

Phylogeny as population history

Joel D. Velasco^{§‡}

The construction and use of phylogenetic trees is central to modern systematics. But it is unclear exactly what phylogenies and phylogenetic trees represent. They are sometimes said to represent genealogical relationships between taxa, between species, or simply between “groups of organisms.” But these are incompatible representational claims. This paper focuses on how trees are used to make inferences and then argues that this focus requires that phylogenies represent the histories of populations.

KEYWORDS

Phylogeny • Population • Species • Systematics • Taxonomy

1. Introduction

The project of this paper is to understand what a phylogenetic tree represents and to discuss some of the implications that this has for the practice of systematics. At least the first part of this task, if not both parts, might appear trivial—or perhaps better suited for a single page in a textbook rather than a scholarly research paper. But this would be a mistake. While the task of interpreting phylogenetic trees is often treated in a trivial way, their interpretation is tied to foundational conceptual questions at the heart of systematics—questions whose answers are hotly disputed. I have previously argued that widely shared ideas about the meaning and interpretation of phylogenetic trees are inconsistent with species concepts other than some genealogical version of a phylogenetic species concept (Velasco 2008). Here I rely on a similar approach and concentrate on the implications of the necessary conditions underlying the inferences that we make using phylogenetic trees. I argue that common practices for the interpretation and use of trees are in conflict and that unacceptable principles about species as units of phylogeny must be given up. According to the view that I will develop, all phylogenetic trees depict the history of populations.¹ The branches on trees represent collections of population lineages through time and the splits represent population lineage splits. This is true regardless of whether the tips of the trees are themselves populations, or are species or higher taxa. Although this conclusion might be paired naturally with a view that species must be monophyletic groups, this population-centric view of trees is independent of that view of species. If we still want to have species that are paraphyletic groups of populations, this is permissible as long as we also do not treat species as the units of phylogeny. This population-centric view opposes a species-centric view of phylogeny and might be called a “rank-free” approach since it entails that we do not need to determine which groups are species (which is partly a ranking question) in order to build a tree. This conclusion and the argument for it are

[§] Division of Humanities and Social Sciences, MC 101-40, Calif. Institute of Technology Pasadena, CA 91125

[‡] Department of Philosophy, Box 43092, Texas Tech University, Lubbock, TX 79409

✉ E-mail: joel@joelvelasco.net

Received 2 October 2012; Revised 1 December 2012; Accepted 4 January 2013



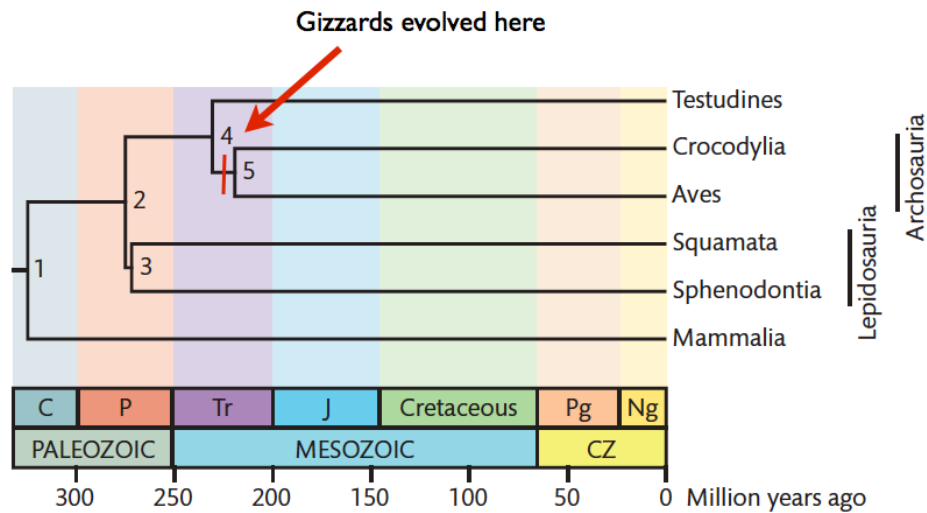


Figure 1 — Time tree of the amniotes (from Shedlock and Edwards 2009). The divergence times (with sources and associated confidence intervals) are found in their Table 1 (not shown here). Abbreviations: C (Carboniferous), CZ (Cenozoic), J (Jurassic), Ng (Neogene), P (Permian), Pg (Paleogene), and Tr (Triassic). The added bar points out where we infer the gizzards to have evolved.

meant to be consistent with, but not require, acceptance of the conclusions of Velasco (2008) regarding species.

2. Reading and Using Phylogenetic Trees

Phylogenetic trees depict genealogical relationships between various groups of organisms. The interpretation and use of trees for inferences is central to systematics and many introductions to the field (or even to biology generally) start with this topic.² Important for us will be how trees are used to make inferences about the past. A straightforward example of this is inferring trait evolution. Figure 1 depicts one reasonable phylogeny of the major amniote clades. Each tip is a monophyletic group any of whose members can stand in for the genealogical relationship of any other member. For example, every bird (in Aves) has the same relationship to any group in any other clade including any crocodile (in Crocodylia), any turtle (in Testudines) and any lizard or snake (in Squamata). The nodes represent real divergence points: birds and crocodiles share a common ancestor about 219 million years ago (mya), birds and turtles about 231 mya, and birds and lizards about 275 mya. More traditional phylogenies put turtles elsewhere, such as sister to the rest of Sauropsida. In that case, the dating of the bird-turtle split is closer to 300 mya (Carroll 1988). The branches represent real lineages that lead to organisms at that tip and none of the other tips. This is essential for tracing the effect of evolutionary changes along the branches. When we study the effects of trait evolution, we must know (or at least have an estimate of) the underlying phylogeny.

For example, birds and crocodiles but no other living reptiles have a gizzard. The gizzard is a kind of specialized stomach, attached to the “true” stomach, which allows these organisms to swallow stones and then use them to aid in the breakdown of food that is passed back and forth between the glandular stomach and the gizzard. How and when did the gizzard evolve? Given the true phylogeny, we can infer that it probably evolved on the indicated branch that leads to both the birds and the crocodiles, but leads to nothing else (Figure 1). That would mean the gizzard evolved between 219 and 231 mya and has been inherited by birds and crocodiles, both of which are descendants of this branch. In other words, the gizzards in birds and crocodiles are homologous. Incidentally, this also would mean that the dinosaurs had gizzards as they are on the branch leading to the birds. This implication could be erroneous; for example, it is possible that the gizzard evolved earlier and it was lost in the branch leading to the turtles. Alternatively, it is possible

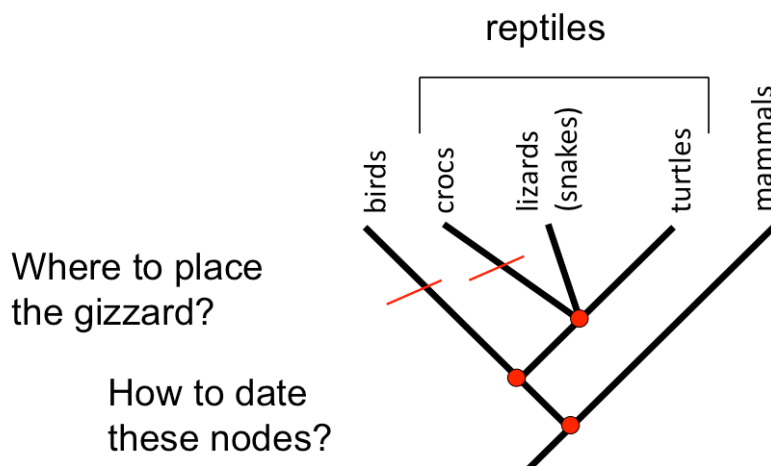


Figure 2 — If we use an incorrect phylogeny (for example, one based on a particular classification) and then attempt to infer facts about evolutionary history, such as the origin of traits, we will make incorrect inferences. In this figure, we have treated the traditional reptiles as a single clade and then expanded the group into some of its taxonomic orders. Relationships on the tree represent classificatory relationships. Using this tree, it would appear that gizzards evolved twice independently, which is incorrect, and we also would infer incorrectly the dates of various nodes.

that the gizzard has evolved twice. Some other organisms have gizzard-like organs such as certain fish and even earthworms. Again, with the proper phylogenetic tree, it is easy to see that either massive numbers of groups have lost the gizzard, or it has independently evolved in the fish, earthworms, and archosaurs. A detailed examination of the “gizzards” in the earthworms and fish makes it clear that many of the similarities are only superficial and these are not really the same character trait being expressed in the different groups, unlike in the crocodiles and the birds where the gizzard truly is the same character.

Thinking about how the tree constrains the possible evolutionary scenarios leads to the conclusion that the phylogeny has to represent the true history of the lineages in question and not simply represent our classification scheme or a similarity metric of some kind. For example, on the traditional picture in which we classify birds in class Aves and reptiles (like crocodiles) in class Reptilia, it is very tempting to place reptiles at the tip of a tree as a sister group to the birds. Depending on our views about classification, it might be permissible to treat Reptilia as a taxon for classification purposes, but it is not permissible to use it as the tip of a phylogenetic tree. We can see why this is problematic when we try to expand the reptiles into their traditional orders as though classification dictated phylogeny (Figure 2). If the gizzard did evolve only once and spread to birds and crocodiles, the proper explanation is blocked since there is no branch on this diagram where the gizzard could have evolved in this way. Instead it looks like a massive coincidence that birds and crocodiles share such detailed similarities like the gizzard and all of its natural components. This is especially odd given the obvious ecological differences between birds and crocodiles, which would lead to the apparent conclusion that they have been under very different selection pressures for quite some time. Issues also will arise when we start thinking about extinct groups such as dinosaurs, which traditionally belong with reptiles, but in terms of traits belong on the branch leading to the birds (witness the recent discoveries of numerous feathered dinosaurs).

The fact that this tree cannot correspond to genealogy is seen also by trying to date the nodes. Exactly when did birds split from reptiles? On this tree, it appears that you could use crocodiles, lizards, or turtles equally as your comparison group since they all appear equally related to birds. Or, perhaps to maximize the strength of your evidence, you should look at all the groups and do a sort of averaging of the evidence. But this would be a mistake since birds are related differently to different groups of reptiles. Averaging the numbers would be meaningless.

We arrived at Figure 2 by subdividing the reptiles into smaller groups and placing those groups in the location where the reptiles were (stemming from the same node). This is called *expansion*. Similarly, we can always collapse sister groups together into a larger clade (merging or *contraction*). We also should be able to add and subtract taxa without changing the topological relationships of the other members of the tree (*addition* and *subtraction*, or pruning). These manipulations will not affect the implied genealogical relationships of the groups and so should not affect our inferences about history. For example, if we altered the tree in Figure 1 by expanding the birds into several subgroups, all of which have gizzards, collapsed the lepidosaurs into a single group stemming from node 3, and removed mammals entirely, our inference regarding the origins of the gizzard would not change. For ease of communication, I will call these *tree manipulation principles*. These are usually assumed to hold and are used without comment, but sometimes they are explicitly described as essential to understanding how trees work (e.g., Baum and Smith 2012) or for understanding how evolutionary explanations work more generally (O’Hara 1997). But tree manipulation principles combined with the placement of arbitrary groups as taxa on trees lead to absurdity.

The misrepresentation that results from phylogenies that do not represent history has dire consequences when we try to use trees to make inferences. Genealogical trees are essential for testing selection hypotheses, testing models of rates of evolutionary change (such as the molecular clock hypothesis), doing comparative biology (are traits X and Y correlated?), and for testing biogeographical hypotheses (did marsupials originate in Australia or migrate from elsewhere?). Any inference that requires knowing the history of any groups will require an estimate of the phylogenetic tree. Knowing a traditional classification of this group or knowing which groups are similar to other groups helps a lot for phylogenetic purposes, but only insofar as it is a good guide to the true genealogy.

We now have an argument that no taxon—a group placed at the tip of a tree—could be a paraphyletic group. For the same reasons, if we are going to ask questions about groups above the species level, such as “when did this group originate?,” “when did this group diverge from this other group?,” or “has group X undergone anagenetic change faster than group Y ?,” the groups in question must be clades or these questions will make no sense. If we add some background assumptions, such as we can always treat taxa (in classification) as taxa (tips of trees in phylogenetics), and we are consistent in how we use trees, then this is an argument for a phylogenetic system of classification as well (Velasco 2008). But even if we don’t accept this last step, at least we must all agree that our phylogenies should not depend on our classification.

While the statement that the correct phylogenetic tree does not depend on how we classify would probably find near universal agreement, our current practices dictate that phylogenies are dependent on classification. This is because on most views of phylogeny, at least according to the definitions typically given (though not the practice as we will see), what constitutes a clade is determined by which groups are species and what counts as speciation. On most views (both by definitions and by practice) this aspect of systematics is determined partly by our classification of what counts as a species and not directly by history.

I have so far only presented claims about what a tree represents and how to read and use phylogenetic trees. What I have said is relatively uncontroversial. However, as I will argue, tree manipulation principles like expansion and contraction or addition and subtraction, which are essential underpinnings for the inferential role that trees play, conflict with other principles of trees—those involving species. Perhaps after further consideration, we will decide to give up certain manipulation principles. But doing so would invite a host of problems that we should avoid. As we have just seen in our example of the gizzard, manipulation principles are tied to the ability to make historical inferences using trees. For this reason, I will proceed as though we should hold on to these principles and see where they lead us.

3. Population Trees

We now have an argument that all groups placed at the tips of trees must be monophyletic clades. A tree then displays the relationships between these tips. For example, according to Figure 1, there is a clade that contains crocodiles and birds but not snakes. But what is a clade? A standard definition is some ancestor and all of its descendants. Traditionally, this means an ancestral species and all the species descended from it. Alternatively, we could say directly that the genealogical claim cited above means that the most recent common ancestor of birds and crocodiles is a descendant of the most recent common ancestor of crocodiles

and snakes. But what has ancestor-descendant relationships in this way? Again, the traditional answer is that this is all done with species. Phylogenetic trees are ultimately about the relationships between species. If we put birds at the tip, it merely represents a collection of all the bird species.

That phylogenetic trees are really species trees in this sense is a common sentiment: “A phylogenetic tree is a representation of the historical course of speciation. ... It is the business of phylogenetic systematists to attempt to recover this history of speciation” (Wiley 1981, 2). That the great Tree of Life is a phylogenetic tree and a tree of species is found throughout the literature (Cracraft and Donoghue 2004; Dawkins 2004; Hodkinson and Parnell 2006). Despite the rhetoric, in practice these authors, along with everyone else, do not treat phylogenetic trees as species trees in this sense. As I will argue, thinking of trees in this way conflicts with the tree manipulation principles and would serve as a faulty foundation for history-based phylogenetic inferences.

If all phylogenetic trees are ultimately species trees, then species are like atoms with no internal parts that have phylogenies. But if our practices are any guide, this is not correct. Though we do primarily talk about *species* trees, we also talk about and use *population* trees. “Species tree” refers to a tree that has species at the tips regardless of how it is interpreted, whereas “population tree” refers to a tree that has at least some subspecific populations as tips.

Phylogenetic trees containing tips that are smaller than species are commonplace in the literature. Sometimes they depict the relationships between named groups, such as subspecies or varieties; other times the tips are different populations of the same species, usually delimited by geography. These trees are used for precisely the same purposes as species trees: representing the history of groups of organisms in a way useful for historical inferences. In his book *Evolutionary Pathways in Nature* (Avice 2006), John Avice explains the phylogenetic approach to inferring evolutionary history and summarizes a large number of examples where trees have played an essential role. Earlier, we saw an example of phylogenetic character mapping above the species level when we examined the gizzard (Figure 1). While the majority of Avice’s examples are at or above the species level, several of them involve subspecific groups. For example, Avice examines the evolution of the bold black and red plumage in poisonous birds—the Hooded and the Variable Pitohui (*Pitohui dichrous* and *P. kirhocephalus*)—using data from Dumbacher and Fleisher (2001) (Avice 2006, 65–68). By finding clades of subspecific populations, we can see that the phylogenetic evidence indicates that the distinct plumage pattern was probably the shared ancestral state of the Pitohuis and not due to convergent evolution. However, in one population on the Wandammen Peninsula of New Guinea’s north coast, this distinct plumage pattern was found embedded in a clade that otherwise lacked the pattern, making it quite likely that in this population the plumage pattern was secondarily gained as an instance of mimicry. It was essential in this case not only to find clades of populations within a single species, but to plot their relationships to other species as well. Thus the phylogeny depicted is not a pure population tree or pure species tree.

Similar studies have been carried out closer to our taxonomic home. Morin et al. (1994), Kaessmann et al. (1999), Won and Hey (2003), and many others have investigated the divergence of chimpanzee species and subspecies and a huge number of studies have been done on the population history of humans (Cavalli-Sforza et al. 1994; Templeton 2007). Research in both groups is usefully summarized in Arnold (2009). Alan Templeton has famously argued against various population trees of humans. But he does not argue that these population trees are meaningless; rather, they are meaningful and have genealogical implications—implications that have been repeatedly tested and found wanting (Templeton 2002, 2005, 2007). Entire subdisciplines such as phylogeography, divergence population genetics, and coalescent methods depend on the study of population histories. To deny that phylogenetic trees and methods are profitably applied below the species level is to claim that each of these subdisciplines is fundamentally flawed in its foundations.

4. The Problem with Population Trees

Although biologists create and use population trees, there are reasons to worry about their interpretation. One reason is that systematists often define phylogenies in terms of species. If we do this, then populations do not have phylogenetic relationships. But we don’t want to simply rely on traditional definitions (which are not all consistent with each other); we want to examine the reasons behind them. And

the primary reason to think that there is something fundamentally wrong with population trees is that populations are not completely isolated from one another—they reticulate.

In many cases, this claim is false. There are a variety of recognized conspecific populations with no gene flow between them. In sexual species, many populations are isolated from others, often for geographical reasons such as a river or mountain range forming external barriers to reproduction. Many species concepts, most clearly internodal concepts (Ridley 1989; Kornet 1993), might raise such groups to the rank of species. If they have been isolated long enough, we might expect that the groups would have their own apomorphies in which case some phylogenetic species concepts would recognize multiple species. But this lack of gene flow is consistent with conspecificity on other species concepts. For example, the biological species concept demands intrinsic reproductive isolation, which these populations might lack and these populations may not have diverged ecologically or morphologically in significant ways and thus are conspecific on various ecological and morphological concepts as well. For many species concepts, population trees can be meaningful, even on the unnecessarily strict interpretation of requiring no gene flow between distinct lineages.

While this suggests that some population trees can be meaningful, to use them as the foundation for phylogenetics one would want a much more robust defense. Phylogenetic differentiation should be meaningful in the face of reticulation. But the natural interpretation of branching on a tree is that there is no more gene flow between the separate branches. Reticulations are more naturally represented by a network with lateral connections as well as vertical branches, rather than a strict tree. While this clarity has advantages, it comes at a high price. Not only do we have to reinterpret a massive amount of good phylogenetic work being done below the species level, but above the traditional species level as well. It is very common for well-defined and accepted species to not be completely isolated.

Mallet (2005) surveys a variety of studies on hybrids and concludes that at least 25% of plant and 10% of animal species form hybrids with other species in nature. This usually leads to introgression and therefore gene flow between species. And this is surely an underestimate as it is based on our own enormously inadequate sampling conducted by examining organisms in the wild. Even in species with relatively contained geographical habitats, we cannot examine every organism. And individual hybrid organisms are not always detectable by way of morphological differences.

Worries about gene flow between different species depend heavily on what groups we classify as species. Different species concepts lead to different species. A major reason for the development of the ecological species concept was that important ecological differences can persist in spite of gene flow, such as in the famous case of oak trees in the genus *Quercus* (Van Valen 1976). The same example shows that morphological differences can persist in spite of gene flow as well. In recent years there has been a trend toward taxonomic inflation as more and more previously recognized species are split into distinct groups—often for conservation purposes (Mallet 2005). For example, according to his version of the phylogenetic species concept, Joel Cracraft estimates that the number of bird species could rise from approximately 9,000 to approximately 18,000 (Cracraft 1997). Many of these distinct subspecies are morphologically distinguishable from others and all are genetically distinguishable. Yet there is frequently substantial gene flow between them, which is why they were not recognized as separate species initially. This naturally leads to increased “hybridization” between groups.

Even when not raised to the rank of species, a number of subspecific groups in nature have been recognized as special subspecies or evolutionarily significant units (Ryder 1986; Moritz 1994; Hey et al. 2003). If we wish to inquire about the history of these groups and ask questions about their origins and relations to other groups, phylogenies are required. So if we deny that population trees make sense in the face of reticulation, we also must deny that many species trees are meaningful. Perhaps for some species there would be no correct phylogenetic tree with them as tips. Not only would this unnecessarily hamper the study of the history of these groups, it would destroy the supposedly solid foundation of phylogenetic trees at higher taxonomic levels. When we reflect on the way that phylogenetic trees are used, clearly species trees can be meaningful and useful even in the face of reticulation. Thus we should accept that population trees are meaningful as well.

5. What are Population Trees?

A problem that the defender of population trees must face is how to interpret branching in a way consistent with reticulation. A helpful start is to think of phylogenetic trees as models that contain idealizations (Velasco 2012). This means that we are not forced to say exactly when and where splits take place, how much reticulation is sufficient for the continued integrity of a lineage, or which group of organisms constitutes a clade. But we still need to say what historical relationships we are trying to model.

One way to do this would be parallel to defining species trees as the history of speciation—we define populations and then use the definition to determine when one population has split into two; population trees model this process. Defining “population” and “population split” is not easy, though some have tried (Millstein 2010). But focusing on the definition of species and the necessity of determining the precise timing of speciation is what gets species trees in trouble. Replacing “species” with “population” is unlikely to help. Rather, we should focus on the pattern of divergence and build our trees from that. We could use this understanding of phylogeny to define populations in terms of gene flow or divergence just as some want to define species that way. I remain agnostic on the issue and instead focus on what a phylogenetic tree could mean regardless of what we call the groups of organisms at the tips.

Our goal is to understand the meaning of phylogeny without resorting to defining it with species. One approach is to define it with population histories. If populations are defined historically, then this might work. But a better approach is to think directly about what the pattern of divergence between groups of organisms could represent in a phylogeny. A useful way to proceed with this problem is to think about what kinds of entities can be on the tips of phylogenetic trees.

We have already said that monophyletic clades are allowed to be placed at the tips of trees. But we have not answered the central question of what a clade is. When dealing directly with groups of organisms, it seems that we would need them to be an exclusive group in the sense that any organism in the population is more closely related to any other than to any outside the population. If we define the genealogical relatedness of organisms using recency of common ancestry (Velasco 2009), we get a very stringent criterion whereby populations would almost never count as exclusive: a single migrant from another population who has offspring would collapse the two separate lineages. This is too extreme. Alternatively, we can define exclusivity in terms of gene trees. The most sensible view is that an exclusive group is a group where the proportion of the genome that forms a clade is higher than any conflicting potential clade (Ané et al. 2007; Baum 2007). This view was designed as a way to talk generally about clades of all sizes. We could use concordance factors to determine subspecies (Avice and Ball 1990) or species and phylogeny more generally (Baum 2009; Velasco 2010). Regardless, phylogeny is not determined by our classification but rather by genealogical history.

Concordance factors are just one way that we might determine phylogeny. Many methods exist for determining species trees or population trees that look at the pattern of divergence and do not define phylogenies in terms of the process of species or population splitting. For example, using coalescent methods we can infer the likely patterns of lineage branching, taking advantage of the depth of gene coalescence and not just the topology (Knowles and Maddison 2002; Edwards 2009). We also can use divergence population genetics to infer levels of past gene flow and depict that pattern with our trees (Won and Hey 2003). What is important in all these cases is that we are inferring the history of groups of organisms *independently* from how we wish to classify them. Perhaps we are examining conspecific populations, or entirely separate species, or even a mix of the two. It is irrelevant. The methods are the same because the phylogenies mean the same thing in any case. Population level phylogenies can be meaningful because populations have genealogical histories that are useful to represent.

6. When Species Trees and Population Trees Disagree

Even if we accept that population trees can be meaningful, this is a long way from showing that they are the foundation for all phylogenetic trees. The key argument for this conclusion is that the truth of species phylogenies (i.e., which species are genealogically more closely related to others) are determined by population level divergences and not by recency of speciation events. I will argue for this by examining an

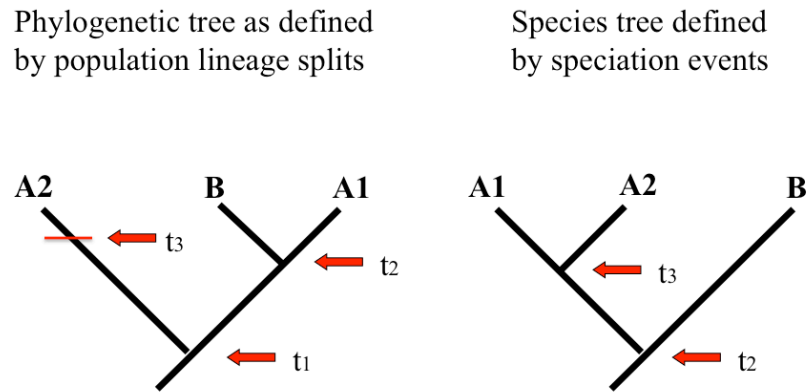


Figure 3 — At t_1 , two populations of the same species split. This is not a speciation event. Then at t_2 , a new species B splits off from A1. After t_3 , A1 and A2 are now considered different species. If we want to use trees for historical inferences, such as reconstructing the history of traits, we will need to use the tree defined by population splits and not speciation events.

instance of the set of common cases where two different ways of understanding the tree lead to different conclusions: the “The Wrong Tree” problem (Velasco 2008). In such cases, phylogenetics should be concerned with population trees.

It is a clear mistake to place paraphyletic groups (e.g., the reptiles or the fishes) at the tips of phylogenetic trees. Such groups do not have a unique history that is not shared by other groups. But many species are paraphyletic groups of populations; some populations in these species are more closely related to other species than to other populations in their own species. By reasoning that is completely parallel with the case of higher taxa, such species cannot serve as phylogenetic units at the tips of trees.

That such groups of paraphyletic species are common is well known. It is trivial for such groups to form (Figure 3). Start with two phylogenetically distinct populations of the same species. Call them A1 and A2. Now imagine that a new population splits off from A1 and the resulting lineage becomes a new species. Call this species B. Now species A ($=A1+A2$) is a paraphyletic group consisting of some but not all the populations descended from the ancestral A population. A phylogenetic tree with A at a tip would be misleading since some of the As (namely, the A1s) would be more closely related to species B than to other members of their own species (the A2s). To make the inconsistency more obvious, just fast-forward in time; A1 and A2 have continued to diverge and are now separate species. Perhaps there has been another lineage split so no anagenesis is required. If the species tree tracks speciation history, A1 will be more closely related to A2 and the relationships will be (B, (A1, A2)). But this tree is not phylogenetically meaningful since tracking the history of populations, organisms, or genes will give us the result that A1 is closer to B than to A2 (Figure 3). It is this tree where the lineage splits are defined by the population level splits and not the speciation events that is meaningful for historical inferences. When they became distinct species according to some particular species concept is irrelevant.³

That such examples of apparent inconsistency are common is not in dispute. What is in dispute is what to do about it. Given the fact that biologists use population trees and define phylogenetic trees as species trees, the answer would appear to be species trees and population trees just represent different things and so they cannot be genuinely inconsistent. Though this would appear to be the obvious answer, I know of no one in the literature who explicitly endorses it. While this is a possible resolution to the problem, I argue that the costs associated with the resolution are too high.

7. Privileging Species Trees

In this section, I will argue against the view that phylogenies are fundamentally species trees as defined by the history of speciation. One consequence of this species-centric view is that population trees are not phylogenies or at least not phylogenies of the same kind. They track the history of populations and lineage

splits while species trees track speciations. But this strict separation between species trees and population trees has numerous odd consequences, some more serious than others. For example, recall that the inference about the history of the plumage in the Variable Pitohuis required examining the relationships of some of the populations not just in relation to other conspecific populations but to other species as well. This particular path of investigation would be unavailable since the trees used in the study are apparently meaningless; it seems incoherent to use a tree with a mix of species and non-species at the tips. What could count as a split between a population and a species? It can't be a speciation event. We could perhaps try to give them meaning by treating them as population trees and treating the species in question (and any other higher taxon as well) as placed on the tree due to the population-level histories of the populations inside it. If the taxon in question happens to be a monophyletic group of populations, this might work out. But on most views of species, there is no problem with a species being a paraphyletic group of populations. In such cases, it is entirely unclear where a single species *qua* group of populations should go on a population tree when the populations inside it belong in different places.

While there are stipulations we could make about how to deal with such cases, they all lead to problems when we try to use the tree for inferences about the past. Recall that paraphyly at the higher level leads to mistaken inferences about evolutionary history. When we insisted on treating reptiles as a single group apart from birds, we were blocked from recognizing the true history of the gizzard—that it is shared between birds and some, but not all, reptiles. Similarly, when we collapse paraphyletic species into a single tip of a tree, we are blocked from recognizing the true history of traits that are shared between one species and some, but not all, populations of the paraphyletic species. Attempts to date nonexistent or misleading nodes would lead to incorrect pictures of the history and lead us astray in inferring biogeography and morphological or molecular histories.

Even when all the populations in a taxon belong in the same place on a tree, the species-centric view can cause problems. The relationship of species to each other due to speciation events can come apart from the relationships of the populations inside them (Velasco 2008). The only reasonable way to interpret a species-centric view is that when the tree in question has only species at its tips, the default is that it is a species tree defined by species-level history while any tree with subspecific populations at the tips, even if some of the tips are species, is defined by population-level history. But then how species are related to each other as part of a species tree could be different from how the very same groups are related to each other on a population tree. This is certain to cause immense confusion even if we can avoid strict inconsistency.

Notice what now follows on this species-centric view. First, generalized manipulation principles fail. A general principle of expansion implies that if we took a species tree and replaced a single species with multiple populations from that species, we would not need to adjust the remainder of the tree in any way. Generalized contraction implies that if we had a population tree consisting of populations in many different species, and we then collapsed the conspecifics into single groups, we would not need to adjust the tree. If we allow paraphyletic species, now or in the past, and use species-centric phylogenies, both are false. Similar problems exist for addition and subtraction of branches. Each of these failures requires that we can have mixed species and population trees, or at least can transition from one to the other without changing topologies. If we can't do this because they mean different things, then we can still hold on to restricted forms of the manipulation principles.

Perhaps more importantly, when we build phylogenetic trees, it would be essential to decide if we are building a species tree or a population tree. The same tips and the same evidence could lead us to produce different trees depending on which we were hoping to build. If we asked, “how are chimps related to gorillas and humans?” the proper response might be to say, “be more specific, do you mean how are the populations related? Or, how are the species related?” If we were to build the tree using data from individual organisms that we take to be representative of larger groups, we would need to know whether we intend these larger groups to be species or populations. It would be essential to ask this question because the correct answer about relationships could vary depending on what type of tree you were interested in (Velasco 2008).

In the desired compatibility situation, we could build a tree of groups and say that it doesn't matter whether the tips are all separate species or (e.g.) whether two tips are different populations of the same species. But according to the view that population trees are interpreted differently than species trees, this would matter. Since how a group is related to another now depends on whether we are asking about species

or population history, it becomes essential to rank groups properly. For example, whether a particular group is a subspecies or a genuine species might make a difference in how it is phylogenetically related to other groups. The topologies of the population and species trees it is a part of might be different.

A related worry is that even if we determined that we were looking for a species tree, we would need to specify some particular species concept by which we were building our tree. Simply pointing out that we have evidence that the genes, organisms, or populations were related in certain ways wouldn't entail how the species were related. And we would have to settle on more than a particular grouping criterion—a ranking criterion would be needed as well.

LaPorte (2005) accepts the “speciation history” definition of species trees and argues that since there is no objectively correct definition of species, there is no objectively correct evolutionary tree. LaPorte is correct; if we define a species tree by speciation history, then different species concepts lead to different trees. One way out of this is to insist that only certain species concepts are correct. But barring agreement on which concepts those are, it would seem impossible for phylogeneticists to agree on the correctness of their trees. Yet in practice, they do agree. This could be explained by some kind of more general agreement about what species are, but this type of proposal faces several problems.

One common reason for thinking that species are special is that they are the units of evolution. This is certainly a popular view; an anthology of important papers on species has been called *The Units of Evolution: Essays on the Nature of Species* (Ereshefsky 1992). It is perfectly reasonable to attempt to figure out whether evolutionary units are objective features of the world and what they are like. What is questionable is whether these evolutionary units uniquely pick out species.

It is unclear what different authors mean by “evolutionary unit” (various possibilities are surveyed in Ereshefsky 1991). Some authors argue that species play a role in the theory of evolution, which requires them to be individuals (Hull 1980; Williams 1985). Others say that they are the unit of selection (Gould 2002) or perhaps objects of selection (Mayr 1997)—this means that species evolve whereas organisms, genes, and other biological entities do not. In some extreme cases, the idea of an evolutionary process is built directly on the idea of what happens to species through time (Eldredge and Cracraft 1980). But whether this is accepted, it is normally thought that to be a unit of evolution means that regarding evolutionary forces—such as a changing environment leading to different selection pressures—the group responds as a single unit with a shared fate (Wiley 1981; Mayden 1997).

While some authors have argued that there are unique, objective evolutionary groups that fit this definition (Simpson 1961; Wiley 1981; Mayden 1997), others point out that there are no such universal, fundamental units. They hold that there are important, identifiable discontinuities in genealogical, interbreeding, genetic, morphological, and ecological groups, but that these do not overlap. Different groups respond differently to different kinds of pressures such that some groups might be a unit regarding some forces but not others (Dupré 1999, 2001; Mishler and Donoghue 1982; Wilson 2005). Being a unit and responding as a whole also seems to come in degrees; there will be nested and perhaps even overlapping groups, which are units to some degree or other. On the face of it, small family units, demes, local populations, and groups of multiple populations all seem to be good candidates for groups that are units with respect to evolutionary theory. Though it is doubtful that there is a single level in nature that represents the unit of evolution, if I were forced to pick such a level, then I would think that the population is the natural level. Conspecific populations are not united regarding the forces of evolution, which is why species typically evolve from single populations and not from entire ancestral species (Mishler and Brandon 1987). The very phenomenon of paraphyletic species serves to reinforce this point. When species trees and population trees disagree, this discordance arises because some species has not functioned as a single evolutionary unit.

While a number of authors have thought that species, unlike subspecies or genera, have unique and important properties, many others have doubted this (Ereshefsky 1991; Mishler 1999; Mallett et al. 2007; Baum 2009). They see species as one level in a hierarchy—a level that is rather arbitrary. Sometimes authors espousing this position hold that it is reasonable to think that species don't exist. Another view is that if species do exist, their boundaries are arbitrary in a way. This view is not new—it goes back to Darwin (Darwin 1859; Mallett 2008a,b; Ereshefsky 2009).

Even if we could specify some reason that species are uniquely important, we should not think that the phylogeny of species are determined by speciation events. Cases of disagreement between the species trees and the population trees arise because our classifications through time do not always correspond to the genealogical histories of the groups. But why should our phylogenies track our classifications? Velasco (2008) describes two types of disagreement: the “No Tree” problem and the “Wrong Tree” problem. The names are appropriate if we assume that a correct phylogeny tracks the history of populations. Imagine it doesn't. Then since our phylogenetic inference methods track the history of genes, organisms, and populations, they will deliver the population trees and not the species trees. This is good because if we attempted to use the species trees as phylogenies, our inferences would be mistaken in the same way as if we artificially placed crocodiles, lizards, and turtles in a clade on a tree because they were all classified as reptiles. In other words, our methods are aimed at recovering genealogy. If we were aiming for the history of speciation, our methods are no good. It would be foolish to say that the history of speciation is just unattainable as a goal, but what we can be sure of is that this is not what systematists are interested in and not what currently used phylogenetic methods can recover.

The view that I am advocating might be construed as rank-free phylogeny. Calls for a rank-free taxonomy are controversial, and they are made even more so by the fact that traditional taxonomy requires ranks. But the call for rank-free phylogeny should not be so controversial. Phylogeneticists have behaved this way for years. Ongoing debates about the species problem have not affected phylogenetics in any way. Debates about classification have focused on how to use a phylogeny if we knew it, and one major position in the species debates (the phylogenetic species concept) requires phylogenies to be in place *first* to define species. While opponents do not think that this is the correct species concept, they do not (typically) argue for their position on the basis of the claim that one cannot make sense of a phylogeny independent of classification.

In the end, the view that population trees and species trees represent fundamentally different things is an unacceptable view. We can use species trees in the sense that we can use phylogenetic trees that have species at their tips. But even if we agreed that species have a unique ontological status and identity through time, and we agreed that populations also had identity through time, what counts as a genealogical relationship between populations or between species for the purposes of phylogenetics requires that the phylogenetic relationships between groups of these different kinds be determined in the same way—by population-level splits. Population trees and species trees must be compatible.

8. What about Species?

So if species trees and population trees must be consistent, what does this mean for species? As the reader may have noticed, the most problematic cases mentioned in Section 7 stem from the fact that species might now be, or have been in the past, a paraphyletic group of populations. Defenders of a history-based version of the phylogenetic species concept do not accept that paraphyletic species exist (Mishler and Donoghue 1982; Mishler 1985; Donoghue 1985; Velasco 2008). Such views avoid many of the above problems. But apart from commitments to the phylogenetic species concept, there are powerful arguments for avoiding species-centric definitions of phylogeny.

First, as noted above, such authors often use population trees to determine which groups are candidates for species. So population trees must at least have some meaning apart from species trees. We can imagine that we have both species trees and population trees and that these are always consistent. One suspects that this is the same as the population-centric view. After all, the species trees would have branches that do correspond to population-level lineages even if they “directly represent” species lineages and their splits are population-level splits even if they directly represent speciation events. The idea that these are somehow metaphysically different and carry a different interpretation than population trees, even though they are—for all practical applications—exactly the same, strikes me as gratuitous and implausible.

On a strict reading of the species-centric view, we don't have exact equivalence of the population trees and the species trees anyway. Imagine a case where we have already had significant phylogenetic branching

of populations within a species. Consider the ladder-shaped pectinate phylogeny of four populations (A1, (A2,(A3, A4))). Now if A4 splits into two groups, call them E and F, including one uncontroversial new species, to avoid paraphyletic species, each group must now be a separate species. Labeling A1=A, A2=B, and A3=C, the population-level phylogeny is still the ladder as before: A, (B, (C, (E, F))). However, the natural way to interpret a species-centric view is that phylogeny depends not just on which groups are species, but also on when they became species. So on a strict reading of a species-centric view, there would be a massive polytomy resulting in a star phylogeny since all such groups speciated from each other at exactly the same time. Because we want to recover the more structured history, no competent systematist would argue for a star phylogeny as the species tree. But this is because they do not really believe that the species tree is determined by speciation events, but rather by the history of the lineages. In the distant past (say at t_1) there was a split between a group whose descendants later became A1 and a group whose descendants form A2, A3, and A4. This was not a speciation event at the time. But once these groups' descendants become independent species of their own, we will want to represent this split on the species tree as we would on any acceptable phylogeny. Thus even with the most narrowly defined phylogenetic species concept, there is still a reason to think of phylogenies as fundamentally about population histories rather than species histories.

I do not want anything significant to hinge on such examples. A species-centric phylogeneticist might attempt to retroactively call the split at t_1 a speciation event or claim that it was a speciation event all along because of what would happen in the future (O'Hara 1993). Determining exactly when speciation occurs is a difficult issue and one that does not have any problem free solutions when combined with history-based phylogenies.⁴ But whatever we want to say about this issue, restricting ourselves to a particular kind of history-based species concept does not immediately make the species-centric view unproblematic. The only way to avoid species trees coming apart from population trees is to have a species concept whereby anything that would count as a population split also counts as speciation. This leads to the view that there can never be two distinct populations of the same species and is so far divorced from taxonomic practice that I consider it to be unacceptable. But it is worth noting that this is perhaps the ultimate expression of the "population-centric" view.

So no matter what our species concept, there is a reason to treat phylogenies as fundamentally about population histories and not species histories. Where does this leave us with respect to species? Harrison (1998) argues that there are good reasons to think that species as a whole have a life history and a species may be paraphyletic for much of its history. As for what happens when we place them on trees, Harrison says that we should not start with the assumption that species must be units of phylogeny. This presumably means that it is not always okay to place species at the tip of a phylogenetic tree and so we would have to eliminate the traditional understanding of species trees. (This is consistent with what I have argued here.)

The solution to these tensions is obvious. Phylogenetic trees should not be defined by the history of speciations but rather by the history of lineages and the pattern of their divergences. This conclusion is independent of any particular view of species. It may be that there are important reasons to recognize the process of speciation and for this we should have some particular non-phylogenetic species definition (Harrison 1998; Coyne and Orr 2004). But then species are not the foundation of phylogeny. It may be that it is so important to recognize phylogenetic patterns that our taxonomies should be phylogenetically based and species, as taxa, should be based on phylogeny as well (Baum 2009). But species still cannot serve as the foundation of phylogeny. The groups that do serve as foundational in either case are subspecific, historically united groups—populations—and thus it is appropriate to think of phylogenetic trees as population trees.

What is not consistent is to hold that (a) higher taxa are clades, (b) only clades can serve as the tips of phylogenetic trees, and (c) clades are defined with species and speciations, and then use a species concept that allows for paraphyletic species. This combination of views is widely held, explicitly so by some defenders of the PhyloCode, such as Kevin de Queiroz who argues for a lineage-based concept of species (de Queiroz 1998, 1999). In the PhyloCode, clades can be specified by species, yet species are treated in a fundamentally different way than clades and can be named in any way consistent with the traditional ICBN and ICZN codes (Dayrat et al. 2008; <http://phylocode.org>). This allows for the possibility of paraphyletic species. Clades (and so phylogenetic trees) can then supposedly be determined from these species, but this can lead to inconsistencies. Criticisms of how the PhyloCode deals with species have been made (e.g.,

Cellinese et al. 2012), but note that this problematic combination of views is held by many systematists who have differing views on classification, species, and nomenclature, and not just those who subscribe to the view espoused in the preface of the PhyloCode.

We want to build and use phylogenetic trees, but we need some notion of what can constitute an acceptable taxon on a tree. In other words, we need to know what a clade is. Traditionally this has required the use of species, but it does not need to be this way. The possibility of building phylogenies without the prior use of species has already been demonstrated with the additional call to get rid of the species category altogether (Pleijel 1999; Pleijel and Rouse 2000a,b; Fisher 2006; Fisher et al. 2007). While this position is consistent with my arguments here, it does not follow from them. We could continue to use “species” for classification and, when a species is a monophyletic group of populations, we could place it on a phylogenetic tree. But whether we keep the species rank or not, species cannot be the foundation of phylogenetics.

9. The Power of Semantics

That phylogenetic trees are species trees is entrenched in the language of biologists, not only because of higher systematics but also because of the enormous literature on the issue of how gene trees relate to species trees, which is often approached using population genetic methods (for a survey, see Edwards 2009). Frequently, questions of definitions and semantics are ignored by biologists who assume that biological practice will determine such issues. Scott Edwards, a researcher who works both above and below the species level, deals directly with these issues. It is worth quoting him in full both for his conclusion and the tone in which it is presented.

WHAT’S IN A NAME? It is a legitimate question to ask, as a colleague of mine did recently, whether species trees have any validity if in fact the definition of species is still in limbo (as they are likely to be for a long time). This colleague suggested that the term “population tree” is better suited to the new paradigm, because it avoids the issue of species validity (notwithstanding the problem of defining populations in nature). I would be happy with this terminology, but defining it this way might seem to exonerate those working at higher taxonomic levels, for whom population processes are minor concerns. Phylogeneticists working on the higher level questions tend not to concern themselves with populations, or their genetics. For this reason, “population trees” might become appropriated solely by phylogeographers and those working near the species level. This would be unfortunate, because gene tree heterogeneity and the species tree problem in principle affects all levels of phylogeny, even if the extent of deep coalescence or branch length heterogeneity is less among higher taxa or sparsely sampled clades. For this reason I suggest we simply exercise a verbal substitution and reserve the term “phylogeny” to refer to species trees. Phylogenies as they have been built in the last few decades would then be called gene trees, which is generally what they are, *sensu stricto* (Edwards 2009, 12).

The questioner is worried that since there is some problem about what species are, then there should be a parallel problem about species trees. Not so, suggests Edwards. This is because there is a large research program dedicated to separating species trees from gene trees and, within that paradigm, it is quite clear what is meant by a species tree—the splits are lineage branchings and not necessarily speciations (as measured by some definition). When his colleague suggests the term “population tree” instead, clearly Edwards thinks that these are the same thing—that it is merely a question of what we ought to call the things that we are inferring, not that species trees and population trees are genuinely different or that there is a genuine question about which of the two we are inferring. The primary reason to continue using the name “species trees” is so that this important research does not get ignored in areas where it is needed.

Much of the literature in this field already speaks of population trees rather than species trees, but the diagrams and the methods are the same—only the labels have changed (Takahata 1989; Knowles and Maddison 2002). Others explicitly identify species trees with population trees (Pamilo and Nei 1988; Avise 1989). The terminology seems to be a matter of taste. The figures in Edwards (2009) suggest that populations and species are different things but that the topologies between these are always consistent.

Whether we simply eliminate talk of species trees, identify species trees with population trees, or view them as distinguishable but always consistent, it is appropriate to think of phylogenetic trees as population trees representing the history of population lineages through time. Rather than being revolutionary, the idea that species trees are population trees is already deeply embedded as part of phylogenetic practice.

10. Conclusion

The divergence of lineages through time is a continual process with no obvious thresholds. At the very bottom we have local demes with little or no mating structure. Moving to larger populations, especially those with geographical patterns, we get mating structure within populations. When this structure persists through time, phylogenetic patterns begin to emerge. Populations split and, over time, gene flow is reduced and the patterns are reinforced; significant differences evolve in one population but not another and some systematists would recognize a new species. Often gene flow between the groups continues to be reduced over time and the groups continue to diverge ecologically, morphologically, and genetically. Eventually, by any species definition, we have distinct species. I told this story in a way familiar to those who study speciation. Indeed, this is the standard model of speciation as a process of divergence through time. While it certainly doesn't always happen this way, it often does.

But it would be a mistake to think of this story as a speciation story with reticulation throughout that ends with distinct species and signals the start of a new story—phylogenetic branching with a lack of reticulation. There is no important, completely general, species boundary in nature. Species are but one level of a continuum. If there is a species rank, it is arbitrary in an important way. If we were to base our understanding of phylogeny and all of systematics on this arbitrary and often times misleading system of ranking, phylogenetics would be a flawed science.

Fortunately, we do not have to base phylogenetics on species. We can, and should, have rank-free phylogenies. It is important to recognize that phylogeny is not dependent on classification. It may be that species are important biological units and there is good reason to preserve the rank of species. It could even be that the higher ranks like genus and family serve important functions for classification. But these conclusions do not affect the practice of phylogenetics. We should recognize that phylogeny and genuine divergence is present from the beginning of our story, albeit in a messy and not so easily tractable way. The story of divergence in the presence of reticulation continues above the traditional species level. It is a story that unites the fields of systematics and population genetics in a way that recognizes the genuine contributions that each field makes to the other (Avice 1989; Baum and Shaw 1995). These contributions are required to truly understand and achieve the goal of recovering the evolutionary history of life.

Literature cited

- Ané, C., B. Larget, D.A. Baum, S.D. Smith, and A. Rokas. 2007. Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution* 24: 412–426. [doi:10.1093/molbev/msl170](https://doi.org/10.1093/molbev/msl170)
- Arnold, M.L. 2009. *Reticulate Evolution in Humans: Origins and Ecology*. Oxford: Oxford University Press.
- Avice, J.C. 1989. A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology and Evolution* 4: 279–281. [doi:10.1016/0169-5347\(89\)90203-6](https://doi.org/10.1016/0169-5347(89)90203-6)
- Avice, J.C. 2006. *Evolutionary Pathways in Nature: A Phylogenetic Approach*. Cambridge: Cambridge University Press. [doi:10.1017/CBO9780511606939](https://doi.org/10.1017/CBO9780511606939)
- Avice, J.C. and R.M. Ball. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* 7: 45–67.
- Baum, D.A. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* 56: 417–426.
- Baum, D.A. 2008. Individuality and the existence of species through time. *Systematic Biology* 47: 641–653. [doi:10.1080/106351598260644](https://doi.org/10.1080/106351598260644)
- Baum, D.A. 2009. Species as ranked taxa. *Systematic Biology* 58: 74–86. [doi:10.1093/sysbio/syp011](https://doi.org/10.1093/sysbio/syp011)



- Baum, D.A. and K.L. Shaw. 1995. Genealogical perspectives on the species problem. In: *Experimental and Molecular Approaches to Plant Biosystematics*. Ed. by P.C. Hoch, A.G. Stephenson. St. Louis, MO: Missouri Botanical Garden.
- Baum, D.A. and S.D. Smith. 2013. *Tree Thinking: An Introduction to Phylogenetic Biology*. Greenwood Village, CO: Roberts and Company Publishers.
- Baum, D.A., S.D. Smith, and S.S.S. Donovan. 2005. The Tree-Thinking Challenge. *Science* 310: 979–980. doi: [10.1126/science.1117727](https://doi.org/10.1126/science.1117727)
- Carroll, R.L. 1988. *Vertebrate Paleontology and Evolution*. New York: W.H. Freeman.
- Cavalli-Sforza, L.P., P. Menozzi, and A. Piazza. 1994. *The History and Geography of Human Genes*. Princeton: Princeton University Press.
- Cellinese, N., D.A. Baum, and B.D. Mishler. 2012. Species and Phylogenetic Nomenclature. *Systematic Biology* 61: 885–891. doi: [10.1093/sysbio/sys035](https://doi.org/10.1093/sysbio/sys035)
- Coyne, J.A. and H.A. Orr. 2004. *Speciation*. Sunderland, MA: Sinauer Associates.
- Cracraft, J. 1997. Species concepts in systematics and conservation biology—an ornithological viewpoint. In: *Species: The Units of Biodiversity*. Ed. by M.F. Claridge, H.A. Dawah, M.R. Wilson. London: Chapman and Hall.
- Cracraft, J., and M.J. Donoghue. 2004. Introduction. In: *Assembling the Tree of Life*. Ed. by J. Cracraft, M.J. Donoghue. New York: Oxford University Press.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.
- Dawkins, R. 2004. *The Ancestor's Tale*. New York: Houghton Mifflin Co.
- Dayrat, B., P.D. Cantino, J.A. Clarke, and K. de Queiroz. 2008. Species names in the PhyloCode: the approach adopted by the International Society for Phylogenetic Nomenclature. *Systematic Biology* 57: 507–514. doi: [10.1080/10635150802172176](https://doi.org/10.1080/10635150802172176)
- de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: *Endless Forms: Species and Speciation*. Ed. by D.J. Howard, S.H. Berlocher. New York: Oxford University Press.
- de Queiroz, K. 1999. The general lineage concept of species and the defining properties of the species category. In: *Species: New Interdisciplinary Essays*. Ed. by R.A. Wilson. Cambridge, MA: MIT Press.
- Donoghue, M.J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist* 88: 172–181. doi: [10.2307/3243026](https://doi.org/10.2307/3243026)
- Dumbacher, J.P. and R.C. Fleisher. 2001. Phylogenetic evidence for colour pattern convergence in toxic pitohuis: Müllerian mimicry in birds? *Proceedings of the Royal Society of London B* 268: 1971–6. doi: [10.1098/rspb.2001.1717](https://doi.org/10.1098/rspb.2001.1717)
- Dupré, J. 1999. On the impossibility of a monistic account of species. In: *Species: New Interdisciplinary Essays*. Ed. by R.A. Wilson. Cambridge, MA: MIT Press.
- Dupré, J. 2001. In defence of classification, *Studies in the History and Philosophy of the Biological and Biomedical Sciences* 32: 203–219. doi: [10.1016/S1369-8486\(01\)00003-6](https://doi.org/10.1016/S1369-8486(01)00003-6)
- Eldredge, N. and J. Cracraft. 1980. *Phylogenetic Patterns and the Evolutionary Process: Method and Theory in Comparative Biology*. New York: Columbia University Press.
- Edwards, S.V. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* 63: 1–19. doi: [10.1111/j.1558-5646.2008.00549.x](https://doi.org/10.1111/j.1558-5646.2008.00549.x)
- Ereshefsky, M. 1991. Species, higher taxa, and the units of evolution. *Philosophy of Science* 58: 84–101. doi: [10.1086/289600](https://doi.org/10.1086/289600)
- Ereshefsky, M. 1992. *The Units of Evolution: Essays on the Nature of Species*. Cambridge, MA: MIT Press.
- Ereshefsky, M. 2009. Darwin's solution to the species problem. *Synthese* 175: 405–425. doi: [10.1007/s11229-009-9538-4](https://doi.org/10.1007/s11229-009-9538-4)
- Fisher, K. 2006. Rank-free monography: A practical example from the moss clade *Leucophanella* (Calymperaceae). *Systematic Botany* 31: 13–30. doi: [10.1600/036364406775971697](https://doi.org/10.1600/036364406775971697)
- Fisher, K.M., D.P. Wall, K.L. Yip, and B.D. Mishler. 2007. Phylogeny of the Calymperaceae with a rank-free systematic treatment. *Bryologist* 110: 46–73. doi: [10.1639/0007-2745\(2007\)110\[46:POTCWA\]2.0.CO;2](https://doi.org/10.1639/0007-2745(2007)110[46:POTCWA]2.0.CO;2)

- Gould, S.J. 2002. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.
- Harrison, R. 1998. Linking evolutionary pattern and process: The relevance of species concepts for the study of speciation. In: *Speciation and its Consequences*. Ed. by D. Otte, J.A. Endler. Sunderland, MA: Sinauer Associates.
- Hey, J., R.S. Waples, M.L. Arnold, R.K. Butlin, and R.G. Harrison. 2003. Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology & Evolution* 18: 597–603. doi:[10.1016/j.tree.2003.08.014](https://doi.org/10.1016/j.tree.2003.08.014)
- Hodkinson, T.R. and J.A.N. Parnell. 2006. *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*. Boca Raton, FL: CRC Press. doi:[10.1201/9781420009538](https://doi.org/10.1201/9781420009538)
- Hull, D.L. 1980. Individuality and selection. *Annual Review of Ecology and Systematics* 11: 311–332. doi:[10.1146/annurev.es.11.110180.001523](https://doi.org/10.1146/annurev.es.11.110180.001523)
- Kaessmann, H., V. Wiebe, and S. Paabo. 1999. Extensive nuclear DNA sequence diversity among chimpanzees. *Science* 286: 1159–1162. doi:[10.1126/science.286.5442.1159](https://doi.org/10.1126/science.286.5442.1159)
- Knowles, L.L. and W.P. Maddison. 2002. Statistical phylogeography. *Molecular Ecology* 11: 2623–2635. doi:[10.1046/j.1365-294X.2002.01637.x](https://doi.org/10.1046/j.1365-294X.2002.01637.x)
- Kornet, D.J. 1993. Permanent splits as speciation events: a formal reconstruction of the internodal species concept. *Journal of Theoretical Biology* 164: 407–435. doi:[10.1006/jtbi.1993.1164](https://doi.org/10.1006/jtbi.1993.1164)
- LaPorte, J. 2005. Is there a single objective, evolutionary tree of life? *The Journal of Philosophy* 102: 357–374.
- Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* 20: 229–237. doi:[10.1016/j.tree.2005.02.010](https://doi.org/10.1016/j.tree.2005.02.010)
- Mallet, J. 2008a. Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society of London B* 363: 2971–2986. doi:[10.1098/rstb.2008.0081](https://doi.org/10.1098/rstb.2008.0081)
- Mallet, J. 2008b. Mayr's view of Darwin: was Darwin wrong about speciation? *Biological Journal of the Linnean Society* 95: 3–16. doi:[10.1111/j.1095-8312.2008.01089.x](https://doi.org/10.1111/j.1095-8312.2008.01089.x)
- Mallet, J., M. Beltrán, W. Neukirchen, and M. Linares. 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology* 7: 28. doi:[10.1186/1471-2148-7-28](https://doi.org/10.1186/1471-2148-7-28)
- Mayden, R.L. 1997. A hierarchy of species concepts: The denouement in the saga of the species problem. In: *Species: The Units of Biodiversity*. Ed. by M.F. Claridge, H.A. Dawah, M.R. Wilson. London: Chapman and Hall.
- Mayr, E. 1997. The objects of selection. *Proceedings of the National Academy of Sciences* 94: 2091–2094. doi:[10.1073/pnas.94.6.2091](https://doi.org/10.1073/pnas.94.6.2091)
- Millstein, R. 2010. The concepts of population and metapopulation in evolutionary biology and ecology. In: *Evolution Since Darwin: The First 150 Years*. Ed. by M.A. Bell, D.J. Futuyma, W.F. Eanes, J.S. Levinton. Sunderland, MA: Sinauer.
- Mishler, B.D. 1985. The morphological, developmental, and phylogenetic basis of species concepts in bryophytes. *The Bryologist* 88: 207–214. doi:[10.2307/3243030](https://doi.org/10.2307/3243030)
- Mishler, B.D. 1999. Getting rid of species. In: *Species: New Interdisciplinary Essays*. Ed. by R.A. Wilson. Cambridge, MA: MIT Press.
- Mishler, B.D. and R.N. Brandon. 1987. Individuality, pluralism, and the phylogenetic species concept. *Biology & Philosophy* 2: 397–414. doi:[10.1007/BF00127698](https://doi.org/10.1007/BF00127698)
- Mishler, B.D. and M.J. Donoghue. 1982. Species concepts: a case for pluralism. *Systematic Zoology* 31: 491–503. doi:[10.2307/2413371](https://doi.org/10.2307/2413371)
- Morin, P.A., J.J. Moore, R. Chakraborty, L. Jin, J. Goodall, and D.S. Woodruff. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* 265: 1193–1201. doi:[10.1126/science.7915048](https://doi.org/10.1126/science.7915048)
- Moritz, C.C. 1994. Defining “evolutionary significant units” for conservation: a critical review. *Molecular Ecology* 3: 401–411. doi:[10.1111/j.1365-294X.1994.tb00080.x](https://doi.org/10.1111/j.1365-294X.1994.tb00080.x)
- O'Hara, R.J. 1993. Systematic generalization, historical fate, and the species problem. *Systematic Biology* 42: 231–246. doi:[10.1093/sysbio/42.3.231](https://doi.org/10.1093/sysbio/42.3.231)

- O'Hara, R.J. 1997. Population thinking and tree thinking in systematics. *Zoologica Scripta* 26: 326–329. [doi:10.1111/j.1463-6409.1997.tb00422.x](https://doi.org/10.1111/j.1463-6409.1997.tb00422.x)
- Pamilo, P. and M. Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5: 568–583.
- Pleigel, F. 1999. Phylogenetic taxonomy, a farewell to species, and a revision of *Heteropodarke* (Hesionidae, Polychaeta, Annelida). *Systematic Biology* 48: 755–789. [doi:10.1080/106351599260003](https://doi.org/10.1080/106351599260003)
- Pleigel, F. and G. Rouse. 2000a. Least-inclusive taxonomic unit: A new taxonomic concept for biology. *Proceedings of the Royal Society of London Series B: Biological Sciences* 267: 627–630. [doi:10.1098/rspb.2000.1048](https://doi.org/10.1098/rspb.2000.1048)
- Pleijel, F. and G. Rouse. 2000b. A new taxon, *Capricornia* (Hesionidae, Polychaeta), illustrating the LITU ('Least-Inclusive Taxonomic Unit') concept. *Zoologica Scripta* 29: 157–168. [doi:10.1046/j.1463-6409.2000.00041.x](https://doi.org/10.1046/j.1463-6409.2000.00041.x)
- Ridley M. 1989. The cladistic solution to the species problem. *Biology & Philosophy* 4: 1–16. [doi:10.1007/BF00144036](https://doi.org/10.1007/BF00144036)
- Ryder, O.A. 1986. Species conservation and systematics and the dilemma of subspecies. *Trends in Ecology and Evolution* 1: 9–10. [doi:10.1016/0169-5347\(86\)90059-5](https://doi.org/10.1016/0169-5347(86)90059-5)
- Shedlock, A.M. and S.V. Edwards. 2009. Amniotes (Amniota). In: *The Timetree of Life*. Ed. by S.B. Hedges, S. Kumar. New York: Oxford University Press.
- Simpson, G.G. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.
- Steel, M.A. 2007. Tools to construct and study big trees: A mathematical perspective. In: *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*. Ed. by T.R. Hodkinson, J.A.N. Parnell. Boca Raton, FL: CRC Press.
- Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* 122: 957–966.
- Templeton, A.R. 2002. Out of Africa again and again. *Nature* 416: 45–51. [doi:10.1038/416045a](https://doi.org/10.1038/416045a)
- Templeton, A.R. 2007. Haplotype trees and modern human origins. *Yearbook of Physical Anthropology* 48: 33–59.
- Templeton A.R. 2007. Genetics and recent human evolution. *Evolution* 61: 1507–19. [doi:10.1111/j.1558-5646.2007.00164.x](https://doi.org/10.1111/j.1558-5646.2007.00164.x)
- Van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon* 25: 233–239. [doi:10.2307/1219444](https://doi.org/10.2307/1219444)
- Velasco, J.D. 2008. Species concepts should not conflict with evolutionary history, but often do. *Studies in the History and Philosophy of Biological and Biomedical Sciences* 39: 407–414. [doi:10.1016/j.shpsc.2008.09.007](https://doi.org/10.1016/j.shpsc.2008.09.007)
- Velasco, J.D. 2009. When monophyly is not enough: exclusivity as the key to defining a phylogenetic species concept. *Biology & Philosophy* 24: 473–486. [doi:10.1007/s10539-009-9151-4](https://doi.org/10.1007/s10539-009-9151-4)
- Velasco, J.D. 2010. Species, genes, and the tree of life. *The British Journal for the Philosophy of Science* 61: 599–619. [doi:10.1093/bjps/axp051](https://doi.org/10.1093/bjps/axp051)
- Velasco, J.D. 2012. The future of systematics: tree thinking without the tree. *Philosophy of Science* 79: 624–636. [doi:10.1086/667878](https://doi.org/10.1086/667878)
- Wiley, E. 1981. *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. New York: Wiley-Interscience.
- Williams, M. 1985. Species are individuals: theoretical foundations for the claim. *Philosophy of Science* 52: 578–90. [doi:10.1086/289276](https://doi.org/10.1086/289276)
- Wilson, R.A. 2005. *Genes and the Agents of Life: The Individual in the Fragile Sciences: Biology*. New York: Cambridge University Press.
- Won Y-J. and J. Hey. 2005. Divergence population genetics of chimpanzees. *Molecular Biology and Evolution* 22: 297–307. [doi:10.1093/molbev/msi017](https://doi.org/10.1093/molbev/msi017)

Notes

1. A reviewer points out that we also have phylogenies of genes (or various molecular bits) and what I say about phylogenies generally will not apply to phylogenies of genes. This is true and I do not mean to suggest that these are not properly called “phylogenies.” This is mainly a semantic issue and some authors (e.g., Hennig) use “phylogeny” in a specific way that would not apply to genes. However, it is so common now to refer to molecular phylogenies that I see no reason to insist on overturning that. But molecular phylogenies are clearly a different kind of thing than a phylogeny of a group of organisms, which is the only thing I will talk about here. Whether there are multiple kinds of phylogenies of organisms and what the relationships between them are is the subject of the paper and I will leave genes for another time. It is generally assumed that phylogenies of genes and phylogenies of groups of organisms are conceptually distinct (but see, e.g., Baum 2009 and Velasco 2010).
2. I assume a basic familiarity with phylogenetic trees and terms like monophyly and paraphyly. The ideal introduction for our purposes is Velasco (2008), but for a brief and more general introduction, see Baum et al. (2005). For a longer one, see Baum and Smith (2012).
3. For a more thorough discussion of variants of this example and what it means for how we should think about species, see Velasco (2008).
4. For careful examinations of these diachronic issues, see (e.g.) Baum and Shaw (1995) and Baum (1998).

ACKNOWLEDGMENTS

I would like to thank the editors at *Philosophy & Theory in Biology* and two anonymous referees for helpful comments on drafts of this paper. I would also like to thank numerous audiences and readers of various papers and talks of mine, many of which overlap with the contents of this paper and helped shape my thinking on this topic.

Copyright © 2013 Author(s).

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs license, which permits anyone to download, copy, distribute, or display the full text without asking for permission, provided that the creator(s) are given full credit, no derivative works are created, and the work is not used for commercial purposes.

ISSN 1949-0739