey016>
- Ajmone-Marsan P, Garcia JF, Lenstra JA (2010) On the origin of cattle: how aurochs became domestic and colonized the world. *Evol Anthropol* 19:148–157. <https://doi.org/10.1002/evan.20267>
- Yeşilbağ K, Förster C, Bank-Wolf B et al (2008) Genetic heterogeneity of bovine viral diarrhoea virus (BVDV) isolates from Turkey: identification of a new subgroup in BVDV-1. *Vet Microbiol* 130:258–267. <https://doi.org/10.1016/j.vetmic.2008.01.016>
- Stalder H, Bachofen C, Schweizer M et al (2018) Traces of history conserved over 600 years in the geographic distribution of genetic variants of an RNA virus: bovine viral diarrhoea virus in Switzerland. *PLoS ONE* 13:e0207604. <https://doi.org/10.1371/journal.pone.0207604>
- Aiewsakun P, Katzourakis A (2016) Time-dependent rate phenomenon in viruses. *J Virol* 90:7184–7195. <https://doi.org/10.1128/jvi.00593-16>
- Flores EF, Ridpath JF, Weiblen R et al (2002) Phylogenetic analysis of Brazilian bovine viral diarrhoea virus type 2 (BVDV-2) isolates: evidence for a subgenotype within BVDV-2. *Virus Res* 87:51–60
- Giammarioli M, Ridpath JF, Rossi E et al (2015) Genetic detection and characterization of emerging HoBi-like viruses in archival foetal bovine serum batches. *Biologicals* 43:220–224. <https://doi.org/10.1016/j.biologics.2015.05.009>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.