

March 14, 2018. **Classification II: Phylogenetic taxonomy including incorporation of fossils; *PhyloCode***

I. Phylogenetic Taxonomy - the argument for rank-free classification

A number of recent calls have been made for the reformation of the Linnaean hierarchy (e.g., De Queiroz & Gauthier, 1992). These authors have emphasized that the existing system is based in a non-evolutionary world-view; the roots of the Linnaean hierarchy are in a specially-created world-view. Perhaps the idea of fixed, comparable ranks made some sense under that view, but under an evolutionary world view they don't make sense. There are several problems with the current nomenclatorial system:

1. The current system, with its single type for a name, cannot be used to precisely name a clade. E.g., you may name a family based on a certain type specimen, and even if you were clear about what node you meant to name in your original publication, the exact phylogenetic application of your name would not be clear subsequently, after new clades are added.
2. There are not nearly enough ranks to name the thousands of levels of monophyletic groups in the tree of life. Therefore people are increasingly using informal rank-free names for higher-level nodes, but without any clear, formal specification of what clade is meant.
3. Most aspects of the current code, including priority, revolve around the ranks, which leads to instability of usage. For example, when a change in relationships is discovered, several names often need to be changed to adjust, including those of groups whose circumscription has not changed. E.g., when it was detected that the Cactaceae is nested inside of the Portulacaceae, one of these well-known family names has to be abandoned. Frivolous changes in names often occur when authors merely change the rank of a group without any change in postulated relationships.
4. While practicing systematists know that groups given the same rank across biology are not comparable in any way (i.e., in age, size, amount of divergence, diversity within, etc.), many users do not know this. For example, ecologists and macroevolutionists often count numbers of taxa at a particular rank as an erroneous measure of "biodiversity." The non-equivalence of ranks means that at best (to those who are knowledgeable) they are a meaningless formality and perhaps not more than a hindrance. At worst, in the hands of a user of classifications who naively assumes groups at the same rank are comparable in some way, formal ranks lead to bad science -- removing the ranks would serve the same purpose as child-proof door locks for the back seat of your car!

II. The *PhyloCode*

How exactly to write the new phylogenetic code of nomenclature is still a bit controversial at this point, especially as applied to species (see draft of the *PhyloCode* and other materials at: <http://www.ohio.edu/phylocode/index.html>), but the basics are clear. Such a new code should maintain the principle of priority (the first name for a lineage should be followed)

and other aspects of the current code that promote effective communication of new names to the community.

However, two major changes would be made:

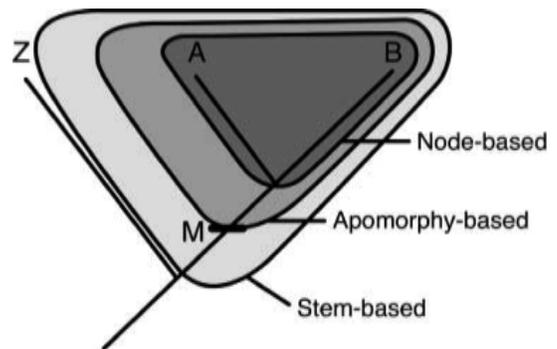
1. The Linnaean ranks would be abandoned, Instead, names of clades would be hierarchically nested uninomials regarded as proper names. A clade would retain its name regardless of where new knowledge might change its phylogenetic position, thus increasing nomenclatorial stability. Furthermore, since clade names would be presented to the community without attached ranks, users would be encouraged to look at the actual attributes of the clades they compare, thus improving research in comparative biology.
2. Two or more types (called "specifiers") would be used, for efficient and accurate representation of phylogenetic relationships. There are two types of specifiers, internal and external. Two or more of the former are used to name node-based groups. Two or more of the former and one of the latter are used to name stem-based groups.

Three ways of defining a name: Node-, Stem-, or Apomorphy-based:

A node-based definition is intended to name a clade with the referent being the most recent common ancestor of the specifiers given, and all descendants of that ancestor. Thus the two or more specifiers are all internal -- A & B in this case.

A stem-based definition is intended to name a clade with the referent being the most recent common ancestor of all the internal specifiers given (and all descendants of that ancestor), **but not including** the common ancestor with the external specifier(s) -- Z in this case.

An apomorphy-based definition is intended to name a clade with the referent being the most recent common ancestor of all the internal specifiers given (and all descendants of that ancestor) **that has the named apomorphy** (e.g., seed plants). It is thus intended to go down the stem to the ancestor in which the apomorphy (e.g., seeds) evolved. N.B., these types of names are controversial because they are ontologically mixed, and many think they should not be used.



III. Incorporation of fossils

Much of the reason why these different types of definitions exist is because of perceived needs when integrating fossils and modern organisms in the same classification, so we need to consider fossils more carefully.

Fossils are in general best treated like any other semaphoront, but require some special consideration given the facts that a few of them might actually represent ancestors, and almost all have a paucity of characteristics. A couple important concepts to mention today:

- The concept of "Form Genus." The current code allows the description of fossils into a practical genus concept that is overtly non-phylogenetic.
- The "incertae sedis" convention. This is used when it is not clear what higher taxon a lower-level taxon belongs to. Kip will talk more about this next week.

- "Crown group" = deepest clade incorporating all extant taxa (e.g., Mammals or Angiosperms)
- "Stem group" = deepest clade incorporating the ancestral lineage leading to the crown group (e.g., stem Mammals or stem Angiosperms)
- How to treat hypothesized ancestors? Would appear as 0-length terminal branches. In a phylogenetic classification should be considered in a special way, as equivalent to the name of the clade extending from that ancestor. E.g., if *Archeopteryx* really is the ancestor of crown group birds (=Aves), it should not be considered as one of the modern birds, but rather as equivalent to Aves at the time slice where it existed.
- Avoid saying "basal taxon" or even "early diverging taxon" -- nodes are ordered in time, for sure, but never terminal taxa -- and fossils should be treated as terminal taxa unless we are really sure they represent an ancestor.
- The question of which terminal taxa retain the most plesiomorphic characters is an important empirical question, addressed in the previous section of the class, when we were talking about ancestral state reconstructions, but it has nothing to do with whether a clade is not very diverse, or is typically shown on the left-hand side of a dichotomy!

IV. How could rank-free classification be applied to terminal taxa?

At the moment, the species rank is the most controversial topic among *PhyloCode* advocates. Some (primarily zoologists) want to retain it as one fundamental rank in an otherwise rank-free system, others (primarily botanists) want to get rid of it. Debate about removing species from the PhyloCode rages at this very moment (see Cellinese, Baum, and Mishler. 2012. Species and phylogenetic nomenclature. *Systematic Biology* 61: 885–891). More next time...

V. An example of publishing a formal name under the *PhyloCode* rules, forthcoming in *Phylonoms*, the companion volume to the *PhyloCode*. This will be the formal name of the Green Plants:

Viridiplantae T. Cavalier-Smith 1981 [B.D. Mishler, J. D. Hall, R. M. McCourt, K. G. Karol, C. F. Delwiche, and L.A. Lewis, this volume], converted clade name.

Registration Number: XXX

Definition: The least inclusive clade containing *Arabidopsis thaliana* (Linnaeus) Heynh. 1842, *Chlamydomonas reinhardtii* P. A. Dangeard 1888, *Ulva intestinalis* Linnaeus 1753, *Palmophyllum umbracola* W.A. Nelson & K.G. Ryan 1986, *Nephroselmis olivacea* F. Stein 1878, and *Mesostigma viride* Lauterborn 1894. This is a node-based name in which all of the specifiers are extant; it is intended to apply to the green plant crown clade. Abbreviated definition: <*Arabidopsis thaliana* (Linnaeus) Heynh 1842 & *Chlamydomonas reinhardtii* P. A. Dangeard 1888 & *Ulva intestinalis* Linnaeus 1753 & *Palmophyllum umbracola* W.A. Nelson & K.G. Ryan 1986 & *Nephroselmis olivacea* F. Stein 1878 & *Mesostigma viride* Lauterborn 1894.

Etymology: Derived from Latin, *viridis* (green) and *planta* (plant).

Reference phylogeny: The primary reference phylogeny is figure 2A in Zechman et al. (2010), a phylogeny of two plastid genes (*rbcL* and *atpB*). See also Zechman et al. (2010: fig. 2B) and Leliaert et al. (2016: figs. 4, 5).

Composition: This is one of the largest and most diverse eukaryotic clades, containing perhaps 500,000 named species ranging from unicells such as *Chlamydomonas* and *Chlorella*, to coenocytes such as *Codium*, colonial forms such as *Volvox*, and multicellular organisms such as *Ulva*, *Spirogyra*, mosses, ferns, conifers, and flowering plants. It includes two major subclades: the total clades of *Charophyta* (called *Streptophyta* in the reference phylogeny) and

Chlorophyta. In addition, the small clade *Palmophyllales*, which was classified with *Prasinococcales* into *Palmophyllophyceae* (Leliaert et al., 2016), may either be part of *Chlorophyta* or sister to *Chlorophyta* plus *Charophyta* (Zechman et al., 2010; Leliaert et al., 2016).

Diagnostic apomorphies: Chlorophyll b, thylakoids stacked into grana, "stellate" structure of the flagellae (as seen under TEM in cross-section), and true starch stored in plastids (Pickett-Heaps, 1982; Mishler and Churchill, 1985).

Synonyms: *Chlorobionta* (sensu Simpson, 2006), *Chlorobiota* (sensu Kenrick and Crane, 1997), *Viridiaeplantae* (sensu Cavalier-Smith, 2004; a linguistically incorrect alternative spelling), and *Plantae* (in some uses of that name; see Comments).

Comments:

We use *Viridiplantae* instead of *Chlorobionta*, *Chlorobiota*, or *Viridiaeplantae* because it has much wider use. *Plantae* has often been used for this clade (e.g., Copeland, 1956), but is also used widely in several other senses: e.g., for just the embryophytes (e.g., Margulis, 1974; Raven et al., 1999), or for green plants (i.e., *Viridiplantae* as defined here) plus red plants (*Rhodoplantae*), or both green and red plants plus glaucophytes (e.g., Rodríguez-Ezpeleta et al., 2005; Weber et al., 2006; Moustafa et al., 2008), or for a highly polyphyletic group consisting of all photosynthetic, multicellular organisms minus their unicellular relatives (e.g., Whittaker and Margulis, 1978; and many popular biology textbooks). Because of this ambiguity in application, we choose not to use *Plantae* for this clade so as to avoid confusion.

It was difficult to decide whether to use a node-based or a branch-based definition. This is because of two empirical problems. One, which argues against a node-based definition, is that there remain a number of understudied unicellular organisms that have been called micromonadophytes (Mattox and Stewart, 1984; Mishler et al., 1994) or prasinophytes (Lewis and McCourt, 2004), some of which branch near the base of *Viridiplantae* (see Guillou et al., 2004; Turmel et al., 2009; Zechman et al., 2010). Some of these may later be found to be sister to the clade comprising *Charophyta* and *Chlorophyta* (or these two subclades plus *Palmophyllales*, if the latter is not part of *Chlorophyta*), and thus not members of *Viridiplantae* under the node-based definition proposed here. The other empirical problem, which cautions against a branch-based definition, is that there is too little certainty about the sister group of *Viridiplantae*. It has been argued that the sister group is either the glaucophytes or the *Rhodoplantae* (Rodríguez-Ezpeleta et al., 2005, 2007; Hackett et al., 2007), but this may be an artifact of gene phylogenies derived from the plastid (which certainly is closely related in *Viridiplantae*, glaucophytes, and *Rhodoplantae*). The host organisms that gave rise to these three clades are possibly only distantly related once the chloroplast signature is removed completely -- it may be a case of one clade of endosymbionts having become established in a polyphyletic set of hosts, as seen for example in modern dinoflagellates, which occur as endosymbionts in such disparate hosts as sea anemones, sponges, corals, jellyfish, clams, radiolaria, and foraminifera (Taylor et al., 2008). We are uncertain enough about this issue to choose not to use a member of *Rhodoplantae* as an external specifier for a branch-based definition, because the name might end up applying to a more inclusive clade than intended. We conclude that the future risks of the node-based name are less than those of a branch-based name. We include multiple internal specifiers representing all the major known subclades, to help ensure that the name applies to the crown clade originating in their most recent common ancestor regardless of future changes in relationships among them.

Literature Cited:

- Copeland, H.F. 1956. The classification of the lower organisms. Pacific Books, Palo Alto, California.
- Cavalier-Smith, T. 1981. Eukaryote kingdoms: seven or nine? *Biosystems* 14:461–481.
- Cavalier-Smith, T. 2004. Only six kingdoms of life. *Proceedings of the Royal Society of London-B* 271:1251-1262.
- Guillou, L., W. Eikrem, M.-J. Chrétiennot-Dinet, F. Le Gall, R. Massana, K. Romari, C. Pedrós-Alió, and D. Vaulot. 2004. Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* 155:193–214.
- Hackett, J.D., H. S. Yoon, S. Li, A. Reyes-Prieto, S.E. Rümmele, and D. Bhattacharya. 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of *Rhizaria* with chromalveolates. *Mol. Biol. Evol.* 24:1702–1713.
- Kenrick P. and P.R. Crane. 1997. The origin and early diversification of land plants. A cladistic study. Smithsonian Institution Press Washington DC.

- Leliaert, F., A. Tronholm, C. Lemieux, M. Turmel, M. S. DePriest, D. Bhattacharya, K. G. Karol, S. Fredericq, F. W. Zechman, and J. M. Lopez-Bautista. 2016. Chloroplast phylogenomic analyses reveal the deepest-branching lineage of the *Chlorophyta*, *Palmophyllophyceae* class. nov. *Scientific Reports* 6:25367.
- Lewis, L.A., and R.M. McCourt. 2004. Green algae and the origin of land plants. *Am. J. Bot.* 91:1535–1556
- Margulis L. 1974. Five-kingdom classification and the origin and evolution of cells. *Evol. Biol.* 7:45–78.
- Mattox, K. R., and K. D. Stewart. 1984. Classification of the green algae: a concept based on comparative cytology. Pages 29–72 in: *Systematics of the green algae* (D. Irvine & D. John, eds.). Systematics Association Special Volume No. 27. Academic Press, London and Orlando.
- Mishler, B.D., and S.P. Churchill. 1985. Transition to a land flora: phylogenetic relationships of the green algae and bryophytes. *Cladistics* 1:305–328.
- Mishler, B.D., L.A. Lewis, M.A. Buchheim, K.S. Renzaglia, D.J. Garbary, C.F. Delwiche, F.W. Zechman, T.S. Kantz, and R.L. Chapman. 1994. Phylogenetic relationships of the "green algae" and "bryophytes." *Ann. Mo. Bot. Gard.* 81:451–483.
- Moustafa A., A. Reyes-Prieto, and D. Bhattacharya 2008. *Chlamydiae* has contributed at least 55 genes to *Plantae* with predominantly plastid functions. *PLoS ONE* 3(5):e2205. doi:10.1371/journal.pone.0002205
- Pickett-Heaps, J. D. 1982. New light on the green algae. *Carolina Biology Reader* 115, Carolina Biological Supply Company, Burlington, North Carolina.
- Raven, P.H., R.F. Evert, and S.E. Eichhorn, 1999. *Biology of plants*, 6th edition, New York: W. H. Freeman and Company.
- Rodríguez-Ezpeleta N., H. Brinkmann, S. C. Burey, B. Roure, G. Burger, W. Löffelhardt, H. J. Bohnert, H. Philippe, and B. F. Lang. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15:1325–1330.
- Rodríguez-Ezpeleta, N., H. Brinkmann, G. Burger, A. J. Roger, M. W. Gray, H. Philippe, and B. F. Lang. 2007. Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr. Biol.* 17:1420–1425.
- Simpson, M.G. 2006. *Plant Systematics*. Elsevier Press. Burlington, Mass.
- Taylor, F. J. R., M. Hoppenrath, and J.F. Saldarriaga. 2008. Dinoflagellate diversity and distribution. *Biodiversity and Conservation* 17:407-418.
- Turmel, M., M.-C. Gagnon, C.J. O'Kelly, C. Otis, and C.Lemieux. 2009. The chloroplast genomes of the green algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of Prasinophytes and the origin of the secondary chloroplasts of euglenids. *Mol. Biol. Evol.* 26:631–648.
- Weber, A.P.M., M. Linka, and D. Bhattacharya. 2006. Single, ancient origin of a plastid metabolite translocator family in *Plantae* from an endomembrane-derived ancestor. *Eukaryotic Cell* 5:609–612
- Whittaker, R.H., and L. Margulis. 1978. Protist classification and kingdoms of organisms. *BioSystems* 10:3–18.
- Zechman, F.W., H. Verbruggen, F. Leliaert, M. Ashworth, M.A. Buchheim, M.W. Fawley, H. Spalding, C.M. Pueschel, J.A. Buchheim, B. Verghese, and M.D. Hanisak. 2010. An unrecognized ancient lineage of green plants persists in deep marine waters. *J. Phycol.* 46:1288–1295.

Authors:

- Brent D. Mishler; University Herbarium, Jepson Herbarium, and Department of Integrative Biology; University of California, Berkeley; CA, 94720-2465; U.S.A. E-mail: bmishler@berkeley.edu.
- John D. Hall; Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, 20742; U.S.A. E-mail: jdhall@umd.edu.
- Richard M. McCourt; Department of Botany; Academy of Natural Sciences of Drexel University; Philadelphia, PA, 19103; U.S.A. E-mail: mccourt@ansp.org.
- Kenneth G. Karol; Cullman Program for Molecular Systematics; The New York Botanical Garden; Bronx, NY, 10458; U.S.A. E-mail: kkarol@nybg.org.
- Charles F. Delwiche; Department of Cell Biology and Molecular Genetics; University of Maryland; College Park, MD 20740; U.S.A. E-mail: delwiche@umd.edu.
- Louise A. Lewis; Department of Ecology and Evolutionary Biology; University of Connecticut; Storrs, CT 06269-3043 U.S.A. E-mail: louise.lewis@uconn.edu.