

## SHORT REPORTS

# “Frozen evolution” of an RNA virus suggests accidental release as a potential cause of arbovirus re-emergence

David J. Pascall<sup>1</sup>✉, Kyriaki Nomikou<sup>2,3</sup>✉, Emmanuel Bréard<sup>4</sup>, Stephan Zientara<sup>4</sup>, Ana da Silva Filipe<sup>2</sup>, Bernd Hoffmann<sup>5</sup>, Maude Jacquot<sup>6,7</sup>, Joshua B. Singer<sup>8</sup>, Kris De Clercq<sup>8</sup>, Anette Bøtner<sup>9,10</sup>, Corinne Sailleau<sup>4</sup>, Cyril Viarouge<sup>4</sup>, Carrie Batten<sup>11</sup>, Giantonella Puggioni<sup>12</sup>, Ciriaco Ligios<sup>12</sup>, Giovanni Savini<sup>13</sup>, Piet A. van Rijn<sup>14,15</sup>, Peter P. C. Mertens<sup>3,11</sup>, Roman Biek<sup>1</sup>‡\*, Massimo Palmarini<sup>12</sup>‡\*



## OPEN ACCESS

**Citation:** Pascall DJ, Nomikou K, Bréard E, Zientara S, Filipe AdS, Hoffmann B, et al. (2020) “Frozen evolution” of an RNA virus suggests accidental release as a potential cause of arbovirus re-emergence. *PLoS Biol* 18(4): e3000673. <https://doi.org/10.1371/journal.pbio.3000673>

**Academic Editor:** Andrew Fraser Read, The Pennsylvania State University, UNITED STATES

**Received:** December 9, 2019

**Accepted:** March 24, 2020

**Published:** April 28, 2020

**Copyright:** © 2020 Pascall et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are contained within the paper and Supporting Information files. Sequences of viruses used in the paper have been submitted to GenBank, and GenBank numbers are available in [S1 Table](#).

**Funding:** This work was supported by the European Union H2020 (<https://ec.europa.eu/programmes/horizon2020/en>) PALE-Blu grant (project no. 727393-2 to PPCM, SZ, BH, GS, PAVR, RB, and MP) and by the Wellcome Trust (<https://wellcome.ac.uk/>) (092806/B/10/Z; 206369/Z/17/Z)

**1** Institute of Biodiversity, Animal Health and Comparative Medicine, Boyd Orr Centre for Population and Ecosystem Health, University of Glasgow, Glasgow, United Kingdom, **2** MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom, **3** The School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington, Leicestershire, United Kingdom, **4** UMR Virologie, INRA, École Nationale Vétérinaire d'Alfort, Laboratoire de Santé Animale d'Alfort, ANSES, Université Paris-Est, Maisons-Alfort, France, **5** Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Greifswald-Insel Riems, Germany, **6** Spatial Epidemiology Lab (SpELL), University of Brussels, Brussels, Belgium, **7** INRAE-VetAgro Sup, UMR Epidemiology of Animal and Zoonotic Diseases, Saint Genès-Champagnelle, France, **8** Infectious Diseases in Animals, Exotic and Particular Diseases, Sciensano, Brussels, Belgium, **9** Section for Veterinary Clinical Microbiology, Department of Veterinary and Animal Sciences, University of Copenhagen, Copenhagen, Denmark, **10** Department of Virus and Microbiological Special Diagnostics, Statens Serum Institut, Copenhagen, Denmark, **11** The Pirbright Institute, Pirbright, Woking, Surrey, United Kingdom, **12** Istituto Zooprofilattico Sperimentale della Sardegna, Via Duca degli Abruzzi, Sassari, Italy, **13** Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise (IZSAM), Teramo, Italy, **14** Department of Virology, Wageningen Bioveterinary Research (WBVR), Lelystad, the Netherlands, **15** Department of Biochemistry, Centre for Human Metabolomics, North-West University, Potchefstroom, South Africa

✉ These authors contributed equally to this work.

‡ These authors are joint senior authors on this work.

\* [massimo.palmarini@glasgow.ac.uk](mailto:massimo.palmarini@glasgow.ac.uk) (MP); [roman.biek@glasgow.ac.uk](mailto:roman.biek@glasgow.ac.uk) (RB)

## Abstract

The mechanisms underlying virus emergence are rarely well understood, making the appearance of outbreaks largely unpredictable. Bluetongue virus serotype 8 (BTV-8), an arthropod-borne virus of ruminants, emerged in livestock in northern Europe in 2006, spreading to most European countries by 2009 and causing losses of billions of euros. Although the outbreak was successfully controlled through vaccination by early 2010, puzzlingly, a closely related BTV-8 strain re-emerged in France in 2015, triggering a second outbreak that is still ongoing. The origin of this virus and the mechanisms underlying its re-emergence are unknown. Here, we performed phylogenetic analyses of 164 whole BTV-8 genomes sampled throughout the two outbreaks. We demonstrate consistent clock-like virus evolution during both epizootics but found negligible evolutionary change between them. We estimate that the ancestor of the second outbreak dates from the height of the first outbreak in 2008. This implies that the virus had not been replicating for multiple years prior to its re-emergence in 2015. Given the absence of any known natural mechanism that could explain BTV-8 persistence over this long period without replication, we hypothesise that the second outbreak could have been initiated by accidental exposure of livestock to frozen material contaminated with virus from

to MP). 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.

**Abbreviations:** AICc, Akaike's Information Criterion corrected for small sample size; BTV-8, bluetongue virus serotype 8; FMDV, foot-and-mouth disease virus; HPD, highest posterior density; ML, maximum likelihood; OIE, World Organisation for Animal Health; UTR, untranslated region.

approximately 2008. Our work highlights new targets for pathogen surveillance programmes in livestock and illustrates the power of genomic epidemiology to identify pathways of infectious disease emergence.

## Introduction

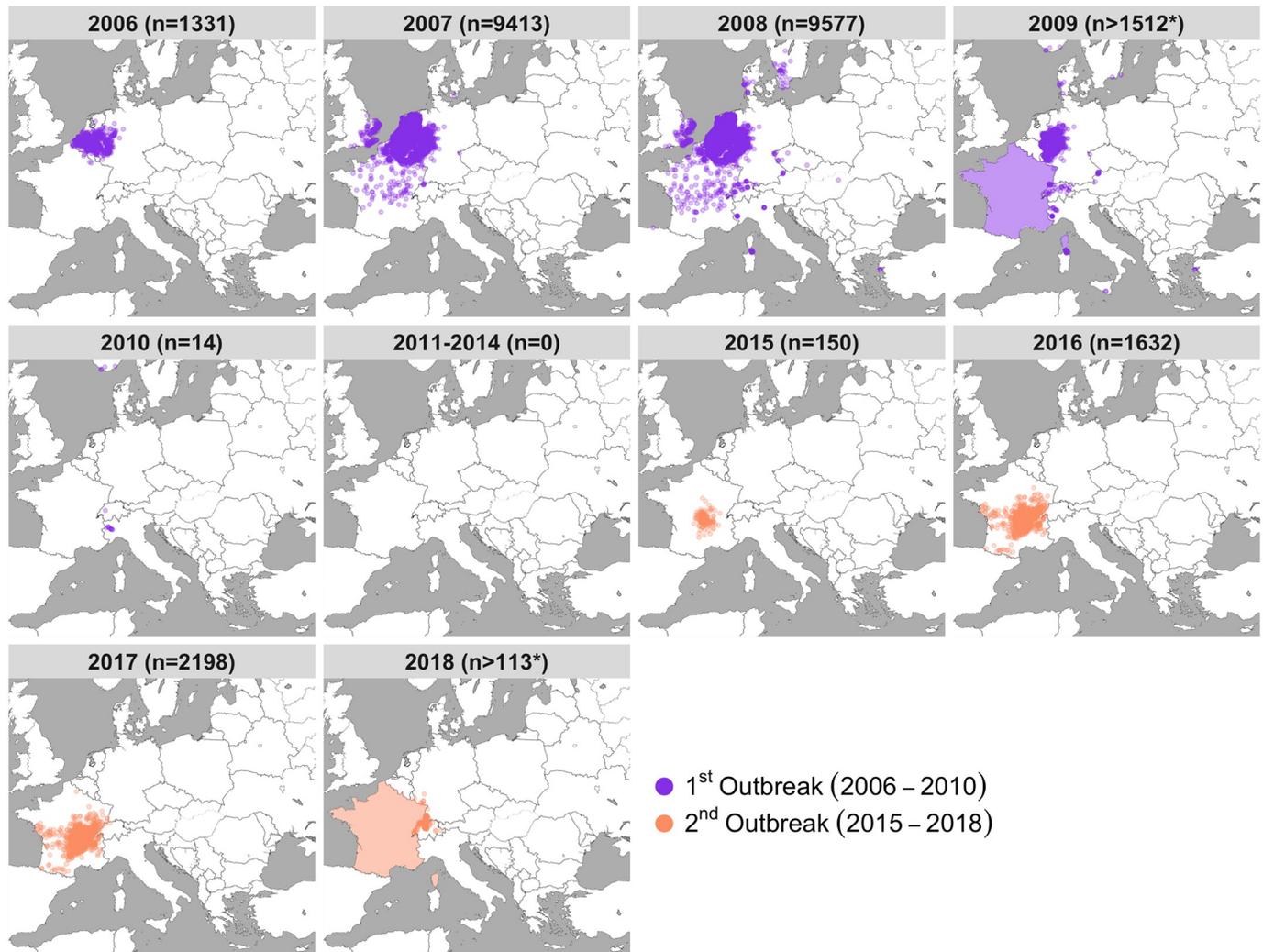
Infectious disease outbreaks are a major burden on human and animal health. They can dramatically reduce the productivity of entire countries due to direct losses, control measures, trade bans, or public fear [1]. Diseases caused by arthropod-borne viruses (arboviruses) in particular have increased substantially in recent decades [2–4], and there is an urgent need to better understand the causes of their emergence in order to devise better control and prevention strategies. The factors leading to disease emergence are often unclear, and case studies of intensely studied outbreaks can therefore provide important wider lessons.

Bluetongue is a major disease of domestic ruminants caused by the bluetongue virus (BTV); an arbovirus transmitted by *Culicoides* midges. BTV is the type species of the genus *Orbivirus*, within the family Reoviridae, and possesses 10 double-stranded RNA genome segments encoding for 7 structural and 4 or 5 nonstructural proteins [5–8]. BTV infection in sheep can induce a variety of clinical outcomes, which in the most extreme cases include a lethal haemorrhagic fever [9–11]. Infection in cows and goats results instead in milder and often subclinical signs [9,10]. BTV can also infect wild ruminants and, more rarely, other mammal species [12–16].

Like many other arboviruses, the geographical spread of BTV has increased significantly in the last 20 years [17–19]. In August 2006, BTV serotype 8 (BTV-8) emerged for the first time in the Netherlands [20–25], leading to dramatic losses of sheep and causing extensive economic damage to farming communities, costing on the order of billions of euros [26–29]. The virus quickly spread across the continent, with confirmed infections in 16 countries by 2008 (Fig 1). The outbreak was ultimately controlled through a pan-European vaccination campaign, using inactivated vaccines, with a few last cases detected in Europe in 2010 [30]. However, after a five-year period with no BTV-8 cases recorded throughout Europe, the virus re-emerged in France [31] and has since continued to spread. France was declared enzootic in 2018 and recent cases reported in adjacent countries, including Germany, Switzerland, and Belgium [32].

The source and mechanism of BTV-8 re-emergence in France remains obscure. Initial genetic data from one isolate suggested the re-emerging virus in France to be a close relative of the lineage causing the 2006–2010 outbreak [31,33,34]. The prevailing theory was that the virus had continued to be transmitted subclinically but remained unrecorded in livestock or wild ruminants after it had been declared absent from Europe in 2011 [31]. However, there is currently little evidence to support this hypothesis. Based on serological evidence, wild ungulates do not appear to have sustained transmission [35,36]. Similarly, serological testing of cattle sampled in 2014 indicated a rapid decline of seropositivity after vaccination ceased in France in 2010, consistent with a new (re-)introduction of the virus in, or just before, 2015 [35,37,38].

We describe the use of phylogenetic and evolutionary analyses of BTV-8 virus samples, collected during the first and second European outbreaks, to gain insights into the mechanisms that allowed BTV-8 to re-emerge in France in 2015. For this, we generated a novel dataset of full genome sequences for more than 150 viruses sampled throughout both outbreaks.



**Fig 1. Emergence and re-emergence of BTV-8 in Europe.** Location and number of premises housing livestock infected with BTV-8 collected from “immediate reports” to the OIE (World Organisation for Animal Health). Data were accessed from the WAHIS database on 12 July 2019. Immediate report data were provided by the OIE. Each point corresponds to one infected premise, with total counts for each year ( $n$ ), with the first outbreak shown in purple and the second outbreak shown in orange. Please note that points on the maps reflect only immediate notification to the OIE, resulting in a smaller number than the totality of affected premises during the outbreaks. In addition, France is shaded in 2009 and 2018 as France stopped providing immediate reporting, requiring location data, to the OIE during those years. As such, the counts of infected premises from 2009 and 2018 are not comparable to previous years as they exclude France, where the virus was widespread. Data are available from the WAHIS database at [https://www.oie.int/wahis\\_2/public/wahid.php/Wahidhome/Home](https://www.oie.int/wahis_2/public/wahid.php/Wahidhome/Home). BTV-8, bluetongue virus serotype 8. Map adapted from tiles by Stamen Design, under Creative Commons (CC BY 3.0) using data by OpenStreetMap, under the Open Database Licence.

<https://doi.org/10.1371/journal.pbio.3000673.g001>

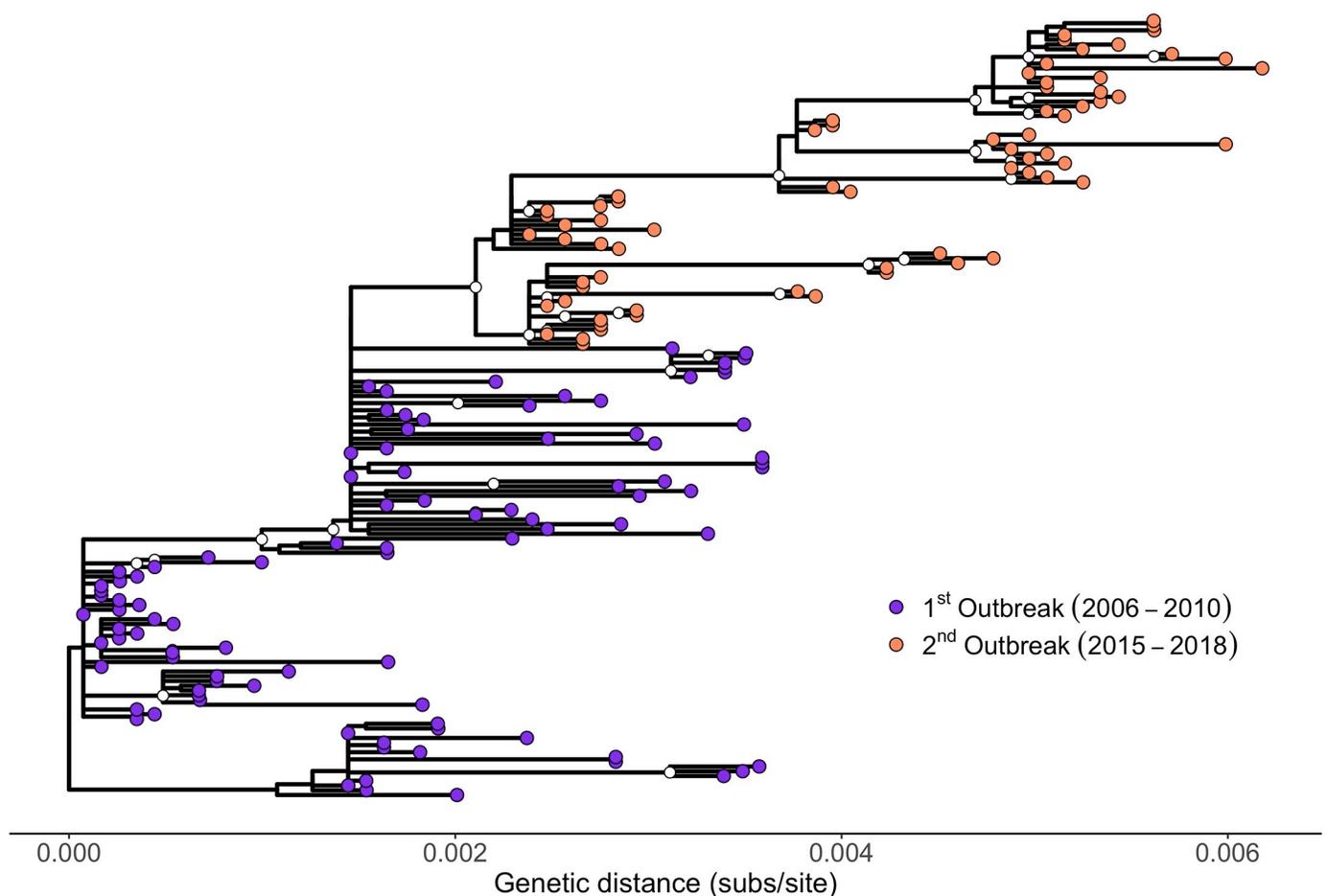
## Results

### Viruses from the first and second European BTV-8 outbreaks form a single monophyletic clade

We analysed newly sequenced full genomes of 153 BTV-8 samples collected from infected sheep and cattle throughout the BTV-8 outbreaks in Europe along with 11 BTV-8 isolates previously published. Samples from the first outbreak were collected from infected animals in 10 different countries between 2006 and 2009, while samples from the second outbreak were collected from France between 2015 and 2018 (S1 Table). To minimise or exclude the possibility

of including genome mutations acquired during extensive passage of the virus in culture, we sequenced the great majority of samples directly from clinical samples (blood) of infected animals or from isolates kept in culture for a minimum number of passages (S1 Table).

A maximum likelihood (ML) tree revealed considerable genetic diversity both within the first (2006–2010) and the second outbreak (2015–2018) (Fig 2 and S1 Fig). The tree showed that all sequences from the second European outbreak form a well-supported monophyletic clade that is nested within the virus lineages circulating during the first outbreak in 2006–2010 (Fig 2 and S1 Fig). Specifically, the clade of the second outbreak derives from a clade from the first outbreak including predominantly viruses from France and Germany collected in 2007 and 2008. The viruses from the second French outbreak can be distinguished into two further clades, one including viruses from 2015 and 2016, the other including samples spanning the entire outbreak (2015–2018). Interestingly, both clades were already present among the eight samples from the farm in France (in Auvergne-Allier) from which the first diagnosis of re-emerged BTV-8 was made in August/September of 2015. Surprisingly, the branch leading to



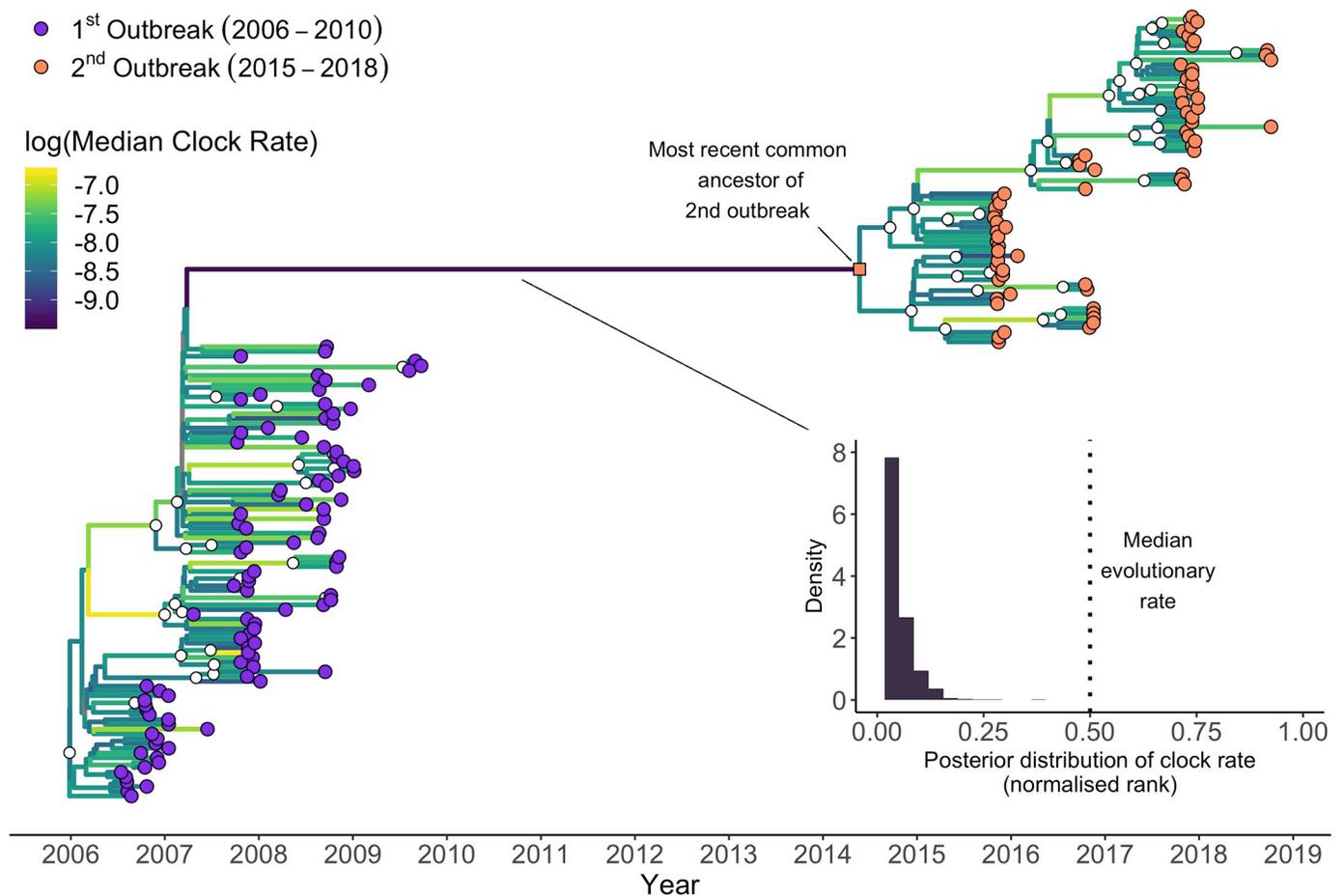
**Fig 2. Phylogenetic tree of 164 BTV-8 samples collected during the European outbreaks between 2006 and 2018.** ML tree estimated in PhyML. The scale bar shows substitutions per site. Internal nodes represented 700 or more times within 1,000 bootstraps are indicated by a white circle. Samples from the first outbreak are shown as purple circles, while samples from the second outbreak are shown as orange circles. Note that an identical tree with labels corresponding to the individual samples is shown as S1 Fig. The tree is available in S3 Data as PhyMLTreeFinal.tree. BTV-8, bluetongue virus serotype 8; ML, maximum likelihood.

<https://doi.org/10.1371/journal.pbio.3000673.g002>

the re-emerging virus appeared short, given the five-year period between the outbreaks, implying a low rate of evolution during this period (Fig 2 and S1 Fig).

### BTV-8 re-emergence associated with exceptionally slow evolution

To test if the lower amount of divergence along the branch separating the two outbreaks was unusual, we estimated the evolutionary rate of BTV-8 from the set of 164 genomes. For this, we applied a lognormal relaxed clock model that allows for branch-specific heterogeneity in clock rates, implemented in the Bayesian phylogenetic software BEAST (Fig 3 and S2 Fig). The mean evolutionary rate estimate was  $4.04 \times 10^{-4}$  substitutions per site per year (95% highest posterior density [HPD]:  $3.37 \times 10^{-4}$ ,  $4.72 \times 10^{-4}$ ), corresponding to an expected 7.76 substitutions per year (95% HPD: 6.47, 9.06) across the entire BTV-8 genome. In contrast, the



**Fig 3. Timescaled phylogenetic tree of BTV-8 samples collected during the European outbreaks between 2006 and 2018.** Maximum clade credibility time-calibrated phylogenetic tree generated in BEAST. The tree is scaled in years, with the final sampling date being October 2018. Clades with posterior support of 0.9 or higher are indicated by a white circle. Samples from the first and second outbreaks are shown as purple and orange circles, respectively. The branches are coloured accordingly to their median evolutionary rate across the posterior (see heatmap within the figure). The long branch leading from the first outbreak to the second (in dark purple) shows the slowest evolutionary rate on the maximum clade credibility tree. The inset shows the posterior distribution of the normalised rank of the evolutionary rate of the long branch relative to the other branches in the tree, as estimated from the lognormal relaxed clock. Relative to the rest of the tree, values close to 0 represent slow evolution, values close to 1 represent fast evolution, and 0.5 represents the median branch-specific evolutionary rate on the tree. This analysis reveals an unusually slow evolution of BTV-8 between outbreaks. See the Methods section for the specifics of this calculation. Note that an identical tree with labels corresponding to the individual samples is shown as S2 Fig. The tree is available in S3 Data as MCC.tree; other presented data can be extracted using code in S1 Data from the trees in S3 Data as AllGMRF\_reduced.trees. BTV-8, bluetongue virus serotype 8.

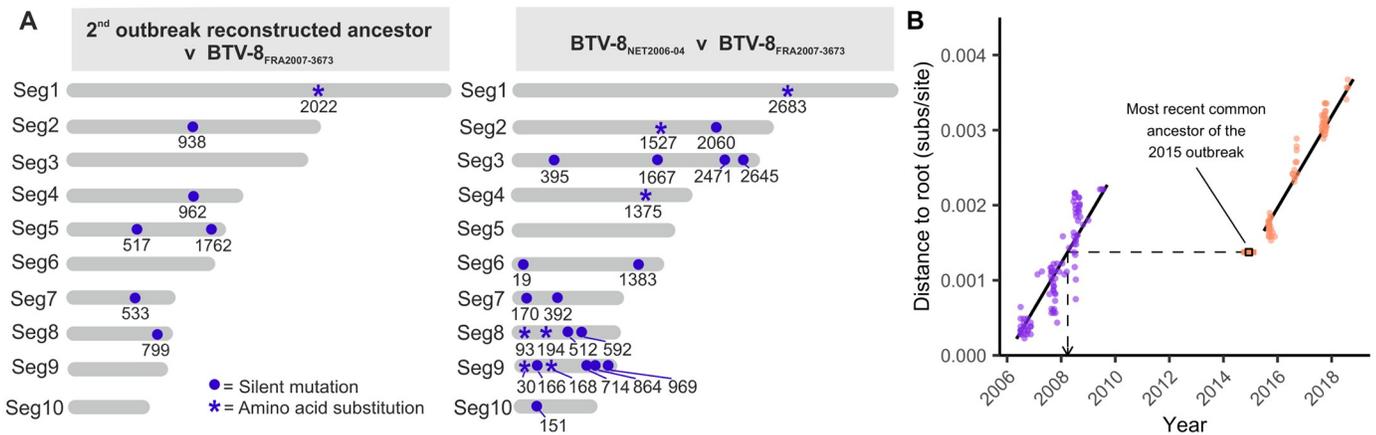
<https://doi.org/10.1371/journal.pbio.3000673.g003>

emerging branch had an estimated mean evolutionary rate that was nearly an order of magnitude slower, at  $8.24 \times 10^{-5}$  substitutions per site per year (95% HPD:  $3.93 \times 10^{-5}$ ,  $1.32 \times 10^{-4}$ ). Indeed, of the branches within the presented maximum clade credibility tree, this branch had the lowest median rate across the posterior distribution of trees.

Using BEAST, we reconstructed the sequence of the most recent common ancestor of the viruses sequenced from the second outbreak. This ancestral sequence displayed only 7 nucleotide substitutions (of which 6 were synonymous or in the untranslated regions [UTRs] of the viral genome) compared to BTV-8<sub>FRA2007-3673</sub>, the genetically closest virus within the dataset, which was collected in France in August 2007 (Fig 4A). In comparison, BTV-8<sub>FRA2007-3673</sub> displayed 23 nucleotide substitutions (of which 16 were synonymous or in the UTR) compared to the first sequence available from the first outbreak and collected in August 2006 in the Netherlands (BTV-8<sub>NET2006-04</sub>) (Fig 4A). The number of mutations of BTV-8<sub>FRA2007-3673</sub> compared to the BTV-8 sequences in the dataset in 2006 ( $n = 23$ ) ranges between 15 and 23, while those compared to the virus sequences collected in 2008 ( $n = 37$ ) varies between 2 and 56. Hence, sequence variation between BTV-8 samples collected only a year apart during the first outbreak is, in general, far higher than that between the ancestor of the re-emerged BTV-8 strain and its closest relative in the first outbreak.

### "Frozen evolution": Clock-like evolution of BTV-8 during, but not in between, outbreaks

Next, we included the reconstructed ancestral sequence in our dataset and re-estimated the ML tree to get measures of genetic distance from the root of the tree. Consistent with clock-like evolution, the genetic distance between virus sample and the root increased linearly with time during both the first (TempEst: slope =  $7.2708 \times 10^{-4}$  subs/site/year,  $r^2 = 0.8113$ ) and the



**Fig 4. Lack of evolution of BTV-8 between the two European outbreaks.** (A) Graphic representation of nucleotide substitutions between the genomes of the earliest BTV-8 collected from the first European outbreak (BTV-8<sub>NET2006-04</sub>), the reconstructed ancestor of the second BTV-8 outbreak (BTV-8<sub>FRA2015</sub>), and the most similar virus to the latter sequence present in our dataset (BTV-8<sub>FRA2007-3673</sub>). Substitutions are shown as a blue circle, with numbers indicating the genomic position for each of the 10 genomic segments. Asterisks indicate those mutations inducing also an amino acid substitution. (B) Genetic divergence of 164 BTV-8 samples collected from the two European outbreaks against their sampling date (circles). The regression lines corresponding to the posterior mean estimate from the best fitting linear model for each outbreak are shown in black. Credible intervals are omitted because nonindependence of points would have made conventional estimate of standard errors invalid. When the day was unknown, the date was fixed to the 16th of the month, and when the month was unknown, the date was fixed to the midpoint of the year. The inferred age of the ancestor of the second outbreak is shown as a square, with a 95% HPD error. The dashed line indicates that the inferred ancestor of the second outbreak has a degree of divergence that is equivalent to a virus from the first outbreak circulating in 2008. Mutations can be derived from the raw sequence data deposited on GenBank under the IDs listed in S1 Table. Distances to the root can be extracted from the tree in S3 Data as MlincludingAncestor.tree, and dates are available in dateswithuncertainty.txt in S3 Data. BTV-8, bluetongue virus serotype 8; HPD, highest posterior density; Seg, virus genome segment.

<https://doi.org/10.1371/journal.pbio.3000673.g004>

second outbreak (TempEst: slope =  $6.9291 \times 10^{-4}$ ,  $r^2 = 0.9605$ ). There was no evidence that the evolutionary rate of the virus differed between the two outbreaks ( $p$ -value for Date:Outbreak interaction = 0.194). However, there is a clear discontinuity in the accumulation of mutations between the two outbreaks, consistent with a period during which clock-like evolution had essentially ceased. Consequently, the reconstructed sequence of the ancestor of the second outbreak, when included in the ML tree, has an inferred distance from the tree root that is consistent with a virus from late March 2008, according to the root-to-tip regression (Fig 4B).

## Discussion

Diseases of livestock can be exceedingly interesting models to study virus emergence, given that harmonised international surveillance systems and regulatory frameworks provide opportunities to access field samples with associated metadata across national borders. Here, the BTV-8 European outbreaks provided us with the opportunity to investigate the mechanisms surrounding arbovirus emergence based on a uniquely rich dataset. Our results indicate that the re-emergence of BTV-8 in France in 2015 was caused by a virus that exhibits a lack of evolutionary changes since the first outbreak. This is inconsistent with the prevalent view of undetected low-level circulation of the virus in wild or domestic ruminants between 2010 and 2015, and instead points to another mechanism of emergence.

We showed a large discontinuity in the number of mutations accumulated by BTV-8 between 2010 and 2015, even though the evolutionary rates of the virus during the first and second outbreak were indistinguishable and of the same order as rates reported in previous BTV studies [39,40]. If the virus had been replicating consistently in an undetected population from 2010 to 2015, we would expect the genetic distance of the isolates from the second European outbreak to continue the trend of increased divergence after the first outbreak. However, the sequences from the second outbreak exhibit genetic divergences that fall considerably below what would be expected if the trend line from the first outbreak was extended, illustrating a paucity of mutations (Fig 4B). Indeed, the divergence of the reconstructed ancestor of the second bluetongue outbreak is consistent with the virus stopping replication in March 2008. The lack of divergence is also illustrated by the fact that the reconstructed ancestor of the BTV-8 outbreak has only 7 mutations separating it from its closest relative in the analysed dataset, a French sample collected in August 2007 (BTV-8<sub>FRA2007-3673</sub>), despite putatively having been replicating for at least half a decade after that sample's collection. In comparison, BTV-8<sub>FRA2007-3673</sub> showed 23 mutations compared to the genome of the earliest BTV-8 sample obtained from the Netherlands in August 2006, only a year earlier. The corresponding rate of evolution estimated for the emerging branch was almost an order of magnitude slower than the mean clock rate, highlighting it as exceptionally low (Fig 3). Moreover, we hypothesise that some or all of the estimated seven mutations on this branch might have accumulated during the first outbreak, given that the emerging branch connects to an internal node in the time-scaled phylogeny with a date of early 2007, at the height of the first outbreak. The subsequent accumulation of seven mutations is consistent with the idea that this virus continued to circulate until early 2008 (the inferred date from the root-to-tip regression) and then ceased to change altogether until its re-emergence in 2015. While previous studies have found apparent evolutionary stasis to be the result of mislabelling [41], this can be ruled out in our case due to the discontinuity applying to all samples from the second outbreak, not just a single isolate. Another hypothetical scenario could be envisaged if BTV-8 was to remain "latent" in midges' eggs for a number of years. However, there is no evidence of vertical transmission in BTV-

infected *Culicoides* [42–45]. This, in conjunction with the need for infected midge eggs to survive for years, rather than a single overwintering season, makes this scenario highly unlikely.

Given the unexpectedly low number of mutations observed between the two outbreaks, our data indicate that the common ancestor of the second European outbreak either ceased or dramatically slowed its replication in early 2008. This is inconsistent with current knowledge and paradigms of the biology of BTV and RNA viruses in general. For example, a potential explanation could be that BTV persistently infected a host for several (5 to 8) years, with little or no replication, before being reactivated and starting the second outbreak. While this may be possible with DNA viruses, or RNA viruses with a DNA intermediate [46–51], a mechanism for this has never been described before for reoviruses such as BTV and in general for other RNA viruses.

Our findings have interesting parallels to puzzling examples from other RNA viruses, such as two Ebola virus outbreaks in the Democratic Republic of Congo in 2014 and 2018 [52,53]. Isolates from both outbreaks were minimally divergent from isolates collected about a decade earlier, resulting in a far lower evolutionary rate than other known lineages. It has been suggested that such slow evolution may be caused by the maintenance of the virus in an animal reservoir, where infection might be associated with lower replication rates compared to human hosts [52,53]. Rabies virus may provide an additional peculiar example based on a handful of reports in human patients of virus reactivation after latency of several years [54], but it has not been documented whether these cases involved a lack of evolutionary changes. In other cases, such as foot-and-mouth disease, viral RNA and infectious virus have been shown to persist in reservoir hosts for multiple years. However, re-isolation of virus (as opposed to detection of viral RNA) indicates that the virus replicates during persistent infection and accumulates nucleotide substitutions at a rate comparable to actively replicating viruses [55,56]. Hepatitis C virus, for example, is also known to persist in a number of patients for a number of years but, again, with continuing viraemia and thus virus replication [57].

Overall, we judge the possibility of persistence of BTV-8 in a mammalian or invertebrate host for longer than five years, in the absence of viral replication, followed by viral reactivation and subsequent onwards spread, to be unlikely given the current understanding of RNA virus biology. We hypothesise that accidental release of frozen material contaminated with BTV-8 could be the cause of the virus re-emergence in France in 2015. Anthropogenic causes of virus outbreaks have been described before. Accidental virus release is thought to have been responsible for the 1977 influenza A H1N1 outbreak, caused by a virus that closely matched a variant circulating in the 1950s [58,59]; likewise, the 1995 Venezuelan equine encephalitis subtype IC epidemic was caused by a virus closely related to a strain circulating in 1962–1964 [60]. For livestock pathogens, a localised outbreak of foot-and-mouth disease virus (FMDV) in the United Kingdom in 2007 was linked to virus escaped from research facilities [61].

Our data cannot reveal the actual source from which BTV-8 was re-introduced in France in 2015. We speculate that laboratory escape of virus preparations, such as the case of FMDV in the UK in 2007, is unlikely, as BTV needs an insect vector for efficient transmission and we are unaware of any *in vivo* insect experiments in France with BTV during that period. However, due to specific animal husbandry procedures, there are important potential sources of frozen virus that apply to viruses of livestock and not viruses of most other animals, specifically the widespread use of bull semen for artificial insemination and embryo transfer in cows [62,63]. BTV has been detected in the semen of viraemic bulls and rams, can initiate infection in the mother, and can be transmitted vertically to the embryo [64,65]. Additionally, contaminated embryos can cause transmission on implantation [66]. As such, both semen and embryos may represent potential sources of BTV infection. Contaminated frozen colostrum may also be a potential source, considering that oral transmission has been shown to be possible with BTV-8

[67]. However, it is not normal practice to keep colostrum frozen for a number of years. Interestingly, while international regulations specify that bull donors and semen that are exported internationally must be screened for various pathogens including BTV [68], this does not apply to premises trading only locally and carrying out private insemination procedures [69]. Thus, semen from a BTV-8–infected bull could have been collected or an embryo generated from an infected but asymptomatic animal and used years later without detection.

We stress that the link between bull semen trade and embryo implantation in France and the BTV-8 re-emergence in 2015 is only speculative. However, we have shown that the re-emergence of BTV-8 in France in 2015 is unlikely to be due to cryptic continuing transmission, and we can exclude a reintroduction from another endemic country. Thus, our data are incompatible with the two current dominant theories for explaining the 2015 outbreak [31]. The lack of accumulated mutations in the virus implies that there was either an ongoing persistent infection in the absence of viral replication for several years, or the virus originated from material that had been frozen during the first outbreak. We argue the second of these explanations to be more likely. Our findings highlight new areas requiring thorough surveillance programmes for the control of infectious disease of livestock. In addition, our approach illustrates how unrecognised pathways of disease emergence can be revealed using pathogen genomic epidemiology.

## Methods

### Samples

Blood samples from animals infected with BTV-8 were received from ten European countries during the bluetongue outbreaks from 2006 to 2018. In some instances, samples analysed were viruses isolated in tissue culture from blood of infected animals. [S1 Table](#) provides the metadata related to the dataset used in this study. These include virus strain names, animal species of origin, geographical location, and date of sampling. In addition, metadata include whether the viral genome sequence was obtained directly from clinical material (blood) or from an isolate in tissue culture, sequencing methods, and GenBank accession numbers.

### RNA extraction and Illumina library preparation

Total RNA was extracted from infected blood samples, and virus isolates using Trizol LS (Invitrogen, Carlsbad, CA) and purified using Direct-zol RNA MiniPrep (Zymo Research, Irvine, CA) as per the manufacturer's protocol. RNA samples were treated with DNase I (Ambion, Austin, TX) and purified with 3× Agencourt RNAClean XP beads (Beckman Coulter, Brea, CA). Total RNA concentration was quantified using the Qubit Fluorometer (Life Technologies, Carlsbad, CA) and Qubit RNA HS Assay (Life Technologies, Carlsbad, CA), while RNA integrity was assessed using Agilent 4200 TapeStation System (Agilent, Santa Clara, CA). In order to avoid cross-contamination, RNA extractions from virus isolates were performed separately from those of infected blood samples. Similarly, RNA extractions were carried out separately on the basis of geographical origin and year of collection of the samples. In addition, library preparations, target enrichment, and sequencing runs (see below) were carried out also on separate days, following the same criteria as above. Libraries from low and without measurable RNA (low input) were also prepared separately from those with measurable RNA (high input). Libraries were prepared for Illumina sequencing using the Illumina TruSeq Stranded mRNA HT kit (Illumina, San Diego, CA) using 5 µL of sample RNA (up to 250 ng of total RNA) according to the manufacturer's instructions. Briefly, after RNA was fragmented, it was reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA) and random hexamers. Single-stranded cDNA was immediately converted to double-stranded

cDNA, cleaned up with Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA), quantified using Qubit Fluorometer and Qubit dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA), and the size distribution was assessed using a 4200 TapeStation System with High Sensitivity D1000 Screen Tape assay (Agilent, Santa Clara, CA). A-tailing was performed, followed by indexed adapter ligation. After a purification step, dual indexed libraries were PCR amplified, and the purified PCR products were pooled in equimolar concentrations and sequenced using 150 paired-end sequencing on MiSeq or NextSeq500/550 sequencers (Illumina, San Diego, CA).

### Targeted enrichment sequencing

We carried out multiplexed viral targeted enrichment followed by Illumina sequencing using the NimbleGen SeqCap EZ system (Roche, Pleasanton, CA) for improved viral detection from clinical material. This approach was followed in order to increase the number of BTV-8 samples from which we could obtain a complete viral genome sequence directly from clinical material, including those with very low amounts of viral RNA. Libraries were prepared following the standard Illumina TruSeq Stranded mRNA protocol, described above. They were quantified using Qubit Fluorometer and Qubit dsDNA high sensitivity (HS) Assay Kit (Life Technologies, Carlsbad, CA). Quality and size distribution were validated using the High Sensitivity D1000 Screen Tape assay (Agilent, Santa Clara, CA) in a 4200 TapeStation System (Agilent, USA) and were normalised according to BTV viral load and mass. A 1- $\mu$ g aliquot of the pooled library was enriched using a SeqCap EZ Developer Probe (Roche/NimbleGen) (see below), according to the manufacturer's protocol. After a 14-cycle post-enrichment PCR amplification, the cleaned PCR products were pooled and were sequenced with a 151-base paired-end reads on a NexSeq500/550 cartridge (Illumina, San Diego, CA). Probes were designed using all BTV sequences available on NCBI Genbank, RefSeq, DNA Data Bank of Japan (DDBJ), and EMBL EBI databases (as accessed by October 2016). The resulting NimbleGen biotinylated soluble capture probe library ("BTV-Cap") contains a probe set of more than 500,000 probes, designed to minimise capture of *Culicoides sonorensis*, *Bos taurus*, *Ovis aries*, *Capra hircus*, and *Mesocricetus auratus* genomes.

### Consensus calling

All consensus calling was performed on a cluster running Ubuntu v. 14.04.5 LTS. In all cases, BAM and SAM files were handled using samtools v. 1.3 [70]. The R packages ggplot2 v. 3.2.1 [71], seqinr v. 3.6-1 [72], stringr v.1.4.0 [73], and vcfR v. 1.8.0 [74] were all used in scripts at various points in the following section. Paired-end raw reads were trimmed with Trim Galore! v. 0.4.0 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)) with a quality cut-off of 30. Any reads below 50 bp were discarded. Following this, any reads that were unpaired were discarded if they were under 100 bp. Overlapping reads were combined using FLASH v. 1.2.11 [75]. Reads were mapped using bowtie2 v. 2.3.4.2 [76] to a reference database containing all the segments of all the described strains of BTV in order to manually check for mixed infections. Reads were allowed to have as many valid maps as could be found. Mapping statistics were generated with weeSAM v. 1.5 (<https://github.com/centre-for-virus-research/weeSAM>) and transcripts per million for each target were then generated using eXpress v. 1.5.1 [77]. Separately, for the consensus generation, reads were mapped using Tanoti v. 9 July 2018 (<https://github.com/vbsreenu/Tanoti/tree/master/src>) against a reference BTV-8 genome from the European BTV-8 outbreak (GenBank accession numbers: JX680447-JX680456). Different software was used for the quality control and consensus building steps, as bowtie2 generates meta-data required for the downstream quality control steps that Tanoti does not.

Variants from the BTV-8 reference were then called from the Tanoti alignment with lofreq\* v. 2.1.2 [78], with the minimum coverage of the filtering step set to 5 and all other parameters at their default values. Any variants from the reference with an allele frequency of greater than 0.5 were replaced into their positions in the reference to build a new reference sequence. Reads were then remapped to this new reference. This process was then repeated either 5 times or until the reference generated after the process was identical to the reference at the start of the last round of mapping. Reads were then mapped again against this new reference, and ambiguities were called. A base was called unambiguously if the allele frequency of the dominant allele was greater than 0.75; otherwise, the base was called ambiguously over all alleles with a frequency of greater than 0.05 using a bespoke script in R (S1 Data). For both of these consensus sequences, positions were masked with "N"s if the coverage at the site was less than 5 separate paired-end reads.

Final sequences were processed and annotated for submission to GenBank using an extension to the BTV-GLUE resource (<http://btv-glue.cvr.gla.ac.uk>). GenBank accession numbers for each sample are in S1 Table.

### Quality control

A sequence showing evidence of mixed infection or contamination was discarded. Potential contamination and/or mixed infection were detected by finding sequences that met the following two criteria: (i) visible numbers of reads mapping to serotypes other than BTV-8 or the closely related BTV-18 in segments 2 and 6; and (ii) the presence of regions that, when aligned, showed large numbers of unique SNPs and ambiguous nucleotides. In total, 8 samples were discarded due to mixed infection with different BTV strains and/or contamination (7 samples from the first outbreak and 1 from the second). During quality control, segment 7 from the sample FRA2008-28 was also removed, as it represented an obvious reassortment from a distinct BTV strain, but the rest of the sample was preserved. We used GiRaF v. 1.02 [79] and MrBayes v 3.2.7a [80] on all the samples to test for the presence of less obvious reassortments between serotypes. Within the GiRaF algorithm, per-segment trees were run for 1,000,000 iterations, with 500,000 iterations discarded as burn-in. All other parameters were left at their default values. No reassortment was detected, so we opted to use all segments in a single concatenated phylogenetic tree. However, it should be noted that, as there is little variation in many segments, our ability to detect reassortment between two distinct but phylogenetically related strains is correspondingly low.

### Phylogenetics

Two separate phylogenetic analyses were performed, a ML analysis performed in PhyML v. 20120412 [81] and a Bayesian analysis performed in BEAST v. 1.10 [82].

### ML analysis

To explore the diversity of the outbreaks, we generated a ML tree. All segments were concatenated into a single sequence, and a phylogeny using the GTR+G+I nucleotide model was run in PhyML v. 20120412 [83]. All parameters were optimised by ML. The algorithm was the best of NNI and SPR moves with 10 random starts, with 1,000 bootstraps being performed on the best tree found. A ML tree containing all sequences was then run using the same settings as the first, and 1,000 bootstraps were performed on the best tree found in those 12 starts. Given the observed short branch between the first and second outbreaks, this tree was then rooted at the optimal root found from the tree containing only the sequences from the first

outbreak, generated under the same settings, as calculated by TempEst v. 1.5.1 [84]. The TempEst rooting procedure also confirmed clock-like evolution for this dataset.

### BEAST analysis

Using known break points, the sequence for each segment was split into the UTR and the first, second, and third codon positions of the coding sequence. In the ninth and tenth segments, there are regions with overlapping open reading frames; these were also placed together in their own partition. Separate evolutionary models, linked across segments, were applied to each of these partitions. The segments shared a lognormal relaxed molecular clock [85]. Given the difficulty, caused by combinatorial explosion, of model selection when there are multiple partitions, we performed a pre-analysis model selection protocol. Each segment was concatenated and a model was chosen for each partition using jModelTest v 2.1.10 [86]. The best model that was implemented in BEAST v. 1.10 was selected based on Akaike's Information Criterion corrected for small sample size (AICc). This was a GTR model for the first codon position, an HKY for the second, GTR+G with 4 gamma categories for the third, K80 for the UTR, and JC for the regions with overlapping ORFs. We used a GRMF skyride model for the tree prior [87]. When the sampling date was not exactly known, the age of the tip was estimated in the MCMC with a uniform prior over the period of uncertainty. This meant that the sampling date was correctly controlled for despite the observed discontinuity [88]. All priors were left at their default values except for the mean of the lognormal distribution for the relaxed molecular clock, which was given a lognormal (-7.6, 3) prior. In all cases, ambiguous nucleotides were used in the tree likelihood. Two trees were run, one containing only the sequences from the 2015 outbreak and one containing all sequences. The tree containing all sequences was used to reconstruct the sequence of the ancestor of all the viruses in the second outbreak. BEAST will reconstruct the sequence even in locations where the majority of sequences show gaps in the alignment. As such, there were three nucleotides that we removed from the final reconstructed sequence, corresponding to locations in the original multiple sequence alignment where all sequences but one had gaps. The BEAST XMLs for the two analyses described above are available as [S2 Data](#).

### Downstream statistical analysis and figure generation

Observed genetic distances from the full ML tree and sampling date were combined in R. When the exact sampling date was unknown, if the day within the sampling month was unknown, the date was fixed to the 16th of the month, and when the month was unknown, the date was fixed to the midpoint of the year. General linear models in base R (glm function) were fitted to test if the evolutionary rate of the virus was the same between two outbreaks. The models used a gamma distribution with an identity link. The regression equation for the first model was as follows: Genetic distance from root ~ Date + Outbreak. The regression equation for the second model was as follows: Genetic distance from root ~ Date + Outbreak + Date:Outbreak. After no evidence was found of differential rates between the two outbreaks, the general linear model without the interaction was run in brms v. 2.10.0 [89]. A normal (0, 10) prior was placed over the intercept, standard normal priors were placed over all regression coefficients, and a gamma (0.01, 0.01) prior was placed over the shape parameter. The normalised rank of the evolutionary rate of the long branch was calculated as follows: for each tree in the posterior, ranking the estimated evolutionary rate of each branch from slowest to fastest, extracting the rank for the long branch, subtracting 1 so that the minimum was 0, then dividing by the number of branches minus 1, so that a number between 0 and 1 was generated. Figures were generated using the following R packages: ggplot2 v. 3.2.1 [71], ggtree v. 1.16.6 [90], ggthemes v. 4.2.0 [91], cowplot v. 1.0.0 [92], ggmap v. 3.0.0 [93], viridis v. 0.5.1 [94], tidybayes v. 1.1.0 [95],

lubridate v. 1.7.4 [96], sp v. 1.3–1 [97], raster v. 3.0–7 [98], maptools v. 0.9–8 [99], rgeos v. 0.5–2 [100], rgdal v. 1.4–6 [101], sf v. 0.8–0 [102], and PBSmapping v. 2.72.1 [103].

## Supporting information

**S1 Data. Scripts for the analyses performed in the paper in a zipped folder.** BTV8analysis\_gz.sh, BTV8analysis.sh, consensusfinal.R, consensusinitial.R, figure.R, finaldepthcorrection.R, namecorrection.R, and sequencecomparison.R are all used in the consensus sequence generation. Splits.R breaks fasta files into separate fasta files for codons 1, 2, and 3, the UTR, and regions of overlapping ORFs in segments 9 and 10. Reconstruction.R puts the reconstructed ancestor back together again. HockeyStickFinal.R, Map.R, MLTree.R, and BEAST-Tree.R all generate figures. ORF, open reading frame; UTR, untranslated region. (CPGZ)

**S2 Data. BEAST 1.10 XMLs for the two Bayesian analyses in a zipped folder.** (CPGZ)

**S3 Data. Data files used in the analyses in this paper in a zipped folder.** AllGMRF\_reduced.trees is the tree posterior for all 164 samples downsampled to 500 trees. MCC.tree is the maximum clade credibility tree for all 164 samples. PhyMLTreeFinal.tree is the ML tree of all 164 samples. dateswithuncertainty.txt contains the sampling date for each sample. ML, maximum likelihood. (CPGZ)

**S1 Fig. Phylogenetic tree of 164 BTV-8 samples collected during the European outbreak between 2006 and 2018 with tip labels.** ML tree estimated in PhyML. The scale shows substitutions per site. Clades represented 700 or more times within 1,000 bootstraps are indicated by a white circle. Samples from the first outbreak are shown with purple circles, while samples from the second outbreak are shown with an orange circle. BTV-8, bluetongue virus serotype 8; ML, maximum likelihood. (TIF)

**S2 Fig. Timescaled phylogenetic tree of BTV-8 samples collected during the European outbreaks between 2006 and 2018 with tip labels.** Maximum clade credibility time-calibrated phylogenetic tree generated in BEAST. The tree is scaled in years, with the final sampling date being October 2018. Clades with posterior support of 0.9 or higher are indicated by a white circle. Samples from the first outbreak are shown with a purple circle, while samples from the second are shown with orange circles. The branches are coloured accordingly to their median evolutionary rate across the posterior (see heatmap within the figure). BTV-8, bluetongue virus serotype 8. (TIF)

**S1 Table. Table of all sample IDs and their associated metadata.** Metadata corresponds to sampling time and location, species of isolation, tissue of isolation, whether the sample was cultured or not, passage history, sequencing method, and GenBank accession numbers. (XLSX)

## Acknowledgments

The Quadro P6000 graphics card used in this research was donated by the NVIDIA Corporation. We are extremely grateful to the World Organisation for Animal Health (OIE) for providing the data included in Fig 1. Results of the study are the sole responsibility of the authors

and not the OIE. We are grateful to local veterinarians, members of our laboratories, and Bethany Clark for helpful comments to the manuscript.

## Author Contributions

**Conceptualization:** Peter P. C. Mertens, Roman Biek, Massimo Palmarini.

**Data curation:** David J. Pascall, Kyriaki Nomikou.

**Formal analysis:** David J. Pascall, Kyriaki Nomikou.

**Funding acquisition:** Peter P. C. Mertens, Roman Biek, Massimo Palmarini.

**Investigation:** David J. Pascall, Kyriaki Nomikou, Emmanuel Bréard, Stephan Zientara, Ana da Silva Filipe, Maude Jacquot, Roman Biek.

**Methodology:** David J. Pascall, Kyriaki Nomikou, Ana da Silva Filipe, Maude Jacquot, Roman Biek.

**Project administration:** Kyriaki Nomikou, Roman Biek, Massimo Palmarini.

**Resources:** Emmanuel Bréard, Stephan Zientara, Bernd Hoffmann, Kris De Clercq, Anette Bøtner, Corinne Sailleau, Cyril Viarouge, Carrie Batten, Giantonella Puggioni, Ciriaco Ligios, Giovanni Savini, Piet A. van Rijn, Peter P. C. Mertens.

**Software:** Joshua B. Singer.

**Supervision:** Stephan Zientara, Peter P. C. Mertens, Roman Biek, Massimo Palmarini.

**Visualization:** David J. Pascall.

**Writing – original draft:** David J. Pascall, Roman Biek, Massimo Palmarini.

**Writing – review & editing:** David J. Pascall, Kyriaki Nomikou, Emmanuel Bréard, Stephan Zientara, Ana da Silva Filipe, Bernd Hoffmann, Maude Jacquot, Joshua B. Singer, Kris De Clercq, Anette Bøtner, Corinne Sailleau, Cyril Viarouge, Carrie Batten, Giantonella Puggioni, Ciriaco Ligios, Giovanni Savini, Piet A. van Rijn, Peter P. C. Mertens, Roman Biek, Massimo Palmarini.

## References

1. Smith K, Machalaba C, Seifman R, Feferholtz Y, Karesh W. Infectious disease and economics: The case for considering multi-sectoral impacts. *One Health*. 2019; 7:100080. <https://doi.org/10.1016/j.onehit.2018.100080> PMID: 30671528
2. Wilder-Smith A, Gubler D, Weaver S, Monath T, Heymann D, Scott T. Epidemic arboviral diseases: priorities for research and public health. *The Lancet Infectious Diseases*. 2017; 17:e101–e6. [https://doi.org/10.1016/S1473-3099\(16\)30518-7](https://doi.org/10.1016/S1473-3099(16)30518-7) PMID: 28011234
3. Mayer S, Tesh R, Vasilakis N. The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. *Acta Trop*. 2017; 166:155–63. <https://doi.org/10.1016/j.actatropica.2016.11.020> PMID: 27876643
4. Weaver S, Reisen W. Present and future arboviral threats. *Antiviral Res*. 2010; 85(2):328–45. <https://doi.org/10.1016/j.antiviral.2009.10.008> PMID: 19857523
5. Ratnien M, Caporale M, Golder M, Franzoni G, Allan K, Nunes S, et al. Identification and characterization of a novel non-structural protein of bluetongue virus. *PLoS Pathog*. 2011; 7(12):e1002477. <https://doi.org/10.1371/journal.ppat.1002477> PMID: 22241985
6. Mellor P, Baylis M, Mertens P. *Bluetongue*. Pastoret P-P, editor. London: Academic Press; 2009.
7. Roy P. Bluetongue virus: dissection of the polymerase complex. *J Gen Virol*. 2008; 89(Pt 8):1789–804. <https://doi.org/10.1099/vir.0.2008/002089-0> PMID: 18632949

8. Stewart M, Hardy A, Barry G, Pinto R, Caporale M, Melzi E, et al. Characterization of a second open reading frame in genome segment 10 of bluetongue virus. *J Gen Virol*. 2015; 96(11):3280–93. <https://doi.org/10.1099/jgv.0.000267> PMID: 26290332
9. Caporale M, Di Gialleonardo L, Janowicz A, Wilkie G, Shaw A, Savini G, et al. Virus and host factors affecting the clinical outcome of bluetongue virus infection. *J Virol*. 2014; 88(18):10399–411. <https://doi.org/10.1128/JVI.01641-14> PMID: 24991012
10. Maclachlan N, Drew C, Darpel K, Worwa G. The pathology and pathogenesis of bluetongue. *J Comp Pathol*. 2009; 141(1):1–16. <https://doi.org/10.1016/j.jcpa.2009.04.003> PMID: 19476953
11. Schwartz-Cornil I, Mertens P, Contreras V, Hemati B, Pascale F, Breard E, et al. Bluetongue virus: virology, pathogenesis and immunity. *Vet Res*. 2008; 39(5):46. <https://doi.org/10.1051/vetres:2008023> PMID: 18495078
12. Clarke L, Ruder M, Kienzle-Dean C, Carter D, Stallknecht D, Howerth E. Experimental Infection of White-Tailed Deer (*Odocoileus virginianus*) with Bluetongue Virus Serotype 3. *J Wildl Dis*. 2019; 55(3):627–36. <https://doi.org/10.7589/2018-06-159> PMID: 30605393
13. Linden A, Gregoire F, Nahayo A, Hanrez D, Mousset B, Massart A, et al. Bluetongue virus in wild deer, Belgium, 2005–2008. *Emerg Infect Dis*. 2010; 16(5):833–6. <https://doi.org/10.3201/eid1605.091217> PMID: 20409376
14. Lopez-Olvera J, Falconi C, Fernandez-Pacheco P, Fernandez-Pinero J, Sanchez M, Palma A, et al. Experimental infection of European red deer (*Cervus elaphus*) with bluetongue virus serotypes 1 and 8. *Vet Microbiol*. 2010; 145(1–2):148–52. <https://doi.org/10.1016/j.vetmic.2010.03.012> PMID: 20395075
15. Mauroy A, Guyot H, De Clercq K, Cassart D, Thiry E, Saegerman C. Bluetongue in captive yaks. *Emerg Infect Dis*. 2008; 14(4):675–6. <https://doi.org/10.3201/eid1404.071416> PMID: 18394296
16. Jauniaux T, De Clercq K, Cassart D, Kennedy S, Vandebussche F, Vandemeulebroucke E, et al. Bluetongue in Eurasian lynx. *Emerg Infect Dis*. 2008; 14(9):1496–8. <https://doi.org/10.3201/eid1409.080434> PMID: 18760034
17. Wilson A, Mellor P. Bluetongue in Europe: past, present and future. *Philos Trans R Soc Lond B Biol Sci*. 2009; 364(1530):2669–81. <https://doi.org/10.1098/rstb.2009.0091> PMID: 19687037
18. Baylis M, Caminade C, Turner J, Jones A. The role of climate change in a developing threat: the case of bluetongue in Europe. *Rev Sci Tech*. 2017; 36(2):467–78. <https://doi.org/10.20506/rst.36.2.2667> PMID: 30152470
19. Maclachlan N, Zientara S, Wilson W, Richt J, Savini G. Bluetongue and epizootic hemorrhagic disease viruses: recent developments with these globally re-emerging arboviral infections of ruminants. *Curr Opin Virol*. 2019; 34:56–62. <https://doi.org/10.1016/j.coviro.2018.12.005> PMID: 30654271
20. Elbers A, Popma J, Oosterwolde S, van Rijn P, Vellema P, van Rooij E. A cross-sectional study to determine the seroprevalence of bluetongue virus serotype 8 in sheep and goats in 2006 and 2007 in the Netherlands. *BMC Vet Res*. 2008; 4:33. <https://doi.org/10.1186/1746-6148-4-33> PMID: 18752681
21. Mehlhorn H, Walldorf V, Klimpel S, Jahn B, Jaeger F, Eschweiler J, et al. First occurrence of *Culicoides obsoletus*-transmitted Bluetongue virus epidemic in Central Europe. *Parasitol Res*. 2007; 101(1):219–28. <https://doi.org/10.1007/s00436-007-0519-6> PMID: 17385085
22. Szmargd C, Wilson A, Carpenter S, Mertens P, Mellor P, Gubbins S. Mortality and case fatality during the recurrence of BTv-8 in northern Europe in 2007. *Vet Rec*. 2007; 161(16):571–2. <https://doi.org/10.1136/vr.161.16.571-e> PMID: 17951569
23. Elbers A, Backx A, Ekker H, van der Spek A, van Rijn P. Performance of clinical signs to detect bluetongue virus serotype 8 outbreaks in cattle and sheep during the 2006-epidemic in The Netherlands. *Vet Microbiol*. 2008; 129(1–2):156–62. <https://doi.org/10.1016/j.vetmic.2007.10.034> PMID: 18164148
24. Elbers A, Backx A, Meroc E, Gerbier G, Staubach C, Hendrickx G, et al. Field observations during the bluetongue serotype 8 epidemic in 2006. I. Detection of first outbreaks and clinical signs in sheep and cattle in Belgium, France and the Netherlands. *Prev Vet Med*. 2008; 87(1–2):21–30. <https://doi.org/10.1016/j.prevetmed.2008.06.004> PMID: 18620767
25. Elbers A, Backx A, Mintiens K, Gerbier G, Staubach C, Hendrickx G, et al. Field observations during the Bluetongue serotype 8 epidemic in 2006. II. Morbidity and mortality rate, case fatality and clinical recovery in sheep and cattle in the Netherlands. *Prev Vet Med*. 2008; 87(1–2):31–40. <https://doi.org/10.1016/j.prevetmed.2008.06.003> PMID: 18620768
26. Velthuis A, Saatkamp H, Mourits M, de Koeijer A, Elbers A. Financial consequences of the Dutch bluetongue serotype 8 epidemics of 2006 and 2007. *Prev Vet Med*. 2010; 93(4):294–304. <https://doi.org/10.1016/j.prevetmed.2009.11.007> PMID: 19962204
27. Rushton J, Lyons N. Economic impact of Bluetongue: a review of the effects on production. *Vet Ital*. 2015; 51(4):401–6. PMID: 26741252

28. Hasler B, Howe K, Di Labio E, Schwermer H, Stark K. Economic evaluation of the surveillance and intervention programme for bluetongue virus serotype 8 in Switzerland. *Prev Vet Med.* 2012; 103(2–3):93–111. <https://doi.org/10.1016/j.prevetmed.2011.09.013> PMID: 22018548
29. Pinior B, Lebl K, Firth C, Rubel F, Fuchs R, Stockreiter S, et al. Cost analysis of bluetongue virus serotype 8 surveillance and vaccination programmes in Austria from 2005 to 2013. *Vet J.* 2015; 206(2):154–60. <https://doi.org/10.1016/j.tvjl.2015.07.032> PMID: 26371833
30. Zientara S, Sanchez-Vizcaino J. Control of bluetongue in Europe. *Vet Microbiol.* 2013; 165(1–2):33–7. <https://doi.org/10.1016/j.vetmic.2013.01.010> PMID: 23462519
31. Sailleau C, Bréard E, Viarouge C, Vitour D, Romey A, Garnier A, et al. Re-Emergence of Bluetongue Virus Serotype 8 in France, 2015. *Transboundary and emerging diseases.* 2017; 64:998–1000. <https://doi.org/10.1111/tbed.12453> PMID: 26617414
32. Gale P, Gauntlett F, Bowen J. Bluetongue virus (BTV-8) in Germany and Belgium. Report of the UK Department for Environment Food and Rural Affairs. 2019 [cited 2019 Dec 1]. <https://www.gov.uk/government/collections/animal-diseases-international-monitoring>
33. Courtejoie N, Durand B, Bournez L, Gorlier A, Breard E, Sailleau C, et al. Circulation of bluetongue virus 8 in French cattle, before and after the re-emergence in 2015. *Transbound Emerg Dis.* 2018; 65(1):281–4. <https://doi.org/10.1111/tbed.12652> PMID: 28464486
34. Breard E, Sailleau C, Quenault H, Lucas P, Viarouge C, Touzain F, et al. Complete Genome Sequence of Bluetongue Virus Serotype 8, Which Reemerged in France in August 2015. *Genome announcements.* 2016; 4(2).
35. Rossi S, Balenghien T, Viarouge C, Faure E, Zanella G, Sailleau C, et al. Red deer (*Cervus elaphus*) did not play the role of maintenance host for bluetongue virus in France: The burden of proof by long-term wildlife monitoring and *Culicoides* snapshots. *Viruses.* 2019; 11:E903. <https://doi.org/10.3390/v11100903> PMID: 31569721
36. Grego E, Sossella M, Bisanzio D, Stella M, Giordana G, Pignata L, et al. Wild ungulates as sentinel of BTV-8 infection in piedmont areas. *Veterinary Microbiology.* 2014; 174:93–33. <https://doi.org/10.1016/j.vetmic.2014.09.015> PMID: 25306211
37. Courtejoie N, Bournez L, Zanella G, Durand B. Quantifying bluetongue vertical transmission in French cattle from surveillance data. *Veterinary Research.* 2019; 50:34. <https://doi.org/10.1186/s13567-019-0651-1> PMID: 31088555
38. Bournez L, Cavalerie L, Sailleau C, Bréard E, Zanella G, Servan de Almeida R, et al. Estimation of French cattle herd immunity against bluetongue serotype 8 at the time of its re-emergence in 2015. *BMC Veterinary Research.* 2018; 14:65. <https://doi.org/10.1186/s12917-018-1388-1> PMID: 29499711
39. Nomikou K, Hughes J, Wash R, Kellam P, Breard E, Zientara S, et al. Widespread Reassortment Shapes the Evolution and Epidemiology of Bluetongue Virus following European Invasion. *PLoS Pathog.* 2015; 11(8):e1005056. <https://doi.org/10.1371/journal.ppat.1005056> PMID: 26252219
40. Carpi G, Holmes E, Kitchen A. The evolutionary dynamics of bluetongue virus. *Journal of Molecular Evolution.* 2010; 70:583–92. <https://doi.org/10.1007/s00239-010-9354-y> PMID: 20526713
41. Worobey M. Phylogenetic evidence against evolutionary stasis and natural abiotic reservoirs of influenza A virus. *J Virol.* 2008; 82(7):3769–74. <https://doi.org/10.1128/JVI.02207-07> PMID: 18234791
42. Osborne C, Mayo C, Mullens B, McDermott E, Gerry A, Reisen W, et al. Lack of Evidence for Laboratory and Natural Vertical Transmission of Bluetongue Virus in *Culicoides sonorensis* (Diptera: Ceratopogonidae). *J Med Entomol.* 2015; 52(2):274–7. <https://doi.org/10.1093/jme/tju063> PMID: 26336312
43. Nunamaker R, Sieburth P, Dean V, Wigington J, Nunamaker C, Mecham J. Absence of transovarial transmission of bluetongue virus in *Culicoides variipennis*: immunogold labelling of bluetongue virus antigen in developing oocytes from *Culicoides variipennis* (Coquillett). *Comp Biochem Physiol A Comp Physiol.* 1990; 96(1):19–31. [https://doi.org/10.1016/0300-9629\(90\)90036-r](https://doi.org/10.1016/0300-9629(90)90036-r) PMID: 1975536
44. White D, Wilson W, Blair C, Beaty B. Studies on overwintering of bluetongue viruses in insects. *J Gen Virol.* 2005; 86(Pt 2):453–62. <https://doi.org/10.1099/vir.0.80290-0> PMID: 15659765
45. Wilson A, Darpel K, Mellor P. Where does bluetongue virus sleep in the winter? *PLoS Biol.* 2008; 6(8):e210. <https://doi.org/10.1371/journal.pbio.0060210> PMID: 18752350
46. Nowak M, Bonhoeffer S, Hill A, Boehme R, Thomas H, McDade H. Viral dynamics in hepatitis B virus infection. *Proc Natl Acad Sci U S A.* 1996; 93(9):4398–402. <https://doi.org/10.1073/pnas.93.9.4398> PMID: 8633078
47. Koyuncu O, MacGibeny M, Enquist L. Latent versus productive infection: the alpha herpesvirus switch. *Future Virol.* 2018; 13(6):431–43. <https://doi.org/10.2217/fvl-2018-0023> PMID: 29967651
48. Cohrs R, Gilden D. Human herpesvirus latency. *Brain Pathol.* 2001; 11(4):465–74. <https://doi.org/10.1111/j.1750-3639.2001.tb00415.x> PMID: 11556692

49. Bangham C. Human T Cell Leukemia Virus Type 1: Persistence and Pathogenesis. *Annu Rev Immunol*. 2018; 36:43–71. <https://doi.org/10.1146/annurev-immunol-042617-053222> PMID: 29144838
50. Kulkarni A, Bangham C. HTLV-1: Regulating the Balance Between Proviral Latency and Reactivation. *Front Microbiol*. 2018; 9:449. <https://doi.org/10.3389/fmicb.2018.00449> PMID: 29615991
51. Coffin J, Swanstrom R. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harbor perspectives in medicine*. 2013; 3(1):a012526. <https://doi.org/10.1101/cshperspect.a012526> PMID: 23284080
52. Mbala-Kingebeni P, Pratt C, Wiley M, Diagne M, Makiala-Mandanda S, Aziza A, et al. 2018 Ebola virus disease outbreak in Equateur Province, Democratic Republic of the Congo: a retrospective genomic characterisation. *Lancet Infect Dis*. 2019; 19(6):641–7. [https://doi.org/10.1016/S1473-3099\(19\)30124-0](https://doi.org/10.1016/S1473-3099(19)30124-0) PMID: 31000465
53. Lam T, Zhu H, Chong Y, Holmes E, Guan Y. Puzzling Origins of the Ebola Outbreak in the Democratic Republic of the Congo, 2014. *J Virol*. 2015; 89(19):10130–2. <https://doi.org/10.1128/JVI.01226-15> PMID: 26202242
54. Gautret P, Carrara P, Parola P. Long incubation in imported human rabies. *Annals of Neurology*. 2014; 75:324–5.
55. Cortey M, Ferretti L, Pérez-Martín E, Zhang F, de Klerk-Lorist L-M, Scott K, et al. Persistent Infection of African Buffalo (*Syncerus caffer*) with Foot-and-Mouth Disease Virus: Limited Viral Evolution and No Evidence of Antibody Neutralization Escape. *Journal of Virology*. 2019; 93:e00563–19. <https://doi.org/10.1128/JVI.00563-19> PMID: 31092573
56. Arzt J, Fish I, Pauszek S, Johnson S, Chain P, Rai D, et al. The evolution of a super-swarm of foot-and-mouth disease virus in cattle. *PLoS ONE*. 2019; 14:e0210847. <https://doi.org/10.1371/journal.pone.0210847> PMID: 31022193
57. Gray R, Parker J, Lemey P, Salemi M, Katzourakis A, Pybus O. The mode and tempo of hepatitis C virus evolution within and among hosts. *BMC Evol Biol*. 2011; 11:131. <https://doi.org/10.1186/1471-2148-11-131> PMID: 21595904
58. Nakajima K, Desselberger U, Palese P. Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature*. 1978; 274(5669):334–9. <https://doi.org/10.1038/274334a0> PMID: 672956
59. Rozo M, Gronvall G. The reemergent 1977 H1N1 strain and the gain-of-function debate. *mBio*. 2015; 6:e01013–15. <https://doi.org/10.1128/mBio.01013-15> PMID: 26286690
60. Brault A, Powers A, Medina G, Wang E, Kang W, Salas R, et al. Potential Sources of the 1995 Venezuelan Equine Encephalitis Subtype IC Epidemic. *Journal of Virology*. 2001; 75:5823–32. <https://doi.org/10.1128/JVI.75.13.5823-5832.2001> PMID: 11390583
61. Anderson I. Foot and mouth disease 2007: a review and lessons learned. The Stationery Office; 2008.
62. Moore S, Hasler J. A 100-Year Review: Reproductive technologies in dairy science. *J Dairy Sci*. 2017; 100(12):10314–31. <https://doi.org/10.3168/jds.2017-13138> PMID: 29153167
63. Stevenson J, Britt J. A 100-Year Review: Practical female reproductive management. *J Dairy Sci*. 2017; 100(12):10292–313. <https://doi.org/10.3168/jds.2017-12959> PMID: 29153166
64. Thomas F, Singh E, Hare W. Control of bluetongue virus spread by embryo transfer. *Progress in clinical and biological research*. 1985; 178:653–4. PMID: 2989915
65. Backx A, Heutink C, van Rooij E, van Rijn P. Clinical signs of bluetongue virus serotype 8 infection in sheep and goats. *Vet Rec*. 2007; 161(17):591–2. <https://doi.org/10.1136/vr.161.17.591> PMID: 17965371
66. Haegeman A, Vandaele L, De Leeuw I, Oliveira A, Nauwynck H, Van Soom A, et al. Failure to remove bluetongue serotype 8 virus (BTV-8) from in vitro produced and in vivo derived bovine embryos and subsequent transmission of BTV-8 to recipient cows after embryo transfer. *Frontiers in Veterinary Science*. 2019; 6(432). <https://doi.org/10.3389/fvets.2019.00432>
67. Backx A, Heutink R, van Rooij E, van Rijn P. Transplacental and oral transmission of wild-type bluetongue virus serotype 8 in cattle after experimental infection. *Vet Microbiol*. 2009; 138(3–4):235–43. <https://doi.org/10.1016/j.vetmic.2009.04.003> PMID: 19419822
68. OIE. Terrestrial Animal Health Code. Paris: OIE; 2019.
69. Ministère de l'agriculture. Fixant les conditions sanitaires exigées pour les agréments visés à l'article L. 222–1 du code rural dans le cadre de la monte publique artificielle des animaux de l'espèce bovine. In: *Journal Officiel de la République Française* 2008. p. 1143.
70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943

71. Wickham H. ggplot2: Elegant Graphics for Data Analysis New York: Springer-Verlag; 2016 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/ggplot2/index.html>
72. Charif D, Lobry J. Seqin{R} 1.0–2: a contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis. Structural approaches to sequence evolution: Molecules, networks, populations. New York: Springer Verlag; 2007. p. 207–32.
73. Wickham H. stringr: Simple, Consistent Wrappers for Common String Operations 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/stringr/index.html>
74. Knaus B, Grünwald N. vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources. 2017; 17:44–53. <https://doi.org/10.1111/1755-0998.12549> PMID: 27401132
75. Magoč T, Salzberg S. FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011; 27:2957–63. <https://doi.org/10.1093/bioinformatics/btr507> PMID: 21903629
76. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012; 9:357–9. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
77. Roberts A, Pachter L. Streaming fragment assignment for real-time analysis of sequencing experiments. Nature Methods. 2013; 10:71–3. <https://doi.org/10.1038/nmeth.2251> PMID: 23160280
78. Wilm A, Aw P, Bertrand D, Yeo G, Ong S, Wong C, et al. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Research. 2012; 40:11189–201. <https://doi.org/10.1093/nar/gks918> PMID: 23066108
79. Nagarajan N, Kingsford C. GiRaF: Robust, computational identification of influenza reassortments via graph mining. Nucleic Acids Research. 2011; 39:e34. <https://doi.org/10.1093/nar/gkq1232> PMID: 21177643
80. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001; 17(8):754–5. <https://doi.org/10.1093/bioinformatics/17.8.754> PMID: 11524383
81. Guindon S, Dufayard J, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology. 2010; 59:307–21. <https://doi.org/10.1093/sysbio/syq010> PMID: 20525638
82. Suchard M, Lemey P, Baele G, Ayres D, Drummond A, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution. 2018; 4:vey016. <https://doi.org/10.1093/ve/vey016> PMID: 29942656
83. Guindon S, Delsuc F, Dufayard J, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol. 2009; 537:113–37. [https://doi.org/10.1007/978-1-59745-251-9\\_6](https://doi.org/10.1007/978-1-59745-251-9_6) PMID: 19378142
84. Rambaut A, Lam T, Max Carvalho L, Pybus O. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evolution. 2016; 2:vew007. <https://doi.org/10.1093/ve/vew007> PMID: 27774300
85. Drummond A, Ho S, Phillips M, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006; 4:699–710.
86. Posada D. jModelTest: phylogenetic model averaging. Mol Biol Evol. 2008; 25(7):1253–6. <https://doi.org/10.1093/molbev/msn083> PMID: 18397919
87. Minin V, Bloomquist E, Suchard M. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Molecular Biology and Evolution. 2008; 25:1459–71. <https://doi.org/10.1093/molbev/msn090> PMID: 18408232
88. Wertheim J. The re-emergence of H1N1 influenza virus in 1977: a cautionary tale for estimating divergence times using biologically unrealistic sampling dates. PLoS ONE. 2010; 5(6):e11184. <https://doi.org/10.1371/journal.pone.0011184> PMID: 20567599
89. Bürkner P. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software. 2017; 80:1–28.
90. Yu G, Smith D, Zhu H, Guan Y, Lam T. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 2017; 8:28–36.
91. Arnold J. ggthemes: Extra Themes, Scales and Geoms for 'ggplot2' 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/ggthemes/index.html>
92. Wilke C. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2' 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/cowplot/index.html>
93. Kahle D, Wickham H. ggmap: Spatial visualization with ggplot2. The R Journal. 2013; 5:144–61.
94. Garnier S. viridis: Default Color Maps from 'matplotlib' 2018 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/viridis/index.html>

95. Kay M. tidybayes: Tidy Data and Geoms for Bayesian Models 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/tidybayes/index.html>
96. Grolemund G, Wickham H. Dates and times made easy with lubridate. *Journal of Statistical Software*. 2011; 40.
97. Pebesma E, Bivand RS. Classes and methods for spatial data in {R}. *R News*. 2005; 5.
98. Hijmans R. raster: Geographic Data Analysis and Modeling 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/raster/index.html>
99. Bivand R, Lewin-Koh N. maptools: Tools for Handling Spatial Objects 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/maptools/index.html>
100. Bivand R, Rundel C. rgeos: Interface to Geometry Engine—Open Source ('GEOS') 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/rgeos/index.html>
101. Bivand R, Keitt T, Rowlingson B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/rgdal/index.html>
102. Pebesma E. Simple features for R: Standardized support for spatial vector data. *The R Journal*. 2018; 10:439–46.
103. Schnute J, Boers N, Haigh R. PBSmapping: Mapping Fisheries Data and Spatial Analysis Tools 2019 [cited 2019 Dec 1]. <https://cran.r-project.org/web/packages/PBSmapping/index.html>