The motivation for the development of phylogenetic nomenclature (originally called “phylogenetic taxonomy”) was to allow biological classification (or “systematization”), to represent phylogenetic relationships, and to embody important principles such as “the untenability of paraphyletic groups” (de Queiroz and Gauthier 1990). From this starting point, de Queiroz and Gauthier developed a creative new basis for systematization in which the entities are not ranked taxa but clades (de Queiroz and Gauthier 1990, 1992, 1994; de Queiroz 1992, 1994). Instead of attaching names to taxa by reference to a type and a rank, as in traditional biological nomenclature, phylogenetic nomenclature labels taxa by the use of multiple “specifiers”—specimens or apomorphetic traits that unambiguously refer to a particular monophyletic taxon. The subsequent development of phylogenetic nomenclature included formation of a scholarly society, the International Society for Phylogenetic Nomenclature, and development of a formal Code, the PhyloCode (http://www.phylocode.org), that is now in final preparation for publication at University of California Press. The hope for many phylogenetic systematists was for a new system of nomenclature that would firmly connect taxonomy to phylogeny and would allow for more stable ways to name the many significant clades that make up the Tree of Life. However, a major challenge for the development of the PhyloCode has been the treatment of species. We argue in this paper that the rank of species (Ereshefsky 2010; Mishler 2010). Under the Darwinian approach, particular species may be meaningful evolutionary groups of organisms, just like particular genera, sections, subspecies, and varieties, but the species rank is not qualitatively distinct from other ranks. The Modern Synthesis reversed Darwin’s advance, attributing renewed significance to the species rank and presenting species as the unit of evolution (see Mishler 2010 for more discussion of this checkered history).

Partly in reaction to the preeminent position given to species during the Modern Synthesis, the past few decades have seen the proliferation of species concepts (Mayden 1997, listed at least 24 different concepts). Although the development of these concepts has helped clarify alternative ways of prioritizing biological concepts and phenomena, there has certainly been no clear winner: We are no closer to finding a single universally accepted definition of species than we were in Darwin’s time. Fortunately, the challenge of designing a system of phylogenetic nomenclature can (and should) be separated from the species concept debate.

The approach to species currently adopted in the PhyloCode was presented and defended by Dayrat et al. (2008). They stressed the incompatibility of Linnaean binomial nomenclature with phylogenetic nomenclature and concluded that species and clades are two different kinds of entities and that the former should not be accommodated within the PhyloCode system of rank-free nomenclature. In taking the stance that species are distinct from clades, Dayrat et al. (2008) adopted the particular conception of species proposed by de Queiroz (2007). According to this view, species are separately evolving but connected subpopulations (metapopulations), and as such species are not clades but rather segments of lineages. Although de Queiroz’s lineage-based species concept has been criticized (e.g., Baum 2009; Hausdorf 2011; see also the critique below), the pros or cons of his species concept are not the most pressing issue here. Accepting that lineages are not the same
entities as clades, and that the PhyloCode is about naming clades, our main disagreement with Dayrat et al. (2008) is that they still retain species in the PhyloCode. In doing so, they effectively impose on all would-be users of the PhyloCode their view that species are lineages. We believe, in contrast, that the only way to improve the PhyloCode, so that it can be used by anyone who wishes to name clades, regardless of what they think about the nature of species, is to remove all mention of species and to treat all clades, from the very smallest to the very largest, equally. This will allow systematists freedom to equate species with clades, or with lineages, or to deny the existence of species entirely.

de Queiroz’s (2007) views on the nature of species have unfortunately (and unnecessarily) played a tremendous role in shaping the PhyloCode resulting in a Code that is not useable by the many individuals who are fully committed to the principles of phylogenetic nomenclature but do not accept this particular view of species. A system of nomenclature should not be tied to a particular philosophical outlook on something as controversial as the nature of species, especially a system focused on the naming of clades.

It is not our intent here to review the history of the debate over species and phylogeny or to anoint any particular species concept as the correct one. Instead, in this paper, we propose that the PhyloCode be amended to remove mention of “species” and to accommodate users who wish to be able to attach names to clades that approximate taxa at the traditional species level. Before doing so, we provide an overview of those principles that guide the system of phylogenetic nomenclature in general that are most relevant to the question of whether and how species names might be accommodated and then a brief history of how species have been treated in this system so far.

BACKGROUND

Phylogenetic Systematics and Nomenclature

The motivation for phylogenetic nomenclature was to facilitate the precise naming of any monophyletic group in such a way that nomenclature can remain stable as long as knowledge of phylogenetic relationships remains stable. For the purposes of this discussion on species, it is important to clarify the concept of monophyly since “ancestral species” are mentioned in some definitions. The distinction between diachronic lineages and synchronic clades is illustrated in Fig. 1. Lineages are relationships through time between ancestors and descendants, whereas clades are composed of sets of tips that are in existence at any one time. A simple definition of monophyly, and hence clades, is: “a monophyletic group is all and only the descendants of a common ancestor” (Mishler 2010). The ancestor in this definition is not a species but rather a part of a lineage, such as an organism, kin group, or population (as discussed by Mishler and Brandon 1987). Also, as pointed out by Baum (2009), monophyly of a group does not (and can not) mean monophyly on every single gene tree—horizontal transfer and incomplete lineage sorting are frequent enough to ensure this. Instead, monophyly refers to an ensemble characteristic of organismic and/or genic descent.

Traditional nomenclature has limitations as applied to the modern understanding of the Tree of Life. First, the Tree of Life is a deeply nested hierarchy of monophyletic groups. There are far too many levels in this hierarchy for the number of ranks available in the traditional Codes of nomenclature. The rules in the traditional Codes that do not allow taxa to be nested inside taxa at the same rank lead to frustrating instability in names as our knowledge of phylogenetic relationships progresses. Second, formulating a taxon concept around a single type does not allow precise reference of a monophyletic group because it is not clear exactly what node on the tree is being named. As when triangulating in navigation, it is necessary to refer to at least two points of reference (“specifiers” in the PhyloCode) to unambiguously attach a name to a clade. Third, the traditional Codes are rank-based, but taxa at the same rank have nothing necessarily in common and are not comparable. Since ranking is highly subjective, whereas assignment of organisms to clades is largely objective (if not always easy), it seems prudent to build a nomenclatural system around clade affiliation rather than rank assignment. Clades exist regardless of rank, and the addition of rank to a clade, for example, whether they are referred to as genera or families, does not add any value to our nomenclatural system. For all these reasons and more, there has been a movement to develop a new code of rank-free phylogenetic nomenclature, the PhyloCode.

The development of a new nomenclature for species names has been almost as controversial as the species problem itself. A major point of contention between the traditional Codes and the PhyloCode is the lack of
mandatory ranks in the latter. As a result, Linnaean species binomials are incompatible with phylogenetic taxonomy because they naturally imply the existence of both a genus rank and a species rank.

A number of initial proposals on how to treat species names in the light of phylogenetic taxonomy have been presented in the literature (de Queiroz and Gauthier 1992; Graybeal 1995; Schander and Thollesson 1995; Cantino 1998; Cantino et al. 1999; Ereshefsky 1999; Mishler 1999; Pleijel 1999; Pleijel and Rouse 2000, 2003; Artois 2001; Hillis et al. 2001; Dayrat et al. 2004; Dayrat 2005; Wolsan 2007; Baum 2009). These approaches worry primarily about how to convert traditional binomials into rank-free names that could be used in phylogenetic nomenclature. Sundberg and Pleijel (1994) proposed retaining the Linnaean species binomials for clades, with the traditional genus name serving as a prenomen: not part of the species name, but added by convention in order to retain historical name legacy. Cantino (1998) proposed, instead, the hyphenation of genus and epithet, thereby conserving the current generic name and including it as a part of the species name. However, this option was criticized because it would naturally lead people to consider species with the same prenomen as being closely related even though they might not be (Schander and Thollesson 1995). From that position, the combination of a higher taxon name (serving as a clade address) with a species uninomial would be preferred (Graybeal 1995; Schander and Thollesson 1995; Mishler 1999; Pleijel 1999). Although unimomials composed of species epithets have many virtues, they pose significant problems regarding their uniqueness within and across biological domains. A series of expedients to avoid name duplication were explored, using a combination of names followed by sequential numbers, registration numbers, etc. (see a full list of proposals in Cantino et al. 1999; Artois 2001; Wolsan 2007). Dayrat et al. (2004) proposed to add three levels of uniqueness to species uninomials: (1) the traditional species epithet, (2) the original author’s name, and (3) the date of publication. In cases where the three levels would still not provide uniqueness, page numbers and letters (a, b) could be added to the name string. Through such measures, it was shown to be possible to develop a clade-based naming system that encompassed taxa at or below the traditionally recognized species rank (Dayrat and Gosliner 2005).

In addition to worrying about the form of species names, some authors considered ways to make it clear from the name itself whether a certain taxon is at the species rank or not. Graybeal (1995) suggested the use of different kinds of parentheses to indicate species uninomials. Pleijel and Rouse (2000) recommended the use of an initial upper case letter for clade names that do not refer to least-inclusive taxonomic units (LITUs), and lower case for uninomials that do refer to LITUs. However, no agreement has been achieved as to the form of species names, mainly because these nomenclatural issues tended to be confounded with philosophical positions on the nature of species, in particular whether species are clades or some other kind of biological entity and, if they are clades, whether the species rank is special.

It is worth highlighting that, regardless of one’s view of species ontology, in many empirical studies that sample a large number of specimens, the set of organisms that were traditionally considered to comprise a single species are often found to approximately correspond to a clade. The same is of course true for sets of organisms that have been traditionally ranked as genera, families, etc. and are often found to correspond to clades. This is one reason why Baum (2009) argued that the PhyloCode ought to at least allow systematists to name clades that happen to approximately correspond in composition to a traditional species.

A Brief History of the PhyloCode with Reference to Species

The first original draft of the PhyloCode was sketched at a workshop held at Harvard University in 1998 and was modified and subsequently made public in April 2000 via the Internet at www.phylocode.org. This draft covered the general application of clade naming rules that revolved around the philosophical principles of phylogenetic nomenclature. The publication of this early draft generated a large body of literature, some of which was supportive (Bremer 2000; Brecher and Sumrall 2001; Ereshefsky 2001; Laurin 2001, 2005; Bertrand and Pleijel 2003; Pleijel and Rouse 2003; Lee 2005 and many others) and some critical (Lucas 1992; Lidén and Oxelman 1996; Lidén et al. 1997; Dominguez and Wheeler 1997; Moore 1998; Benton 2000; Nixon and Carpenter 2000; Nixon et al. 2003, among others).

In 2002, a second workshop held at Yale University focused more on species names and the best approach to handle their nomenclature. Much disagreement characterized the debate and the issue was left open for further discussion, although the clear majority felt that a separate Code covering species would ultimately be necessary.

The First International Phylogenetic Nomenclature Meeting was held in Paris in July 2004. Further discussion about species and their names took place among the over 70 attendees. Different approaches were proposed, including two by Dayrat and Clarke (ironically, both authors in Dayrat et al. 2008) who argued for a uninominal system within the PhyloCode. Under both proposals species were treated differently to clades: the species names being defined using the formula “the species that includes specimen X,” with the author clarifying the kind of species entity to which the name refers. It was then agreed that Dayrat and Clarke should form a committee with the specific charge of drafting a Code for species along these lines.

The attempt to develop a phylogenetic code of nomenclature for species faced major challenges and ultimately failed due to the difficulty of reconciling several important issues including: (1) the realization that species defined under the traditional Codes vary widely.
in concept, (2) difference between ICZN and ICBN in how they handle species, (3) the conviction that every species name established under the current Codes would have to be converted into a PhyloCode name, and (4) the realization that lineages are nested inside of each other in complicated patterns meaning that if species are to be defined as lineages, they will nest at several hierarchical levels. Given these issues, the committee proposed at the Second International Phylogenetic Nomenclature Meeting held at Yale University in 2006 that species names continue to be established via the traditional Codes (ICBN, ICZN, and ICNB) and that subspecific groups be viewed as kinds of species. This proposal was favored by a clear majority of the meeting attendees, but with several nays (including the two authors of this paper who attended the meeting). This approach was then adopted by the committee that oversees the PhyloCode as described positively by Dayrat et al. (2008) in their summary of the recommendations: 

On May 24, 2007, the Committee on Phylogenetic Nomenclature (CPN) [. . .] adopted a new article in the PhyloCode addressing the naming of species in the context of phylogenetic nomenclature. This vote [. . .] represents a major step in the development of the PhyloCode (Dayrat et al. 2008). This statement was meant by its authors to portray the ultimate achievement of the PhyloCode, but we feel that this vote represents instead a conclusive defeat for phylogenetic nomenclature. After years of debates in the literature, and at meetings and workshops, the difficulty of dealing with naming species in a phylogenetic context culminated with a surrender of species names to the traditional Codes.

Several revisions to the PhyloCode were made based on Dayrat et al. (2008). Notably, in the current PhyloCode draft (version 4c), art. 21 states that This code does not govern the establishment or precedence of species names. To be considered available (ICZN) or validly published (ICBN and ICNB), a species name must satisfy the provisions of the appropriate rank-based code (e.g., ICNB, ICBN, and ICZN) This article describes how species names governed by the rank-based codes are to be interpreted and used under this code. Although acknowledging the value of democratic decision making, we remain deeply disappointed that the difficulty of dealing with naming species in a phylogenetic context culminated with a surrender not just of species taxa (which we are happy to let go) but species names as well. We here wish to argue that this poor decision should be revisited and reversed.

**The General Lineage Species Concept and the Problems It Poses for the PhyloCode**

The de Queiroz (2007) approach to species is implicitly enshrined in the current version of the PhyloCode, which clearly states that clades and species do not equate (Art. 1.1) and implies that species are lineages (Rec. 21.4C). Many different species concepts are currently applied across the different biological domains, and many scientists find some of these useful whether conceptually, practically, or both. Many practitioners believe that species are ranked or unranked clades; many do not. But this disagreement is not relevant to a code of nomenclature that is rank agnostic. Rather than trying to attain a unified species concept, a sort of biological holy grail, we may have to just accept and embrace the diversity of thoughts about species and agree that phylogenetic nomenclature should remain logically independent from the philosophical debate about species. It is not constructive to assert an “official” species concept in the PhyloCode, which is only a nomenclatural tool for naming clades not a ground for reconciliation.

We understand de Queiroz’s position on species as segments of lineages and we agree with him that these are not the same kind of entity as clades. Lineages are diachronic, whereas clades are synchronic. Many authors since Hennig have appreciated this distinction, although it is sometimes muddled. Our disagreement with de Queiroz boils down to these points: (1) many currently named species are certainly not lineages. We have evidence for only a small number of them being actual lineages (it is empirically difficult to demarcate lineages as compared to the well-worked out evidential procedures for finding clades); (2) even when we can detect lineages, it is clear that nested levels of lineages can be discovered and it is unclear how one privileged level should be recognized as species. In other words, the species-ranking problem is in no way solved by equating species with lineages (regardless of exactly how lineages are defined). However, this disagreement should be irrelevant to the PhyloCode. The PhyloCode governs only clades. Other important biological entities such as lineages, genes, trophic levels must be named in other ways preferably in such a way that they are clearly distinguished from clade names (even in cases where they have the same nomenclatural root).

The PhyloCode as currently constructed works under the assumption that species are not clades, yet paradoxically brings species into the Code by stipulating that species names are disallowed for clades (Dayrat et al. 2008). When a clade happens to approximate a traditional species in content, the PhyloCode mandates that clade be given a new name, distinct from the traditional species name. However, when a clade happens to approximate a traditional genus or family (or any higher rank), it strongly recommends retaining and converting the traditional name whenever possible. This constraint is incompatible with several reasonable positions one might take on the nature of species and on the goals of phylogenetic nomenclature.

Some reasonable biologists believe in the existence of clades and doubt that species are meaningful real entities that warrant inclusion in a formal taxonomic system. For these individuals, the current PhyloCode poses severe problems because it inhibits their ability to name clades at or below the traditional species level. Likewise, a biologist who accepts that species are distinct from clades but who believes that a certain traditional species or subspecies name is best viewed as referring to a clade cannot apply that name to that clade.
Even more so, the *PhyloCode* discriminates against individuals who believe that species correspond to clades of some (perhaps arbitrarily determined) rank. These individuals not only have trouble naming clades at or below the species level but, when they discover a clade that closely approximates a named species, they are forbidden from giving that clade its logical name—the relevant species epithet.

The current *PhyloCode* requires systematists to refer to species and subspecific taxa whose names are regulated by the traditional codes. The Code allows for no other way to utilize such taxa. Furthermore, the *PhyloCode* also requires systematists to refer to *PhyloCode* and may not be given clade discriminates against in-Linnaean binomials. Similarly, a nested taxon *S* should be museum ought to be legally independent of the existing Codes (although some similarities in spirit are common to all, of course). In other words, no rules in the *PhyloCode* should depend on rules in the traditional Codes, including for species names. However, the *PhyloCode* as it stands does depend on traditional Codes, and these are not controlled by proponents of phylogenetic nomenclature. For example, ICBN could amend their Code to say that species names are invalid if, when published, there was intent to treat the genus name as a “prenomen” under the *PhyloCode*.

(2) The current Code calls on systematists to name new species under the traditional Codes even if they intend only to use the names as specifiers within phylogenetic nomenclature (Art. 21.4). We feel that it is intellectually dishonest to assign a genus name to a new species (as required under the traditional codes) when in fact the “genus” is believed to be an artificial concept and its use would be later dropped by the author of the name. Even more dramatic is the case where a new species does not fall under any of the traditionally described genera, implying that the author should first establish a new genus under the traditional Codes before creating a new species, also subjected to traditional rules.

(3) Species established under the traditional Codes and clades named under the *PhyloCode* are different entities, hard to precisely compare because they are fundamentally defined by a single-type and multiple specifiers, respectively. Therefore, adopting traditional binomials to label clades and their specifiers generates natural inconsistencies between clades and what traditional names actually represent. This is a fundamental reason why specifiers under the *PhyloCode* should be museum specimens (physical reference objects) at all levels, never Linnaean binomials. Similarly, a nested taxon can be equally used as a specifier of a larger clade, as long as the reference taxon has been previously defined according to the rules of the *PhyloCode* with reference to specimens.

(4) In many cases, biologists discover clades with finer phylogenetic structure within a currently terminal taxon that approximately compare to species as circumscribed based on one of the many species concepts. In these instances, it seems logical to retain a clade reference to the original species name and allow that it contains smaller subclades requiring new names. This contrasts with what happens when a species is split under the traditional Codes, where the species name must be transferred down to one of the newly discovered subclades (the one containing the type), resulting in serious difficulties keeping track of synonymy in databases. Therefore, treating names of current terminal taxa just like clade names at all levels has a decided benefit for stability and biodiversity informatics. Yet this is banned under the current *PhyloCode* by rules that disallow from assigning any species name to a clade.

**A Solution to Naming Species in Phylogenetic Nomenclature**

We suggest that the rules governing names of clades at and around the traditional species level should be identical to the rules that apply at higher levels. As there are no ranks under the *PhyloCode*, it seems illogical to have an implicit rank of species. However, Dayrat et al. (2008) define a position that is laid out in the current version of the *PhyloCode*, which effectively applies special rules at the “species” level in the *PhyloCode* hierarchy. Such an implicit ranking exists due to articles in the *PhyloCode* that imply or overtly specify that clades at or below the level of species are not governed by the *PhyloCode* and may not be given clade names. For example, it is allowed that the specifiers of clade names be species rather than particular type specimens. Because it is wrong to use species as specifiers for clades that approximate to traditional species or subclades thereof, this approach implies that clades cease to exist or cease to be nameable at the approximate level of species.

Furthermore, Article 10.9 of the *PhyloCode* states: “A clade name may not be converted from a preexisting specific or infraspecific epithet (ICBN and ICNB) or from a name in the species group (ICZN).” This implies that if we discover a clade that happens to conform in content to a traditional species, and we have developed a suitable definition for that clade, we are nonetheless forbidden from using the traditional species name for that clade. Such a restriction does not apply to genus-approximating clades, showing that the *PhyloCode* applies different nomenclatural rules to clades whose content approximates species than to other hierarchical levels.

We believe that in order to maintain continuity with past literature, systematists should be allowed to
recognize clades that roughly correspond in content to traditional species (or subspecies) and to use their former epithets as uninnominal clade names. We certainly do not advocate mass conversion of all existing species names established under the traditional Codes to clade names, as many of these in fact, do not even correspond to clades. Furthermore, just as for higher clades, only relevant species names should be converted into clades names at the discretion of expert systematists. Fisher (2006) provides an example of a monograph treating clades that approximate species using rigorously defined uninnominal clade names that would be invalid under the current draft of the PhyloCode.

We are fully aware that many epithets are very common within and between domains. However, homonyms can be fully resolved by the addition of references to the name string, for example, clade name author, year, registration number, and/or clade address, following up the strategy articulated by Dayrat et al. (2004) but applying it to all levels instead of only species. It is already determined that valid clade names will be assigned registration numbers by RegNum (Cellinese 2012), an on-line database, currently in its final stage of development that stores and manages phylogenetic definitions. This mechanism will readily allow users to redeploy a former species (or subspecific) epithet to clade without leading to ambiguity due to homonyms. For example, the epithet digitata is common to 299 plant species (IPNI). However, Adansonia digitata L. could be converted to Digitata Linnaeus 1759, 2:1144 [New Author] (Adansonia) RegNum-number, whereas Ceinfugosia digitata Cav. would convert to Digitata Cavanilles 1787, 3:174 [New Author] (Ceinfugosia) RegNum-number. We believe that naming the clade (Adansonia) Digitata need not prevent individuals recognizing the traditional species Adansonia digitata L. Even if the composition of these two taxa were to diverge over time, for example, because the species A. digitata came to be divided into two while the clade continued to refer to the larger grouping, we are confident that systematists would be able to communicate without confusion. After all, several competing ideas on species limits may coexist in the current system, yet systematists are quite able to indicate the taxon of interest in a particular case using conventions such as sensu. Although our proposal would result in a name string containing a few more components than traditional species names, these measures would largely remain behind the scenes in databases, ensuring the required uniqueness of clade names, and providing a stable link to important legacy data.

CONCLUSIONS

We do not agree with the assessment of Dayrat et al. (2008) that the PhyloCode’s development is complete. Much progress still needs to be made before it is published. In particular, we cannot support the adopted approach for “resolving” the species problem by tying the PhyloCode to Linnaean binomials named under the traditional Codes. This approach, if enforced, will result in many previous supporters of Phylogenetic Nomenclature refraining from using the PhyloCode. The PhyloCode needs to be extended to cover all clades, including those that approximate to or are within traditional species and should be made philosophically inclusive so that it can be adopted by anyone regardless of their preferred species concept.

If monophyly provides the basis for classification, then there is no reason why a clade-naming system should not be extended to cover all clades even those that may have similar content to someone’s concept and/or rank of species. Therefore, it seems clear to us that phylogenetic nomenclature must be able to handle clades at all levels and, for reasons of clear communication and fair reference to the legacy literature, that one ought to be able to assign them the traditional names (or some derivative thereof).

The species problem does not need to be resolved for the purposes of finalizing the PhyloCode. No species concept (including the ones advocated by the authors of the present paper) should be enshrined in the PhyloCode. Since its purpose is to name clades, any mention of “species” in the PhyloCode should be removed. We need to streamline the PhyloCode to focus solely on rational rules for naming clades at any level including the traditional species level.

The authors of this paper have submitted the needed modifications to the PhyloCode in a proposal to the Committee of Phylogenetic Nomenclature (Appendix 1, Dryad doi:10.5061/dryad.sr2hd7md) and sincerely hope that they, or some equivalent set of changes, will be adopted before the official publication of the PhyloCode. We hope that this paper will stimulate discussion and result in a more open-minded Code that can be applied to all clades in the Tree of Life.

SUPPLEMENTARY MATERIAL

Supplementary material including an online-only appendix can be found in the Dryad data repository at http://datadryad.org. doi: doi:10.5061/dryad.sr2hd7md.

FUNDING

N.C. is very grateful for support from the National Science Foundation (DEB-0953677) for work that fostered these ideas and D.B. acknowledges the National Science Foundation (DEB-0949121).

ACKNOWLEDGMENTS

We would like to thank Torsten Eriksson for feedback on an early draft of the manuscript, and Dan Faith, Matt Haber, Dick Olmstead, and an anonymous reviewer for constructive criticisms and valuable suggestions that helped improve our paper.