Joel Velasco

Department of Philosophy

Stanford University

Draft as of 8-14-09 (words in text 10,241)

Phylogeny as Population History

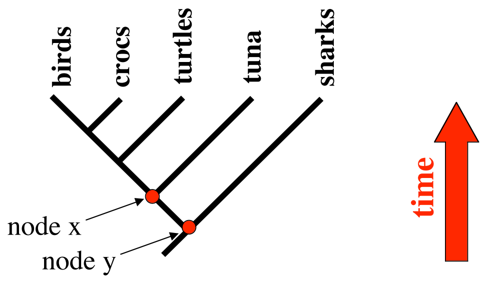
1. Introduction

The project of this chapter 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 this trivial way, in fact their interpretation is tied to foundational conceptual questions at the heart of systematics – questions whose answers are hotly disputed. As I will argue, common practices for the interpretation and use of trees are actually in conflict and therefore some must be given up. Here I defend the view 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. The branches on trees represent collections of population lineages through time and the splits represent population lineage splits. I will argue for this view by focusing on how we use phylogenetic trees for inferences about the evolutionary past.

2. Reading a tree

Figure 1 displays the relationships of several chordate groups placed at the tips of an evolutionary, or phylogenetic, tree. The tips of a phylogenetic tree are usually referred to as “taxa” (singular taxon). The word “taxa” is also used to describe the formally named groups, such as *Homo sapiens*, in our classification system. This is, of course, not a coincidence, as trees are typically understood to display the relationships between the taxa in our taxonomies. Any set of taxa shares a common ancestor deep in the past. As time passes, we move up the tree and single lineages divide into descendant lineages. The nodes represent points of divergence – for example, at the earliest node, node y, one of these lineages eventually leads to the sharks while the other leads to the other four groups.

Figure 1. A phylogenetic tree depicting the genealogical relationships between various groups of vertebrate animals.



Groups shown on a phylogenetic tree are understood to be related to each other genealogically in the same way that individual organisms within a family tree are. You are more closely related to your first cousins than to your second cousins since you and your first cousins share grandparents rather than just great-grandparents. Similarly, tuna are more closely related to birds than they are to sharks since tuna and birds share a common ancestor at node x, while tuna and sharks share an ancestor at node y (which is earlier in time). There is an ancestor shared by tuna and birds (at node x) that is not an ancestor of sharks; on the other hand, every ancestor shared by tuna and sharks is also an ancestor of birds. By generalizing this example, one can see that horizontal distance among the tips on a tree means nothing when it comes to degree of relatedness. Though it may seem that tuna and sharks share many properties that might lead you to believe that they are closely related, it is recency of common ancestry, not morphological similarity, which defines genealogical relationships.

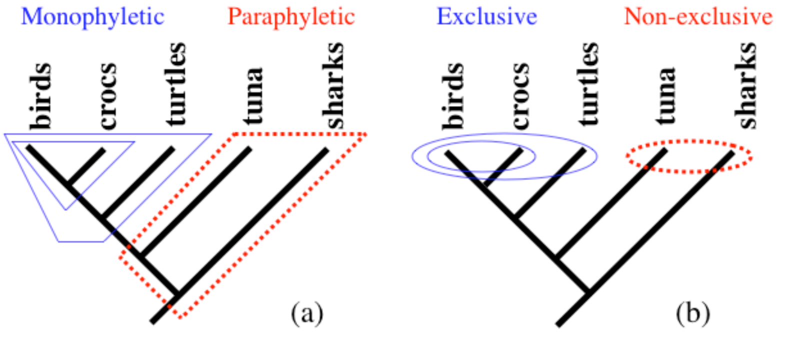
An essential concept for phylogenetics is the idea of a monophyletic group. A monophyletic group consists of an ancestor and all of its descendants. Two nested monophyletic groups are depicted in Figure 2a. Birds + crocs form a monophyletic group within the larger monophyletic group composed of birds + crocs + turtles.[1] Because monophyletic groups start from a particular node in a tree and then spread, on any tree there is a unique branch leading to them and only them. Thus they are also called clades (from the Greek *klados* for branch).

---

[1] There is some ambiguity in this standard definition of monophyly. The group that is truly monophyletic is the most recent common ancestor of birds, crocs, and turtles together with all of its descendants. This group definitely includes various extinct taxa (such as the dinosaurs) and depending on evolutionary history of which we are uncertain, it might include the lepidosaurs (lizards, snakes, and tuataras) as well. For a detailed, but early discussion of the debates about this phylogeny, see Laurin and Gauthier (1996). For a more recent survey focusing on molecular evidence, see Shedlock and Edwards (2009). In any case, there are additional taxa in this monophyletic group, but they are not on this tree. What matters for this example is that neither tuna nor sharks are nested in this group.

Groups are called "paraphyletic" when they include an ancestor, but not all of its descendants. Groups that include multiple groups of organisms, but that do not include their common ancestor are called "polyphyletic." The most recent common ancestor of tuna and sharks is also an ancestor of birds. This means that tuna + sharks is not monophyletic.

Figure 2. A monophyletic group is composed of an ancestor and all of its descendants. A paraphyletic group includes only some of those descendants. An exclusive group is a contemporaneous group of organisms more closely related to each other than to anything outside the group.



A tree is just a representation of which groups are monophyletic groups. A set of monophyletic groups yields a unique phylogenetic tree and vice versa. This allows for a compact representation of a tree. By using nested parentheses, we can depict the nested set of clades. For example, we can represent Figure 1 as (((birds, crocs), turtles), tuna), sharks.

A key fact about monophyletic groups that partially explains their importance is that in a purely diverging tree (i.e., one in which branches split but never join), a single time-slice of a monophyletic group forms a group that is genealogically <u>exclusive</u> – members of the group at that time are more closely related to each other than any of them is to anything outside the group that exists at the same time (see Figure 2b). Non-monophyletic groups do not lead to exclusive groups. Exclusive groups can serve as valid tips of trees since there is a branch on the Tree of Life that leads to them and only to them.

3. The implications of a tree

Phylogenetic trees attempt to depict actual evolutionary history. Since the relationships of the taxa depicted in Figure 1 are purported to be objective facts about the world, any particular tree can be correct or incorrect. Thus it is sometimes important to distinguish a particular tree from the true tree, which we may not know.

Because these are genealogical relationships, these facts should stay the same when we add or remove taxa to the tree. For example, if Figure 1 is correct, then birds are more closely related to crocs than to tuna. This means that we can build a tree of just these three taxa that looks like this: (birds, crocs), tuna. This is called a subtree of the larger tree, Figure 1. All of the relationships implied by a subtree follow from the larger tree. So if a tree is correct, then all of its subtrees are too. And it is just a mathematical theorem that the full collection of subtrees trees for all sets of three taxa determines the full phylogenetic tree containing all these taxa (Steel 2007: 99).

And of course Figure 1 is itself a subtree of various larger trees. A single tree does not need to exhaustively depict all of the subgroups of a single clade to be correct. For example, if we wanted to depict the history of mammals as well, we could add that as a tip. Our best evidence indicates that the most recent common ancestor of mammals and birds was an ancestor of turtles but not of tuna. This would mean that the node leading to the mammal branch would be between the nodes leading to turtles and tuna. This example illustrates a general principle. Some taxa are already on a tree implicitly because they are part of a larger taxon at a tip (like snapping turtles are part of turtles). But for any taxon that is genuinely not on a correct tree, there is a unique place that it can be added to the tree to make the new tree correct.

The idea that any group can be consistently added to a tree depends both on the tips, and the group being added, being monophyletic groups. For example, "sharks" are a paraphyletic group. Some sharks are more closely related to rays and skates than to other sharks. So if we attempted to add rays to Figure 1, they would have no place to go. They would have to go inside the shark tip. When we discover that our taxa are paraphyletic in this way, we have to be more careful about which groups are on the tips. So far, we have been using informal names like "shark" as taxa. But we want taxa to be carefully delineated. In a situation like this, there are two alternatives (other than abandoning the tree or previous groupings). One is to say that "shark" just represents some taxon that includes all sharks – perhaps they stood for all the cartilaginous fishes, the Chondrichthyes. In that case, rays are already on the tree (as part of the sharks). This kind of move is very controversial when common names like "shark" are involved. People tend to think they know what a shark is (compare the claim that birds are reptiles). If we don't want to go that route, we would have to restrict the shark tip to a specific group such as the mackerel sharks (Lamniformes). Then the rays would be placed branching out from the lineage leading to sharks.

Some trees are subtrees of others and are directly contained inside the larger trees. By combining principles of adding and subtracting branches (subtraction is often called "pruning") we can show that there is a more general notion of two trees being consistent with each other. That is, they do not imply any contradictory relationships. As a general example, Figure 1 is consistent with the tree that has ((birds, crocs), mammals), tuna. But it is inconsistent with ((birds, mammals), tuna), crocs. This is because this latter tree implies that birds are closer to tuna than to crocs, which contradicts Figure 1. But notice that mammals could go anywhere and the tree could be consistent with Figure 1. Consistency is not a matter of truth directly – though of course if two trees are inconsistent than at least one of them implies some false relationships.

Consistency between trees involves looking at the genealogical implications of a tree. But not all such consequences are equally obvious. Because monophyletic groups share a unique history not shared by anything outside the group, it is possible to take a monophyletic group and "collapse" it into a single exclusive tip with a single lineage leading to it. For example, the "birds" tip in Figure 1 is already a collapsed lineage of around 9,000 living species, but it could be further combined with the crocodiles to form the Archosauria – a group that includes the dinosaurs (see Figure 3a). This is acceptable since everything outside of the archosaurs (for example, the turtles) is equally related to every archosaur. Monophyletic groups form clades and so can be depicted as single branches on the Tree of Life.

Just as a monophyletic group can be collapsed, it can be expanded as well. A tip that forms an exclusive group can become multiple tips if there are multiple exclusive groups inside it. Since the original tip was an exclusive group, doing this does not disturb the rest of the tree in any way. For example, in Figure 3a below, two groups of turtles are placed on the tree. These two groups do not exhaustively cover all turtles. Since turtles form an exclusive group, every non-turtle is equidistantly related to every turtle, whether a snapping turtle or a sea turtle – just as is true of the groups depicted in Figure 4a.

The logic of expansion dictates that only exclusive groups can be placed at the tips of trees. If a non-exclusive group such as the traditional Pisces (fish - including tuna and sharks) is placed on the tip of a tree as in Figure 3b, the tree will not properly represent evolutionary history. This is because the branch leading to this group cannot represent real historical lineages since there is no lineage that leads to both tuna and sharks without also leading to birds and a host of other groups. Placing Pisces at the tip implies that all non-fish, like turtles, are equally related to every fish. But this is not true; turtles (and all of the Sarcopterygii) are more closely related to tuna than they are to sharks. Just as "fish" is a paraphyletic group, "reptiles" (if it includes crocodiles and turtles but not birds) is too, so we cannot put reptiles at the tip of a tree unless we include birds in the reptiles.

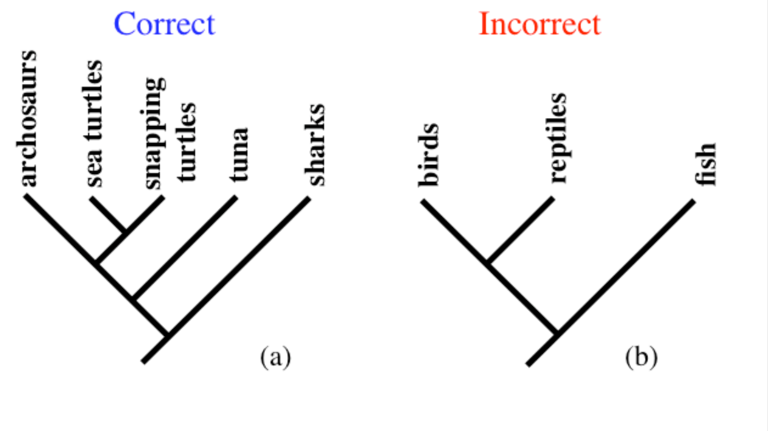


Figure 3 The tree on the left properly combines the birds and the crocodiles into the archosaurs while splitting the turtles into two (non-exhaustive) groups. The tree on the right attempts to place the paraphyletic groups "reptiles" and "fish" onto tips, but is an incorrect tree.

This means that classification does not dictate phylogeny. It may be that there are good reasons to continue to use Pisces and Reptilia as taxa in our classification system. But just because a group is united by a name doesn't imply that it is united in the right way genealogically. Therefore it doesn't follow that all taxa (in our classification system) can serve as taxa (the tips of a phylogenetic tree).

Conceptually, this is a simple point and follows from the idea that trees represent history. Whether we think classification must be based solely on history is a separate issue (Hull 1970). But it is easy to be misled, especially with common names like sharks. For instance, in his popular book *Why Evolution is True* (Coyne 2009), Jerry Coyne begins chapter 1, "What is Evolution" by discussing the bird/reptile history. But he places lizards at the tip of a tree as a branch separate from snakes. This is a mistake since the lizards do not form a monophyletic clade. The snakes are nested inside lizards.

One reason that this issue is so confusing is that it is clear that a tree structure corresponds to a hierarchical groups-within-groups structure. Taxonomists since Linnaeus (with some exceptions) have classified hierarchically this way. Species are within genera, genera within families, etc. with no overlap. Since a classification produces a hierarchical structure and an evolutionary tree produces a hierarchical structure, it is natural to think that there is just one structure. This is the thought of those who classify phylogenetically. On this view, all taxa are clades and the proper classification is determined by the phylogeny. Other views of classification such as evolutionary taxonomy say that while phylogeny is a part of classification, it does not determine it. A third view, phenetics, says that classification is completely independent of phylogeny (Hull 1970).

But what is clear on any view is that phylogenies are not dependent on classification. That is, we don't classify first and then build the phylogeny based on that. For example, one might think that paraphyletic groups are just fine as taxa. Then it might be permissible to continue to use Reptilia in its traditional sense as a taxon. If correct, this does not imply that we can build a correct tree with Reptilia at the tips. Such a tree would be an incorrect phylogeny regardless of our classification.

4. Classification does not dictate phylogeny

To see the importance of using only clades on trees, it is worth looking at one controversial example in classification. It shows that is not simply a convention that we read trees the way we do – rather, it is necessary if we want to use trees for inference problems that involve the genealogical history of various groups. Figure 4 depicts one reasonable phylogeny of a few of the major clades of the amniotes (the timetree is from Shedlock and Edwards 2009):

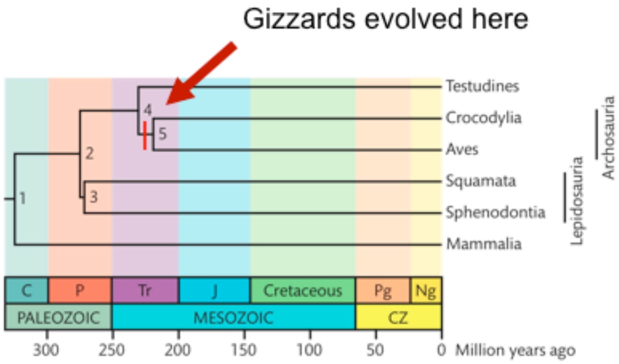


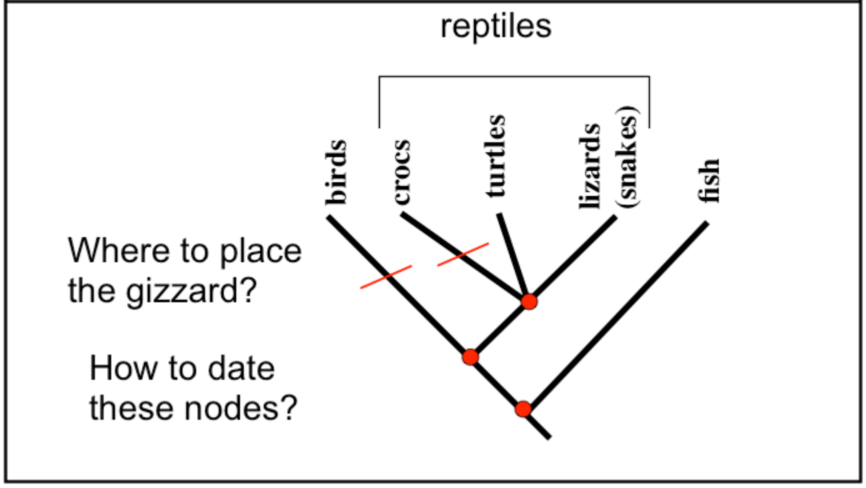
Fig. 2 A timetree of amniotes. The divergence times are from Table 1. *Abbreviations*: C (Carboniferous), CZ (Cenozoic), J (Jurassic), Ng (Neogene), P (Permian), Pg (Paleogene), and Tr (Triassic).

Here, each tip is an exclusive group each 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 croc (in Crocodylia), any turtle (in Testudines) and any lizard or snake (in Squamata). Here, the nodes represent real divergence points – birds and crocs share a common ancestor about 219 million years ago (mya), birds and turtles about 231 mya, and birds and lizards about 275 mya. The more traditional phylogenies will put turtles elsewhere such as sister to the rest of Sauropsida. Here the dating of the bird-turtle split might be 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 effect of trait evolution, it is essential that we know the underlying phylogeny.

For example, the gizzard is an organ that both birds and crocodiles have, but that no other living reptiles have. It 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 which can pass 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 as in Figure 4. That would mean that the gizzard evolved between 219 and 231 million years ago and has simply been inherited by both birds and crocodiles, both of which are descendants of this branch. Incidentally, this would mean that the dinosaurs also had gizzards as they are on the branch leading to the birds. Of course this is subject to error – for example, it is possible that the gizzard evolved earlier and it was simply lost in the branch leading to the turtles. Alternatively, it is possible that the gizzard has evolved twice. In fact, 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 the archosaurs. In fact, a detailed examination of the "gizzards" in the earthworms and the fish makes it clear that many of the similarities are only superficial and in an important sense, 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.

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 as in Figure 3b. As noted above, this is incorrect. It is easier to see this when we try to expand the reptiles into their traditional orders as though classification dictated phylogeny as in Figure 5.

Figure 5 If we use an incorrect phylogeny based on classification and then attempt to infer facts about evolutionary history from this tree, we will arrive at incorrect inferences. For example, if attempt to unite the reptiles in a single clade, it would appear that gizzards evolved twice independently and we would be wrong when attempting to date various nodes.



Here, assuming that the gizzard did evolve only once and spread to birds and crocodiles, that proper explanation is blocked since there is no branch on this diagram where the gizzard could have evolved. Instead, it looks like a massive coincidence that birds and crocodiles would 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. Similar problems arise when we attempt to infer the history of the endochondral bones that give the bony fish (like tuna) their name since they shared with the reptiles and birds but not the sharks.

The fact that this tree cannot correspond to genealogy is also easily seen by trying to date the nodes. Exactly when did birds split from reptiles? On this tree, it appears as though 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 of 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.

This misrepresentation 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,Y correlated), and for testing biogeographical hypotheses (did marsupials originate in Australia or North America). Essentially, 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, but only insofar as it is a good guide to the true genealogy – where the latter is what is actually required.

We now have arguments that it is necessary that any taxon, a group placed at the tip of a tree, must be a monophyletic clade. If we are going to ask questions such as "when did this group originate", "when did this group diverge from this other group", "has group X undergone anagenetic change faster than group Y has", etc, the groups in question must be clades or these questions will make no sense.

I said that there is no controversy that phylogenies should not be dependent on classification. While this statement would probably find near universal agreement, as a matter of fact, our current practices dictate that phylogenies <u>are</u> 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 by our classifications 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. I believe what I have said so far is relatively uncontroversial. However, as I will argue, principles like expansion and contraction, addition and subtraction actually conflict with other principles of trees – those involving species. For ease of communication, I will call these principles tree manipulation principles. Perhaps after further consideration, we will decide to give up certain manipulation principles like addition. But doing so would invite a host of problems that we should avoid. As we have just seen, 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 the principles that we have laid out so far to see where they lead us.

5. How we determine which taxa are closely related

I previously described taxa as being genealogically related to each other in an analogous way to how organisms are related to each other. But organisms have ancestor/descendent relationships. It isn't clear that taxa do. For instance, when I say that crocs are closer to birds than to snakes, how is this determined? It is naturally read as a claim that individual crocodiles are more closely related to individual birds than to individual snakes (Velasco 2008). But this places a severe restriction on what can count as a lineage split – complete and total cessation of reproductive connections. While such a system can be worked out (Velasco 2009), for most systematists, this is simply too strong.

Another way of understanding this genealogical claim about taxa is that there is a clade that contains crocs and birds but not snakes. But what is a clade? We said before it is some ancestor and all of its descendants. Traditionally, this means an ancestral species and all of the species descended from it. Alternatively, we could just say directly that it means that the most recent common ancestor of birds and crocs is a descendant of the most recent common ancestor of crocs and snakes. But what has ancestor/descendant relationships in this way? Again, the traditional answer is that this is all done in terms of species. Phylogenetic trees are ultimately species trees. If we put birds at the tip, it merely represents a collection of all the bird species.

That phylogenetic trees are really species trees is a common sentiment. As E.O. Wiley puts it in the introduction to his important book *Phylogenetics*, "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 is a tree of species is found throughout the literature (Cracraft and Donoghue 2004, Dawkins 2004, Hodkinson and Parnell 2006). Despite the rhetoric, I will show that in practice, phylogenetic trees are not species trees and in particular, thinking of trees in this way actually conflicts with the important tree manipulation principles and would serve as a faulty foundation for history-based phylogenetic inferences.

6. Population trees

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.

Phylogenetic trees that contain 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 just different populations of the same species usually delimited by geography. These trees are used for precisely the same purposes as species trees – as representing the history of groups of organisms in a way useful for historical inferences. In his book *Evolutionary Pathways in Nature* (Avise 2006), John Avise 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. While the majority of Avise's examples are at or above the species level as in the gizzard example, several of the examples involve subspecific groups. For example, Avise examines the evolution of the bold black and red plumage in the poisonous birds the Hooded and the Variable Pitohui (*Pitohui dichrous* and *P. kirhocephalus*) using data from Dumbacher and Fleisher (2001) (Avise 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 lacked the pattern making it quite likely that in this population the plumage pattern was secondarily gained as an instance of mimicry. It is worth noting that 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 or pure species tree, but is mixed.

And of course studies have been carried out closer to our taxonomic home. Morin et al. (1994), Kaessmann et al. (1999), Won and Hey (2003), and a host of others have studied 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 different population trees of humans. But he does not argue that these population trees are meaningless, but rather, they are meaningful and have genealogical implications – implications that Templeton has repeatedly tested and found wanting (Templeton 2002, 2005, 2007). Entire subdisciplines such as phylogeography, divergence population genetics, and coalescent methods broadly speaking 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 their foundations.

7. The problem with population trees

Despite the fact that biologists create and use population trees, there are reasons to worry about their interpretation. One reason is just the obvious fact that systematists often define phylogenies in terms of species. If we stick to that, then of course populations do not have phylogenetic relationships. Besides just blindly sticking to traditional definitions (which in any case are not all consistent with each other), 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.

The first thing to say about this claim is that in many cases, it is false. There are many currently recognized conspecific populations with no gene flow between them. Asexual species are the obvious groups to point to, and asexuality is the dominant form of life on the planet (Hull 1988: 429). But even 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) would 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 as well. But this lack of gene flow is consistent with conspecificity on other 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 so are conspecific on various ecological and morphological concepts as well. The important point here is that with many species concepts, population trees can be meaningful, even on the unnecessarily strict interpretation of requiring no gene flow between distinct lineages.

While the above point means that some population trees can be meaningful, in order to use them as the foundation for phylogenetics, one would want a much more robust defense. We want phylogenetic differentiation to 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 there certainly is something to be said for this clarity, 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 simply 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 via 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 all others and all are genetically distinguishable. Yet there is often 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 enquire about the history of these groups and ask questions about the origins and relations to other groups, phylogenies are required.

So if we deny that population trees make sense in the face of reticulation, we must also deny that many species trees are meaningful. Perhaps for some species there would simply 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 supposed solid foundation of phylogenetic trees at the higher taxonomic levels. When we reflect on the way that phylogenetic trees are used, it is clear that species trees can be meaningful and useful even in the face of reticulation. Thus we should accept that population trees are meaningful as well.

8. What are population trees?

An obvious problem that the defender of population trees must face is to interpret branching in a way consistent with reticulation. One way to do this would be parallel to defining species trees as the history of speciation – namely, we define populations and then use the definition to determine when one population has split into two and population trees can model this process. Defining population and population split is certainly not easy, though some have tried (Millstein forthcoming). But focusing on the definition of species and the necessity of carefully timing 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 wish to remain agnostic on the issue and instead again focus on what a phylogenetic tree could mean regardless of what we call the groups of organisms at the tips.

Our goal now is to understand the meaning of phylogeny without resorted to defining it in terms of species. One way to go is to define it in terms of population histories. If populations are defined historically, this could work. But a better way to think about phylogeny is to think directly about what the pattern of divergence between groups of organisms could represent. 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 it is exclusive groups that must go at the tips of trees. But what is an exclusive group? When dealing with populations as exclusive, it seems that we need them to be an exclusive group of organisms. If we define the genealogical relatedness of organisms in the obvious way using recency of common ancestry (Velasco 2009), we get a very stringent criterion whereby populations would almost always not count as exclusive since just a single migrant from another population who has offspring would collapse the two separate lineages. This is too extreme. On the other hand, we can define exclusivity in terms of gene trees. Here, 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 clade (Ané et al. 2007, Baum 2007). This view was designed as a way to talk about all clades of all sizes generally. We could use concordance factors to determine subspecies (Avise and Ball 1990) or to determine species (Baum 2009). But whether we wish to do this or not, 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 patters of lineage branching taking advantage of the depth of gene coalescence and not just the topology (Knowles and Maddison 2002, Edwards 2009). We can also 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. Perhaps entirely separate species. Perhaps 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.

9. 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. But this is precisely what I will argue. The key argument for this conclusion is that the truth of species phylogenies, that is, 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 the common case where two different ways of understanding the tree lead to different conclusions. In such cases, it is the population trees which phylogenetics should be concerned with.

It is clear that it is a mistake to place paraphyletic groups such as 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 a great many species are paraphyletic groups of populations. That is, 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. 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 quickly becomes a new species. Call this species B. Now species A (=A1+A2) is a paraphyletic group consisting of some, but not all of 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. Let’s say that after some time A1 and A2 have continued to diverge and are now separate species. If the species tree tracks speciation history, A1 will be more closely related to A2 and the tree will look like (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. This is what is meaningful for historical inferences. When they became distinct species according to some particular species concept is irrelevant.

That such examples of apparent inconsistency are common is not in dispute. What is in dispute is what to do about it. Given that fact that biologists use population trees and define phylogenetic trees as species trees, the obvious answer would appear to be species trees and population trees just represent different things and so they cannot genuinely be 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 of course a possible resolution to the problems, it comes at a high cost. Here I will outline just some of these costs and argue that the price to pay is too high.

10. Privileging species trees

In this section, I will argue against the view that species trees as defined by the history of speciation is fundamentally what phylogenies represent. One immediate 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.

Notice what follows. First, the failure of generalized manipulation principles. 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, both are false. Similar problems exist for addition and subtraction of branches. Of course 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 that because they simply mean different things, then we can still hold on to restricted forms of the manipulation principles.

Second, 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 non-existent or misleading nodes would lead to incorrect pictures of the history and lead us astray in inferring biogeography, or morphological or molecular histories.

Blocking the investigation of interaction between species and populations by using phylogenies has consequences even when there aren't any paraphyletic species involved. For example, recall that the inference about the history of the plumage in the Variable Pitohuis required the examining the relationships of some of the populations not just 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.

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 answer might be to say, "be more specific, do you mean how are the populations related? Or how are the species related?" 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 desirable compatibility situation, we can build a tree of groups and say that it doesn't matter whether the tips are all separate species or perhaps two tips are different populations of the same species. But now it would matter. For it is 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. But if it represents an ancestral population split, then apparently we are dealing with population trees and not species trees. We could imagine defining population trees in such a way as to allow species (and any other higher taxa as well) at the tips as long as we understand the branchings to be population level lineage branchings. But then how species are related to each other as part of a species tree could be inconsistent with how the very same groups are related to each other on a population tree. This is certain to cause immense confusion.

Since we can't mix populations and species, it now 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 wouldn't just have to settle on 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 they do agree. Perhaps this can be explained by some kind of more general agreement about what species are. I will examine one such proposal which will serve as an example of the problems with this view.

One common reason for thinking that species are special is that they are the units of evolution. This is certainly a popular view – for example, an anthology of important papers on species has been called *The Units of Evolution: Essays on the Nature of Species* (Ereshefsky 1992). Of course 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 that these evolutionary units uniquely pick out species.

It is unclear exactly what different authors mean by "evolutionary unit". Some authors argue that species play a theoretical 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 or not this is accepted, it is normally thought that to be a unit of evolution means that with respect to 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 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 with respect to 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 such that 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 good candidates for groups that are units with respect to evolutionary theory. Though it is doubtful that there is a single natural level in nature that represents the unit of evolution, if I were forced to pick such a level, I would think that the population is the natural level. Conspecific populations are not united with respect to 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 precisely 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 simply one level in a hierarchy – a level that is rather arbitrary. Sometimes authors with this position hold that it is reason to think that species don't exist. Another view is that while perhaps species do exist, their boundaries are arbitrary in a way. This view is not new – it goes back to Darwin (Darwin 1859, Mallett 2008a, 2008b, 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 cases 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 just the same way as if we artificially placed crocs, lizards, and turtles in a clade on a tree just 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 right now, that is not what systematists are interested in and that is not what currently used phylogenetic methods can recover.

The view that I am advocating might be thought of 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 nearly 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 define species. While opponents do not think that this is the correct species concept, they do not 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. Even if we agreed that species have unique ontological statuses 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 building phylogenetic trees and the use of trees requires that the phylogenetic relationships between groups of these different kinds be determined in the same way. Population trees and species trees must be compatible.

11. What about species?

So if species trees and population trees must be consistent, what does this mean for species? Harrison (1998) argues that there are good reasons to think that species as a whole have a life history and for much of its history, a species may be paraphyletic. 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 get rid of the traditional understanding of species trees. This view is consistent with what I have argued here. Defenders of a history-based version of the phylogenetic species concept simply do not accept that paraphyletic species exist (Mishler and Donoghue 1982, Mishler 1985, Donoghue 1985, Velasco 2008). However, if we continued to define species trees by speciation history, these species trees would not be exactly what these systematists want since instead of a paraphyletic species, we would have many new species coming into existence at the same time. This would lead to a polytomy on such a species tree. But what we want is to recover the more structured population history.

The solution 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 well 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 foundations of phylogeny. It may be that it is so important to recognize phylogenetic patters that our taxonomies should be phylogenetically based and species, as taxa, should be based on phylogeny as well (Baum 2009). But then again, species cannot serve as the foundation of phylogeny. The groups that do serve as foundational in either case are subspecific, historically united groups. Such groups are populations and thus it is appropriate to think of phylogenetic trees as population trees.

The fact that building phylogenies without the prior use of species is easily demonstrated to be possible – some systematists have already done so with a call to get rid of the species category all together (Pleijel 1999, Pleijel and Rouse 2000a and 200b, Fisher 2006, Fisher et al. 2007). While this position is of course consistent with my position here, it does not follow from it. 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 foundations of phylogenetics.

12. The power of semantics

That phylogenetic trees are species trees is now entrenched in the language not only because of higher systematics, but partly because of the enormous literature on the issue of how gene trees relate to species trees, often approached using population genetics methods. For a recent survey of this literature, see Edwards (2009).

Often, questions of definitions and semantics are simply ignored by biologists with the idea that biological practice will determine such issues, but in this survey piece, Scott Edwards, a researcher who works both above and below the species level, deals directly with these issues. It is worth quoting one section titled “What’s in a name” in full because it is useful to look not just at the conclusion, but 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 pg. 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 – here the splits are lineage branchings and not necessarily speciations as measured by some definition. When his colleague suggests the term “population tree” instead, it is clear that Edwards things that these are the same thing – that it is merely a question of what we ought to call these things that we are inferring, not that species trees and population trees are genuinely different and there is a genuine question about which of the two we are inferring. The primary reason to continue the using the name “species trees” is so that this important research does not get ignored in areas where it is needed.

In fact, 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 get rid of talk of species trees, we identify species trees with population trees, or whether they are distinguishable but always consistent, it is appropriate to think of phylogenetic trees as fundamentally population trees. Rather than being revolutionary, the idea that species trees are population trees is already deeply embedded as part of phylogenetic practice.

13. 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 structure, 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. Eventually 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 ecological, 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 through it, all of systematics, on this arbitrary and often times misleading system of ranking, phylogenetics would be a flawed science.

Luckily, 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 well 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 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. And 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 that each field has genuine contributions to make to the other (Avise 1989, Baum and Shaw 1995). And the contributions of each field are required to truly understand the goal of recovering the evolutionary history of life.

REFERENCES

Ané C., Larget B., Baum DA, Smith SD, Rokas A (2007). Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution*, **24**: 412–426.

Arnold ML (2009). *Reticulate Evolution in Humans: Origins and Ecology*. Oxford University Press, Oxford.

Avise, JC (1989). A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology and Evolution* **4**: 279-281.

Avise, JC (2006). *Evolutionary Pathways in Nature: A Phylogenetic Approach*. Cambridge University Press, Cambridge UK.

Avise, JC & Ball, RM (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**(2), 417–426.

Baum, D.A. (2009). Species as ranked taxa. *Systematic Biology* **58**(1), 74-86.

Baum, D. A. & Shaw, K. L. (1995). Genealogical perspectives on the species problem. In P. C. Hoch & A. G. Stephenson (Eds.), Experimental and Molecular Approaches to Plant Biosystematics (pp. 289-303). St. Louis, MO: Missouri Botanical Garden.

Carroll RL (1988). *Vertebrate Paleontology and Evolution*. W.H. Freeman, New York.

Cavalli-Sforza LP, Menozi P, and Piazza A (1994). *The History and Geography of Human Genes*. Princeton University Press, Princeton NJ.

Coyne, JA (2009). *Why Evolution is True*. Viking Penguin, New York, USA.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.

Cracraft, J. (1997). Species concepts in systematics and conservation biology — an ornithological viewpoint. Pp. 325-339 in Species The Units of Biodiversity (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.) Chapman and Hall, London

Cracraft, J, & Donoghue, M. J. (2004). Introduction. In J. Cracraft and M. J. Donoghue (Eds.), *Assembling the Tree of Life* (pp. 1-4). Oxford University Press, USA.

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. Houghton Mifflin Company, New York.

Donoghue, M. J. (1985). A critique of the biological species concept and recommendations for a phylogenetic alternative. The Bryologist, **88**(3), 172-181.

Dumbacher JP and Fleisher RC (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.

Dupré, J. (1999). On the impossibility of a monistic account of species. In R. A. Wilson (Ed.), Species: New Interdisciplinary Essays (pp. 3-29). 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.

Eldredge, N. & Cracraft, J. (1980). Phylogenetic patterns and the evolutionary process : Method and theory in comparative biology. New York: Columbia University Press.

Edwards, SV (2009). Is a new and general theory of molecular systematics emerging? *Evolution* **63**(1): 1-19.

Ereshefsky, M. (1991). Species, higher taxa, and the units of evolution, *Philosophy of Science,* **58**, 84-101.

Ereshefsky, M. (1992). *The Units of Evolution: Essays on the Nature of Species*. MIT Press, Cambridge, Massachusetts.

Ereshefsky M. (2009). Darwin's solution to the species problem. *Synthese*. doi: 10.1007/s11229-009-9538-4.

Fisher, K. (2006). "Rank-Free Monography: A Practical Example from the Moss Clade *Leucophanella* (Calymperaceae)," *Systematic Biology* 31: 13-30.

Fisher K.M., Wall D.P., Yip K.L., Mishler B.D. (2007). Phylogeny of the Calymperaceae with a rank-free systematic treatment. *Bryologist*. **110**:46–73.

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 D. Otte & J. A. Endler, (Eds). *Speciation and its Consequences*, (pp. 19-31). Sunderland, England: Sinauer Associates.

Hey, J., Waples, RS, Arnold ML, Butlin RK, Harrison RG (2003). Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology & Evolution*, **18**: 597–603.

Hodkinson TR and Parnell JAN (2006). *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*. CRC Press Boca Raton, FL.

Hull, DL (1970). Contemporary systematic philosophies. *Annual Review of Ecology and Systematics*, **1**: 19-54.

Hull DL (1980). Individuality and selection. *Annual Review of Ecology and Systematics*, **11**, 311-32.

Hull DL (1988). *Science as a Process*. Chicago: Chicago University Press.

Kaessmann H, Wiebe V, and Paabo S (1999). Extensive nuclear DNA sequence diversity among chimpanzees. *Science* **286**: 1159-62.

Knowles LL and Maddison WP (2002). Statistical Phylogeography. *Molecular Ecology* **11**: 2623-35.

Kornet DJ (1993). Permanent splits as speciation events: a formal reconstruction of the internodal species concept. *Journal of Theoretical Biology*, **164**: 407-435.

LaPorte, J. (2005). Is there a single objective, evolutionary tree of life? *The Journal of Philosophy*, **102**(7), 357-374.

Laurin M and Gauthier JA (1996). Amniota. Mammals, reptiles (turtles, lizards, Sphenodon, crocodiles, birds) and their extinct relatives. Version 01 January 1996. http://tolweb.org/Amniota/14990/1996.01.01 in The Tree of Life Web Project, http://tolweb.org/

Mallet J (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**: 229–237.

Mallet J, Beltrán M, Neukirchen W, Linares M (2007). Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*. **7:** 28.

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.

Mallet J (2008b). Mayr's view of Darwin: was Darwin wrong about speciation? *Biological Journal of the Linnean Society*

Mayden, R. L. (1997). A hierarchy of species concepts: The denouement in the saga of the species problem. In M. F. Claridge, H. A. Dawah, & M. R. Wilson (Eds.), *Species: The Units of Biodiversity* (pp. 381-424). London: Chapman and Hall.

Mayr, E (1997). The objects of selection. *Proceedings of the National Academy of Sciences*, **94**: 2091-4.

Millstein, RL (forthcoming). Populations as Individuals. *Biological Theory: Integrating Development, Evolution and Cognition. Special edition on The Edges and Boundaries of Biological Objects*, guest edited by M. Haber and J. Odenbaugh.

Mishler, B. D. (1985). The morphological, developmental, and phylogenetic basis of species concepts in bryophytes. *The Bryologist*, **88**(3), 207-214.

Mishler, B. D. (1999). Getting rid of species, in R.A. Wilson (ed.), Species: New Interdisciplinary Essays. Cambridge, MA: MIT Press.

Mishler, B. D., & Brandon, R. N. (1987). Individuality, pluralism, and the phylogenetic species concept. Biology and Philosophy, 2, 397-414.

Mishler, B. D., & Donoghue, M. J. (1982). Species concepts: A case for pluralism. Systematic Zoology, 31(4), 491-503.

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, and Woodruff DS (1994). Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265, 1193-201.

Moritz CC (1994). Defining "evolutionary significant units" for conservation: A critical review. *Molecular Ecology* **3**: 401-411.

Pamilo, P and Nei M (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.

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.

Pleijel, F. and G. Rouse (2000b). "A new taxon, *capricornia* (*Hesiondae*, *Polychaeta*), illustrating the LITU ('Least-Inclusive Taxonomic Unit') concept." *Zoologica Scripta* 29: 157-168.

Ridley M. 1989. The cladistic solution to the species problem. *Biology and Philosophy*. 4:1–16.

Ryder OA (1986). Species conservation and systematics and the dilemma of subspecies. *Trends in Ecology and Evolution* **1**: 9-10.

Shedlock AM and Edwards SV (2009). Amniotes (Amniota). Pp 375-379 in *The Timetree of Life*, SB Hedges and S Kumar, Eds. Oxford University Press 2009.

Simpson GG (1961). *Principles of Animal Taxonomy*. New York: Columbia University Press.

Steel MA (2007). Tools to construct and study big trees: A mathematical perspective. Pp 97-112 in *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*, TR Hodkinson and JAN Parnell Eds. CRC Press Boca Raton, FL.

Takahata, N (1989). Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* **122**:957–966.

Templeton AR (2002). Out of Africa again and again. Nature 416, 45-51.

Templeton AR (2007). Haplotype trees and modern human origins. Yearbook of Physical Anthropology 48, 33-59.

Templeton AR (2007). Genetics and recent human evolution. Evolution 61, 1507-19.

Van Valen, L. (1976). Ecological species, multispecies, and oaks. Taxon, 25(2/3), 233-239.

Velasco, JD (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.

Velasco, JD (2009). When monophyly is not enough: Exclusivity as the key to defining a phylogenetic species concept. *Biology and Philosophy* **24**: 473-486.

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.

Wilson, R. A. (2005). *Genes and the Agents of Life: The Individual in the Fragile Sciences*: Biology. New York, NY: Cambridge University Press.

Won Y-J and Hey J (2005). Divergence population genetics of chimpanzees. Molecular Biology and Evolution 22, 297-307.