What can phylodynamics bring to animal health research?

Claire Guinat, Timothee Vergne, Arthur Kocher, Debapryio Chakraborty, Mathilde C. Paul, Mariette Ducatez, Tanja Stadler

What can phylodynamics bring to animal health research?

Review Article

Author(s):
Guinat, Claire; Vergne, Timothée; Kocher, Arthur; Chakraborty, Debapryio; Paul, Mathilde C.; Ducatez, Mariette; Stadler, Tanja

Publication date:
2021-09

Permanent link:
https://doi.org/10.3929/ethz-b-000487306

Originally published in:
Trends in Ecology & Evolution 36(9), https://doi.org/10.1016/j.tree.2021.04.013

Funding acknowledgement:
842621 - Digging deeper into genes to track infectious disease outbreaks (EC)

Title: What can phylodynamics bring to animal health research?

Authors names: Claire Guinat[1,2*], Timothee Vergne[3], Arthur Kocher[4], Debapryio Chakraborty[3], Mathilde C. Paul[3], Mariette Ducatez[3], Tanja Stadler[1,2]

Affiliations:

[1]Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 4058 Basel, Switzerland

[2]Swiss Institute of Bioinformatics (SIB), Switzerland

[3]IHAP, Université de Toulouse, INRAE, ENVT, 23 chemin des cappelles, 31300 Toulouse, France

[4]Transmission, Infection, Diversification & Evolution (tide) group, Max Planck Institute for the Science of Human History, Kahlaische str. 10, 07745 Jena, Germany

***Corresponding author**
Dr. Claire Guinat
DVM, MSc, PhD
Computational Evolution group
Mattenstrasse 26
4058 Basel
Switzerland
claire.guinat@bsse.ethz.ch

Keywords: phylodynamics, molecular epidemiology, population dynamics, pathogen evolution and transmission, wildlife and livestock populations

Summary (120 words) 116

Infectious diseases are a major burden to global economies, public and animal health. To date, quantifying the spread of infectious diseases to inform policy making has traditionally relied on epidemiological data collected during epidemics. However, interest has grown on recent phylodynamic techniques to infer pathogen transmission dynamics from genetic data. Here, we provide examples where this new discipline has enhanced disease management in public health and illustrate how it could be further applied in animal health. In particular, we describe how phylodynamics can address fundamental epidemiological questions, such as inferring key transmission parameters in animal populations and

quantifying spill-over events at the wildlife-livestock interface, and generate important insights for the design of more effective control strategies.

Text (3500 words) 3492

Controlling the spread of zoonotic diseases

Infectious diseases represent a major concern for public and animal health and global economy, as recent epidemics have resulted in vast socioeconomic damages and in a severe loss of human and animal life. Most of emerging infectious diseases that are reported globally in humans are **zoonotic** (see Glossary), that is, they are naturally transmitted from animals to humans [1,2]. Major examples in the recent past include globally devastating epidemics such as pandemic influenza H1N1 in 2009, Ebola in 2014, coronavirus diseases (i.e. SARS in 2003, MERS in 2012 and COVID-19 in 2020) and Zika in 2015 [1,2].

Experiences from most recent outbreaks indicate that an effective control of zoonotic diseases requires an integrated collaboration between the animal and human health sectors under the One Health concept [3]. For some diseases, human infections can be prevented through interventions in animal populations, for instance, in rabies with vaccination campaigns for dogs and control of free-roaming canid populations [4,5]. Also, active disease surveillance in animal populations is essential in providing early warning for public health authorities, as is the case for the West Nile fever [6]. Thus, prior knowledge on how infectious diseases spread and can be controlled in animal populations remains crucial for reducing the risk of animal-to-human transmission.

Quantifying the transmission dynamics of animal infectious diseases represents one of our longstanding interests to inform policy making. This is usually addressed by fitting mathematical models to epidemiological data (number, date and location of infected animals

or premises) collected during outbreaks to estimate key transmission parameters, such as the **basic reproduction number** (R_0, expected number of secondary infections generated by an infected epidemiological unit in a totally susceptible population) [7]. However, recent extensions in phylogenetics, coined as phylodynamics [8], mean that such question can now be addressed based on the integrated use of epidemiological and genetic data, expanding our capacity to understand disease transmission dynamics and identify suitable interventions.

While phylodynamic approaches have been fruitfully applied to a range of global public health threats, such as COVID-19 [9–11], Ebola [12–14], Zika [15,16], Hepatitis C [17], HIV infection [17–21] and dengue [22], they remain overlooked in animal health research, restricting our ability to gain as much insight as possible from epidemics to design more effective control strategies. We provide examples where this discipline has enhanced human infectious disease management and illustrate how they can be further extended to the field of animal infectious diseases. In particular, we describe how phylodynamics can address fundamental epidemiological questions in animal health research, such as inferring key transmission parameters in animal populations and quantifying **spill-over events** at the wildlife-livestock interface.

Inferring disease transmission dynamics with epidemiological data

One of the key transmission parameters to estimate during epidemics is the basic reproduction number (R_0) (or the **effective reproduction number** (R_e), which captures the number of secondary infections generated at any time during an epidemic from a partially immune population, also expressed as R_t to reflect changes over time) [7]. The magnitude of R estimates is a useful indicator of the disease transmissibility, epidemic sustainability and control strategies effectiveness. If $R < 1$, the disease will die out in the population because, on average, an infectious individual infects less than one individual; if $R > 1$, the disease will

spread exponentially. The objective of control programmes is thus to identify and implement strategies that decrease the *R* values to less than 1 such that the disease cannot persist in the population. Alternatively, predictive mathematical models use *R* estimates as inputs to simulate and assess the effect of control strategies during an epidemic [23–25]. Thus, accurate estimation of *R* values is crucial for disease policy formulation (Box 1).

The 2001 foot-and-mouth disease (FMD) epidemic in the UK provides a noteworthy example of how such transmission parameters have played a central role in guiding the development of policies in animal health [26]. Outbreaks of FMD caused severe disruption and economic losses for the country as a whole [27,28] and raised concerns about the management and control of infectious diseases in livestock [29]. Extensive epidemiological data generated during outbreak investigations stimulated the development of a variety of mathematical models that aimed at estimating *R* values to quantify the disease transmission in different domestic animal populations [30,31], including cattle [32], sheep [33,34] and pigs [33,35] and to evaluate the effectiveness of strategies, including vaccination [36–39], culling [31,36,37] or ban of livestock movement [40] on the disease spread.

Although epidemiological data is of paramount importance for quantifying key transmission parameters during epidemics, genetic data is another extremely valuable source of information, which has become increasingly available with the recent advances of sequencing technologies. The growing accessibility of genetic data has subsequently led to recent extensions in **phylogenetics**, coined as **phylodynamics** [8], that aim to generate precise insights into disease transmission dynamics based on the integration of various sources of information, including genetic and epidemiological data.

Inferring disease transmission dynamics with genetic data

First steps in phylogenetics

Genetic data has been specifically used to infer **phylogenetic trees** that show the ancestral relationships among pathogens. In their simplest forms, reconstructing phylogenetic trees requires a sequence alignment and a nucleotide substitution model that describes the nature of nucleotide substitutions (Figure 1) [41]. The model is used to calculate the likelihood of various possible phylogenetic trees, from which the most likely tree is identified. Parts of a phylogenetic tree include: the tips representing the sequences, the branch lengths representing the genetic change (number of nucleotide substitutions) between sequences and the nodes joining two branches representing the most recent common ancestor (MRCA) of the subsequent sequences. A set of sequences with relatively small genetic differences compared to differences between other sequences will be thus clustered together in the tree.

In animal health, phylogenetic analyses have been traditionally performed to determine the geographical origin of emerging outbreaks. For example, African swine fever (ASF) virus strains sampled from infected pigs in Georgia in 2007 were identified as closely related to strains circulating in Mozambique, Madagascar and Zambia [42], pointing to an introduction from this region. The underlying explanations were that ASF virus, that is currently responsible of severe outbreaks in domestic pig and wild boar populations in Asia and Europe, entered Georgia via imports of contaminated pig products from Eastern Africa or Madagascar. Phylogenetic analyses also helped detect pathogen genetic diversity during outbreaks. Rift Valley Fever (RVF) is a zoonotic mosquito-borne disease associated with severe economic losses due to death and abortion in infected livestock. During the sequential RVF epidemics in Eastern Africa in 2006, phylogenetic analyses pointed to the existence of multiple viral lineages prevailing in different geographical locations [43]. This suggested the maintenance of RVF virus in these endemic regions through the hatching of infected *Aedes* mosquito eggs or from the low-level cycling of the virus in livestock and humans. Likewise, phylogenetic analyses helped identifying spill-overs of pathogens between species. Examples

include recent high-profile subtypes of avian influenza viruses (AIV), such as H5N1 which continues to cause major economic costs with the loss of millions of poultry across Asia and Europe and has raised concern as a potential pandemic threat [44]. Phylogenetic analysis of the 1997 highly pathogenic avian influenza H5N1 outbreak in Honk Kong showed that the circulating viruses were reassortants from H5N1 viruses in poultry and H9N2 and/or H6N1 viruses in quail [45]. This led to strict restrictions in Hong Kong's live poultry markets, such as banning the sale of quail with other poultry species together and thorough cleaning of the market on a monthly 'rest-day'.

The new era of phylodynamics

One of the aims of phylodynamics is to infer disease transmission dynamics based on the genome of pathogens sampled during epidemics and associated epidemiological data. Phylodynamics have been mostly developed for the study of rapidly evolving pathogens, especially RNA viruses, for which a large amount of genetic variation occurs over the course of a single epidemic. In such cases, pathogen evolutionary and epidemiological timescales are likely similar, meaning that phylogenetic trees can be used as proxies of the transmission trees [46]. Phylodynamics rely on **molecular clock models** [47,48] and on **population dynamic models** (57) (Figure 1). Molecular clock models describe the rate at which nucleotide substitutions occur over time. They allow to convert the branch lengths of a phylogenetic tree from units of nucleotide substitution to units of time and thus, to estimate the time of the MRCA (tMRCA), providing a lower bound estimate for the onset of the epidemic. Population dynamic models describe the history of population dynamics based on genetic sequences sampled the population (Box 1).

Thus, phylodynamics have proven useful in infectious disease epidemiology and have been fruitfully applied to diverse global public health threats. Phylodynamic analyses have revealed temporal changes in R_e estimates, providing insights into the epidemiological

160 determinants of disease spread, such as past public health interventions. Phylodynamic
161 inferences from Egyptian Hepatitis C virus (HCV) sequences showed that the increase in R_e
162 estimates coincided with the time when the antischistosomal injection campaigns started,
163 around 1920 [17], supporting the hypothesis that HCV epidemics were caused by viral
164 contamination of antischistosomal injections. Moreover, phylodynamic analyses have helped
165 to highlight heterogeneity in R_e estimates between individuals, allowing the identification of
166 transmission differences within and between risk groups. As an example, the estimates from
167 genetic data showed that HIV transmission dynamics in Latvia were driven by injecting drug
168 user groups between 1998 and 2002 and that the heterosexual groups played a smaller role in
169 HIV transmission [19]. Phylodynamic inferences have also demonstrated the possible
170 estimation of the timeline and the number of unreported cases, since the sampling rate is a
171 parameter that can be estimated in BD models. During the 2014 Ebola virus outbreak in Sierra
172 Leone, the sampling proportion was estimated at around 70% based on phylodynamic
173 inferences, which was in line with previous reports [12].

174 ***Applying phylodynamics in animal health research***

175 *Quantifying disease transmission dynamics within animal populations*

176 First, phylodynamic approaches can be valuable to quantify disease transmission
177 dynamics within animal populations in animal health. Similar to public health, this could
178 provide opportunities to generate hypotheses related to the impact of past interventions.

179 Peste des petits ruminants (PPR) is a highly contagious viral disease of small
180 ruminants, causing severe mortality and production loss in developing countries [50]. Despite
181 control efforts including live animal movement control and mass annual vaccination, the
182 disease continues to spread across Asia and Africa and threatens farmers' livelihoods and
183 international trade in affected countries. Given the rapid and on-going spread of PPR across
184 Asia and Africa, it was crucial to determine whether past interventions had any significant

effect on disease transmission dynamics. Phylodynamic inferences using PPR virus sequences from different regions in Asia and Africa showed a decrease in N_e estimates since late 1990s, which was likely associated to effective control strategies, such as the start of mass vaccination campaign in different regions [51]. Other phylodynamic inferences based on Chinese PPR viral sequences revealed an increase in N_e estimates at the beginning of 2014, which was likely associated with increasing lamb meat trade during Spring Festival [52]. Similarly, phylodynamic approaches were used to support the paramount importance of vaccination strategies during outbreaks of Infectious Bronchitis virus (IBV), a highly infectious respiratory disease of poultry, responsible for substantial economic loss for the poultry industry [53]. The high-profile of IBV is mainly linked to its high variability and rapid evolution, leading to a large number of strains circulating during outbreaks and limited vaccine effectiveness due to poor cross-protection. Analyses of IBV sequences from Italian poultry farms demonstrated a decrease in N_e estimates between 2013 and 2014, reflecting the implementation of the large-scale vaccination trial, followed by a sharp increase in 2015, likely associated with stop of the trial since the reduction in the number of outbreaks was considered unsatisfactory by the authorities [54]. Similar analyses were applied to ASF virus sequences in Eurasia and Africa and showed a continuous increase in N_e estimates from the 18^{th} to the 21^{th} century, which, according to the authors, was likely attributed to the continuous viral circulation in the region as well as the growing intensity of pig trade during the 19^{th} century [55]. As a last example, phylodynamic analysis was applied to West Nile virus (WNV) sequences obtained from birds, horses, mosquitoes and humans in Italy between 2011 and 2018 [56], with results showing an increase in R_e estimates in 2015-16 and 2017-18. Since humans and horses are dead-end hosts in the epidemiological cycle of WNV, this suggested that the increase in R_e was linked to an increase in reservoir birds or vectors during this period.

Quantifying disease transmission dynamics between animal populations

With the recent development of structured population dynamic models, phylodynamics also represent a promising technique to quantify disease transmission dynamics between animal populations in animal health, in particular at the wildlife-livestock interface (Figure 3).

While a number of phylodynamic studies have been attempted, including examples with FMD, ASF, avian influenza and Newcastle, most of them have relied on **discrete trait analysis** (DTA) **models** [51,57–65], rather than structured population dynamic models. Originating from the field of phylogeography, DTA models were adapted to describe the rate at which pathogen lineages transmit between different species as analogous to a nucleotide substitution process, independently from the population dynamics [66]. These approaches provided quantitative insights into disease transmission dynamics (i) between different bird populations during H5N8 outbreaks in Europe, North America and Asia [57,58], H9N2 and H5N1 outbreaks in Asia and Africa [59,60,67,68] and Newcastle disease outbreaks in China [69,70], (ii) between domestic pigs, wild suids, and *Ornithodoros* ticks during ASF outbreaks in Eurasia and Africa [55], between different pig farming systems during porcine reproductive and respiratory syndrome virus (PRRSV) outbreaks in North America [64,71] and (iii) between cattle and African buffalo (*Syncerus caffer*) or water buffalo (*Bubalus bubalis*) during FMD outbreaks in Eastern Africa [61] or Asia [62], respectively. However, despite their higher computational efficiency as compared to structured population dynamic models, DTA models have been identified as very sensitive to heterogeneous sampling across species, which can result in biased estimates of transmission parameters [72,73].

Very few research studies have implemented structured population dynamic models as an alternative approach to DTA. Examples concern the H5N2 outbreak in North America [63,74], in which transmission parameters were estimated between different bird species using

a multi-type BD model. Authors showed that the wild bird-to-poultry and wild bird-to-wild bird transmission rates were greater than the poultry-to-wild bird and poultry-to-poultry transmission rates, respectively [63], suggesting that the transmission could not be maintained in poultry without spill-overs from wild birds (Figure 4). They also showed that the transmission rate from layer chicken to turkey populations was greater than the reverse [74], indicating that the poultry type played a role in the H5N2 outbreak dynamics. Another example concerns PRRSV outbreaks [71] that remains the most important pig disease in the United States, due to the continuous emergence of new outbreaks resulting in severe economic losses for the pig industry [75]. Based on a structured coalescent model, authors investigated the transmission events leading to the occurrence of PRRSV outbreaks in different pig farming systems in North America and showed a clear transmission dominance occurring from farms of non-commercially related pig systems compared to farms from the same pig system.

Expanding the field of phylodynamics to slowly evolving pathogens

Although phylodynamics are best applied to rapidly evolving pathogens, the development of whole genome sequencing technologies has also made possible the expansion of this field to the study of slowly evolving pathogens [76,77].

Mycobacterium bovis (*M. bovis*), the aetiological agent of bTB, represents a key example in animal health to illustrate how phylodynamics offer promises to quantify disease transmission dynamics, even for slowly evolving pathogens. *M. bovis* is thought to be maintained in a variety of wild animals, from which spill-over infections are associated with outbreaks in livestock and humans [78]. bTB in cattle has been effectively controlled in many countries thanks to regular tuberculin testing programmes associated with slaughter of infected herds. In other countries such as the UK, Ireland and New Zealand, eradication efforts have been hampered by wildlife reservoirs, that are suspected to continually re-infect

cattle populations. Again, the majority of *M. bovis* cross-species transmission parameters have been estimated using DTA models. They allowed inferring *M. bovis* transmission rates between elk (*Cervus elaphus nelsoni*), white tailed deer (*Odocoileus virginianus*) and cattle populations in North America [79]. Estimates showed that the source of bTB infection for cattle was more likely for deer than elk populations, suggesting that bTB eradication efforts should be targeted at deer populations rather than elk. DTA models based on *M. bovis* sequences from cattle and different wildlife populations in New Zealand provided high estimates of transmission rates between wildlife and cattle [80]. Despite being unable to estimate the transmission direction due to low genetic variability of *M. bovis* and asymmetric sampling across species, the low bTB prevalence in New Zealand cattle populations supported the fact that wildlife populations are likely acting as reservoirs and should remain the target of control campaigns. One recent phylodynamic study based on a structured coalescent model allowed providing more robust estimates of *M. bovis* cross-species transmission in the UK [81]. The directional cross-species transmission rates indicated that badger-to-cattle transmission occurred more frequently than cattle-to-badger and that within-species transmission was greater than between-species transmission, highlighting the importance of targeting both populations for effective control.

Concluding remarks and future directions

Infectious diseases are important concerns in animal health, and their effective prevention and control is widely expected to improve human health. Existing studies in public health have demonstrated how insightful phylodynamic approaches could be, highlighting the value of genetic data for a broader understanding of infectious disease transmission. To date, phylodynamics are not common in animal health research, impacting our collective ability to effectively manage infectious diseases at the wildlife-livestock interface. Although there are a

number of applications, they are not systematically based on population dynamic models (structured or not) that allow for the direct estimation of key epidemiological parameters (i.e., R_e, β, δ or ρ). Moreover, the majority of studies do not fully leverage all benefits of phylodynamic tools, while more detailed information can be obtained, including estimates of the timeline and the total number of cases/first imported cases or estimates of the types of transmission events that gave rise to them (i.e., local transmission versus importations), that are crucially important to guide disease control efforts. Promoting phylodynamics in the field of animal health research should minimize the impact of epidemics by generating a greater understanding of pathogens evolution and transmission and control, in which human-animal interactions are embedded.

Despite the great achievements of phylodynamics to inform disease control, major challenges include missing and inaccurate data, as well as a lack of data sharing (see Outstanding Questions). This is due to various factors, such as the difficulty of collecting samples, costs of sequencing and proprietary interests [82]. Several collaborative efforts have been done to make sequences and associated epidemiological data publicly and timely available through web-based platforms, such as COVID-19 Data Portal, IRD [83] and GISAID [84,85], during recent high-profile outbreaks. However, there is currently no guidelines to ensure that data are shared as widely and as quickly as possible, emphasizing the fact that the global research community should agree on standard tools and practices that promote cooperation to address the barriers to data sharing. Moreover, pathogen evolutionary complexities, such as **recombination** or within-host diversity, pose significant challenges to the field [86,87]. One problem arising from recombination is that the ancestral relationships among genetic sequences have to be represented by a phylogenetic network instead of a tree. Due to within-host diversity, pathogen isolates sampled from a given host are not necessarily genetically identical to those transmitted by the host, obscuring the relationship between the

phylogenetic and transmission trees. However, phylodynamics is a growing field and novel phylodynamic models that are able to accommodate such complexities [88,89] have been developed to efficiently infer pathogen transmission dynamics between hosts.

Genetic data combined with epidemiological data offers the potential for unprecedented insight into infectious diseases spread between human and animal populations. However, this potential can be compromised when a poorly sampled population contributes to transmission, as strong biases in the data are inevitable. Experiences have shown that most emerging infectious disease outbreaks involve wildlife as maintenance hosts or reservoirs [90], such as badgers for bTB, wild migratory birds for HPAI and West Nile, wild boars for ASF, chimpanzees and gorillas for Ebola, wet market animals for SARS and COVID-19, making it sometimes challenging to generate data from these hard-to-sample populations. Detection and diagnosis of infectious diseases in wildlife compartments have been considerably hampered by lack of existing government wildlife surveillance, difficulties in deploying teams for field investigation, lack of sustainable funding and insufficient laboratory capacity [91,92]. Thus, prevention and control of infectious diseases in humans must involve an interdisciplinary and holistic approach that acknowledges the importance of wildlife surveillance.

Acknowledgements

The authors are grateful to Karin Bernodt (SVA, Sweden), Ariane Payne, Jean Hars (OFB, France), Gilles Balança (CIRAD, France), Chloe LeGall and Mattias Delpont (ENVT, France) for sharing the pictures for Figure 3. They also gratefully acknowledge help by Sarah Nadeau (ETH Zurich, Switzerland) and Daniel Grear (NWHC, USA) in producing an adapted version of a phylogenetic tree for Figure 4. They also thank the cEvo group (ETH Zurich, Switzerland) for providing useful comments on this project. This project has received funding

from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 842621.

Competing interests: The authors declare no competing interests.

References (80-100) 98 references

1 Taylor, L.H. *et al.* (2001) Risk factors for human disease emergence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 983–989

2 Jones, K.E. *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451, 990–993

3 Gebreyes, W.A. *et al.* (2014) The global one health paradigm: challenges and opportunities for tackling infectious diseases at the human, animal, and environment interface in low-resource settings. *PLoS Negl Trop Dis* 8, e3257

4 Fooks, A.R. *et al.* (2014) Current status of rabies and prospects for elimination. *The Lancet* 384, 1389–1399

5 Zinsstag, J. *et al.* (2017) Vaccination of dogs in an African city interrupts rabies transmission and reduces human exposure. *Sci. Transl. Med.* 9, eaaf6984

6 Gossner, C.M. *et al.* (2017) West Nile virus surveillance in Europe: moving towards an integrated animal-human-vector approach. *Eurosurveillance* 22, 30526

7 Anderson, R.M. and May, R.M. (1979) Population biology of infectious diseases: Part I. *Nature* 280, 361–367

8 Grenfell, B.T. *et al.* (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *science* 303, 327–332

9 Nie, Q. *et al.* (2020) Phylogenetic and phylodynamic analyses of SARS-CoV-2. *Virus Res.* 287, 198098

10 Benvenuto, D. *et al.* (2020) The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog. Glob. Health* 114, 64–67

11 Nadeau, S.A. *et al.* (2021) The origin and early spread of SARS-CoV-2 in Europe. *Proc. Natl. Acad. Sci.* 118,

12 Stadler, T. *et al.* (2014) Insights into the early epidemic spread of Ebola in Sierra Leone provided by viral sequence data. *PLoS Curr.* 6,

13 Alizon, S. *et al.* (2014) Quantifying the epidemic spread of Ebola virus (EBOV) in Sierra Leone using phylodynamics. *Virulence* 5, 825–827

14 Volz, E. and Pond, S. (2014) Phylodynamic analysis of Ebola virus in the 2014 Sierra Leone epidemic. *PLoS Curr.* 6,

15 Faria, N.R. *et al.* (2016) Zika virus in the Americas: early epidemiological and genetic findings. *Science* 352, 345–349

16 Boskova, V. *et al.* (2018) The influence of phylodynamic model specifications on parameter estimates of the Zika virus epidemic. *Virus Evol.* 4, vex044

17 Stadler, T. *et al.* (2013) Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl. Acad. Sci.* 110, 228–233

18 Stadler, T. *et al.* (2012) Estimating the basic reproductive number from viral sequence data. *Mol. Biol. Evol.* 29, 347–357

19 Kühnert, D. *et al.* (2016) Phylodynamics with migration: a computational framework to quantify population structure from genomic data. *Mol. Biol. Evol.* 33, 2102–2116

20 Stadler, T. and Bonhoeffer, S. (2013) Uncovering epidemiological dynamics in heterogeneous host populations using phylogenetic methods. *Philos. Trans. R. Soc. B Biol. Sci.* 368, 20120198

21 Kühnert, D. *et al.* (2014) Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth–death SIR model. *J. R. Soc. Interface* 11, 20131106

22 Rasmussen, D.A. *et al.* (2014) Reconciling phylodynamics with epidemiology: the case of dengue virus in southern Vietnam. *Mol. Biol. Evol.* 31, 258–271

23 Fournié, G. *et al.* (2013) Interventions for avian influenza A (H5N1) risk management in live bird market networks. *Proc. Natl. Acad. Sci.* 110, 9177–9182

24 Fournié, G. *et al.* (2011) Impact of the implementation of rest days in live bird markets on the dynamics of H5N1 highly pathogenic avian influenza. *J. R. Soc. Interface* 8, 1079–1089

25 Turner, J. *et al.* (2012) Modelling bluetongue virus transmission between farms using animal and vector movements. *Sci. Rep.* 2, 1–7

26 Kao, R.R. (2002) The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. *Trends Microbiol.* 10, 279–286

27 Thompson, D. *et al.* (2002) Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001. *Rev. Sci. Tech.-Off. Int. Epizoot.* 21, 675–687

28 Knight-Jones, T. and Rushton, J. (2013) The economic impacts of foot and mouth disease–What are they, how big are they and where do they occur? *Prev. Vet. Med.* 112, 161–173

29 Woolhouse, M. and Donaldson, A. (2001) Managing foot-and-mouth. *Nature* 410, 515–516

30 Haydon, D.T. *et al.* (2003) The construction and analysis of epidemic trees with reference to the 2001 UK foot–and–mouth outbreak. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 121–127

31 Ferguson, N.M. *et al.* (2001) Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature* 413, 542–548

32 Orsel, K. *et al.* (2007) The effect of vaccination on foot and mouth disease virus transmission among dairy cows. *Vaccine* 25, 327–335

33 Hu, B. *et al.* (2017) Bayesian inference of epidemiological parameters from transmission experiments. *Sci. Rep.* 7, 1–13

34 Streftaris, G. and Gibson, G.J. (2004) Bayesian analysis of experimental epidemics of foot–and–mouth disease. *Proc. R. Soc. Lond. B Biol. Sci.* 271, 1111–1117

35 Eblé, P. *et al.* (2006) Quantification of within-and between-pen transmission of foot-and-mouth disease virus in pigs. *Vet. Res.* 37, 647–654

36 Ferguson, N.M. *et al.* (2001) The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions. *Science* 292, 1155

37 Keeling, M.J. *et al.* (2001) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* 294, 813–817

38 Keeling, M.J. *et al.* (2003) Modelling vaccination strategies against foot-and-mouth disease. *Nature* 421, 136–142

39 Tildesley, M.J. *et al.* (2006) Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* 440, 83–86

40 Morris, R.S. *et al.* (2001) Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain, 2001. *Vet. Rec.* 149, 137–144

41 Whelan, S. *et al.* (2001) Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends Genet. TIG* 17, 262–272

42 Rowlands, R.J. *et al.* (2008) African swine fever virus isolate, Georgia, 2007. *Emerg. Infect. Dis.* 14, 1870–1874

43 Nderitu, L. *et al.* (2011) Sequential Rift Valley fever outbreaks in eastern Africa caused by multiple lineages of the virus. *J. Infect. Dis.* 203, 655–665

44 Peiris, J.S.M. *et al.* (2007) Avian influenza virus (H5N1): a threat to human health. *Clin. Microbiol. Rev.* 20, 243–267

45 Guan, Y. *et al.* (2002) Emergence of multiple genotypes of H5N1 avian influenza viruses in Hong Kong SAR. *Proc. Natl. Acad. Sci. U. S. A.* 99, 8950–8955

46 Pybus, O.G. and Rambaut, A. (2009) Evolutionary analysis of the dynamics of viral infectious disease. *Nat. Rev. Genet.* 10, 540–550

47 Drummond, A.J. *et al.* (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88

48 Drummond, A.J. *et al.* (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973
49 Nee, S. *et al.* (1994) The reconstructed evolutionary process. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 344, 305–311
50 Bardhan, D. *et al.* (2017) The economic impact of peste des petits ruminants in India. *Rev. Sci. Tech. Int. Off. Epizoot.* 36, 245–263
51 Padhi, A. and Ma, L. (2014) Genetic and epidemiological insights into the emergence of peste des petits ruminants virus (PPRV) across Asia and Africa. *Sci. Rep.* 4, 7040
52 Bao, J. *et al.* (2017) Evolutionary dynamics of recent peste des petits ruminants virus epidemic in China during 2013-2014. *Virology* 510, 156–164
53 Ignjatović, J. and Sapats, S. (2000) Avian infectious bronchitis virus. *Rev. Sci. Tech. Int. Off. Epizoot.* 19, 493–508
54 Franzo, G. *et al.* (2016) Effect of different vaccination strategies on IBV QX population dynamics and clinical outbreaks. *Vaccine* 34, 5670–5676
55 Alkhamis, M.A. *et al.* (2018) Phylodynamics and evolutionary epidemiology of African swine fever p72-CVR genes in Eurasia and Africa. *PloS One* 13, e0192565
56 Veo, C. *et al.* (2019) Evolutionary Dynamics of the Lineage 2 West Nile Virus That Caused the Largest European Epidemic: Italy 2011-2018. *Viruses* 11,
57 Global Consortium for H5N8 and Related Influenza Viruses (2016) Role for migratory wild birds in the global spread of avian influenza H5N8. *Science* 354, 213–217
58 Venkatesh, D. *et al.* (2020) Regional Transmission and Reassortment of 2.3.4.4b Highly Pathogenic Avian Influenza (HPAI) Viruses in Bulgarian Poultry 2017/18. *Viruses* 12,
59 Trovão, N.S. *et al.* (2015) Bayesian Inference Reveals Host-Specific Contributions to the Epidemic Expansion of Influenza A H5N1. *Mol. Biol. Evol.* 32, 3264–3275
60 Yang, J. *et al.* (2019) Inferring host roles in bayesian phylodynamics of global avian influenza A virus H9N2. *Virology* 538, 86–96
61 Omondi, G. *et al.* (2019) Phylogeographical and cross-species transmission dynamics of SAT1 and SAT2 foot-and-mouth disease virus in Eastern Africa. *Mol. Ecol.* 28, 2903–2916
62 Brito, B. *et al.* (2017) Phylodynamics of foot-and-mouth disease virus O/PanAsia in Vietnam 2010-2014. *Vet. Res.* 48, 24
63 Grear, D.A. *et al.* (2018) Inferring epidemiologic dynamics from viral evolution: 2014-2015 Eurasian/North American highly pathogenic avian influenza viruses exceed transmission threshold, R0 = 1, in wild birds and poultry in North America. *Evol. Appl.* 11, 547–557
64 Alkhamis, M.A. *et al.* (2016) Applications of Bayesian Phylodynamic Methods in a Recent U.S. Porcine Reproductive and Respiratory Syndrome Virus Outbreak. *Front. Microbiol.* 7, 67
65 Kamath, P.L. *et al.* (2016) Genomics reveals historic and contemporary transmission dynamics of a bacterial disease among wildlife and livestock. *Nat. Commun.* 7, 11448
66 Kühnert, D. *et al.* (2011) Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. *Infect. Genet. Evol.* 11, 1825–1841
67 Alkhamis, M.A. *et al.* (2015) Phylodynamics of H5N1 Highly Pathogenic Avian Influenza in Europe, 2005-2010: Potential for Molecular Surveillance of New Outbreaks. *Viruses* 7, 3310–3328
68 Arbi, M. *et al.* (2020) Historical origins and zoonotic potential of avian influenza virus H9N2 in Tunisia revealed by Bayesian analysis and molecular characterization. *Arch. Virol.* 165, 1527–1540
69 Xiang, B. *et al.* (2020) Insights into Genomic Epidemiology, Evolution, and Transmission Dynamics of Genotype VII of Class II Newcastle Disease Virus in China. *Pathogens* 9, 837
70 Chen, L. *et al.* (2020) Phylodynamic analyses of class I Newcastle disease virus isolated in China. *Transbound. Emerg. Dis.* n/a,
71 Jara, M. *et al.* (2020) Porcine reproductive and respiratory syndrome virus dissemination across pig production systems in the United States. *Transbound. Emerg. Dis.* n/a,
72 De Maio, N. *et al.* (2015) New Routes to Phylogeography: A Bayesian Structured Coalescent Approximation. *PLoS Genet.* 11, e1005421
73 Müller, N.F. *et al.* (2017) The Structured Coalescent and Its Approximations. *Mol. Biol. Evol.* 34, 2970–2981

74 Hicks, J.T. *et al.* (2020) Agricultural and geographic factors shaped the North American 2015 highly pathogenic avian influenza H5N2 outbreak. *PLoS Pathog.* 16, e1007857
75 Neumann, E.J. *et al.* (2005) Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. *J. Am. Vet. Med. Assoc.* 227, 385–392
76 Kao, R.R. *et al.* (2014) Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends Microbiol.* 22, 282–291
77 Biek, R. *et al.* (2015) Measurably evolving pathogens in the genomic era. *Trends Ecol. Evol.* 30, 306–313
78 Olea-Popelka, F. *et al.* (2017) Zoonotic tuberculosis in human beings caused by Mycobacterium bovis-a call for action. *Lancet Infect. Dis.* 17, e21–e25
79 Salvador, L.C.M. *et al.* (2019) Disease management at the wildlife-livestock interface: Using whole-genome sequencing to study the role of elk in Mycobacterium bovis transmission in Michigan, USA. *Mol. Ecol.* 28, 2192–2205
80 Crispell, J. *et al.* (2017) Using whole genome sequencing to investigate transmission in a multi-host system: bovine tuberculosis in New Zealand. *BMC Genomics* 18, 180
81 Crispell, J. *et al.* (2019) Combining genomics and epidemiology to analyse bi-directional transmission of Mycobacterium bovis in a multi-host system. *eLife* 8,
82 Yozwiak, N.L. *et al.* (2015) Data sharing: Make outbreak research open access. *Nature* 518, 477–479
83 Zhang, Y. *et al.* (2017) Influenza Research Database: An integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res.* 45, D466–D474
84 Bogner, P. *et al.* (2006) A global initiative on sharing avian flu data. *Nature* 442, 981–981
85 Shu, Y. and McCauley, J. (2017) GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill. Bull. Eur. Sur Mal. Transm. Eur. Commun. Dis. Bull.* 22,
86 Frost, S.D. *et al.* (2015) Eight challenges in phylodynamic inference. *Epidemics* 10, 88–92
87 Firestone, S.M. *et al.* (2019) Reconstructing foot-and-mouth disease outbreaks: a methods comparison of transmission network models. *Sci. Rep.* 9, 4809
88 Maio, N.D. *et al.* (2016) SCOTTI: Efficient Reconstruction of Transmission within Outbreaks with the Structured Coalescent. *PLOS Comput. Biol.* 12, e1005130
89 Müller, N.F. *et al.* (2020) Bayesian inference of reassortment networks reveals fitness benefits of reassortment in human influenza viruses. *Proc. Natl. Acad. Sci.* 117, 17104–17111
90 Bengis, R.G. *et al.* (2004) The role of wildlife in emerging and re-emerging zoonoses. *Rev. Sci. Tech. Int. Off. Epizoot.* 23, 497–511
91 Schwind, J.S. *et al.* (2014) Capacity building efforts and perceptions for wildlife surveillance to detect zoonotic pathogens: comparing stakeholder perspectives. *BMC Public Health* 14, 684
92 Kuiken, T. *et al.* (2005) Pathogen surveillance in animals. *Science* 309, 1680–1681
93 Volz, E.M. (2012) Complex population dynamics and the coalescent under neutrality. *Genetics* 190, 187–201
94 Drummond, A.J. *et al.* (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192
95 Frost, S.D.W. and Volz, E.M. (2010) Viral phylodynamics and the search for an "effective number of infections." *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365, 1879–1890
96 Stadler, T. (2010) Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404
97 Kendall, D.G. (1948) On the generalized" birth-and-death" process. *Ann. Math. Stat.* 19, 1–15
98 Vaughan, T.G. *et al.* (2014) Efficient Bayesian inference under the structured coalescent. *Bioinforma. Oxf. Engl.* 30, 2272–2279

Figures legends

Figure 1. **From field samples to epidemiological parameters.** Recent extensions of phylogenetics, coined as phylodynamics, mean that disease transmission parameters can be

now inferred based on genetic data. First, pathogen field samples are collected during an epidemic, from which genetic sequences are obtained. Then sequences are aligned based on their nucleotides to identify homologous genetic positions. Aligned genetic sequences are usually represented as rows in a matrix, with gaps (-) inserted so that homologous nucleotides are aligned in successive columns. A nucleotide substitution model that describes the nature of nucleotide substitutions is then used to calculate the likelihood of various possible phylogenetic trees, from which the most likely tree is identified. The molecular clock models describe the rate at which nucleotide substitutions occur over time. They allow to convert the branch lengths of a phylogenetic tree from units of nucleotide substitution to units of calendar time and thus to estimate the time of the MRCA (t_0) providing a lower bound estimate for the onset of the epidemic. A population dynamic model that describes the history of population dynamics based on genetic sequences sampled the population is then fitted to the reconstructed phylogenetic tree which allows to quantify epidemiological parameters such as the basic reproduction number (R_0).

Figure 2. Example of multi-type birth-death model to study avian influenza virus transmission between wild bird, domestic poultry and human populations. Under this model, transmission can occur within wild bird ($\beta_{\text{wild bird} \rightarrow \text{wild bird}}$), domestic poultry ($\beta_{\text{poultry} \rightarrow \text{poultry}}$) and human ($\beta_{\text{human} \rightarrow \text{human}}$) compartments. In addition, transmission is also possible between compartments: from domestic poultry to human ($\beta_{\text{poultry} \rightarrow \text{human}}$) and vice versa ($\beta_{\text{human} \rightarrow \text{poultry}}$), and from wild bird to domestic poultry ($\beta_{\text{wild bird} \rightarrow \text{poultry}}$) and vice versa ($\beta_{\text{poultry} \rightarrow \text{wild bird}}$). This model also assumes a becoming non-infectious rate for the human (δ_{human}), domestic poultry (δ_{poultry}) and wild bird ($\delta_{\text{wild bird}}$) compartments and a sampling probability with which infected humans (ρ_{human}), domestic poultry (ρ_{poultry}) and wild birds ($\rho_{\text{wild bird}}$) are sampled and sequenced.

Figure 3. Wildlife-livestock interface for the transmission of infectious diseases. A) European wild boar (*Sus scrofa*) populations are known to spread African swine fever to domestic pigs (*Sus scrofa domesticus*) by indirect contact (such as B) when drinking from water throughs) or C) by direct contact (D) European badgers (*Meles meles*) can become infected with bovine tuberculosis (bTB) and transmit the disease to cattle by indirect contact, such as when E) drinking from water thoughs or F) eating from cattle fee. G) Wild birds, such as greylag goose (*Anser anser*), can act as vectors of avian influenza and can infect domestic poultry (such as H,I) cattle egret (*Bubulcus ibis*) by coming in close proximity to poultry on free range farms. [Image credits: (A) Karin Bernodt, SVA (B) Ariane Payne, OFB (C) Jean Hars, OFB (D) BadgerHero Wikimedia Creative Commons (E,F) Ariane Payne, OFB (G) G.Balança CIRAD (H) C. LeGall ENVT (I) M. Delpont ENVT]

Figure 4. Time-calibrated maximum clade credibility (MCC) phylogenetic tree of H5N2 sequences from wild birds and poultry during the 2014–2015 outbreak in North America [63]**.** MCC tree showing ancestral hosts for HPI H5N2 recovered using the multi-type birth-death model. Branch tree colors represent the most probable host types of the isolate ancestors (domestic poultry or wild bird). Authors were able to infer cross-species transmission at a coarse scale: notably, H5N2 was likely introduced into poultry farms by transmission events from infected wild birds, and that once the virus entered the poultry production system, transmission was mainly driven by poultry production-related mechanisms illustrated by close phylogenetic distance among sequences from poultry populations, with a few transmission events back to wild birds.

Text Box - 411 words

Box 1. Epidemiological and phylodynamic models

In animal health, *R* estimates have been traditionally inferred by fitting mathematical models to epidemiological data collected during outbreak investigations. The susceptible-infected-recovered (SIR) model is one of the simplest and most frequently used mathematical models. It divides the population into three compartments, each containing individuals that are identical in terms of disease status, i.e. susceptible (S), infectious (I) and recovered (R), and is dynamic in that the individuals move from one compartment to another over time. A variety of other models are derivatives from it: the susceptible-infectious-susceptible (SIS) model assumes no immunity against re-infection, the susceptible-exposed-infectious-recovered (SEIR) model assumes a latent period before becoming infectious and the susceptible-infectious-recovered-susceptible (SIRS) models assumes a temporary immunity against re-infection but no latent period.

In public health, *R* estimates have been increasingly inferred by fitting population dynamic models to genetic data collected during outbreak investigations. **Coalescent models** [93,94] were the first applied to epidemics. They describe the rate at which a subset of pathogen lineages of a phylogenetic tree merge backward in time, until a common ancestor is reached. In that sense, the coalescent allows to estimate the **effective population size** over time (N_e), which describes the level of genetic diversity within a pathogen population over time and can be translated to the effective number of infections over the course of the epidemic [95]. Other population models were developed as alternatives to coalescent models, named as **Birth-Death (BD) models** [17,18]. They describe the rate at which pathogen lineages of a phylogenetic tree split and go extinct forward in time and can be translated to the rate at which infected hosts transmit and die/recover over the course of the epidemic. Therefore, in comparison to coalescent models, BD models allow for the direct estimation of

key epidemiological parameters, such as the birth/transmission (β) and death/recovery (δ) rates, and provide R (β/δ) estimates comparable to those inferred by mathematical models (Figure 2) [96,97]. They also allow to explicitly model the sampling process through a sampling (ρ) rate parameter, while coalescent models assume a small random sampling from a large population, which might not be realistic in the case of well-monitored epidemics.

More recently, these population dynamic models have been extended to account for structured populations, be they time-, species- or geographic-related, referred to as **structured population dynamic models**, such as the multi-type BD [19,20] or the structured coalescent [72,73,98] models, allowing to quantify transmission rates in structured populations and making the estimations more robust to heterogenous sampling.