Bayesian Inference of Ancestral Host–Parasite Interactions under a Phylogenetic Model of Host Repertoire Evolution

MARIANA P. BRAGA[1,2,*], MICHAEL J. LANDIS[2,3], SÖREN NYLIN[1], NIKLAS JANZ[1] AND FREDRIK RONQUIST[4]

[1]*Department of Zoology, Stockholm University, Stockholm, SE-10691, Sweden;* [2]*Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA;* [3]*Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06511, USA; and* [4]*Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden*
Correspondence to be sent to: Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA;
E-mail: mariana@wustl.edu.

Abstract.—Intimate ecological interactions, such as those between parasites and their hosts, may persist over long time spans, coupling the evolutionary histories of the lineages involved. Most methods that reconstruct the coevolutionary history of such interactions make the simplifying assumption that parasites have a single host. Many methods also focus on congruence between host and parasite phylogenies, using cospeciation as the null model. However, there is an increasing body of evidence suggesting that the host ranges of parasites are more complex: that host ranges often include more than one host and evolve via gains and losses of hosts rather than through cospeciation alone. Here, we develop a Bayesian approach for inferring coevolutionary history based on a model accommodating these complexities. Specifically, a parasite is assumed to have a host repertoire, which includes both potential hosts and one or more actual hosts. Over time, potential hosts can be added or lost, and potential hosts can develop into actual hosts or vice versa. Thus, host colonization is modeled as a two-step process that may potentially be influenced by host relatedness. We first explore the statistical behavior of our model by simulating evolution of host–parasite interactions under a range of parameter values. We then use our approach, implemented in the program RevBayes, to infer the coevolutionary history between 34 Nymphalini butterfly species and 25 angiosperm families. Our analysis suggests that host relatedness among angiosperm families influences how easily Nymphalini lineages gain new hosts. [Ancestral hosts; coevolution; herbivorous insects; probabilistic modeling.]

Extant ecological interactions, such as those between parasites and hosts, are often the result of a long history of coevolution between the involved lineages (Elton 1946; Klassen 1992). Specialization is predominant among parasites, including parasitic herbivorous insects, which complete all their larval development on an individual host (Thompson 1994; Forister et al. 2015). But host–parasite interactions are not static, they continuously evolve over time via gains and losses of hosts (Janz and Nylin 2008; Nylin et al. 2018). The colonization of new hosts and loss of old hosts not only shape the evolutionary trajectories of the interacting lineages but can also have large effects at ecological timescales (Nosil 2002; Calatayud et al. 2016). These effects are evident, for example, with emerging infectious diseases and zoonotic diseases (Acha and Szyfres 2003), which involve the colonization of new hosts within and among groups of domesticated species (Subbarao et al. 1998), wildlife (Fisher et al. 2009), and humans (Hahn et al. 2000). Unraveling the processes that underlie changes in species interactions is thus key to understanding evolutionary and ecological phenomena at various timescales, such as the emergence of infectious diseases, community assembly, and parasite diversification (Hoberg and Brooks 2015).

Many methods developed to study historical associations focus on congruence between host and parasite phylogenies (Brooks 1979; Huelsenbeck et al. 1997; de Vienne et al. 2013). Such methods largely fall into two main classes of cophylogenetic approaches: 1) topology- and distance-based methods, which estimate the congruence between two phylogenies (Legendre et al. 2002) and 2) event-based methods, which map the parasite phylogeny onto the host phylogeny using evolutionary events (Ronquist 2003). Typically, cospeciation is the null hypothesis in these methods, where host shifts are invoked only to explain deviations from cospeciation (de Vienne et al. 2013). Moreover, most of these methods do not allow ancestral parasites to be associated with more than one host lineage, thus failing to account for a potentially important driver of parasite diversification (Janz and Nylin 2008).

An alternative approach to studying coevolving host–parasite interactions is to perform ancestral state reconstructions of individual host taxa onto the parasite phylogeny and combine the ancestral host states *a posteriori* into inferred host ranges (e.g., Nylin et al., 2014). Even though this approach allows ancestral parasites to have multiple hosts, it assumes that the interactions between the parasite and each host evolve independently. This has a number of serious drawbacks. For instance, ancestral parasites may be inferred to have an unrealistically high number of hosts, or no host at all. Furthermore, the more narrowly circumscribed the host taxa are, the more likely it is that ancestral parasite lineages are reconstructed as having no hosts. In addition, the independence assumption causes the phylogenetic relationships among hosts to be ignored, meaning that the model assigns equal rates to all colonizations of new hosts regardless of how closely related the new host is to the current hosts being used by the parasite.

A desirable model of host usage should therefore allow parasites to have multiple hosts, while also

allowing for among-host (or context-dependent) effects to influence ancestral host use estimates and gain and loss rates in whatever manner explains the biological data best. One possible solution is to restate the problem of host–parasite coevolution in terms of historical biogeography. For instance, the Dispersal-Extirpation-Cladogenesis model of Ree et al. (2005) allows species ranges to stochastically evolve as a set of discrete areas over time through area gain events (dispersal), area loss events (extirpation), and cladogenetic events (range inheritance patterns that reflect speciational models). Although these methods are designed for biogeographic inference, a similar approach is clearly suitable for more realistic modeling of host–parasite coevolution dynamics, where colonization and loss of hosts (instead of discrete areas) is modeled as a continuous-time Markov process (e.g., Hardy, 2017). In biogeography, the colonization of a new area or the disappearance from a previously occupied area is modeled as a binary trait: the species is either present or absent in the area. While this binary view might be adequate in biogeography, it may be too simplistic for use in the coevolution between hosts and parasites. For instance, it is known that butterflies can utilize a range of plants that they do not regularly feed on in the wild, and it has been suggested that these potential hosts have played an important role in the evolution of host use in butterflies, by increasing the variability in host use through time and across clades (Janz et al. 2016; Braga et al. 2018). This hypothesis can only be directly tested, however, if we explicitly model the evolution of host use as a two-step process, which cannot be done with the binary methods that are used today to study host–parasite coevolution or biogeography.

Here, we propose a model where a parasite is assumed to have a *host repertoire*, defined as the set of all potential and actual hosts for that parasite. In this model, the colonization of a new host involves two steps: first, the parasite gains the ability to use the new host (it becomes a potential host), and then starts actually using it in nature (it becomes an actual host). These two steps can be interpreted as the inclusion of the new host into the fundamental host repertoire of the parasite, and then into the realized host repertoire of the parasite— analogous to the concepts of fundamental and realized niches (Nylin et al. 2018; Larose et al. 2019) and to the compatibility and encounter filters of Combes (2001). Similarly, the complete loss of a host from a parasite's realized repertoire involves two steps. First, it changes from an actual to a potential host, and then it is lost completely from the host repertoire. For example, if the geographic range of a host contracted to become allopatric with respect to a parasite's geographic range, the host would remain as part of the fundamental repertoire until the parasite completely lost the ability to use the host, in which case the host would be lost from the repertoire. Even when in sympatry, the evolution of a new defense mechanism by the host may prevent the parasite from using that host. However, since host use

is a complex and multidimensional trait, it is unlikely that a parasite loses all the machinery necessary to use a host in one single event, and it may well retain some ability to survive on the host. Thus, a complete gain or loss of a host lineage requires at least two events, and three host–parasite interaction states are necessary for such a two-step model: the host is used (actual host), the parasite has some ability to use the host but does not use it in nature (potential host), and the parasite cannot use the host (nonhost).

In this article, we develop a Bayesian approach to coevolutionary inference based on such a model of host repertoire evolution, inspired by the previous work on similar biogeographic inference problems by Landis et al. (2013). The basic two-state biogeographic model, when applied to coevolution, accommodates both multiple ancestral hosts and changes in host configurations over time that correspond to evolutionary changes in host usage. We extend this model to also include a two-step host colonization process, such that the fundamental host repertoire can persist over time and affect the evolution of the realized repertoire. We have implemented the model in RevBayes (Höhna et al. 2016), allowing us to simulate data as well as perform Bayesian Markov chain Monte Carlo (MCMC) inference under the model. This Bayesian framework allows one to estimate the joint distribution of host gain and loss rates, the effect (if any) of phylogenetic distances among hosts upon host gain rates, and the historical sequences of evolving host repertoires among the parasites. Using simulations, we explore the statistical behavior of our approach and demonstrate its empirical application with an analysis of the coevolution between Nymphalini butterflies and their angiosperm hosts.

METHODS

Model Description

We are interested in modeling the evolution of ecological interactions between M extant parasite taxa and N host taxa, where each parasite uses one or more hosts. Rooted and time-calibrated phylogenetic trees describe the evolutionary relationships among the M parasite taxa and among the N host taxa. In this study, the trees are considered to be known without error. In principle, it would be straightforward for the model to accommodate phylogenetic uncertainty in the host or parasite trees but MCMC inference may prove challenging under such conditions.

Each parasite taxon has a host repertoire, which is represented by a vector of length N that contains the information about which hosts the given parasite uses. The interaction between the mth parasite and the nth host is denoted $x_{m,n}$. At any given time, each host taxon can assume one of three states with respect to a parasite lineage: $x_{m,n}$ is equal to 0 (nonhost), 1 (potential host), or 2 (actual host). Criteria for how to code nonhost, potential

host, and actual host states will depend on the host–parasite system under study; below, we provide criteria for our Nymphalini data set that may act as guidelines. We allow all host repertoires in which the parasite has at least one actual host. Thus, the state space, S, includes $3^N - 2^N$ host repertoires for N hosts.

Here, we define the transition from state 0 to state 1 as the gain of the ability to use the host, and the transition from state 1 to state 2 as the time when the parasite actually starts to use the host in nature. If we assume that gains and losses of hosts occur according to a continuous-time Markov chain, the probability of a given coevolutionary history between a parasite clade and their hosts can be easily calculated (Ree and Smith 2008). This calculation is based on a matrix, $\mathbf{Q}$, containing the instantaneous rates of change between all pairs of host repertoires, and thus describing the Markov chain. Based on the $\mathbf{Q}$ matrix, it is possible to calculate the transition probability of the observed host repertoires at the tips of the parasite tree by marginalizing over the infinite number of histories that could produce the observed host repertoires. Unfortunately, computing these transition probabilities becomes intractable as the number of host repertoire configurations, S, grows large. Modeling host repertoire evolution for host repertoire size $N = 7$ requires an $S \times S$ rate matrix defined for $S = 3^7 - 2^7 = 2059$, causing $\mathbf{Q}$ to be too large for efficient inference. In order to handle large host repertoires, we numerically integrate over possible histories along the parasite tree using data augmentation and MCMC rather than analytically computing the probabilities using matrix exponentiation. This data augmentation approach has been used to model sequence evolution for protein-coding genes (Robinson 2003) and historical biogeography (Landis et al. 2013; Quintero and Landis 2020), suggesting the framework may be useful to model host–parasite interactions as well. In this study, we assume that both daughter lineages identically inherit their host repertoires from their immediate ancestor at the time of cladogenesis. It is possible to extend the model to consider more complex speciation scenarios, such as daughter lineages dividing the host repertoire between them, but we leave the exploration of such model variants for future studies.

We define a model where the gain of a host (both $0 \rightarrow 1$ and $1 \rightarrow 2$) depends on the phylogenetic distance between the available hosts and those currently used by a lineage. Figure 1 schematically illustrates the evolutionary dynamics of the model using $M = 4$ parasite species and $N = 5$ host species, while assuming that host gain rates are independent (Fig. 1a,c) or dependent (Fig. 1b,d) of phylogenetic distances among hosts. To formalize these dynamics, let $q_{\mathbf{y},\mathbf{z}}^{(a)}$ be the rate of change from host repertoire $\mathbf{y}$ to repertoire $\mathbf{z}$ by changing the state of host a. Also, let λ_{ij} be the rate at which an individual host changes from state i to state j, and $\eta(\mathbf{y}, a, \beta)$ be a phylogenetic-distance rate modifier. In practice, we constrain the host change rates such that

$0 \le \lambda_{ij} \le 1$, and estimate the parameter, μ, which we use to rescale all rates in λ. We refer to μ as the maximum rate (or rate, for short) of host repertoire evolution, since $0 \le \mu\lambda_{ij} \le \mu$. The phylogenetic-distance rate modifier function, η, further rescales the rates of host gain to allow new hosts that are closely related to the parasite's current hosts to be colonized at higher rates than distantly related hosts. Poulin (2010) reported a negative exponential relationship between host relatedness and parasite faunas they share, emulating the relationship between dispersal rates and the geographical distances among islands (MacArthur and Wilson 1967) by treating hosts as islands and host ancestry as space (Janzen 1968). Taking this as inspiration, we defined the phylogenetic-distance rate modifier function to impose a negative exponential relationship between the host gain rates and the phylogenetic distances of newly acquired hosts. We define the instantaneous rate of change as

$$q_{\mathbf{y},\mathbf{z}}^{(a)} = \begin{cases} \mu\lambda_{10}, & \text{if potential host loss } (y_a = 1 \text{ and } z_a = 0) \\ \mu\lambda_{01}\eta_1 & \text{if potential host gain } (y_a = 0 \text{ and } z_a = 1) \\ (\mathbf{y}, a, \beta) & \\ \mu\lambda_{21}, & \text{if actual host loss } (y_a = 2 \text{ and } z_a = 1) \\ \mu\lambda_{12}\eta_2 & \text{if actual host gain } (y_a = 1 \text{ and } z_a = 2) \\ (\mathbf{y}, a, \beta) & \\ 0, & \text{if direct transition between states 0 and} \\ & 2(|y_a - z_a| > 1) \\ 0, & \text{if } \mathbf{y} \text{ and } \mathbf{z} \text{ differ at more than one host} \\ 0 & \text{if } \mathbf{y} \text{ does not contain at least one actual} \\ & \text{host} \end{cases}$$

and the phylogenetic-distance rate modifier function as

$$\eta(\mathbf{y}, a, \beta) = e^{-\beta d/\bar{d}}, \tag{0.1}$$

where β controls the effect of d, the average pairwise phylogenetic distance between the new host, a, and the hosts currently occupied in $\mathbf{y}$; and $\bar{d}$ is the average phylogenetic distance between all pairs of hosts. Pairwise phylogenetic distance is defined as the sum of branch lengths separating two leaf nodes. The difference between η_1 and η_2 is that for the first function, pairwise distances are calculated between the new host and all potential and actual hosts, while in the second only actual hosts are included. This allows for a model formulation where the effect of host distances on λ_{01} and on λ_{12} are independent, while still allowing a formulation where they are equal. Regardless of its formulation, the effect of host distances depends on how it is scaled by β. If $\beta > 0$, the gain rate of phylogenetically close hosts is higher than distant hosts. If $\beta = 0$, the gain rate of host a is equal to the unmodified gain rate, λ_{01} or λ_{12}.

The process of host repertoire evolution with different combinations of μ and β will produce data sets with different levels of phylogenetic conservatism. When μ is low, the host repertoire evolves slowly,
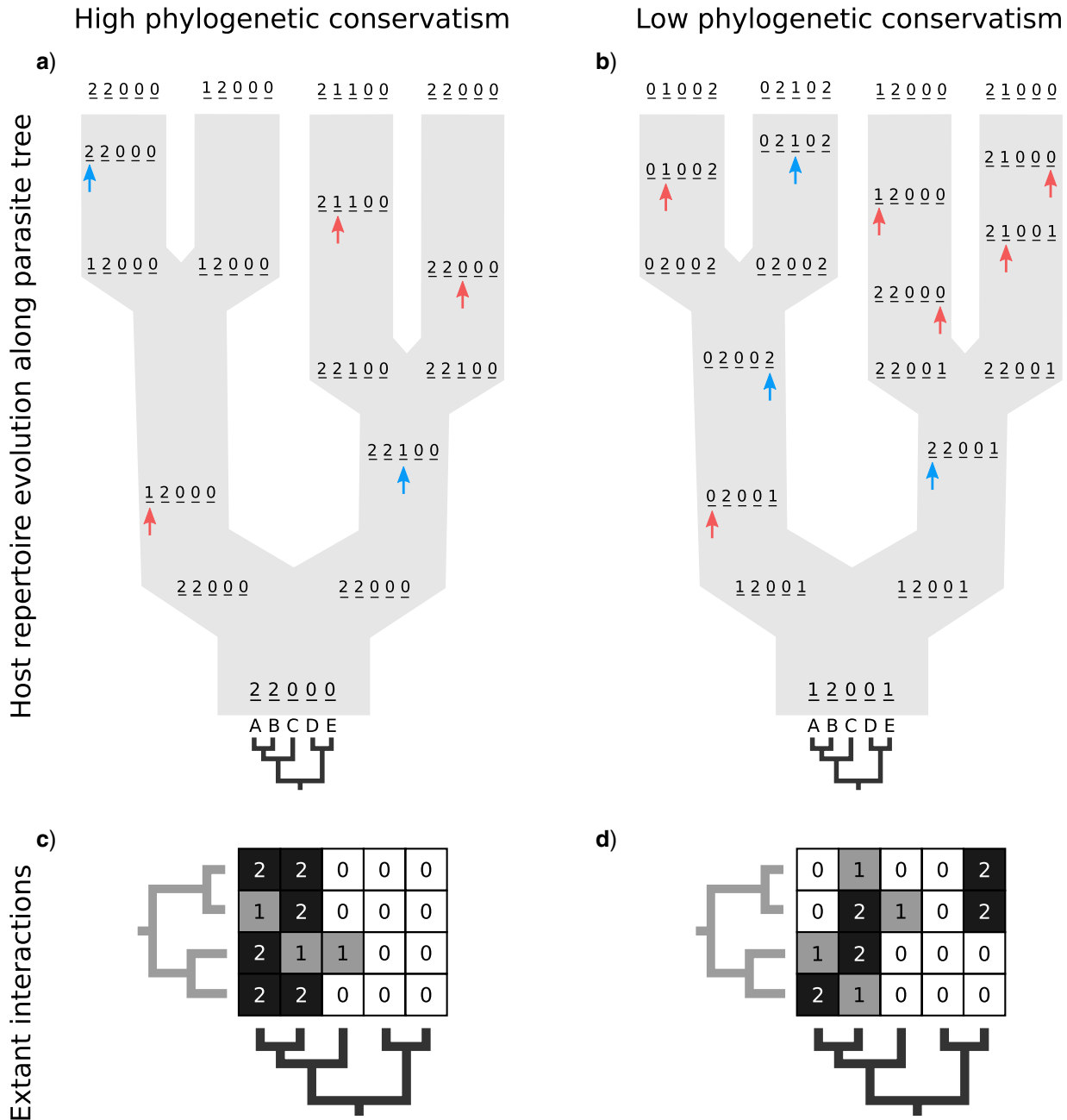
FIGURE 1.   Host repertoire evolution within a hypothetical parasite clade and resulting host-parasite interactions. Two examples of coevolutionary histories between four parasites and five hosts are shown to illustrate how the model works. Host repertoires evolve by gains (0 → 1 and 1 → 2, blue arrows) and losses (1 → 0 and 2 → 1, red arrows). Coevolutionary histories in (a) and (b) produce the interactions in (c) and (d), respectively. In (c) and (d), each column represents one host, and each row represents the host repertoire of one parasite. High phylogenetic conservatism is produced when the rate of repertoire evolution, μ, is low and the effect of the phylogenetic distance between hosts, β, is high. Conversely, low phylogenetic conservatism is produced when μ is high and β is low.

resulting in similar repertoires between closely related butterflies, that is, high phylogenetic conservatism. We use phylogenetic distances between hosts as a proxy for differences in traits related to host use. Thus, β = 0 is equivalent to fast trait evolution, where phylogenetic distance is not a good proxy for trait distance. On the other hand, when β is high, phylogenetic and trait distances match (high phylogenetic conservatism),

hence closely related hosts are more similar then distantly related hosts.

We fit this model using the Bayesian data augmentation strategy described in Landis et al. (2013). The method estimates the joint posterior probability of model parameters, $\theta = (\mu, \lambda, \beta)$, and data-augmented evolutionary histories, X_{aug}, conditional on the observed host repertoire data, X_{obs}, the rooted time-calibrated

tree for the parasites, Ψ_p, and rooted time-calibrated tree for the hosts, Ψ_h, using MCMC. To sample values from the posterior, $P(X_{aug}, \theta | X_{obs}, \Psi_p, \Psi_h)$, new parameter values for μ, λ, and β are proposed using standard Metropolis–Hastings proposals for updating simple parameters (Hastings 1970). Analogously, our MCMC stochastically proposes and/or accepts new augmented host repertoire histories using the Metropolis–Hastings algorithm. Augmented histories are proposed using two types of MCMC moves: branch-specific moves and node-and-branch moves. Branch-specific moves propose a new augmented history by sampling a branch from the phylogeny uniformly at random, then proposing new histories for a subset of host characters using the rejection sampling method of Nielsen (2002) under the assumption that all host characters evolved under mutual independence ($\beta = 0$); this assumption allows us to rapidly propose new augmented histories. Although augmented histories are proposed assuming host characters evolve independently, we compute the acceptance probability for the branch-specific move by considering the full-featured model probability that allows for nonindependent rates of character change when calculating the Metropolis–Hastings ratio. Thus, the augmented histories are sampled in proportion to their posterior probabilities under the full model. Node-and-branch moves involve sampling new host repertoire states for a node sampled uniformly at random within the parasite tree, along with the three branches incident to the node. Together, the branch-specific moves, the node-and-branch moves, and the parameter moves allow MCMC to estimate the posterior probability of combinations of host repertoire histories and evolutionary parameters. Further details are provided in Landis et al. (2013) and Quintero and Landis (2020).

Model Selection

When $\beta = 0$, the phylogenetic-distance dependent model, M_D becomes a mutual-independence model, M_0, where the interaction between the parasite and each host evolves independently. These models are therefore nested ($M_0 \subseteq M_D$), and we can compute Bayes factors for model M_D over model M_0 using the Savage–Dickey ratio (Verdinelli and Wasserman 1995; Suchard et al. 2001; Marin et al. 2010), defined as

$$B_{D,0} = \frac{P(\beta = 0 | M_D)}{P(\beta = 0 | \mathbf{x}_{obs}, M_D)}, \qquad (0.2)$$

where $P_D(\beta = 0 | M_D)$ is the prior probability and $P(\beta = 0 | \mathbf{x}_{obs}, M_D)$ is the posterior probability, both defined in terms of the phylogenetic-distance dependent model, M_D, at the restriction point $\beta = 0$, where M_D and M_0 are equivalent. In practice, we are careful to assign β a prior that is independent of other model parameters to satisfy the conditions of the "naïve" Savage–Dickey ratio (Heck 2019). While we could directly compute the prior probability of $\beta = 0$, we approximated the

posterior probability of $\beta = 0$ using the kernel density estimator from the R package *kdensity* (Chen 2000) with a gamma function, which only takes positive values, with a bandwidth of 0.02. To interpret if and how Bayes factors favored the phylogenetic-distance dependent model, M_D, we followed the guidelines of Jeffreys (1961): model M_0 is favored for Bayes factors with values less than 1, insubstantial support is awarded to model M_D for values between 1 and 3, substantial support for values between 3 and 10, strong support for values between 10 and 30, very strong support for values between 30 and 100, and decisive support for values greater than 100.

Data Analysis

Simulation study We simulated 50 data sets for each of nine combinations of values for the rate of host repertoire evolution, μ (0.1, 0.5, and 1.0), and values of β (0, 1, and 4). These parameter combinations produce data sets with varying degrees of phylogenetic conservatism for both parasites and hosts: when μ is low, closely related butterflies have similar repertoires; and when β is high, closely related hosts are used by the butterflies (Fig. 2). Each data set contained 34 insects and 25 hosts and was produced by simulating host repertoire evolution in the parasite tree used in the empirical study (see below). Host gain and loss rates were chosen to resemble the rates inferred from the empirical analysis. This simulation was designed to assess our statistical power to detect the effect of phylogenetic distance among hosts upon host gain rates given the size of our empirical data set and the type of variation we expected it to contain.

We ran independent MCMC analyses for each set of 50 data sets, under the phylogenetic-distance dependent model. We then quantified how well the posterior probabilities of coevolutionary histories corresponded to the true history known from each simulation. Specifically, we first computed the posterior probability of interaction between each host and each internal node in the butterfly tree, for states 1 and 2 separately. Then, we compared the true coevolutionary history of each simulation to the corresponding posterior distribution of the sampled coevolutionary histories (observed probability). In order to estimate how much of the observed accuracy in ancestral state estimation should be expected by chance, we calculated, for each simulated coevolutionary history, the mean posterior probability for each true state across the 50 replicates with the same parameter combination (probability expected by chance). This shows what is the expected posterior probability for each true ancestral state when the data set used for inference has similar properties to the true data set, but not the same tip states.

Empirical study In order to validate our method, we compiled data from the literature for butterflies from the tribe Nymphalini (Nymphalidae) and their host plants (see Supplementary Material available on Dryad at https://doi.org/10.5061/dryad.x95x69pdw
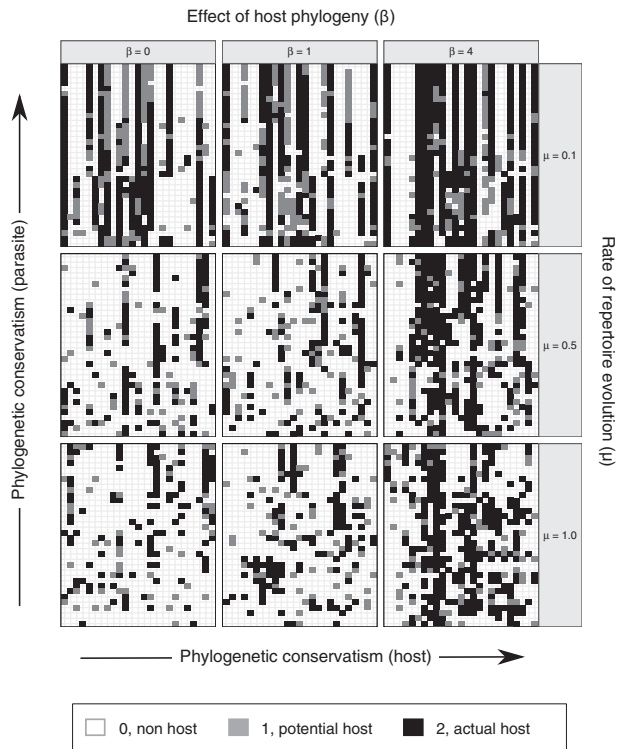
FIGURE 2. Simulated data sets for nine parameter combinations. Interactions between Nymphalini butterflies and their host plants for one of 50 simulations with each parameter combination. In each of the nine data sets, each column represents one host in the repertoire and each row shows the host repertoire of one butterfly species. Rows and columns are ordered by butterfly and plant phylogeny, respectively. When phylogenetic conservatism in host–parasite interactions is low for both hosts and parasites, the interactions are more randomly spread (matrix at bottom-left corner). As phylogenetic conservatism among parasites increases, host repertoires (rows) become more similar (upper matrices). When there is phylogenetic conservatism among hosts, host repertoires include more closely related hosts (neighboring columns; matrices to the right).

for reference list). We chose this butterfly clade because we expect that a large fraction of the real potential hosts are known, as there have been systematic experimental studies of larval feeding ability. The data set included 34 butterflies species and plants from 16 angiosperm families (Supplementary Figs. S1 and S2 available on Dryad). For each butterfly species, host plants commonly used in nature were coded as "actual hosts" and plants never used were coded as "nonhosts." Plants that are not commonly used in nature, but for which there is strong evidence (field observation or experiment) that the larvae can feed upon them, were coded as "potential hosts."

Because we lack the information on potential hosts for most host–parasite systems (i.e., hosts are usually only classified as hosts or nonhosts), we tested whether our model is able to recover the same parameter estimates and coevolutionary histories when all the potential hosts are recoded as nonhosts. For that, we ran the same analysis as for the full data set, but replaced all the 1s

from the empirical data set with 0s. Then, we compared the posterior probabilities inferred from the full data set and the recoded data set. To assess the similarities between the coevolutionary histories inferred using the different data sets, we calculated summary statistics for the absolute difference in probability of each interaction between hosts and internal nodes in the butterfly tree.

For both the simulation and empirical studies, we used the phylogenetic relationships between butterfly species in the Nymphalini tribe as proposed by Chazot et al. (2020, Supplementary Fig. S3 available on Dryad) and the phylogenetic relationships between angiosperm families proposed by Magallón et al. (2015). Although our framework allows the inclusion of a large number of hosts in the same analysis, computational time increases significantly with the size of the host repertoire. We therefore chose to include 25 plant families, which allows the inclusion of all the 16 host lineages used by any of the butterflies. For biologists interested in applying our method to their data sets, we found our data set of 34 parasite species and 25 host lineages took approximately 1 week to analyze, while a larger data set comprised of 66 parasite species and 50 host lineages took about 3 weeks to complete the same number of generations.

To ensure the inclusion of all plant lineages that might have been used as hosts in the past, we pruned the angiosperm phylogenetic tree so that all 16 families in the data set were included, and the remaining branches were collapsed to more ancestral nodes until only 25 tips were left. We then pruned all the branches leading up to the tips to the time of origin of the butterfly clade (approx. 22 Ma), and this pruned tree was then used to calculate phylogenetic distances between hosts. To simplify the analysis, we hold the phylogenetic distances between plant families constant through geological time, even though the distances would be expected to increase as evolution proceeds towards the recent. After preliminary analyses, however, we found that time-calibrated distances between host lineages (i.e., phylogenetic distance calculated as described above) have a weak effect on host gains. On the other hand, "cladogenetic" distances (calculated as the phylogenetic distance between host lineages after setting each branch length equal to 1) had a much stronger effect. Thus, we present the results using cladogenetic distances in the main paper and results using the untransformed "anagenetic" distances in the Supplementary Figs. S4–S6 available on Dryad).

We summarized inferred coevolutionary histories in two ways. First, we calculated the posterior probability for fundamental and realized host repertoires at internal nodes of the Nymphalini phylogeny based on the frequency with which states 1 and 2 were sampled for each host during MCMC. Second, in order to facilitate the visualization of ancestral state reconstructions, we reduced the dimensionality of the host repertoire by
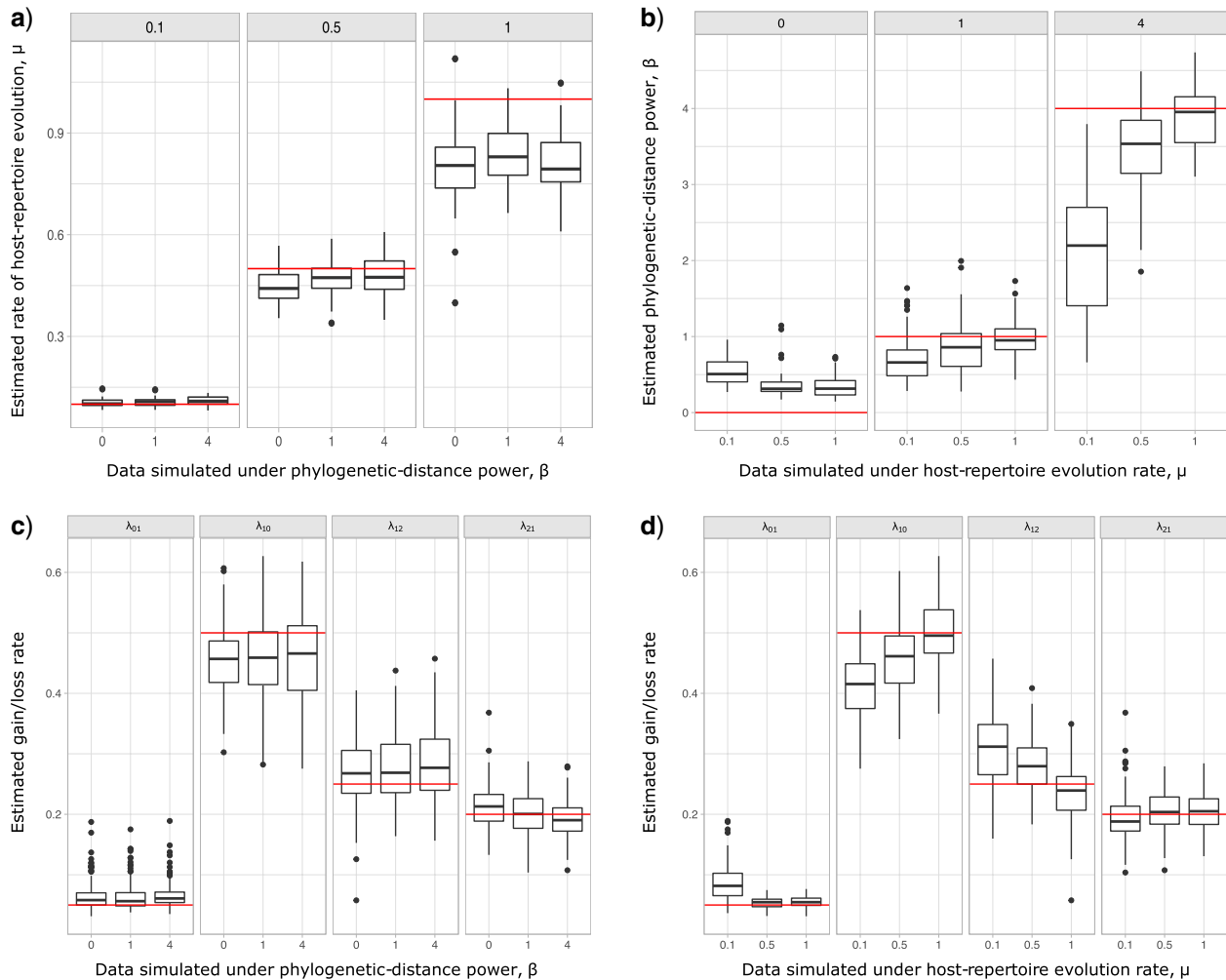
FIGURE 3.   Posterior probabilities of parameters in the simulation study. a) and b) Faceted by true parameter values of μ and β, respectively. Fifty data sets were simulated for each combination of $\beta \in \{0, 1, 4\}$ and $\mu \in \{0.1, 0.5, 1\}$, while $\lambda_{01} = 0.05$, $\lambda_{10} = 0.5$, $\lambda_{12} = 0.25$, and $\lambda_{21} = 0.2$ were held constant. Boxplots show the distribution of mean estimated parameter values for each of the 50 data sets for each parameter combination. Medians are represented by thick black lines, boxes are delimited by first and third quartiles, black vertical lines extend at most $1.5 \times$ IQR (i.e., the interquartile range), and dots show outliers. Red horizontal lines mark the true parameter value used in the simulations.

assigning hosts to modules based on extant butterfly–plant interactions (Supplementary Fig. S2 available on Dryad). Modules are groups of plants and butterflies that interact more with each other than with other taxa, thus host plants are assigned to the same module when they are used by the same butterflies. To identify the modules, we used a simulated annealing algorithm that maximizes the index of modularity. Specifically, we used Newman and Girvan's metric (Newman and Girvan 2004) modified for bipartite networks (Barber 2007) as implemented in the software MODULAR (Marquitti et al. 2014).

Software configuration   All analyses were performed in RevBayes (Höhna et al. 2016). For the simulated data, we ran two independent MCMC analyses for 10^5 cycles, sampling parameters and node histories every 50 cycles, and discarding the first 10^4 cycles as burnin. For the empirical data, we ran three independent MCMC analyses, each set to run for 5×10^5 cycles, sampling every 50 cycles, and discarding the first 10^5 cycles as burnin. To verify that MCMC analyses converged to the same posterior distribution, we applied the Gelman diagnostic (Gelman and Rubin 1992) provided through the R package *coda* (Plummer et al. 2006), and used a threshold of 1.1. For both simulated and empirical data sets, we used the following priors: $\beta \sim$ Exponential(1), $\mu \sim$ Exponential(10), and $\lambda \sim$ Dirichlet(1, 1, 1, 1). A RevBayes tutorial for the empirical analysis is available at https://revbayes.github.io/tutorials/host_rep/host_rep.html. Analysis scripts and data files are available in the Supplementary material available on Dryad.
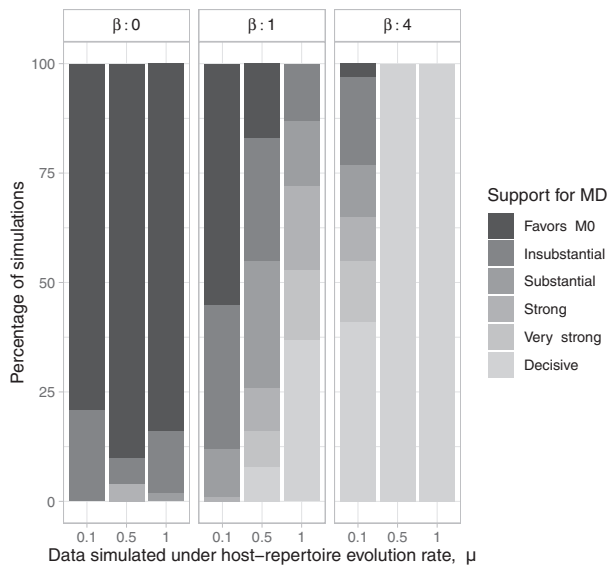
FIGURE 4.   Distribution of Bayes factors for the simulation study. Each column corresponds to the strength of support per 2×50 MCMC analyses.

RESULTS

Simulation Study

Posterior distributions of parameter values for the 9×100 MCMC analyses are shown in Figure 3. Overall, the model was able to accurately recover the true simulation parameters. When estimating β, accuracy was lowest when $\beta = 0$, possibly because the density of the prior distribution of β is highest close to zero. Estimating $\beta = 4$ was most difficult when μ was small, that is, when η produces the smallest host gain rate multipliers. Also, accuracy was lower when estimating the highest rate of host repertoire evolution, $\mu = 1$, which is likely due to higher rates of evolution generating more uncertainty in ancestral state estimates.

We performed model selection based on Bayes factors (Fig. 4). Considering that the prior distribution is $\beta \sim$ Exp(1), a high density for $\beta = 0$ under M_D is necessary to result in a Bayes factor < 1 and thus selection of M_0. For simulations with $\beta = 0$, the correct model, M_0, was selected in the majority of simulations. When $\beta = 1$, Bayes factors correctly selected M_D in many cases, but strong support for M_D was only achieved in simulations with high μ. Finally, the large majority of simulations with $\beta = 4$ supported M_D.

Figure 5 shows that accuracy in ancestral state estimation was affected by both β and μ, increasing with phylogenetic conservatism on both the butterfly (low μ) and the plant phylogenies (high β). Overall, posterior probability was higher on the estimation of actual hosts (State 2) than potential hosts (State 1), but for both, the observed accuracy was higher than the accuracy expected by chance across all parameter combinations.

Empirical Study

The estimated mean rate of host repertoire evolution for Nymphalini was $\mu = 0.86$, the mean phylogenetic-distance parameter value was $\beta = 1.31$, and the mean gain/loss rates were $\lambda_{01} = 0.014$, $\lambda_{10} = 0.71$, $\lambda_{12} = 0.13$, and $\lambda_{21} = 0.14$ (Fig. 6, blue). Our method recovered similar parameter estimates for the empirical data set when omitting the intermediate state at the tips—that is, recoding all potential hosts (State 1) as nonhosts (State 0): $\mu = 1.25$, $\beta = 1.32$, $\lambda_{01} = 0.007$, $\lambda_{10} = 0.73$, $\lambda_{12} = 0.17$, and $\lambda_{21} = 0.096$ (Fig. 6, orange). Although the μ and λ estimates differ somewhat between the analyses, both analyses produced similar mean total numbers of events per unit time (1.35 and 1.17 events/Myr for the observed and recoded data sets, respectively). We note, however, that the rates into State 1 were underestimated, and the rates out of State 1 were overestimated when 1s were removed from the data set. For both the full data set and the recoded data set, Bayes factors (BF) selected the dependence model, M_D (BF > 100). For comparison, when using the time-calibrated tree for host plants, the independence model, M_0 was selected (BF < 1) for both full and recoded data sets.

Finally, we reconstructed the fundamental and realized host repertoires at internal nodes of the Nymphalini phylogeny based on the sampled histories during MCMC. Coevolutionary histories inferred using the full and recoded data sets were very similar, as well as when using the branch-length-transformed (cladogenetic) host phylogeny and when using the time-calibrated (anagenetic) host phylogeny to measure phylogenetic distance. Thus, we only display the ancestral states inferred from the three-state data set and the cladogenetic host phylogeny (all branch lengths = 1) in the results (Fig. 7). Supplementary Figure S5 available on Dryad shows the posterior distribution of ancestral states for all the four model variants investigated in this study. To facilitate visualization of the ancestral state reconstruction, we grouped the 16 parasitized host families into five modules, as identified by the simulated annealing algorithm (Supplementary Fig. S2 available on Dryad). Eleven families (representing three modules) were inferred to be used by ancestral Nymphalini species with probability > 0.9.

We found strong support for the interaction between the ancestor of all Nymphalini butterflies and Urticaceae hosts. All other host families have been colonized in the last 10 Myr, after the divergence of the two largest clades within Nymphalini, *Vanessa*, and *Nymphalis + Polygonia*. Most species within *Vanessa*, both extant and ancestral, are specialists on Urticaceae. *Vanessa virginiensis* and *Vanessa cardui* are the only extant species that use more than two host families, and these hosts have likely been colonized by their most recent common ancestor (node 38 in Fig. 8). On the other hand, the variation in host use in the *Nymphalis + Polygonia* clade seems to be the result of host colonizations by multiple species along the diversification of the clade. For example, in Figure 8 we can see the colonization of new hosts by the ancestor of
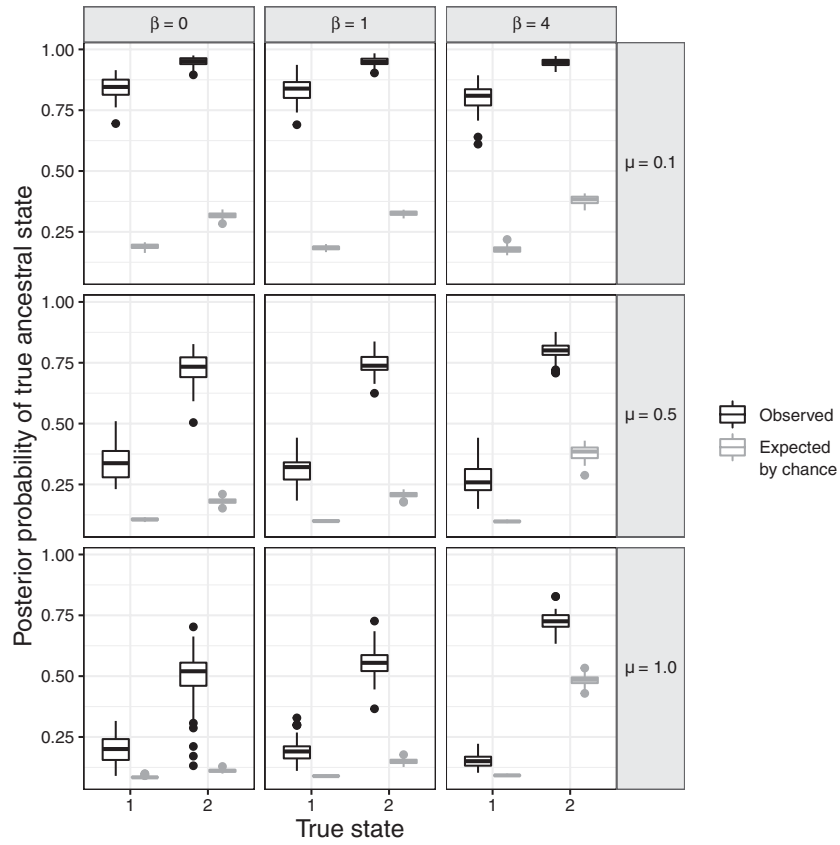
FIGURE 5.   Posterior probability of true ancestral states in the simulation study. Each boxplot shows the distribution of average posterior probability for the true state across all interactions between hosts and internal nodes in the butterfly tree for each batch of 50 simulated data sets. Each panel shows the observed and expected probabilities for States 1 and 2, for each combination of μ and β.
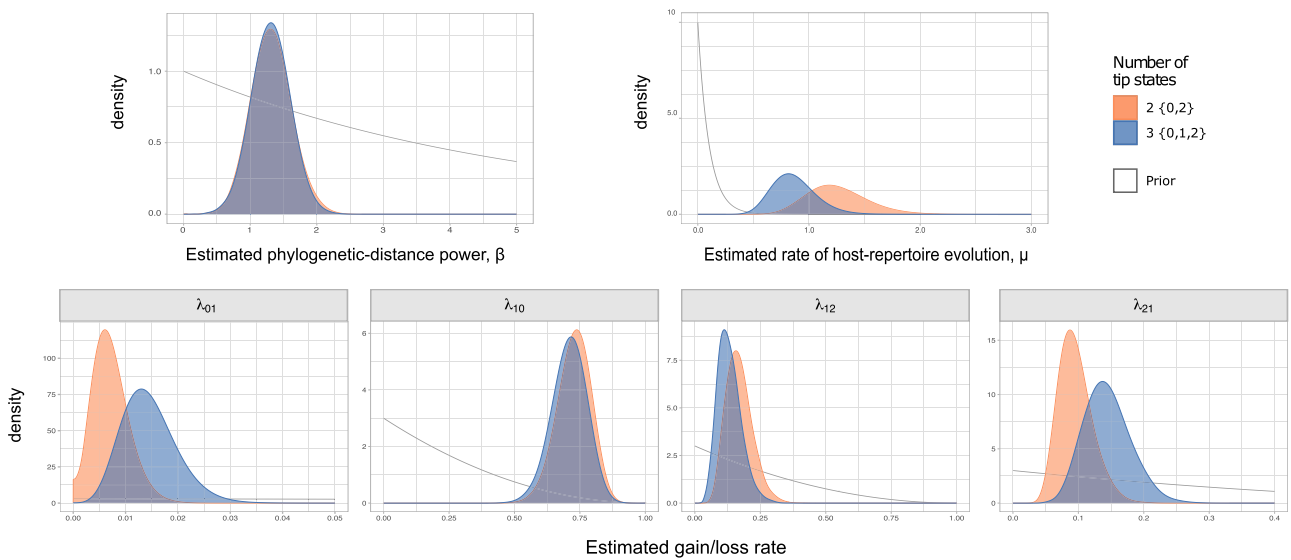

FIGURE 6.   Marginal posterior densities for parameters in the Nymphalini-Angiosperms study for both the full data set (three states at tips) and the data set omitting the intermediate state (two states at tips). Gray lines correspond to the priors $\beta \sim \text{Exponential}(1)$, $\mu \sim \text{Exponential}(10)$, and $\lambda \sim \text{Dirichlet}(1,1,1,1)$, where elements of λ follow the marginal distribution, $\lambda_{i,j} \sim \text{Beta}(1,3)$.
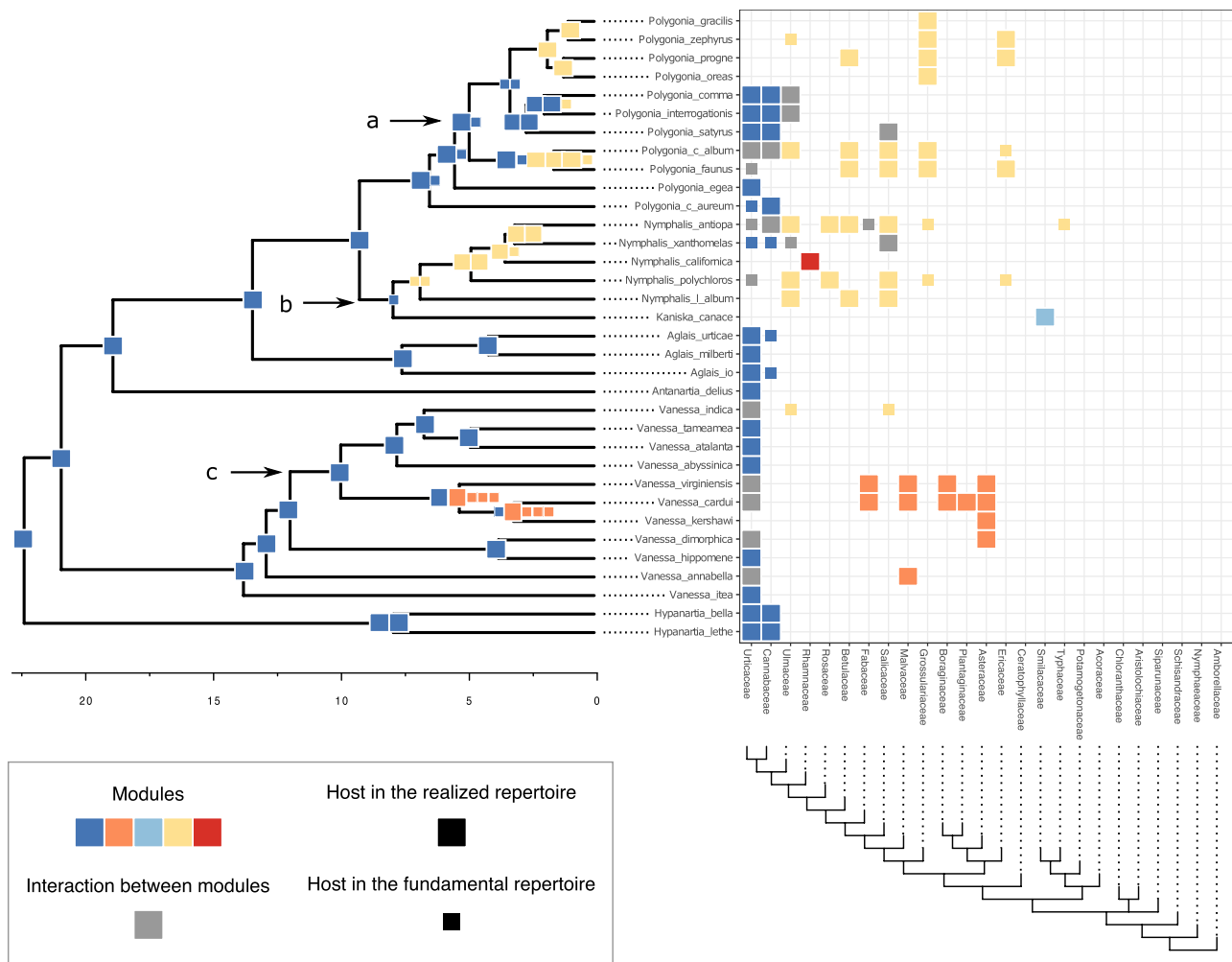
FIGURE 7. Evolution of butterfly–plant interactions through time. Ancestral state estimates (left) of host repertoire across the Nymphalini phylogeny are shown for interactions with more than 90% posterior probability. The *x*-axis under shows time before present in millions of years. Extant species interactions (right) between Nymphalini and their host plants are presented as a raster, where each square represents one interaction between a butterfly species and a host family. Colors represent different modules, that is, groups of plants that are often hosts to the same butterflies at present time. Square size was used to differentiate between actual and potential hosts. Arrows indicate nodes shown in Figure 8.

Polygonia c-album and *Polygonia faunus* (node 53) as well as strong specialization on a new host by *Kaniska canace* and a less pronounced shift in host preference by its sister species (node 51).

DISCUSSION

The method we develop here to infer the evolutionary history of ecological interactions has many advantages over previous approaches. First, it is based on stochastic models and on established principles of statistical inference, which means that it provides a robust framework for characterizing the evolutionary processes that shape ecological interaction networks and for selecting among alternative coevolutionary models. Second, our model introduces the concept of a host repertoire as the evolving trait, which we think is an important step forward. Besides accounting for the

possibility of parasites having more than one host over time scales of macroevolutionary significance, we can now directly infer the influence of host relatedness on the process of gaining new hosts. Third, the stochastic model of host–parasite coevolution that we introduce here is, to our knowledge, the first that explicitly accounts for evolution of the fundamental host repertoire. By recognizing the fact that a parasite may have potential hosts in addition to its actual hosts, and that the set of potential hosts may persist over time, the dynamic of the model changes. What would otherwise have appeared as remarkable repeated patterns of colonization of the same host lineages can now be explained as the effect of frequent transitions between potential and actual hosts in an otherwise conserved host repertoire.

Our model can readily be extended in many interesting ways. The version we present here accounts
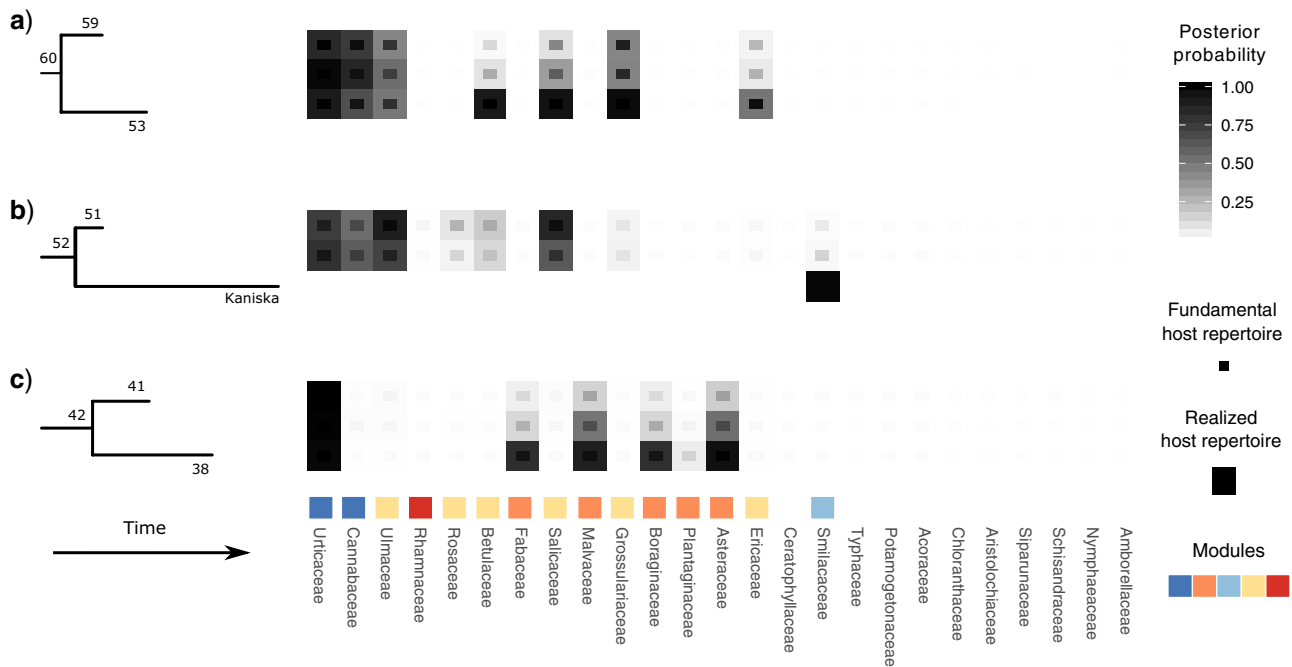
FIGURE 8.   Host repertoires at selected nodes of the Nymphalini tree (arrows in Fig. 7). Numbers indicate the node index (compatible with Supplementary Fig. S3 available on Dryad). For the only terminal taxa depicted, *Kaniska canace*, the observed host repertoire is shown. For all other repertoires, the posterior probabilities for hosts being in the fundamental (States 1 or 2) and the realized (State 2) host repertoire are shown. Colors indicate module affiliation of each host family and are consistent with Figure 7.

for the effect of host phylogeny by allowing the rate of host gain to depend on host relatedness. For simplicity, we assumed that the number of available hosts and host relatedness remain constant over geological time. This would be appropriate for a group of parasites that radiated after the relevant host lineages had been formed, which is arguably the case for the empirical example we chose. However, it should be relatively straightforward to extend our framework to account for more complex dependencies on host phylogeny. For instance, the host configurations could be modeled as changing over time, reflecting host cladogenesis.

Another interesting direction for future research would be to modify the particular ways in which hosts and parasites coevolve. We note, for example, that Figure 7 shows that host repertoires of *Vanessa* species overlap very little with the host repertoires of *Nymphalis* + *Polygonia* species, but it is not immediately clear what drives this pattern. One could design a model that allows the rates of host gain and loss to be influenced by evolving host traits—like secondary metabolites, growth form, or phenology, to mention a few examples relevant for insect–host plant interactions—in addition to relatedness among hosts. Or, one might extend the model to allow closely related parasite lineages to competitively exclude one another from host usage, similar to how competing lineages might exclude one another from geographical regions (Quintero and Landis 2020). Finally, one might introduce a biogeographical component to the coevolutionary process, requiring parasites to be in sympatry with their actual hosts, while allowing parasites to be in sympatry or allopatry with their potential hosts. Statistically comparing such model

variants will help illuminate drivers of host–parasite coevolution.

A potential concern with our approach is that already the basic version of the model is fairly parameter-rich. Given the type and amount of data that we can likely collect on host–parasite interaction, is there enough statistical power to select among the models of interest? And is it possible to infer the model parameters of interest with a reasonable degree of accuracy?

Overall, our results are encouraging in this respect. The simulations indicate that it is possible to infer the true parameter values of the basic model regardless of the level of phylogenetic conservatism in both parasites and hosts (Fig. 3). We were also able to distinguish between models with or without host relatedness effects using Bayes factors (Fig. 4). Further studies will have to show to what extent the sensitivity of the model test can be increased by selecting appropriate priors and improving the sampling of parameter space close to the boundary condition satisfying the restricted model. One option is to relax the assumption that β is nonnegative, which would simplify the sampling of values close to β = 0. It will also be important to explore how data set sizes and tree shapes, for both hosts and parasites, influence our ability to distinguish the models when the effect of host phylogeny is small.

Importantly, the empirical analysis indicates that the method is able to model the evolution of fundamental and realized host repertoires even when the information about potential hosts is lacking. This significantly increases the applicability of our method, as information about fundamental host repertoires is missing for most host–parasite systems. Potential host data is difficult to

collect, as it requires experimental testing of a large number of potential host–parasite pairs. A possible improvement of our method, which we did not explore here, would be to model uncertainty in the observations of nonhosts when data on potential hosts are missing. That is, if we had no information about a host species being used by a particular parasite, we would translate that to a certain probability p of the species actually being a nonhost, and a complementary probability $1-p$ of it being a potential host (Kuhner and McGill 2014). Modeling this observational uncertainty could help reduce the bias in parameter estimates that we observed when data on potential hosts were missing and all 0-valued states in the data set were inappropriately treated as true nonhosts. This extension would also allow us to make predictions about host use abilities in extant parasites that could then inform experiments that aim to characterize fundamental host repertoires.

We demonstrated the empirical application of our approach with a Bayesian inference of the coevolutionary history between 34 Nymphalini butterflies and 25 angiosperm families. We estimated the rate of host repertoire evolution along each branch of the butterfly tree as being 1.35 events per million years. Bayes factors favored the dependence model, M_D, where the probability of gaining a given hosts is affected by the phylogenetic distance between hosts, thus highlighting the importance of accounting for shared evolutionary history among host lineages. It is also interesting to note that this effect was not large enough to support M_D when phylogenetic distances between host plants were proportional to divergence time (anagenetic distances). As expected, β tends to be lower and harder to infer in star-like host phylogenies, whose divergence times are clustered near its root node. Perhaps of greater interest, because the estimates of β were lower when estimated using anagenetic, rather than the cladogenetic, host distances, the result suggests those traits that determine how easily a plant family is colonized by Nymphalini originated close in time to the origins of the plant families themselves.

Estimates of gain and loss rates were not symmetric, and the rates also varied between states. According to our results, gain of the ability to use a host, λ_{01}, is very rare (0.8–2% of overall rate), whereas loss is common (59–83% of overall rate). On the other hand, transition rates between states 1 and 2 were more symmetric (λ_{12} between 5% and 22%; λ_{21} between 8% and 21% of overall rate). These rate estimates support the idea that the use of the same host lineage by multiple, phylogenetically widespread butterfly lineages is more likely explained by recolonization of hosts that have been used in the past (recurrence homoplasy), that is, by transitions between actual and potential hosts, rather than by completely independent colonizations of the same host (Janz et al. 2001). Note that alternative scenarios that have been proposed in the literature to explain the evolution of Nymphalini host plant preferences, for instance by involving narrow ancestral host plant ranges and repeated independent colonization events (Hardy and Otto, 2014, e.g.), are also allowed by our model, but they are inferred to be much less likely than the conservative host repertoire scenario. Yet, because the potential host state is exited at the highest rate, the rate estimates also suggest that parasites do not retain their potential host relationships for prolonged periods of time. The moderate rates of transitions between potential and actual host states and the high departure rate from the potential host state together help explain why phylogenetic "pulses" of recurrent host acquisition manifest in some lineages but not others.

For example, the use of Grossulariaceae by two nonsister clades within *Polygonia* is best explained by a scenario where Grossulariaceae was a potential host for the ancestral species (node 60 in Fig. 8) and was subsequently gained as an actual host twice (at nodes 53 and 58, Fig. S5 available on Dryad). The ability to use Salicaceae host plants seems to be even older. Salicaceae was likely a potential host for the ancestor of *Nymphalis + Polygonia* and later became an actual host in three different parts of the clade. If potential hosts were not explicitly modeled here, these transitions would look like three independent colonizations of a plant group that is very distant from the ancestral host (Salicaceae and Urticaceae diverged about 90 Ma). Instead, we could show that what might appear as big and sudden host shifts are in fact the result of retention of ancestral host use abilities.

Understanding how ecological interactions change is crucial if we want to predict both short- and long-term consequences of global mixing of biota (Hoberg and Brooks 2015). Host–parasite interactions are of particular interest given the risk of emerging diseases, which can affect human populations directly and indirectly through their effects on crop species and wildlife (Brooks et al. 2014). Our method was designed to quantify changes in host–parasite interactions by modeling the process of gaining and losing hosts, thus allowing us to make predictions based on host–parasite history. We hope our approach will not only generate deeper insights into the evolutionary dynamics of host–parasite interactions but also help humankind mitigate some of the risks incurred by current environmental change.

REFERENCES

Acha P. N., Szyfres B. 2003. Zoonoses and communicable diseases common to man and animals, vol. 580. Washington, DC: Pan American Health Organization.

Barber M.J. 2007. Modularity and community detection in bipartite networks. Phys. Rev. E 76:066102.

Braga M.P., Guimarães Jr P.R., Wheat C.W., Nylin S., Janz N. 2018. Unifying host-associated diversification processes using butterfly-plant networks. Nat. Commun. 9:5155.

Brooks D.R. 1979. Testing the context and extent of host-parasite coevolution. Syst. Biol. 28:299–307.

Brooks D.R., Hoberg E.P., Boeger W.A., Gardner S.L., Galbreath K.E., Herczeg D., Mejia-Madrid H.H., Racz S.E., Dursahinhan A.T. 2014. Finding them before they find us: informatics, parasites, and environments in accelerating climate change. Comp. Parasitol. 81:155–164.

Calatayud J., Hórreo J.L., Madrigal-González J., Migeon A., Rodríguez M.Á., Magalhães S., Hortal J. 2016. Geography and major host evolutionary transitions shape the resource use of plant parasites. Proc. Natl. Acad. Sci. USA 113:201608381–9845.

Chazot N., Condamine F., Dudas G., Pena C., Matos-Maravi P., Freitas A.V., Willmott K.R., Elias M., Warren A., Aduse-Poku K., Lohman D.J., Penz C.M., DeVries P., Kodandaramaiah U., Fric Z.F., Nylin S., Muller C., Wheat W.C., Kawahara A.Y., Silva-Brandao K.L., Lamas G., Zubek A., Ortiz-Acevedo E., Vila R., Vane-Wright R.I., Mullen S.P., Jiggins C.D., Slamova I., Wahlberg N. 2020. The latitudinal diversity gradient in brush-footed butterflies (Nymphalidae): conserved ancestral tropical niche but different continental histories. *BioRxiv*.

Chen S.X. 2000. Probability density function estimation using gamma kernels. Ann. Inst. Stat. Math. 52:471–480.

Combes C. 2001. Parasitism: the ecology and evolution of intimate interactions. Chicago:University of Chicago Press.

de Vienne D.M., Refregier G., Lopez-Villavicencio M., Tellier A., Hood M.E., Giraud T. 2013. Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. New Phytol. 198:347–385.

Elton C. 1946. Competition and the structure of ecological communities. J. Anim. Ecol. 15:54–68.

Fisher M.C., Garner T.W., Walker S.F. 2009. Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. Annu. Rev. Microbiol. 63:291–310.

Forister M.L., Novotny V., Panorska A.K., Baje L., Basset Y., Butterill P.T., Cizek L., Coley P.D., Dem F., Diniz I.R., Drozd P., Fox M., Glassmire A.E., Hazen R., Hrcek J., Jahner J.P., Kaman O., Kozubowski T.J., Kursar T.A., Lewis O.T., Lill J., Marquis R.J., Miller S.E., Morais H.C., Murakami M., Nickel H., Pardikes N.A., Ricklefs R.E., Singer M.S., Smilanich A.M., Stireman J.O., Villamarín-Cortez S., Vodka S., Volf M., Wagner D.L., Walla T., Weiblen G.D., Dyer L.A. 2015. The global distribution of diet breadth in insect herbivores. Proc. Natl. Acad. Sci. USA 112:442–447.

Gelman, A., Rubin D.B. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7:457–472.

Hahn B.H., Shaw G.M., De Cock K.M., Sharp P.M. 2000. Aids as a zoonosis: scientific and public health implications. Science 287:607–614.

Hardy N.B. 2017. Do plant-eating insect lineages pass through phases of host-use generalism during speciation and host switching? Phylogenetic evidence. Evolution 71:2100–2109.

Hardy N.B., Otto S.P. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. Proc. R. Soc. B 281: 20132960.

Hastings W.K. 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

Heck D.W. 2019. A caveat on the Savage–Dickey density ratio: the case of computing Bayes factors for regression parameters. Br. J. Math. Stat. Psychol. 72:316–333.

Hoberg E.P., Brooks D.R. 2015. Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. Philos. Trans. R. Soc. Lond. Ser. B, 370:–20130553.

Höhna S., Landis M.J., Heath T.A., Boussau B., Lartillot N., Moore B.R., Huelsenbeck J.P., Ronquist F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. Syst. Biol. 65:726–736.

Huelsenbeck J.P., Rannala B., Yang Z.H. 1997. Statistical tests of host-parasite cospeciation. Evolution 51:410–419.

Janz N., Braga M.P., Wahlberg N., Nylin S. 2016. On oscillations and flutterings—a reply to Hamm and Fordyce. Evolution 70:1150–1155.

Janz N., Nyblom K., Nylin S. 2001. Evolutionary dynamics of host-plant specialization: a case study of the tribe Nymphalini. Evolution 55:783–796.

Janz N., Nylin S. 2008. The oscillation hypothesis of host-plant range and speciation. In: Tilmon K.J., editor. Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects. California: California University Press, p. 203–215.

Janzen D.H. 1968. Host plants as islands in evolutionary and contemporary time. Am. Nat. 102:592–595.

Jeffreys H. 1961. The Theory of Probability. Oxford: OUP Oxford.

Klassen G.J. 1992. Coevolution: a history of the macroevolutionary approach to studying host-parasite associations. J. Parasitol. 78:573.

Kuhner M.K., McGill J. 2014. Correcting for sequencing error in maximum likelihood phylogeny inference. G3: Genes, Genomes, Genet. 4:2545–2552.

Landis M.J., Matzke N.J., Moore B.R., Huelsenbeck J.P. 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62:789–804.

Larose C., Rasmann S., Schwander T. 2019. Evolutionary dynamics of specialisation in herbivorous stick insects. Ecol. Lett. 22:354–364.

Legendre P., Desdevises Y., Bazin E. 2002. A statistical test for host–parasite coevolution. Syst. Biol. 51:217–234.

MacArthur R.H., Wilson E.O. 1967. The theory of island biogeography, vol. 1. Princeton:Princeton University Press.

Magallón S., Gómez-Acevedo S., Sánchez-Reyes L.L., Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. 207:437–453.

Marin J.-M., Robert C.P. 2010. On resolving the Savage–Dickey paradox. Electron. J. Stat. 4:643–654.

Marquitti F. M.D., Guimarães P.R., Pires M.M., Bittencourt L.F. 2014. MODULAR: software for the autonomous computation of modularity in large network sets. Ecography 37:221–224.

Newman M. E.J., Girvan M. 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69:026113.

Nielsen R. 2002. Mapping mutations on phylogenies. Syst. Biol. 51:729–739.

Nosil P. 2002. Transition rates between specialization and generalization in phytophagous insects. Evolution 56:1701–1706.

Nylin S., Agosta S., Bensch S., Boeger W.A., Braga M.P., Brooks D.R., Forister M.L., Hambäck P.A., Hoberg E.P., Nyman T., Schäpers A., Stigall A.L., Wheat C.W., Österling M., Janz N. 2018. Embracing colonizations: a new paradigm for species association dynamics. Trends Ecol. Evol. 33:4–14.

Nylin S., Slove J., Janz N. 2014. Host plant utilization, host range oscillations and diversification in nymphalid butterflies: a phylogenetic investigation. Evolution 68:105–124.

Plummer M., Best N., Cowles K., Vines K. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

Poulin R. 2010. Decay of similarity with host phylogenetic distance in parasite faunas. Parasitology 137:733–741.

Quintero I., Landis M.J. 2020. Interdependent phenotypic and biogeographic evolution driven by biotic interactions. Syst. Biol. doi: 10.1093/sysbio/syz082.

Ree R.H., Moore B.R., Webb C.O., Donoghue M.J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59:2299–2311.

Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57:4–14.

Robinson D.M. 2003. Protein evolution with dependence among codons due to tertiary structure. Mol. Biol. Evol. 20:1692–1704.

Ronquist F. 2003. Parsimony analysis of coevolving species associations. In: Page R. D. M., editor. Tangled trees: phylogeny, cospeciation and coevolution. Chicago: University of Chicago Press, p. 22–64.

Subbarao K., Klimov A., Katz J., Regnery H., Lim W., Hall H., Perdue M., Swayne D., Bender C., Huang J., Hemphill M., Rowe T., Shaw M., Xu X., Fukuda K., Cox N. 1998. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science 279:393–396.

Suchard M.A., Weiss R.E., Sinsheimer J.S. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. 18:1001–1013.

Thompson J. 1994. The coevolutionary process. Chicago: University of Chicago Press.

Verdinelli I., Wasserman L. 1995. Computing Bayes factors using a generalization of the Savage–Dickey density ratio. J. Am. Stat. Assoc. 90:614–618.