

INVITED REVIEWS AND SYNTHESSES

Inferences from tip-calibrated phylogenies: a review and a practical guide

ADRIEN RIEUX¹ and FRANÇOIS BALLOUX*Department of Genetics, Evolution and Environment, UCL Genetics Institute, University College London, Gower Street, London WC1E 6BT, UK*

Abstract

Molecular dating of phylogenetic trees is a growing discipline using sequence data to co-estimate the timing of evolutionary events and rates of molecular evolution. All molecular-dating methods require converting genetic divergence between sequences into absolute time. Historically, this could only be achieved by associating externally derived dates obtained from fossil or biogeographical evidence to internal nodes of the tree. In some cases, notably for fast-evolving genomes such as viruses and some bacteria, the time span over which samples were collected may cover a significant proportion of the time since they last shared a common ancestor. This situation allows phylogenetic trees to be calibrated by associating sampling dates directly to the sequences representing the tips (terminal nodes) of the tree. The increasing availability of genomic data from ancient DNA extends the applicability of such tip-based calibration to a variety of taxa including humans, extinct megafauna and various microorganisms which typically have a scarce fossil record. The development of statistical models accounting for heterogeneity in different aspects of the evolutionary process while accommodating very large data sets (e.g. whole genomes) has allowed using tip-dating methods to reach inferences on divergence times, substitution rates, past demography or the age of specific mutations on a variety of spatiotemporal scales. In this review, we summarize the current state of the art of tip dating, discuss some recent applications, highlight common pitfalls and provide a 'how to' guide to thoroughly perform such analyses.

Keywords: Bayesian phylogenetics, calibration, divergence time and substitution rate inferences, measurably evolving populations, population dynamics, tip-dating

Received 21 September 2015; revision received 1 February 2016; accepted 4 February 2016

Introduction

The idea of molecular dating was first proposed in 1962 by Zuckerkandl & Pauling (1962) when they suggested that the divergence time between two species could be measured by the number of mutations accumulated between molecular sequences (in their case, protein sequences). As molecular sequence divergence can only provide a relative timescale, calibration using an

external source of information is required to convert relative into absolute divergence times. Unless one assumes the substitution rate to be known (which is a strong hypothesis irrespective of the molecular sequence under scrutiny), two sources of independent age-related information can be exploited. One approach is to assign dates to internal nodes representing the most recent common ancestors (MRCAs) between lineages using information from the fossil record or dated biogeographical event (see box 1 in Ho *et al.* 2011a), which are rarely known with great precision. An alternative strategy, which is the focus of this review, takes advantage of the information about the age of the sequenced samples themselves to calibrate the phy-

Correspondence: Adrien Rieux, Fax: +0262 499 293;

E-mail: adrien.rieux@cirad.fr

¹Present address: CIRAD, UMR PVBMT, 97410, Saint-Pierre, La Réunion, France.

logeny by assigning dates to the tips (sometimes also called terminal nodes) of the tree, hence the term tip dating. Tip dating is only possible when there is sufficient spread in the age of the samples analysed and is ideally suited for data sets of serially sampled fast-evolving taxa, or those including ancient DNA sequences (Drummond *et al.* 2003b).

The conceptual bases of tip dating were laid out in the late 1980s when sequence data from samples with associated dates of isolation started to accumulate in public databases (Rambaut 2000). In order to obtain accurate divergence time estimates for fast-evolving genomes such as RNA viruses, the date of isolation must be accounted for. Indeed, the number of new mutations accumulated in each sequence is expected to correlate with the date of isolation. The idea of exploiting known isolation dates to conjointly estimate the rate of evolution with the time since the divergence of other internal nodes emerged by turning this reasoning around (see principle in Fig. 1). With this principle, formalized and applied to HIV data by Li *et al.* (1988), pairs of sequences must be chosen to be independent of each other by ensuring that no two pairs share any evolutionary history (branches on the tree) since their respective common ancestors. Subsequent analyses performed on hepatitis B virus data showed that the stochastic nature of the substitution process can lead to one sequence sampled earlier to exhibit more divergence from the outgroup than one sampled later and thus inflate the variance of the estimate (Bollyky & Holmes 1999). To realistically express such uncertainty,

Rambaut (2000) and Drummond *et al.* (2001) introduced new methods incorporating sequence dates into a maximum-likelihood (ML) tree reconstruction framework. Those approaches were subsequently embedded into a Bayesian statistical inference framework to jointly estimate substitution rates, divergence times and demography while accounting for the uncertainty in the genealogy by using Markov chain Monte Carlo (MCMC) integration (Drummond *et al.* 2002). Such developments ultimately led to the release of BEAST (Drummond & Rambaut 2007), which has been the most popular tip-dating program developed so far (Table 1). Early implementations were assuming a constant rate of evolution throughout the tree. However, the analysis of numerous data sets showing considerable departures from clocklike evolution led to the development and incorporation of methods accounting for rate variation among lineages (Bromham & Penny 2003; Welch & Bromham 2005; Ho & Duchêne 2014).

Concomitant with the development of increasingly sophisticated and computationally efficient phylogenetic inference algorithms during the last decade, the field of ancient DNA (aDNA) research underwent its own revolution (Millar *et al.* 2008; Der Sarkissian *et al.* 2015). Although the field's focus was initially limited to mitochondrial DNA and a few nuclear markers, numerous whole-genome sequences from the deep past have now been retrieved. This breakthrough is tightly connected to the massive sequence throughput of modern sequencing technologies and the ability to target short and degraded DNA molecules. Additionally, the analytical power

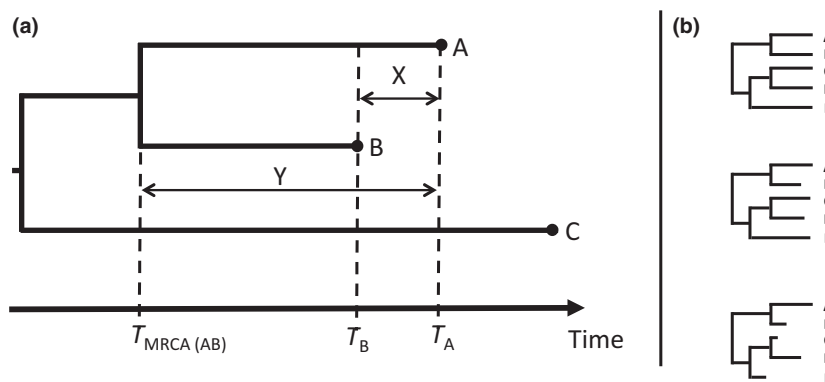


Fig. 1 Tip-dating principle. (a) In this simplified theoretical situation adapted from Rambaut (2000), sequences A and B were isolated at different points in time (T_A and T_B , respectively) and C is an outgroup sequence. If we assume the rate of evolution to be the same in lineages A and B, then the amount of molecular evolution expected to have occurred between T_A and T_B is equal to $d_{AC} - d_{BC}$ (d_{AC} and d_{BC} being the genetic distance between A&C and B&C, respectively). If the time X between T_A and T_B represents a significant proportion of the time Y since A and B last shared a common ancestor, then one can use tip dates to conjointly estimate the rate of evolution $\mu = (AC-BC)/(T_A-T_B)$ and extrapolate the age of $T_{MRCA(AB)}$. (b) Top: Tree with modern samples only for which no divergence time estimate is possible without calibrations on internal nodes or a strong prior on the rate of molecular clock. Middle: Tree where tip dates may not be widely spread enough for accurate inferences. Bottom: Tree where tip date width should be sufficiently broad to allow divergence time and rate of evolution estimates with a good degree of certainty, since the sample dates cover a relatively large fraction of the total age of the tree.

Table 1 List of available and useful programs/packages to perform tip dating

Software to perform tip dating	Method	Tree topology	Models of rate variation [†]	Tip date uncertainty (as in Fig. 4) [‡]	Ref./Source	
BEAST	Bayesian	Estimated/Fixed	SC, LC, ARC, URC	a, b, c, d, e	[1]	
R8s	Nonparametric	Fixed	SC, LC, DC, ARC	a	[2]	
PAML (MCMCTREE)	Bayesian	Fixed	SC, ARC, URC	a, b, d, e	[3]	
PHYSHER	ML	Fixed	LC, DC	a	[4]	
Mr Bayes	Bayesian	Estimated/Fixed	SC, ARC, URC	a, b, d, e	[5]	
TipDate	ML	Estimated	SC, DC	a	[6]	
DAMBE	Distance based	Fixed	SC, DC, ARC	a	[7]	
Multidivtime	Bayesian	Fixed	SC, ARC	a	[8]	
LSD	Least squares	Fixed	SC	a	[9]	
Software/packages to assist tip dating						
ClonalFrameML, SplitsTree4, RDP4, LDhat	Various tools to detect recombining individuals/genomic regions					[10,11,12,13]
MapDamage2	R & Python framework to detect patterns of damage in ancient DNA reads					[14]
PartitionFinder	Python program to select best-fit partitioning schemes and models of molecular evolution					[15]
BEAUTi	Java program assisting the preparation of infiles for running BEAST					[16]
TipDatingBeast	R package assisting the implementation of the date-randomization and the leave-one-out cross-validation tests in BEAST					[17]
SiteSampler	Java application to sample the sites of a sequence alignment to produce replicate data sets					[18]
Murray <i>et al.</i> 's R scripts	R scripts to test for temporal signal and confounding between temporal and genetic structures					[19]
Tracer	Java program to analyse the trace (log) files generated by Bayesian MCMC programs					[20]
LogCombiner	Java program to combine log and tree files from multiple independent runs of Bayesian MCMC programs					[16]
TreeAnnotator	Java program to assist summarizing the information from a sample of trees onto a consensus tree					[16]
FigTree	Java graphical viewer of phylogenetic trees					[21]
DensiTree	Java graphical viewer for sample of (multiple) phylogenetic trees					[22]
Path-O-gen	Java program to plot root-to-tip distances as a function of sampling time					[23]
adephylo	R package to perform various phylogenetic analyses such as computing the distance of a set of tips to the root					[24]
TreesStat	Java program to calculate summary statistics (including root-to-tip lengths) from a set of trees					[16]
BEAGLE	High-performance library to make use of highly parallel processors and speed up inferences					[25]
PHYLOCH	R package to read, manipulate and extract information from annotated trees					[26]

[1] (Drummond & Rambaut 2007), [2] (Sanderson 1997), [3] (Yang 2007), [4] (Fourment & Holmes 2014), [5] (Ronquist *et al.* 2012b), [6] (Rambaut 2000), [7] (Xia 2013), [8] (Thorne *et al.* 1998), [9] <http://www.atgc-montpellier.fr/LSD/>, [10] (Didelot & Wilson 2015), [11] (Huson & Bryant 2006), [12] (Martin *et al.* 2015), [13] (McVean *et al.* 2002), [14] (Jonsson *et al.* 2013), [15] (Lanfear *et al.* 2012), [16] Distributed in the BEAST package, [17] <https://cran.r-project.org/web/packages/TipDatingBeast/index.html>, [18] <https://github.com/simonho/sitesampler/>, [19] (Murray *et al.* 2015) [20] <http://tree.bio.ed.ac.uk/software/tracer/>, [21] <http://tree.bio.ed.ac.uk/software/figtree/>, [22] <https://www.cs.auckland.ac.nz/~remco/DensiTree/>, [23] <http://tree.bio.ed.ac.uk/software/pathogen/>, [24] (Jombart *et al.* 2010), [25] <http://beast.bio.ed.ac.uk/beagle>, [26] <http://www.christophheibl.de/Rpackages.html>.

[†]Models of rate variation among branches: strict clock (SC), local multirate clock (LC), discrete multirate clock (DC), autocorrelated relaxed clock (ARC) and uncorrelated relaxed clock (URC). See Ho & Duchêne (2014) for more details.

[‡]Available distribution to model tip date uncertainty (see also Fig. 4): a: point values (no uncertainty), b: normal distribution, c: empirical description of the probability density function directly measured on the calibrated sample, d: uniform distributions with hard minimum and maximum bounds, e: uniform distribution with hard minimum and soft maximum bounds.

obtained through the analysis of billions of sequence reads allows quantifying contamination issues that have haunted aDNA research for decades. Whole genomes have now been sequenced from ancient anatomically modern humans (Fu *et al.* 2013a,b), archaic humans (Briggs *et al.* 2009; Krause *et al.* 2010), ancient pathogens (Bos *et al.* 2011, 2014; Rasmussen *et al.* 2015) and many animals including megafaunal species (Shapiro *et al.* 2004; Gilbert *et al.* 2008).

Following the detection and removal of potential recombining sites (see Section “To date or not to date – when is tip dating appropriate?” and Appendix S1-A, Supporting information for more details), such time-stamped sequence data can be integrated into phylogenetic reconstructions and used as calibration points to estimate the timing of divergence events, the rate at which substitutions accumulate and reconstruct the past demography of many species for which fossil data are not available. Tip-based calibration represents far more than a substitute when information to calibrate internal nodes is not available, as it is expected to improve inference accuracy and robustness (Rieux *et al.* 2014; Sheng *et al.* 2014). Indeed, the uncertainty associated with the dates of time-stamped sequences simply mirrors the uncertainty in the estimated age of the sequences (the error associated with sampling time or the C-14 radiocarbon dating). This is a well-characterized source of error that can be integrated into phylogenetic inference (Drummond *et al.* 2006; Ho & Phillips 2009). In contrast, calibration of internal nodes relies on the assumption of an association between an MRCA in the phylogenetic tree and information such as fossil data or biogeographical evidence. Those indirect hypothesized associations come at a cost of considerable uncertainty which is additionally much more difficult to model satisfyingly.

Besides allowing to estimating substitution rates and divergence times, thoroughly time-calibrated phylogenetic trees represent a powerful tool for hypothesis testing. As such, it is not surprising that tip-dating methods have been widely adopted in the fields of ancient DNA and microbial genomics (Biek *et al.* 2015). However, these methods can be complex to implement and many tip-based inferences in the recent literature are reported with erroneously narrow confidence intervals or are even wholly incorrect. In the following, we review the current state of the art of molecular tip dating. We start with a general overview of tip-dating methodologies and their applicability. We then review some recent results obtained through tip dating and discuss various questions that this methodology can address. Finally, we outline some future research perspectives and provide in Appendix a ‘how to’ guide to assist users performing thorough tip-dating analyses (see major steps summarized in Fig. 5).

To date or not to date – when is tip dating appropriate?

In this review, we restrict the term tip dating to analyses in which dated sequences contain sufficient temporal information for populations to be characterized as ‘measurably evolving’ (see Drummond *et al.* 2003b and description below) and can hence be used as the principal source for molecular clock calibration. Numerous studies included ancient DNA sequences from extant or extinct species into phylogenetic inference by both fixing the age of the tips (to the sample radiocarbon ages) and incorporating external information by specifying the age of one or several internal node(s) (based on the fossil record) and/or the rate of evolution (e.g. Briggs *et al.* 2009; Bunce *et al.* 2009; Rieux *et al.* 2014; Sheng *et al.* 2014; Heintzman *et al.* 2015). Among these studies, we only consider as ‘tip dating’ those that performed specific analyses to disentangle the signal coming from the different calibration sources and were able to demonstrate that tip dates had sufficient temporal spread to inform on the molecular clock (e.g. Rieux *et al.* 2014; Sheng *et al.* 2014). This definition of tip dating entails us to focus this review on studies performed at the population (i.e. intraspecies) scale. Indeed, despite recent progress in the field of ancient DNA, aDNA sequences remain generally too scarce and recent for thorough dating of ancient events (i.e. millions of years) from molecular data and tip-dating methods alone. In this context, it is worth mentioning recent methodological developments named ‘fossil tip dating’ or ‘total evidence dating’ that allow combining morphological and molecular data to simultaneously infer the placement of a fossil in the phylogeny and calibrate trees at deeper geological times (Pyron 2011; Ronquist *et al.* 2012a). A major challenge in applying these approaches is the requisite to compile morphological character data for both extant and fossil taxa, which is complicated by the fact that for many groups the fossil record is extremely scarce and fragmentary (Arcila *et al.* 2015).

Measurably evolving populations

Tip-dating calibration requires working on measurably evolving populations (MEPs), a concept introduced over a decade ago by Drummond *et al.* (2003b). MEPs are populations exhibiting detectable amounts of *de novo* nucleotide changes among DNA sequences sampled at different time points. Our ability to capture measurable amounts of evolutionary change from sequence data depends on the evolutionary rate of the DNA/RNA sequence analysed (μ), its length (L) and the duration over which samples were collected (t). The pioneering research in tip dating that started around 20 years ago

focused on fast-evolving RNA viruses as at the time molecular sequences rarely exceeded 1000 bp and data sets covering sampling over 20 years were scarce (Drummond *et al.* 2003b). Progress in DNA sequencing of both modern and ancient material has led to a massive increase in both sequence length (L) and time span covered by the sequences (t). As a result, many populations from a diverse range of taxa can now be treated as MEPs (Biek *et al.* 2015).

Before performing any calibration, it is crucial to test whether there is temporal signal in the molecular data (Firth *et al.* 2010). Many studies have relied on the ‘regression method’ (Buonagurio *et al.* 1986; Shankarappa *et al.* 1999; Korber *et al.* 2000; Drummond *et al.* 2003a) which is based on the fit of a linear regression between the age of the samples and their root-to-tip distance (i.e. the number of substitutions separating each sample from the hypothetical ancestor at the root of the tree) (Fig. 2). Assuming statistical independence, a significant positive correlation between root-to-tip distances and sampling times would indicate the presence of detectable amounts of *de novo* mutations within the data set timescale. However, this is a problematic test as there is extensive pseudoreplication between samples. Indeed, the same branches in the phylogeny will contribute to multiple root-to-tip distances. While this problem can be partly alleviated by the use of a nonparametric test when assessing significance of the relationship, the non-

independence between distances cannot be completely controlled for (Drummond *et al.* 2003a). Additionally, the regression method can be misleading when there is substantial rate variation among lineages. The test is also likely to produce a spurious signal in the case of nonuniform distribution of the sampling times (Ho *et al.* 2007a, 2011b), which can often be the case for studies relying on aDNA for calibration where a handful of ancient sequences are typically combined with a larger number of modern samples.

A more robust method to investigate the extent of genetic and temporal signal within a data set is the ‘date-randomization test’ (Ramsden *et al.* 2008; Duffy & Holmes 2009). This test involves generating multiple randomized data sets by permutation of sampling times, and comparing parameter estimates obtained with the initial data set vs. the randomized ones (See Fig. 2). A recent simulation-based study (Duchêne *et al.* 2015b) recommends conducting at least 20 randomizations (this number is arbitrary and in principle it would be preferable to run a far larger number of replicates). By performing tip dating on the date-randomized data sets, expectations of the rate and divergence time estimates can be computed in the absence of temporal signal. To determine whether there is sufficient temporal signal in a data set, one thus needs to verify that the mean rate (or TMRCA) estimated with the original sampling times is not contained within any of the 95%

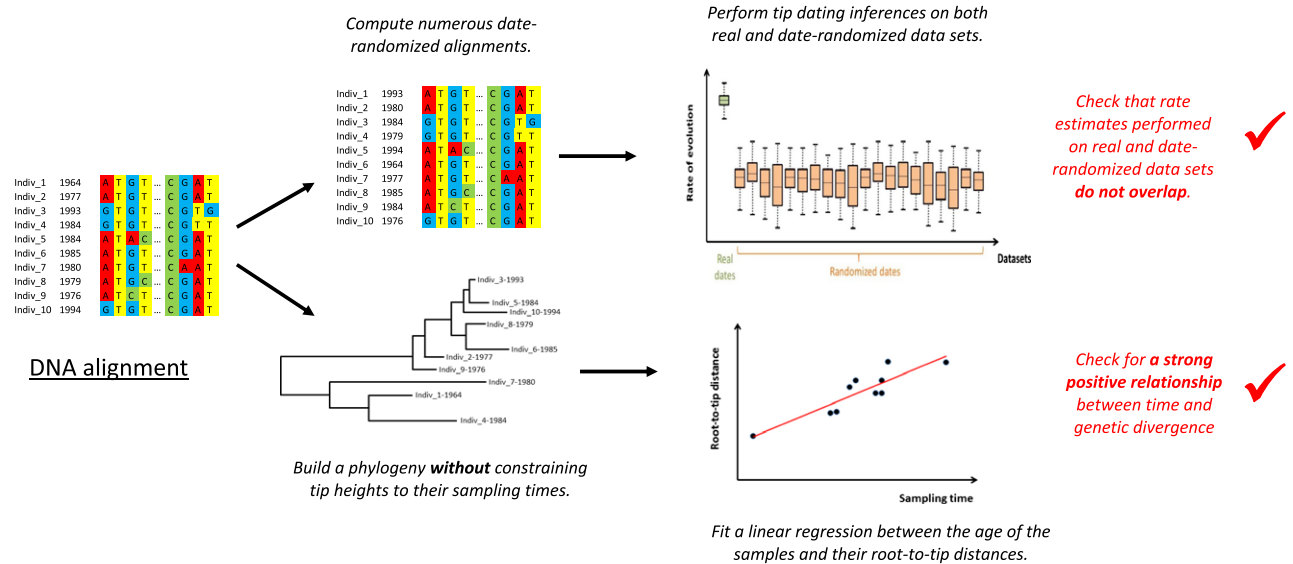


Fig. 2 Testing for temporal signal. Flow chart for testing measurable evolutionary change in a data set prior to any tip-dating analysis. The most robust method existing so far is the ‘date-randomization test’ which involves generating multiple randomized data sets by permutation of sampling times, and comparing parameter estimates obtained with the initial data set *vs* the randomized ones (see Section To date or not to date – when is tip dating appropriate? in the text for more details on how to perform this test and interpret the results); visual evidence for a temporal signal can also be obtained by fitting a linear regression between the age of the samples and their root-to-tip distances, which has to be computed from a tree built without constraining tip heights to their sampling times. Different tools allowing computing date-randomized data sets and root-to-tip distances are listed in Table 1.

credible interval of those obtained from the date-randomized data sets. A more stringent criterion is to compare the 95% credible interval of the original rate (or TMRCA) instead of its mean. Importantly, the 'date-randomization test' can accommodate nonrandom sampling resulting in a nonuniform distribution of sampling times by a modification of the randomization procedure. The latter involves identifying clusters of samples with the same, or very similar, sampling times and permuting the sampling times among but not within these clusters (Duchêne *et al.* 2015b).

Finally, a distinct approach uses model selection and compares the fit of models with the sampling dates included or excluded (Rambaut 2000; Drummond *et al.* 2003a; Baele *et al.* 2012). Temporal signal is confirmed if the inclusion of the sampling dates improves the fit. In practice, to keep the two cases as similar as possible, tip dates for the 'no dates' model are set to the most recent sampling date in the original data set (Murray *et al.* 2015).

Two recent studies investigated the performance of several of the above tests using both simulated and empirical data sets. Murray *et al.* (2015) showed that all of the standard tests of temporal signal can be misleading for data where temporal and genetic structures are confounded (i.e. where related sequences are more likely to have been sampled at similar times). Duchêne *et al.* (2015b) demonstrated that the performance of the date-randomization test can be affected when the sampling times are not uniformly distributed throughout the tree. On a more positive note, both studies show that the 'clustered permutation' date-randomization approach can successfully correct for such confounders. Consequently, we highly recommend performing the clustered date-randomization test using the most stringent criterion prior to any study aiming to applying tip calibrations. While the regression approach is statistically unsatisfying, it can still be used as a visual 'sanity check' for the reliability of rate estimates since the slope coefficient corresponds to the substitution rate under the assumption of a strict molecular clock, and the R^2 coefficient of determination indicates the degree to which sequence evolution followed a clocklike rate. For data sets failing to pass the DRT, the sampling and/or the sequencing strategy should be modified to widen the sampling time window (e.g. by including older samples, if possible) and/or increasing the number of sites in the alignment (e.g. by sequencing more nucleotides, if possible). Investigating the temporal signal of a data set is crucial because in the absence of a temporal signal in the data, the result will be driven by the prior, and is thus likely to be misleading.

Nonrecombining sequences

Central assumptions behind all phylogenetic models developed so far are that sequences evolve neutrally, and have not undergone genetic recombination. Although deviation from neutrality can satisfyingly be accommodated (see Appendix S1-C, Supporting information), accounting for recombination in phylogenetic inferences remains a difficult task and ignoring it leads essentially systematically to misleading inferences (Hedge & Wilson 2014). The rate of homologous recombination has been shown to vary significantly between species (Smukowski & Noor 2011), between lineages within species (Didelot *et al.* 2012; Bauer *et al.* 2013) as well as between regions within genomes (Everitt *et al.* 2014; Yahara *et al.* 2014). The influence of recombination will hence vary from data set to data set. In addition to its potential to distort tree topology, recombination can also create a false signal of apparent mutational evolution by introducing additional divergence between samples taken at different time points (Sanchez-Buso *et al.* 2014). The easiest strategy to avoid such bias when analysing MEPs is to consider recombination-free genomic regions such as mitochondrial DNA. An alternative approach is to perform a preliminary analysis to detect recombination events in the alignment based on conflicts in the topologic placement in the phylogeny and exclude incriminated individuals and/or genomic regions from subsequent dating inferences or partition the data around recombination breakpoints (see Appendix S1-A, Supporting information).

Applications of tip dating: hypotheses testing and potential sources of errors

Once nonrecombining genomic regions have been identified, tip-dating inferences can be performed to test various hypotheses while estimating evolutionary rates and timescale-related parameters. This usually involves considering phylogenetic methods based on molecular clocks, which make assumptions about patterns of rate variation among lineages (see Ho & Duchêne 2014; Welch & Bromham 2005 for previous reviews on this particular topic). In the following, we discuss various recent tip-dating applications.

Estimating the age of internal nodes in a phylogenetic tree

One of the most popular aims of dating analyses is to convert the genetic divergence measured between sequences of DNA into an absolute age for any internal node of a phylogenetic tree. This estimated node age is

referred to as the time to the most recent common ancestor (TMRCA). Internal phylogenetic nodes represent such putative ancestors for all sampled individuals within a clade defined by a node and can as such sometimes be associated with key chapters in a species' evolutionary history. In humans for instance, dating events such as the dawn of humankind millions of years ago or the expansion of anatomically modern humans from an African cradle some 100 k years ago have received tremendous interest. By using mitochondrial genomes from ancient archaic and modern humans spanning 65 000 years as calibration points for the mitochondrial clock, recent studies were able to date the divergence between Chimpanzee, *Homo neanderthalensis* and *H. sapiens* as well as the TMRCA of all modern human mtDNAs with an unprecedented precision (Fu *et al.* 2013b; Rieux *et al.* 2014). Also, by assuming that the coalescence of certain candidate haplotypes coincides with discrete human expansion or migrations, tip-dating analyses allowed to date various other events such as the 'out of Africa' exit or the initial colonization of different islands such as Sahul, Japan, Remote Oceania, Madagascar, New Zealand (Rieux *et al.* 2014).

Beyond simply reflecting a divergence event between lineages and taxa, age estimates for internal nodes can also provide information about the timing of emergence of novel specific genetic variants (Holden *et al.* 2013; Spagnoletti *et al.* 2014). For example, Eldholm *et al.* (2015) used whole-genome sequences of the nonrecombining bacteria *Mycobacterium tuberculosis* (*Mtb*) obtained from isolates sampled over 13 years to reconstruct the dynamic of antibiotic resistance-conferring mutations during a major ongoing outbreak in Argentina. Their results indicate that a multidrug-resistant *Mtb* strain had been circulating for 15 years before the outbreak was initially detected and about one decade before the earliest documented transmission of *Mtb* strains with such extensive resistance profile.

Dating internal nodes can also provide meaningful information about pathogen host species jumps such as those leading to novel emerging zoonotic, agronomic and wildlife diseases (Parrish *et al.* 2008; Engering *et al.* 2013). By using a panel of whole-genome sequences of *Staphylococcus aureus* representing the breadth of the species' diversity with known dates of isolation, Weinert *et al.* (2012) were able to infer and map the different host switches on the phylogenetic tree while accounting for uncertainty about ancestral host associations. Their results point to multiple jumps back and forth between human and bovids with the first switch from humans to bovids dating back to around 5500 BP, which coincides with the time when cattle domestication started expanding throughout the Old World, thus suggesting

a central role for anthropogenic change in the emergence of new endemic diseases.

Estimation of evolutionary timescales from tip-calibrated phylogenies has become routine in biology, forming the basis of a wide range of evolutionary and ecological studies as illustrated above. However, it is important to remember that various sources of error can affect these estimates, including incorrect calibration dates (Ho *et al.* 2008; Molak *et al.* 2013), misspecification of the demographic, substitution and molecular clock models (Navascues & Emerson 2009; Ho *et al.* 2011b; Duchêne *et al.* 2015c) or the presence of tree imbalance (Duchêne *et al.* 2015a). Whereas the effect of model misspecification is a common issue in biology, the influence of tree imbalance is worth considering as in addition to being affected by biological factors such as past demography or natural selection, it is also directly related to sampling effort and strategy. Tree imbalance refers to the relative number of tips descending from internal nodes in a phylogenetic tree. In a completely balanced tree, each of the two lineages descending from any internal node leads to the same number of tips. A completely imbalanced tree has a pectinate or comb-like arrangement of branches. Imbalanced ultrametric trees are characterized by an excess of long branches because some of the lineages will have few descendants in the sample. In their study based on both simulated and empirical data sets, Duchêne *et al.* (2015a) found that tree imbalance had a detrimental impact on dating precision and produced a systematic bias with overall timescales being underestimated. A pronounced effect was observed in analyses with shallow calibrations. The greatest decrease in accuracy usually occurred in the age estimates for medium and deep nodes of the tree. Those results indicate that in case of tree imbalance, molecular clock analyses can be improved by increasing taxon sampling, with the specific aims of including deeper calibrations, breaking up long branches and reducing tree imbalance.

Estimating rates of evolution and their variation

Conjointly with the estimation of the TMRCA of two or more sampled genomes, dating a phylogenetic tree also allows inferring the rate at which mutations accumulate on their connecting lineage (Fig. 1). Substitution rates can be heterogeneous both between genomic regions (Bromham & Penny 2003) and along evolutionary lineages (Lanfear *et al.* 2010). Various partitioning algorithms allow accounting for heterogeneity within genomes (e.g. Lanfear *et al.* 2012) and dedicated molecular clock models have been developed to deal with variable substitution rates along lineages (see Ho &

Duchêne 2014; Welch & Bromham 2005 for more details on this topic).

Such developments allowed, for example revising the rate at which the human mitochondrial DNA accumulates mutations (Fu *et al.* 2013b; Rieux *et al.* 2014). By pre-estimating the optimal partitioning scheme and the best-fit nucleotide substitution model for each partition of the mtDNA molecule (see Appendix S1-C, Supporting information), and taking advantage of the wealth of sequenced ancient human mitochondrial genomes from the last 65 000 years, it was possible to refine the mitochondrial substitution rates for different genomic regions and codon positions (Rieux *et al.* 2014). In addition, the use of a relaxed molecular clock allowed investigating rate variation along evolutionary lineages. There is an extensive debate in the literature on the existence and significance of time-dependent rates of molecular evolution (Ho *et al.* 2011a). In particular, one common observation is for substitution rates to accelerate in recent generations in several species including humans (Ho *et al.* 2007b, 2011a; Henn *et al.* 2009; Duchêne *et al.* 2014). This phenomenon is generally ascribed to the time needed for natural selection to weed out deleterious mutations from the population. Consistent with this pattern, the recent rates obtained for human mitochondrial DNA recovered a subtle but significant negative correlation between the age of the ancient genome used for calibration and the substitution rate estimated (Rieux *et al.* 2014). Stronger evidence for purifying selection was found in the form of a stark difference in the substitution rates of first and second codons (PC1 + 2) versus third codon (PC3). Although PC3 mutations accumulate linearly with time in humans, a clear acceleration in the rate starting at around 30 000 years was recorded for the mostly non-synonymous mutations at PC1 + 2 (Rieux *et al.* 2014).

In addition to allowing to explore patterns of rate variation in time, tip dating combined with the relaxed molecular clocks also permits investigating variation in substitution rates across lineages. One of the clearest evidence for dramatic heterogeneity in substitution rates has been reported for historical variation in substitution rate in *Yersinia pestis*, the aetiologic agent of plague (Cui *et al.* 2013) with nearly 40-fold difference between the slowest and the fastest evolving branches.

Like for any model-based inference procedure, inference quality of tip-based Bayesian coalescent analyses will depend on how well the model captures the underlying processes that generated the data. In this context, it has been shown using simulations that molecular rate estimates obtained from ancient DNA-based calibrations can be upwardly biased for populations evolving under complex demographic scenarios and/or populations with large effective sizes, as both situations tend to pro-

duce genealogies where ancient and modern DNA samples segregate in different lineages (Navascues & Emerson 2009; Emerson *et al.* 2015).

Reconstructing past demography

Conjointly with rates of evolution and divergence times, effective population sizes (EPSs) through time can be estimated from a time-calibrated tree using Bayesian skyline plots (BSP) and related methods implemented in coalescent-based algorithms such as BEAST (see Ho & Shapiro 2011 for a recent review). The EPS corresponds to the number of idealized individuals that contribute offspring to the descendent generation and is almost always smaller than the census population size. One interesting feature of such reconstruction methods is that they are well suited to detect population bottlenecks. Such estimates of EPS variation have notably allowed investigating the relative impacts of climatic and anthropogenic factors on the widespread extinctions of large mammals such as the steppe bison (Shapiro *et al.* 2004) or the cave bear (Stiller *et al.* 2010) at the end of the Pleistocene epoch.

Skyline-plot-based demographic inferences also intensively contributed to the field of molecular epidemiology since EPS plots can help detecting when epidemics took off. In a recent study published thirty years after the discovery of HIV-1, Faria *et al.* (2014) made use of more than 800 concatenated RNA sequences obtained from individuals infected by the virus between 1959 and 2004 in Africa and the USA to shed light on the early transmission, dissemination and establishment of the virus in human populations. Their findings indicate that Kinshasa in the 1920s was the epicentre of early transmission and the source of pre-1960 pandemic viruses elsewhere and that the demographic history of group M and non-pandemic group O was similar until ~1960, when the M group underwent an epidemiological transition and outpaced regional population growth. Their results emphasize the likely role of iatrogenic interventions and changes in sexual behaviour in the emergence of HIV-1 group M. In addition to shedding light on the origins of an epidemic, the combination of tip dating and demographic reconstruction also allows testing whether policy changes for managing an epidemic have been effective. This has notably been the case with the United Kingdom HIV-1 epidemic and Egyptian hepatitis C virus (HCV) epidemic (Stadler *et al.* 2013), 2009 influenza A (H1N1) pandemic (Fraser *et al.* 2009) or 2014–2015 Ebola epidemics (Stadler *et al.* 2014).

The use of BSPs and related methods for the estimation of past EPS has become increasingly common in the literature, covering a wide spectrum of taxa from viruses to large mammals. Despite their appeal, EPSs are not

always straightforward to interpret and to equate to actual census sizes (Frost & Volz 2010). It is also important to remember that BSP models are assuming a single panmictic population and violation of this assumption can lead to misleading inferences (Stack *et al.* 2010; de Silva *et al.* 2012; Heller *et al.* 2013). Damage in ancient DNA has also been shown to potentially distort inferences of demographic histories (Axelsson *et al.* 2008; Rambaut *et al.* 2009). Particular attention should thus be given to those factors in future studies making use of BSP models to study past demography (more details in Appendix S1-B, Supporting information).

Estimating the age of a sample

Another application of tip-dating analyses is the estimation of the age of a sample for which this information is missing using phylogenetic molecular clock-based methods. There are a number of circumstances under which the leaf ages of sequences may be unknown or at best, highly uncertain. First, ancient DNA sequences may be amplified from specimens older than the 50–55 000 years BP radiocarbon limits. For example, nearly 100 of the bison sequences reported in Shapiro *et al.* (2004) were too old to be assigned finite radiocarbon ages. Second, for rapidly evolving taxa, the date of sampling may be unknown due to the loss or absence of accurate archival information. Even if the sampling date is known to the nearest year, it may be critical to estimate the isolation date more accurately. Finally, it may also be important to independently assess the authenticity of posited sampling dates due to their extreme age or because they are contentious.

Shapiro *et al.* (2011) developed and implemented a leaf-dating method that estimates the age or date of isolation of individual sequences within the Bayesian MCMC framework provided by the software package

BEAST (Drummond & Rambaut 2007). In this method, the sequence’s leaf age is treated as a random variable and an additional parameter for the age of the terminal node is introduced and treated identically to the internal node age parameters in terms of proposals made by the MCMC kernel. Methodologically, the leaf-dating method is similar to relaxing the constraints of the molecular clock on specific lineages within a phylogenetic tree. The method has been validated using both simulated and empirical data sets (Shapiro *et al.* 2011) and applied to different data sets to estimate the unknown age of some specimens before integrating them into classical tip-dating analyses (Gray *et al.* 2013; Stadler *et al.* 2013; Alter *et al.* 2015). For example, Faria *et al.* (2014) included a historical HIV strain from 1959 as an internal control and estimated its age, recovering an estimate centred on 1958 (95% BCI: 1946–1970). Although the estimated ages recovered are often associated with wide credible intervals, the leaf-dating method provides a means to include in molecular clock analyses sequenced samples for which little or no temporal information is available. Incorporating additional sequence data can improve the resolution of the phylogenetic, demographic and geographic history of the sampled sequences and can significantly extend the temporal range of the analysis.

Reconstructing transmission chains

Reconstructing transmission chains (i.e. who infected whom) using phylogenetic reconstruction methods has received considerable interest over recent time (Croucher & Didelot 2015). Although it should be stressed that while a phylogenetic tree and a transmission chain are both mathematical graphs representing the sequences as nodes connected by edges, they have radically different properties and are not interchange-

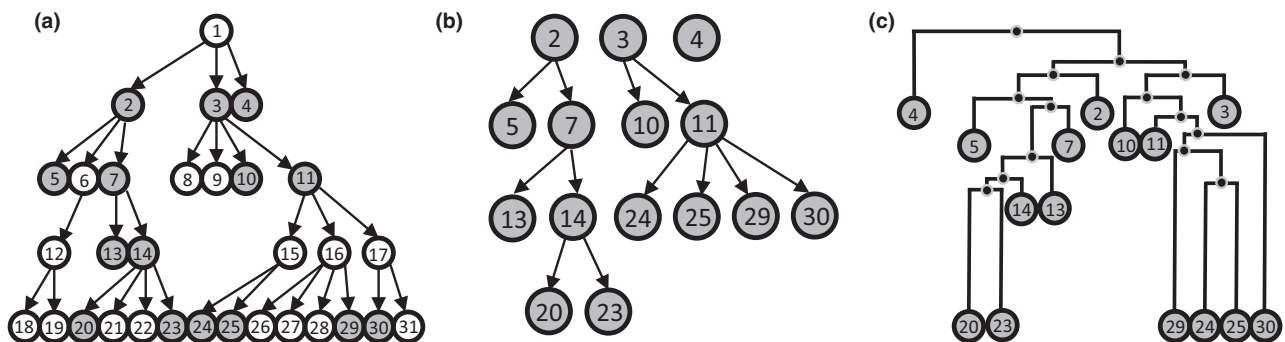


Fig. 3 Transmission graph vs. Phylogenetic tree. This figure adapted from Jombart *et al.* (2011) illustrates the difference between a transmission chain and a phylogenetic reconstruction. Panel a represents the transmission chain of a pathogen as arrows connecting hosts represented as circles, with grey circles representing sampled hosts. In panel (a, b) transmission graph (or network) is correctly reconstructed from the sampled hosts. In panel (c), a time-structured phylogeny is reconstructed using the same samples with black dots representing hypothetical ancestral isolates.

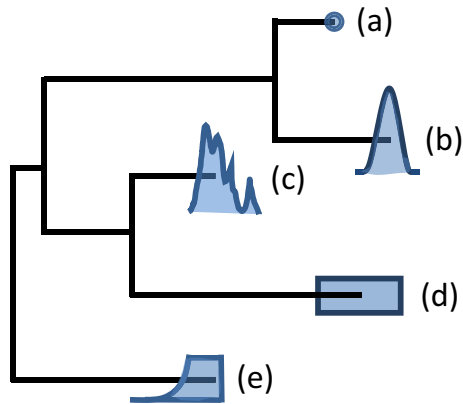


Fig. 4 Different statistical distributions to model uncertainty in tip calibrations inferences. Different distributions can be used to model the error associated with sampling dates. Choosing the best-suited one depends on the type of sample and the information associated with the dating method (Ho & Phillips 2009). Point values (a) can be used if the age of a sample is exactly known (e.g. sampling date). Modelling radiocarbon dating errors with a normal distribution (b) is common practice in ancient DNA studies even though recent improvement allow to use empirical description of the probability density function directly measured on the calibrated sample (c) (see Molak *et al.* 2015 for more details on this topic). Uniform distributions with hard minimum and maximum bounds (d) are suited to samples obtained from a well-defined stratum [e.g. ancient DNA retrieved from ice cores (Willerslev *et al.* 2007) or from samples associated with archaeological horizons (Edwards *et al.* 2007)] or to model uncertainty in sampling time accuracy (e.g. if the sampling month is known for some samples but not for others). Finally, uniform distribution with hard minimum and soft maximum bounds (e) can be suited to ancient DNA samples beyond the 45–50 ka resolution limit of radiocarbon dating (thus yielding a minimum age) for which additional information (e.g. from fossil data) exists and justifies the use of a soft maximum bound. This figure is adapted from Ho & Duchêne (2014).

able (Fig. 3) (Jombart *et al.* 2011; Didelot *et al.* 2014; Hartfield *et al.* 2014). Phylogenetic trees are binary graphs where the samples occupy the tips (terminal nodes) with each pair of samples connected by one node representing their putative most recent common ancestor (MRCA). Conversely, transmission chains are networks where each sample is represented by one node (internal or terminal) that can be connected to an arbitrary number of other nodes by edges representing between-host transmission events of the pathogen.

Phylogenetic reconstruction can in some circumstances inform on the transmission chain. However, a

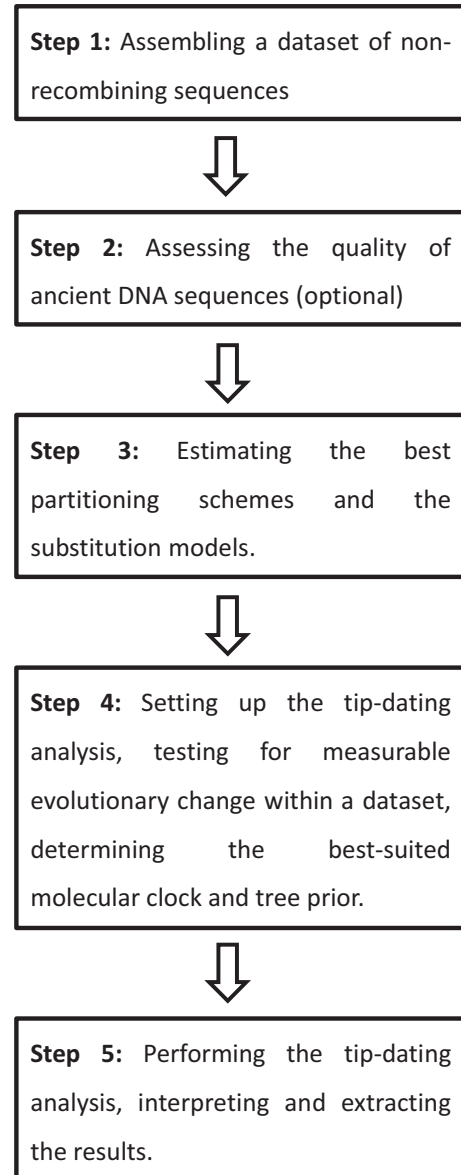


Fig. 5 Major steps to conduct accurate tip dating. This figure summarizes the five main steps that ought to be conducted when performing tip-dating analyses. For each of those steps, additional advices such as the important choices that must be made or the software to be used are given in the form of a practical guide available in Appendix S1 (Supporting information).

naïve phylogenetic analysis can lead to highly inaccurate inference of transmission under a series of realistic scenarios. If sampling is shallow so that only a small proportion of infected individuals are sampled, isolates connected in a phylogenetic framework are indirectly linked via a chain of unsampled hosts, and the branches in the tree will not represent direct transmission events. Conversely, in the case of densely sampled outbreaks and epidemics, another problem arises as some samples

can be both ancestors and descendants of other samples. This pattern is not well captured by a classical phylogenetic tree which considers all samples as descendants (tips) connected by inferred ancestors (nodes) (Fig. 3c). Additional problems arise for pathogens that remain infectious over a period of time (i.e. chronic infections), when a pathogen lineage continues to evolve within a host after the latter has infected other individuals. Finally, it has been shown that in the presence of within-host pathogen genetic diversity, a hallmark of most viruses and an increasingly widely recognized pattern in bacteria alike, phylogenetic reconstruction struggles to inform on the transmission chain (Worby *et al.* 2014).

While these represent major challenges, the field has seen some recent exciting developments aiming at adapting tip-based phylogenetic reconstruction to the reconstruction of transmission chains. Didelot *et al.* (2014) introduced a Bayesian inference scheme which allows superimposing the transmission network upon a phylogenetic tree applicable to well-sampled outbreaks while accounting for within-host evolution. Also of note, is the recently developed Bayesian MCMC algorithm by Gavryushkina *et al.* (2014) to infer sampled ancestor trees, that is trees in which sampled individuals can be direct ancestors of other sampled individuals. When analysing an HIV data set from the United Kingdom, they were able to detect sampled ancestors and estimate the probability that an individual will be removed from the process upon sampling (i.e. diagnosis). They could also demonstrate that even if sampled ancestors are not of specific interest in the analysis, failing to account for them leads to significant bias in the branching model and clock rate estimates.

Concluding remarks and future perspectives

Introduced over a decade ago, the concept of measurably evolving populations (MEPs), along with the analytical methodology it has spawned, has revolutionized our ability to study population dynamic and evolutionary processes using genetic data. Although only fast-evolving taxa such as RNA viruses were initially classifiable as MEPs, the recent rise in our ability to sequence DNA at high throughput both from modern and ancient material has opened up the field to a variety of additional organisms (Biek *et al.* 2015). Crucially, the sampling dates need to have sufficient temporal spread to capture measurable amounts of evolutionary change and perform tip dating thoroughly. In this context, we hope this review will encourage users of tip-dating methodologies (Fig. 5) to systematically investigate and report the temporal signal existing in their data set (see Section “To date or not to date – when is tip dating appropriate?”

and Appendix S1-D, Supporting information) to avoid rate and divergence time estimates to be mostly driven by prior information. We also hope to see more studies investigating the performance of the currently available tests to identify MEPs, such as the ones recently performed by Duchêne *et al.* (2015b) and Murray *et al.* (2015). Compelling directions for future research may target data sets with large number of modern sequences and a small set of ancient sequences, for which the date-randomization test might not be appropriate.

It is important to acknowledge that substantial advancements have been made in the field of molecular dating using heterochronous samples. Since the first attempts of incorporating noncontemporaneous sequences into maximum-likelihood (Rambaut 2000) and Bayesian (Drummond *et al.* 2002) frameworks made 15 years ago, many refinements have allowed for improved inference in a variety of biological systems. Those comprise relaxed molecular clocks, more realistic demographic models including the Bayesian skyline plot, the possibility to explicitly model the error associated with sampling times and the accumulation of post-mortem damage of DNA with time. Current molecular-dating models still fail to fully allow for joint reconstruction of the phylogeny and recombination patterns even though progress is being made towards that direction (McGill *et al.* 2013; O’Fallon 2013). However, such ancestral recombination graph (ARG) models will likely always be limited to situations where the rate of genetic recombination remains low as the phylogenetic paradigm is in essence not applicable to taxa undergoing extensive recombination. Additional compelling directions for future refinements may focus on spatially explicit modelling of population structure as the latter has been shown to be a factor distorting evolutionary rate and divergence time estimates. Finally, computational developments allowing for an improved use of highly parallel processors are also required to reduce calculation time that can be enormous when dealing with parameter-rich models and huge genomic data sets.

Acknowledgements

We thank Mark Achtman, Stephane Hue, Dan Jeffares, Gael Ker-goat, Angelis Konstantinos, Pierre Lefevre and Marie Pages for helpful discussions and constructive comments on previous versions of this manuscript. The work was funded by an ERC (ERC 260801) and a NERC (NE/M000338/1) grant to FB and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

References

- Alter SE, Meyer M, Post K *et al.* (2015) Climate impacts on transoceanic dispersal and habitat in gray whales from the Pleistocene to 2100. *Molecular Ecology*, **24**, 1510–1522.

- Arcila D, Pyron RA, Tyler JC, Orti G, Betancur-R R (2015) An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Molecular Phylogenetics and Evolution*, **82**, 131–145.
- Axelsson E, Willerslev E, Gilbert MTP, Nielsen R (2008) The effect of ancient DNA damage on inferences of demographic histories. *Molecular Biology and Evolution*, **25**, 2181–2187.
- Baele G, Lemey P, Bedford T *et al.* (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.
- Bauer E, Falque M, Walter H *et al.* (2013) Intraspecific variation of recombination rate in maize. *Genome Biology*, **14**, R103.
- Biek R, Pybus OG, Lloyd-Smith JO, Didelot X (2015) Measurably evolving pathogens in the genomic era. *Trends in Ecology & Evolution*, **30**, 306–313.
- Bollyky PL, Holmes EC (1999) Reconstructing the complex evolutionary history of hepatitis B virus. *Journal of Molecular Evolution*, **49**, 130–141.
- Bos KI, Schuenemann VJ, Golding GB *et al.* (2011) A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature*, **478**, 506–510.
- Bos KI, Harkins KM, Herbig A *et al.* (2014) Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature*, **514**, 494–497.
- Briggs AW, Good JM, Green RE *et al.* (2009) Targeted retrieval and analysis of five neandertal mtDNA genomes. *Science*, **325**, 318–321.
- Bromham L, Penny D (2003) The modern molecular clock. *Nature Reviews Genetics*, **4**, 216–224.
- Bunce M, Worthy TH, Phillips MJ *et al.* (2009) The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20646–20651.
- Buonagurio DA, Nakada S, Parvin JD *et al.* (1986) Evolution of human influenza-A viruses over 50 years – rapid, uniform rate of change in NS gene. *Science*, **232**, 980–982.
- Croucher NJ, Didelot X (2015) The application of genomics to tracing bacterial pathogen transmission. *Current Opinion in Microbiology*, **23**, 62–67.
- Cui Y, Yu C, Yan Y *et al.* (2013) Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 577–582.
- Der Sarkissian C, Allentoft ME, Avila-Arcos MC *et al.* (2015) Ancient genomics. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, **370**, 20130387.
- Didelot X, Wilson DJ (2015) ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Computational Biology*, **11**, e1004041.
- Didelot X, Eyre DW, Cule M *et al.* (2012) Microevolutionary analysis of *Clostridium difficile* genomes to investigate transmission. *Genome Biology*, **13**, R118.
- Didelot X, Gardy J, Colijn C (2014) Bayesian inference of infectious disease transmission from whole-genome sequence data. *Molecular Biology and Evolution*, **31**, 1869–1879.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Drummond A, Forsberg R, Rodrigo AG (2001) The inference of stepwise changes in substitution rates using serial sequence samples. *Molecular Biology and Evolution*, **18**, 1365–1371.
- Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, **161**, 1307–1320.
- Drummond A, Pybus OG, Rambaut A (2003a) Inference of viral evolutionary rates from molecular sequences. *Advances in Parasitology*, **54**, 331–358.
- Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG (2003b) Measurably evolving populations. *Trends in Ecology & Evolution*, **18**, 481–488.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, 699–710.
- Duchêne S, Holmes EC, Ho SYW (2014) Analyses of evolutionary dynamics in viruses are hindered by a time-dependent bias in rate estimates. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 1786.
- Duchêne D, Duchêne S, Ho SYW (2015a) Tree imbalance causes a bias in phylogenetic estimation of evolutionary timescales using heterochronous sequences. *Molecular Ecology Resources*, **15**, 785–794.
- Duchêne S, Duchêne D, Holmes EC, Ho SYW (2015b) The performance of the date-randomization test in phylogenetic analyses of time-structured virus data. *Molecular Biology and Evolution*, **32**, 1895–1906.
- Duchêne S, Ho SYW, Holmes EC (2015c) Declining transition/transversion ratios through time reveal limitations to the accuracy of nucleotide substitution models. *BMC Evolutionary Biology*, **15**, 36.
- Duffy S, Holmes EC (2009) Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *Journal of General Virology*, **90**, 1539–1547.
- Edwards CJ, Bollongino R, Scheu A *et al.* (2007) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1377–1385.
- Eldholm V, Monteserin J, Rieux A *et al.* (2015) Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. *Nature Communications*, **6**, 7119. doi: 10.1038/ncomms8119.
- Emerson BC, Alvarado-Serrano DF, Hickerson MJ (2015) Model misspecification confounds the estimation of rates and exaggerates their time dependency. *Molecular Ecology*, **24**, 6013–6020.
- Engering A, Hogerwerf L, Slingenbergh J (2013) Pathogen-host-environment interplay and disease emergence. *Emerging Microbes & Infections*, **2**, e5.
- Everitt RG, Didelot X, Batty EM *et al.* (2014) Mobile elements drive recombination hotspots in the core genome of *Staphylococcus aureus*. *Nature Communications*, **5**, 3956.
- Faria NR, Rambaut A, Suchard MA *et al.* (2014) The early spread and epidemic ignition of HIV-1 in human populations. *Science*, **346**, 56–61.
- Firth C, Kitchen A, Shapiro B *et al.* (2010) Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Molecular Biology and Evolution*, **27**, 2038–2051.
- Fourment M, Holmes EC (2014) Novel non-parametric models to estimate evolutionary rates and divergence times from heterochronous sequence data. *BMC Evolutionary Biology*, **14**, 163.

- Fraser C, Donnelly CA, Cauchemez S *et al.* (2009) Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*, **324**, 1557–1561.
- Frost SDW, Volz EM (2010) Viral phylodynamics and the search for an ‘effective number of infections’. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1879–1890.
- Fu Q, Meyer M, Gao X *et al.* (2013a) DNA analysis of an early modern human from Tianyuan Cave, China. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 2223–2227.
- Fu Q, Mittnik A, Johnson PLF *et al.* (2013b) A revised timescale for human evolution based on ancient mitochondrial genomes. *Current Biology*, **23**, 553–559.
- Gavryushkina A, Welch D, Stadler T, Drummond AJ (2014) Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Computational Biology*, **10**, e1003919.
- Gilbert MTP, Drautz DI, Lesk AM *et al.* (2008) Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 8327–8332.
- Gray RR, Tanaka Y, Takebe Y *et al.* (2013) Evolutionary analysis of hepatitis C virus gene sequences from 1953. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**, 20130168.
- Hartfield M, Murall CL, Alizon S (2014) Clinical applications of pathogen phylogenies. *Trends in Molecular Medicine*, **20**, 394–404.
- Hedge J, Wilson DJ (2014) Bacterial phylogenetic reconstruction from whole genomes is robust to recombination but demographic inference is not. *MBio*, **5**, e02158–14.
- Heintzman PD, Zazula GD, Cahill JA *et al.* (2015) Genomic data from extinct North American camelops revise camel evolutionary history. *Molecular Biology and Evolution*, **32**, 2433–2440.
- Heller R, Chikhi L, Siegmund HR (2013) The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS One*, **8**, e62992.
- Henn BM, Gignoux CR, Feldman MW, Mountain JL (2009) Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. *Molecular Biology and Evolution*, **26**, 217–230.
- Ho SYW, Duchêne S (2014) Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology*, **23**, 5947–5965.
- Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.
- Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.
- Ho SYW, Kolokotronis S-O, Allaby RG (2007a) Elevated substitution rates estimated from ancient DNA sequences. *Biology Letters*, **3**, 702–705.
- Ho SYW, Shapiro B, Phillips MJ, Cooper A, Drummond AJ (2007b) Evidence for time dependency of molecular rate estimates. *Systematic Biology*, **56**, 515–522.
- Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS One*, **3**, e1615.
- Ho SYW, Lanfear R, Bromham L *et al.* (2011a) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**, 3087–3101.
- Ho SYW, Lanfear R, Phillips MJ *et al.* (2011b) Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Systematic Biology*, **60**, 366–374.
- Holden MTG, Hsu L-Y, Kurt K *et al.* (2013) A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Research*, **23**, 653–664.
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.
- Jombart T, Balloux F, Dray S (2010) adephylo: new tools for investigating the phylogenetic signal in biological traits. *Bioinformatics*, **26**, 1907–1909.
- Jombart T, Eggo RM, Dodd PJ, Balloux F (2011) Reconstructing disease outbreaks from genetic data: a graph approach. *Heredity*, **106**, 383–390.
- Jonsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, **29**, 1682–1684.
- Korber B, Muldoon M, Theiler J *et al.* (2000) Timing the ancestor of the HIV-1 pandemic strains. *Science*, **288**, 1789–1796.
- Krause J, Fu Q, Good JM *et al.* (2010) The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*, **464**, 894–897.
- Lanfear R, Welch JJ, Bromham L (2010) Watching the clock: studying variation in rates of molecular evolution between species. *Trends in Ecology & Evolution*, **25**, 495–503.
- Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.
- Li WH, Tanimura M, Sharp PM (1988) Rates and dates of divergence between AIDS virus nucleotide sequences. *Molecular Biology and Evolution*, **5**, 313–330.
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, **1**, vev003.
- McGill JR, Walkup EA, Kuhner MK (2013) GraphML specializations to codify ancestral recombinant graphs. *Frontiers in Genetics*, **4**, 146-Article No.: 146.
- McVean G, Awadalla P, Fearnhead P (2002) A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics*, **160**, 1231–1241.
- Millar CD, Huynen L, Subramanian S, Mohandesan E, Lambert DM (2008) New developments in ancient genomics. *Trends in Ecology & Evolution*, **23**, 386–393.
- Molak M, Lorenzen ED, Shapiro B, Ho SYW (2013) Phylogenetic estimation of timescales using ancient DNA: the effects of temporal sampling scheme and uncertainty in sample ages. *Molecular Biology and Evolution*, **30**, 253–262.
- Molak M, Suchard MA, Ho SYW, Beilman DW, Shapiro B (2015) Empirical calibrated radiocarbon sampler: a tool for incorporating radiocarbon-date and calibration error into Bayesian phylogenetic analyses of ancient DNA. *Molecular Ecology Resources*, **15**, 81–86.
- Murray GGR, Wang F, Harrison EM *et al.* (2015) The effect of genetic structure on molecular dating and tests for temporal signal. *Methods in Ecology and Evolution*, **7**, 80–89.

- Navascues M, Emerson BC (2009) Elevated substitution rate estimates from ancient DNA: model violation and bias of Bayesian methods. *Molecular Ecology*, **18**, 4390–4397.
- O'Fallon BD (2013) ACG: rapid inference of population history from recombining nucleotide sequences. *BMC Bioinformatics*, **14**, 40.
- Parrish CR, Holmes EC, Morens DM *et al.* (2008) Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews*, **72**, 457–470.
- Pyron RA (2011) Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia. *Systematic Biology*, **60**, 466–481.
- Rambaut A (2000) Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics*, **16**, 395–399.
- Rambaut A, Ho SYW, Drummond AJ, Shapiro B (2009) Accommodating the effect of ancient DNA damage on inferences of demographic histories. *Molecular Biology and Evolution*, **26**, 245–248.
- Ramsden C, Melo FL, Figueiredo LM *et al.* (2008) High rates of molecular evolution in hantaviruses. *Molecular Biology and Evolution*, **25**, 1488–1492.
- Rasmussen S, Allentoft ME, Nielsen K *et al.* (2015) Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell*, **163**, 571–582.
- Rieux A, Eriksson A, Li M *et al.* (2014) Improved calibration of the human mitochondrial clock using ancient genomes. *Molecular Biology and Evolution*, **31**, 2780–2792.
- Ronquist F, Klopfstein S, Vilhelmsen L *et al.* (2012a) A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. *Systematic Biology*, **61**, 973–999.
- Ronquist F, Teslenko M, van der Mark P *et al.* (2012b) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.
- Sanchez-Buso L, Comas I, Jorques G, Gonzalez-Candelas F (2014) Recombination drives genome evolution in outbreak-related *Legionella pneumophila* isolates. *Nature Genetics*, **46**, 1205–1211.
- Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, **14**, 1218–1231.
- Shankarappa R, Margolick JB, Gange SJ *et al.* (1999) Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *Journal of Virology*, **73**, 10489–10502.
- Shapiro B, Drummond AJ, Rambaut A *et al.* (2004) Rise and fall of the Beringian steppe bison. *Science*, **306**, 1561–1565.
- Shapiro B, Ho SYW, Drummond AJ *et al.* (2011) A Bayesian phylogenetic method to estimate unknown sequence ages. *Molecular Biology and Evolution*, **28**, 879–887.
- Sheng G-L, Soubrier J, Liu J-Y *et al.* (2014) Pleistocene Chinese cave hyenas and the recent Eurasian history of the spotted hyena, *Crocuta crocuta*. *Molecular Ecology*, **23**, 522–533.
- de Silva E, Ferguson NM, Fraser C (2012) Inferring pandemic growth rates from sequence data. *Journal of the Royal Society Interface*, **9**, 1797–1808.
- Smukowski CS, Noor MAF (2011) Recombination rate variation in closely related species. *Heredity*, **107**, 496–508.
- Spagnoletti M, Ceccarelli D, Rieux A *et al.* (2014) Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic *Vibrio cholerae* lineage. *MBio*, **5**, e01356–14.
- Stack JC, Welch JD, Ferrari MJ, Shapiro BU, Grenfell BT (2010) Protocols for sampling viral sequences to study epidemic dynamics. *Journal of the Royal Society Interface*, **7**, 1119–1127.
- Stadler T, Kuehnert D, Bonhoeffer S, Drummond AJ (2013) Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 228–233.
- Stadler T, Kuehnert D, Rasmussen DA, du Plessis L (2014) Insights into the early epidemic spread of ebola in sierra leone provided by viral sequence data. *PLoS Currents*, **6**.
- Stiller M, Baryshnikov G, Bocherens H *et al.* (2010) Withering away-25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution*, **27**, 975–978.
- Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution*, **15**, 1647–1657.
- Weinert LA, Welch JJ, Suchard MA *et al.* (2012) Molecular dating of human-to-bovid host jumps by *Staphylococcus aureus* reveals an association with the spread of domestication. *Biology Letters*, **8**, 829–832.
- Welch JJ, Bromham L (2005) Molecular dating when rates vary. *Trends in Ecology & Evolution*, **20**, 320–327.
- Willerslev E, Cappellini E, Boomsma W *et al.* (2007) Ancient biomolecules from deep ice cores reveal a forested Southern Greenland. *Science*, **317**, 111–114.
- Worby CJ, Lipsitch M, Hanage WP (2014) Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. *PLoS Computational Biology*, **10**, e1003549.
- Xia X (2013) DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution*, **30**, 1720–1728.
- Yahara K, Didelot X, Ansari MA, Sheppard SK, Falush D (2014) Efficient inference of recombination hot regions in bacterial genomes. *Molecular Biology and Evolution*, **31**, 1593–1605.
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.
- Zuckerkandl E, Pauling L (1962) *Molecular Disease, Evolution, and Genetic Heterogeneity*. Academic Press, New York.

A.R. and F.B. contributed equally to the writing of this review.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1. A step-by-step practical guide to conduct accurate tip-dating.

Appendix S2. Manual modifications to perform to the xml-input files.