

Widespread Historical Contingency in Influenza Viruses

Jean Claude Nshogozabahizi,^{*1} Jonathan Dench,^{*1} and Stéphane Aris-Brosou^{*1,2}

^{*}Department of Biology, and [†]Department of Mathematics and Statistics, University of Ottawa, Ontario K1N 6N5, Canada

ORCID IDs: 0000-0001-6337-8949 (J.D.); 0000-0003-4987-0296 (S.A.B.)

ABSTRACT In systems biology and genomics, epistasis characterizes the impact that a substitution at a particular location in a genome can have on a substitution at another location. This phenomenon is often implicated in the evolution of drug resistance or to explain why particular “disease-causing” mutations do not have the same outcome in all individuals. Hence, uncovering these mutations and their locations in a genome is a central question in biology. However, epistasis is notoriously difficult to uncover, especially in fast-evolving organisms. Here, we present a novel statistical approach that relies on a model developed in ecology and that we adapt to analyze genetic data in fast-evolving systems such as the influenza A virus. We validate the approach using a two-pronged strategy: extensive simulations demonstrate a low-to-moderate sensitivity with excellent specificity and precision, while analyses of experimentally validated data recover known interactions, including in a eukaryotic system. We further evaluate the ability of our approach to detect correlated evolution during antigenic shifts or at the emergence of drug resistance. We show that in all cases, correlated evolution is prevalent in influenza A viruses, involving many pairs of sites linked together in chains; a hallmark of historical contingency. Strikingly, interacting sites are separated by large physical distances, which entails either long-range conformational changes or functional tradeoffs, for which we find support with the emergence of drug resistance. Our work paves a new way for the unbiased detection of epistasis in a wide range of organisms by performing whole-genome scans.

KEYWORDS correlated evolution; epistasis; networks; influenza

ONE of the most fundamental questions in biology concerns the emergence of new structures and new functions, in particular at the molecular and genetic level (Lynch 2007). As such, a large body of experimental work has accumulated over the past decade to unravel the mutational history at the origin of simple phenotypes. For instance, one particular bacterial drug resistance is conferred by five mutations, but out of the $5! = 120$ possible ways in which these mutations can accumulate, only a handful of mutational trajectories are evolutionarily accessible (Weinreich *et al.* 2006). This line of work suggests that some mutations are required for subsequent mutations to occur. Further work demonstrates that such permissive mutations are not limited to

bacteria, as they are also found in vertebrate (Ortlund *et al.* 2007), yeasts (Sorrells *et al.* 2015), and viral systems (Gong *et al.* 2013). However, while such chains of dependent or conditional substitutions—called historical contingency (Mohrig *et al.* 1995; Harms and Thornton 2014)—are expected to lead to mutational trajectories, their shape and ramifications are not completely elucidated.

As the experimental determination of these trajectories can be tedious, taking >20 years in the case of the long-term evolution experiment (Blount *et al.* 2008), computational solutions were sought to reconstruct historical contingencies and the mutational correlations they imply. Initial solutions relied on protein sequence alignments to compute site-wise vectors of amino acid (AA) frequencies, from which pairs of coevolving residues could be identified (Neher 1994; Taylor and Hatrick 1994). While this general approach is still used in statistical physics to predict protein folds (Shindyalov *et al.* 1994; Sutto *et al.* 2015), numerous refinements were brought either through the use of metrics such as mutual information (Korber *et al.* 1993; Atchley *et al.* 2000; Gloer *et al.* 2005) or by correcting for shared evolutionary history.

Copyright © 2017 by the Genetics Society of America
doi: 10.1534/genetics.116.193979

Manuscript received July 18, 2016; accepted for publication November 4, 2016; published Early Online November 9, 2016.

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.193979/-/DC1.

¹These authors contributed equally to this work.

²Corresponding author: 30 Marie-Curie Private, Department of Biology, University of Ottawa, Ottawa, ON K1N 6N5, Canada. E-mail: sarisbro@uottawa.ca

One of the first methods to detect correlated evolution while accounting for phylogeny was given in the general context of the evolution of discrete morphological characters (Pagel 1994). Further leveraging on Schöniger and von Haeseler (1994), the method was quickly extended to analyze RNA molecules by modeling dinucleotides (Muse 1995; Rzhetsky 1995) and to map the correlated residues hence detected onto a three-dimensional protein structure (Pollock *et al.* 1999; Poon *et al.* 2007a,b). More recently, a full evolutionary model was proposed to detect epistatic sites in a Bayesian framework (Nasrallah and Huelsenbeck 2013). However, such approaches rely on complex models and Bayesian computing tools may scale up poorly with increasingly large data sets (Poon *et al.* 2008; Aris-Brosou and Rodrigue 2012), and are also generally geared toward detecting positive epistasis (when two mutations increase fitness values).

Here we build on these developments to describe a novel, yet intuitive, statistical method for detecting correlated evolution among pairs of AAs in the typically fast-evolving influenza A virus (Worobey *et al.* 2014). Our method takes inspiration from an approach developed in ecology that is aimed at detecting correlated evolution among phenotypic traits (Pagel 1994; Pagel and Meade 2006). We validate our approach using a two-pronged procedure based on both extensive simulations and analyses of experimentally validated data sets (both in viral and eukaryotic systems). The analysis of several large data sets leads us to reconsider the nature of epistasis in influenza viruses. We find evidence that interacting AAs form networks of sites undergoing substitutions that are most likely to be permissive (*i.e.*, contingent), as they occur in a temporal sequence. These networks cover large physical distances among interacting AAs, suggesting long-range structural and/or functional effects.

Materials and Methods

General approach to detect epistasis

Repurposing of BayesTraits: To model correlated evolution at the molecular level, we employed the maximum likelihood model implemented in BayesTraits (Pagel 1994; Pagel and Meade 2006). See Talavera *et al.* (2015) for a similar model. This model was originally developed as a time-homogeneous Markov process with discrete states in continuous time to investigate the coevolution of discrete binary traits on phylogenetic trees. This is achieved by testing whether a dependent model of trait evolution fits the data better than an independent model. The dependent model allows two traits to coevolve as the rate of change at one trait depends on the state at the other trait, while the independent model does not place any restriction on rates of change (Pagel 1994) (see Supplemental Material, Figure S1 in File S1). In both cases, the likelihood function is optimized by summing (integrating) over all unobserved pairs of character states at internal nodes. As these two models are nested, a likelihood ratio test can be employed for model selection. The test statistic, twice

the log-likelihood difference, is assumed to follow a χ^2 distribution with four degrees of freedom under the null hypothesis (independence).

This general framework further assumes a phylogenetic tree with a known topology and branch lengths proportional to the amount of evolution separating each node. Both assumptions can be addressed as described below, either by a bootstrap analysis, or by resorting to trees sampled from their posterior distribution.

Data recoding: The model described above was implemented for binary traits. Here, however, our goal is to analyze the coevolution of pairs of sites (DNA or AA) along a sequence alignment. In the case of proteins, each site has 20 possible AAs, which represent the states of our system. To avoid resorting to tensor kernels, data recoding is therefore necessary to reconcile the data with the approach. In this context, we evaluated two strategies. First, AAs were partitioned according to their physicochemical properties. Binary properties (side-chain group types) were naturally recoded “0”/“1.” For those with $k > 2$ states, we compared each of the k_i states against the other ones (k_j where $i \neq j$). For instance, in the case of charge, we first assigned state “0” to negative and state “1” to nonnegative AAs, and circled through the two other states (Figure S2 in File S1). This recoding was based on the physicochemical properties as implemented in the R package *protr* version 0.2-1 (Xiao *et al.* 2014). However, results from this first method failed to find any significantly correlated pairs of mutations when applied to our first data set. Hence, a second recoding method was devised and used thereafter.

In our second recoding method, AAs were classified as either being in the outgroup or ingroup consensus state: at each position of an alignment, the outgroup state was defined as the majority-rule consensus AA present in a clade used to root the tree. Here, this rooting clade was defined either as the one containing the oldest sequences, and this clade was removed for downstream analysis, or as the basal clade in a relaxed molecular clock analysis (see below). The ingroup state was then defined as any AA that differs from the consensus AA in the outgroup. Within each column of the alignment, sites sharing the outgroup state were all recoded as “0,” while those sharing the ingroup state were recoded as “1.”

Code optimization: To discover pairs of AAs that are potentially interacting, we need to run the above model on all $[n(n-1)/2]$ pairs of sites in an alignment of length n . This computation can be prohibitively long even with an alignment of modest size. For instance, the influenza H3N2 nucleoprotein (NP) has 498 AAs, which leads to analyzing 123,753 pairs of sites under both the dependent and the independent model (*i.e.*, 247,506 models need to be run). To decrease the computational cost, we first compressed the alignment into site patterns (Yang 2006, p. 105). Then, as pairs of sites can be independently compared, we parallelized the code using R's *foreach* version 1.4.2 and *doMC* (version 1.3.3)

(Revolution Analytics and Weston 2013) packages to take advantage of multicore/multiprocessor architectures (Figure S3 in File S1). To account for multiple testing, we computed the false discovery rate (FDR) according to the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995). We used R (version 3.0.2) for all analyses (R Development Core Team 2013). In the analyses presented below, the pairs of AAs identified to be interacting were subsequently mapped on three-dimensional protein models predicted by homology modeling with SWISS-MODEL (Biasini *et al.* 2014), and plotted using KiNG (Chen *et al.* 2009).

Validation based on simulations

To validate our method for detecting correlated pairs of sites, we followed two approaches: an extensive simulation study and the analysis of data sets in which epistasis was experimentally confirmed. We first present our simulation strategies that, to avoid biasing our results, were based on two different frameworks of simulating correlated evolution. File S1 presents additional simulation results.

Simulations with PHASE: First, sequences were simulated with PHASE 2.0 (Gowri-Shankar and Jow 2006), which allowed us to simulate two categories of sites: those that evolve independently and those that evolve in a correlated manner. Sites that evolve independently (main sequence of length l_i) were simulated using the general time-reversible (GTR) model of nucleotide substitution (Tavaré 1986; see also Aris-Brosou and Rodrigue 2012). Both nucleotide frequencies and transition probabilities were arbitrarily set to reflect published values for *Pseudomonas aeruginosa* (Garrity *et al.* 2004), although this particular setting has no impact on downstream analyses. With PHASE, sites that evolve in a correlated manner (epistatic sites l_e) were simulated under the RNA7D model of dinucleotide substitution (Tillier and Collins 1998). We modified the transition probabilities such that double substitutions occurred for 95% of changes, single substitutions for 5%, while mismatch substitutions were not permitted. In the results presented below, equilibrium dinucleotide frequencies were all assumed to be equal (set to 1/16). Because influenza A viruses typically exhibit dinucleotide bias (Greenbaum *et al.* 2008), which could lead to false positives, we also performed simulations under the RNA16 model with a dinucleotide bias matching that of influenza A viruses (see File S1). As total sequence length was determined ahead of time (part of our factorial design), when simulating correlated evolution, we first simulated a number l_e of pairs of epistatic sites, which were then concatenated to the nucleotide sites simulated under the independent model of length l_i to obtain a final sequence of length $l_s = 2l_e + l_i$. Where correlated evolution was not simulated $l_s = l_i$.

Simulations with Coev: Our second simulation strategy employed Coev (Dib *et al.* 2014), which produces correlated evolution by randomly defining a dinucleotide profile (two-letter nucleotide states, *e.g.*, AA, AT, CG, *etc.*) and simulating

evolution under the Jukes–Cantor substitution model (Jukes and Cantor 1969). The transition rates for each nucleotide of the dinucleotide may be set independently; however, as we have no *a priori* reason to assume either evolves with a different rate, we left these rates equal to their default setting. The strength of selection for the defined profile is governed by the ratio of the parameters d and s (*i.e.*, d/s), which represent the likelihood of a dinucleotide evolving toward ($d/s > 1$) or away ($d/s < 1$) from the defined dinucleotide profile. When $d/s = 1$, there is no selection and sites evolve independently; we varied the strength of selection from independent ($d/s = 1$) to the default setting offered by Coev-web (Dib *et al.* 2015) ($d/s = 100$), with intermediate strengths ($d/s = \{2, 33, 66\}$). PHASE was still used to simulate sites that evolved independently (l_i) to be consistent in our use of a GTR-substitution model.

Sensitivity and specificity: In both sets of simulations, we assessed the impact of branch lengths, tree shape, and number of sequences, both in the presence and absence of epistasis (Figure S4 in File S1). With PHASE, branch lengths (b) were varied across a \log_2 scale for $b \in (-12, -1)$. To reduce possible factorial combinations, all branches of a single simulated tree were the same length. Two tree shapes were used, simulated tree topologies (τ) being either symmetrically bifurcating or pectinate (Figure S4 in File S1). All simulated trees were ultrametric. Each tree contained a number of sequences (n_s) equal to 16, 32, 64, or 128. This led to a full factorial design containing 192 simulation conditions ($12b \times 2\tau \times 4n_s \times 2$ for with or without epistasis). Each simulation condition was replicated 100 times and l_s was set to 100 bp. When epistasis was simulated, the number of epistatic pairs was set to 3. An ANOVA was used to assess the significance of each of these factors. All tests were conducted at the $\alpha = 0.01$ (1%) significance threshold.

With Coev, branch lengths (b) were varied across a \log_2 scale for $b \in \{-12, -10, -8, -6 - 4\}$. Each tree contained a number of sequences (n_s) equal to 32 or 128. The same two tree shapes (τ) were used, being either symmetrically bifurcating or pectinate. This led to a full factorial design containing 100 simulation conditions ($5b \times 2\tau \times 2n_s \times 5d/s$). Each simulation condition was replicated 100 times and l_s was set to 100 bp. When epistasis was simulated, the number of epistatic pairs was set to 3, as under PHASE. All tests were here again conducted at the $\alpha = 0.01$ (1%) significance threshold unless otherwise stated.

Validations based on previous evidence

Previous computational analyses: As a first validation of our approach on actual data, we reanalyzed those studied in a previous computational study (Kryazhimskiy *et al.* 2011). The four data sets obtained from S. Kryazhimskiy (personal communication) consist of: 1219 HA and 1836 NA sequences from H1N1 viruses, as well as 2149 HA and 2339 NA sequences from H3N2 viruses. These data sets are here denoted KDBP11-H1 (HA in H1N1), KDBP11-N1 (NA in H1N1), KDBP11-H3 (HA in H3N2), and KDBP11-N2 (NA in H3N2), respectively.

Experimental evidence: As computational studies make predictions that are not always tested or validated, we reanalyzed data sets in which epistasis was experimentally confirmed. A number of recent studies reported evidence for epistasis and we present results on three of these.

First, we reanalyzed data published by Gong *et al.* (2013), which comprised 424 H3N2 human influenza A NP sequences spanning 42 years between 1968 and 2010. This gene was originally chosen because it evolves relatively slowly and is hence amenable to experimental validation, as all substitutions can be easily tested by site-directed mutagenesis (Gong *et al.* 2013). The viral sequences that they used were downloaded from the Influenza Virus Resource database (Bao *et al.* 2008). This data set is here denoted Gong13NP.

We also retrieved a second data set as analyzed by Duan *et al.* (2014), who used an alignment that contained 1366 human influenza A H1N1 neuraminidase (NA) collected between 1999 and 2009; as in their analysis, the 2009 H1N1 pandemic sequences were excluded. This data set is denoted Duan14NA.

Finally, a recent study performed a combinatorial analysis based on mutations of the eukaryotic transfer RNA-ARG (CCT) gene (tRNA^{Arg^{ccr}}) to detect epistasis (Li *et al.* 2016). We ran our computational analysis on an alignment of the eukaryotic tRNA gene constructed in a manner similar to Li and collaborators (see File S1, “Phylogenetic data of the tRNA genes”): we downloaded all available eukaryotic tRNA^{Arg^{ccr}} genes from GtRNadb (Chan and Lowe 2016), aligned the sequences with the *cmalign* tool from Infernal (Nawrocki and Eddy 2013), and kept the 75-nt region that spanned the conserved portion within *Saccharomyces sp.* As our method requires a phylogenetic tree, we used a phylogeny of eukaryotes (Hedges *et al.* 2015), keeping only those tips corresponding to the taxa in our alignment. Lastly, as the branch lengths of this tree were in millions of years, we rescaled these to expected number of substitutions using the best available eukaryotic tRNA molecular clock (Soares *et al.* 2009). While Li *et al.* (2016) did not perform a fully exhaustive analysis of all pairs of sites, and our analysis requires sites to be polymorphic, we present our comparison for those pairs of sites tested in both of the analyses. For this analysis only, we used a statistical threshold of $\alpha = 0.05$ as in Li *et al.* (2016). This data set is denoted Li2016.

Nature of correlations in influenza evolution

After performing these validations, both on simulations and on experimentally validated data, we set out to investigate the nature of these interactions by testing two hypotheses. First, we revisited the work by Koel *et al.* (2013) that experimentally validated the existence of AA substitutions involved in changes of antigenic clusters. For this, we used 877 H3N2 human influenza hemagglutinin (HA) sequences as in Koel *et al.* (2013), collected between 1968 and 2003, to test if these substitutions responsible for antigenic changes also showed evidence for correlated evolution. This data set is denoted Koel13HA.

Second, we tested if our statistical approach could detect some evidence for correlated evolution at pairs of sites involving the S31N substitution, which is responsible for conferring resistance to the antiinfluenza drug adamantane (Abed *et al.* 2005). For this, we retrieved 668 H3N2 human influenza A matrix protein 2 (M2) sequences that were collected between 1968 and 2003. This data set is denoted Adam03M2.

Phylogenetic analyses

Maximum likelihood was employed to estimate phylogenetic trees for the KDBP11-H1, KDBP11-N1, KDBP11-H3, KDBP11-N2, Gong13NP, Duan14NA, and Adam03M2 alignments with FastTree version 2.1.7 (Price *et al.* 2010) under the WAG + Γ_4 model to account for among-site rate variation. Trees were rooted using the earliest sequences, which were AAD17229/USA/1918, CAA24269/Japan/1968, AAF77036/USA/1918, ABI92283/Australia/1968, Aichi68-NPA/Aichi/2/1968, A/Victoria/JY2/1968, and A/Albany/17/1968, respectively. The R package APE was used to visualize the trees (Paradis 2006). For the Koel13HA data set, a rooted tree was reconstructed using BEAST version 1.8.0 under a relaxed molecular clock assuming an uncorrelated lognormal prior (Drummond *et al.* 2006) and a constant-size coalescent prior under the FLU + Γ_4 substitution model. Analyses were run in duplicate to check convergence, for a total of 100 million steps with a thinning of 5000; log files were combined with LogCombiner after conservatively removing the first 10% of each chain as a burn-in period, as checked with Tracer version 1.5 (<http://tree.bio.ed.ac.uk/software/tracer>).

Phylogenetic uncertainty was taken into account by running our algorithm on bootstrapped trees (or trees sampled from the posterior distribution). Because only the Gong13NP data set showed a large proportion of Shimodaira–Hasegawa-like, approximate likelihood ratio test (Anisimova and Gascuel 2006) node support values in the low (0.0, 0.8) range (Figures S5–S12 in File S1), the results of the bootstrap analyses are only shown for this case.

Data availability

All data sets and the Analysis of Epistasis and Genomic Interacting Sites (AEGIS) script used in this work are available at <http://github.com/sarisbro/AEGIS>.

Results and Discussion

Simulation studies

Excellent specificity but mediocre sensitivity: As a first means to validating our approach, we conducted a fully factorial simulation study. To avoid biasing our results, the simulation models differ from the analysis models. Our simulation results with PHASE demonstrate that alignments containing <32 sequences have a poor ability to detect epistasis, with a sensitivity generally $\leq 20\%$ (Figure S13 in File S1). For this reason, we henceforth focus on alignments of at least 32 sequences.

The specificity (S_p) of our approach is never $<99.9999\%$, which occurs for the symmetric tree shape with the largest number of sequences and rather short branch length of 2^{-6} substitutions per site (Figure 1A). Variation in specificity was so low that we report the value on a $-\log_{10}(1 - S_p)$ scale to emphasize the performance of our approach against thresholds set by the minimum, mean, and maximum number of pairwise comparisons among replicates. These thresholds represent the number of true negatives calculable in our analyses [*i.e.*, (number of pairwise comparisons – number of epistatic pairs)/number of pairwise comparisons] and thus performance above these thresholds demonstrates that high specificity results from low false positive detection rate, and not simply from a large data set with unduly large numbers of true negatives. In spite of the low variation in specificity, the number of sequences (n_s) explains most of the variance (Table S1 in File S1). While we find a two-way interaction with tree shape (τ) and branch length, there was no direct interaction between n_s and either branch length or τ (Table S1 in File S1). We find a three-way interaction between branch length, τ , and n_s , which reinforces the idea that while the number of sequences dominates the specificity of our method, the amount of evolutionary time (sum of branch lengths over topology) remains an important factor (see Figure 1A).

While specificity is excellent, we find that sensitivity is mediocre (Figure 1B), which is consistent with previous studies (Poon *et al.* 2007b). Our approach's sensitivity is a function of branch lengths and the number of sequences (Table S1 in File S1), with the existence of an optimal branch length where sensitivity could become excellent, reaching close to 100%, before degrading quickly again (on a log scale). This optimal response reflects that short branch lengths carry no information while long ones have random site patterns and thus convey no information (Yang 1998). Pectinate trees show an optimum for shorter branch lengths than symmetric trees, and an ANOVA confirms an interaction between tree shape and branch length with respect to sensitivity (Table S1 in File S1). Lastly, we note that larger alignments appear to have better sensitivity (Figure 1A) and that this does have an interaction with tree shape. We note that there exists large variance in sensitivity, due to our method of simulating epistasis, but have omitted error bars from Figure 1 for reasons of clarity. Due to the low sensitivity of our method, we further investigated the prevalence of false positives by plotting precision (Figure 1C). From this figure we find that $<1\%$ of positives are false except when branches are long.

Because dinucleotide bias, as typically found in influenza A viruses, could lead to false positives, we also assessed the impact of this parameter on the performance of our approach. We mimicked bias as found in the four KDBP11 data sets, and employed the full RNA16 model in PHASE to perform simulations similar to those under GTR as above. Our results show that both specificity and precision are unchanged, while sensitivity appears to be decreased. As we did not change the method simulating dependent sites, we did not expect

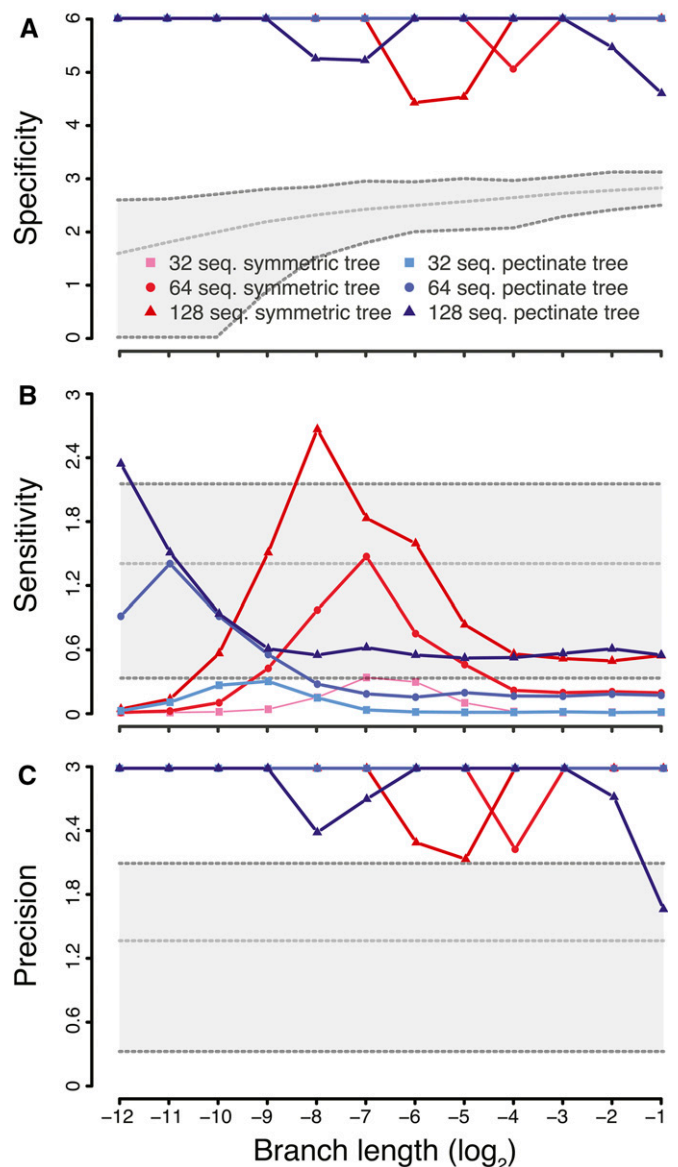


Figure 1 Specificity, sensitivity, and precision results, from simulated data, of our novel epistasis detection method. Results are shown for alignments with 32 (■), 64 (○) and 128 (▲) sequences. Tree shapes are color coded (symmetric in red; pectinate in blue). Branch lengths were varied on a log₂ scale. All y-axes show the mean of $-\log_{10}(1 - \text{summary statistic})$ to highlight the performance of our method as a summary statistic approaches 1; when this value is 1, we arbitrarily assigned it a value 10% larger than the largest finite value within the set. (A) shows specificity (true negative rate) with a gray shaded polygon which illustrates thresholds for excellent specificity in parameter space. The thresholds were established by subtracting three (the number of epistatic pairs simulated) from the minimum (bottom dashed line), mean (middle dashed line), and maximum (top dashed line) number of pairwise comparisons performed across all simulations with the same branch length. These thresholds represent the number of calculable true negatives and allow us to demonstrate our method's excellent specificity. (B and C) show sensitivity (true positive rate) and precision (positive predictive value), respectively. Each panel includes a gray shaded polygon which illustrates thresholds for 50% (bottom dashed line), 95% (middle dashed line), and 99% (top dashed line) detection. These thresholds were arbitrarily chosen to demonstrate idealized benchmarks of performance. Seq, sequence.

sensitivity to be affected. After a thorough review of our data we can only conclude that this difference results from the large variance in sensitivity resulting from the method by which epistasis is simulated. Altogether, dinucleotide bias does not lead to an increase in false positives (Figure S14 in File S1).

To further confirm that our simulation results are not biased by our method of simulating epistasis, we compared our detection method to simulated evolution generated by Coev (Dib *et al.* 2014). These results (see File S1) show that both specificity and precision are comparable with the PHASE simulations, while sensitivity is reduced (Figure S15 in File S1). These findings reflect the particularities of how epistasis was simulated and confirm the excellent specificity and precision of our method. Importantly, these properties are found even when epistasis is simulated at very weak levels (Figure S15 in File S1).

Note that unlike the real data sets analyzed in the next subsection, all these simulations are based on a state space of size four (DNA) and not 20 (AAs): the effect of this reduction of the state space in the simulations is to lead to fewer unique site patterns, and hence larger proportions of false positives in DNA data than in AA data. Our DNA-based simulations show, however, that false positives are already well controlled.

Real data analyses

Limited overlap with previous computational results: As a first evaluation of our algorithm on real data, we reanalyzed four large data sets (>2000 sequences) previously analyzed with another computational method designed to detect epistasis (Kryazhimskiy *et al.* 2011). Briefly, this method consists of mapping mutations on the phylogeny by parsimony, to then infer pairs of sites that are candidates for correlated evolution, before determining if mutations at each pair are temporally “close enough” to be considered as interacting. Unlike our approach, no data recoding is necessary; conversely, no mapping is involved in our method as the likelihood function describing changes of states at pairs of sites plays this role.

Here, the AA data were recoded as outgroup/ingroup character states to match the above simulations: at each position of the alignment, the “outgroup” state represents the majority-rule consensus AA in the outgroup sequences and the “ingroup” state represents all the types of AAs that differ from the outgroup consensus AA. Overall, our approach not only detects fewer epistatic sites than the previous method, but the two approaches also show limited overlap (Figure S16 in File S1). This marginal overlap is even found before FDR correction (Figure S17 and Tables S3–S6 in File S1), so lack of power is an unlikely explanation of the difference. Our extensive simulations suggest that this difference may be the result of the low-to-average sensitivity and of the excellent specificity/precision of our method, so that the AA pairs that we detect may actually be coevolving sites among the truly epistatic pairs. This result begs the question as to whether the few pairs of sites we detect would have any experimental evidence supporting epistasis.

Detected pairs are almost all experimentally validated:

Because the previous data sets were only examined from a computational point of view, we turned to additional data that have been experimentally validated.

First, we analyzed the Gong13NP data set of 424 influenza NP protein sequences (Gong *et al.* 2013). As this is the smallest data set in our study, we assessed two ways of recoding the data into binary character states. We first partitioned AAs according to their physicochemical properties. With this recoding strategy, four pairs of sites were detected before FDR correction (Table S7 in File S1), but none of them matching the pairs detected by Gong and collaborators. Note that after FDR, no interactions were significant (Table S7 in File S1).

The estimated phylogenetic tree for this data set shows a pectinate-like (asymmetric) shape and short branch lengths (Figure 2)—even if H3N2 viruses exhibit a more punctuated mode of evolution than H1N1 viruses (Sandie and Aris-Brosou 2014). Our simulation results show that, under these conditions, with >100 sequences, we can expect a sensitivity $\geq 80\%$ (Figure 1A), so that the Gong13NP data set fulfills all the conditions for detecting true interactions. This suggests that even if the physicochemical recoding seemed *a priori* to be a good idea, capitalizing on the chemistry of life, it wastes statistical power on multiple three-way tests (Figure S2 in File S1).

To better mimic our simulation conditions, we then recoded AAs as outgroup/ingroup states. With this recoding strategy, seven pairs of sites were detected: 259:334, 421:425, 246:470, 217:334, 217:343, 259:421, and 186:259 (Figure 2A and Table S8 in File S1). Gong *et al.* (2013) detected 259:334 as their strongest signal, as well two additional pairs (384:65 and 280:312) with a weaker signal. However, we detected six additional pairs of sites that were never shown to be epistatic. These could either be false positives—but again our simulations suggest that our approach is extremely specific (Figure 1B)—or simply coevolving pairs of sites that are not epistatic (coevolution is necessary but not sufficient for epistasis to exist), or that they are evolving under negative epistasis.

While we find multiple pairs of correlated sites, almost all these pairs are linked to the experimentally confirmed L259S and N334H substitutions (Figure 2B), hereby forming a network of interacting sites. Can we say anything about the nature of these interactions? If they were physical, we would expect that interacting sites would be in close spatial proximity on the folded protein, as in the case of compensatory mutations in RNA molecules (Kimura 1985; Chen *et al.* 1999). However, the spatial distribution of these epistatic pairs of sites on the protein structure does not conform to this prediction: while two residues are considered physically linked when their distance is $\leq 8.5 \text{ \AA}$ (Atilgan *et al.* 2004), we find that the average distance between interacting AAs is 25.9 \AA (SD = 18.1; Figure 2C). This distribution strongly suggests that chained epistasis is not linked by spatially-close physical interactions (Figure S18 in File S1), so that

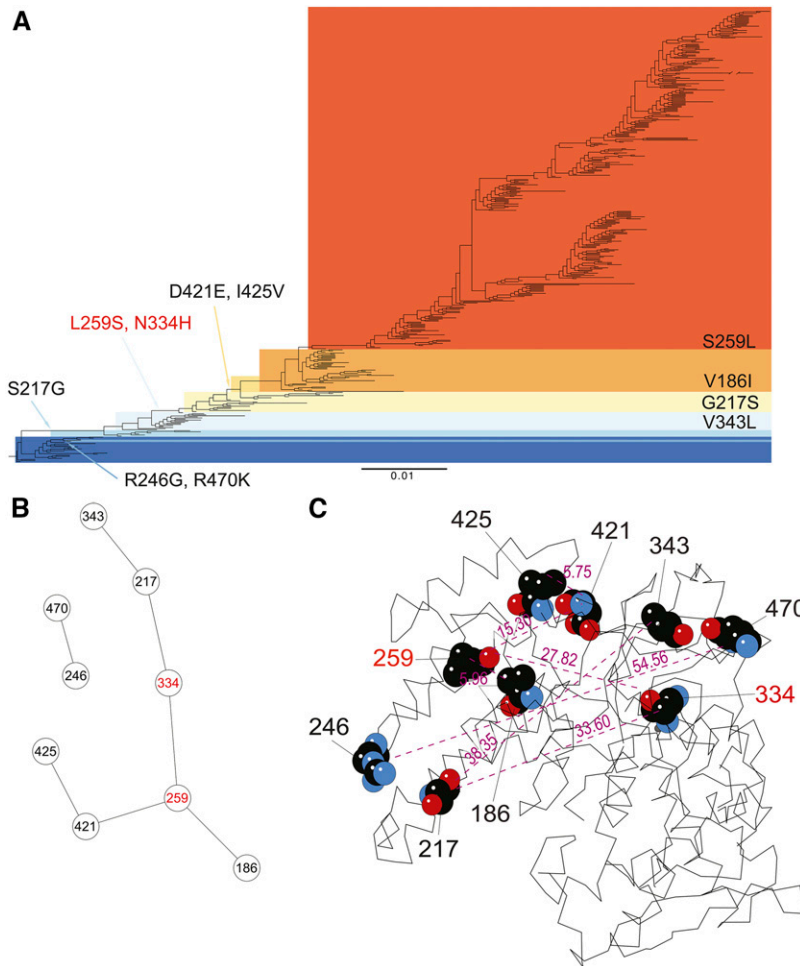


Figure 2 Epistatic pairs of AAs detected in the Gong13NP data set with the outgroup /ingroup recoding. (A) The epistatic mutations that we detected are plotted on the NP phylogenetic tree. The substitutions in red were experimentally validated (Gong *et al.* 2013). (B) Chained epistasis of interacting AAs. (C) The epistatic sites are mapped on a three-dimensional NP protein structure (based on template 3ZDP). The numbers show the AA positions experimentally validated (in red) and those detected only in this study (in black). The numbers in purple show the physical distance between epistatic sites (in Å).

thermodynamic (Thomas *et al.* 2010) or compensatory changes of the three-dimensional structure (Weinreich *et al.* 2006) act at very long spatial ranges.

To further validate our approach, we also analyzed the Duan14NA data set, comprising 1366 NA sequences of pre-pandemic H1N1 viruses (Duan *et al.* 2014). We identify the epistatic pair 275:354 (Figure 3 and Table S9 in File S1), which was experimentally proven to confer oseltamivir resistance and that dominated the population in 2008–2009 (Duan *et al.* 2014). In their study, Duan and collaborators showed that D354G was the main mutation responsible for maintaining the function of NA after alteration of enzyme activity by H275Y (H274Y in N2 numbering), so this is potentially the strongest existing interaction. However, our approach fails to identify the five other mutations (V234M, R222Q, K329E, D344N, and D354G) that were further identified by Duan and collaborators to be interacting with H275Y. The estimated H1N1 tree is more symmetrical in shape than the one estimated for the Gong13NP data and has shorter average branch lengths (see scale bar in Figure 3). Our simulation results suggest that in this case, the sensitivity of our approach can be very low (Figure 1A). This low sensitivity might explain why we fail to detect the five additional sites interacting with position 275.

In spite of this negative result, we find that the interacting pair is, again, a long-range interaction (23.7 Å; Figure 3, inset). The objective of the next two sections is to explore more systematically the nature of epistasis in influenza A viruses, focusing more specifically on (i) the prevalence of chained long-range epistasis, and (ii) the potential nature of these long-range interactions.

Lastly, we ran our approach on the Li2016 data set. While our analysis found a signal for correlated evolution in 25.7% of the pairs of sites tested by Li *et al.* (2016), 90.0% of our significant pairs of sites were identified as epistatic by Li and collaborators (Figure S19A in File S1). We also found a small number of site pairs (37) that were not detected to be epistatic by Li and collaborators. We note, however, that (i) four of these 37 site pairs could not be statistically tested as the original study had only one biological replicate, (ii) all the site pairs we identify as correlated have epistasis measures that fall within the range of those site pairs deemed significant by Li and collaborators (Figure S19B in File S1), and (iii) one site pair we identified forms a Watson–Crick base pair in the folded tRNA molecule. It is therefore possible that the variability in fitness measures in Li *et al.*'s high throughput experiment could be at least partially responsible for these discrepancies. In any case, this comparison supports the

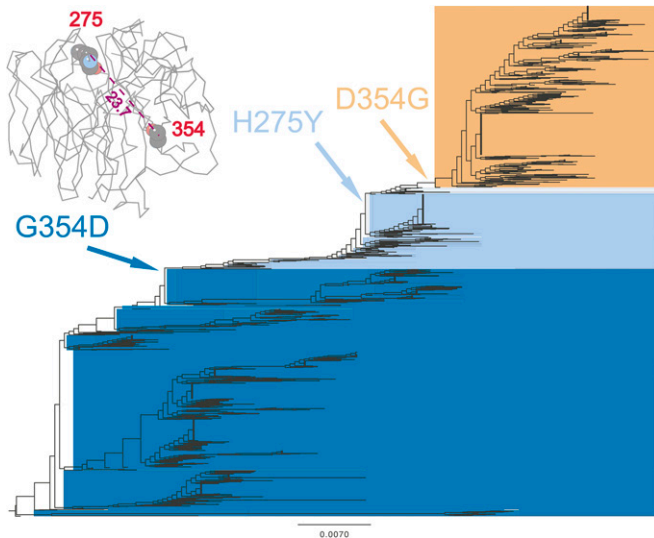


Figure 3 Epistatic pairs of AAs detected in the Duan14NA data set. The epistatic mutations that we detected are plotted on the NA phylogenetic tree. Inset: the epistatic sites are mapped on a three-dimensional NA protein structure (based on template 1HA0). The numbers show the AA positions experimentally validated (in red). The numbers in purple show the physical distance between epistatic sites (in Å).

results of our simulation studies that show that our method has excellent specificity and mediocre sensitivity.

Correlated networks follow a temporal pattern: In a fifth study, the mutations involved in changes of antigenic clusters were investigated, as it was found that double mutations could suffice to explain such cluster changes (Koel *et al.* 2013). The high mutation rate of this virus' antigen, however, does not explain why these cluster changes do not occur more often than the observed 3.3 years, which led Koel and collaborators to postulate that “comutations” may be required to maintain viral fitness, and hence accelerate evolutionary trajectories when one of the two mutations is neutral or even deleterious (Drake 2007). In light of our study, the immediate interpretation of these results would be that correlated evolution is involved during such cluster changes. We reanalyzed these data to test this hypothesis.

By doing so, we find that some of the substitutions previously (and experimentally) implicated in cluster change are indeed involved in epistasis (Figure 4, A and B, and Table S10 in File S1). In particular, as in the Gong13NP data set, we find evidence for networks of correlations, either as short chains such as position 155 interacting with both 158 and 146, which are involved in two consecutive cluster changes; but we also find a much larger network of interactions involving 20 sites, some of which are also involved in the last four cluster changes (Figure 4B). Again, as in the Gong13NP data set, a temporal sequence of substitutions along this network can be found: G124D, found at the SI87/BE89 transition, interacts with K299R, G172D, and E82K; G172D interacts with sites involved in the next transition, BE89/BE92, such as G135K, which is again involved in the next transition, BE92/WU95, where G172E interacts with N262S and

V196A, which is itself interacting with K156Q, involved in the WU95/SY97 transition; finally, K156Q interacts with T192I, involved in the SY97/FU02 transition. It is tempting to propose that such chained interactions reflect permissive substitutions and hence may provide an explanation, as an evolutionary constraint, to the paradox of high mutation rate and slow antigenic evolution. Yet, can we delve further into the nature of these constraints?

At first inspection, Figure 4C suggests that all these interactions are located in the head of the HA protein and hence might respond to steric constraints, *i.e.*, physically mediated. However, Figure S20 in File S1 shows that the strength of these associations is not related to physical distance between pairs of interacting sites. Again, the average distance between pairs of interacting sites is 23.7 Å (SD = 10.4), which is much larger than the canonical 8.5 Å for close proximity. Can we obtain some evidence about the nature of such long-range interactions?

Long-range interactions can be functionally mediated:

The analysis of a sixth data set, Adam03M2, sheds some light on this question. Influenza A viruses are resistant to M2 inhibitors, such as adamantane, and this resistance is associated with the S31N AA substitution, which is found in >95% of the currently circulating viruses (Wang *et al.* 2013; Garcia and Aris-Brosou 2014). There is evidence supporting that the spread of S31N may be unrelated to drug selective pressure, but instead results from its interaction with advantageous mutations located elsewhere in the viral genome (Simonsen *et al.* 2007)—or maybe just in the M2 gene. To test this epistatic hypothesis, we used our approach to analyze a data set of M2 sequences.

The results show that only one pair of epistatic sites (S31N×V51I) is detected (Figure 5, Table S11 in File S1). Again, it is a long-range interaction (35.96 Å), but the reason why this data set is illuminating is that the mutation at position 51 has been shown to play a role in virus replication by stabilizing the amphipathic helices of the M2 protein (Stewart and Pekosz 2011). Thus, V51I may enhance the fitness of M2 protein to increase the frequency of adamantane resistance associated with the S31N mutation. The reversion I51V that appeared in a few sequences in 2000 (red clade in Figure 5) was apparently quickly lost, which supports the hypothesis that the V51I mutation is permissive of S31N. This reversion also supports that, even in the face of high mutation rates, (i) our algorithm still maintains high specificity (Figure 1), and (ii) epistasis can be a very powerful force.

Conclusions

With historical contingency, the accumulation of epistatic substitutions can be seen as a coevolutionary process, where what happens at one AA site is conditional on what happened at another site. Here, it is this very idea of coevolution that we harnessed by coopting a method developed in ecology to test for the correlated evolution of phenotypic traits (Pagel 1994; Pagel and Meade 2006): instead of treating pairs of phenotypic traits as such, we repurposed the method to deal with pairs of

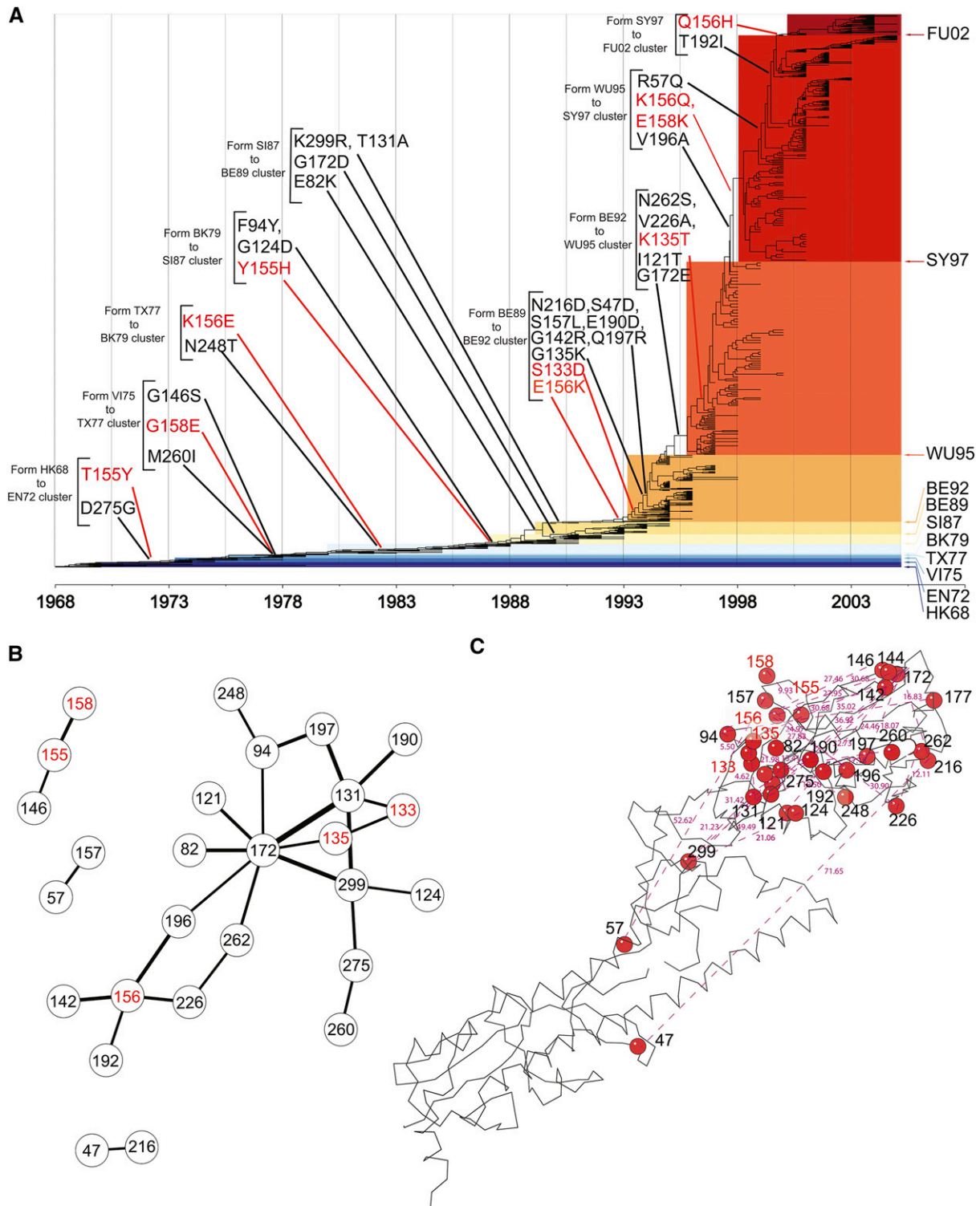


Figure 4 Epistatic pairs of AAs detected in the Koel13HA data set. (A) The epistatic mutations that we detected are plotted on the HA phylogenetic tree. The substitutions in red were experimentally validated to be responsible for cluster change (Koel *et al.* 2013). The antigenic clusters are named after the first vaccine strain in the cluster, with letters and digits referring to location and year of isolation (HK, Hong Kong; EN, England; VI, Victoria; TX, Texas; BK, Bangkok; SI, Sichuan; BE, Beijing; WU, Wuhan; SY, Sydney; FU, Fujian). (B) The thickness of the links is proportional to $-\log_{10} P$ -values, the strength of evidence supporting the interaction. (C) The epistatic sites are mapped on a three-dimensional HA protein structure (based on template 3WHE). The numbers show the AA positions experimentally validated (in red) and those detected only in this study (in black). The numbers in purple show the physical distance between epistatic sites (in Å).

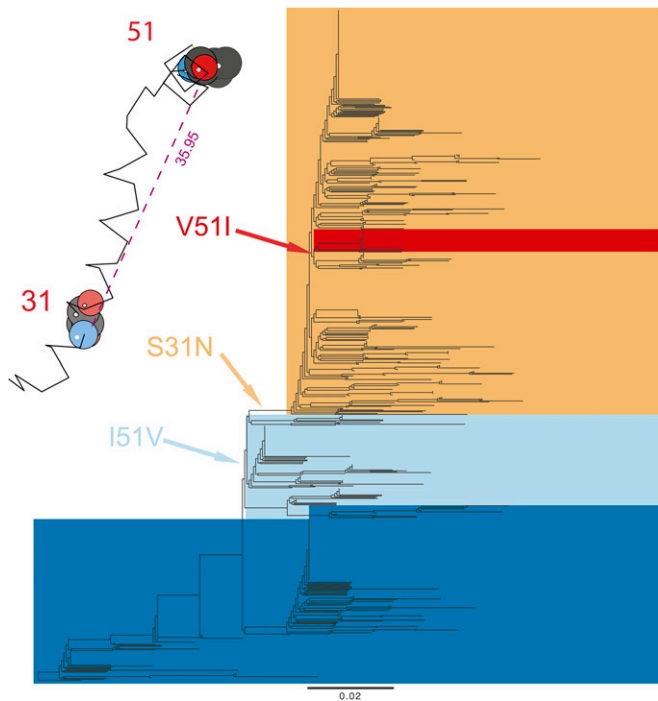


Figure 5 Epistatic pairs of AAs detected in the Adam03M2 data set. The epistatic mutations that we detected are plotted on the M2 phylogenetic tree. Inset: the epistatic sites are mapped on a three-dimensional M2 protein structure (based on template 2KIH). The numbers show the AA positions detected (in red). The numbers in purple show the physical distance between epistatic sites (in Å).

AA sites. Because the original method was developed to handle binary traits, we explored two ways of recoding data and showed that treating AAs as outgroup/ingroup consensus states was a more sensible (albeit less intuitive) option than using physicochemical properties. Extensive simulations demonstrated mediocre sensitivity, but excellent specificity and precision, even in the face of dinucleotide bias, so that pairs of AAs that are detected can be assumed to be actually coevolving. We then validated our approach against other computational results, showing marginal overlap, and against experimentally validated results, showing extensive overlap; hereby suggesting that detected pairs of AAs are genuinely interacting.

With this good statistical behavior, further analyses of independent influenza data sets showed a consistent pattern: (i) many pairs of correlated sites are involved in epistatic interactions; (ii) these pairs of sites form extensive networks of sites, consistent with Poon *et al.* (2007b), but they are also affected by substitutions that occur sequentially—showing evidence for historical contingency in fast-evolving organisms; and (iii), more intriguingly, that these epistatic pairs of sites form long-range spatial interactions. This latter point precludes the idea of a close physical link as in the case of tRNA molecules (Kimura 1985; Chen *et al.* 1999), so that these long-range interactions must bring about stability (Thomas *et al.* 2010) and/or conformational (Mitraki *et al.* 1991; Newcomb *et al.* 1997; Harms and Thornton 2014) and/or functional changes, as in the case of the M2 data set

or in the case of the Ebola virus (Ibeh *et al.* 2016). While there is evidence that epistasis can be prevalent in RNA viruses (at least 31% in Shapiro *et al.* 2006) and in bacteria (15% in Weinreich *et al.* 2006), it is not impossible that epistasis reflects an evolutionary constraint stronger in RNA viruses than in organisms with larger and more redundant genomes: because these viruses have a small genome, mutations are expected to have large fitness effects, which can be alleviated by compensatory mutations (Sanjuán and Elena 2006).

Our choice of focusing on a segmented RNA virus such as influenza may be problematic, in particular in the case of H3N2 viruses, which show a pattern of punctuated evolution that can be interpreted as the result of clonal interference (Illingworth and Mustonen 2012; Strelkova and Lässig 2012). In the absence of recombination, each segment of the virus evolves as a clone, and each clone accumulates different beneficial mutations, only one of which becomes fixed, alongside hitchhiking deleterious mutations that, in our context, would show a pattern of correlated evolution. While (i) this process has to date only been found in H3N2 viruses, and (ii) we also find evidence for chained epistasis in H1N1 viruses (KDBP11-H1, KDBP11-N1, and Duan14NA) as well as in eukaryotic tRNA genes (Li2016), clonal interference remains a problem for our approach. On the other hand, the use of influenza has allowed us to limit our analysis to searching for evidence of epistasis within genes, *contra* among genes. This way of analyzing data, intragenically, might make sense when different segments code for proteins involved in relatively different functions. However, many RNA viruses have overlapping reading frames (such as the M and NS segments in influenza A), so that we can also expect coevolution across viral genes (Neverov *et al.* 2015). Furthermore, experimental evidence in other organisms, such as yeasts, shows that epistatic interactions can involve multiple genes and hence be *intergenic* (Sorrells *et al.* 2015). Although a method to detect epistasis in segmented genomes was proposed (Neverov *et al.* 2015), the computational costs of whole-genome scans can seem prohibitive. Assessing the prevalence of epistasis over entire genomes remains an unexplored area, but is one that we are currently investigating.

In this context, how can we explain the existence of large networks of interacting sites? One possibility would be that a changing environment creates new adaptive landscapes, and that natural populations (*contra* those from *in vitro* studies) do not climb peaks on the landscape but rather chase moving targets (Gavrilets 2004, p. 36). While it is not clear whether such landscapes are robust to changing environments (Hartl 2014), they are certainly a reality in the world of viruses, where vaccination regularly alters the adaptive landscape, thereby leading to chained networks of epistatic interactions, as a mere by-product of evolution. But then, one can wonder if the metaphor of a landscape itself is appropriate when all mutational trajectories are not accessible from specific genomic backgrounds (Weinreich 2010; Sorrells *et al.* 2015). This may be one of the reasons why evolution is so difficult to predict (Weinreich 2010; Sandie and Aris-Brosou 2014; Natarajan *et al.* 2016).

Acknowledgments

We thank Jeremy Dettman, Neke Ibeh, Rees Kassen, Nicolas Rodrigue, and Alex Wong for fruitful discussions; Sergey Kryazhimskiy and Joshua Plotkin as well as Chuan Li and Jianzhi Zhang for sharing data with us; and two anonymous reviewers for comments that improved the article. We are grateful to Compute Canada and to Ontario's High Performance Computing Virtual Laboratory for giving us compute time. This work was supported by the Natural Sciences Research Council of Canada, the Canada Foundation for Innovation (S.A.B.), and by the University of Ottawa (J.C.N. and J.D.).

Literature Cited

- Abed, Y., N. Goyette, and G. Boivin, 2005 Generation and characterization of recombinant influenza A (H1N1) viruses harboring amantadine resistance mutations. *Antimicrob. Agents Chemother.* 49: 556–559.
- Anisimova, M., and O. Gascuel, 2006 Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55: 539–552.
- Aris-Brosou, S., and N. Rodrigue, 2012 The essentials of computational molecular evolution. *Methods Mol. Biol.* 855: 111–152.
- Atchley, W. R., K. R. Wollenberg, W. M. Fitch, W. Terhalle, and A. W. Dress, 2000 Correlations among amino acid sites in bhlh protein domains: an information theoretic analysis. *Mol. Biol. Evol.* 17: 164–178.
- Atilgan, A. R., P. Akan, and C. Baysal, 2004 Small-world communication of residues and significance for protein dynamics. *Biophys. J.* 86: 85–91.
- Bao, Y., P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky *et al.*, 2008 The influenza virus resource at the national center for biotechnology information. *J. Virol.* 82: 596–601.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.
- Biasini, M., S. Bienert, A. Waterhouse, K. Arnold, G. Studer *et al.*, 2014 Swiss-model: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42: W252–W258.
- Blount, Z. D., C. Z. Borland, and R. E. Lenski, 2008 Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 105: 7899–7906.
- Chan, P. P., and T. M. Lowe, 2016 GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* 44: D184–D189.
- Chen, V. B., I. W. Davis, and D. C. Richardson, 2009 King (kinemage, next generation): a versatile interactive molecular and scientific visualization program. *Protein Sci.* 18: 2403–2409.
- Chen, Y., D. B. Carlini, J. F. Baines, J. Parsch, J. M. Braverman *et al.*, 1999 RNA secondary structure and compensatory evolution. *Genes Genet. Syst.* 74: 271–286.
- Dib, L., D. Silvestro, and N. Salamin, 2014 Evolutionary footprint of coevolving positions in genes. *Bioinformatics* 30: 1241–1249.
- Dib, L., X. Meyer, P. Artimo, V. Ioannidis, H. Stockinger *et al.*, 2015 Coev-web: a web platform designed to simulate and evaluate coevolving positions along a phylogenetic tree. *BMC Bioinformatics* 16: 394.
- Drake, J. W., 2007 Too many mutants with multiple mutations. *Crit. Rev. Biochem. Mol. Biol.* 42: 247–258.
- Drummond, A. J., S. Y. Ho, M. J. Phillips, and A. Rambaut, 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4: e88.
- Duan, S., E. A. Govorkova, J. Bahl, H. Zaraket, T. Baranovich *et al.*, 2014 Epistatic interactions between neuraminidase mutations facilitated the emergence of the oseltamivir-resistant H1N1 influenza viruses. *Nat. Commun.* 5: 5029.
- Garcia, V., and S. Aris-Brosou, 2014 Comparative dynamics and distribution of influenza drug resistance acquisition to protein M2 and neuraminidase inhibitors. *Mol. Biol. Evol.* 31: 355–363.
- Garrity, G. M., J. A. Bell, and T. G. Lilburn, 2004 *Taxonomic Outline of the Prokaryotes. Bergey's Manual of Systematic Bacteriology*. Springer, New York.
- Gavrilets, S., 2004 *Fitness Landscapes and the Origin of Species*, Vol. 41. Princeton University Press, Princeton, NJ.
- Gloor, G. B., L. C. Martin, L. M. Wahl, and S. D. Dunn, 2005 Mutual information in protein multiple sequence alignments reveals two classes of coevolving positions. *Biochemistry* 44: 7156–7165.
- Gong, L. I., M. A. Suchard, and J. D. Bloom, 2013 Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife* 2: e00631.
- Gowri-Shankar, V., and H. Jow, 2006 PHASE: a software package for phylogenetics and sequence evolution. University of Manchester, Manchester, United Kingdom. <http://www.bioinf.man.ac.uk/resources/phase/>.
- Greenbaum, B. D., A. J. Levine, G. Bhanot, and R. Rabadan, 2008 Patterns of evolution and host gene mimicry in influenza and other RNA viruses. *PLoS Pathog.* 4: e1000079.
- Harms, M. J., and J. W. Thornton, 2014 Historical contingency and its biophysical basis in glucocorticoid receptor evolution. *Nature* 512: 203–207.
- Hartl, D. L., 2014 What can we learn from fitness landscapes? *Curr. Opin. Microbiol.* 21: 51–57.
- Hedges, S. B., J. Marin, M. Suleski, M. Paymer, and S. Kumar, 2015 Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* 32: 835–845.
- Ibeh, N., J. C. Nshogozabahizi, and S. Aris-Brosou, 2016 Both epistasis and diversifying selection drive the structural evolution of the Ebola virus glycoprotein mucin-like domain. *J. Virol.* 90: 5475–5484.
- Illingworth, C. J. R., and V. Mustonen, 2012 Components of selection in the evolution of the influenza virus: linkage effects beat inherent selection. *PLoS Pathog.* 8: e1003091.
- Jukes, T. H., and C. R. Cantor, 1969 Evolution of protein molecules, pp. 132 in *Mammalian Protein Metabolism*, Vol. III. Academic Press, New York.
- Kimura, M., 1985 The role of compensatory neutral mutations in molecular evolution. *J. Genet.* 64: 7–19.
- Koel, B. F., D. F. Burke, T. M. Bestebroer, S. van der Vliet, G. C. Zondag *et al.*, 2013 Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* 342: 976–979.
- Korber, B. T., R. M. Farber, D. H. Wolpert, and A. S. Lapedes, 1993 Covariation of mutations in the V3 loop of human immunodeficiency virus type 1 envelope protein: an information theoretic analysis. *Proc. Natl. Acad. Sci. USA* 90: 7176–7180.
- Kryazhimskiy, S., J. Dushoff, G. A. Bazykin, and J. B. Plotkin, 2011 Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS Genet.* 7: e1001301.
- Li, C., W. Qian, C. J. Maclean, and J. Zhang, 2016 The fitness landscape of a tRNA gene. *Science* 352: 837–840.
- Lynch, M., 2007 *The Origins of Genome Architecture*. Sinauer Associates, Sunderland, MA.
- Mitraki, A., B. Fane, C. Haase-Pettingell, J. Sturtevant, and J. King, 1991 Global suppression of protein folding defects and inclusion body formation. *Science* 253: 54–58.

- Mohrig, J. R., K. A. Moerke, D. L. Cloutier, B. D. Lane, E. C. Person *et al.*, 1995 Importance of historical contingency in the stereochemistry of hydratase-dehydratase enzymes. *Science* 269: 527–529.
- Muse, S. V., 1995 Evolutionary analyses of DNA sequences subject to constraints of secondary structure. *Genetics* 139: 1429–1439.
- Nasrallah, C. A., and J. P. Huelsenbeck, 2013 A phylogenetic model for the detection of epistatic interactions. *Mol. Biol. Evol.* 30: 2197–2208.
- Natarajan, C., F. G. Hoffmann, R. E. Weber, A. Fago, C. C. Witt *et al.*, 2016 Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* 354: 336–339.
- Nawrocki, E. P., and S. R. Eddy, 2013 Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29: 2933–2935.
- Neher, E., 1994 How frequent are correlated changes in families of protein sequences? *Proc. Natl. Acad. Sci. USA* 91: 98–102.
- Neverov, A. D., S. Kryazhimskiy, J. B. Plotkin, and G. A. Bazykin, 2015 Coordinated evolution of influenza A surface proteins. *PLoS Genet.* 11: e1005404.
- Newcomb, R. D., P. M. Campbell, D. L. Ollis, E. Cheah, R. J. Russell *et al.*, 1997 A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. *Proc. Natl. Acad. Sci. USA* 94: 7464–7468.
- Ortlund, E. A., J. T. Bridgham, M. R. Redinbo, and J. W. Thornton, 2007 Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* 317: 1544–1548.
- Pagel, M., 1994 Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B Biol. Sci.* 255: 37–45.
- Pagel, M., and A. Meade, 2006 Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* 167: 808–825.
- Paradis, E., 2006 *Analysis of Phylogenetics and Evolution With R*. Springer, New York.
- Pollock, D. D., W. R. Taylor, and N. Goldman, 1999 Coevolving protein residues: maximum likelihood identification and relationship to structure. *J. Mol. Biol.* 287: 187–198.
- Poon, A. F. Y., F. I. Lewis, S. L. K. Pond, and S. D. W. Frost, 2007a Evolutionary interactions between N-linked glycosylation sites in the HIV-1 envelope. *PLOS Comput. Biol.* 3: e11.
- Poon, A. F. Y., F. I. Lewis, S. L. K. Pond, and S. D. W. Frost, 2007b An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. *PLOS Comput. Biol.* 3: e231.
- Poon, A. F., F. I. Lewis, S. D. Frost, and S. L. K. Pond, 2008 Spidermonkey: rapid detection of co-evolving sites using Bayesian graphical models. *Bioinformatics* 24: 1949–1950.
- Price, M. N., P. S. Dehal, and A. P. Arkin, 2010 Fasttree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490.
- R Development Core Team, 2013 R: A language and environment for statistical computing. The R Foundation for Statistical Computing, Vienna, Austria.
- Revolution Analytics and S. Weston, 2013 foreach: Foreach looping construct for r. R package version 1. <http://CRAN.R-project.org/package=foreach>.
- Rzhetsky, A., 1995 Estimating substitution rates in ribosomal RNA genes. *Genetics* 141: 771–783.
- Sandie, R., and S. Aris-Brosou, 2014 Predicting the emergence of H3N2 influenza viruses reveals contrasted modes of evolution of HA and NA antigens. *J. Mol. Evol.* 78: 1–12.
- Sanjuán, R., and S. F. Elena, 2006 Epistasis correlates to genomic complexity. *Proc. Natl. Acad. Sci. USA* 103: 14402–14405.
- Schöniger, M., and A. von Haeseler, 1994 A stochastic model for the evolution of autocorrelated DNA sequences. *Mol. Phylogenet. Evol.* 3: 240–247.
- Shapiro, B., A. Rambaut, O. G. Pybus, and E. C. Holmes, 2006 A phylogenetic method for detecting positive epistasis in gene sequences and its application to RNA virus evolution. *Mol. Biol. Evol.* 23: 1724–1730.
- Shindyalov, I. N., N. A. Kolchanov, and C. Sander, 1994 Can three-dimensional contacts in protein structures be predicted by analysis of correlated mutations? *Protein Eng.* 7: 349–358.
- Simonsen, L., C. Viboud, B. T. Grenfell, J. Dushoff, L. Jennings *et al.*, 2007 The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. *Mol. Biol. Evol.* 24: 1811–1820.
- Soares, P., L. Ermini, N. Thomson, M. Mormina, T. Rito *et al.*, 2009 Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* 84: 740–759.
- Sorrells, T. R., L. N. Booth, B. B. Tuch, and A. D. Johnson, 2015 Intersecting transcription networks constrain gene regulatory evolution. *Nature* 523: 361–365.
- Stewart, S. M., and A. Pekosz, 2011 Mutations in the membrane-proximal region of the influenza A virus M2 protein cytoplasmic tail have modest effects on virus replication. *J. Virol.* 85: 12179–12187.
- Strelkowa, N., and M. Lässig, 2012 Clonal interference in the evolution of influenza. *Genetics* 192: 671–682.
- Sutto, L., S. Marsili, A. Valencia, and F. L. Gervasio, 2015 From residue coevolution to protein conformational ensembles and functional dynamics. *Proc. Natl. Acad. Sci. USA* 112: 13567–13572.
- Talavera, D., S. C. Lovell, and S. Whelan, 2015 Covariation is a poor measure of molecular coevolution. *Mol. Biol. Evol.* 32: 2456–2468.
- Tavaré, S., 1986 Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math. Life Sci.* 17: 57–86.
- Taylor, W. R., and K. Hatrick, 1994 Compensating changes in protein multiple sequence alignments. *Protein Eng.* 7: 341–348.
- Thomas, V. L., A. C. McReynolds, and B. K. Shoichet, 2010 Structural bases for stability-function tradeoffs in antibiotic resistance. *J. Mol. Biol.* 396: 47–59.
- Tillier, E. R. M., and R. A. Collins, 1998 High apparent rate of simultaneous compensatory base-pair substitutions in ribosomal rna. *Genetics* 148: 1993–2002.
- Wang, J., Y. Wu, C. Ma, G. Fiorin, J. Wang *et al.*, 2013 Structure and inhibition of the drug-resistant S31N mutant of the M2 ion channel of influenza A virus. *Proc. Natl. Acad. Sci. USA* 110: 1315–1320.
- Weinreich, D. M., 2010 Predicting molecular evolutionary trajectories in principle and in practice, pp. 1–9 in *Encyclopedia of Life Sciences*, John Wiley & Sons, Chichester, United Kingdom.
- Weinreich, D. M., N. F. Delaney, M. A. DePristo, and D. L. Hartl, 2006 Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312: 111–114.
- Worobey, M., G.-Z. Han, and A. Rambaut, 2014 A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* 508: 254–257.
- Xiao, N., Q. Xu, and D. Cao, 2015 protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequences. *Bioinformatics* 31: 1857–1859.
- Yang, Z., 1998 On the best evolutionary rate for phylogenetic analysis. *Syst. Biol.* 47: 125–133.
- Yang, Z., 2006 *Computational Molecular Evolution*, Vol. 21. Oxford University Press, Oxford.

Communicating editor: J. Lawrence

Supplementary text

Comparison of methods to simulate epistasis. To ensure that our simulation results were not biased by the method used to simulate epistasis, we ran a second fully factorial simulation study on a representative subset of the parameter space as described in the main text (*Materials and Methods: Validation based on simulations*). The following modifications were made: (i) Sites that evolve in a correlated manner were simulated using Coev [1] for five different strengths of correlation. This is controlled by the d/s rate parameter that represents the likelihood of a pair of sites to evolve toward (d) or away (s) from a correlated profile. (ii) Branch lengths (b) were varied across a \log_2 scale for all even integer values in $(-12, -4)$. (iii) Each tree contained a number of sequences n_s equal to 32 or 128.

With these Coev simulations, both specificity and precision (Figure S15) are higher than with the PHASE simulations, even at the lowest strength of correlation. Sensitivity appears to have an optimal branch length, similar to what we found with PHASE. With Coev however, the decline in sensitivity results from monomorphic site patterns being generated. From our data recoding and analysis, monomorphic sites are not considered since they carry no evolutionary information. The Coev model defines a unique two character profile as the coevolutionary pair to be simulated. Thus, as evolutionary time (or strength (d/s) of correlated evolution) increases, more sequences adopt this unique correlation profile until sites become monomorphic. The Coev method of simulating epistasis reflects correlated evolution when only a single, correlated, adaptive response is observed for adaptation of an entire population; our PHASE simulations did not enforce unique profiles.

That strong selective pressures can lead to unique adaptive profiles reinforces our claim that alignments should be constructed to include taxa that have not undergone the selection, which would generate epistatic response. Comparison of our method’s performance while simulating correlated evolution under two different profiles of adaptive response demonstrates that our approach is applicable to detecting epistasis during various adaptive responses.

Signal strength for detecting epistasis. To determine the optimal phylogenetic distribution of epistatic sites, we generated columns in our alignments with varying proportions of “positively” and “negatively” correlated site pairs. These simulations were done with PHASE. We define positively correlated sites as those that have identical character states (*i.e.*, (1, 1) and (0, 0)), and negatively correlated site as those that have opposed character states (*i.e.*, (1, 0) and (0, 1)). These negatively correlated sites are equivalent to noise. Note that this designation of positively and negatively correlated pairs should not be confused with the sign of epistatic interactions (positive or negative epistasis).

Intuitively, this combination leads to entirely symmetric outcomes, so the simulations proceeded by first generating a certain proportion of negatively correlated pairs; this proportion was varied between 1 and 31% of the alignment by increments of 3%. The remaining of each alignment was then filled with “positively” correlated site pairs of type (0, 0) in a proportion ranging from 0 to 100% by increments of 10%. What remained, if any, was filled with the other positive state pair (1, 1) (Figure S4B5). We did this for sequence alignments comprising 32, 64 and 128 sequences, hereby leading to 604, 726 and 763 unique simulation conditions, respectively. Each of these simulation conditions was assayed for both tree shapes (symmetric and

pectinate) with 500 replicates. Each replicate randomly shuffled the assignment of site pairs at tree tips. This resulted in nearly 2.1 million runs and thus sequence length (l_s) was fixed to one single epistatic pair to reduce computational burden. Root depth (tip-to-root distance) of trees was held constant across all conditions to allow for better comparison of the impact of additional sequences.

We find that balanced sequence alignments, that contain an equal proportion of correlated binary states, are optimal and that such alignments permit detection of correlated evolution even with high levels of noise ($\gg 31\%$; Figure S21). Consistent with our results on sensitivity, only alignments with 64 or more sequences show a high proportion of significant P -values, that is, a high ability to detect epistasis (Figure S21: panels A, D *vs.* B, C and E, F).

Tree shape, number of sequences, balance of positively correlated pairs as well as the proportion of negatively correlated pairs all affect our ability to detect epistasis (Table S1). While all the terms are significant, and have high, similar AIC values (Table S2), the balance of sequence distribution explains more of the variance than tree shape, n_s or proportion of negatively correlated pairs (Table S2). We compared the variance explained by our full model and a stepwise reduction of terms. The shape of a tree and n_s have a very weak interaction, while both terms have strong interactions with the balance of positively correlated pairs (Table S2). Notably, tree shape and the proportion of negatively correlated pairs have an equally strong interaction, suggesting that certain phylogenies would be more susceptible to noise.

Assessing the effect of dinucleotide bias. Dinucleotide bias can be seen as a form of correlated evolution, that can therefore interfere with our algorithm and generate “false positives” – pairs of sites that evolve in a correlated manner but that are not epistatic. As our method was mostly applied to influenza A viruses, which display dinucleotide bias [2], we assessed the impact that dinucleotide bias would have upon the statistical performance of our method. We performed simulations where coevolving sites had dinucleotide bias calculated from the influenza A nucleic acid sequences used in [3]. Similar to our simulations comparing our approach with coevolution simulated by Coev, we performed simulations in a subset of the most interesting parameter space of our larger simulation study: branch lengths (b) varied across a \log_2 scale for even integer values in $(-12, -4)$, and the number of sequences n_s was either 32 or 128. As per the simulations presented in the main text, sites evolving epistatically were simulated under the RNA7D model however, independently evolving sites were simulated using the RNA16 model with dinucleotide bias. We set the probability of single mutations to be 95% and double mutations to 5%. These simulations show that specificity was significantly affected by the inclusion of dinucleotide bias (see Fig. S14A; $F = 21.922$, $df = 1, P < 1 \times 10^{-5}$). While sensitivity and precision (see Fig. S14C) are lower when dinucleotide bias is included in simulations, we highlight that specificity remains above the maximum threshold, and precision remains above 95%.

We did not expect sensitivity to be different in this simulation study since epistatic sites are simulated using the same model. However, in Fig. S14B, we observe that sensitivity appears reduced. After carefully verifying this was not generated due an error in the simulation protocols, we performed statistical analysis to test if this difference was significant. Statistical analysis was done using a Type III SS anova of a general linear model with independent model of site evolution as the first term (see Table 3). We find that many parameters are significant despite our having compared identical parameter spaces in our analysis. To explain this difference we note that

standard error bars for sensitivity were huge in both these simulations and the larger study in our main text (these bars were not included for reasons of clarity). This information combined with the significance of parameters not varied between simulations suggests that the stochasticity of generating epistatic signal best explains the difference in sensitivity between simulation studies.

Altogether, these simulations provide significant evidence that a dinucleotide bias affects the specificity and possibly precision of our approach. Despite this impact the general performance of exceptional specificity and mediocre sensitivity are consistent.

Detection of chained correlations. In essence, BayesTraits [4, 5] detects pairs of traits (sites) that evolve in a correlated manner on a phylogenetic tree. However, as we did not simulate chained correlations, one could wonder if our use of BayesTraits can actually detect such patterns and tell a situation where sites 1 and 2 on the one hand and 2 and 3 on the other are correlated, *versus* a situation where 1 and 2, and 1 and 3 are correlated. Considering pairs of states at two sites, the simulations performed above demonstrate that if two sites evolve alongside each other, the chance of detecting this signal of correlation is very high. In other words, correlation is detected if each site undergo mutations together. If pairs of mutations at sites 1 and 2 occur along a different branch than those at sites 1 and 3, then we can clearly reconstruct chained interactions. It is because we find such a temporal pattern (see in particular Figure 4 in main text) that we can reconstruct such chained interactions.

Interaction between sequence length and number of sequences. Our general algorithm employs the BH correction for controlling FDR. As a result, the length of a sequence alignment can be expected to affect the ability to detect epistasis. To test this expectation, we simulated alignments of varying sequence lengths $l_s \in (10, 1280)$ on a \log_2 (doubling) scale. As our previous simulations showed little impact of tree shape (Results), we limited this simulation to symmetric trees and used PHASE here again. The number of sequences n_s was set to 32, 64 or 128 sequences. As above, root depth was held constant. For each n_s , the phylogenetic pattern that yielded the strongest signal for correlated evolution was used in the analysis. In this way, we could establish expectations of the likelihood to detect correlated evolution as a function of l_s and n_s . Each combination was replicated only 10 times, as this simulation does not assess detection performance, but rather establishes a trend in P -value after FDR: the only factor changing is the number of pairwise comparisons, which is a function of sequence length.

Both the sequence length and number of sequences have significant effects on P -values after correcting for the False Discovery Rate (FDR; Table S1; Figure S22). As expected, FDR-corrected P -values show a power law distribution as a function of sequence length (Figure S22): doubling sequence length halves our chance of finding a significant result, so that it becomes more difficult to detect any signal for epistasis with long sequences. Epistasis is not recoverable from alignments of only 32 sequences containing > 1000 sites (Figure S22).

References

- [1] Dib L, Silvestro D, Salamin N. Evolutionary footprint of coevolving positions in genes. *Bioinformatics*. 2014 May;30(9):1241–9.

- [2] Greenbaum BD, Levine AJ, Bhanot G, Rabadan R. Patterns of Evolution and Host Gene Mimicry in Influenza and Other RNA Viruses. *PLoS Pathogens*. 2008 06;4(6):e1000079. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2390760/>.
- [3] Kryazhimskiy S, Dushoff J, Bazykin GA, Plotkin JB. Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS genetics*. 2011;7(2):e1001301.
- [4] Pagel M. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London B: Biological Sciences*. 1994;255(1342):37–45.
- [5] Pagel M, Meade A. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist*. 2006;167(6):808–825.
- [6] Dib L, Meyer X, Artimo P, Ioannidis V, Stockinger H, Salamin N. Coev-web: a web platform designed to simulate and evaluate coevolving positions along a phylogenetic tree. *BMC Bioinformatics*. 2015;16:394.
- [7] Li C, Qian W, Maclean CJ, Zhang J. The fitness landscape of a tRNA gene. *Science*. 2016 May;352:837–840.

Supplementary figures

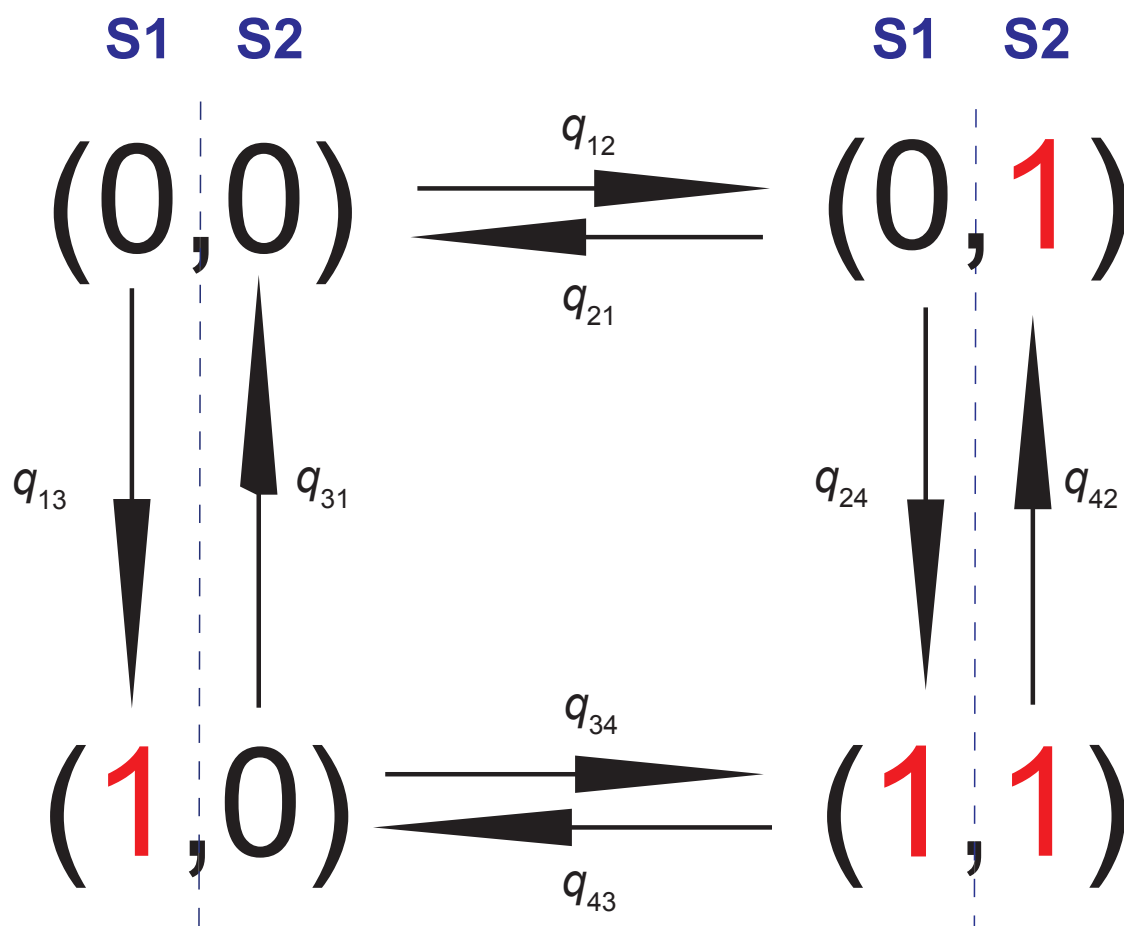


Figure S1: **Transitions among the four combinations of states resulting from two binary variables.** Pair (i, j) identifies the states of a particular pair of traits, that here are amino acid positions S1 and S2. Transitions are represented by arrows. Instantaneous rates of change are denoted $q_{i,j}$.

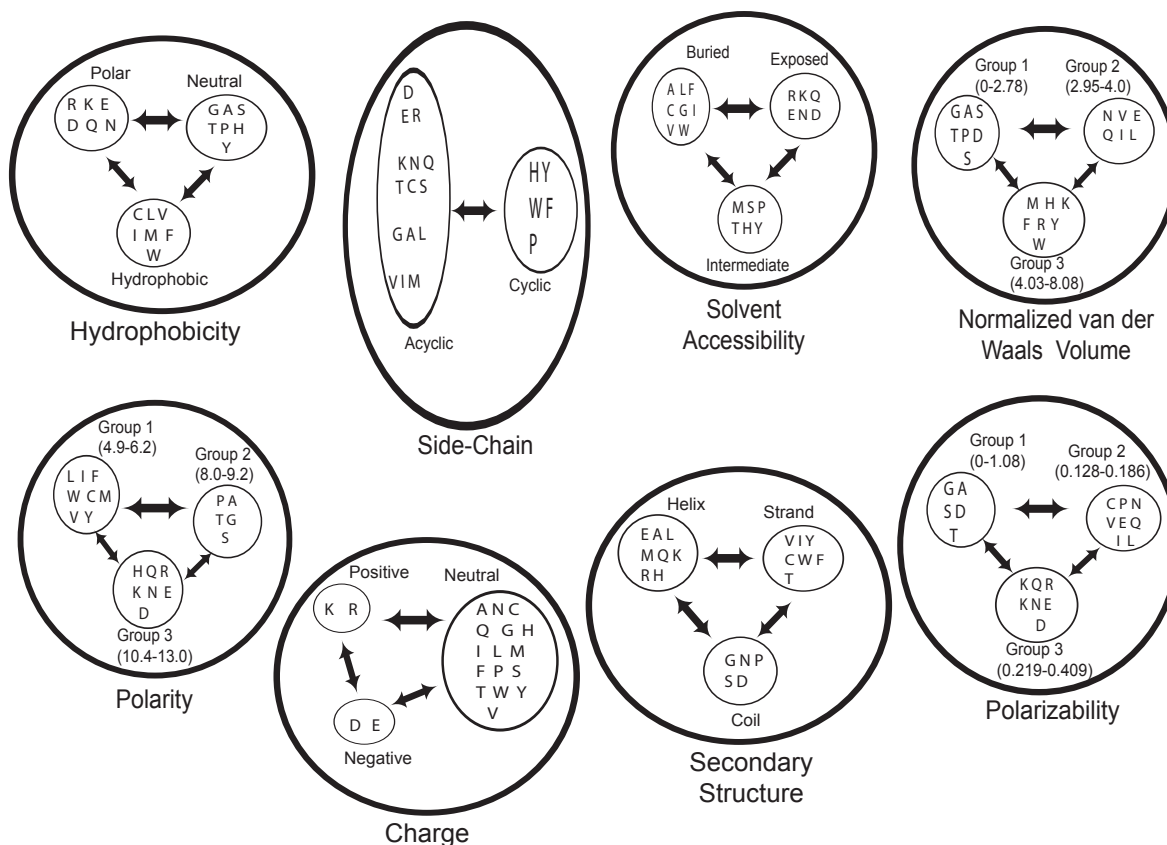


Figure S2: **The amino acids properties and their categories (states) used to recode protein alignments.** For each category, the arrows show the different transitions among physiochemical properties. For continuous categories, the numbers between bracket indicate the range of values employed in the discretization process as per the `protr` R package.

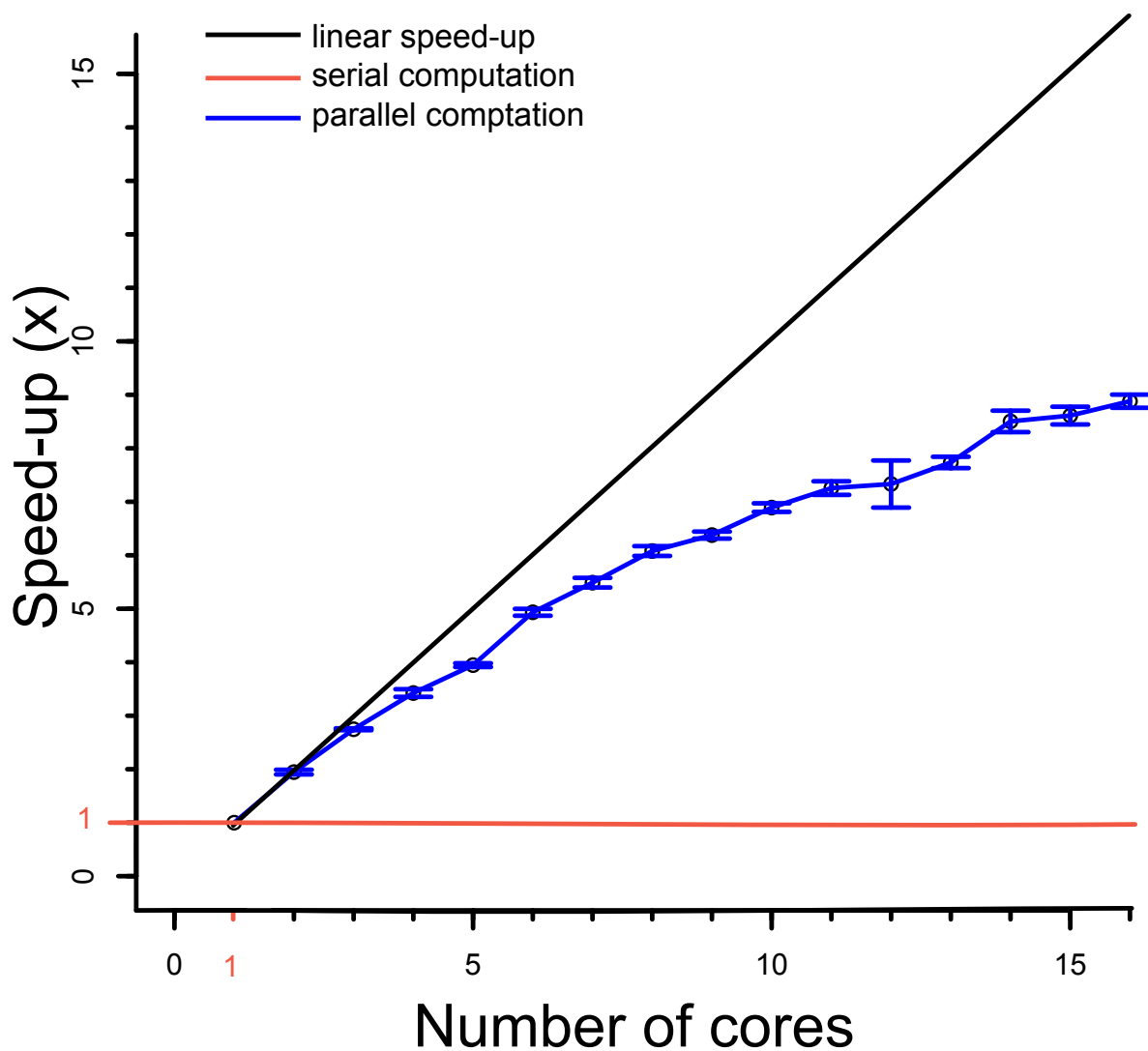


Figure S3: **Speed-up as a function of number of cores requested.** Speed-up is a measure of “acceleration” (how many times faster is the computation with n cores) compared to a serial computation (on a single core; in red). The black line indicates the ideal situation of linear scaling of performance with resource. The actual speed-up curve is shown in blue. Error bars show one standard error, based on four replicates. Overheads due to latency (among others) lead to a non-linear speed-up curve.

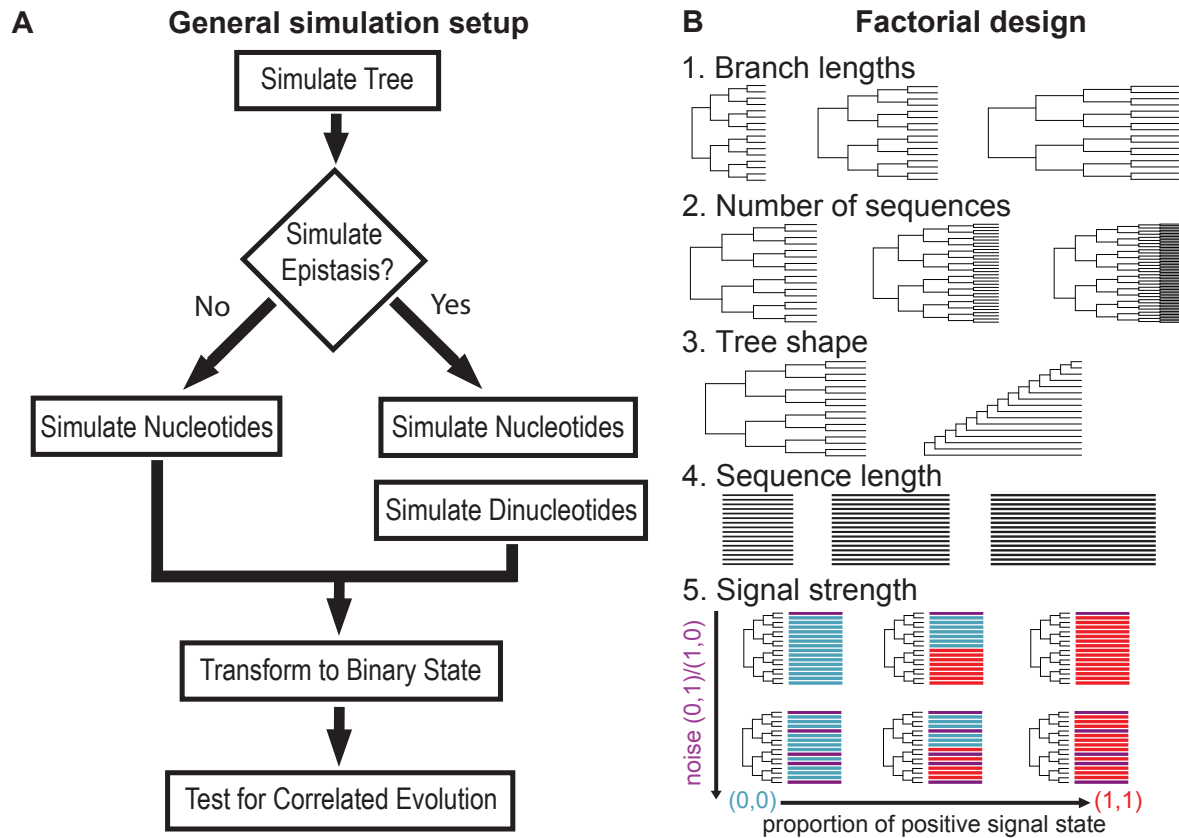


Figure S4: **Design of the simulation experiments.** (A) The general simulation setup starts by simulating a tree of a particular tree shape (symmetric / pectinate) on which sequences are generated. The program PHASE 2.0 is used to generate either single nucleotides (independent model) or dinucleotides (dependent model). In both cases, nucleotides are recoded as 0's and 1's before being analyzed with BayesTraits, hereby testing for correlated evolution. (B) A fully factorial design was adopted to test for the effect of: 1. branch lengths, 2. the number of sequences, 3. tree shape, 4. sequence length and 5. signal strength. See Methods for details.

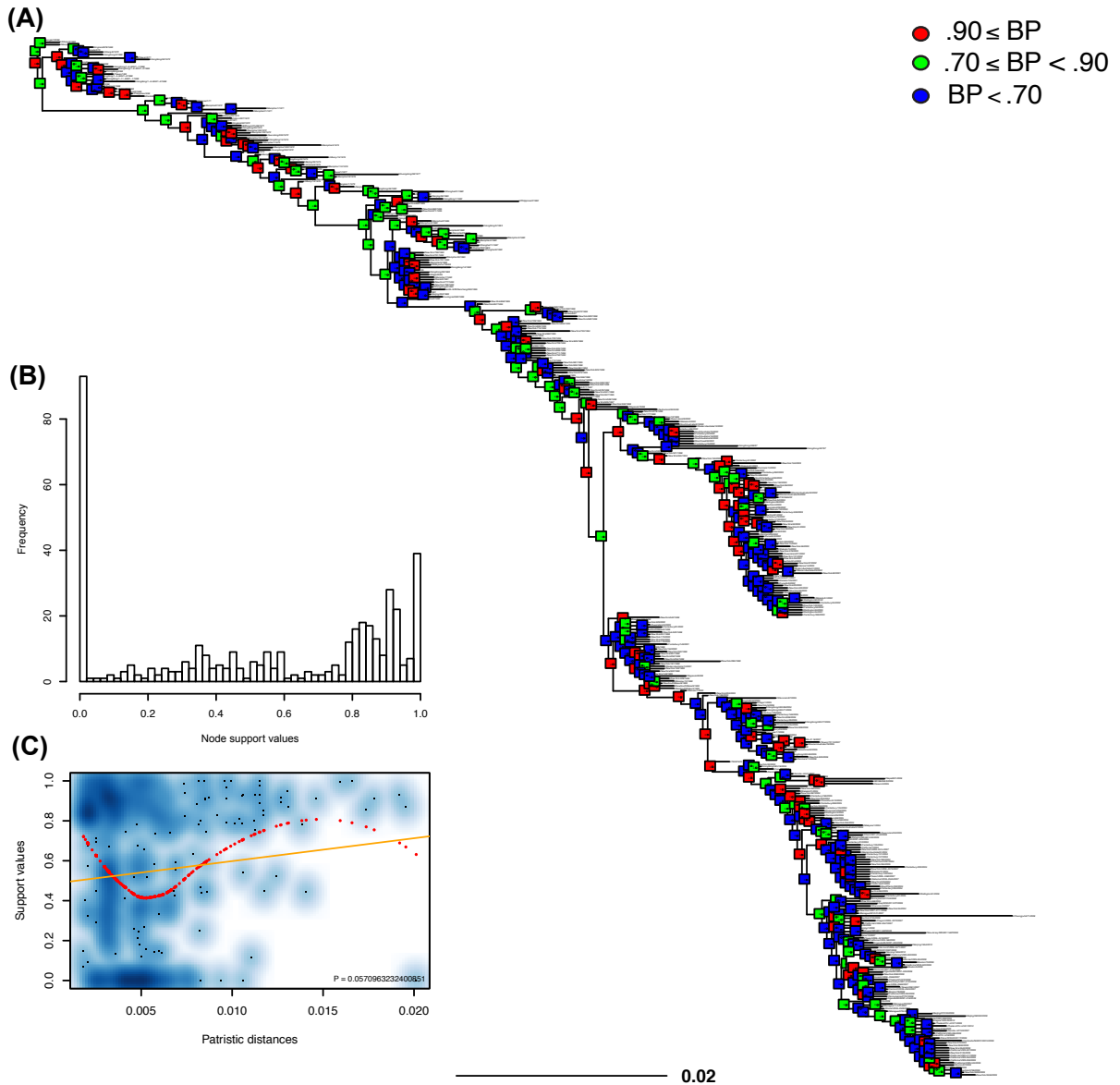


Figure S5: **Node support values for the phylogenetic analysis of the Gong13NP data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

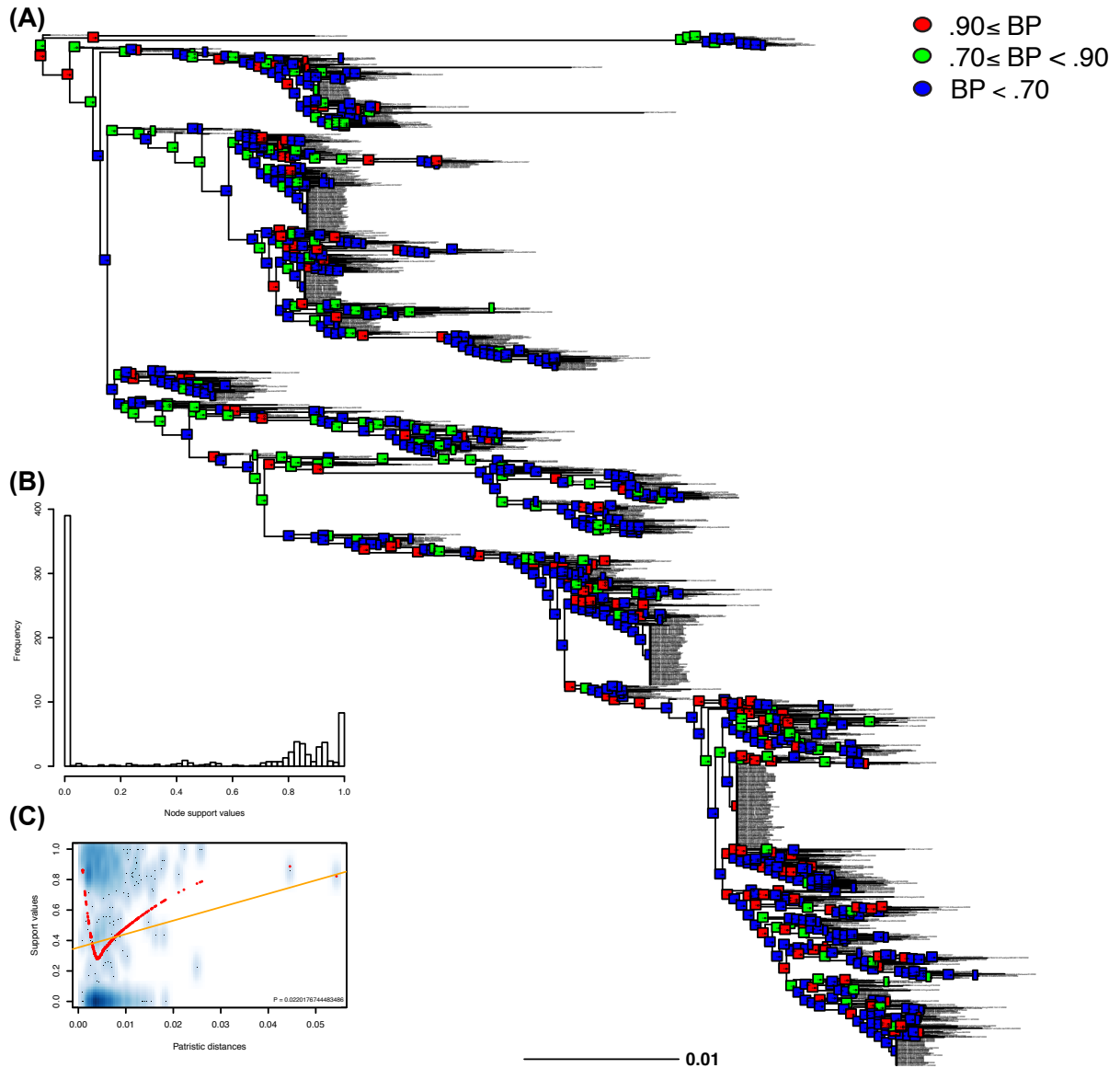


Figure S6: **Node support values for the phylogenetic analysis of the Duan14NA data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

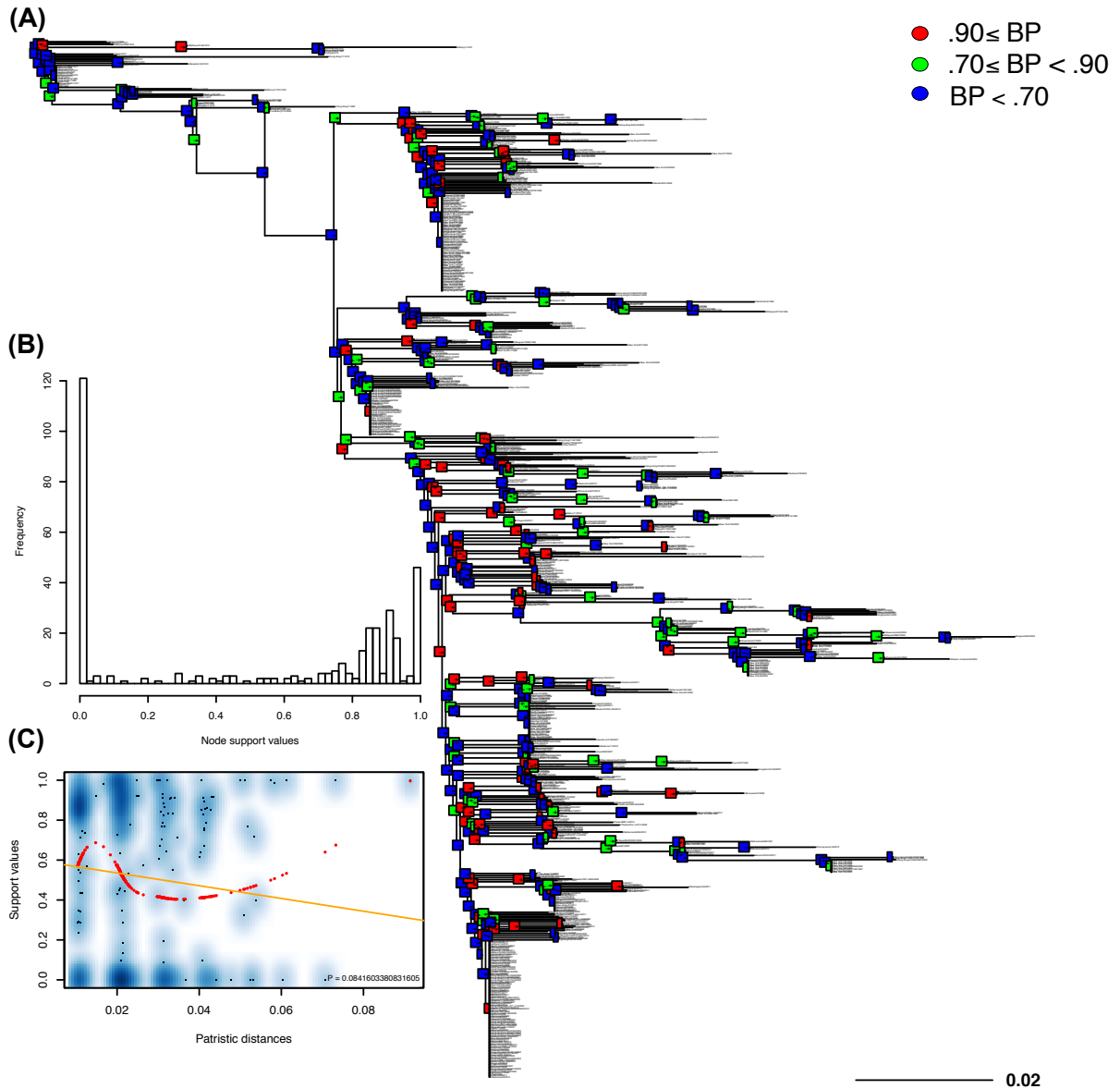


Figure S7: **Node support values for the phylogenetic analysis of the Adam03M2 data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

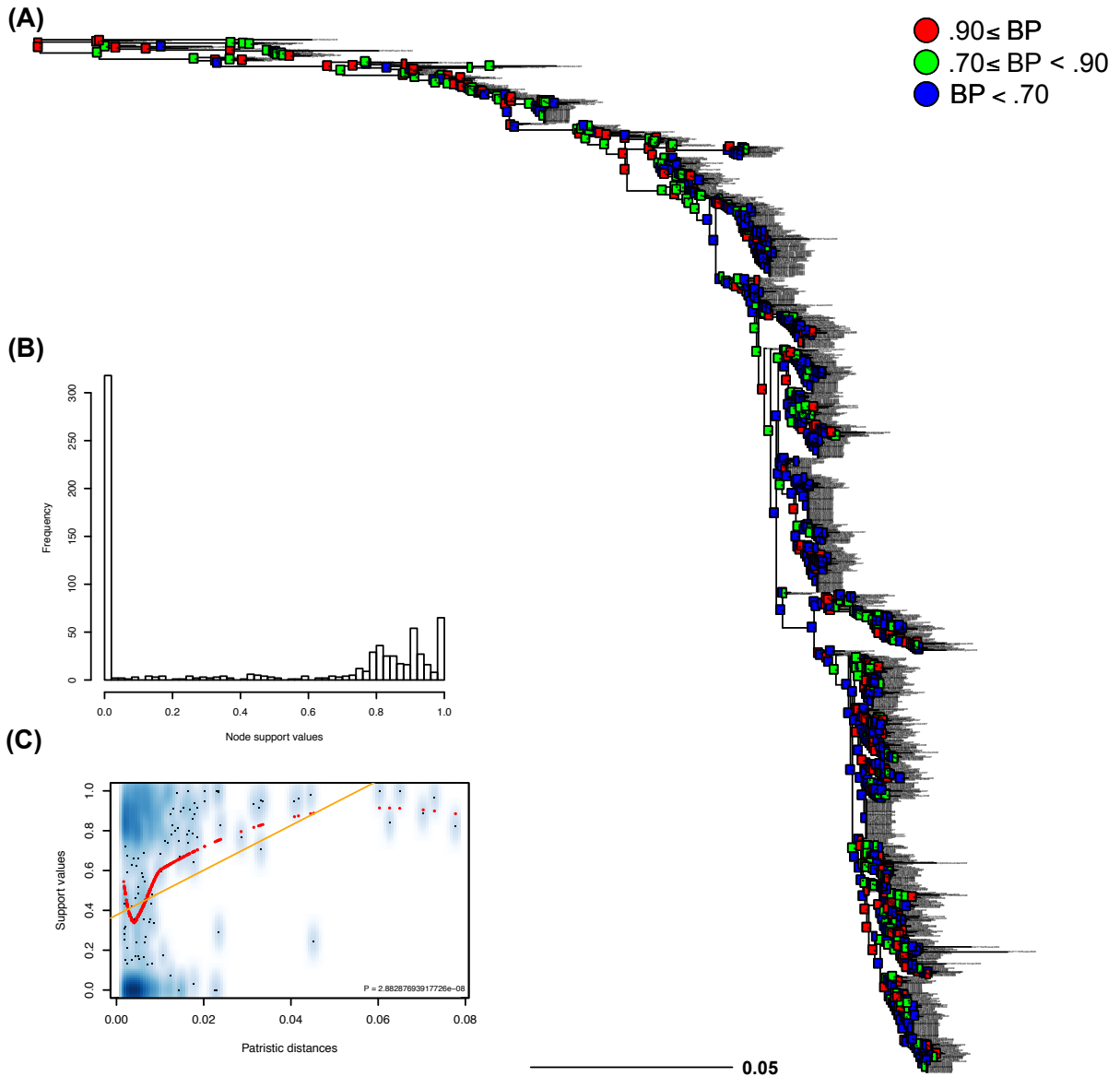


Figure S8: **Node support values for the phylogenetic analysis of the KDBP11 H1 data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

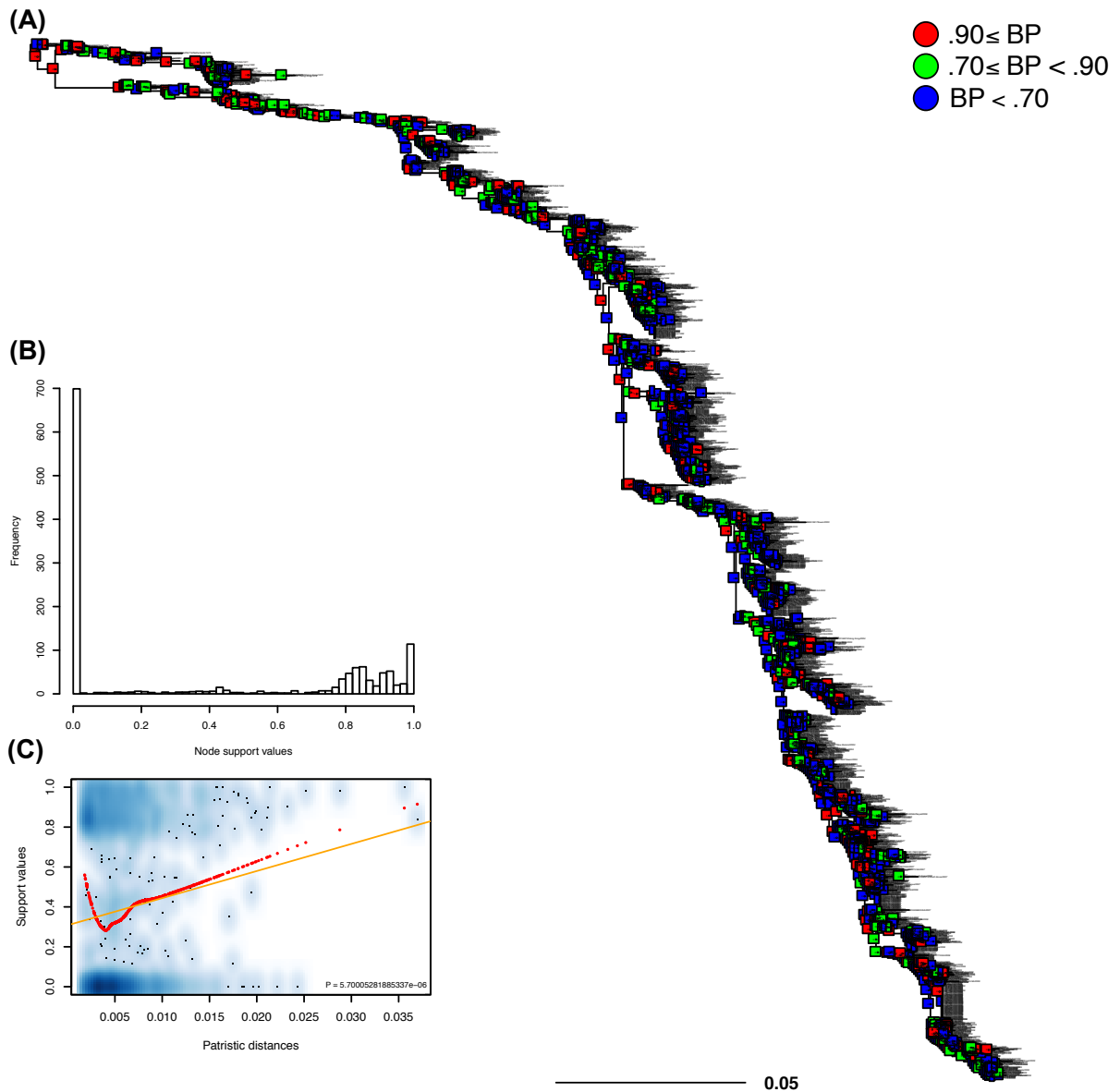


Figure S9: **Node support values for the phylogenetic analysis of the KDBP11 H3 data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

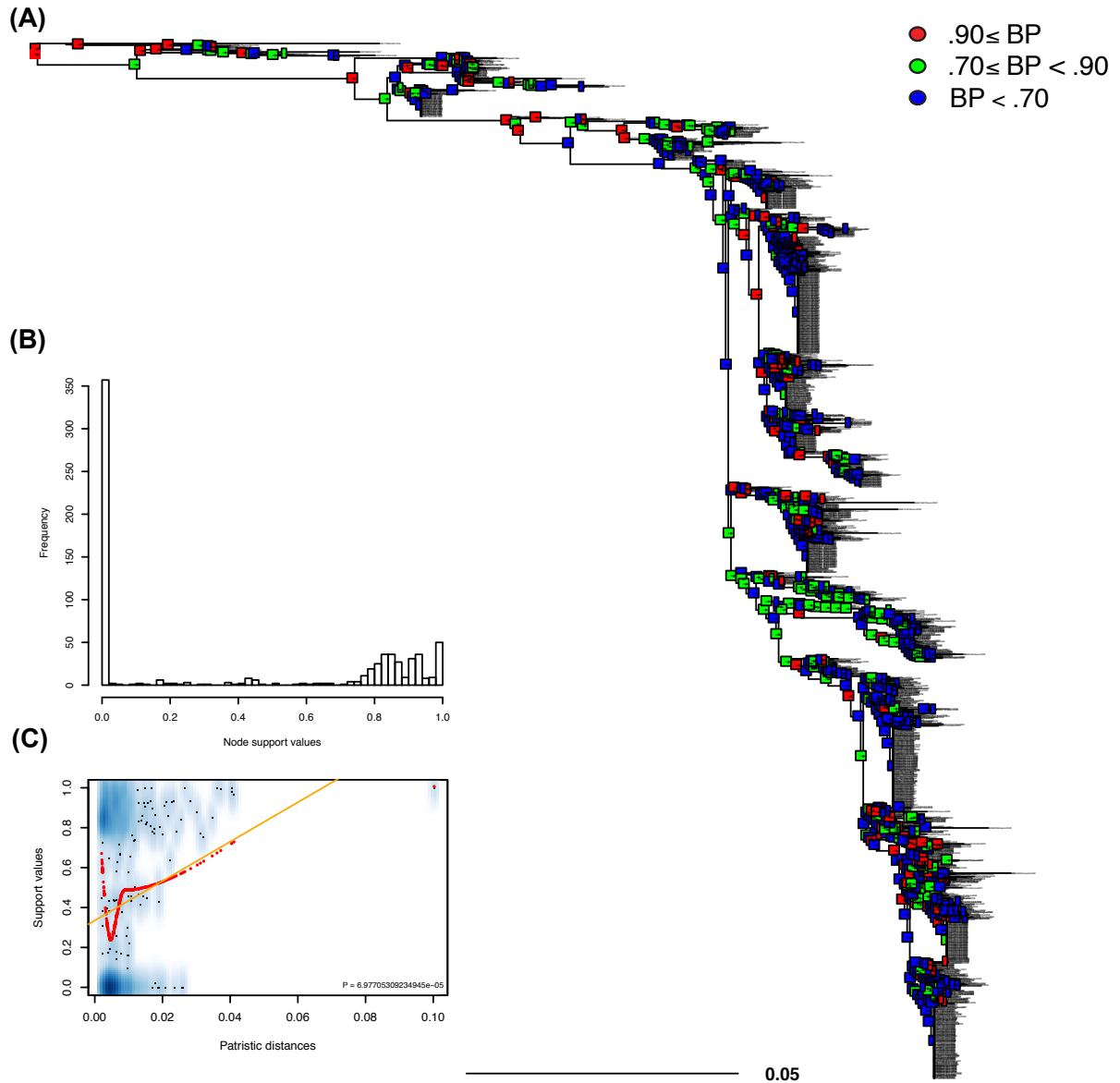


Figure S10: **Node support values for the phylogenetic analysis of the KDBP11 N1 data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

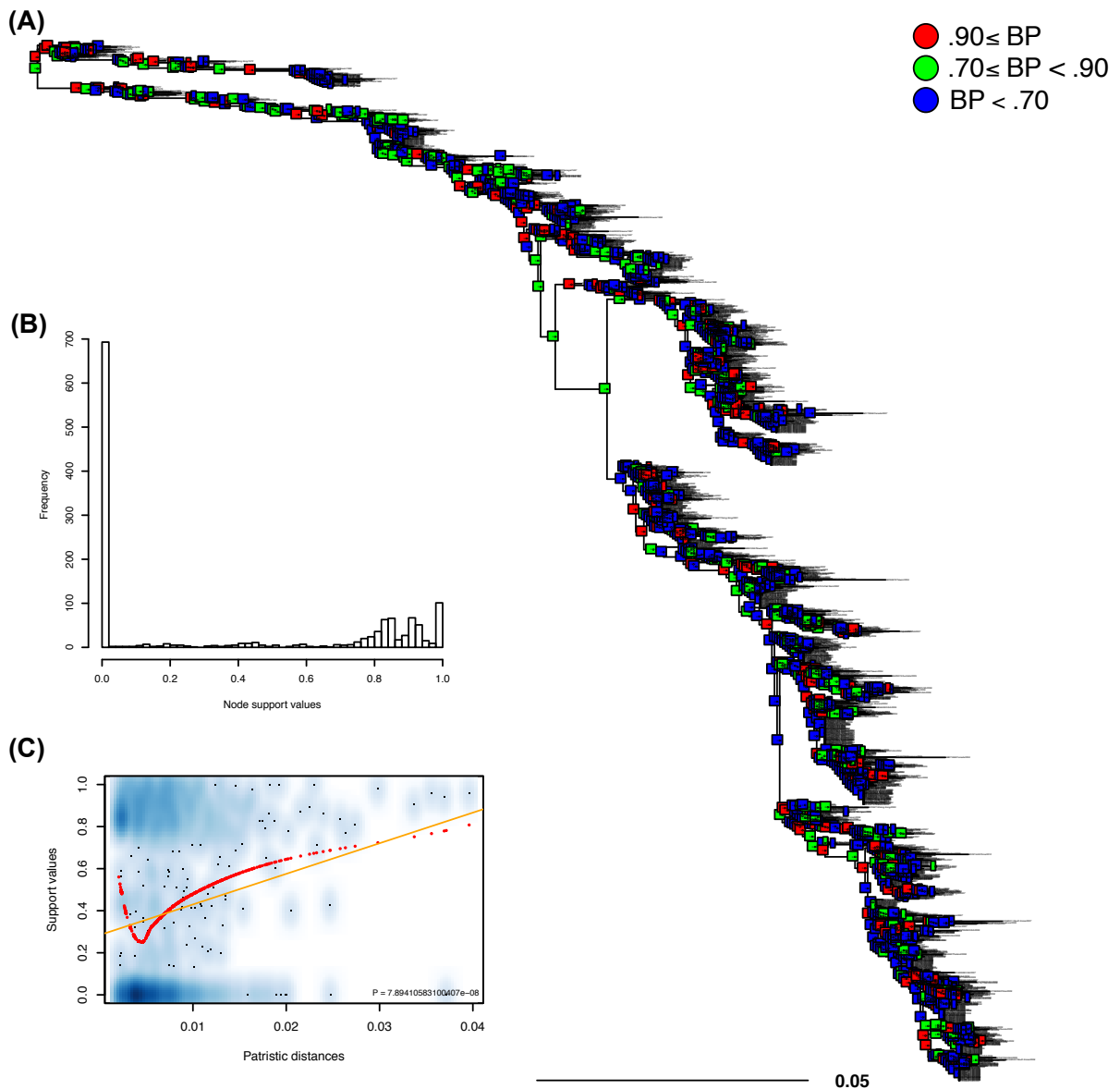


Figure S11: **Node support values for the phylogenetic analysis of the KDBP11 N2 data set.** (A) The estimated maximum likelihood phylogenetic tree. Support values (SH-like aLRT) are color-coded as shown in the top right corner of the panel. Scale bar is in expected number of substitutions per site. (B) The distribution of node support values over the entire tree. (C) Relationship between support values and patristic distance for each pair of sequence. The significance of fitted linear regression (orange line) is shown in the bottom right corner of the panel. The loess (locally estimated scatterplot smoothing) is shown in red.

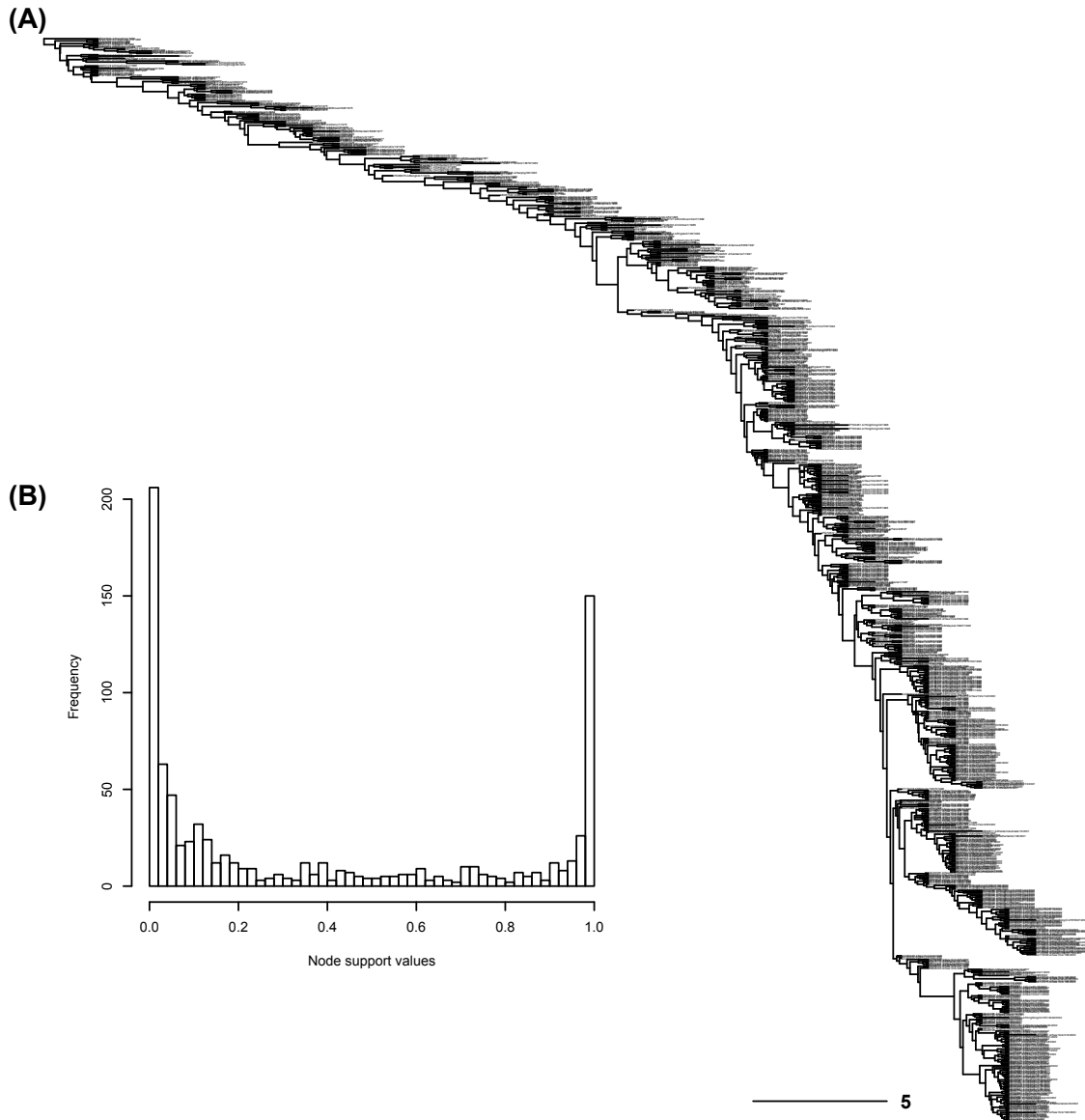


Figure S12: **Node support values for the phylogenetic analysis of the Koel13HA data set.** (A) The estimated maximum a posteriori phylogenetic tree from the BEAST relaxed clock analysis. Scale bar is in years. (B) The distribution of node support values (posterior distributions) over the entire tree.

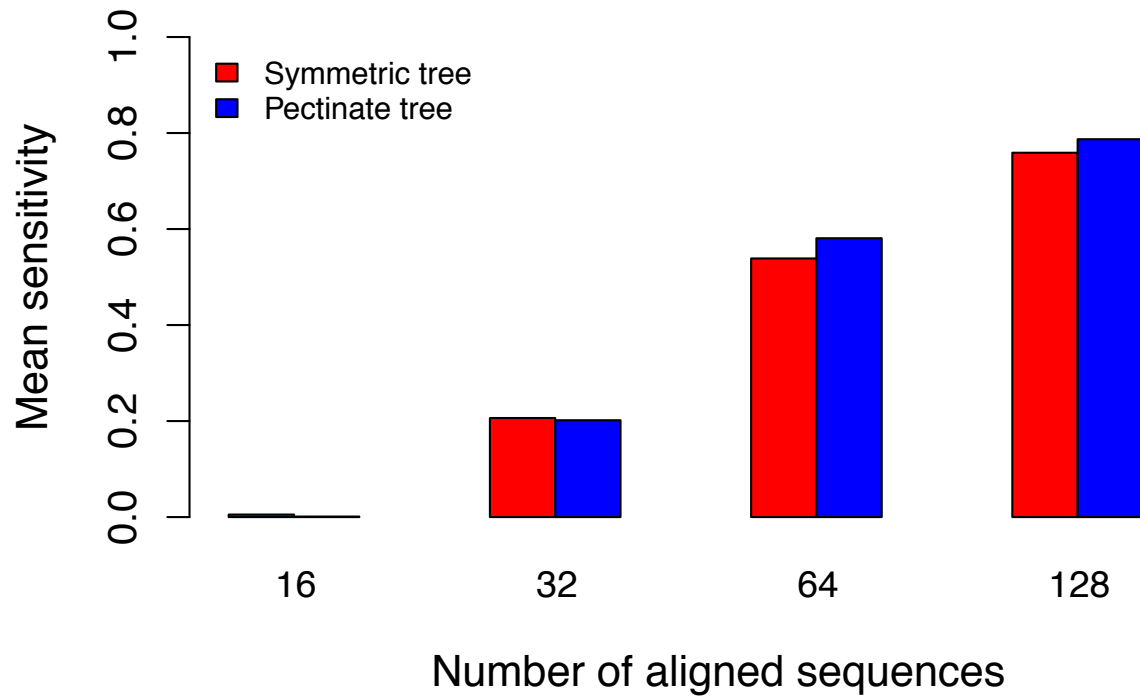


Figure S13: **The sensitivity to detect epistasis based on tree shape and number of aligned sequences.** The figure presents the mean sensitivity across all analyses branch lengths. Since sensitivity varies for each branch lengths tested, errors bars are not presented as they distract from the trends. From these results we decided to ignore alignments with fewer than 16 sequences.

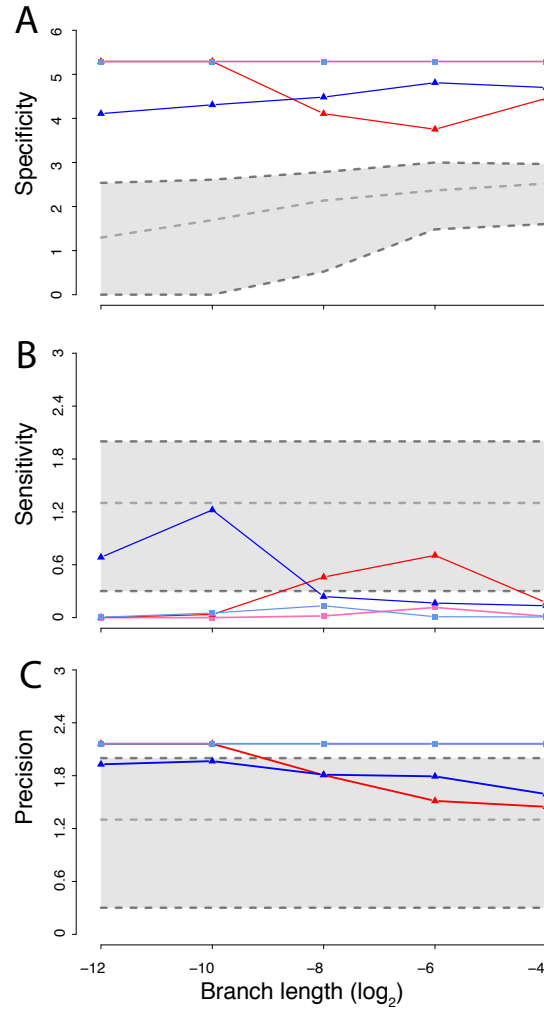


Figure S14: **Performance of our novel epistasis detection method, when dinucleotide bias is included in the simulation of independently evolving sites.** This figure is to be compared with Figure 1 (main text). Two tree shapes, symmetrically bifurcating (red) and pectinate (blue), were used by varying branch lengths (x -axis, on a \log_2 scale) and the number of aligned sequences (square: 32, and triangles: 128 sequences). All y -axes show the mean of $-\log_{10}(1 - \text{summary statistic})$ to highlight performance of our method as values approach unity. When the mean of a summary statistic was unity (meaning our $\log_{10}(1 - \text{summary statistic})$ was infinity), we arbitrarily assigned it a value 10% larger than the largest value in the set. Panel A shows specificity (true negative rate); the polygon shaded in gray shows thresholds for excellent specificity. The thresholds were established by subtracting three (the number of epistatic pairs simulated) from the minimum (lower dashed line), mean (middle dashed line) and maximum (upper dashed line) number of pairwise comparisons performed across all simulations with the same branch length. These thresholds represent the number of calculable true negatives and allow us to demonstrate our method's excellent specificity. Panels B and C show sensitivity (true positive rate) and precision (positive predictive value) respectively. Polygons show the thresholds for 50% (lower dashed line), 95% (middle dashed line) and 99% (upper dashed line) detection.

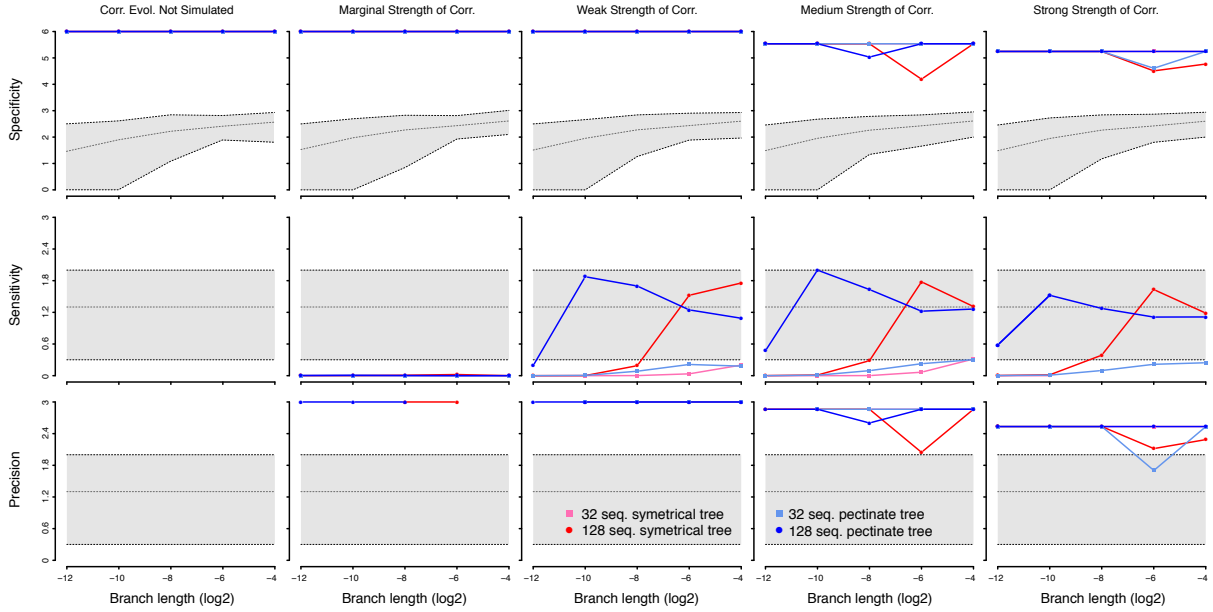


Figure S15: **Specificity, sensitivity and precision results of our novel epistasis detection method, when epistasis was simulated with Coev.** This figure is to be compared with Figure 1 (main text). Results for five levels of correlation are shown with $d/s \in \{1, 2, 33, 66, 100\}$. These levels represent no correlated evolution, marginal, weak, medium, and strong correlations, respectively. Our strong value ($d/s = 100$) was based on the default of Coev-web [6]. Two tree shapes, symmetrically bifurcating (red) and pectinate (blue), were used by varying branch lengths (x -axis, on a \log_2 scale) and the number of aligned sequences (square: 32, circles: 64 and triangles: 128 sequences). All y -axes show the mean of $-\log_{10}(1 - \text{summary statistic})$ to highlight performance of our method as values approach unity. When the mean of a summary statistic was unity (meaning our $\log_{10}(1 - \text{summary statistic})$ was infinity), we arbitrarily assigned it a value 10% larger than the largest finite value in the set. Panel (A) shows specificity (true negative rate); the polygon shaded in gray shows thresholds for excellent specificity. The thresholds were established by subtracting three (the number of epistatic pairs simulated) from the minimum (lower dashed line), mean (middle dashed line) and maximum (upper dashed line) number of pairwise comparisons performed across all simulations with the same branch length. These thresholds represent the number of calculable true negatives and allow us to demonstrate our method's excellent specificity. Panels (B) and (C) show sensitivity (true positive rate) and precision (positive predictive value) respectively. Polygons show the thresholds for 50% (lower dashed line), 95% (middle dashed line) and 99% (upper dashed line) detection.

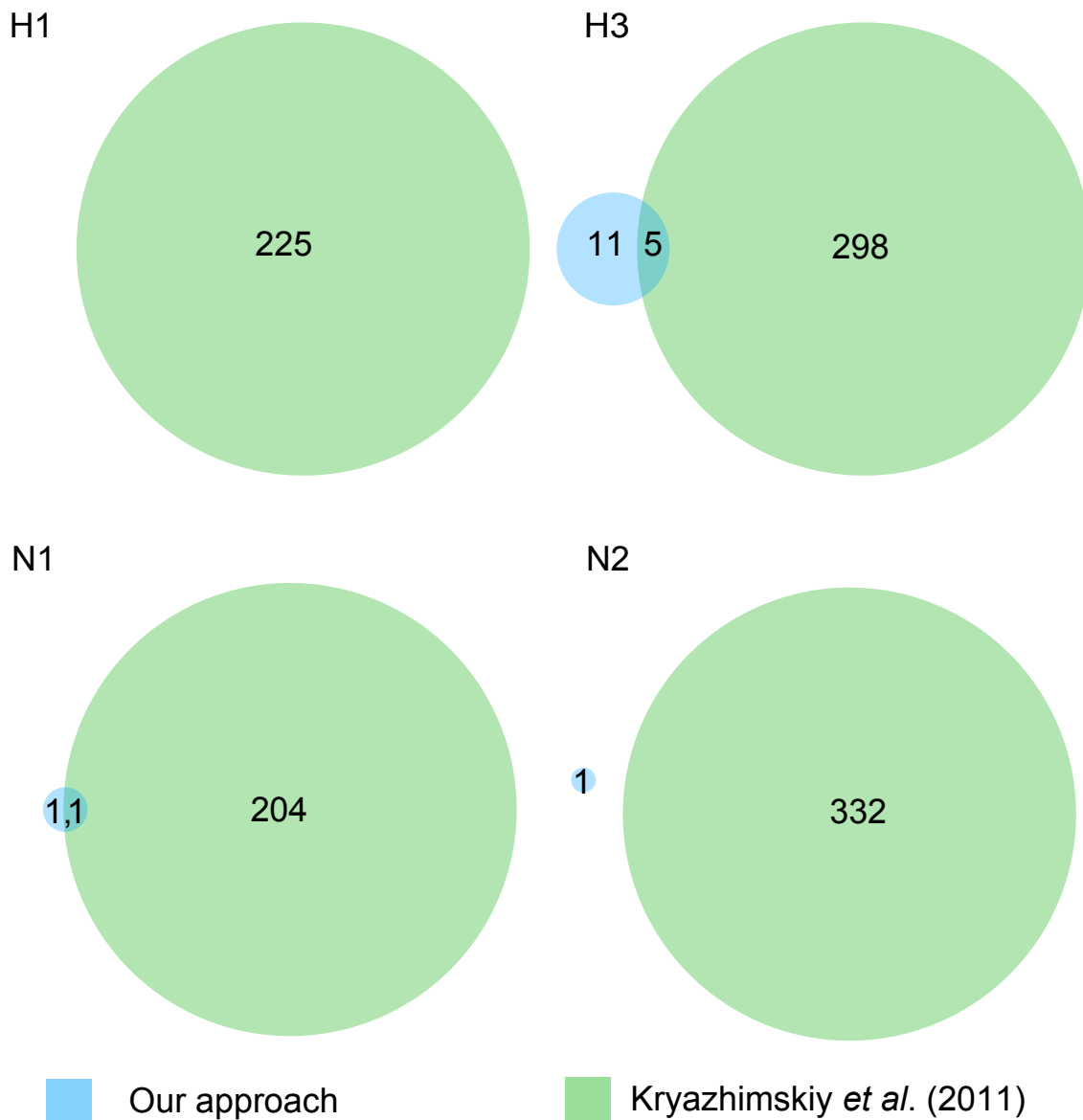


Figure S16: **Comparison with the method by Kryazhimskiy *et al.* (2011): after FDR.** The numbers in the Venn diagrams show the total of pairs of epistatic sites detected by our method (in blue) and the previous method (in green) using the four large KDBP11 data sets. The diameter of each circle is indicative of the number of pairs of epistatic sites detected by each method.

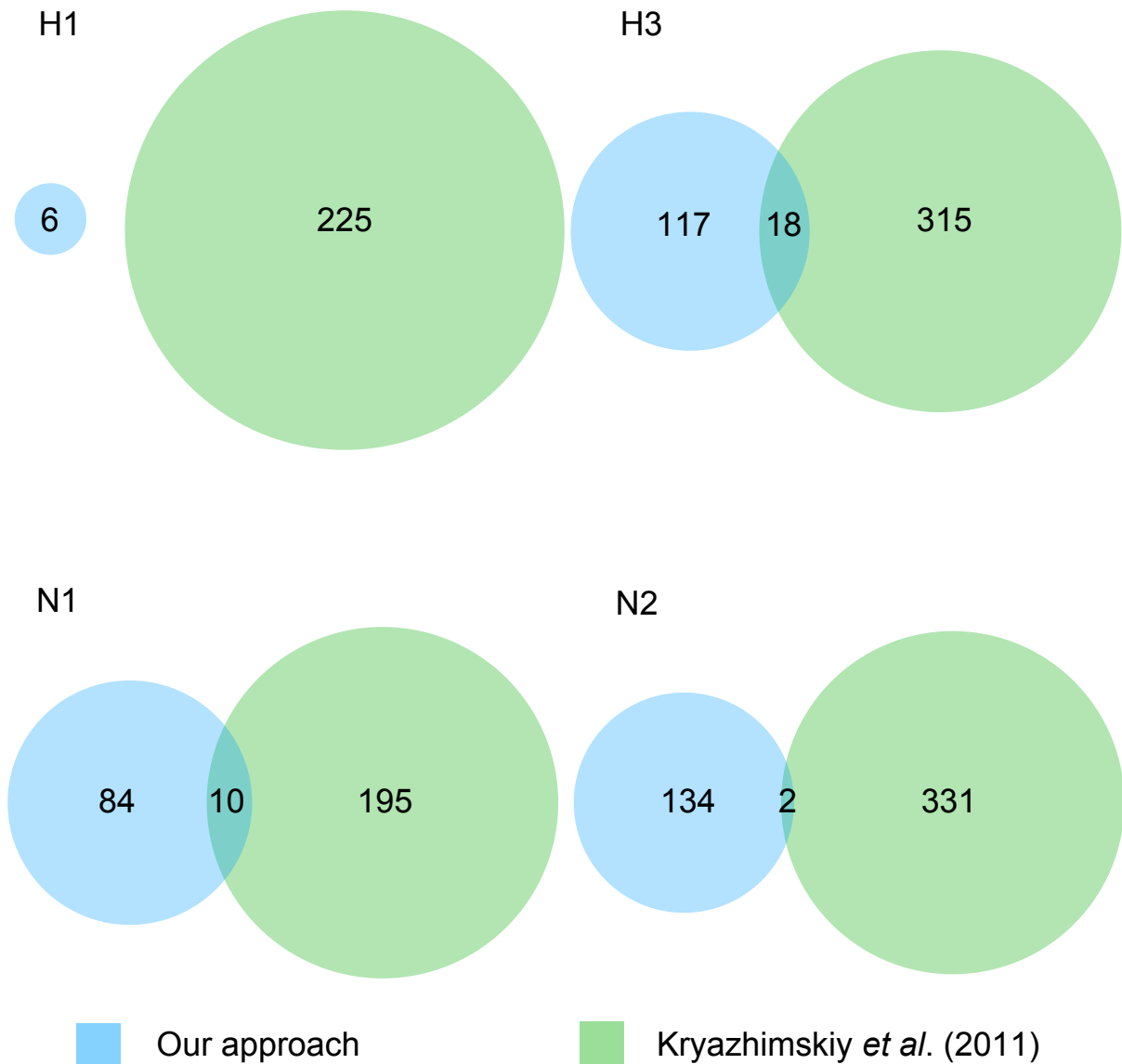


Figure S17: **Comparison with the method by Kryazhimskiy *et al.* (2011): before FDR.** The numbers in the Venn diagrams show the total of pairs of epistatic sites detected by our method (in blue) and the previous method (in green) using the four large KDBP11 data sets. The diameter of each circle is indicative of the number of pairs of epistatic sites detected by each method.

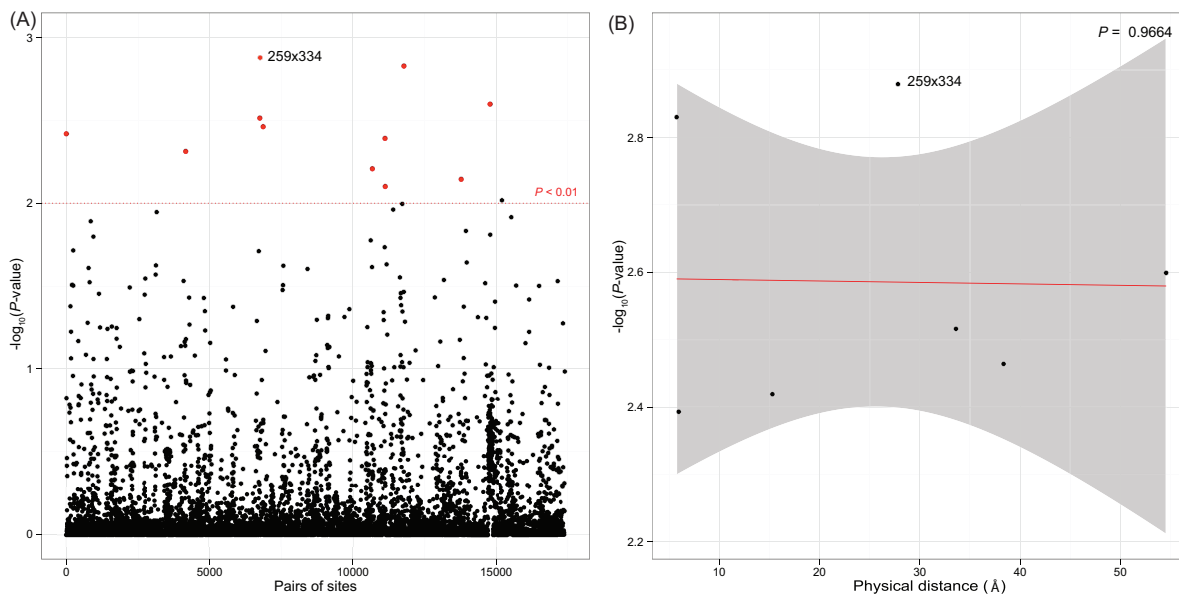


Figure S18: **The correlated sites detected in human influenza nucleoprotein and the physical distance among them.** (A) Pairs of AA sites analyzed using the Gong13NP data set. Each pair is indexed using an arbitrary ordering. Pair of sites detected as epistatic with a $P < 0.01$ before FDR are shown in red. (B) The strength of the association is not related to physical distance (in Å) between sites of each epistatic pair. The significance of the regression (slope) is shown in the top right corner of the panel.

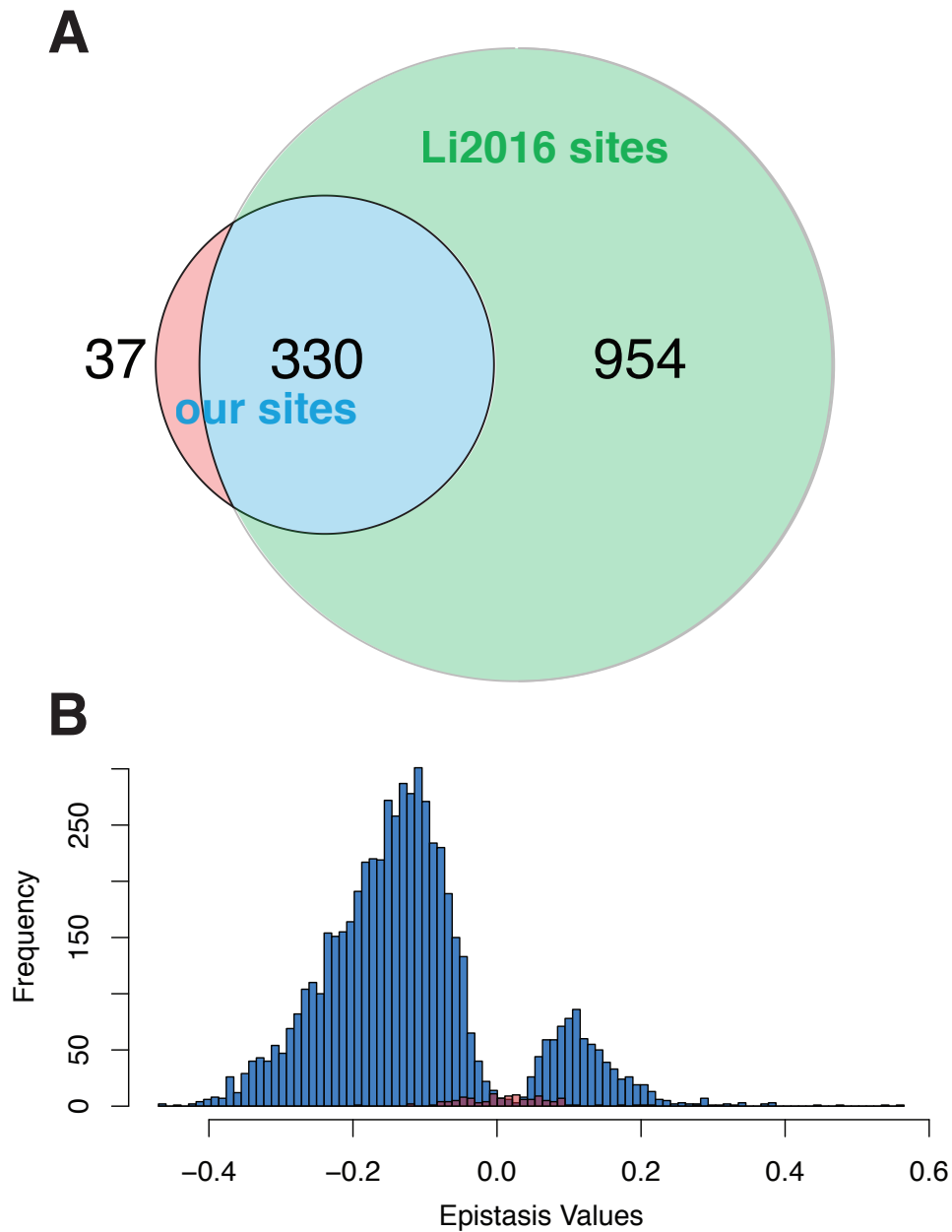


Figure S19: **Comparison of significant site pairs identified between our approach and Li *et al.* (2016).** We analyzed an alignment of eukaryotic tRNA^{ArgCCT} genes and compared our predictions to the empirical evidence in [7]. (A) Comparison of significant site pairs detected by [7] and our analysis. (B) The distribution of epistasis values from all biological replicates in [7] which resulted in significant epistasis. Bars in red are for biological replicates belonging to the 37 pairs of sites that we (but not [7]) found to be significantly correlated. Significance values for both analyses used $\alpha = 0.05$.

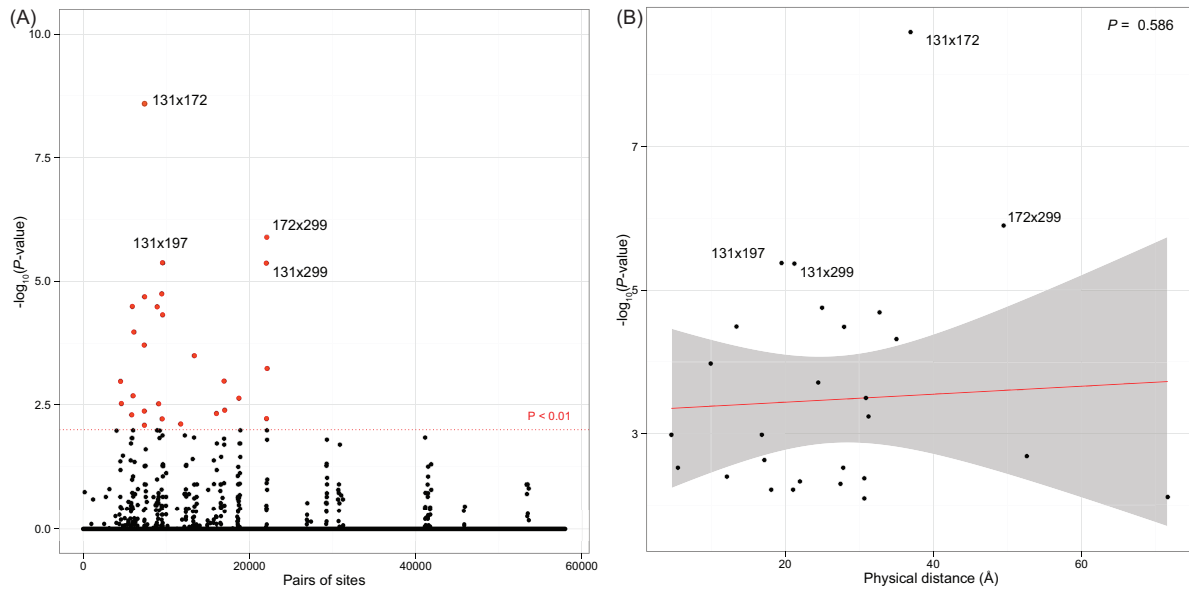


Figure S20: **The interacting sites detected in human influenza hemagglutinin and the physical distance among them.** (A) Pairs of AA sites analyzed using the Koel13HA data set. Each pair is indexed using an arbitrary ordering. Pair of sites detected as epistatic with a $P < 0.01$ after FDR are shown in red. (B) The strength of the association is not related to physical distance (in Å) between sites of each epistatic pair. The significance of the regression (slope) is shown in the top right corner of the panel.

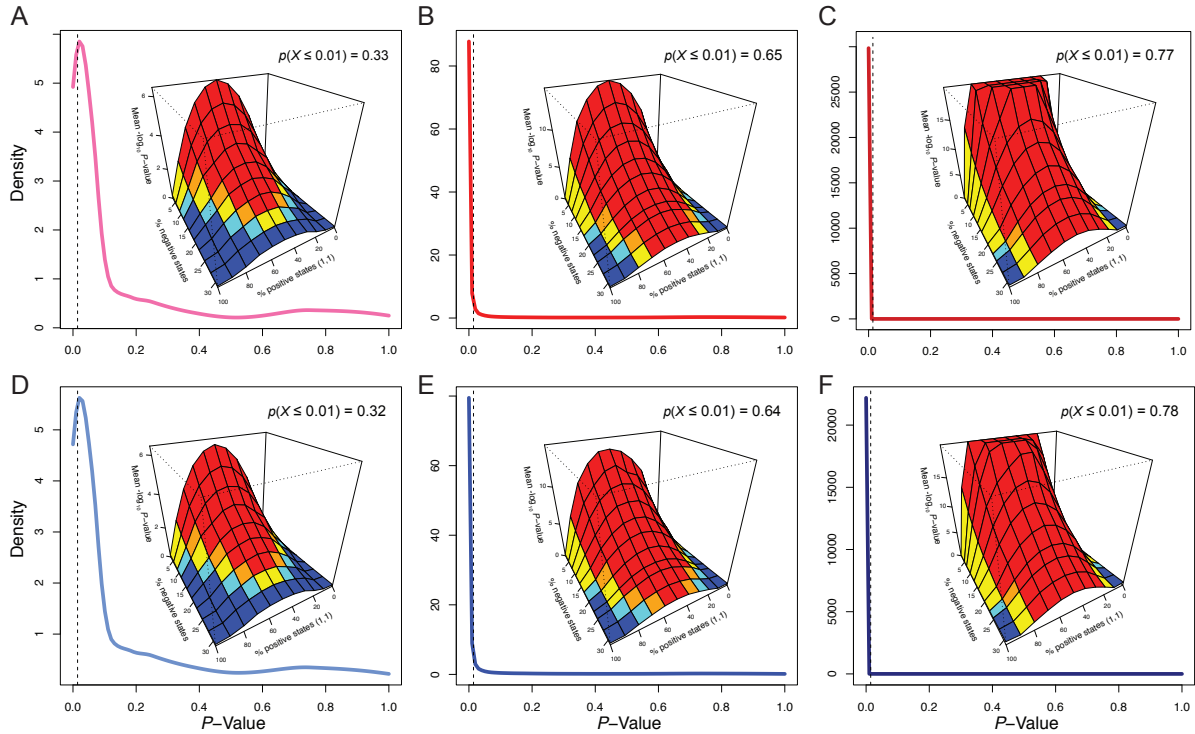


Figure S21: **Optimal signal strength for detecting epistasis.** Two tree shapes, symmetrically bifurcating (panels A, B, C) and pectinate (panels D, E, F), were studied for alignments of 32 (panels A, D), 64 (panels B, E) and 128 (panels C, F) sequences. In each panel, the main plot shows the density of P -values; the dashed vertical black line shows the significance cutoff; the proportion of significant P -values is shown above each inset. Inset plots show significance ($-\log_{10} P$) plotted as a function of the proportion of identical positively correlated states and the proportion of total sequences that bear negatively correlated states. The proportion of significant results is color-coded: red: 100%; orange: 75–100%; yellow: 50–75%; light blue: 25–50%; dark blue: 0–25%.

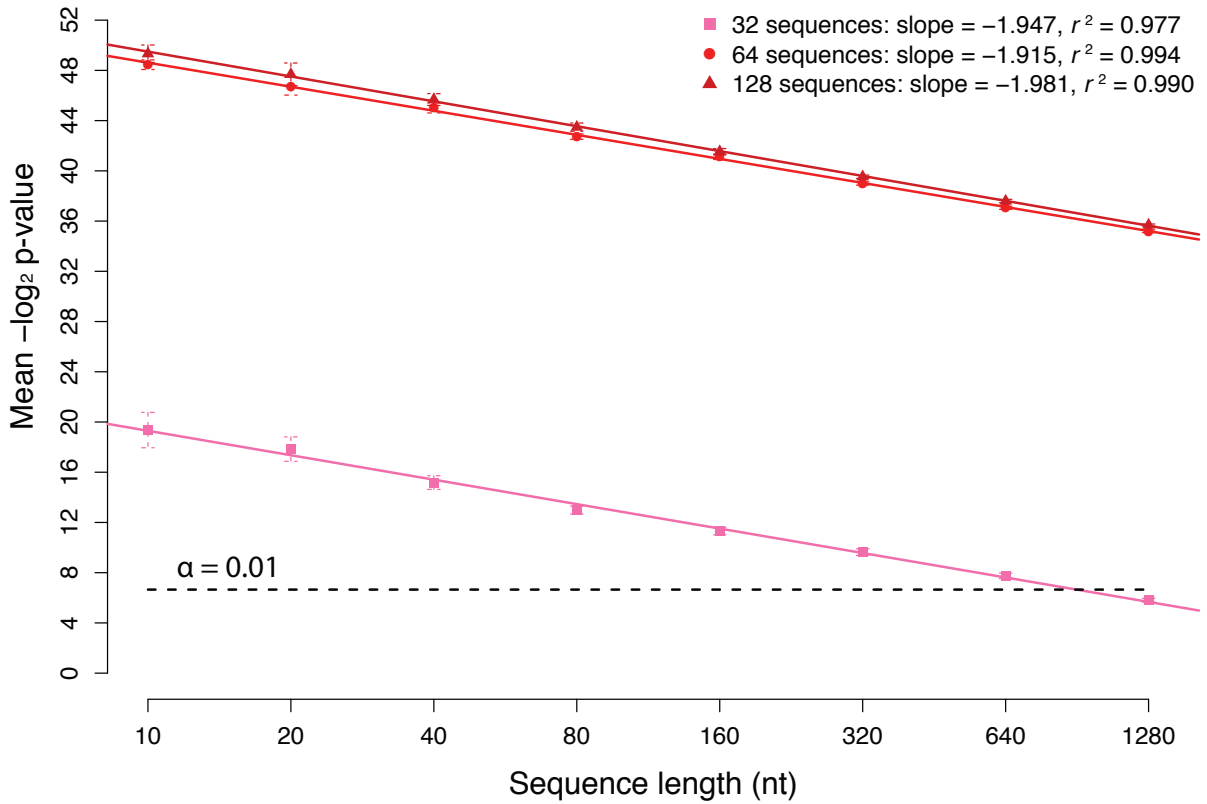


Figure S22: **The impact on detection of correlated evolution after false discovery rate (FDR) correction.** This plot illustrates the relationship between the q -value (FDR-corrected P -value) and length of the sequence alignment being analyzed with our method. This analysis was done for the symmetrically bifurcating tree shape with 32, 64 and 128 sequences. The simulated epistatic signal was generated to match the most significant sequence distribution previously identified. The black horizontal dashed line represents a significance level of 0.01, below which results are no longer significant. Slopes and r^2 values of linear models are shown; all linear fits are highly significant.

Supplementary tables

Table S1: **ANOVA tables for the analysis of specificity, sensitivity, optimal phylogenetic distribution of epistatic pairs and sequence length.** *P*-values in bold are significant. As design was fully factorial all possible interactions (colon notation) among factors were tested and are presented here.

	Df	Sum Sq	Mean Sq	<i>F</i> value	Pr(> <i>F</i>)
Branch length specificity					
τ	1	1.000E-09	9.810E-10	1.015	0.31373
n_s	1	2.000E-08	2.003E-08	20.736	5.30E-06
b	1	1.000E-09	1.325E-09	1.372	0.24150
$\tau:n_s$	1	1.000E-09	1.185E-09	1.226	0.26814
$\tau:b$	1	7.000E-09	7.434E-09	7.696	0.00554
$n_s:b$	1	4.000E-09	4.002E-09	4.143	0.04183
$\tau:n_s:b$	1	2.100E-08	2.073E-08	21.460	3.64E-06
Residuals	18649	1.802E-05	9.660E-01		
Branch length sensitivity					
τ	1	0.3	0.3	4.151	0.041635
n_s	1	733.1	733.1	9224.143	<2.23E-308
b	1	17.2	17.2	216.730	<2.23E-308
$\tau:n_s$	1	1.7	1.7	21.883	2.94E-06
$\tau:b$	1	0.9	0.9	11.830	0.000585
$n_s:b$	1	0.0	0.0	0.000	0.991888
$\tau:n_s:b$	1	0.4	0.4	4.729	0.029681
Residuals	9395	746.7	0.1		
Optimal phylogenetic distribution					
τ	1	36.36	36.36	2,208.72	<2.23E-308
n_s	1	5,232.70	5,232.70	317,874.80	<2.23E-308
+prop	10	130,007.73	13,000.77	789,768.31	<2.23E-308
-perc	10	5,672.48	567.25	34,459.10	<2.23E-308
$\tau:n_s$	1	0.47	0.47	28.47	9.50E-08
τ :+prop	10	284.65	28.46	1,729.18	<2.23E-308
n_s :+prop	10	3,511.84	351.18	21,333.64	<2.23E-308
τ :-perc	10	122.51	12.25	744.20	<2.23E-308
n_s :-perc	10	1,052.12	105.21	6,391.40	<2.23E-308
+prop:-perc	100	8,152.99	81.53	4,952.76	<2.23E-308
$\tau:n_s$:+prop	10	4.27	0.43	25.95	5.52E-50
$\tau:n_s$:-perc	10	0.61	0.06	3.72	5.24E-05
τ :+prop:-perc	100	524.02	5.24	318.33	<2.23E-308
n_s :+prop:-perc	100	1,741.81	17.42	1,058.11	<2.23E-308
$\tau:n_s$:+prop:-perc	100	17.02	0.17	10.34	5.13E-155
Residuals	2092516	34,445.96	0.02		
Effect of sequence length					
n_s	1	0.00	0.00	58.99	4.20E-13
l_s	1	0.00	0.00	174.58	3.32E-30
$n_s:l_s$	1	0.00	0.00	199.52	3.04E-33
Residuals	236	0.00	0.00		

Notes— τ : tree shape (symmetric or pectinate); n_s : number of aligned sequences; b : branch length; +prop: balance of one positively correlated state; -perc: percentage of total taxa with negatively correlated state pairs. l_s : sequence length.

Table S2: **Type III Sum of Squares and F tests for factors studied in our analysis of the phylogenetic distribution of epistatic pairs.** A fully factorial experimental design was used. Likely due to the size of the data set all terms are significant (P -values in bold) so interpretation focuses on the F values to seek out terms that explain more variance.

	Df	Sum of Sq	RSS	AIC	F value	Pr(> F)
τ	1	34.44	34,480.40	-8,592,813.91	2,092.30	< 2.23E-308
n_s	1	42.86	34,488.81	-8,592,303.24	2,603.43	< 2.23E-308
+prop	10	1,114.63	35,560.59	-8,528,269.41	6,771.13	< 2.23E-308
-perc	10	1.08	34,447.04	-8,594,857.90	6.57	2.93E-10
$\tau:n_s$	1	0.29	34,446.25	-8,594,887.78	17.87	2.36E-05
τ :+prop	10	54.52	34,500.48	-8,591,613.49	331.20	< 2.23E-308
n_s :+prop	10	58.45	34,504.41	-8,591,375.04	355.08	< 2.23E-308
τ :-perc	10	57.70	34,503.66	-8,591,420.69	350.51	< 2.23E-308
n_s :-perc	10	31.95	34,477.91	-8,592,983.25	194.09	< 2.23E-308
+prop:-perc	100	284.16	34,730.12	-8,577,908.39	172.62	< 2.23E-308
$\tau:n_s$:+prop	10	6.96	34,452.92	-8,594,500.91	42.27	1.42E-84
$\tau:n_s$:-perc	10	2.13	34,448.09	-8,594,793.99	12.96	5.56E-23
τ :+prop:-perc	100	69.43	34,515.39	-8,590,889.13	42.18	< 2.23E-308
n_s :+prop:-perc	100	147.44	34,593.40	-8,586,163.99	89.57	< 2.23E-308
$\tau:n_s$:+prop:-perc	100	17.02	34,462.98	-8,594,069.54	10.34	5.13E-155

Notes— τ : tree shape (symmetric or pectinate); n_s : number of aligned sequences; +prop: balance of one positively correlated state; -perc: percentage of total taxa with negatively correlated state pairs.

Table S3: **Type III Sum of Squares and F tests for differences in sensitivity between our full simulation study (main text), and dinucleotide-biased simulations.** Both simulation studies used the same model (RNA7D) to simulated epistatic sites, thus a difference was not expected. Significance was found in a subset of our original simulation study when parameters, other than model for independent site evolution, were identical.

	Sum Sq	Df	F value	$\text{Pr}(> F)$
(Intercept)	0.09	1	1.0948	0.29548
Independent Site Model	29.09	1	336.8009	< 2.2E-16
τ	0.24	1	2.7449	0.09764
b	0.08	1	0.8956	0.34402
n_s	1.53	1	17.7263	2.609E-05
$\tau:b$	0.01	1	0.1693	0.68077
$\tau:n_s$	7.62	1	88.2200	< 2.2E-16
$b:n_s$	3.12	1	36.0820	2.066E-09
$\tau:b:n_s$	3.54	1	40.9376	1.759E-10
Residuals	335.48	3884		

Notes— τ : tree shape (symmetric or pectinate); n_s : number of aligned sequences; b : branch length.

Table S4: Detailed results generated using the KDBP11 data set for HA in H1N1 viruses. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The P -values are computed using the standard χ^2 approximation. FDR-corrected P -values follow BH. Only pairs significant at the nominal 1% level (P -value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	P -value	P -value (FDR)
71	73	-38.485092	-29.055355	18.859474	0.000837511178384709	0.984913145780418
12	73	-58.238767	-49.80791	16.861714	0.00205625270630705	0.999999998318065
73	74	-56.071941	-48.042412	16.059058	0.00294093184608624	0.999999998318065
11	73	-35.737211	-28.307707	14.859008	0.00500275725669519	0.999999998318065
41&45&54&84	65	-23.662771	-16.698284	13.928974	0.00752522577487336	0.999999998318065
73	88	-29.379655	-22.670017	13.419276	0.00939886101619647	0.999999998318065

Table S5: Detailed results generated using the KDBP11 data set for NA in H1N1 viruses. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The P -values are computed using the standard χ^2 approximation. FDR-corrected P -values follow BH. Only pairs significant at the nominal 1% level (P -value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	P -value	P -value (FDR)
275	354	-151.206402	-122.116619	58.179566	7.00E-12	2.34E-07
461	466	-80.750412	-63.80729	33.886244	7.86E-07	0.0131356962707871
234	275	-140.234912	-128.581571	23.306682	0.00010995716148976	0.9999999999955
50	286	-56.127815	-45.461299	21.333032	0.000271976976207289	0.9999999999955
200	388	-45.296802	-35.134716	20.324172	0.000430930754599101	0.9999999999955
254	270	-64.855428	-54.755728	20.1994	0.000456107550876772	0.9999999999955
249	275	-155.496669	-145.664956	19.663426	0.000581886713368229	0.9999999999955
336	427	-55.900157	-46.137891	19.524532	0.000619734205504541	0.9999999999955
70	466	-51.543637	-41.860559	19.366156	0.000665869398006613	0.9999999999955
435	461	-85.07601	-75.485225	19.18157	0.000723940959468261	0.9999999999955
136	270	-65.039998	-55.451621	19.176754	0.000725521312396338	0.9999999999955
23	274	-71.48771	-61.947853	19.079714	0.000758101690989998	0.9999999999955
34	336	-41.971628	-32.49925	18.944756	0.00080583070915563	0.9999999999955
200	336	-32.754213	-23.344507	18.819412	0.000852815473945556	0.9999999999955
52	466	-49.232332	-40.012794	18.439076	0.00101259734365267	0.9999999999955
105	466	-41.824315	-32.72081	18.20701	0.00112426857912984	0.9999999999955
263	336	-61.402675	-52.397998	18.009354	0.00122891413529491	0.9999999999955
357	383	-36.893395	-27.896693	17.993404	0.00123776646505713	0.9999999999955
200	274	-26.322141	-17.541908	17.560466	0.00150363511573715	0.9999999999955
336	388	-58.414566	-49.684818	17.459496	0.00157333138969318	0.9999999999955
45	275	-140.885071	-132.221152	17.327838	0.00166901818191334	0.9999999999955
220	336	-55.603238	-46.94017	17.326136	0.00167029202285462	0.9999999999955
200	466	-45.927375	-37.268853	17.317044	0.00167711308491059	0.9999999999955
270	396	-86.099665	-77.448379	17.302572	0.00168802709857463	0.9999999999955
23	452	-111.291565	-102.649438	17.284254	0.00170194209259467	0.9999999999955
248	466	-118.425817	-109.808757	17.23412	0.00174060696636791	0.9999999999955
383	466	-46.488327	-37.899067	17.17852	0.00178450131560515	0.9999999999955
93	275	-131.342994	-122.778536	17.128916	0.00182458253941953	0.9999999999955
275	396	-150.982066	-142.567216	16.8297	0.00208588559096545	0.9999999999955
23	173	-72.887506	-64.473463	16.828086	0.00208739064264929	0.9999999999955
334	466	-53.692311	-45.366264	16.652094	0.00225809043765124	0.9999999999955
143	467	-22.125427	-13.864434	16.521986	0.00239306161668063	0.9999999999955
382	466	-40.933061	-32.681373	16.503376	0.00241300606556916	0.9999999999955
332	467	-13.418112	-5.178711	16.478802	0.00243959327709231	0.9999999999955
173	275	-121.008318	-112.792149	16.432338	0.00249065498094114	0.9999999999955
52	461	-57.912071	-49.789049	16.246044	0.00270617009931429	0.9999999999955
390	467	-27.201104	-19.132944	16.13632	0.00284159256104222	0.9999999999955
454	466	-58.950808	-50.92717	16.047276	0.00295637796235626	0.9999999999955
34	287	-36.310054	-28.313782	15.992544	0.00302918482382142	0.9999999999955

Continued on next page

Table S4 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
86	467	-27.507537	-19.537999	15.939076	0.00310201359905926	0.999999999955
105	461	-50.425315	-42.465231	15.920168	0.00312817848327285	0.999999999955
23	220	-87.652607	-79.760612	15.78399	0.00332314529606204	0.999999999955
15	200	-40.968346	-33.088335	15.760022	0.00335867675914581	0.999999999955
130	331	-68.45982	-60.657796	15.604048	0.00359920495061494	0.999999999955
383	461	-55.168084	-47.384994	15.56618	0.0036601113151743	0.999999999955
388	466	-71.514138	-63.817153	15.39397	0.00395012290198515	0.999999999955
15	274	-47.6502	-39.963419	15.373562	0.00398595416950454	0.999999999955
17&163	466	-39.08149	-31.412476	15.338028	0.00404910689811722	0.999999999955
336	366&367	-27.081351	-19.422294	15.318114	0.00408492800602434	0.999999999955
232	467	-32.514391	-24.881063	15.266656	0.00417893881066467	0.999999999955
16	466	-58.934043	-51.335155	15.197776	0.00430811378552565	0.999999999955
334	461	-62.37204	-54.783778	15.176524	0.00434875529997469	0.999999999955
254	336	-49.567648	-41.996368	15.14256	0.0044144902055856	0.999999999955
101	336	-35.215942	-27.696958	15.037968	0.00462311335217314	0.999999999955
57	466	-66.016384	-58.50624	15.020288	0.0046593229258225	0.999999999955
82	130	-71.473515	-63.99239	14.96225	0.00478015697097867	0.999999999955
262	274	-49.90798	-42.455595	14.90477	0.00490286108222038	0.999999999955
307	466	-69.509585	-62.078536	14.862098	0.00499594947834114	0.999999999955
23	275	-158.091078	-150.690456	14.801244	0.00513171047589567	0.999999999955
275	352	-137.045128	-129.652756	14.784744	0.00516914075797925	0.999999999955
14	336	-35.207757	-27.822389	14.770736	0.00520112818749763	0.999999999955
200	250	-31.51041	-24.131537	14.757746	0.0052309646028319	0.999999999955
84	287	-51.4864	-44.112225	14.74835	0.00525265069975567	0.999999999955
95	396	-86.471666	-79.157217	14.628898	0.00553616002222657	0.999999999955
106	466	-54.834282	-47.543352	14.58186	0.00565187543337498	0.999999999955
70	461	-60.223376	-52.946299	14.554154	0.00572114045480998	0.999999999955
15	388	-66.626424	-59.386456	14.479936	0.00591081509238112	0.999999999955
34	386	-34.010318	-26.778058	14.46452	0.00595097969768998	0.999999999955
211	274	-28.759868	-21.562656	14.394424	0.00613700829742769	0.999999999955
270	434	-62.234339	-55.133788	14.201102	0.00668012274939944	0.999999999955
336	418	-47.091265	-40.052959	14.076612	0.00705451363202247	0.999999999955
275	418	-127.272988	-120.235474	14.075028	0.00705940740052502	0.999999999955
15	262	-64.553026	-57.518177	14.069698	0.00707589887701154	0.999999999955
15	250	-52.840018	-45.970497	13.739042	0.00817623254556632	0.999999999955
427	466	-68.999729	-62.132703	13.734052	0.0081940589981907	0.999999999955
52	105	-18.865615	-12.01964	13.69195	0.00834598796723063	0.999999999955
23	336	-77.920897	-71.082516	13.676762	0.00840146912942918	0.999999999955
382	461	-49.517328	-42.720275	13.594106	0.00870979383594128	0.999999999955
173	274	-34.404951	-27.611158	13.587586	0.00873458058217291	0.999999999955
220	466	-68.70281	-61.910932	13.583756	0.00874917305685707	0.999999999955
79	201	-56.900371	-50.132123	13.536496	0.00893121045800649	0.999999999955
27	200	-17.748473	-10.982961	13.531024	0.00895252549151837	0.999999999955
95	275	-141.70149	-134.940735	13.52151	0.00898970381832009	0.999999999955
95	270	-76.81909	-70.063706	13.510768	0.00903186238930109	0.999999999955
287	466	-53.307035	-46.564777	13.484516	0.00913570767452088	0.999999999955

Continued on next page

Table S4 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
34	250	-40.728247	-33.988081	13.480332	0.00915236589278601	0.9999999999955
220	386	-47.628595	-40.892513	13.472164	0.00918497169373855	0.9999999999955
64	275	-111.303341	-104.594838	13.417006	0.0094081495237065	0.9999999999955
6	45	-54.820441	-48.117229	13.406424	0.00945156809500369	0.9999999999955
93	95	-66.832594	-60.144784	13.37562	0.00957907695899363	0.9999999999955
23	254	-81.617017	-74.937806	13.358422	0.00965099565284477	0.9999999999955
130	344	-88.20565	-81.54131	13.32868	0.00977661863990886	0.9999999999955
101	466	-48.315515	-41.66894	13.29315	0.0099287850007318	0.9999999999955
14	466	-48.307329	-41.663856	13.286946	0.00995559156861858	0.9999999999955

Table S6: Detailed results generated using the KDBP11 data set for HA in H3N2 viruses. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The *P*-values are computed using the standard χ^2 approximation. FDR-corrected *P*-values follow BH. Only pairs significant at the nominal 1% level (*P*-value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
172	212	-95.260701	-66.640486	57.24043	1.10E-11	7.77E-07
172	278	-86.257381	-59.220203	54.074356	5.08E-11	1.79E-06
172	213	-102.53157	-76.376144	52.310852	1.19E-10	2.77E-06
172	242	-90.576939	-64.712052	51.729774	1.57E-10	2.77E-06
172	206	-77.13814	-52.08129	50.1137	3.42E-10	4.82E-06
172	292	-79.869861	-55.371555	48.996612	5.85E-10	6.87E-06
149	172	-142.326561	-118.68048	47.292162	1.33E-09	1.33E-05
137	172	-98.903071	-75.740267	46.325608	2.11E-09	1.86E-05
158	172	-126.00687	-105.252318	41.509104	2.11E-08	0.000165161699520559
172	315	-89.372521	-69.349836	40.04537	4.24E-08	0.00027882614746233
98	172	-82.689173	-62.746316	39.885714	4.57E-08	0.00027882614746233
151	172	-87.87812	-67.97483	39.80658	4.75E-08	0.00027882614746233
172	173	-140.454483	-121.475613	37.95774	1.14E-07	0.000620009254642246
73	172	-103.779205	-85.255579	37.047252	1.76E-07	0.000886978319149443
172	208	-135.618671	-117.172459	36.892424	1.90E-07	0.000890932476549366
110	172	-75.266596	-58.856524	32.820144	1.30E-06	0.00572883065011465
140	172	-107.194851	-92.150417	30.088868	4.69E-06	0.0194692816187225
172	468	-106.731919	-93.090927	27.281984	1.74E-05	0.0682778067287818
172	202	-142.37568	-129.171471	26.408418	2.62E-05	0.0971264383410068
147	236	-133.393355	-122.281267	22.224176	0.000180842506048995	0.637469833822707
236	402	-157.755002	-146.913319	21.683366	0.000231702291790925	0.777857693869534
11	172	-85.524571	-74.819616	21.40991	0.000262583155412055	0.817110082323141
218	236	-131.09927	-120.410813	21.376914	0.000266574920474216	0.817110082323141
149	188	-99.854232	-89.369955	20.968554	0.00032124768836983	0.943665084586376

Continued on next page

Table S5 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
172	241	-86.603477	-76.275782	20.65539	0.000370578823137624	0.999999999755
236	241	-123.959968	-113.711205	20.497526	0.00039822044147797	0.999999999755
171	236	-118.15418	-107.907915	20.49253	0.000399127786084086	0.999999999755
171	245	-157.753608	-147.576446	20.354324	0.000425056094516707	0.999999999755
154	213	-197.645557	-187.505926	20.279262	0.000439830166879762	0.999999999755
171	209	-80.154649	-70.265916	19.777466	0.000552529122753942	0.999999999755
2	377	-132.82316	-123.057211	19.531898	0.000617667224629415	0.999999999755
236	238	-118.153036	-108.440276	19.42552	0.000648191457816139	0.999999999755
208	262	-151.535949	-141.930562	19.210774	0.000714430439784697	0.999999999755
2	91	-138.037954	-128.438014	19.19988	0.000717963619603745	0.999999999755
243	245	-160.870729	-151.288714	19.16403	0.000729713083036265	0.999999999755
236	243	-121.367667	-111.865979	19.003376	0.000784744795860992	0.999999999755
172	402	-120.398298	-110.939689	18.917218	0.000815928741833449	0.999999999755
153	205	-24.643147	-15.201595	18.883104	0.000828611814868485	0.999999999755
172	466	-157.397841	-147.982333	18.831016	0.00084835434424424	0.999999999755
172	238	-80.796828	-71.453865	18.685926	0.000905832076696589	0.999999999755
15	233	-70.853696	-61.536033	18.635326	0.000926769459038801	0.999999999755
172	546	-189.734083	-180.459916	18.548334	0.000963889098073767	0.999999999755
147	209	-95.395213	-86.201372	18.387682	0.00103634281682119	0.999999999755
91	236	-145.73871	-136.637576	18.202268	0.00112667292298252	0.999999999755
69	205	-121.873963	-112.777017	18.193892	0.00113093223673189	0.999999999755
236	546	-227.090789	-218.053187	18.075204	0.00119302430998547	0.999999999755
41	236	-125.973027	-116.954557	18.03694	0.00121375112605182	0.999999999755
172	175	-88.807657	-79.810745	17.993824	0.00123753255760162	0.999999999755
79	153	-33.498754	-24.656956	17.683596	0.00142275580259887	0.999999999755
15	159	-37.517793	-28.759554	17.516478	0.00153361582691103	0.999999999755
41	172	-88.616345	-79.881762	17.469166	0.0015665207568073	0.999999999755
65	122	-102.978119	-94.30307	17.350098	0.00165244594718095	0.999999999755
172	262	-153.063464	-144.390213	17.346502	0.00165511207631241	0.999999999755
79	99	-25.345169	-16.69659	17.297158	0.00169212803581087	0.999999999755
149	154	-237.440547	-228.843112	17.19487	0.00177148141516981	0.999999999755
73	108	-115.35578	-106.801142	17.109276	0.00184069638482554	0.999999999755
171	172	-80.797011	-72.251506	17.09101	0.00185580885609293	0.999999999755
91	209	-107.740568	-99.199889	17.081358	0.00186384383996996	0.999999999755
15	16&294	-36.476618	-27.980262	16.992712	0.0019392618772972	0.999999999755
69	79	-130.729512	-122.265429	16.928166	0.00199606187258083	0.999999999755
172	218	-93.743876	-85.28066	16.926432	0.00199761019179456	0.999999999755
172	391	-128.171689	-119.717041	16.909296	0.00201297504854159	0.999999999755
41	209	-87.974885	-79.53564	16.87849	0.00204089059165746	0.999999999755
16&294	153	-23.713307	-15.294666	16.837282	0.00207882969689865	0.999999999755
91	245	-184.795639	-176.381198	16.828882	0.0020866482405304	0.999999999755
241	245	-163.739881	-155.336381	16.807	0.00210715192978539	0.999999999755
16&294	69	-120.944089	-112.559717	16.768744	0.00214347647400648	0.999999999755
154	212	-190.374688	-182.016027	16.717322	0.00219327641052169	0.999999999755
2	209	-166.160652	-157.850034	16.621236	0.00228940680823264	0.999999999755
41	245	-165.029956	-156.72176	16.616392	0.00229436147719242	0.999999999755
172	188	-94.673858	-86.429946	16.487824	0.00242979879049932	0.999999999755

Continued on next page

Table S5 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
245	402	-196.811909	-188.679286	16.265246	0.00268313322906089	0.999999999755
236	391	-165.528372	-157.46435	16.128044	0.0028520733753099	0.999999999755
238	245	-158.160035	-150.114325	16.09142	0.00289891315078628	0.999999999755
236	377	-140.523917	-132.487123	16.073588	0.00292199249379554	0.999999999755
16&294	79	-24.885007	-16.863386	16.043242	0.00296168483869053	0.999999999755
2	243	-115.11334	-107.124119	15.978442	0.00304822810272354	0.999999999755
69	99	-121.404162	-113.41741	15.973504	0.00305492417351449	0.999999999755
15	79	-46.26147	-38.306894	15.909152	0.00314352243424154	0.999999999755
91	172	-108.382028	-100.430462	15.903132	0.0031519388392941	0.999999999755
99	153	-24.173396	-16.232303	15.882186	0.00318139608280876	0.999999999755
202	209	-141.73422	-133.800234	15.867972	0.0032015399706149	0.999999999755
15	153	-45.089704	-37.168746	15.841916	0.00323879264171545	0.999999999755
218	245	-170.86688	-163.033675	15.66641	0.00350106845546394	0.999999999755
15	363	-58.456346	-50.681616	15.54946	0.00368732505795455	0.999999999755
158	191	-89.091102	-81.393364	15.395476	0.00394749133007244	0.999999999755
73	391	-94.804708	-87.148318	15.31278	0.00409457552700854	0.999999999755
237	546	-195.14541	-187.5139	15.26302	0.00418566143234478	0.999999999755
147	245	-172.450284	-164.83437	15.231828	0.00424377090065453	0.999999999755
154	242	-186.387587	-178.823874	15.127426	0.00444409450263472	0.999999999755
69	142	-122.232786	-114.672781	15.12001	0.00445867239861419	0.999999999755
65	545	-130.41267	-122.858469	15.108402	0.00448158503491347	0.999999999755
209	546	-189.092623	-181.581983	15.02128	0.00465728390114206	0.999999999755
128	313	-82.258413	-74.770147	14.976532	0.00475013965260518	0.999999999755
142	153	-25.001957	-17.553857	14.8962	0.00492141870197038	0.999999999755
73	173	-107.087502	-99.705605	14.763794	0.00521705225442526	0.999999999755
15	205	-37.405868	-30.030795	14.750146	0.00524849869616506	0.999999999755
172	235	-141.653506	-134.310796	14.68542	0.00540018592006486	0.999999999755
79	205	-25.814931	-18.491374	14.647114	0.00549197516844835	0.999999999755
15	264	-36.099708	-28.795231	14.608954	0.00558493640637592	0.999999999755
15	69	-142.320462	-135.018307	14.60431	0.00559635447213735	0.999999999755
15	99	-36.936173	-29.656677	14.558992	0.00570898572169087	0.999999999755
73	188	-61.306877	-54.083756	14.446242	0.0059989478750111	0.999999999755
188	213	-60.059242	-52.839282	14.43992	0.0060156272313775	0.999999999755
172	179	-78.89473	-71.708123	14.373214	0.00619441330064496	0.999999999755
154	217	-231.698276	-224.538646	14.31926	0.00634282405485831	0.999999999755
212	262	-111.177979	-104.037812	14.280334	0.00645205687136396	0.999999999755
91	402	-91.63414	-84.524912	14.218456	0.00662950374489812	0.999999999755
209	218	-93.1015	-86.026803	14.149394	0.00683320368361695	0.999999999755
151	213	-53.263503	-46.203451	14.120104	0.0069214361698553	0.999999999755
138	223	-33.911451	-26.881617	14.059668	0.00710703496939846	0.999999999755
15	229	-43.520539	-36.496346	14.048386	0.00714221807017035	0.999999999755
360	366	-23.046363	-16.058858	13.97501	0.00737524206945139	0.999999999755
209	402	-119.756839	-112.786258	13.941162	0.00748522941595586	0.999999999755
2	238	-112.256468	-105.303071	13.906794	0.00759854979001506	0.999999999755

Continued on next page

Table S5 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
15	176	-38.773334	-31.830015	13.886638	0.0076657892646832	0.999999999755
15	179	-39.249057	-32.318012	13.86209	0.00774846751649405	0.999999999755
554	562&565	-35.843729	-28.949767	13.787924	0.00800360001575418	0.999999999755
262	468	-122.649197	-115.787454	13.723486	0.00823193099993291	0.999999999755
160	400	-34.253878	-27.405005	13.697746	0.00832491007154013	0.999999999755
312	493	-22.524909	-15.678997	13.691824	0.00834644676191187	0.999999999755
2	171	-112.289281	-105.456631	13.6653	0.00844357817895147	0.999999999755
209	243	-83.369645	-76.565585	13.60812	0.00865675023049517	0.999999999755
79	162	-36.750957	-29.966928	13.568058	0.00880923276612045	0.999999999755
209	495	-101.320696	-94.54879	13.543812	0.00890279004421735	0.999999999755
79	176	-27.181719	-20.431972	13.499494	0.00907631684629118	0.999999999755
12	557	-72.257759	-65.510564	13.49439	0.00909651270958733	0.999999999755
18	79	-34.203613	-27.511071	13.385084	0.00953972412230608	0.999999999755
391	466	-148.423343	-141.739556	13.367574	0.0096126583070183	0.999999999755
433&449	514	-22.064224	-15.401217	13.326014	0.00978795692183365	0.999999999755
79	276	-67.062134	-60.400447	13.323374	0.00979919729328571	0.999999999755
154	158	-221.120857	-214.461438	13.318838	0.00981853976584857	0.999999999755
16&294	205	-16.029383	-9.371419	13.315928	0.00983096827222008	0.999999999755
209	238	-80.155108	-73.505347	13.299522	0.00990132596967297	0.999999999755
306	403	-22.358672	-15.713962	13.28942	0.00994489331108972	0.999999999755

Table S7: Detailed results generated using the KDBP11 data set for NA in H3N2 viruses. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The *P*-values are computed using the standard χ^2 approximation. FDR-corrected *P*-values follow BH. Only pairs significant at the nominal 1% level (*P*-value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
4	5	-140.551535	-115.819373	49.464324	4.67E-10	2.52E-05
2	4	-113.963582	-97.896387	32.13439	1.80E-06	0.0348153544795633
30	215	-191.722818	-175.885065	31.675506	2.23E-06	0.0348153544795633
18	215	-141.968795	-126.28716	31.36327	2.58E-06	0.0348153544795633
42	215	-142.317227	-126.887929	30.858596	3.27E-06	0.0353044758154209
215	216	-155.86052	-140.880835	29.95937	4.99E-06	0.0448602999915356
23	215	-153.032804	-138.311892	29.441824	6.36E-06	0.0490014117590225
2	5	-141.640317	-128.095522	27.08959	1.91E-05	0.128596052265248
267	338	-111.370643	-97.965384	26.810518	2.17E-05	0.13016322352628
215	387	-150.643803	-137.556013	26.17558	2.92E-05	0.157372062936907
215	372	-175.876948	-162.935756	25.882384	3.34E-05	0.16323158527533
215	437	-159.054349	-146.202314	25.70407	3.63E-05	0.16323158527533

Continued on next page

Table S6 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
215	310	-181.656122	-169.582247	24.14775	7.46E-05	0.306265388827206
215	265	-170.896434	-158.890888	24.011092	7.95E-05	0.306265388827206
147	310	-168.960636	-157.463052	22.995168	0.000126907935161791	0.456496303305972
23	30	-114.499431	-103.25457	22.489722	0.000160093160150421	0.539874159317256
194	215	-159.70242	-148.579431	22.245978	0.000179042843320709	0.568260920836011
151	208	-298.020159	-286.967044	22.10623	0.000190894074553305	0.572215593699896
147	387	-137.948316	-127.159345	21.577942	0.000243156189340787	0.689103997241034
150	215	-153.709274	-142.974131	21.470286	0.000255431832323016	0.689103997241034
50	208	-128.316646	-117.840993	20.951306	0.000323786781901036	0.803369469548565
221	332	-97.455311	-87.072967	20.764688	0.000352563215472612	0.803369469548565
172	221	-81.930036	-71.559024	20.742024	0.000356226180617258	0.803369469548565
221	399	-106.187484	-95.819908	20.735152	0.000357344266979864	0.803369469548565
267	339	-131.381651	-121.137016	20.48927	0.000399720952128035	0.862693747720811
147	372	-163.181462	-153.039367	20.28419	0.000438844807024497	0.910704246454374
147	194	-147.006933	-136.934545	20.144776	0.000467583059149623	0.934404131091743
208	215	-181.839584	-171.88551	19.908148	0.000520686430556361	0.999999999875
155	390	-49.067913	-39.428541	19.278744	0.000692771360469524	0.999999999875
151	437	-275.23493	-265.88771	18.69444	0.000902355388088139	0.999999999875
223	234	-44.100347	-34.755491	18.689712	0.000904284433107305	0.999999999875
47	267	-99.701921	-90.428847	18.546148	0.000964840550960244	0.999999999875
151	216	-272.031728	-262.847282	18.368892	0.00104516099261065	0.999999999875
208	338	-112.919637	-103.829411	18.180452	0.00113779992965746	0.999999999875
120	267	-163.302895	-154.224436	18.156918	0.00114992491961707	0.999999999875
147	215	-217.560704	-208.519712	18.081984	0.00118938836160598	0.999999999875
82	370	-127.044629	-118.03821	18.012838	0.00122698883474071	0.999999999875
248	249	-105.019997	-96.015487	18.00902	0.00122909886439504	0.999999999875
458	459	-50.477343	-41.617396	17.719894	0.00139974359347339	0.999999999875
2	208	-124.23767	-115.415743	17.643854	0.00144837988963775	0.999999999875
267	400	-82.484163	-73.810887	17.346552	0.00165507497646944	0.999999999875
221	338	-96.858553	-88.214593	17.28792	0.00169914824242634	0.999999999875
42	151	-258.496695	-250.021015	16.95136	0.00197546496346734	0.999999999875
4	30	-133.032124	-124.591141	16.881966	0.00203772173065686	0.999999999875
460	462	-57.896555	-49.533401	16.726308	0.00218449237417284	0.999999999875
151	265	-287.076993	-278.747965	16.658056	0.00225208872620708	0.999999999875
2	30	-134.120903	-125.81302	16.615766	0.00229500254974246	0.999999999875
221	267	-115.812898	-107.51043	16.604936	0.00230612129517527	0.999999999875
367	426	-85.840734	-77.551632	16.578204	0.00233379375398302	0.999999999875
3	5	-114.796378	-106.517702	16.557352	0.00235560629205034	0.999999999875
18	151	-258.149354	-249.890235	16.5182380000001	0.00239706526466199	0.999999999875
30	151	-307.903387	-299.654305	16.498164	0.00241862109504687	0.999999999875
47	338	-80.747576	-72.549737	16.395678	0.00253168438019746	0.999999999875
150	151	-269.889838	-261.71149	16.3566959999999	0.00257604302664682	0.999999999875
328	334	-79.102208	-70.930845	16.342726	0.00259212548843057	0.999999999875
47	208	-101.250916	-93.09991	16.302012	0.00263956227260953	0.999999999875
151	338	-277.516803	-269.38117	16.271266	0.00267595081040717	0.999999999875
467	469	-50.458942	-42.381348	16.155188	0.00281783950547188	0.999999999875
82	267	-112.040302	-103.962751	16.155102	0.00281794732667073	0.999999999875
52	370	-112.243004	-104.167152	16.151704	0.0028222107706144	0.999999999875
4	42	-83.625444	-75.583108	16.084672	0.00290762566973246	0.999999999875
147	150	-141.013787	-132.990365	16.046844	0.00295694582647521	0.999999999875

Continued on next page

Table S6 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
2	42	-84.714222	-76.731945	15.964554	0.00306709755514389	0.999999999875
148	265	-154.588796	-146.624072	15.929448	0.00311530979260433	0.999999999875
47	339	-100.758584	-92.848993	15.819182	0.00327164401657698	0.999999999875
82	172	-78.15744	-70.259604	15.795672	0.00330596158678165	0.999999999875
468	469	-78.861775	-70.974619	15.774312	0.00333744774713474	0.999999999875
52	267	-97.238677	-89.396752	15.68385	0.00347409805255838	0.999999999875
23	151	-269.213363	-261.383175	15.660376	0.0035104477183352	0.999999999875
4	23	-94.34211	-86.512211	15.659798	0.00351134745896731	0.999999999875
4	216	-97.160477	-89.335684	15.649586	0.00352728137476244	0.999999999875
127	221	-94.72738	-86.927875	15.59901	0.00360725014316809	0.999999999875
82	151	-278.186472	-270.400877	15.57119	0.00365199548860917	0.999999999875
241	287	-70.044727	-62.261923	15.565608	0.00366103903998682	0.999999999875
331	435	-52.400018	-44.629912	15.540212	0.00370246262268259	0.999999999875
2	216	-98.249254	-90.486048	15.526412	0.00372516482327112	0.999999999875
2	23	-95.430889	-87.692935	15.475908	0.0038094215182668	0.999999999875
155	328	-56.759264	-49.047166	15.424196	0.00389763482251559	0.999999999875
248	267	-135.975774	-128.313444	15.32466	0.00407311889187079	0.999999999875
47	249	-68.746144	-61.193065	15.106158	0.00448602771443796	0.999999999875
127	285	-65.277967	-57.73946	15.077014	0.00454412162507856	0.999999999875
4	18	-83.278102	-75.767891	15.020422	0.00465904744183665	0.999999999875
359	419&420	-22.759934	-15.258199	15.00347	0.0046940255441954	0.999999999875
199	267	-166.807698	-159.307124	15.001148	0.00469883671018834	0.999999999875
221	437	-94.576657	-87.11821	14.916894	0.00487672489268509	0.999999999875
2	18	-84.366883	-76.913846	14.906074	0.0049000434090325	0.999999999875
267	385	-74.1345	-66.768057	14.732886	0.00528853411749663	0.999999999875
18	30	-103.435422	-96.078339	14.714166	0.00533229497135612	0.999999999875
334	384	-72.339251	-64.983297	14.711908	0.00533759735197847	0.999999999875
89	322	-22.493935	-15.217846	14.552178	0.00572611215121643	0.999999999875
149	390	-39.239054	-32.011958	14.454192	0.00597803773405736	0.999999999875
194	335	-94.541538	-87.341945	14.399186	0.00612419180716595	0.999999999875
199	215	-216.7733	-209.575485	14.39563	0.00613375997292154	0.999999999875
65	381	-85.121412	-77.963863	14.315098	0.00635441619646426	0.999999999875
267	381	-85.467318	-78.315731	14.303174	0.00638774244647655	0.999999999875
151	172	-262.588286	-255.441413	14.2937459999999	0.0064142139127632	0.999999999875
51	431	-150.81514	-143.689073	14.252134	0.0065323406301101	0.999999999875
159	322	-22.114499	-14.992086	14.244826	0.006553305353589	0.999999999875
52	199	-133.721387	-126.615921	14.210932	0.00665140378507478	0.999999999875
199	331	-144.286475	-137.186706	14.199538	0.00668470326733372	0.999999999875
467	468	-79.489709	-72.425517	14.128384	0.00689638112617863	0.999999999875
260&452	261	-21.77714	-14.727302	14.099676	0.00698363284295656	0.999999999875
267	342	-135.50191	-128.473591	14.056638	0.00711646736040761	0.999999999875
257	282&352	-22.199851	-15.184335	14.031032	0.00719667016713788	0.999999999875
329	356	-32.774535	-25.765332	14.018406	0.00723654248121752	0.999999999875
149	370	-105.774035	-98.7817	13.98467	0.00734414381931736	0.999999999875
52	62	-101.5899	-94.629657	13.920486	0.00755320371228818	0.999999999875
268	282&352	-21.514353	-14.561359	13.905988	0.00760122744767489	0.999999999875
267	462	-110.321548	-103.372697	13.897702	0.00762880846715241	0.999999999875
338	370	-126.37497	-119.429498	13.890944	0.00765137588288978	0.999999999875
16	21	-65.851742	-58.90748	13.888524	0.00765947302770498	0.999999999875
52	151	-263.384837	-256.46375	13.842174	0.00781618608652146	0.999999999875

Continued on next page

Table S6 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
215	249	-149.334812	-142.465949	13.737726	0.00818093018169386	0.999999999875
153	368	-19.522221	-12.667663	13.709116	0.00828371284576168	0.999999999875
155	248	-106.249358	-99.395435	13.707846	0.00828830455355267	0.999999999875
227	228	-21.372586	-14.519539	13.706094	0.0082946430297719	0.999999999875
47	344	-55.414002	-48.597998	13.632008	0.00856706076175984	0.999999999875
197	462	-94.856671	-88.054269	13.604804	0.00866927281861274	0.999999999875
69	308	-21.946527	-15.156018	13.581018	0.00875961957864446	0.999999999875
165	274	-28.831479	-22.046053	13.570852	0.00879851364628936	0.999999999875
11	355	-29.487763	-22.70613	13.563266	0.00882764692170956	0.999999999875
43	53	-62.18546	-55.418248	13.534424	0.00893927565536456	0.999999999875
346	390	-34.096673	-27.337333	13.51868	0.00900079186093183	0.999999999875
47	127	-78.616404	-71.858896	13.515016	0.00901516741845965	0.999999999875
267	331	-107.803764	-101.056404	13.49472	0.00909520561465604	0.999999999875
141	151	-254.977731	-248.235514	13.484434	0.00913603386506068	0.999999999875
106	368	-23.105528	-16.365453	13.48015	0.00915309118315888	0.999999999875
56	267	-91.562243	-84.838056	13.448374	0.00928058750487426	0.999999999875
71	127	-86.15186	-79.442996	13.417728	0.00940519423249664	0.999999999875
356	466	-34.056448	-27.348043	13.41681	0.00940895194781366	0.999999999875
432	462	-109.814891	-103.10911	13.411562	0.00943046216358712	0.999999999875
153	344	-33.436465	-26.766585	13.33976	0.00972963369836866	0.999999999875
151	385	-240.280661	-233.617236	13.32685	0.00978440009898351	0.999999999875
18	23	-64.745408	-58.082295	13.326226	0.00978705483280917	0.999999999875

Table S8: **Epistatic pairs of sites of amino acid detected before FDR in the Gong13NP data set of NP proteins in H3N2 human influenza A virus.** The residues were recoded according to their physicochemical properties as in Figure S2.

Epistatic pair of Amino acids (years)	Amino acid physicochemical properties groups	Detected by Gong and coworkers	<i>P</i> -value after FDR **
A131S×T373N (2007–1969)	- Hydrophobicity - Normalized van der Waals Volume - Polarity - Solvent Accessibility	No	1
A131S×E375G (2007–1991)	- Hydrophobicity - Normalized van der Waals Volume - Polarity - Solvent Accessibility - Charge - Polarizability	No	1
A131S×R384G* (2007–1991)	- Polarity - Solvent Accessibility - Charge - Polarizability	No	1
A373S×L259S* (1969–1973)	- Polarizability - Normalized van der Waals Volume - Polarity - Secondary structure	No	1

Notes—* Deleterious mutation; **FDR: False Discovery Rate.

Table S9: Detailed results generated using the Gong13NP data set. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The P -values are computed using the standard χ^2 approximation. FDR-corrected P -values follow BH. Only pairs significant at the nominal 1% level (P -value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	P -value	P -value (FDR)
259	334	-24.112557	-15.110641	18.003832	0.00123197177337009	1
421	425	-41.33507	-32.643448	17.383244	0.00162806970000984	1
259	421	-28.433097	-20.76289	15.340414	0.00404483575372983	1
186	259	-52.01484	-44.549463	14.930754	0.0048470139256368	1
259	411	-43.108741	-35.931834	14.353814	0.00624738040528983	1
286	421	-21.478869	-14.567582	13.822574	0.00788339582234365	1
246	470	-11.518628	-4.755542	13.526172	0.00897146705998841	1

Table S10: Detailed results generated using the Duan14NA data set. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The P -values are computed using the standard χ^2 approximation. FDR-corrected P -values follow BH. Only pairs significant at the nominal 1% level (P -value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	P -value	P -value (FDR)
275	354	-136.541395	-116.877463	39.327864	5.96E-08	0.0020535384925
151	450	-139.615983	-128.751821	21.728324	0.000226982022307198	0.9999999999955
151	332	-159.091073	-148.686575	20.808996	0.000345509589050774	0.9999999999955
69	352	-55.379769	-44.989787	20.779964	0.000350115379225246	0.9999999999955
287	354	-63.891067	-54.806851	18.168432	0.00114397690881296	0.9999999999955
151	173	-137.888164	-128.945172	17.885984	0.00129904744386045	0.9999999999955
82	151	-146.103838	-137.666486	16.874704	0.00204434758921312	0.9999999999955
151	453	-144.681804	-136.271468	16.820672	0.00209431794817483	0.9999999999955
68	332	-95.891365	-87.504276	16.774178	0.00213827945675782	0.9999999999955
23	151	-160.568287	-152.31876	16.499054	0.0024176613631598	0.9999999999955
93	95	-53.766448	-45.618077	16.296742	0.00264576463273458	0.9999999999955
23	452	-93.45056	-85.319271	16.262578	0.00268632246945488	0.9999999999955
130	267	-121.808649	-113.710559	16.19618	0.00276690614373509	0.9999999999955
151	262	-136.41958	-128.732775	15.37361	0.00398586951995639	0.9999999999955
83	151	-154.666562	-147.216912	14.8993	0.00491469796176947	0.9999999999955
151	367	-151.047796	-143.685161	14.72527	0.00530629497982016	0.9999999999955
354	454	-78.108496	-70.771995	14.673002	0.00542977590314364	0.9999999999955
100	151	-130.504031	-123.195851	14.61636	0.00556677477166456	0.9999999999955
68	78	-106.513245	-99.36027	14.30595	0.00637996856547018	0.9999999999955
6	329	-100.48581	-93.444414	14.082792	0.00703545228325753	0.9999999999955
149	151	-156.535571	-149.509767	14.051608	0.00713215283964763	0.9999999999955
151	452	-164.589539	-157.573367	14.032344	0.00719253929803354	0.9999999999955
151	393	-145.958782	-138.989157	13.93925	0.0074914900717713	0.9999999999955
130	239	-53.068511	-46.106562	13.923898	0.00754194493815419	0.9999999999955
6	45	-49.538259	-42.597146	13.882226	0.00768058503537805	0.9999999999955
234	382	-52.245962	-45.500083	13.491758	0.00910694436005499	0.9999999999955

Table S11: Detailed results generated using the Koel13HA data set. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The P -values are computed using the standard χ^2 approximation. FDR-corrected P -values follow BH. Only pairs significant at the nominal 1% level (P -value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	P -value	P -value (FDR)
131	172	-56.325232	-90.641292	68.63212	4.41E-14	2.56E-09
172	299	-63.869614	-91.063367	54.387506	4.37E-11	1.27E-06
131	197	-54.723437	-80.250126	51.053378	2.18E-10	4.20E-06
131	299	-52.796784	-27.588776	50.416016	2.96E-10	4.28E-06
156	196	-149.015007	-125.514329	47.001356	1.52E-09	1.77E-05
82	172	-54.520069	-77.674156	46.308174	2.12E-09	2.05E-05
131	190	-42.767375	-65.287127	45.039504	3.90E-09	3.23E-05
142	156	-144.5229	-122.150763	44.744274	4.49E-09	3.26E-05
94	197	-44.415095	-66.252821	43.675452	7.49E-09	4.83E-05
155	158	-23.289118	-44.194232	41.810228	1.83E-08	0.000105870061375635
121	172	-81.928955	-102.095109	40.332308	3.69E-08	0.000194716683130647
156	226	-182.029994	-162.474344	39.1113	6.61E-08	0.000319184934128395
275	299	-72.503845	-91.350619	37.693548	1.30E-07	0.000577962847914969
172	262	-76.631249	-94.724662	36.186826	2.65E-07	0.00104000518981192
131	133	-89.996826	-108.07345	36.153248	2.69E-07	0.00104000518981192
57	157	-71.313012	-54.028379	34.569266	5.69E-07	0.00206323564386095
260	275	-84.652373	-101.740308	34.17587	6.86E-07	0.00233855447015707
156	192	-159.698581	-142.985686	33.42579	9.77E-07	0.0029905995524521
133	135	-92.222911	-108.932543	33.419264	9.80E-07	0.0029905995524521
226	262	-121.085486	-104.732525	32.705922	1.37E-06	0.00397691373256576
135	172	-56.241943	-72.48434	32.484794	1.52E-06	0.00420337615276976
94	248	-30.942845	-14.856788	32.172114	1.76E-06	0.00464875522733443
146	155	-44.326771	-60.286742	31.919942	1.99E-06	0.00500694975261313
172	196	-85.20475	-100.881214	31.352928	2.59E-06	0.00604869206333314
124	299	-58.586432	-74.256782	31.3407	2.61E-06	0.00604869206333314
47	216	-57.331129	-41.954512	30.753234	3.44E-06	0.00766433062728581
94	172	-46.654893	-61.943841	30.577896	3.73E-06	0.00801358453394216
226	276	-110.892793	-95.935017	29.915552	5.09E-06	0.0103080793574517
124	190	-55.926525	-70.837009	29.820968	5.32E-06	0.0103080793574517
133	157	-101.076796	-115.984826	29.81606	5.34E-06	0.0103080793574517
197	299	-60.894938	-75.768088	29.7463	5.51E-06	0.0103080793574517
190	193	-67.907935	-53.10205	29.61177	5.87E-06	0.0104351974002796
82	124	-52.888221	-67.681553	29.586664	5.94E-06	0.0104351974002796
146	219	-93.45458	-107.94919	28.98922	7.86E-06	0.0130615143536782
82	197	-56.427429	-70.918084	28.98131	7.89E-06	0.0130615143536782
157	226	-110.350463	-95.997502	28.705922	8.97E-06	0.0144137159332662
124	450	-84.747798	-99.073679	28.651762	9.20E-06	0.0144137159332662

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
133	157	-101.076796	-115.984826	29.81606	5.34E-06	0.0103080793574517
197	299	-60.894938	-75.768088	29.7463	5.51E-06	0.0103080793574517
190	193	-67.907935	-53.10205	29.61177	5.87E-06	0.0104351974002796
82	124	-52.888221	-67.681553	29.586664	5.94E-06	0.0104351974002796
146	219	-93.45458	-107.94919	28.98922	7.86E-06	0.0130615143536782
82	197	-56.427429	-70.918084	28.98131	7.89E-06	0.0130615143536782
157	226	-110.350463	-95.997502	28.705922	8.97E-06	0.0144137159332662
124	450	-84.747798	-99.073679	28.651762	9.20E-06	0.0144137159332662
156	193	-141.906891	-127.689876	28.43403	1.02E-05	0.0147925885126242
131	156	-115.085158	-129.301989	28.433662	1.02E-05	0.0147925885126242
57	156	-132.329673	-118.114981	28.429384	1.02E-05	0.0147925885126242
131	262	-53.772767	-67.932639	28.319744	1.07E-05	0.0151901024859121
196	375	-88.241385	-102.325965	28.16916	1.15E-05	0.015908596206199
219	299	-84.39992	-98.45473	28.10962	1.19E-05	0.0159765755133827
54	155	-44.92534	-58.755838	27.660996	1.46E-05	0.0189908627484213
197	276	-64.598716	-78.409744	27.622056	1.49E-05	0.0189908627484213
124	276	-64.414062	-78.209269	27.590414	1.51E-05	0.0189908627484213
172	248	-47.379421	-61.153635	27.548428	1.54E-05	0.0189908627484213
156	262	-136.522344	-122.848716	27.347256	1.69E-05	0.0200587546743246
347	384	-45.407296	-59.078106	27.34162	1.70E-05	0.0200587546743246
155	173&323	-21.437778	-34.852155	26.828754	2.15E-05	0.0249596943834507
133	138	-233.09576	-220.018628	26.154264	2.95E-05	0.0334827438576916
146	260	-68.844774	-55.848314	25.99292	3.17E-05	0.0353938089077092
219	276	-90.87916	-103.849445	25.94057	3.25E-05	0.0355802907411882
155	225	-35.075461	-47.921275	25.691628	3.65E-05	0.0391978185272695
156	246	-221.264234	-208.527253	25.473962	4.04E-05	0.0418124876197024
94	133	-81.269174	-93.938962	25.339576	4.30E-05	0.0437190995686334
54	159	-50.291102	-62.915365	25.248526	4.48E-05	0.0448167713614273
156	375	-139.427401	-126.910084	25.034634	4.95E-05	0.0486452103978388
156	197	-135.830135	-123.368725	24.92282	5.21E-05	0.0500389568285764
375	453	-75.964483	-63.513684	24.901598	5.27E-05	0.0500389568285764
196	220	-190.197331	-177.794496	24.80567	5.50E-05	0.0514666003308989
124	197	-74.605806	-86.927028	24.642444	5.94E-05	0.0534856947229171
135	173&323	-32.68001	-20.37102	24.61798	6.00E-05	0.0534856947229171
142	220	-188.033821	-175.725054	24.617534	6.00E-05	0.0534856947229171
124	375	-77.707519	-90.001173	24.587308	6.09E-05	0.0534856947229171
163	172	-48.55901	-60.814908	24.511796	6.31E-05	0.054559298682351
190	197	-59.452012	-71.688215	24.472406	6.42E-05	0.054744735840454
229	452	-226.236426	-214.036553	24.399746	6.64E-05	0.0557938638811007
146	275	-78.195493	-90.366425	24.341864	6.82E-05	0.0564876048789984
159	248	-28.837736	-40.832342	23.989212	8.03E-05	0.0654365820654832
121	133	-118.085097	-130.066287	23.96238	8.13E-05	0.0654365820654832
124	201	-89.206528	-101.002581	23.592106	9.64E-05	0.0755854137686689
156	276	-126.358671	-114.563495	23.590352	9.65E-05	0.0755854137686689
50	275	-110.836472	-122.471673	23.270402	0.000111809463839929	0.0864212615840093
124	157	-66.812128	-78.413177	23.202098	0.000115381222295596	0.0879987775898886

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
135	197	-59.645916	-71.232886	23.17394	0.000116886421846152	0.0879987775898886
220	452	-172.908979	-161.347677	23.122604	0.000119680940907085	0.0889474890305608
229	375	-236.730719	-225.246181	22.969076	0.00012844020645264	0.0942490983298678
121	190	-62.812756	-74.191632	22.757752	0.000141546180741137	0.10174518241206
261	299	-49.605782	-38.231659	22.748246	0.000142165944029271	0.10174518241206
121	156	-142.863567	-131.514102	22.69893	0.000145424681587181	0.10280815599523
124	262	-77.39017	-88.543916	22.307492	0.000174060335175197	0.120835928031344
142	299	-76.920183	-65.772891	22.294584	0.000175094323868086	0.120835928031344
135	276	-45.598181	-56.70918	22.221998	0.000181023273510994	0.123457872534498
192	246	-181.232002	-170.157115	22.149774	0.000187120238953398	0.126132095954982
172	384	-51.709551	-62.747294	22.075486	0.000193603939211195	0.12778228350828
192	229	-254.692613	-243.665571	22.054084	0.000195512862955383	0.12778228350828
159	172	-56.782493	-67.781472	21.997958	0.000200608073363862	0.12778228350828
220	529	-182.672498	-171.678073	21.98885	0.000201447241126806	0.12778228350828
220	375	-183.67905	-172.702215	21.95367	0.000204721378502803	0.12778228350828
50	260	-94.689275	-83.712958	21.952634	0.000204818592827949	0.12778228350828
135	275	-70.650232	-81.625592	21.95072	0.000204998315788685	0.12778228350828
133	452	-111.861384	-100.91372	21.895328	0.000210267879833248	0.128754390559956
159	173&323	-24.001045	-34.937096	21.872102	0.000212517188936223	0.128754390559956
190	226	-102.042375	-91.120654	21.843442	0.000215325661845345	0.128754390559956
49	529	-83.631148	-72.719162	21.823972	0.000217254543595446	0.128754390559956
63	260	-53.36382	-42.453885	21.81987	0.000217663106345967	0.128754390559956
453	529	-75.705336	-65.007195	21.396282	0.000264224599794471	0.154718182324096
155	163	-25.280744	-35.935942	21.310396	0.000274805841863901	0.159304946528503
49	106	-57.463109	-46.821678	21.282862	0.000278286158348906	0.159725233658278
106	220	-156.810892	-146.1897	21.242384	0.000283482072796493	0.161112311372673
121	196	-91.506957	-102.10514	21.196366	0.000289505995278105	0.162938471323027
196	242	-69.271627	-58.710905	21.121444	0.000299586106181193	0.164549176608459
384	453	-41.724796	-52.280959	21.112326	0.000300836354965583	0.164549176608459
222	299	-51.144795	-40.592114	21.105362	0.000301794731485505	0.164549176608459
54	384	-45.764409	-56.310121	21.091424	0.000303721957859326	0.164549176608459
172	173&323	-46.653453	-57.104388	20.90187	0.000331175002466932	0.176130411862459
133	156	-172.823132	-162.484654	20.676956	0.00036695331471015	0.190004515112401
172	450	-88.325033	-98.655943	20.66182	0.000369494164581097	0.190004515112401
133	384	-88.779026	-99.107331	20.65661	0.000370372782606543	0.190004515112401
172	192	-97.567101	-87.261729	20.610744	0.000378197508699829	0.192316750695869
190	452	-57.611482	-47.322632	20.5777	0.000383936076347235	0.193182273008617
190	375	-64.673245	-54.391881	20.562728	0.000386564493168873	0.193182273008617
106	529	-57.730217	-47.499146	20.462142	0.000404690950940867	0.200258011824123
94	219	-71.216443	-81.439564	20.446242	0.000407632316633544	0.200258011824123
202	222	-54.493339	-44.291284	20.40411	0.000415529163839023	0.20119445225455
155	159	-32.086038	-42.28304	20.394004	0.000417445760909296	0.20119445225455
57	220	-176.494371	-166.303938	20.380866	0.000419950469601527	0.20119445225455
131	276	-42.673108	-52.800566	20.254916	0.000444730422040629	0.211319857095863
216	385	-51.12756	-41.018044	20.219032	0.000452051719206814	0.213052342783894
186	220	-232.61022	-222.575661	20.069118	0.000483950986062465	0.225227730365345
156	157	-125.887003	-115.856311	20.061384	0.000485655792576645	0.225227730365345

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
156	452	-128.94249	-118.936546	20.011888	0.000496707861006196	0.22852503732166
133	229	-270.049502	-260.056591	19.985822	0.000502627916866838	0.228899715244312
82	133	-89.803667	-99.772844	19.938354	0.000513588728797698	0.228899715244312
163	173&323	-16.360794	-26.326642	19.931696	0.000515144927566547	0.228899715244312
94	201	-60.913487	-70.875467	19.92396	0.000516958954721258	0.228899715244312
82	94	-31.19643	-21.238934	19.914992	0.000519069791021454	0.228899715244312
190	192	-83.70289	-73.76272	19.88034	0.000527306458202315	0.228899715244312
82	275	-66.816087	-76.753238	19.874302	0.000528754848769131	0.228899715244312
138	197	-196.176423	-186.240013	19.87282	0.000529110951228873	0.228899715244312
172	375	-80.616478	-90.527502	19.822048	0.000541455136314095	0.231027861632655
49	186	-133.668109	-123.758194	19.81983	0.000542000848405055	0.231027861632655
248	275	-60.168155	-70.032178	19.728046	0.000565066964782535	0.239101693054332
121	261	-62.405206	-72.25909	19.707768	0.000570292590757138	0.239564213668053
193	226	-125.692221	-115.873978	19.636486	0.00058904432396012	0.245661147194015
94	275	-60.533342	-70.31553	19.564376	0.000608634590062218	0.252018194185048
33	54	-54.146174	-44.396762	19.498824	0.000627001926607584	0.257782281457033
157	386	-92.544435	-82.813833	19.461204	0.000637789452684889	0.258428360570166
92	192	-114.465663	-124.189099	19.446872	0.000641947281733724	0.258428360570166
213	275	-71.171593	-80.799194	19.255202	0.000700198623615145	0.27993458076531
63	146	-45.974074	-36.384591	19.178966	0.000724795029350189	0.287783341448154
164	174	-56.108263	-46.53173	19.153066	0.000733344249998602	0.289197048791966
157	192	-90.895589	-81.382307	19.026564	0.000776555386258315	0.304168349604017
57	375	-74.320637	-64.815217	19.01084	0.000782099423448535	0.304283916626252
57	190	-56.288048	-46.791331	18.993434	0.000788282162159093	0.304644779602417
54	260	-68.566102	-59.081965	18.968274	0.000797304697197632	0.304852908491239
121	155	-58.531093	-68.012414	18.962642	0.000799338314484532	0.304852908491239
131	375	-58.396569	-67.869367	18.945596	0.000805524635998256	0.305204334306006
213	347	-51.981455	-61.436208	18.909506	0.000818779048583518	0.308211827573939
260	384	-49.10544	-58.545712	18.880544	0.000829571409825203	0.310259707274626
157	375	-68.291231	-58.86465	18.853162	0.000839904310122042	0.312110595242146
142	229	-241.330907	-231.919572	18.82267	0.000851560607108226	0.314426550280662
124	163	-51.157226	-60.491179	18.667906	0.000913234268536511	0.335064497133301
262	275	-88.295839	-97.621699	18.65172	0.000919934053339233	0.335399855799216
53	189	-86.800404	-77.48348	18.633848	0.000927388166094945	0.336004324928275
5	190	-65.780774	-56.490441	18.580666	0.000949924463000906	0.342031808199767
155	196	-62.001647	-71.268394	18.533494	0.00097036645025228	0.346443758690093
63	201	-69.798372	-60.539949	18.516846	0.000977684168894877	0.346443758690093
192	262	-100.991766	-91.739234	18.505064	0.000982895942460371	0.346443758690093
83	155	-19.979145	-29.22809	18.49789	0.000986082804620758	0.346443758690093
94	299	-34.225275	-24.98394	18.48267	0.000992877745520149	0.346729656071103
197	260	-78.239601	-69.033342	18.412518	0.00102479996611049	0.355734455301947
135	260	-57.909944	-48.723937	18.372014	0.00104369071405264	0.35971801330216
220	226	-226.004872	-216.82416	18.361424	0.0010486862902892	0.35971801330216
213	479	-67.793672	-76.960265	18.333186	0.00106212267358552	0.362183831692662
192	196	-109.842098	-100.684069	18.316058	0.00107035536081723	0.362856726705117
155	156	-112.859521	-121.987746	18.25645	0.00109950091122679	0.369693737437346
124	133	-113.111748	-122.236168	18.24884	0.00110327784330966	0.369693737437346
156	450	-146.622227	-137.504625	18.235204	0.00111007763935278	0.36983448708782
173&323	453	-35.626106	-26.537061	18.17809	0.001139011135188	0.377305574324846

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
124	260	-73.026933	-63.964393	18.12508	0.00116653159437441	0.384226343897071
133	450	-122.819352	-131.850334	18.061964	0.00120015628095649	0.39306813337315
133	299	-95.015347	-104.009823	17.988952	0.00124024857017624	0.402372830676089
138	196	-206.507396	-197.515936	17.98292	0.00124361941014584	0.402372830676089
131	144	-40.354927	-31.368611	17.972632	0.00124938950356557	0.402372830676089
133	246	-196.588891	-187.611804	17.954174	0.00125980814428805	0.402712445219168
53	137	-108.100392	-99.133508	17.933768	0.00127142621777521	0.402712445219168
196	276	-74.911702	-83.872319	17.921234	0.00127861478978042	0.402712445219168
160	213	-33.435388	-42.39487	17.918964	0.00127992097941532	0.402712445219168
121	220	-186.736444	-177.781519	17.90985	0.00128517858143085	0.402712445219168
63	173&323	-23.277442	-14.337132	17.88062	0.00130218482225264	0.404442993617244
142	222	-71.580004	-62.641801	17.876406	0.00130465481812014	0.404442993617244
57	155	-47.248426	-56.170924	17.844996	0.00132321182656014	0.408013774391976
75	157	-66.618	-57.708256	17.819488	0.00133847351204264	0.410536029064083
131	453	-51.268371	-42.37988	17.776982	0.00136429230267776	0.416252762032788
135	299	-42.898676	-34.04359	17.710172	0.00140587086861943	0.426692849496693
143	189	-28.579923	-19.735735	17.688376	0.00141970415627557	0.428647135100494
146	299	-53.891888	-45.069154	17.645468	0.00144733042595702	0.434724066283566
135	155	-32.919557	-41.736097	17.63308	0.00145540468664551	0.434895926210517
131	193	-60.007467	-51.202484	17.609966	0.00147058900589381	0.43505936742826
25	172	-58.717814	-49.913113	17.609402	0.00147096146310055	0.43505936742826
197	384	-51.281896	-60.064081	17.56437	0.001501002402438	0.441592177612016
143	275	-65.141732	-73.918526	17.553588	0.0015082844776122	0.441592177612016
7	479	-64.977813	-56.212322	17.530982	0.00152366577224905	0.443853793051645
92	375	-103.791523	-95.032004	17.519038	0.00153185496336772	0.444008161132133
222	225	-42.124864	-33.396397	17.456934	0.00157514071132425	0.453956964945181
260	262	-74.779231	-66.055495	17.447472	0.0015818407265642	0.453956964945181
126	155	-23.389175	-32.0967	17.41505	0.00160501264800983	0.456748998247154
57	229	-229.585554	-220.879636	17.411836	0.00160732785306916	0.456748998247154
158	275	-59.902418	-68.590982	17.377128	0.0016325406786849	0.460698883376356
155	375	-57.672848	-66.358287	17.370878	0.00163712213171519	0.460698883376356
156	220	-241.394328	-232.731327	17.326002	0.00167039235423028	0.467790554467291
229	529	-235.990236	-227.416409	17.147654	0.00180933848399223	0.504266114985719
163	201	-63.964852	-72.494195	17.058686	0.00188285276345668	0.522243898074563
186	453	-122.060406	-113.554376	17.01206	0.00192254842967854	0.528417986456766
121	197	-81.257522	-89.762545	17.010046	0.00192428154500135	0.528417986456766
33	192	-80.528773	-72.034745	16.988056	0.00194330522641417	0.528889220540984
33	197	-59.591614	-51.105486	16.972256	0.0019570882852008	0.53015143875276
192	347	-94.65143	-86.174938	16.952984	0.00197403068322821	0.532253761426695
222	452	-56.415417	-47.962534	16.905766	0.0020161546394134	0.541094835401829
82	122	-27.678802	-19.279528	16.798548	0.00211512461679475	0.56504043334374
142	529	-89.436843	-81.047975	16.777736	0.00213488335633827	0.566818104947187
157	260	-65.310951	-56.92546	16.770982	0.0021413345693192	0.566818104947187
155	452	-44.859465	-53.222155	16.72538	0.00218539790898842	0.574152125687997
63	248	-26.197581	-17.836658	16.721846	0.00218884974602462	0.574152125687997
192	220	-201.376513	-193.029727	16.693572	0.00221666057942971	0.578827990042974
201	226	-135.545613	-127.21645	16.658326	0.00225181730023905	0.582758253994899
173&323	248	-19.387097	-11.072093	16.630008	0.00228046122106784	0.587548164379124

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
193	452	-79.057276	-70.74903	16.616492	0.00229425908578562	0.588487607092887
159	220	-161.136734	-152.863073	16.547322	0.00236616975457205	0.604259298116923
275	450	-98.261946	-106.518177	16.512462	0.00240324816724158	0.608368105917006
63	142	-63.156418	-54.927657	16.457522	0.00246284967085508	0.620745197475953
57	159	-57.333136	-49.112254	16.441764	0.0024802118342877	0.622415065080772
7	155	-31.195166	-22.993921	16.40249	0.00252401043749195	0.630676228712968
173&323	452	-42.941814	-34.75868	16.366268	0.00256508039109282	0.636613946642521
160	163	-26.946379	-18.771702	16.349354	0.00258448294827329	0.636613946642521
121	450	-94.003495	-102.172221	16.337452	0.00259822267386667	0.636613946642521
173&323	275	-55.955738	-64.123671	16.335866	0.00260005898393745	0.636613946642521
163	192	-73.318484	-65.151683	16.333602	0.00260268251430529	0.636613946642521
124	135	-55.953698	-64.075774	16.244152	0.00270845046910417	0.659701149974659
196	246	-169.998866	-161.881753	16.234226	0.00272044501650903	0.659850199192588
163	220	-153.555007	-145.443602	16.22281	0.00273430476553704	0.660448530242427
172	217	-66.506317	-58.402986	16.206662	0.0027540282814581	0.662204794132106
190	196	-70.953145	-79.052244	16.198198	0.00276442229049456	0.662204794132106
157	347	-63.957638	-55.865698	16.18388	0.00278209317697564	0.66369523238386
452	529	-73.13894	-65.082841	16.112198	0.00287224723577861	0.679608866359536
135	242	-36.093561	-28.052785	16.081552	0.00291166269367737	0.686134497367794
106	146	-48.26613	-40.234117	16.064026	0.00293444267257603	0.688703002952359
142	172	-86.1186	-94.111434	15.985668	0.00303845546913251	0.705678222040266
131	275	-65.896907	-73.889338	15.984862	0.00303954399487061	0.705678222040266
260	347	-69.098607	-61.107562	15.98209	0.00304329059013397	0.705678222040266
124	384	-51.382023	-59.367177	15.970308	0.00305926575782245	0.706556318649272
10	275	-97.113217	-89.146958	15.932518	0.00311106406638095	0.713800871972872
163	375	-55.088634	-47.123892	15.929484	0.00311525997255713	0.713800871972872
57	226	-117.124914	-109.166056	15.917716	0.00313158743115161	0.714717021196294
209	356	-63.999716	-56.060599	15.878234	0.00318698425077535	0.722812995418078
135	160	-34.795395	-26.858048	15.874694	0.00319199804773207	0.722812995418078
124	146	-67.947634	-60.030045	15.835178	0.00324849527740056	0.732744246034672
121	157	-71.215349	-79.122179	15.81366	0.00327967285186859	0.736909438848147
142	246	-167.870281	-160.005454	15.729654	0.00340423315700045	0.761581606145444
133	190	-97.999831	-105.86085	15.722038	0.0034157532792447	0.761581606145444
173&323	197	-47.027883	-54.875769	15.695772	0.00345577878323144	0.76636948952601
33	260	-55.754481	-47.911097	15.686768	0.00346960546209574	0.76636948952601
121	124	-81.04241	-88.88343	15.68204	0.0034768876271406	0.76636948952601
155	244	-30.361867	-38.194334	15.664934	0.00350336047307531	0.769279570546121
157	248	-42.982537	-35.198709	15.567656	0.00365771846383844	0.80014316735364
82	131	-34.458742	-26.688049	15.541386	0.00370053758236977	0.800743716843867
121	131	-59.62614	-67.390754	15.529228	0.003720521168737	0.800743716843867
137	159	-47.792996	-55.556355	15.526718	0.00372465994570903	0.800743716843867
57	92	-88.344692	-80.582992	15.5234	0.00373013799529642	0.800743716843867
82	375	-57.992648	-50.234937	15.515422	0.00374334219880434	0.800743716843867
83	275	-56.443374	-64.190091	15.493434	0.00377997240351158	0.805606618498405
172	260	-75.107703	-67.368664	15.478078	0.00380576304390134	0.808132174560297
146	347	-61.728124	-53.999149	15.45795	0.00383983052194214	0.812390421010898
137	145	-72.056847	-64.335375	15.442944	0.00386542373527898	0.81483132339681

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
138	156	-257.771099	-250.056935	15.428328	0.0038905130048531	0.817148691635269
94	159	-30.908048	-23.204889	15.406318	0.00392859691159841	0.817506550518829
121	229	-240.06565	-232.367633	15.396034	0.00394651672317781	0.817506550518829
94	479	-55.115062	-62.812782	15.39544	0.00394755421606141	0.817506550518829
202	299	-54.730142	-47.033058	15.394168	0.00394977681998065	0.817506550518829
142	175	-74.83521	-67.141825	15.38677	0.00396272797474195	0.817506550518829
94	242	-25.239887	-17.563186	15.353402	0.00402166367405121	0.826722848172868
145	189	-50.169128	-42.522699	15.292858	0.00413080656592058	0.846158503980268
82	479	-63.230761	-70.866305	15.271088	0.00417075880973317	0.847728038740991
54	242	-45.197408	-37.563457	15.267902	0.00417663751623298	0.847728038740991
219	450	-110.984447	-118.616855	15.264816	0.00418233947006941	0.847728038740991
304	452	-79.56105	-71.934924	15.252252	0.00420563286092079	0.849479222813862
160	190	-32.237537	-39.858996	15.242918	0.00422302040104172	0.85002948836246
131	142	-62.348153	-69.959288	15.22227	0.00426173524646933	0.851919733793854
202	386	-94.210581	-86.599464	15.222234	0.00426180304985713	0.851919733793854
135	189	-31.296099	-23.698522	15.195154	0.00431310774307248	0.859212563113098
137	275	-79.92529	-87.51611	15.18164	0.00433893736023183	0.860330350979068
106	173&323	-25.962985	-18.376294	15.173382	0.00435479582101495	0.860330350979068
163	275	-61.425488	-69.009986	15.168996	0.00436324173172065	0.860330350979068
25	155	-28.866135	-21.287199	15.157872	0.00438473493320102	0.860910419108941
124	226	-120.400477	-127.976538	15.152122	0.00439588552796699	0.860910419108941
226	231	-108.30332	-100.737435	15.13177	0.00443557710712461	0.8657397383858
201	452	-87.528315	-79.967121	15.122388	0.00445399276918035	0.8657397383858
131	163	-31.985212	-24.427338	15.115748	0.00446707161298188	0.8657397383858
201	225	-79.175126	-71.623801	15.10265	0.00449298153648703	0.8657397383858
92	196	-103.242272	-110.793034	15.101524	0.00449521582291057	0.8657397383858
82	156	-114.448471	-106.905559	15.085824	0.00452648279179213	0.868874859073476
51	140	-27.99493	-20.456644	15.076572	0.00454500834659843	0.869551596872314
163	197	-51.216915	-58.746908	15.059986	0.00457840562108636	0.8730597824115711
146	172	-69.695959	-62.172157	15.047604	0.00460349484717526	0.87436753389091
226	229	-279.335161	-271.816238	15.0378460000001	0.00462336226496685	0.87436753389091
193	202	-74.95527	-67.438098	15.034344	0.00463051290157857	0.87436753389091
54	450	-86.146326	-78.658884	14.974884	0.00475359387992746	0.894694276686348
83	190	-26.728921	-34.208945	14.960048	0.00478480158961803	0.897653553884003
155	453	-42.586713	-50.05972	14.946014	0.0048145074367405	0.900312890670474
275	384	-64.530687	-71.990203	14.919032	0.00487213013360921	0.908065314243136
124	219	-100.050685	-107.506677	14.911984	0.00488729304888491	0.908065314243136
57	223	-84.14084	-76.698188	14.885304	0.00494511270361342	0.912955998179841
172	276	-65.09222	-72.5239	14.86336	0.00499317170162905	0.918902106487099
63	299	-37.507042	-30.080465	14.853154	0.00501567956499505	0.920123241717604
57	186	-124.452017	-131.833345	14.762656	0.00521966724385314	0.95452400670715
82	106	-35.67608	-28.314866	14.722428	0.00531293764369445	0.968525142153985
33	271	-35.925134	-28.574097	14.702074	0.0053607505954899	0.972459371626231
220	248	-152.637438	-145.287951	14.698974	0.00536806967259607	0.972459371626231
50	172	-102.095461	-94.753884	14.683154	0.00540557358263227	0.973502643058585
146	197	-70.425473	-63.084283	14.68238	0.00540741506063247	0.973502643058585
299	452	-57.488652	-50.15792	14.661464	0.00545741157288859	0.979461761239479
173&323	213	-30.960142	-23.641103	14.638078	0.00551384952353073	0.986536595305792
157	163	-42.781998	-35.474937	14.614122	0.00557225687010077	0.993919171568435
83	122	-16.342897	-9.042956	14.599882	0.00560726279606694	0.997095166527608

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
157	453	-59.019873	-51.728437	14.582872	0.00564936107038549	0.998627784066229
155	160	-25.028745	-32.319985	14.58248	0.00565033488310718	0.998627784066229
33	92	-72.218032	-64.932801	14.570462	0.00568027025204132	0.999999999955
143	248	-27.267079	-19.987242	14.559674	0.00570727434003271	0.999999999955
220	453	-169.875305	-162.59974	14.55113	0.00572875068182932	0.999999999955
193	219	-105.948183	-113.213606	14.530846	0.00578005494456268	0.999999999955
196	529	-90.932077	-83.678513	14.507128	0.00584061667350455	0.999999999955
275	278	-60.252977	-67.505748	14.505542	0.00584468850953412	0.999999999955
164	172	-77.033793	-69.812251	14.443084	0.00600727396046807	0.999999999955
196	347	-82.136695	-74.922972	14.427446	0.00604867095490935	0.999999999955
135	541	-28.532892	-21.328924	14.407936	0.00610071037092097	0.999999999955
75	146	-67.529937	-60.33342	14.393034	0.00614075431316519	0.999999999955
124	220	-179.802458	-172.608846	14.387224	0.00615643640992691	0.999999999955
133	163	-85.137387	-92.326747	14.37872	0.00617946086947774	0.999999999955
219	426	-99.656838	-92.480819	14.352038	0.00625225148398223	0.999999999955
155	172	-54.02349	-61.190159	14.333338	0.00630376701491686	0.999999999955
54	146	-60.15157	-52.998617	14.305906	0.00638009171022702	0.999999999955
63	160	-26.006799	-18.858222	14.297154	0.00640463271205716	0.999999999955
135	219	-82.619265	-89.739675	14.24082	0.00656482548319859	0.999999999955
122	384	-19.61421	-26.730605	14.23279	0.00658797731585026	0.999999999955
220	384	-156.740571	-149.626297	14.228548	0.0066002399998899	0.999999999955
138	142	-204.381318	-197.280048	14.20254	0.0066759139649003	0.999999999955
197	242	-56.047905	-48.964596	14.166618	0.00678183420708611	0.999999999955
122	160	-21.372697	-14.294871	14.155652	0.0068144956456776	0.999999999955
172	347	-71.159305	-64.088703	14.141204	0.00685776341484157	0.999999999955
190	248	-28.417982	-35.486246	14.136528	0.00687182432451994	0.999999999955
244	276	-48.214431	-41.146367	14.136128	0.0068730284506795	0.999999999955
53	278	-89.373057	-82.309062	14.12799	0.0068975713423286	0.999999999955
196	453	-77.860993	-70.799972	14.122042	0.00691556385341208	0.999999999955
172	226	-120.953241	-128.005738	14.104994	0.00696738871790314	0.999999999955
131	304	-65.466472	-58.417651	14.097642	0.00698985561474619	0.999999999955
57	121	-77.922398	-70.874583	14.09563	0.00699601642551784	0.999999999955
106	135	-36.089592	-29.047467	14.08425	0.00703096264802816	0.999999999955
160	219	-74.502493	-81.542656	14.080326	0.00704305225809942	0.999999999955
155	384	-27.9318	-34.962247	14.060894	0.00710322191274571	0.999999999955
83	145	-49.581557	-42.556249	14.050616	0.00713525027446926	0.999999999955
172	529	-81.075068	-74.052081	14.045974	0.0071497620426999	0.999999999955
2	155	-37.541461	-44.560506	14.03809	0.00717447520251169	0.999999999955
242	331	-23.549119	-30.546877	13.995516	0.00730938055947039	0.999999999955
126	193	-52.890505	-45.899921	13.981168	0.00735540286939074	0.999999999955
106	155	-33.576133	-26.591511	13.969244	0.00739386587593449	0.999999999955
172	219	-101.114381	-108.09349	13.958218	0.00742960749346167	0.999999999955
331	384	-28.597205	-35.570754	13.947098	0.0074658252008365	0.999999999955
156	384	-112.828685	-119.796767	13.936164	0.00750160569561331	0.999999999955
160	201	-63.752392	-70.700896	13.897008	0.00763112297862256	0.999999999955
82	242	-32.522742	-25.591032	13.86342	0.00774396575789293	0.999999999955

Continued on next page

Table S10 - Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	P-value	P-value (FDR)
25	193	-59.454494	-52.53825	13.832488	0.0078493296237222	0.999999999955
275	509	-57.4627	-50.551978	13.821444	0.00788728787322224	0.999999999955
146	384	-41.72992	-48.639894	13.819948	0.0078924434293498	0.999999999955
199	247	-65.406967	-58.507465	13.799004	0.00796496857918394	0.999999999955
82	226	-99.63263	-92.734582	13.796096	0.00797508988390194	0.999999999955
138	226	-242.381419	-235.484437	13.793964	0.00798251831619945	0.999999999955
142	452	-81.076979	-74.185702	13.782554	0.00802238894678764	0.999999999955
194	226	-145.435549	-138.544347	13.782404	0.00802291439669367	0.999999999955
126	244	-27.433497	-34.32111	13.775226	0.00804809834599707	0.999999999955
160	384	-25.484145	-32.36505	13.76181	0.00809537591497678	0.999999999955
220	450	-190.460282	-183.585038	13.750488	0.0081354857282594	0.999999999955
106	142	-65.317115	-58.454194	13.725842	0.00822347151126335	0.999999999955
138	248	-168.733679	-161.872729	13.7219	0.00823763051129656	0.999999999955
159	453	-46.472956	-53.324344	13.702776	0.00830666006438074	0.999999999955
160	347	-44.814802	-37.964372	13.70086	0.00831360713550122	0.999999999955
155	189	-21.071808	-27.91832	13.693024	0.00834207829043065	0.999999999955
144	158	-27.55264	-34.385354	13.665428	0.00844310679467353	0.999999999955
426	427	-61.807167	-54.976437	13.66146	0.00845773168901576	0.999999999955
94	260	-44.074939	-37.244767	13.660344	0.00846184940515771	0.999999999955
83	375	-51.343718	-44.527575	13.632286	0.00856602235870729	0.999999999955
347	479	-79.971252	-86.782867	13.62323	0.00859991231801649	0.999999999955
207	534	-49.252066	-42.451926	13.60028	0.00868638590255322	0.999999999955
63	472	-21.409695	-14.612078	13.595234	0.00870551256512819	0.999999999955
57	312	-68.943744	-75.736721	13.585954	0.00874079566293218	0.999999999955
83	278	-20.432775	-13.655385	13.55478	0.00886034838818517	0.999999999955
94	155	-25.260075	-32.034964	13.549778	0.00887967946810242	0.999999999955
275	347	-82.205447	-88.976397	13.5419	0.00891020902684159	0.999999999955
131	450	-65.606709	-72.374406	13.535394	0.00893549907317859	0.999999999955
138	375	-199.776984	-193.016337	13.521294	0.00899054964298773	0.999999999955
49	328	-57.073112	-50.313366	13.519492	0.00899760904776725	0.999999999955
54	193	-74.387307	-81.14572	13.516826	0.00900806315130598	0.999999999955
271	299	-46.499555	-39.741173	13.516764	0.00900830641129602	0.999999999955
94	192	-69.05857	-62.303105	13.51093	0.00903122516243138	0.999999999955
194	220	-204.799787	-198.048616	13.502342	0.00906506675038055	0.999999999955
202	452	-62.267793	-55.521928	13.49173	0.00910705539793133	0.999999999955
196	229	-243.459521	-236.717326	13.48439	0.00913620889882683	0.999999999955
190	229	-214.765556	-208.041104	13.448904	0.00927844678736212	0.999999999955
146	450	-78.511167	-71.791849	13.438636	0.00932000622995632	0.999999999955
57	217	-65.101674	-58.383247	13.436854	0.00932723733819962	0.999999999955
156	347	-130.002894	-123.288134	13.42952	0.00935705548519916	0.999999999955
271	529	-61.85077	-55.1434	13.41474	0.00941743061537137	0.999999999955
260	452	-69.189044	-62.482096	13.413896	0.00942088976152655	0.999999999955

Continued on next page

TableS10 -Continued from previous page

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
142	201	-99.514968	-92.836301	13.357334	0.0096555631683235	0.9999999999955
157	222	-54.129263	-47.451169	13.356188	0.00966037645785334	0.9999999999955
197	386	-103.343478	-96.674705	13.337546	0.00973900452533127	0.9999999999955
53	384	-92.442923	-99.11064	13.335434	0.00974795185568822	0.9999999999955
10	34	-61.255058	-54.591963	13.32619	0.00978720801200927	0.9999999999955
131	529	-59.485435	-52.834357	13.302156	0.00988999689741887	0.9999999999955
260	299	-58.561035	-51.915353	13.291364	0.00993649480352199	0.9999999999955
63	124	-51.664336	-45.019507	13.289658	0.00994386472616582	0.9999999999955

Table S12: Detailed results generated using the Adam03M2 data set. Site-1 and site-2 show the AA positions along the alignment. The lkl terms show the log-likelihood values after optimization under the independent and the dependent model. The test statistic is twice the log-likelihood difference between the two models. The *P*-values are computed using the standard χ^2 approximation. FDR-corrected *P*-values follow BH. Only pairs significant at the nominal 1% level (*P*-value) are shown; those above the horizontal line inside the table are also significant after FDR.

site-1	site-2	lkl-indep	lkl-dep	test-stat	<i>P</i> -value	<i>P</i> -value (FDR)
31	51	-215.103387	-197.277061	35.652652	3.41E-07	0.00113281616036631
12	53	-77.799957	-65.876321	23.847272	8.57E-05	0.142316180090207
13	56	-46.536317	-40.493356	12.085922	0.0167236344812858	0.999999964628273
11	58	-29.888689	-23.892009	11.99336	0.0174007103450287	0.999999964628273
13	54	-71.386993	-66.032614	10.708758	0.0300397186352879	0.999999964628273
6	70	-17.476614	-12.309403	10.334422	0.0351558484101363	0.999999964628273
5	8	-12.900464	-7.964634	9.87166	0.0426459429703552	0.999999964628273
13	51	-168.765715	-164.127949	9.275532	0.0545696117650336	0.999999964628273
63	75	-16.839233	-12.370825	8.936816	0.0626983299054582	0.999999964628273
40	63	-17.108515	-12.661735	8.89356	0.0638157412109532	0.999999964628273
24	52	-58.933899	-54.596959	8.67388	0.0697883949162066	0.999999964628273
4	90	-32.302657	-27.98877	8.62777400000001	0.0711074320946067	0.999999964628273
75	96	-17.939049	-13.701758	8.474582	0.0756614405570145	0.999999964628273
5	10	-24.371314	-20.161276	8.420076	0.0773472039538335	0.999999964628273
16	17	-39.158191	-34.956543	8.403296	0.0778732712198347	0.999999964628273
75	77	-24.809162	-20.607664	8.402996	0.0778827071170962	0.999999964628273
50	88	-56.048288	-51.902973	8.290629999999999	0.0814937940632067	0.999999964628273
14	68	-47.801338	-43.671019	8.260638	0.0824839822106254	0.999999964628273
12	51	-173.532531	-169.559639	7.945784	0.0935845065252358	0.999999964628273
15	94	-27.233707	-23.320416	7.826581999999999	0.0981412984366585	0.999999964628273
9	55	-22.540834	-18.64306	7.795548	0.0993612397204912	0.999999964628273