

Locating Uncertainty in Stochastic Evolutionary Models: Divergence Time Estimation

Charles H. Pence

Received: date / Accepted: date

Keywords probability · uncertainty · evolutionary theory · divergence time · scientific modeling · stochastic models · Bayesian models

Abstract Philosophers of biology have worked extensively on how we ought best to interpret the probabilities which arise throughout evolutionary theory. In spite of this substantial work, however, much of the debate has remained persistently intractable. I offer the example of Bayesian models of divergence time estimation (the determination of when two evolutionary lineages split) as a case study in how we might bring further resources from the biological literature to bear on these debates. These models offer us an example in which a number of different sources of uncertainty are combined to produce an estimate for a complex, unobservable quantity. These models have been carefully analyzed in recent biological work, which has determined the relationship between these sources of uncertainty (their relative importance and their disappearance in the limit of increasing data), both quantitatively and qualitatively. I suggest here that this case shows us the limitations of univocal analyses of probability in evolution, as well as the simple dichotomy between “subjective” and “objective” probabilities, and I conclude by gesturing toward ways in which we might introduce more sophisticated interpretive taxonomies of probability (modeled on some recent work in the philosophy of physics) as a path toward advancing debates on probability in the life sciences.

1 Introduction

One of the oldest and most persistent debates in the philosophy of biology concerns the status of the probabilities that evolutionary theory seems constantly to employ. Are they objective or subjective? And whichever answer to that question we choose, how are they grounded – on the basis of which facts in the biological world (or about human knowers) do they rest? The problem has been tackled in a variety of ways. First, we have the exploration of probabilities as they arise in various kinds of biological processes – to name only a few examples, there has been significant discussion of the “randomness” of mutation (Stamos

C.H. Pence
Institut supérieur de philosophie, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium,
E-mail: charles@charlespence.net

2001; Merlin 2010, 2016), of the influence (or lack thereof) of probabilistic causation in the processes of natural selection and genetic drift (Matthen and Ariew 2002; Millstein 2006; Brandon and Ramsey 2007; Millstein 2016; Walsh et al 2017), the nature of “historicity” or path-dependence in biological systems (Beatty and Desjardins 2009; Desjardins 2011, 2016), and the way in which factors external to an evolving system, such as the environment, should be considered (Brandon 1990; Abrams 2009; Lenormand et al 2009).

Another literature has attempted to apply the extensive discussion of the interpretation of probability in general philosophy of science to evolution in particular (Millstein 2011; Drouet and Merlin 2013). Several articles have turned to recent work on “mechanistic” interpretations of probability (to use Marshall Abrams’s apt phrase) as a tool for understanding biological probabilities in a way which is neither classically subjective (i.e., not an ignorance interpretation deriving only from our lack of detailed understanding of biological systems) nor classically objective (in the sense, for example, of “brute” probabilities arising from quantum mechanics). This work has shown some real promise, for instance, when applied to the cases of fitness and genetic drift (Abrams 2012b,a; Strevens 2016), detailing analyses of micro-causal structure that could produce the observed patterns of probabilistic causation that systems experiencing selection and drift exhibit. (We will return to these interpretations in section 4.)

One feature of this body of work, however, is troubling. While obtaining the correct interpretation of these probabilities is doubtless an important enterprise, it is often a distinct challenge to find cases where the biological literature makes genuine contact with the question of interpretations of probability in evolution. In the absence of this interaction, philosophers are often “on our own” in pursuing questions of chance and evolution. To take just one example, the debate over the interpretation of the probabilities at work in natural selection, genetic drift, and fitness – often called the “causalist vs. statisticalist” debate – has tended to be cashed out with highly abstracted examples from roulette wheels (Strevens 2016) or coin tosses (Walsh 2007). Only occasionally has an author argued for a position’s superiority from biological examples (e.g., Ariew and Lewontin 2004; Millstein 2008), and the correct reading of these examples tends to be hotly disputed (Otsuka 2016).

We would be well served, then, by a search for case studies in which biological practice gives us a window into the source and status of evolutionary probabilities, and it is precisely my aim in this paper to explore such a case. When we examine the ways in which biologists estimate the divergence times of lineages (a major endeavor in the last decade of biological research), we see that we can, in fact, clearly distinguish the impact of probabilities which are the result of our ignorance of contemporary sequence data from those that are the result of our (ineliminable) ignorance about the deep evolutionary past. My hope here, then, is twofold. First, this case shows us a way in which we can precisely analyze the source and status of biological probabilities in a real-world example with empirical relevance. And second, I am cautiously optimistic that the general approach which I pursue here, which draws lessons for our understanding of biological probability directly from biological models, offers a way in which we might build a pluralistic and empirically informed method for understanding the role of probability in evolutionary biology.

I will begin by introducing the basic idea of divergence time estimation, describing the history of increasing sophistication in these models. I then lay out the results concerning the sources of uncertainty in our data that can be derived from one set of these models.¹ Next,

¹ I thank a reviewer for noting that the mere use of the term ‘uncertainty’ here, which I have imported from the biological literature in an effort to avoid confusion with these sources, carries a strong philosophical implication that these probabilities are merely subjective; I will argue against this interpretation in what follows.

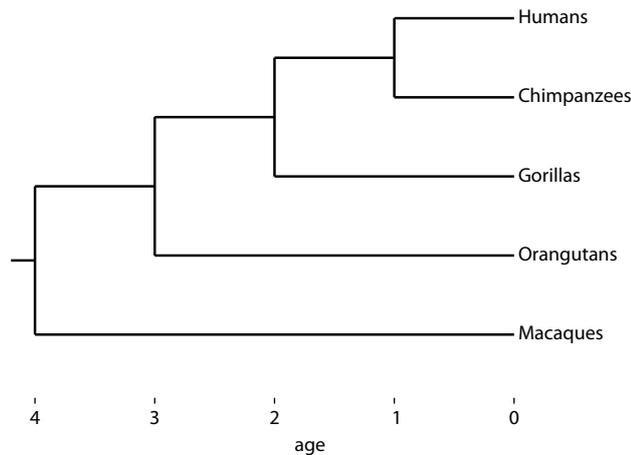


Fig. 1 A basic phylogenetic tree for humans and their closest relatives, with arbitrary branch-lengths.

I turn to the lessons on interpretation that we might draw from the probabilities attached to their results, and I conclude with some thoughts on the analysis of probabilities in biology more generally.

2 Building Models of Divergence Time Estimation

Imagine that we are interested in the evolutionary relationship between humans and our closest relatives. On the basis of speculation and some morphological data, we might well infer a basic tree structure like that found in figure 1. But this tree is significantly underspecified. In particular, we would like to obtain data that can confirm the vertical *distance* between the nodes (ensuring that our branches are in fact correct, e.g., that chimpanzees really are more closely related to humans than are gorillas), and we would like to know the *times* at which those divergences took place (e.g., whether humans diverged from chimpanzees much more recently than their common ancestor did from gorillas).

There are two primary sources of data to which we can appeal in this case. First, we have fossil specimens. The difficulty of dating phylogenetic trees using sporadic fossil data is not to be underestimated.² The dating of the fossils themselves can be difficult, and the dates ascribed to multiple instances of fossils from the same clade need to be combined to offer a best estimate for the time of the clade's appearance. Each fossil's date, in turn, bears at least some error due to the inherently imperfect nature of the fossil record and dating methods, and there is an active debate over whether and how to take account of estimates of this error.

Importantly, it arises from different sources on either side of a hypothesized branching event. For the minimum age of a clade, we have simply the inherent error in determining the age of a particular rock formation and the fossils which it contains. For the maximum age of a clade, on the other hand, we add to this first source of error the likelihood that the

² For a host of further details, as well as worked-out examples for a number of clades, the interested reader can consult Benton et al (2009).

branching event was in fact earlier than the appearance of the earliest fossil at issue – as it is vanishingly improbable that any one particular fossil was the first representative of its clade. As Benton et al (2009, p. 40) put it, this results in “relatively secure ‘hard’ minimum constraints on a particular branching point, and a soft maximum constraint.” This asymmetry makes epistemic sense. If we see a fossil at a given date, this can let us confidently assert a minimum date for the clade at issue (there is no way that the clade could have appeared *after* one of its members had lived!). But there simply is no fossil evidence which *could* speak to the maximum age of the clade. Our fossil evidence for the clade’s existence slowly becomes less common as we move backwards in time, and ultimately disappears. We are forced, then, to use probability densities for these maximum ages, estimated from models of fossil preservation, for instance.

What we are left with, in the end, are bounds which tend to be very imprecise. For example, our best current data indicates that the divergence of the Hominidae (the clade including orangutans, gorillas, chimpanzees, and humans, i.e., the divergence at the point labeled ‘3’ in figure 1) may be constrained by fossil evidence alone to somewhere between 33.7 and 11.2 million years ago (Benton et al 2009, p. 48), and the human-chimpanzee divergence between 10 and 5.7 million years ago (Benton et al 2009, p. 46).

The second source of data that we use to calibrate these trees comes from the genetic sequences of the extant organisms from the tips of each of the tree’s branches. These sequences should allow us to determine how “far apart” these organisms are. In turn, because we can plausibly infer that the farther apart two organisms are, the more time it should have taken for evolution to produce the apparent divergence, these sequences should be able to inform our estimation of divergence times. In order to make this vague notion more precise, however, we need to make clear these vague notions of distance, and control for the obvious interfering effect of natural selection in driving evolutionary outcomes.

Here enters the concept of the molecular clock. The fundamental idea, as originally proposed by Zuckerkandl and Pauling (1965), rests on the observation that, if mutations in areas of the genome that are not under active selection accumulate at a roughly constant rate over time, the genetic distance between organisms at such loci should offer us a rough measure of the time since they diverged from their last common ancestor. Many of the models for the molecular clock which have been proposed have, in turn, relied on a significant number of assumptions, which I lack the space to explore in detail here – suffice it to say that it remains a topic of open debate whether we can in fact make sense of a “molecular clock” at all (for a review of the argument, see Schwartz and Maresca 2006).

While these general critiques will not be my focus, one of the problems inherent to any understanding of the molecular clock will be important for us going forward. Consider how we might represent divergence as a simple, dynamical equation in this context. We hypothesize that there has been a split between two lineages at some unknown time in the past, and each of these lineages, then, has continued on diverging until the present, resulting in them becoming some distance apart from one another (the sense in which we should interpret this distance will be made more precise in a moment). We can thus say that the distance of divergence = $2 \times$ the divergence rate \times the elapsed time since divergence.³ The issue, then, is this: divergence time only appears in any dynamical model like this when multiplied by the rate of change. We cannot ever observe it directly – only its effects in terms of distance between extant organisms, accumulated at some also-unknown rate of change.⁴

³ The constant factor of two, here, indicates that *both* lineages have continued to diverge since the original branching event.

⁴ And, as already discussed, its traces in the fossil record, though we are not considering those at the moment. We will return to the question of combining fossil and molecular data shortly.

The rate and the time, then, are what is known as *nonidentifiable* – it is impossible to obtain precise values for rate and time, even in the presence of an infinite number of observations of distances.

How, then, are these distances and rates to be defined and understood? The distance, here, is taken to be the true number of substitutions that brought us from the original point of divergence to one or the other of the extant organisms that we are now able to sequence. The rate, in turn, is the speed at which these substitutions took place over evolutionary time. How quickly, that is, will portions of the sequence of an organism tend to change in the absence of any particular selective pressure?

The challenge here lies in accurately estimating the rate of divergence. The first element of a model for the rate is a substitution model – a model for the manner in which genetic substitutions will accumulate in a particular sequence, in the absence of selection. There is a massive diversity in these models, dating back a number of decades – far more than I could enumerate in a brief discussion here (see, e.g., Felsenstein 2004, ch. 13). The simplest such model, which dates to Jukes and Cantor (1969), assumes that all base pairs within a sequence hold an equal probability of substitution to any other base pair, at any time. There are a variety of ways to make this model more complex. To take just two examples, we might permit the rate of substitution to vary according to the type of substitution, accounting for the biochemical fact that transitions (when a locus moves from A to G or from C to T) occur significantly more often than transversions (all other types of substitution; see Hasegawa et al 1985). At the cost of even more complexity, we might allow the parameters of the substitution model to vary based on the particular locus considered (Halpern and Bruno 1998; Yang and Rannala 2006).

A complete model for the rate, however, can be even more complex than a simple single substitution model, because it may be the case that the substitution model itself varies across time and across the tree of life. A number of models have attempted to work with a “relaxed clock” (Drummond et al 2006; Lepage et al 2007), on which the clock rate is permitted to vary with time or by clade. Data from a variety of lineages suggests that the rate of the molecular clock is itself a parameter which is under the control of evolution (Gillespie 1984, 1991), and lineage-specific molecular clock rates have been modeled in a variety of different ways (Rannala and Yang 2007). The possibility that these rates are correlated between neighboring clades has also been evaluated, without a clear result (for two opposing perspectives, see Lepage et al 2007; Linder et al 2011).

As already noted, however, because rates and times are confounded in expressions for distance, we will require probabilistic models for estimating the true value of the divergence time. In the last two decades, work has centered on Bayesian models for divergence time estimation (Nascimento et al 2017). Our best resource to minimize the impact of confounding involves, then, the combination of all three of the major resources we have discussed so far – fossil data, current genetic distances, and models of the molecular clock – to produce an overall model for estimating divergence times. In particular, if we use a Bayesian model for the inference of divergence times, we need to be able to set our prior probabilities on the various possible values for divergence time. It is here that the fossil data can enter, providing values for our priors which should permit the inference of divergence times with significantly higher quality.

Methods for setting these priors have seen impressive improvement in recent years. Early models simply took the fossil dates to provide priors with zero uncertainty – the oldest extant fossil known *just is* our best prior for the age of the given node (Thorne and Kishino 2002). As discussed above with regard to fossil calibration, however, this significantly misrepresents the quality and nature of the fossil evidence, conflating the “hard” estimate of a

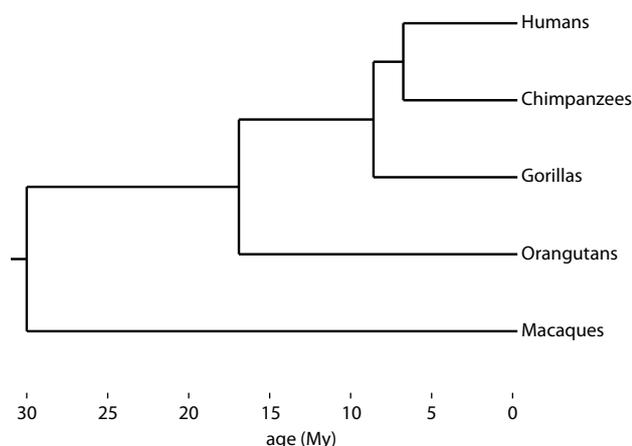


Fig. 2 A phylogenetic tree for humans and their closest relatives, with accurate branch-lengths. The values of and 95% confidence intervals for each node, from most recent to oldest, are: $8.06 > t_1 = 7.05 > 5.79$; $10.2 > t_2 = 8.89 > 7.31$; $19.7 > t_3 = 17.2 > 14.1$; $34.7 > t_4 = 30.3 > 24.9$. Data from dos Reis and Yang (2013, p. 38, table 5).

minimum age with a singular estimate of clade appearance. From a technical perspective, as well, the absence of uncertainties in these priors significantly affects the quality of the divergence times estimated. A next generation of models, then, took the fossil record to provide hard bounds on node ages – either only a hard minimum bound on the age of a node (Sanderson 1997), or both minimum and maximum bounds on the age of a node (Thorne et al 1998) – outside of which we set the probability of finding the true divergence to zero.

More sophisticated models, in turn, have investigated using “soft bounds” (priors with non-zero probabilities outside the fossil calibration range) as a way to better account for fossil uncertainty, both with a per-locus-varying molecular clock (Yang and Rannala 2006) and a per-lineage varying molecular clock (Rannala and Yang 2007). In particular, these priors allow us to account for the fact, also discussed above, that maximum ages for a clade have at least some chance to be larger (or *far* larger) than any of the fossils which we have currently discovered. (The result of applying the per-locus-varying model of the rate from Yang and Rannala (2006), along with fossil priors from Benton et al (2009), to the primate phylogenetic tree in figure 1 is provided in figure 2.)

Recent work has endeavored to integrate even more information into these estimates, as there is no in-principle reason not to move beyond the twin sources of fossil and sequence data. Wilkinson et al (2011), for example, use extant data about clade diversity to add information to our model about the likely process of fossil deposition and recovery within each lineage. Ronquist et al (2012), most ambitiously, construct what they call a “total-evidence” model, integrating genetic sequences, coded morphological data, and fossil evidence to produce divergence-time estimates.

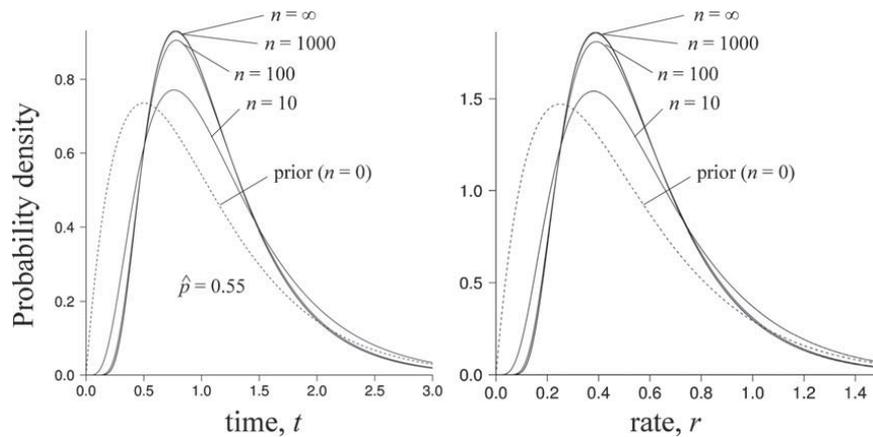


Fig. 3 Graph of the prior probability and the posterior probability distributions for both time t and rate r , given an increasing number n of sites. Lines for $n = 1000$ and $n = \infty$ are almost identical. For the details of derivation, which are not important to the point here, see dos Reis and Yang (2013, fig. 4), appears courtesy John Wiley and Sons.

3 Understanding Uncertainty

As can be readily seen from the brief history presented above, one of the main drivers of increasing complexity in models of divergence time estimation has been a persistent desire to understand the *uncertainty* in fossil calibrations, clock rates, and resultant divergence times. In a scathing critique entitled “Reading the Entrails of Chickens,” Graur and Martin “conclude that all divergence estimates discussed here are without merit” (2004, p. 85), offering a list of thirteen references, each of which fails to establish its conclusions, they argue, “mainly through the conversion of statistical estimates (which by definition possess standard errors, ranges and confidence intervals) into errorless numbers” (p. 80). Their paper concludes by advising the reader: “whenever you see a time estimate in the evolutionary literature, demand uncertainty!” (Graur and Martin 2004, p. 85).

At the very least, there is clearly an opening here for philosophical analysis. We have, as we often do, a process of biological inference which results in estimates for a quantity with significant errors attached. Further, we have a wide variety of sources of uncertainty that drive this need for probabilistic modeling, each of which have been approached by biologists in a variety of different ways, across an active and highly theoretically self-reflective literature. Let’s now turn to some results due to Yang, Rannala, and dos Reis, which allow us to not just qualitatively investigate some of these sources of uncertainty, but quantify them, along with their interrelationships.

If we want to more carefully analyze the probabilities at issue here, one fruitful first question to ask would be the following: how much of the uncertainty in our resultant estimations of divergence times is the result of our ignorance of contemporary sequence data, and how much is the result of the fossil calibrations? The first paper to extensively consider this question was Yang and Rannala (2006), who derived that, even in the presence of infinite sequence data, the uncertainty in divergence times does *not* fall to zero, as any error in fossil calibration remains impossible to eradicate. While this paper uses a molecular clock that varies only per-locus, they followed up a year later with a per-lineage-varying clock (Rannala and Yang 2007), and obtained roughly the same infinite-sites results.

More detail was then offered by dos Reis and Yang (2013), who were the first to investigate the approach to the infinite-sites limit. How, that is, would our estimates for divergence times change as we move from a small quantity of sequence data toward an infinite amount? They explored this both in the context of a single, correct fossil calibration (with uncertainty), and given multiple, possibly mutually inconsistent fossil calibrations.⁵ The convergence, as it turns out, occurs quite quickly – the $n = 100$ plot is very near the infinite-sites limit, and the $n = 1000$ plot is indistinguishable from the infinite-sites plot. The plot converges not to a single point, but rather a somewhat more precise ridge. Their results may be seen in figure 3, which shows the approach to the case of infinite sites for $n = 10, 100,$ and 1000 .

The precise derivation of these graphs is not important for our purposes. Rather, let's consider what we might learn from the fact that this sort of analysis is possible here. The fact that the uncertainty in these models is ineliminable is, on the one hand, an unsurprising consequence of the nonidentifiability of rates and times. But to stop the analysis at this point significantly underplays the interesting nature of these results. Several points are worth examining in some detail.

First, because we have both the infinite-sites model and the approximations to it via finite data, we are, in essence, able to partition the uncertainty which appears in these divergence time estimates. The approach to the infinite-sites model is carried out by adding further contemporary sequence data. Whatever uncertainty is *eliminated* in the infinite-sites model, therefore, is straightforwardly a result of our ignorance of further sequence information. The remaining uncertainty, on the other hand (i.e., the failure of the $n = \infty$ curves in figure 3 to collapse to single values), is a direct result of the historical inaccessibility of the evolutionary past and the intrinsic difficulties in accurately identifying and dating fossils, along with the effect of nonidentifiability.

As was discussed briefly above (see also Yang and Rannala 2006, p. 214), the process of accurately dating a fossil comes with a number of potential hazards. Simple experimental error may cause problems in radiometric dating or an incorrect geological assignment. More importantly, the relationship between a particular fossil and the clade under investigation may be difficult to discern. A fossil might lie on an extinct side branch, rather than being ancestral to the clade at issue, or it may be hard to determine whether a fossil in fact precedes a divergence. Lastly, it is likely that even if we were in possession of a complete and accurate catalog of the available fossil evidence, we would still face uncertainty arising from the epistemic inaccessibility of the deep evolutionary past.

We thus have, at the very least, the following sources of uncertainty at work in divergence time estimations: the lack of further contemporary sequence data, the various factors that lead to uncertainty in fossil priors, and the mathematical fact of nonidentifiability itself. While I want to reserve further discussion of this point for the next section (in the context of the interpretation of the probabilities that arise in this context), it is worth noting briefly now that these various sources of uncertainty are clearly not identical, and we thus have the opportunity to interpret each of them in different ways. In particular, the uncertainty arising from the fossil record is, I will argue, more than simply a source of subjective probabilities best described by a traditional ignorance interpretation.

Second, because the finite-data approach can give us an idea of how quickly the system approaches the infinite-data limit, this allows us, in effect, to determine the relative importance of each of these two types of uncertainty. The fact that adding further sequence

⁵ Their model uses a per-locus-varying molecular clock, though, as they note, a relaxed-clock model is only likely to add further uncertainty.

data very rapidly converges upon the infinite-data limit indicates that any contemporary sequence uncertainty is, in general, dominated by the uncertainty in the fossil data. The extent to which this is true, in fact, has been surprising to practitioners. Mulcahy et al, for example, noted that there was only a minor difference between a divergence time estimate with 2 and 25 sites, which they referred to as “a somewhat disturbing result that should be investigated further” (2012, p. 989). There is an expectation among biologists, it seems, that the overall uncertainty would be more sensitive to the addition of contemporary sequence data than it in fact is.

Finally, it is worthy of note that – while the calculation for real cases of divergence times is exceptionally difficult – we can quantify precisely this partitioning by determining how much of the variance in the posterior distribution of divergence times arises from the lack of sequence data and how much arises from the fossil priors. Dos Reis and Yang estimate this value using a generalized Bayesian heuristic and predict that the proportion u_F of variance due to fossil calibration will equal the ratio of the squares of the confidence interval widths w for the infinite- and finite-sites models, $u_F = w_\infty^2/w^2$, and the proportion of variance due to sequence data will thus be $u_S = 1 - w_\infty^2/w^2$. While the details are not of vital importance here, the fact that the source of these probabilities can be not only conceptually and qualitatively but also quantitatively assessed is an exciting possibility for philosophical analysis.

4 Model Uncertainties and Biological Probabilities

So much for the formal results concerning the sources of uncertainty in divergence-time models. It's now time to turn to the upshot of this debate for the philosophical analysis of the probabilities that occur in our resulting estimations of divergence times. Here, I think the lessons that divergence time estimation teaches us have an impact beyond our understanding of the particular model of dos Reis and Yang.

I should begin with an effort to clarify my terms, as the debate over the interpretation of probability (despite its age) has yet to settle on a common terminology. A view commonly found in the literature separates probabilities into exactly two categories, which I have been referring to as ‘subjective’ and ‘objective’ thus far. Subjective probabilities, then, are intended to quantify our uncertainty about a concrete fact in the world. Because they are linked to our uncertainty, they respond to the addition of further data about the phenomenon concerned. Objective probabilities, on the other hand, quantify probabilistic facts that are genuinely parts of the structure of the world. No amount of further data one might provide could, then, reduce their influence. On this common (perhaps cartoon) view, these two categories exhaust the space of probabilities. My goal in this section, then, is to argue for the claim that the case study of divergence time estimation can be added to the list of those which make trouble for this classic classification of probability.

As a starting point for making my case, consider a fairly common argument for the claim that the probabilities we find in evolutionary biology are, and must be, subjective. A number of authors, but perhaps most especially Alex Rosenberg (Rosenberg 1994; Graves et al 1999), have argued that all of the probabilities deployed by evolution are subjective, because “agents with our cognitive limitations and our interests can only use the theory for explanations and predictions by imposing [subjective] probabilities on them” (Graves et al 1999, p. 141). The biological world, the argument goes, is inescapably messy. Biological explanations must abstract away from this detail, “imposing” probabilities on various quantities in order to cover our ignorance of the fine details of the biological data (Matthen 2009).

As a result, we should by no means be surprised that probabilities feature in biological explanations. If they did not, there simply would *be no* biological explanations.

Without a doubt, some probabilities in biology are amenable to this sort of analysis. Consider, for example, large-scale modeling work in ecology, which aims to integrate data from a wide variety of sources to make generalized conclusions about real-world populations. Such models inevitably rely on precisely the sort of simplification-over-messiness that this argument invokes (Weisberg 2014). The kinds of causal interactions present in any real-world ecological system are quite probably permanently beyond our power to comprehend in anything like full detail. We have no choice, then, but to generalize away from them, summarizing them via a variety of probabilistic influences (likely rates of predation, expected carrying capacities, and so forth).

But should we accept the conclusion, as some have urged, that this is therefore the correct interpretation of *all* the probabilities in evolutionary theory? Divergence time estimation provides an impressive example of a case where finer-grained analysis offers fruitful results, results that would be missed by a univocal analysis of biological probability. As we saw above, precisely the advantage of this case is that we can readily analyze a variety of sources of uncertainty in the biological world which each have an effect on our final divergence time estimates. This analysis lets us extract (at least) two interesting morals.

First, whatever the merits in particular cases of an argument concerning abstraction in the face of biological messiness, appealing to any such argument in this case (and, by extension, as a general panacea for all biological probabilities) seems to be a mistake. We've seen here a vast amount of research on the various sources of uncertainty which have been elucidated by the biologists working in these cases, the tracking of those sources' respective impact on final estimates, and the relationship between the fossil-prior uncertainty and the addition of further contemporary sequence data as we move toward the infinite-sites limit. A view that looked solely toward abstraction would declare all of this work to be only of secondary importance, perhaps useful for a reconstruction of biological practice but not informative with regard to the status or nature of the probabilities themselves. I hope to have shown here that at least for this case, such a move both forecloses on much interesting reasoning to be done concerning the probabilities themselves, and does not align with the practice of the biologists concerned.

One might push the moral even more broadly. A general analysis of the notion of probability across evolutionary biology is doubtless a desirable philosophical result. But it seems as though it would be difficult to offer a single interpretation of probability that could encompass probabilities arising from sources as diverse as (i) ignorance of further contemporary sequence data, (ii) the mathematical constraint of nonidentifiability, and (iii) facts about the destructive nature of the process of fossilization in the evolutionary long term. The sophistication of biological modeling work here, and the corresponding breadth of philosophical interpretation available, seems to resist the imposition of a single interpretation of probability – and I suspect that this point will generalize to further areas of biological theory.

The second moral, then, brings us back to the main argument of the section. We can, I claim, also use this example to make trouble for the dichotomy between subjective and objective probability in the first place. As we saw above, these models of divergence time estimation have uncertainty which arises from at least three sources. The first is the lack of further data about contemporary sequences of extant organisms. This, it seems, is a very clear instance of subjective probability – sequence data serves as a source of uncertainty in our divergence time estimates only because we have failed to collect further sequence data.

The latter two sources of uncertainty, on the other hand, are more difficult to analyze. First, we have the uncertainty which arises from the (presumably fundamentally mathemat-

ical) fact that the model is nonidentifiable. This appears to arise simply as a matter of the mathematical formalism which must be used to express models of divergence time, and we will return to it below. While I lack the space to fully pursue the parallel here, it may perhaps, in turn, be comparable to recent discussions of purely mathematical explanation in both evolutionary and non-evolutionary sciences (Sober 2011; Baker 2017). Second, we have the uncertainty which arises from the fossil priors, which in more complex models are expressed as probabilistic ranges rather than single points. Portions of this uncertainty, as described above, are certainly due to our ignorance. We will never have a complete catalog of all of the available fossil data relevant to a particular divergence, for instance – and the same holds for cases of experimental error.

But several of the other sources of uncertainty here seem different. There's nothing about us as subjective knowers that affects the process of fossilization itself, and factors of this very process give rise to several of the sources of uncertainty at issue here. As Derek Turner has influentially argued, not only do we deal with the consequences of the products of fossilization being destroyed over time, “our background theories [about the way fossilization works, e.g.] give us reason to believe that they have been destroyed” (Turner 2005, p. 216). That is, the objective way in which the various processes of fossil deposition (e.g., the fairly unusual conditions required for the successful production of a fossil), geologic change (e.g., the destruction of fossils by subduction at plate boundaries), etc. actually operate entail that the incompleteness of the fossil record and the destruction of fossils are objective features of the natural world.

When uncertainty in our divergence time estimates can be directly traced to these kinds of sources, we have an instance where, I claim, we are not quantifying the uncertainty of a particular knower in a particular context (or that of any particular epistemic community). They are thus not, on the cartoon view, obviously subjective probabilities. At the same time, it is not clear that they are, in a traditional sense, objective probabilities – which are most commonly associated with theories like quantum mechanics in which probabilistic structure is an irreducible feature of the world.

It might be objected here that I have misinterpreted the notion of subjective probability.⁶ Surely, the objection goes, these probabilities are and could only be subjective. There is a concrete fact – the time at which a fossil was laid down – and we desire to know it. We could have known it, had we been present when the fossil was deposited, but we were not, and hence insofar as we need to reason about these matters probabilistically, we are doing so in the context of our ignorance of this fact. Quantifying ignorance, thus subjective probability.

I disagree. Probabilities are not free-floating features of the world (except perhaps in a metaphysically sensitive analysis of objective probabilities as they might arise in a domain like quantum mechanics). They arise, rather, within a particular theoretical context. We could, of course, have been present at the time of deposition of a fossil. Alternative humans with time-travel at their disposal would construct a very different theory of paleobiology, and the probabilities which appeared within it would need to be interpreted in a dramatically different manner. But as it sits, the probabilities that appear in divergence time estimates need to be interpreted as arising (at the very least) partly from features of the external world (like the processes of fossilization and the destructive forces that impinge on fossils), partly from our epistemic limitations (e.g., lack of time travel, which is not obviously a variety of “ignorance”), and partly from features of our theories about that external world. This last category might include features that pertain directly to our uncertainty concerning the data (e.g., lack of further sequence information), or features that are

⁶ A reviewer has urged the point.

more difficult to categorize (e.g., mathematical limitations that could, plausibly, be taken to apply to *any* theory that could be constructed in the domain at issue). In any event, if we wish to analyze probabilities within biology, we ought, as I hope to have done here, perform that analysis by analyzing, as directly as we can, the manner in which those theories do and must invoke those probabilities.

The correct analysis in our case, at least of the more complex sources of probability at work here, thus seems to land in a third category, between the classical notions of objective and subjective probability. There is, it seems, no reason to think that the processes of fossilization could not be represented in a deterministic way, or derived from a hypothesized deterministic fundamental physics – no quantum effects control whether or not a particular organism will become fossilized and preserved until the present day – but at the same time, it seems as though these probabilities simply are (and we have reason to believe would remain) features of our best theories of the relevant processes occurring in the world.

Comparison to probabilistic reasoning in other domains may be helpful to elucidate the sense of probability that I am claiming applies in this case. Consider the discussion of roulette wheels in Abrams (2012b). Even in a world where the underlying microphysics are entirely deterministic, there remains a sense in which the fact that there is a 50-percent chance that a spin of the roulette wheel turns up red is a true, objective fact about the world. There are a variety of ways in which to cash this out – Abrams’s own involves describing features of the causal mapping between “input” microstates and “output” states of the wheel as a whole, and Strevens (2011) offers a different but similarly inspired reconstruction. The overall intuition, however, remains constant: these probabilities are genuine features of the world *even though* they are in fact instances of underlying deterministic causal pathways.

Notably, this is far from the only place in which this class of probability has been noted in the life sciences. Millstein (2003) considers the interplay across evolutionary theory between what she terms “deterministic” and “indeterministic” notions of probability (both being “objective,” at least in the sense of existing independently of human knowers), and similar moves are made by Abrams (2012a) and Strevens (2016) in the cases of fitness and genetic drift, respectively.

But in many cases, we are still not given by these authors the kinds of tools needed to analyze the uncertainty in the fossil case. For example, Millstein concludes that the primary interpretation of probabilities in evolution is likely to be population-level propensities (2003, pp. 1326–1327), and Abrams and Strevens both push for their particular formulations of mechanistic probability, which produce objective, reliable probabilities in certain kinds of highly structured deterministic causal systems (Abrams 2012b). Both these general accounts seem inapplicable to our case here – Millstein’s applies only to evolving populations, and the work of Abrams and Strevens requires analysis of certain kinds of recurring features in the causal structures of the probabilistic system that seems to be impossible given the complex interplay involved in the process of fossilization. This is, of course, no indictment against any of the work accomplished by these authors – rather, their focus lies elsewhere, often on the probabilities that arise in fitness, natural selection, and genetic drift.

An example of the sort of more detailed meta-level framework that I claim would yield more fruitful analyses in the case of divergence time estimation models may be found in an entirely different area of the philosophy of science. In discussing the question of observational equivalence between deterministic and stochastic models in the philosophy of physics, Werndl (2013) has noted that we must be cautious in specifying the kinds of observations that we take to be relevant to a decision about a stochastic model. Translating her examples into our context, when we discuss the sources of uncertainty which cause that model to be probabilistic, we might be considering (1) uncertainty which arises with respect

to the *actually observed* data, (2) uncertainty with respect to “*currently possible observations*, i.e., currently possible observations given the available technology,” (3) with respect to “*observations which are possible in principle where there are limits, in principle, on observational accuracy*,” or (4) with respect to “*observations which are possible in principle where there are no limits, in principle, on observational accuracy*. (Here, although observation can never be made with infinite precision, observations which are possible in principle allow the researcher to come arbitrarily close to these infinitely precise values)” (Werndl 2013, p. 2253). As she notes, for researchers in practice, the type of uncertainty relevant for model-selection choice is (1), but if our goal is to analyze the uncertainty within these models, such uncertainty might well fall into any one of these four classes – while the standard “subjective/objective” distinction usually considers only one of these levels.⁷

Even Werndl’s analysis, though, doesn’t seem to quite fit the distinction between sequence data and fossil data being drawn in the divergence time case. Werndl is right to draw our attention to the limits on observations which give rise to the stochasticity present in our models, but it’s not clear that limits *on accuracy* are the ones relevant here.⁸ Her analysis does seem to draw our focus to the right features of the biological model: the differing ways in which uncertainty is generated by sequence data (resource limitations, experimental error, etc.) and fossil data (resources limitations as well, but also facts about the fossilization process, not least among them the epistemic inaccessibility of evolutionary deep time), and how that uncertainty is combined in the derivation of divergence times.

In that light, then, we could read these sources of uncertainty in at least three different ways, each with their own attendant in-practice and in-principle limitations, and each deserving of a separate variety of analysis. The lack of contemporary sequence data arises thanks to experimental error and contemporary resource limitations, but is a circumstance in which complete data (or, at least, given the speed of the approach to the infinite-sites limit, what *amounts to* complete data) is available – something like case (4) in Werndl’s taxonomy. Fossil uncertainty certainly arises, in part, from the same sources. But the fact that objective features of the process of fossilization play a role here seems to move us toward Werndl’s class (3), where while some further observations are possible, there are in-principle limits, here deriving from features of the processes of fossilization. Finally, we have the peculiar case of the nonidentifiability of rate and time, which, as noted above, only appear multiplied together in models for divergence time. What makes this an interesting aspect of divergence-time models is that, while it does not introduce any “real-world” uncertainty into divergence time estimates (i.e., it is not as though the fact that rate and time only appear multiplexed that prevents us from gathering more empirical data), it nevertheless remains responsible for setting an absolute limit on our ability to accurately estimate divergence times. This seems to give rise to a sort of probability that does not cleanly fit into Werndl’s taxonomy, necessitating the addition of (5*) uncertainty which enters into our measurements of a quantity as a feature of the models for estimating that quantity, and is not eliminable in principle.⁹

⁷ Which one will often depend on the kind of analysis and the theory being considered. For instance, analyses of quantum mechanics tend to focus on (4), as QM offers us strict limits on what can be observed in principle (Earman 2007). In evolutionary biology, the focus of philosophers of biology on practical considerations tends to shift the analysis to (2), though some work (such as that on the empirical methods of measuring natural selection found in Endler 1986) focuses on instances of (1).

⁸ To be clear, this is not a criticism of Werndl, for whom measure-theoretic models of particle location are the central example, and hence limits on accuracy are precisely the relevant focus of her work.

⁹ Symmetry would argue for the introduction of (6*) uncertainty which enters into our measurements of a quantity as a feature of the models for estimating that quantity, which would be eliminable in principle (i.e., would disappear given the use of more sophisticated models). No (6*)-type uncertainty seems to me to be present in the case study here, as it seems impossible to construct an identifiable model for divergence time.

While it is difficult to pursue the precise details of a taxonomy such as this on the basis of a single case study, the broader point remains, I hope, clear: examples such as these are well served by a significantly more detailed analysis of how such probabilities arise within particular, real-world biological models, transcending the basic subjective/objective distinction. Even better if, as in this case, these uncertainties have empirical upshot and can be explored quantitatively. I want to conclude by considering how these broader lessons might be applied in examples beyond the one I have described here.

5 A Way Forward

As I noted in the introduction, discussions of probability in evolutionary theory have frequently focused on extremely generalized cases – such as the broad interpretation of probabilities as they occur throughout fitness, natural selection, and genetic drift – and, as a result, required equally general philosophical analysis. I should make it clear that such analysis is important, and I have written on these topics myself (Pence and Ramsey 2013; Pence 2017). But the persistent intractability of these debates should give us pause, and cause to reevaluate our approach. The causalist-statisticalist debate, for instance, has continued for fifteen years (Matthen and Ariew 2002; Walsh et al 2002), and shows no signs of abating (Otsuka 2016; Walsh et al 2017).

I would suggest that one way, at least, to move forward on these debates is exemplified here. Many biological models are probabilistic, and for many, this is due to the need to quantitatively track uncertainties arising from the vicissitudes of data collection. This means that there is a (broadly untapped) collection of empirical work which is directly relevant to our understanding of the interpretation of these probabilities. Such an approach has several benefits. The presence of active modeling work often means, as in the example here, that we can manipulate model parameters with the goal of seeing the interaction between various sources of model uncertainty. The topic also garners significant interest from practicing biologists, making active dialogue with scientists more likely and more fruitful.¹⁰

As I have argued here, one impact of such an expansion of scope in our analysis of probabilities in evolution is likely to be the progressive realization that our analytical frameworks for understanding probability (often bequeathed to us via discussions of determinism and indeterminism in classical and quantum physics over the first part of the twentieth century) are inadequate to the task of clearly describing the work of biologists in these arenas. Philosophers of biology, given the robust use of probability in the variety of stochastic modeling contexts that biology provides us, may be uniquely positioned to advance research on the interpretation of probability in this way.¹¹

And while I do not claim to have definitively argued for the superiority of a taxonomy like Werndl's (as expanded) from only one example, such a taxonomy does have a number of very significant advantages for the analysis of biological models. Focusing our work on the sources of uncertainty and the ways in which these translate into stochastic or probabilistic models has both allowed for maximal engagement with the work of biologists, as well as drawn our attention not to the abstract causal structure of the model, but rather to the precise relationship between the various pieces of data utilized by the model and the conclusions drawn from them – the relationship between statistical theory and world. It is my hope

¹⁰ And not just biologists, either – see, for example, the robust literature on the tracking and management of uncertainty in climate science (e.g., Foley 2010).

¹¹ The already much-mentioned work of authors like Strevens and Abrams on the role of mechanistic probability provides one excellent example of this (Strevens 2011, 2013, 2016; Abrams 2012a,b, 2015).

that redirecting our work in this way could offer us true progress on our understanding of probabilities in evolutionary biology.

Acknowledgements My sincere thanks to two anonymous reviewers for this journal, who very dramatically improved this paper (and caught a few serious errors!). For comments on a very early version of this project, thanks to an audience at the Models and Simulations 6 conference, at the University of Notre Dame. Many thanks also to Mario dos Reis for the initial inspiration behind the project, which was born at NESCent – still inspiring interdisciplinary work years after its unfortunate closure.

References

- Abrams M (2009) What determines biological fitness? The problem of the reference environment. *Synthese* 166(1):21–40
- Abrams M (2012a) Measured, modeled, and causal conceptions of fitness. *Frontiers in Genetics* 3:196
- Abrams M (2012b) Mechanistic probability. *Synthese* 187(2):343–375
- Abrams M (2015) Probability and manipulation: Evolution and simulation in applied population genetics. *Erkenntnis* 80(S3):519–549
- Ariew A, Lewontin RC (2004) The confusions of fitness. *British Journal for the Philosophy of Science* 55(2):347–363
- Baker A (2017) Mathematical spandrels. *Australasian Journal of Philosophy* pp 1–15
- Beatty JH, Desjardins EC (2009) Natural selection and history. *Biology & Philosophy* 24(2):231–246
- Benton MJ, Donoghue PCJ, Asher RJ (2009) Calibrating and constraining molecular clocks. In: Hedges SB, Kumar S (eds) *The Timetree of Life*, Oxford University Press, Oxford, pp 35–86
- Brandon RN (1990) *Adaptation and Environment*. Princeton University Press, Princeton, NJ
- Brandon RN, Ramsey G (2007) What's wrong with the emergentist statistical interpretation of natural selection and random drift? In: Hull DL, Ruse M (eds) *The Cambridge Companion to the Philosophy of Biology*, Cambridge University Press, Cambridge, pp 66–84
- Desjardins E (2011) Reflections on path dependence and irreversibility: Lessons from evolutionary biology. *Philosophy of Science* 78(5):724–738
- Desjardins E (2016) Contingent evolution: Not by chance alone. In: Ramsey G, Pence CH (eds) *Chance in Evolution*, University of Chicago Press, Chicago, pp 223–243
- dos Reis M, Yang Z (2013) The unbearable uncertainty of Bayesian divergence time estimation. *Journal of Systematics and Evolution* 51(1):30–43
- Drouet I, Merlin F (2013) The propensity interpretation of fitness and the propensity interpretation of probability
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4(5):e88
- Earman J (2007) Aspects of determinism in modern physics. In: Butterfield J, Earman J (eds) *Handbook of the Philosophy of Science: Philosophy of Physics*, North-Holland, Amsterdam, pp 1369–1434
- Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ
- Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates, Sunderland, MA
- Foley A (2010) Uncertainty in regional climate modelling: A review. *Progress in Physical Geography* 34(5):647–670
- Gillespie JH (1984) Molecular evolution over the mutational landscape. *Evolution* 38(5):1116–1129
- Gillespie JH (1991) *The Causes of Molecular Evolution*. Oxford University Press, Oxford
- Graur D, Martin W (2004) Reading the entrails of chickens: Molecular timescales of evolution and the illusion of precision. *Trends in Genetics* 20(2):80–86
- Graves L, Horan BL, Rosenberg A (1999) Is indeterminism the source of the statistical character of evolutionary theory? *Philosophy of Science* 66(1):140–157
- Halpern AL, Bruno WJ (1998) Evolutionary distances for protein-coding sequences: Modeling site-specific residue frequencies. *Molecular Biology and Evolution* 15(7):910–917
- Hasegawa M, Kishino H, Yano Ta (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22(2):160–174
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian Protein Metabolism*, vol 3, Academic Press, New York, pp 21–132
- Lenormand T, Roze D, Rousset F (2009) Stochasticity in evolution. *Trends in Ecology & Evolution* 24(3):157–165
- Lepage T, Bryant D, Philippe H, Lartillot N (2007) A general comparison of relaxed molecular clock models. *Molecular Biology and Evolution* 24(12):2669–2680
- Linder M, Britton T, Sennblad B (2011) Evaluation of Bayesian models of substitution rate evolution—parental guidance versus mutual independence. *Systematic Biology* 60(3):329–342
- Matthen M (2009) Drift and “statistically abstractive explanation”. *Philosophy of Science* 76(4):464–487
- Matthen M, Ariew A (2002) Two ways of thinking about fitness and natural selection. *Journal of Philosophy* 99(2):55–83
- Merlin F (2010) Evolutionary chance mutation: A defense of the Modern Synthesis' consensus view. *Philosophy and Theory in Biology* 2:e103
- Merlin F (2016) Weak randomness at the origin of biological variation: The case of genetic mutations. In: Ramsey G, Pence CH (eds) *Chance in Evolution*, University of Chicago Press, Chicago, pp 176–195

- Millstein RL (2003) Interpretations of probability in evolutionary theory. *Philosophy of Science* 70:1317–1328
- Millstein RL (2006) Natural selection as a population-level causal process. *British Journal for the Philosophy of Science* 57(4):627–653
- Millstein RL (2008) Distinguishing drift and selection empirically: “The Great Snail Debate” of the 1950s. *Journal of the History of Biology* 41(2):339–367
- Millstein RL (2011) Chances and causes in evolutionary biology: How many chances become one chance. In: Illari PM, Russo F, Williamson J (eds) *Causality in the Sciences*, Oxford University Press, Oxford, pp 425–444
- Millstein RL (2016) Probability in biology: The case of fitness. In: Hájek A, Hitchcock C (eds) *The Oxford Handbook of Probability and Philosophy*, Oxford University Press, Oxford, pp 601–622
- Mulcahy DG, Noonan BP, Moss T, Townsend TM, Reeder TW, Sites JW, Wiens JJ (2012) Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. *Molecular Phylogenetics and Evolution* 65(3):974–991
- Nascimeto FF, dos Reis M, Yang Z (2017) A biologist’s guide to Bayesian phylogenetic analysis. *Nature Ecology & Evolution* 1(10):1446–1454
- Otsuka J (2016) A critical review of the statisticalist debate. *Biology & Philosophy* 31(4):459–482
- Pence CH (2017) Is genetic drift a force? *Synthese* 194(6):1967–1988
- Pence CH, Ramsey G (2013) A new foundation for the propensity interpretation of fitness. *British Journal for the Philosophy of Science* 64(4):851–881
- Rannala B, Yang Z (2007) Inferring speciation times under an episodic molecular clock. *Systematic Biology* 56(3):453–466
- Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* 61(6):973–999
- Rosenberg A (1994) *Instrumental Biology, or the Disunity of Science*. University of Chicago Press, Chicago
- Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14(12):1218–1231
- Schwartz JH, Maresca B (2006) Do molecular clocks run at all? A critique of molecular systematics. *Biological Theory* 1(4):357–371
- Sober E (2011) A priori causal models of natural selection. *Australasian Journal of Philosophy* 89(4):571–589
- Stamos DN (2001) Quantum indeterminism and evolutionary biology. *Philosophy of Science* 68(2):164–184
- Strevens M (2011) Probability out of determinism. In: Beisbart C, Hartmann S (eds) *Probabilities in Physics*, Oxford University Press, Oxford, pp 339–364
- Strevens M (2013) *Tychomancy: Inferring Probability from Causal Structure*. Harvard University Press, Cambridge, MA
- Strevens M (2016) The reference class problem in evolutionary biology: Distinguishing selection from drift. In: Ramsey G, Pence CH (eds) *Chance in Evolution*, University of Chicago Press, Chicago, pp 145–175
- Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51(5):689–702
- Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15(12):1647–1657
- Turner D (2005) Local underdetermination in historical science. *Philosophy of Science* 72(1):209–230
- Walsh DM (2007) The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science* 74(3):281–303
- Walsh DM, Lewens T, Ariew A (2002) The trials of life: Natural selection and random drift. *Philosophy of Science* 69(3):429–446
- Walsh DM, Ariew A, Matthen M (2017) Four pillars of statisticalism. *Philosophy, Theory, and Practice in Biology* 9
- Weisberg M (2014) Understanding the emergence of population behavior in individual-based models. *Philosophy of Science* 81(5):785–797
- Wernli C (2013) On choosing between deterministic and indeterministic models: Underdetermination and indirect evidence. *Synthese* 190(12):2243–2265
- Wilkinson RD, Steiper ME, Soligo C, Martin RD, Yang Z, Tavaré S (2011) Dating primate divergences through an integrated analysis of palaeontological and molecular data. *Systematic Biology* 60(1):16–31
- Yang Z, Rannala B (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* 23(1):212–226
- Zuckermandl E, Pauling L (1965) Molecules as documents of evolutionary history. *Journal of Theoretical Biology* 8(2):357–366