

Phylodynamic Inference and Model Assessment with Approximate Bayesian Computation: Influenza as a Case Study

Oliver Ratmann^{1,2*}, Gé Donker³, Adam Meijer⁴, Christophe Fraser², Katia Koelle^{1,5}

1 Department of Biology, Duke University, Durham, North Carolina, United States of America, **2** Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **3** NIVEL, Netherlands Institute for Health Services Research, Utrecht, The Netherlands, **4** RIVM, National Institute for Public Health and the Environment, Centre for Infectious Disease Control, Bilthoven, The Netherlands, **5** Fogarty International Center, National Institutes of Health, Bethesda, Maryland, United States of America

Abstract

A key priority in infectious disease research is to understand the ecological and evolutionary drivers of viral diseases from data on disease incidence as well as viral genetic and antigenic variation. We propose using a simulation-based, Bayesian method known as Approximate Bayesian Computation (ABC) to fit and assess phylodynamic models that simulate pathogen evolution and ecology against summaries of these data. We illustrate the versatility of the method by analyzing two spatial models describing the phylodynamics of interpandemic human influenza virus subtype A(H3N2). The first model captures antigenic drift phenomenologically with continuously waning immunity, and the second epochal evolution model describes the replacement of major, relatively long-lived antigenic clusters. Combining features of long-term surveillance data from the Netherlands with features of influenza A (H3N2) hemagglutinin gene sequences sampled in northern Europe, key phylodynamic parameters can be estimated with ABC. Goodness-of-fit analyses reveal that the irregularity in interannual incidence and H3N2's ladder-like hemagglutinin phylogeny are quantitatively only reproduced under the epochal evolution model within a spatial context. However, the concomitant incidence dynamics result in a very large reproductive number and are not consistent with empirical estimates of H3N2's population level attack rate. These results demonstrate that the interactions between the evolutionary and ecological processes impose multiple quantitative constraints on the phylodynamic trajectories of influenza A(H3N2), so that sequence and surveillance data can be used synergistically. ABC, one of several data synthesis approaches, can easily interface a broad class of phylodynamic models with various types of data but requires careful calibration of the summaries and tolerance parameters.

Citation: Ratmann O, Donker G, Meijer A, Fraser C, Koelle K (2012) Phylodynamic Inference and Model Assessment with Approximate Bayesian Computation: Influenza as a Case Study. *PLoS Comput Biol* 8(12): e1002835. doi:10.1371/journal.pcbi.1002835

Editor: Sergei L. Kosakovsky Pond, University of California San Diego, United States of America

Received: March 19, 2012; **Accepted:** October 19, 2012; **Published:** December 27, 2012

Copyright: © 2012 Ratmann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: We gratefully accept financial support from the Wellcome Trust (<http://www.wellcome.ac.uk>), grant WR092311MF, the National Science Foundation (<http://www.nsf.gov>), grant NSF-EF-08-27416, and through the RAPIDD program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health (<http://www.fic.nih.gov>). 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.

* E-mail: oliver.ratmann@imperial.ac.uk

Introduction

Many infectious pathogens, most notably RNA viruses, evolve on the same time scale as their ecological dynamics [1]. One of the perhaps best documented examples are human influenza A viruses, which cause substantial morbidity and mortality as they escape host immunity predominantly through the evolution of their surface antigens [2]. The resulting, dynamical interaction between the ecological and evolutionary processes can be better understood through the formulation and simulation of so-called “phylodynamic” mathematical models, e.g. [3–8]. However, while data on disease incidence as well as viral genetic and antigenic variation are increasing for many viruses, e.g. [9–13], fitting and assessing phylodynamic models to these data is still not commonly done.

Historically, epidemiological time series data have been pervasively used to analyze hypotheses of host-pathogen interactions at the population level [14–17]. However, time series data capture the underlying evolutionary processes of pathogens only

very indirectly. For flu, this has limited the type of infectious disease models that can be statistically interfaced with time series data, and the number of epidemiological parameters that can be simultaneously estimated [18,19]. Consequently, the disease behavior of rapidly evolving pathogens is increasingly studied under additional, complementary data sets [1], most typically in ways that attempt to qualitatively reproduce prominent disease attributes [3–8].

More recently, coalescent-based statistical methods have been used to elucidate the disease dynamics of RNA viruses from molecular genetic data alone [20]. These methods have been particularly useful to reconstruct epidemiological transmission histories, identifying when and where transmission occurred and how viral populations change over time. For example, coalescent-based analyses have highlighted the importance of the tropics in the complex circulation dynamics of human influenza A (H3N2) virus (in short: H3N2) [9,21,22]. However, most coalescent methods estimate past population dynamics within a class of

Author Summary

The infectious disease dynamics of many viral pathogens like influenza, norovirus and coronavirus are inextricably tied to their evolution. This interaction between evolutionary and ecological processes complicates our ability to understand the infectious disease behavior of rapidly evolving pathogens. Most statistical methods for the analysis of these “phylo-dynamics” require that the likelihood of the data can be explicitly calculated. Currently, this is not possible for many phylo-dynamic models, so that questions on the interaction between viral variants cannot be well-addressed within this framework. Simulation-based statistical methods circumvent likelihood calculations. Considering inter-pandemic human influenza A virus subtype H3N2, we here illustrate the effectiveness of these methods to fit and assess complex phylo-dynamic models against both sequence and surveillance data. We find that combining molecular genetic and epidemiological data is key to estimate phylo-dynamic parameters reliably. Moreover, the information in the available data taken together is enough to expose quantitative model inconsistencies. Methods such as ABC which can combine sequence and surveillance data appear to be well-suited to fit and assess mechanistic hypotheses on the phylo-dynamics of RNA viruses.

flexible demographic functions including exponential and logistic growth as well as the nonparametric Bayesian skyride [23,24]; but see also [25]. These demographic functions do not explicitly describe the non-linear population dynamics of RNA viruses. Thus, assessing which ecological interactions underlie observed patterns of sequence diversity, and estimating the respective strength of these interactions, is difficult within this framework.

Because of these limitations, we adopt a different statistical approach known as Approximate Bayesian Computation (ABC) to infer the phylo-dynamics of RNA viruses. ABC allows mechanistic phylo-dynamic models to be simultaneously fitted against both sequence and surveillance data. This method circumvents explicit likelihood calculations by simulating instead from the stochastic model that defines the likelihood [26]. Recent extensions of ABC allow for model assessment to be carried out at no further computational cost [27]. We further suggest incorporating variable selection procedures to quantify if and to what extent the data provide support for the inclusion of specific model components [28].

To demonstrate the utility of our approach, we consider the phylo-dynamics of inter-pandemic H3N2. We obtained weekly reports of H3N2 incidence in the Netherlands from 1994–2009 by combining influenza-like-illness (ILI) surveillance data with detailed records of associated, laboratory-confirmed cases of flu by type and subtype [29,30], and similarly for France and the USA; see Figure 1 and the supplementary online material (Text S1). In addition, we reconstructed the ladder-like phylogeny of H3N2’s haemagglutinin gene (HA) from dated European sequences collected in 1968–2009 (see Figure 1 and Text S1). To represent H3N2’s global phylo-dynamics, we focus on a class of spatially structured phylo-dynamic compartmental models that formalize probabilistically how evolving, antigenic variants interact epidemiologically. These antigenic variants might correspond to the major antigenic clusters that are distinguishable in H3N2 antigenic maps [31], but can in principle also represent a different phenotypic resolution. The evolutionary dynamics of viral genotypes are separately formulated for each antigenic phenotype because genetic distances do not necessarily easily translate into

phenotypic relationships [5]. Spatial substructure has been incorporated in several models of H3N2 phylo-dynamics to reflect the global circulation of the virus [4,8,32]. We adopt here a simple source-sink framework, where the sink is thought of as the Netherlands into which viral genetic diversity and antigenic strains are imported on a seasonal scale from a source population where the virus persists [9,33]. We fit and assess two distinct models to the combined features of sequence and incidence data described in Figure 1 and Table 1. The first model captures H3N2’s antigenic drift phenomenologically through gradual loss of immunity, and the second model describes the antigenic evolution of the virus explicitly with particular assumptions on the tempo of antigenic change.

Methods

Approximate Bayesian Computation

To perform phylo-dynamic inference and goodness-of-fit analyses for complex phylo-dynamic models, we adopt a simulation-based approach that has become known as Approximate Bayesian Computation (ABC) [26]. Our first goal is to estimate the posterior density

$$\pi(\theta|x) \propto \ell(x|\theta)\pi(\theta) \quad (1)$$

of epidemiological and evolutionary model parameters $\theta \in \Theta$ under approximations to the likelihood $\ell(x|\theta)$ of observed population incidence and phylogenetic data x . The prior density $\pi(\theta)$ can be used to incorporate existing information or limit the range of plausible values of model parameters. Our second goal is to assess fitted phylo-dynamic models based on a recent extension of ABC [27].

ABC methods circumvent computations of the likelihood $\ell(x|\theta)$ by comparing the observed data x to simulated data y in terms of many, lower-dimensional summary statistics S_1, \dots, S_K such as those in Figure 1. Using a distance function ρ_k that compares summaries, each simulation y is weighted according to the magnitude of the summary error $\varepsilon_k = \rho_k(S_k(y), S_k(x))$ under a weighting scheme $w(\varepsilon_k; \tau_k)$, and this value is used in place of the likelihood term in Monte Carlo algorithms. In essence, ABC is a particular auxiliary variable Monte Carlo method, where the K summary errors take on the role of auxiliary variables. Integrating these errors out, the ABC likelihood approximation $\ell_\tau(x|\theta)$ adopted here is

$$\ell(x|\theta) \approx \ell_\tau(x|\theta) = \int \prod_k w(\rho_k(S_k(y), S_k(x)); \tau_k) \ell(dy|\theta), \quad (2)$$

where the weighting scheme is typically the Indicator

$$w(\varepsilon_k; \tau_k) = 1/(\tau_k^+ - \tau_k^-) \mathbb{1}\{\tau_k^- \leq \varepsilon_k \leq \tau_k^+\} \quad (3)$$

with tolerance parameter $\tau_k^- < 0 < \tau_k^+$ or the Exponential

$$w(\varepsilon_k; \tau_k) = \frac{1}{\tau_k} \exp\left(-\frac{|\varepsilon_k|}{\tau_k/2}\right) \quad (4)$$

with $\tau_k > 0$. Intuitively, the summary errors indicate how well a parameterized model reproduces the observed data. Once Monte Carlo algorithms such as the Markov Chain Monte Carlo (MCMC) sampler proposed by Marjoram et al. [34] have converged, the magnitude of the summary errors can be used to diagnose goodness-of-fit with respect to each of the summaries S_k .

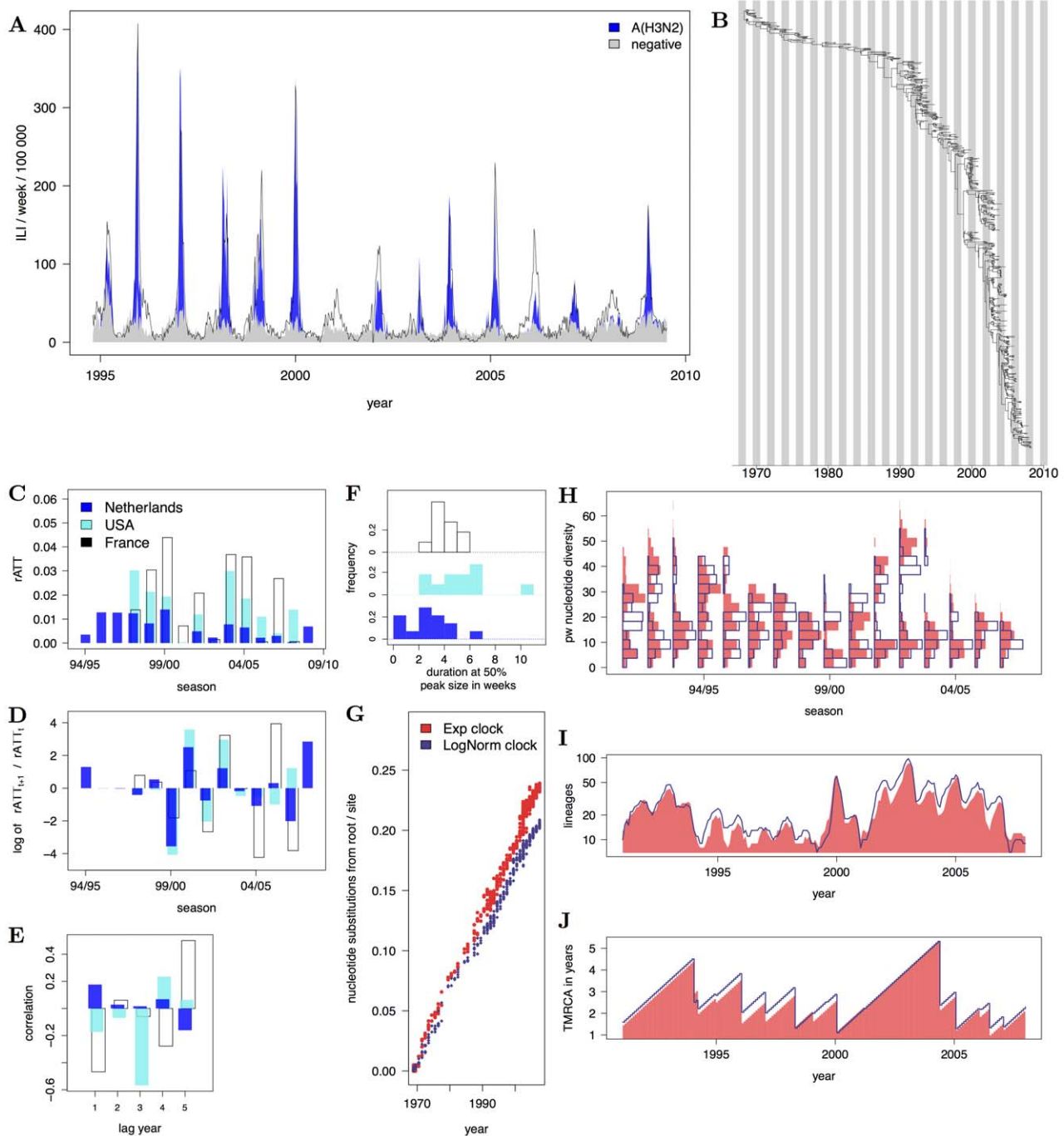


Figure 1. Features of H3N2 sequence and incidence data. (A) Weekly ILI time series from the Netherlands, and estimated time series of influenza A(H3N2) from weekly virological data. Type and subtype specific time series were estimated under an additive Negative Binomial regression model; see Text S1. (B) Reconstructed HA phylogeny from 776 European sequences with known times of isolation. The phylogeny was inferred with the BEAST program under a relaxed Exponential clock; see Text S1. (C) H3N2 seasonal attack rates (rATT), calculated from estimated H3N2 case report times series in the Netherlands in 1994–2009 (blue), and the USA (cyan) as well as France in 1997–2008 (black). (D) Ratio of consecutive case report attack rates on the log scale. (E) Autocorrelation of case report peaks. (F) Histogram of the duration of seasonal epidemics at half their peak size. (G) Number of estimated nucleotide substitutions of dated HA sequences from the root A/Bilthoven/16190/68 as in Smith et al. [31]. Nucleotide substitutions were estimated with BEAST under an Exponential clock (red) and Lognormal clock model (violet). (H) Histograms of pairwise nucleotide diversity among sequences collected in the same season. (I) Time series of the number of phylogenetic lineages circulating within the same month. (J) Time series of the time to the most recent common ancestor of phylogenetic lineages circulating within the same month. Colors from H to J are as in G.

doi:10.1371/journal.pcbi.1002835.g001

Table 1. Basic phylogenetic summaries of H3N2 surveillance data and phylogenies, and calibrated weighting schemes.

shorthand	summary	data	distance	summary values and distances*			weighting scheme†	
				Netherlands	France	USA	under the SEIRS model	under the epochal evolution model
μ -attack	average $rATT_t$, where $rATT_t$ is the total case report incidence in season t		log ratio	0.56%	1.9% (-1.26)	1.4% (-0.97)	Indicator (3) $\tau^- = -1.3$ $\tau^+ = 0.35$	Indicator (3) $\tau^- = -1.3$ $\tau^+ = 0.35$
σ -attack	standard deviation in $\log(rATT_{t+1}/rATT_t)$		log ratio	1.68	2.78 (-0.5)	2.24 (-0.28)	Indicator (3) $\tau^- = -0.7$ $\tau^+ = 0.35$	Indicator (3) $\tau^- = -0.7$ $\tau^+ = 0.35$
explosiveness	average duration of reported seasonal epidemics at half their peak size	time series 1994–2009	log ratio	3.2	4.54 (-0.32)	5.81 (-0.57)	Indicator (3) $\tau^- = -0.6$ $\tau^+ = 0.35$	Indicator (3) $\tau^- = -0.6$ $\tau^+ = 0.35$
correlation	Pearson autocorrelation of case report peaks at a lag of 2 & 4 years		largest difference	0.07 & 0	0.06 & -0.27 (-0.27)	-0.06 & 0.23 (0.23)	Exponential (4) $\tau = 15$	Indicator (3) $\tau^- = -0.6$ $\tau^+ = 0.35$
pop-attack	largest seasonal population-level attack rate	Ref. [2]	difference $S_k(x) - S_k(y)$		20%		Indicator (3) $\tau^- = -0.1$ $\tau^+ = 0.05$	Exponential (4) $\tau = 8$
divergence	average substitution rate per genome per year		phylogeny 1968–2009		6.1	5.3 (-0.13)	Indicator (3) $\tau^- = -0.4$ $\tau^+ = 0.4$	Indicator (3) $\tau^- = -0.2$ $\tau^+ = 0.2$
diversity	average pairwise diversity between any two sequences sampled in the same season		log ratio		15.4	15.13 (0.01)	Indicator (3) $\tau^- = -0.6$ $\tau^+ = 0.6$	Indicator (3) $\tau^- = -0.35$ $\tau^+ = 0.35$
lineages	average LIN_t , where LIN_t is the number of lineages in month t		log ratio		23.6	26.7 (-0.125)	Indicator (3) $\tau^- = -1.3$ $\tau^+ = 1.3$	Indicator (3) $\tau^- = -0.8$ $\tau^+ = 0.8$
TMRCAs	largest $TMRCAt$, where $TMRCAt$ is the time to the most recent common ancestor of all lineages in month t		log ratio		5.34	5.32 (≈ 0)	Exponential (4) $\tau = 6$	Indicator (3) $\tau^- = -0.4$ $\tau^+ = \infty$

*Distances between summaries derived from the first listed and subsequent data sets are given in brackets.

†Weighting schemes differ across models to accommodate weak or strong inconsistencies; see also Table 3.

‡The number of dated HA sequences available before 1990 is very small, so that these years effectively do not contribute to the diversity. To make this sampling effect more apparent, all phylogenetic summaries except the divergence are only computed on the period 1991–2009.

doi:10.1371/journal.pcbi.1002835.t001

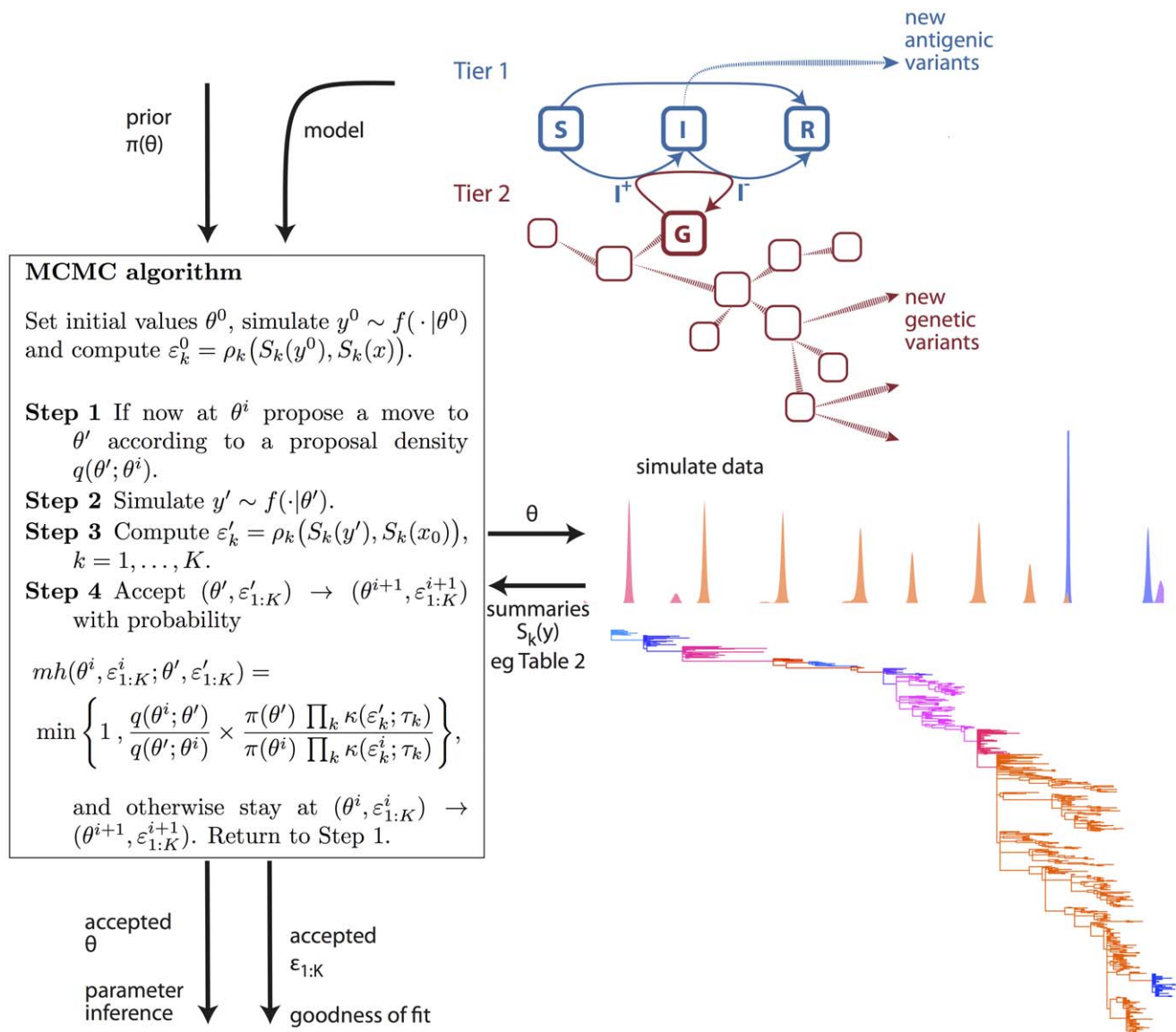


Figure 2. Overview of simulation-based phylodynamic inference and model assessment. Phylodynamic hypotheses are formulated into evolving, dynamical systems models. We used a two-tier model formulation whose genetic component is tied to its ecological component through the flows through the prevalence class. Existing knowledge on model parameters is incorporated through the prior $\pi(\theta)$, and Monte Carlo algorithms such as MCMC are used to fit the model to different types of data, e.g. incidence time series and reconstructed phylogenies (see Figure 1) with an ABC approach. ABC is based on likelihood approximations such as (2), which requires a specification of phylodynamic summaries (e.g. Table 1). The summary errors are used to diagnose if the fitted phylodynamic model is consistent with available data in terms of the specified summaries. doi:10.1371/journal.pcbi.1002835.g002

To use this detailed information on each summary, we prefer using (2) to the Mahalanobis approximation (see [26]). Although uncommon, we typically use the log ratio $\rho_k(S_k(x), S_k(y)) = \log(S_k(x)/S_k(y))$ so that the errors ε_k can be uniformly interpreted as fold-deviations. Parameter inference using ABC is approximate in that the ABC target density $\pi_\tau(\theta|x) \propto \ell_\tau(x|\theta) \pi(\theta)$ approaches the posterior density (1) as τ tends to zero if the summaries are sufficient for θ [26]. We use a Monte Carlo algorithm that is very similar to the MCMC sampler in Figure 2. A full specification of the algorithm is given in Text S1.

It is typically difficult to establish the sufficiency of phylodynamic summaries analytically, and instead a small set of summaries is chosen such that model parameters of interest can be estimated [26]. Table 1 lists basic features of H3N2 epidemiological and phylogenetic data

that were primarily considered in this study. Phylodynamic models were fitted and assessed against the features of the Dutch incidence data and the viral phylogeny derived under the Exponential clock model. The differences between these summaries and those derived from the remaining data in Figure 1 were used to set the ABC tolerances large enough so that inference is robust to the choice of phylogenetic reconstruction method and reporting country. Although smaller tolerances can be computationally feasible, these were not supported by the additional data considered. We typically use the Indicator weighting scheme (3) with tolerances τ_k that encompass differences in summary values across reporting countries and/or reconstruction methods, see Table 1. When a model never fits a particular summary well, we use (4) to give a mild prior preference to small errors [27]. See Text S1 for further details.

Spatial two-tier models to represent H3N2 phylodynamics

Deterministic skeleton. ABC methods require that each phylodynamic simulation must run on the order of tens of seconds. To meet this computational requirement while still allowing for flexible modeling [6,35], we adopt a two-tier approach that separates the genotypes of rapidly evolving viruses from their antigenic phenotypes [7]. The underlying rationale is that differences in genotype are only relevant from a population dynamic perspective if they translate into perceivable phenotypic differences. The first tier describes the dynamic interactions of antigenic variants in the host population, here in terms of coupled susceptible-exposed-infected-recovered-susceptible (SEIRS) equations that are further spatially structured into a strongly seasonally forced sink population and a re-seeding, weakly seasonally forced source population (denoted by \downarrow and \circ respectively). The second tier simulates a phylogeny that is consistent with the prevalence and incidence dynamics of each antigenic unit in the first tier. Assuming polarized immunity [3], the deterministic skeleton for the i th antigenic unit is

$$\begin{aligned}\frac{dS_i^\downarrow}{dt} &= \mu^\downarrow(N^\downarrow - S_i^\downarrow) - \beta_i^\downarrow \frac{S_i^\downarrow}{N^\downarrow} \sum_{j=1}^n \sigma_{ij}(I_j^\downarrow + M_j^\downarrow) + \gamma R_i^\downarrow \\ \frac{dE_i^\downarrow}{dt} &= \beta_i^\downarrow \frac{S_i^\downarrow}{N^\downarrow} (I_i^\downarrow + M_i^\downarrow) - (\mu^\downarrow + \phi) E_i^\downarrow \\ \frac{dI_{1i}^\downarrow}{dt} &= \phi E_i^\downarrow - (\mu^\downarrow + 2\nu) I_{1i}^\downarrow - h(t - t_i^e) I_{1i}^\downarrow \\ \frac{dI_{2i}^\downarrow}{dt} &= 2\nu I_{1i}^\downarrow - (\mu^\downarrow + 2\nu) I_{2i}^\downarrow - h(t - t_i^e) I_{2i}^\downarrow \\ \frac{dR_i^\downarrow}{dt} &= 2\nu I_{2i}^\downarrow - (\mu^\downarrow + \gamma) R_i^\downarrow\end{aligned}\quad (5a)$$

$$\begin{aligned}\frac{dS_i^\circ}{dt} &= \mu^\circ(N^\circ - S_i^\circ) - \beta_i^\circ \frac{S_i^\circ}{N^\circ} \left(\sum_{j=1}^n \sigma_{ij} I_j^\circ + M^\circ \right) + \gamma R_i^\circ \\ \frac{dE_i^\circ}{dt} &= \beta_i^\circ \frac{S_i^\circ}{N^\circ} (I_i^\circ + M^\circ) - (\mu^\circ + \phi) E_i^\circ \\ \frac{dI_{1i}^\circ}{dt} &= \phi E_i^\circ - (\mu^\circ + 2\nu) I_{1i}^\circ - h(t - t_i^e) I_{1i}^\circ \\ \frac{dI_{2i}^\circ}{dt} &= 2\nu I_{1i}^\circ - (\mu^\circ + 2\nu) I_{2i}^\circ - h(t - t_i^e) I_{2i}^\circ \\ \frac{dR_i^\circ}{dt} &= 2\nu I_{2i}^\circ - (\mu^\circ + \gamma) R_i^\circ\end{aligned}\quad (5b)$$

$$\frac{dG_{ik}^\downarrow}{dt} = \frac{(1 + s_{\rho_{k0}}) I_i^{+\downarrow}}{\sum_l (1 + s_{\rho_{l0}}) G_{il}^\downarrow} G_{ik}^\downarrow - \frac{I_i^{-\downarrow}}{\sum_l G_{il}^\downarrow} G_{ik}^\downarrow - \zeta G_{ik}^\downarrow \quad (5c)$$

$$\frac{dG_{ik}^\circ}{dt} = \frac{(1 + s_{\rho_{k0}}) I_i^{+\circ}}{\sum_l (1 + s_{\rho_{l0}}) G_{il}^\circ} G_{ik}^\circ - \frac{I_i^{-\circ}}{\sum_l G_{il}^\circ} G_{ik}^\circ - \zeta G_{ik}^\circ, \quad (5d)$$

where all model parameters are described in Table 2 or below. Two infectious subcompartments I_{1i} , I_{2i} are employed to obtain more realistic infectiousness profiles [36]. G_{ik} is the number of individuals infected with the k th genotype of the i th antigenic unit, $I_i^\downarrow = I_{1i}^\downarrow + I_{2i}^\downarrow$, $I_i^\circ = I_{1i}^\circ + I_{2i}^\circ$ for convenience, and $N^\downarrow = S_i^\downarrow + E_i^\downarrow + I_i^\downarrow + R_i^\downarrow$ and $N^\circ = S_i^\circ + E_i^\circ + I_i^\circ + R_i^\circ$ for all i .

In the first tier (5a–5b), competition between two antigenic variants i, j arises through resource depletion via partial cross-immunity σ_{ij} that decays multiplicatively with kinship level k_{ij} , $\sigma_{ij} = \sigma_{i-1,i}^{k_{ij}}$, where $0 \leq \sigma_{i-1,i} \leq 1$ is the degree of cross-immunity between mother-daughter variants [3]. The emergence of antigenic variants is described phenomenologically with a per capita hazard function h after the emergence time t_i^e of the resident phenotype i [7]. The hazard function is parameterized with a scale parameter λ and a shape parameter κ . The strength of sinusoidal seasonal forcing in the source population in the transmission parameter, $\beta_i^\circ = \beta(1 + \varphi^\circ \sin(2\pi(t - t_i^\circ)))$, is assumed to be much smaller than in the asynchronously forced sink population, $\varphi^\circ \ll \varphi^\downarrow$, and t^\downarrow is set so that transmission peaks at the winter solstice in the Northern hemisphere. M_j^\downarrow is the number of infected visiting travelers from the source, while M° is the number of individuals that re-seed the source population. Thus, the source population can be interpreted as an interconnected, re-seeding tropical region whose population size N° is to be estimated. We further calibrate the sink population to represent the Netherlands, using demographic data to specify N^\downarrow and μ^\downarrow over the study period 1968–2009. To fit model (5), we transform β into R_0 at disease equilibrium of a single variant [14], and define M_i^\downarrow , M° by $M_i^\downarrow = m^\downarrow I_i^\circ / N^\circ$ and $M^\circ = m^\circ \hat{I}^\circ$ where \hat{I}° is the number of infected individuals at disease equilibrium of a single variant.

In the second tier (5c–5d), the instantaneous loss in G_{ik} is proportional to genotype frequency, while the gain in G_{ik} is weighted by the fitness advantage of each genotype. As before [7], fitness is assumed to increase linearly with the number ρ_{k0} of nucleotide mutations between the k th genotype and the founder genotype of the i th antigenic variant. The total number of infections and losses I_i^+ and I_i^- are the simulated transitions in and out of I_i at time t , so that (5c–5d) are tied to (5a–5b). New genotypes evolve at a rate ζ , and a genealogy of the i th antigenic unit is generated by recording the emergence times of each genotype along with their kinships. The branch length between offspring and parental genotype is always one. After extinct genotypes are pruned, the branch length between two genotypes gives the number of nucleotide substitutions between them. These genealogies are concatenated by connecting the root genotype with a genotype of the parental antigenic unit that is randomly drawn according to genotype frequencies at time t_i^e . The residual selection parameter s accounts phenomenologically for selection pressures between genetic variants that are evident from the shape of the virus phylogeny, but remain unexplained by a particular ecological model of antigenic variants (5a–5b). Once the distribution of s has been inferred from population incidence and genetic data, we can then quantify how well a phylodynamic model describes patterns of continual immune selection mechanistically, and also compare alternative phylodynamic models in this respect.

Stochastic process model. To account for demographic stochasticity, Markov transition probabilities are derived from (5), assuming that the per capita rates are constant over a small time interval Δt , and that transitions out of any state are independent and multinomially distributed. Generally, consider a state X and

Table 2. Phylogenetic model parameters, prior and estimated densities.

symbol	description	prior density	mean \pm std. dev., 95% conf. interval of	
			posterior density under the SEIRS model	posterior density under the epochal evolution model
R_0	Basic reproductive number	uninformative	3.03 ± 0.55 , [1.77, 4.14]	18.7 ± 5.3 , [9.2, 26.8]
R_{eff}	effective reproductive number	-	1.26 ± 0.05 , [1.17, 1.35]	1.42 ± 0.12 , [1.27, 1.51]
$1/\phi$	Average incubation period in days	0.9		
$1/\nu$	Average infectiousness period in days	1.8		
$1/\gamma$	Average duration of immunity in years	uninformative	9.8 ± 1.8 , [6.5, 12.2]	206 ± 103 , [46, 380]
ρ	Reporting rate	uninformative	0.15 ± 0.06 , [0.06, 0.26]	0.56 ± 0.23 , [0.25, 0.95]
s	Residual selection	Exponential slab with mean 0.007 & Gaussian pseudo-prior centered at 0.09 [28]	0.1 ± 0.16 , [0.01, 0.44]	0.04 ± 0.07 , [0.001, 0.12]
i_s	Inclusion probability of s	uninformative	1 ± 0 , [1,1]	1 ± 0 , [1,1]
ζ	Mutation rate, genome ⁻¹ year ⁻¹	uninformative	1.32 ± 0.3 , [1.0, 1.9]	3.38 ± 1.2 , [1.8, 5.4]
N^{\downarrow}	Size of sink population	fixed to Dutch demographic data, http://statline.cbs.nl		
N°	Size of source population	uninformative	1.28 ± 0.95 , $[0.43, 3.6] \times 10^8$	2.9 ± 1.6 , $[0.7, 5.7] \times 10^8$
μ^{\downarrow}	Birth/death rate in the sink population	fixed to Dutch demographic data		
μ°	Birth/death rate in the source population, years ⁻¹	1/50; average lifespan of 60 years adjusted by net fertility rate in South East Asia		
ϕ^{\downarrow}	Seasonal forcing in the sink population	$\mathcal{U}(0.07, 0.6)$; see Text S1	0.42 ± 0.14 , [0.3, 0.6]	0.35 ± 0.15 , [0.12, 0.58]
ϕ°	Seasonal forcing in the source population	$\mathcal{U}(0, 0.02)$; key assumption, see Text S1	0.01 ± 0.007 , [0.002, 0.02]	0.013 ± 0.006 , [0.008, 0.02]
m^{\downarrow}	Number of travelers visiting the sink population	$\mathcal{U}(3 \times 10^6, 15 \times 10^6)$; encompassing lowest & highest annual records; http://statline.cbs.nl	8.5 ± 2.8 , $[3.6, 14.1] \times 10^6$	9.9 ± 3.4 , $[3.8, 14.6] \times 10^6$
m°	Fraction of \hat{I}° re-seeding the source population	$\mathcal{U}(0, 0.1)$	0.06 ± 0.03 , [0.01, 0.09]	0.06 ± 0.03 , [0.02, 0.09]
$\sigma_{i-1,i}$	Partial cross-immunity of mother-daughter variants	uninformative	-	0.76 ± 0.05 , [0.67, 0.85]
λ	Scale parameter of the antigenic emergence rate	uninformative	-	386 ± 97 , [247, 533]
κ	Shape parameter of the antigenic emergence rate	2; Ref. [7]		

doi:10.1371/journal.pcbi.1002835.t002

all $n(X)$ per capita rates r_{xy} out of X . The state transitions out of X into states $Y_1, \dots, Y_{n(X)}$ are

$$\Delta N_{xy_1}, \dots, \Delta N_{xy_{n(X)}} \sim \text{Multinomial}(Y; \frac{r_{xy_1}}{\sum_y r_{xy}}, \dots, \frac{r_{xy_{n(X)}}}{\sum_y r_{xy}}),$$

where $Y \sim \text{Binomial}(X, 1 - \exp(-\sum_y r_{xy} \Delta t))$ is the total number of individuals leaving X at time t .

For the application to H3N2 phylogenetics, simulations were started in $t_0=1968$ at the disease equilibrium of a single antigenic variant and generated under a multinomial Euler scheme with $\Delta t=1/8$ days. After the simulations of the first tier completed, the corresponding phylogeny was simulated based on the flows in and out of the prevalence compartments [7]. Simulated data were recorded after $t=1990$ to match the time range of the observed summaries in Figure 1. We do not estimate the initial conditions of the state variables and assume that by 1990, the phylogenetic processes do not depend any longer on the initial values in 1968.

Stochastic observation model. To interface the two-tier model with observed case report data and phylogenies, we simulated reported incidence under a Poisson model with mean $\rho \sum_i I_i^+$ and drew a requested number of genotypes at specified sampling times without replacement according to the genotype frequencies at those times. Replacing the genotype emergence times with the corresponding sampling times and pruning non-sampled genotypes, we obtained a dated phylogeny with branch lengths encoding nucleotide substitution distances.

Inferring inclusion probabilities of model parameters

A frequent problem in phylogenetic modeling is to determine if a specific model parameter should be included. For example, it can be unclear which types of ecological interactions between antigenic variants underlie pathogen phylogenetics, or if the residual selection parameter s in (5c–5d) is required in addition to a given ecological mechanism that induces immune selection. Following existing variable selection procedures [28], we use an additional indicator variable i_θ to denote whether a single model

parameter θ is present ($t_\theta = 1$) or absent ($t_\theta = 0$) and estimate its posterior probability under equation (2). Here, we use a standard spike-and-slab variable selection procedure [28] to estimate inclusion probabilities of the residual selection parameter s .

Results

Basic geographic framework for modeling H3N2 phylodynamics

To illustrate ABC methodology with the summaries in Table 1, we begin with a classical phenomenological model that implicitly accounts for antigenic drift through gradual loss of immunity [37]. H3N2 phylodynamics are represented with a spatial two-tier system of equations that is a special case of (5) when the antigenic emergence rate is set to

$$h(t - t_i^e) = 0. \quad (6)$$

For simplicity, we will refer to (5) without antigenic variants as the SEIRS model.

Simulated data. We first tested ABC on simulated data generated under the SEIRS model and found that the subset R_0 , $1/\gamma$, ρ , ι_s , s , ζ of model parameters can be reliably estimated with ABC tolerances τ_k^{sim} that are smaller than those in Table 1 (see Text S1). Tighter tolerances on the population level attack rate contributed most to more reliable estimates of R_0 .

Parameter inference. The behavior of the spatial SEIRS model, when fitted to the case report and phylogenetic summaries in Table 1, is illustrated in Figure 3 with parameter estimates given in Table 2. On real data, the summary errors were considerably larger than on simulated data, so that the τ_k^{sim} could not be used. Instead, we chose ABC tolerances with a data-driven approach that compares summary errors across different empirical data sets (see the Methods section and Table 1). Overall, we can simultaneously infer the epidemiological and evolutionary parameters R_0 , $1/\gamma$, ρ , ι_s , s , ζ , N° . As shown in Figure 3A, the MCMC algorithm may get occasionally stuck in the tails of the target density (see Text S1 for further discussion). The posterior mean and standard deviation of R_0 , 3.03 ± 0.55 , are relatively large in comparison to estimates from previous studies [36,38–40], and R_0 is positively correlated with the average duration of infection $1/\gamma$ to yield realistic incidence time series. We back-calculated the effective reproductive number R_{eff} from the prevalence growth rate at the beginning of each season (see Text S1), and find that many combinations of R_0 and $1/\gamma$ give a tight mean posterior $R_{\text{eff}} = 1.26 \pm 0.05$ in agreement with these studies. In the absence of any ecological mechanisms inducing strain competition, the mean residual selection parameter is large $s = 0.1 \pm 0.16$ and always included in the SEIRS model. Generally, the sequence divergence imposes negative correlations between ζ , s (Figure 3E) and ζ , N° (Figure 3F), and the posterior mean mutation rate $\zeta = 1.32 \pm 0.3/\text{genome}/\text{year}$ is much smaller than H3N2's substitution rate, $5.3\text{--}6.1/\text{genome}/\text{year}$, when selection is incorporated into the model. Figure 4 illustrates that the fraction of susceptible individuals ranges within 15–40% and changes smoothly under seasonal forcing, thus leading to sustained oscillations in disease incidence. We failed to estimate m^\dagger , φ^\dagger , φ° , m° and recovered distributions close to the prior. Our prior assumptions are summarised in Table 2 and more fully discussed in Text S1.

Sensitivity of parameter estimates to phylodynamic summaries. The extent to which phylodynamic parameters can be estimated depends mainly on the type of information that underlies the ABC summaries. As described more fully in Text S1,

a broad range of epidemiological parameters are quantitatively consistent with summaries of the H3N2 case report data in Figure 1A because variable reporting rates can mask the true extent of population incidence when immunity is not permanent [19]. Detailed studies of closely monitored populations and serological data suggest interpandemic seasonal H3N2 attack rates between 10–20% in temperate regions [2], and we found that conditioning on a broad window of maximum seasonal population incidence attack rates (*pop-attack*) between 15–30% ensures that key epidemiological parameters can be well estimated (Figure S4 in Text S1).

Moreover, while the sequence divergence and diversity are standard descriptors of viral phylogenies [1], we found that they are not sufficient to infer the size of the source population N° when the mutation rate ζ and the residual selection parameter s are simultaneously estimated. Considering the narrowness of the phylogeny in terms of the number of circulating lineages, we could estimate the source population size (Figures S5–6 in Text S1). We can use the number of lineages despite their dependence on sampling effort because with ABC, we are free to sample simulated sequences exactly as in the observed data set, see the Methods section. Finally, the time to the most recent common ancestor (TMRCA) links the evolutionary dynamics with the ecological interactions between antigenically distinct viral variants because weak selective advantages invariably lead to coexistence and deep phylogenies in the face of high ζ and weak φ . In the absence of sufficiently strong ecological interactions, the TMRCA's favor a larger residual selection parameter (Figure S7 in Text S1).

Goodness of fit. The summary errors reveal that the SEIRS model fails to reproduce the irregular interannual variability in winter season epidemics, and the narrowness and limited diversity of the HA phylogeny despite large s (Figure 3G–I). However, the model can reproduce H3N2's high divergence rate. This is not the case for the SEIRS model without a separate, weakly seasonally forced source population (see Text S1).

Epochal evolution model of H3N2 phylodynamics

While several models have been able to simulate phylodynamics that are consistent with some aspects of the observed data, most notably the ladder-like phylogeny of H3N2's haemagglutinin gene [4,5,41], none have been quantitatively fitted and tested against a set of epidemiological and molecular genetic features such as those in Figure 1. Here, we focus on the epochal evolution model as formulated in [7] within the above spatial framework, which is identical to (5) when antigenic variants are interpreted as major antigenic clusters. To fit (5) to the serial replacement of 11 major antigenic clusters within 1968–2002 [31], we define an antigenic cluster as any antigenic unit that survives for at least 2 years and use the summaries in Table 1 as well as the number of antigenic clusters generated in 1968–2002 (*nclust*). Following [7], the emergence rate is set to increase with age,

$$h(t - t_i^e) = \frac{\kappa}{\lambda} \left(\frac{t - t_i^e}{\lambda} \right)^{\kappa-1}, \quad (7)$$

$\kappa = 2$, and the scaling parameter λ is estimated. For simplicity, we refer to (5) with this antigenic emergence rate and an antigenic resolution that is determined by *nclust* as the epochal evolution model.

Simulated data. We generated data under the epochal evolution model and fitted both models with the summaries in Table 1. The summary errors deviated from zero only when the SEIRS model was fitted, indicating that ABC can correctly and readily identify model mismatch (see Text S1).

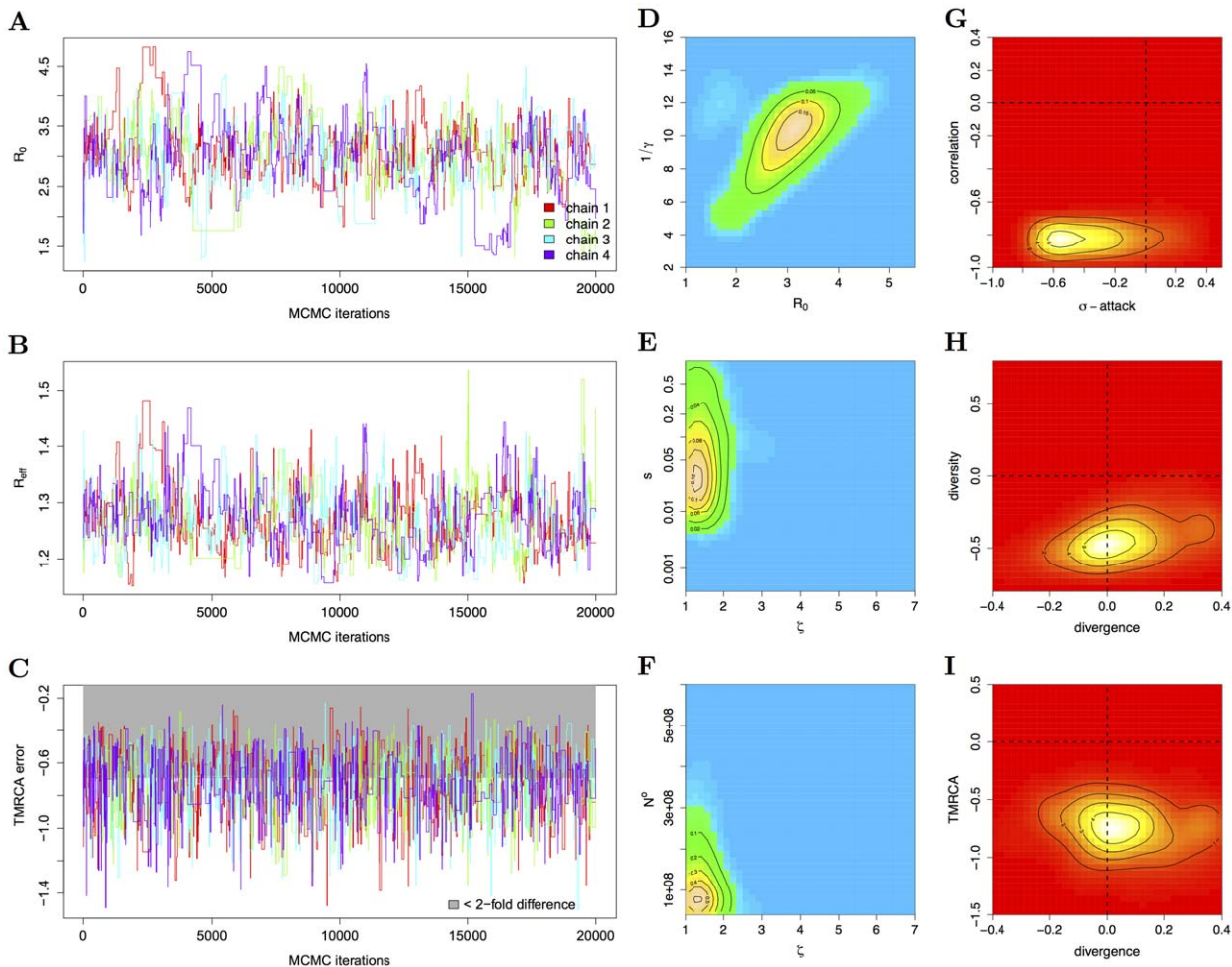


Figure 3. Phylodynamic inference and goodness-of-fit analysis of the spatial SEIRS model. (A–C) MCMC trajectories of the estimated R_0 , the calculated R_{eff} , and the TMRCA summary error of four chains that were started at overdispersed starting values (see Methods). Samples before iteration 1000 were discarded. (D–F) Two-dimensional histograms of parts of the ABC fit, illustrating the correlations between the estimated parameter pairs (R_0 , $1/\gamma$), (s , ζ) and (N° , ζ). Throughout, histograms were computed from all samples across the four chains after burn-in. Color codings are separate for each subplot, with respective density values indicated in the contours. (G–I) Two-dimensional histograms of parts of the joint density of summary errors, illustrating goodness-of-fit with respect to the *correlation* and interannual variability of the case report data, as well as the *divergence*, *diversity* and the *TMRCA*'s of the HA phylogeny. doi:10.1371/journal.pcbi.1002835.g003

Parameter inference. We could fit and assess the epochal evolution model against summaries of H3N2 surveillance and sequence data with ABC (Figure 5 and Tables 2–3). Partial cross-immunity between mother-daughter antigenic clusters is relatively weak ($\sigma_{i-1,i} = 0.76 \pm 0.05$), leading to abrupt changes in herd immunity and weak competition between clusters because few susceptibles are cross-depleted (see also Figure 6). In contrast to the SEIRS model, the fitted epochal evolution model excites irregular viral dynamics and reproduces the limited diversity and the small TMRCA's of H3N2's HA phylogeny (Figures 3G–I versus Figures 5I–K). This enabled us to use tighter weighting schemes for several summaries under the epochal evolution model (see Table 1). The choice of summaries influences parameter estimates in a similar manner as for the SEIRS model, see Text S1.

Goodness of fit. In the absence of strong seasonal forcing in the source population, infrequent cluster invasions often excite large invasion waves and refractory oscillations [32], as well as pronounced genetic bottlenecks that are inconsistent with the HA phylogeny, see Figure 6. More importantly, when the epochal evolution model is fitted to H3N2's narrow phylogeny, ABC

reveals that aspects of disease incidence cannot be quantitatively reproduced at the same time. In particular, the estimated average duration of intra-cluster immunity is $1/\gamma = 205 \pm 100$ years, which in turn implies a mean posterior R_0 of 19 ± 5 in order to reproduce H3N2's explosiveness; see Figures 5G,H. The effect of such high values of R_0 is hard to discern on interpandemic case report data without much stronger assumptions on the reporting rate than in this study. However, an R_0 around 20 implies long term population level attack rates well below 10%, which is not compatible with epidemiological estimates (see Figure 5L and [2]). To match aspects of H3N2's HA phylogeny, unrealistically low *pop-attack* rates are further compensated by high $N^\circ = 2.9 \times 10^8 \pm 1.6 \times 10^8$. If strong seasonal forcing is assumed in the source population ($\varphi^\circ > 0.15$ but see [42]), the epochal evolution model produces a much better fit in line with previous work [7] (see Text S1).

Variable selection. Finally, we identify significant levels of unexplained selection pressures in the HA phylogeny under the epochal evolution model. While the mean posterior residual

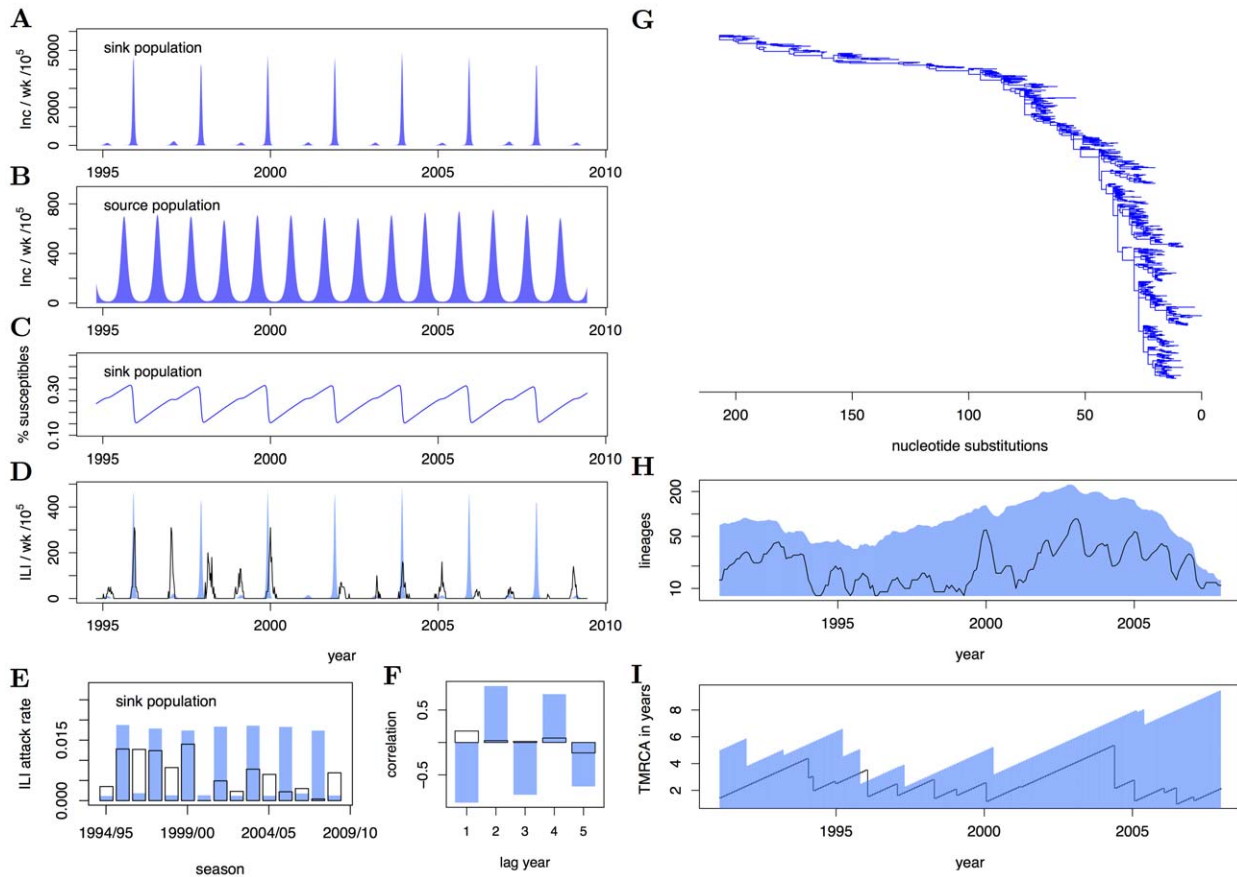


Figure 4. Phylodynamics arising under the spatial SEIRS model. (A–B) Population-level weekly incidence in the sink and source population, respectively. (C) Corresponding weekly time series of the percentage of susceptible individuals in the sink population. (D) Simulated H3N2 weekly surveillance time series in the sink population (blue) and reconstructed H3N2 time series in the Netherlands (black). (E) Simulated and observed case report seasonal attack rates, and (F) autocorrelation function of case report peaks. Typically, simulations under the fitted model show sustained oscillations that follow a clear biennial pattern. (G) Simulated HA phylogeny under a large, estimated residual selection parameter. (H) Simulated and observed lineage profile, and (I) simulated and observed time series of the time to the most recent common ancestor of extant phylogenetic lineages. Despite a relatively high selection parameter, the number of lineages and the time to the most recent common ancestor are overall too high when compared to data. Model parameters are $R_0 = 3$, $1/\phi = 0.9$, $1/\nu = 1.8$, $1/\gamma = 8.5$, $\rho = 0.11$, $N^c = 1.2 \times 10^8$, $1/\mu^c = 50$, $\phi^l = 0.4$, $\phi^c = 0.01$, $m^l = 7.7 \times 10^6$, $m^c = 0.04$, $\zeta = 1.2$, $s = 0.04$.
doi:10.1371/journal.pcbi.1002835.g004

selection parameter s is smaller than under the SEIRS model, the posterior probability of including s is still 1. Typically, small s are confounded with the corresponding inclusion probability because $i_s = 0$ or $i_s = 1$ and small s are almost equally likely [28]. Under both models, the inclusion probability is unambiguously estimated, indicating that the estimated residual selection parameter is too large to be ignored and that selection occurs not only between antigenic clusters but also within them.

Discussion

ABC for phylodynamic inference and model assessment

Fitting mechanistic models to infectious disease dynamics of RNA viruses that may escape immunity is notoriously difficult, and key epidemiological parameters such as R_0 can be estimated only under tacit assumptions from incidence time series [18,19,43]. Currently, alternative statistical synthesis approaches are explored to harness the information in complementary data sources [25,44,45]. Considering summaries of interpandemic H3N2 sequence and surveillance data, we show here that ABC can be used to fit and assess complex phylodynamic models which describe how evolutionary and ecological processes of the

influenza virus may interact. Key phylodynamic parameters could be estimated under relatively weak assumptions (Table 2), and ABC diagnosed readily if and in which direction the two considered models deviate from all the available data taken together.

Phylodynamic parameter inference and goodness-of-fit analyses rely critically on the possibility to combine epidemiological and molecular genetic data. In particular, H3N2 case report data were not sufficient to disentangle the reporting rate from epidemiological parameters, and measures of sequence divergence and diversity were not sufficient to separate the population size from evolutionary parameters. To the extent that other RNA viruses are characterized by different phylodynamic behavior, different sets of summaries must be identified in each case to replace likelihood calculations.

ABC relates evolutionary and epidemiological data mechanistically through an evolving dynamic system and thereby allows us to investigate empirical phylodynamic hypotheses more directly than is possible with other statistical data synthesis approaches [44,45]. Whenever the evolution and ecology of the virus are inseparably linked [1], case report and phylogenetic summaries are co-dependent. In general, this reduces the degrees of freedom

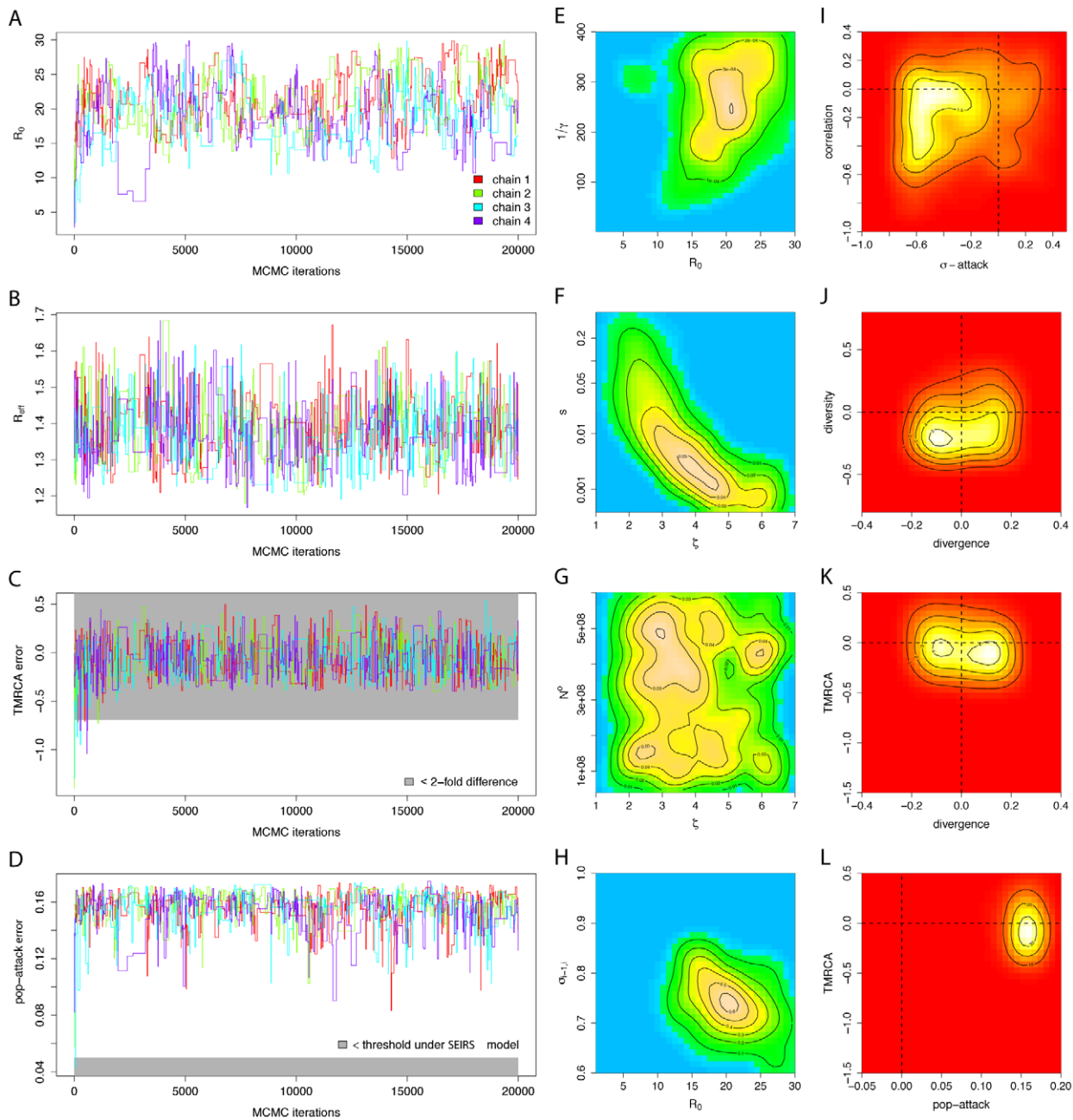


Figure 5. Phyldynamic inference and goodness-of-fit analysis of the spatial epochal evolution model. (A–D) MCMC trajectories as in Figure 3. The summaries *TMRCA* and *pop-attack* are in conflict and cannot be simultaneously fitted, so that the tolerance for *pop-attack* was relaxed. Samples before iteration 5000 were discarded. (E–H) Two-dimensional histograms of parts of the ABC fit as in Figure 3. Partial cross-immunity is relatively low and correlates negatively with R_0 . (I–L) Two-dimensional histograms of parts of the joint density of summary errors as in Figure 3. The epochal evolution model captures the irregularity in H3N2 case report attack rates, and the divergence, diversity and narrowness of the HA phylogeny well, albeit at a high residual selection parameter that is essentially always included. However, under this fitted model, population-level attack rates are in conflict with the *TMRCA*'s and cannot be simultaneously reproduced when ϕ° is weak. doi:10.1371/journal.pcbi.1002835.g005

of a phyldynamic model in reproducing features of both types of data simultaneously, and may reveal model inconsistencies. For example, the fitted epochal evolution model could not reproduce the *TMRCA*'s and the population attack rates at the same time (Figure 5L).

The reported parameter estimates and summary errors are derived by conditioning only on the phyldynamic summaries and

weighting schemes described in Table 1. ABC is sensitive to the chosen summary statistics and the tolerances τ_k since they determine how the prior $\pi(\theta)$ is re-weighted in light of the presented evidence (see for example Table S3 and Figure S4 in Text S1) [26]. Here, we chose broad enough tolerances τ_k such that the weighting schemes are robust to differences in surveillance time series from the Netherlands, France and the US. This

Table 3. Goodness of fit to summaries of H3N2 surveillance data and phylogenies.

summary	mean \pm std. dev., 95% conf. interval of			
	posterior density under the SEIRS model	comments	posterior density under the epochal evolution model	comments
μ -attack	-0.67 ± 0.42 , [-1.23, 0.22]	encompassing values for all countries	-0.54 ± 0.45 , [-1.18, 0.24]	encompassing values for all countries
σ -attack	-0.27 ± 0.27 , [-0.65, 0.18]	encompassing values for all countries	-0.29 ± 0.2 , [-0.67, 0.24]	encompassing values for all countries
explosiveness	-0.12 ± 0.14 , [-0.42, 0.09]	explosiveness of Dutch data not matched well	-0.06 ± 0.23 , [-0.48, 0.27]	encompassing values for all countries
correlation	-0.84 ± 0.08 , [-0.85, -0.79]	inconsistent	-0.15 ± 0.25 , [-0.48, 0.16]	encompassing values for all countries
pop-attack	0.03 ± 0.02 , [-0.04, 0.05]	consistent	0.19 ± 0.04 , [0.1, 0.17]	inconsistent in conflict with TMRCA
divergence	0.06 ± 0.15 , [-0.12, 0.34]	consistent	0 ± 0.12 , [-0.18, 0.18]	consistent
diversity	-0.43 ± 0.12 , [-0.59, -0.21]	inconsistent by a factor < 2	-0.08 ± 0.17 , [-0.33, 0.23]	consistent
lineages	-1.2 ± 0.07 , [-1.29, -1.08]	inconsistent by a factor ≥ 2	-0.48 ± 0.18 , [-0.2, -0.73]	inconsistent by a factor < 2
TMRCA	-0.73 ± 0.19 , [-1.04, -0.41]	inconsistent by a factor ≥ 2	-0.06 ± 0.19 , [-0.35, 0.27]	consistent
nclust	-		-0.03 ± 0.15 , [-0.24, 0.2]	consistent

doi:10.1371/journal.pcbi.1002835.t003

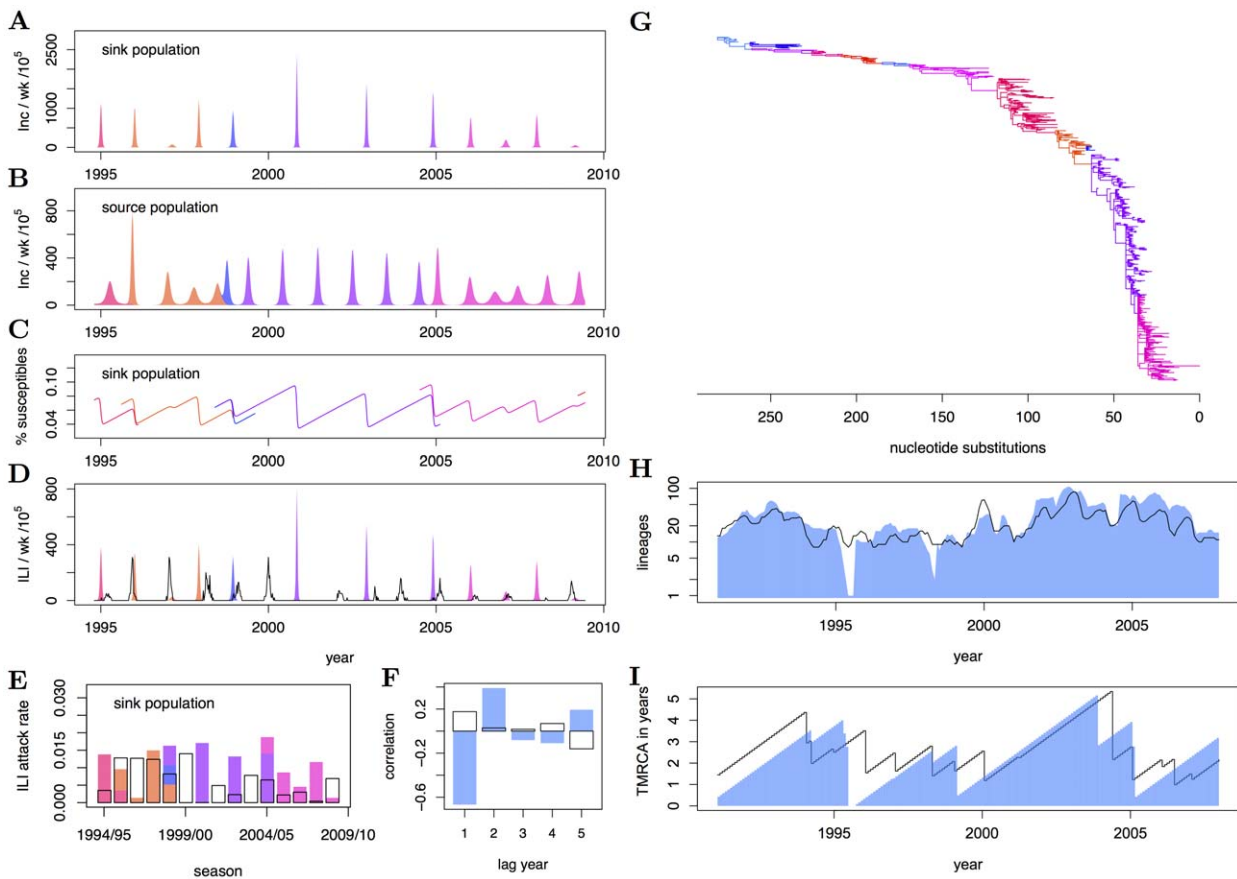


Figure 6. Phylodynamics arising under the fitted, spatial epochal evolution model. (A–I) Subplots are as in Figure 4. Typically, simulations under the fitted model display large infection waves and strong genetic bottlenecks at antigenic cluster invasions that are followed by refractory dynamics (A,B,H). These intrinsic dynamics result in irregular variation in case report attack rates that is well in line with H3N2 time series of several countries (E,F). Model parameters are $R_0=24$, $1/\phi=0.9$, $1/\nu=1.8$, $1/\gamma=340$, $\sigma_{i-1,i}=0.76$, $\rho=0.7$, $N^c=2.7 \times 10^8$, $1/\mu^c=50$, $\phi^l=0.46$, $\phi^c=0.012$, $m^l=10 \times 10^6$, $m^c=0.02$, $\zeta=4.54$, $s=2.5 \times 10^{-3}$, $\kappa=2$, $\lambda=0.78$, and the intrinsic dynamics are generally less pronounced when ϕ^c is higher; see Text S1.

doi:10.1371/journal.pcbi.1002835.g006

approach seems appropriate to avoid overfitting in the context of the limitations of syndromic influenza surveillance, but may be less suited in the analysis of other viral infectious diseases. Figures 3 and 5 illustrate that the resulting dimension reduction regularizes the underlying, intractable likelihood into a smooth, yet well-defined surface such that key phylogenetic model parameters are identifiable and goodness-of-fit can be characterized. It remains unclear to what degree the use of sufficient statistics or the full historical data would be desirable. Indeed, when infectious pathogens escape immunity, the likelihood surface can be especially complex [18,43]. Likelihood-based inference is then sensitive to small changes in the complete historical data [46,47], which can be problematic when the reported incidence time series or viral phylogeny is itself subject to considerable uncertainty and/or bias [2,48].

Application to H3N2 phylodynamics

We used ABC to fit mechanistic phylodynamic models of inter-pandemic influenza A(H3N2) to summaries of surveillance data from the Netherlands and sequence data from Northern Europe. Influenza is a globally circulating virus, and the mechanistic models considered must account for the replenishment of genetic variants from outside Northern Europe in order to reproduce features of influenza's phylogeny. In contrast, semi- or non-parametric models of population dynamics that are used in coalescent methods do not necessarily require this layer of spatial complexity [9,20]. Here, the mechanistic structure of Eqns. (5) constrains the set of possible phylodynamic trajectories in such a way that influenza's global disease dynamics must be explicitly accounted for. Put more generally, the quantitative features of H3N2 sequence and incidence data contain sufficient information to determine at least some basic aspects of phylodynamic process models statistically.

The two models we analyzed show clear limitations in their ability to replicate features of H3N2 sequence and surveillance data simultaneously, and the ABC error diagnostics give some indication how these models could be refined (Figures 3 and 5). For example, the phylogenies generated under the SEIRS and the epochal evolution models have, across time, more lineages than the observed HA phylogeny (Table 3). One possible explanation is that localized extinctions may not occur sufficiently often under the re-seeding source-sink framework, suggesting that models with more detailed population structure, either in space or by age, may result in thinner phylogenies. Accounting for these types of population structure can be critical for understanding viral phylodynamics; here we showed that the fit of the epochal evolution model to both sequence and incidence data depends critically on the assumed spatial model structure and the associated ϕ° (see Text S1).

The SEIRS model could not generate the irregularity in observed incidence data. In comparison, our analysis of the epochal evolution model demonstrates that epochal evolutionary processes can easily excite irregular between-season dynamics that match observed data (see Figures 5L and S18 in Text S1). Since the virus is known to be under intense immune selection [2], it seems plausible that antigenic evolution is an important co-factor in explaining influenza's irregular seasonality in temperate regions [49].

Several alternative models have been proposed to reproduce H3N2's narrow HA phylogeny. Here, we identified an additional,

testable constraint for these models on surveillance data, that arises through the phylodynamic interactions in Eqns. (5). The cluster-specific duration of immunity $1/\gamma$ must be sufficiently long to avoid deep phylogenetic branching. If the fitted values of $1/\gamma$ and R_0 are correlated, this in turn implies a characteristic range of population level attack rates that can be tested against available data as in Figure 5L. In particular, the duration of immunity can be lower if the time between replacement events is shorter. Thus, while H3N2's limited standing genetic diversity provides information on the strength of immune interactions between H3N2 antigenic variants, this second constraint may help identify the tempo of antigenic evolution.

For the epochal evolution model with source-sink migration dynamics, the average simulated waiting time is 3.3 ± 0.7 years from the emergence of the current antigenic cluster to the next successfully invading offspring antigenic cluster, and this implies phylodynamics that are inconsistent with the molecular genetic and epidemiological summaries in Table 1 taken together. More frequent and more gradual transitions between antigenic variants that are smaller than H3N2's antigenic clusters would allow for lower estimates of R_0 that are more in line with observed population level attack rates, break weaker refractory oscillations in their onset, and might also provide sufficient, continual selection pressures to explain the fast divergence in H3N2's HA phylogeny [8]. In this case, sequence and surveillance data would point to a finer antigenic resolution than the one suggested through antigenic map analyses [31]. Alternatively, it is also possible that finer population structure, either in space or by age, could increase extinction rates and thereby allow for a narrow HA phylogeny under a broader, more realistic set of epidemiological parameters without accelerating the tempo of antigenic evolution per se.

More broadly, both types of data are now increasingly becoming available for RNA viruses [9–13]. This study indicates that these data, when considered simultaneously, may drastically constrain parameter space and readily expose model deficiencies, so that ABC appears as a well-suited tool to explore the phylodynamics of RNA viruses.

Supporting Information

Text S1 Supplementary online text describing the influenza A (H3N2) sequence and surveillance data used, ABC algorithms and summary statistics, ABC analyses on simulated data and sensitivity analyses.
(PDF)

Acknowledgments

We thank Steven Riley, Simon Cauchemez, Anton Camacho, David Rasmussen and Neil Ferguson as well as three reviewers for their time and thoughtful comments. Computations were performed at the Imperial College High Performance Computing Service (<http://www.imperial.ac.uk/ict/services/teachingandresearchservices/highperformancecomputing/>), and we thank Simon Burbidge and Matt Harvey.

Author Contributions

Conceived and designed the experiments: OR. Performed the experiments: OR. Analyzed the data: OR CF KK. Contributed reagents/materials/analysis tools: GD AM. Wrote the paper: OR CF KK.

References

- Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, et al. (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303: 327–332.
- Cox N, Subbarao K (2000) Global epidemiology of influenza: past and present. *Annual Review of Medicine* 51: 407–421.

3. Gog JR, Grenfell BT (2002) Dynamics and selection of many-strain pathogens. *Proc Natl Acad Sci USA* 99: 17209–17214.
4. Ferguson NM, Galvani AP, Bush RM (2003) Ecological and immunological determinants of influenza evolution. *Nature* 422: 428–433.
5. Koelle K, Cobey S, Grenfell B, Pascual M (2006) Epochal Evolution Shapes the Phylogenetics of Interpandemic Influenza A (H3N2) in Humans. *Science* 314: 1898–1903.
6. Gog J (2008) The impact of evolutionary constraints on influenza dynamics. *Vaccine* 26: C15–C24.
7. Koelle K, Khatri P, Kamradt M, Kepler T (2010) A two-tiered model for simulating the ecological and evolutionary dynamics of rapidly evolving viruses, with an application to influenza. *Journal of The Royal Society Interface* 7: 1257–1274.
8. Bedford T, Rambaut A, Pascual M (2012) Canalization of the evolutionary trajectory of the human influenza virus. *BMC Biology* 10. doi:10.1186/1741-7007-10-38
9. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, et al. (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453: 615–619.
10. Siebenga J, Vennema H, Renckens B, De Bruin E, Van Der Veer B, et al. (2007) Epochal evolution of GGI.4 norovirus capsid proteins from 1995 to 2006. *Journal of Virology* 81: 9932–9941.
11. Donaldson E, Lindesmith L, LoBue A, Baric R (2010) Viral shape-shifting: norovirus evasion of the human immune system. *Nature Reviews Microbiology* 8: 231–241.
12. Fischer W, Ganusov V, Giorgi E, Hraber P, Keele B, et al. (2010) Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS One* 5: e12303.
13. Jabara C, Jones C, Roach J, Anderson J, Swanstrom R (2011) Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc Natl Acad Sci USA* 108: 20166–20171.
14. Keeling MJ, Rohani P (2008) Modeling infectious diseases in humans and animals. Princeton: Princeton Univ Press. 408p.
15. Becker N, Britton T (1999) Statistical studies of infectious disease incidence. *Journal of the Royal Statistical Society: Series B* 61: 287–307.
16. Ionides EL, Breto C, King AA (2006) Inference for nonlinear dynamical systems. *Proc Natl Acad Sci USA* 103: 18438–18443.
17. Andrieu C, Doucet A, Holenstein R (2010) Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B* 72: 2693–2742
18. Finkenstädt B, Morton A, Rand D (2005) Modelling antigenic drift in weekly flu incidence. *Statistics in Medicine* 24: 3447–3461.
19. Cauchemez S, Ferguson N (2008) Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *Journal of the Royal Society Interface* 5: 885–897.
20. Pybus O, Rambaut A (2009) Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics* 10: 540–550.
21. Bedford T, Cobey S, Beerli P, Pascual M (2010) Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). *PLoS Pathogens* 6: e1000918.
22. Bahl J, Nelson M, Chan K, Chen R, Vijaykrishna D, et al. (2011) Temporally structured metapopulation dynamics and persistence of influenza A H3N2 virus in humans. *Proc Natl Acad Sci U S A* 108: 19359–19364.
23. Drummond A, Rambaut A, Shapiro B, Pybus O (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22: 1185–1192.
24. Minin V, Bloomquist E, Suchard M (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular Biology and Evolution* 25: 1459–1471.
25. Rasmussen D, Ratmann O, Koelle K (2011) Inference for nonlinear epidemiological models using genealogies and time series. *PLoS Computational Biology* 7: e1002136.
26. Marin JM, Pudlo P, Robert C, Ryder R (2011) Approximate Bayesian computational methods. *Statistics and Computing*. In press
27. Ratmann O, Andrieu C, Wiuf C, Richardson S (2009) Model criticism based on likelihood-free inference, with an application to protein network evolution. *Proc Natl Acad Sci U S A* 106: 10576–10581.
28. O'Hara R, Sillanpää M (2009) A review of Bayesian variable selection methods: what, how and which. *Bayesian Analysis* 4: 85–118.
29. Dijkstra F, Donker G, Wilbrink B, Van Gageldonk-Lafeber A, Van Der Sande M, et al. (2009) Long time trends in influenza-like illness and associated determinants in The Netherlands. *Epidemiol Infect* 137: 473–9.
30. Meijer A, Rimmelzwaan G, Dijkstra F, Donker G (2009) Actuele ontwikkelingen betreffende influenza; griepspotters in actie. *Tijdschr Infect* 4: 176–84.
31. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305: 371–376.
32. Koelle K, Kamradt M, Pascual M (2009) Understanding the dynamics of rapidly evolving pathogens through modeling the tempo of antigenic change: influenza as a case study. *Epidemics* 1: 129–137.
33. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, et al. (2008) The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320: 340–346.
34. Marjoram P, Molitor J, Plagnol V, Tavaré S (2003) Markov Chain Monte Carlo without likelihoods. *Proc Natl Acad Sci U S A* 100: 15324–15328.
35. Arinaminpathy N, Ratmann O, Koelle K, Epstein S, Prince G, et al. (2012) Impact of cross-protective vaccines on epidemiological and evolutionary dynamics of influenza. *PLoS* (2011). *Proc Natl Acad Sci U S A* 109: 3173–3177.
36. Ferguson N, Cummings D, Cauchemez S, Fraser C, Riley S, et al. (2005) Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437: 209–214.
37. Pease C (1987) An evolutionary epidemiological mechanism, with applications to type A influenza. *Theoretical Population Biology* 31: 422–452.
38. Monto A, Koopman J, Longini Jr I (1985) Tecumseh study of illness. xiii. Influenza infection and disease, 1976–1981. *American Journal of Epidemiology* 121: 811–822.
39. Mills C, Robins J, Lipsitch M (2004) Transmissibility of 1918 pandemic influenza. *Nature* 422: 904–906.
40. Cauchemez S, Valleron A, Boëlle P, Flahault A, Ferguson N (2008) Estimating the impact of school closure on influenza transmission from sentinel data. *Nature* 452: 750–754.
41. Recker M, Pybus OG, Nee S, Gupta S (2007) The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc Natl Acad Sci USA* 104: 7711–7716.
42. Viboud C, Alonso W, Simonsen L (2006) Influenza in tropical regions. *PLoS medicine* 3: e89.
43. Truscott J, Fraser C, Cauchemez S, Meechai A, Hinsley W, et al. (2011) Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza. *Journal of The Royal Society Interface* 9: 304–312.
44. Ades A, Sutton A (2006) Multiparameter evidence synthesis in epidemiology and medical decision-making: current approaches. *Journal of the Royal Statistical Society: Series A* 169: 5–35.
45. Birrell P, Ketschis G, Gay N, Cooper B, Presanis A, et al. (2011) Bayesian modelling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proc Natl Acad Sci U S A* 108: 18238–18243.
46. Ramsay J, Hooker G, Campbell D, Cao J (2007) Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B* 69: 741–796.
47. Wood S (2010) Statistical inference for noisy nonlinear ecological dynamic systems. *Nature* 466: 1102–1104.
48. Stack J, Welch J, Ferrari M, Shapiro B, Grenfell B (2010) Protocols for sampling viral sequences to study epidemic dynamics. *Journal of the Royal Society Interface* 7: 1119–1127.
49. Lipsitch M, Viboud C (2009) Influenza seasonality: lifting the fog. *Proc Natl Acad Sci U S A* 106: 3645.