

Towards a comprehensive diatom classification and phylogeny (Bacillariophyta)

David W. Williams^{1,*} & J. Patrick Kociolek²

¹Botany Department, The Natural History Museum, Cromwell Road, London UK-SW7 5BD, United Kingdom

²University of Colorado Museum of Natural History, University of Colorado, Boulder CO 80309, U.S.A.

*Author for correspondence: dmw@nhm.ac.uk

Background and aims – At one time it was believed that without reasonable fossil evidence the phylogeny of any group of organisms was almost unknowable. And so fossils were sought, found and their significance duly proclaimed. Such a dominating viewpoint slowly disappeared in the 1970s and 1980s and data derived from the structure of extant organisms were given more attention. At the same time, molecular data became accessible as another source of phylogenetic information and, eventually, it was believed – as it was with fossils before – that without reasonable molecular evidence the phylogeny of any group of organisms is almost unknowable. Justification for this view has been sought in the explanation that, in general, morphology is misleading (convergence, parallelisms, etc.). These same views have come to prevail in diatom phylogenetic studies. A problem with this constellation of viewpoints is that if morphology is so misleading then so too are fossils (which are, after all, morphological specimens). This paper addresses some of these issues.

Methods – Discussion and analysis of classification, past and present, and discussion and analysis of the role played by characters (homologies), taxa and fossils in the search for a comprehensive classification.

Conclusion – Classification is equivalent to phylogeny. Sources of data (evidence) are morphology, palaeontology and molecules. Morphology and palaeontology, when understood as derived from specimens (character-bearers), are the same thing. No data source has any unique privilege over another. Stories explaining away character conflict have no place in scientific studies.

Key words – diatoms, Bacillariophyta, classification, fossils, phylogenetics, characters (homology).

INTRODUCTION

This paper is a brief exploration of the interrelationships between phylogeny, classification and fossils and how that interplay may be used with profit to explore diatom relationships.

PHYLOGENY: THE PAST AND THE PRESENT

Some time ago, investigation into the phylogenetic relationships of any group of organisms was frowned upon, considered a dull and possibly irrelevant subject. For example, an early critique, published in 1917 in the *New Phytologist* (Ernst Haeckel was still alive, Richards 2008) under the title of *The Reconstruction of Elementary Botanical Teaching* (Blackman et al. 1917) became part of an extended and at times acrimonious discussion which eventually became known as the ‘Tansley Manifesto affair’, named after the British ecologist Arthur Tansley (Godwin 1957, 1985), one of the five signatories (Blackman et al. 1917, the ‘manifesto’ and the entire

two year correspondence was later published as a pamphlet, Anonymous 1919; see Boney 1991 for a detailed account of the affair). The ‘manifesto’ had as its aim to set the teaching of botany on a new course in the UK. They noted:

“Botany in this country [UK] is still largely dominated by the morphological tradition, founded on the attempt to trace the phylogenetic relationships of plants, which began as the result of the general acceptance of the doctrine of descent. Elementary teaching is mainly occupied with the endless facts of structure and with their interpretation from the phylogenetic standpoint. Side by side with this there generally goes a discussion of function which is often limited by a crude Darwinian teleology.” (Blackman et al. 1917: 242).

The phrase “...endless facts of structure and with their interpretation from the phylogenetic standpoint”, meant morphology explained by some story concerning how one structure changed into another over time. It is quite possible that Tansley’s article set back taxonomy, systematics and classification for decades in the UK – its message possibly still

having an effect, lingering on in succeeding generations who bought into their ‘vision’ as if it was accepted wisdom.

Whatever truth there is in the Blackman et al. article, what they were describing was not the pitfalls of morphology or even those of phylogeny but what eventually – in the 1970s – corresponded to the narrative phase of phylogeny: that is telling stories about changes in taxa, about transformations in characters. Inspiration for such stories usually came from the fossil record, with diagrams linking one group to another via their stratigraphically placed fossil remains.

Remarkably, narrations still exist. Nearly 90 years after Tansley’s broadside, again in the pages of the *New Phytologist*, and somewhat ironically as an invited Tansley Review, is an article that exemplifies this approach – but this time using molecular evidence rather than fossils as data that requires phylogenetic ‘interpretation’:

“It is of interest that the molecular phylogeny of diatoms also favours the centric forms as ancestral...with some molecular evidence agreeing [data] that the earliest diatoms could have been neritic...This suggests that silicification evolved late in the evolutionary history of diatoms. It is possible that any earlier silicified diatoms have failed to be preserved for some environmental or taphonomic reasons...” (Raven & Waite 2004: 46).

One might examine some of those phrases for meaning: “...favours the centric forms as ancestral...” Ancestral to what?; “...the earliest diatoms could have been neritic...” Could have been?; “It is possible that any earlier silicified diatoms have failed to be preserved...” Well, anything is possible in this world of make believe. These are simply stories, inventions with no substance in empirical science: data, evidence. More importantly, they are not empirically constrained – the data hardly matter.

The problem Tansley and his co-authors should have been addressing was not the source of evidence (morphology in their case) but the method of analysis and interpretation, or lack of it. In this sense, one might view the history of diatom classification as a series of changes in the source of data (live specimens for Agardh; frustule and valve morphology for H. L. Smith; plastids for Mereshkowsky, and so on) or else some general appeal for vast quantities of data, evidence enough to arrive at a sensible answer. Of course, to search for a sensible answer, you need a sensible question.

DIATOM CLASSIFICATIONS AND CHANGING EVIDENCE

Here we address the issue of data (evidence) with reference to past diatom classifications. A widely accepted classification was that presented by H.L. Smith, who divided diatoms into three groups he named as tribes, Raphidieae, Pseudoraphidieae and Crytoraphidieae (Smith 1872). Smith’s work was published as a series of articles in *The Lens*. Most copies were lost in a fire in Chicago (Van Heurck tells the story in *Le Microscope*, Van Heurck 1878). Henri Van Heurck, who thought highly of Smith’s work, reprinted his classification in the third edition of *Le microscope* (Van Heurck 1878), using it to organise the *Synopsis des diatomées de Belgique* (Van Heurck 1880–1885) and *A treatise on the Diatomaceae* (Van Heurck 1896). It was via Van Heurck’s work that Smith’s classification was promoted.

Smith’s three tribes correspond to today’s raphid, ‘araphid’ and ‘centric’ diatoms – not exactly but close enough. What evidence characterises each group? For raphid diatoms, it is the presence of a raphe on at least one valve; for pseudoraphid (‘araphid’) diatoms, it is a combination of their lack of a raphe and bilaterally symmetrical valves; for cryptoraphids, it is the valve symmetry, which is more or less circular, radially symmetrical (both raphid and ‘araphid’ diatom valves are bilaterally symmetrical). Effectively, Smith divided diatoms up on the basis of a combination of two characters: valve symmetry and whether they possess a raphe or not.

Another, different classification, from the same time, was offered by Pfitzer but based on a different character, the plastids (Pfitzer 1871). Pfitzer proposed that diatoms could be divided into two groups, Coccochromaticae, with many small plastids, and Placochromaticae with a few large ones. Taxa within Coccochromaticae are themselves subdivided on the basis of valve symmetry (bilateral versus radial), presenting a conflict between the evidence supporting the sub-division using plastids and the sub-division using valve symmetry, as bilaterally symmetrical valves occur in both of Pfitzer’s primary groups, Placochromaticae and Coccochromaticae (table 1).

Interestingly, Pfitzer (1871: 148) also offered a diagram which includes various generic names enclosed in boxes and arranged in a scheme designed to indicate what he thought their relationships might be, the nearer the boxes the closer

Table 1 – Summary of Pfitzer’s 1871 classification of diatoms with the main two subgroups and divisions within.

Coccochromaticae		
I. Bilaterale Formen	a. Nach der Querebene symmetrisch	b. Nach der Querebene asymmetrisch
II. Centrische Formen	a. Schalen mit theilweise zygomorpher Gestaltung	b. Schalen rein centrisch
Placochromaticae		
a. Mit Knoten	α. asymmetrische Formen	β. symmetrische oder diagonal gebaute Formen
b. Ohne Knoten		

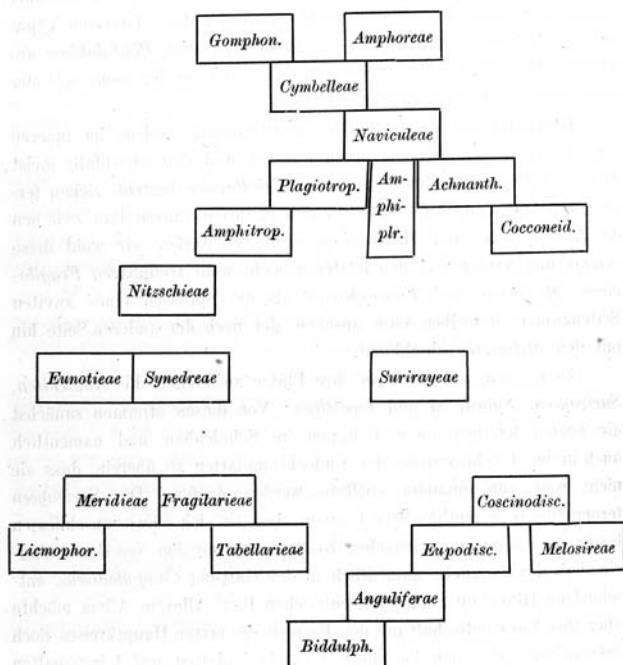


Figure 1 – Reproduction of a figure from Pfitzer (1871: 148). Pfitzer's figure includes taxon names (most as families) enclosed in boxes and arranged in a scheme designed to indicate their relationships; the nearer the boxes, the closer the relationship.

the relationship, although certain parts of the diagram are difficult to interpret (fig. 1). For example, the implication is that Eunotieae and Synedrae are very closely related as the boxes are linked together, whereas Surirayeae seems to exist in some phylogenetic limbo, implying its relationships are unknown (fig. 1). Nevertheless, the interesting aspect of Smith's and Pfitzer's classifications is not whether they are now of any import but as an illustration of conflict in certain characters: that is, evidence that may point in two (or more) directions, evidence that support different outcomes as only part of Pfitzer's Coccochromaticae is equivalent to Smith's Cryptoraphideae ('Schalen centrisch gebaut'), the remaining sub-section ('Schalen nach Umriss und Structur bilateral gebaut') plus the Placochromaticae being equivalent to Raphidieae plus Pseudoraphideae (table 2).

How might one deal with character conflict? As far as we can tell, there are three ways, the first, in our view, the least successful.

1. Choose a classification arbitrarily

A classification proposed some 20 years after Smith and Pfitzer was that of Schütt's in his contribution to *Die Natürlichen Pflanzenfamilien*; his classification has also been widely used (Schütt 1896). Schütt discussed both Smith's and Pfitzer's classification and the evidence used to support each. He offered his own new classification dividing diatoms into two primary groups, Centricae and Pennatae, the former more or less equivalent to Smith's Cryptoraphideae and Pfitzer's 'Schalen centrisch gebaut' (the first sub-division of Coccochromaticae), the latter equivalent to Smith's Raphidieae plus

Pseudoraphideae and Pfitzer's 'Schalen nach Umriss und Structur bilateral gebaut' (the second sub-division of Coccochromaticae) and Placochromaticae. Schütt's classification was based not just on evidence but on argument. That is, Schütt ignored the conflict by simply assuming that symmetry was of greater significance. Interestingly, Schütt did divide his Centricae into two groups based more or less on symmetry, recognising that within centric diatoms, there were some 'non-centric' centric diatoms: Eucyclicae represented what might be thought of as those Centricae that had valves with well defined radial symmetry and the Hemicyclicae which included many groups that had bilaterally symmetrical valves but lacked a central sternum, such as species in Rutilariaceae. Thus, Schütt's argument in defence of symmetry was qualified by other considerations leaving one with the conclusion that his classification was somewhat arbitrary, relative to the available evidence.

One problem with a classification that is created (as opposed to supported) is that it often assumed to represent some aspect of reality and interpreted as such. For example, the botanist Charles Bessey (Pool 1915; Overfield 1993; Cuerrier et al. 1996) in a discussion of diatom evolution (Bessey 1900), used a modified version of Schütt's classification and from it drew a phylogenetic tree (derived in the way that Tansley and his associates were probably grumbling about, fig. 2). Bessey noted the major subdivision between the centric and pennate diatoms. He also noted the symmetry of Rutilariaceae that in his diagram appear highly derived at the tip of the centric diatom lineage. Bessey promotes the idea that within the centric and pennate diatom lineages, both were evolving towards some form of bilaterally symmetrical valves (see fig. 2). Much later, Chin proposed a similar scheme but added a geological scale for diatom fossils, which offered a more precise temporal dimension (Chin 1978: table 1, see also Chin 1991: 93, fig. 76, for a more graphic diagram).

Bessey's and Chin's interpretation might appear considered, as that was the evidence available at the time. But that evidence was summarised by the classification, and the classification included characters that conflicted with the subdivisions, hence Bessey's and Chin's interpretation of the apparently 'derived' nature of Rutilariaceae relative to other centric diatoms and first appearances in the fossil record. Their interpretation is executed to make a story fit an artificial result (reflecting "a crude Darwinian teleology", see above). Thus, as we stated above, phylogenetic narratives are often derived from fossil record evidence, and diatom narratives are no exception.

We close this section with a few comments on fossils (as opposed to the fossil record) and their value. Colin Patterson, a vertebrate palaeontologist, wrote the following in 1987:

"By about 1960 palaeontology had achieved such a hold on phylogeny reconstruction that there was a commonplace belief that if a group had no fossil record its phylogeny was totally unknown and unknowable" (Patterson 1987: 8).

What did he mean? More or less, that if the fossil record, the stratigraphic record, wasn't taken into account, then, so it was believed, phylogenetic relationships were forever elusive. This turned out to be untrue as phylogenetic relationships can be determined more directly from evidence at hand:

Table 2 – Comparison of Pfitzer’s (1871) and Smith’s (1872) classification; Schütt’s (1896) later solution is added as a third column.

Pfitzer (1871)	H.L. Smith (1872)	Schütt (1896)
I. Coccochromaticae		
a. Schalen centrisch gebaut	Cryptoraphideae	Centricae
b. Schalen nach Umriss und Structur bilateral gebaut	Raphideae +	Pennatae
II. Placochromaticae		
	Pseudoraphideae	

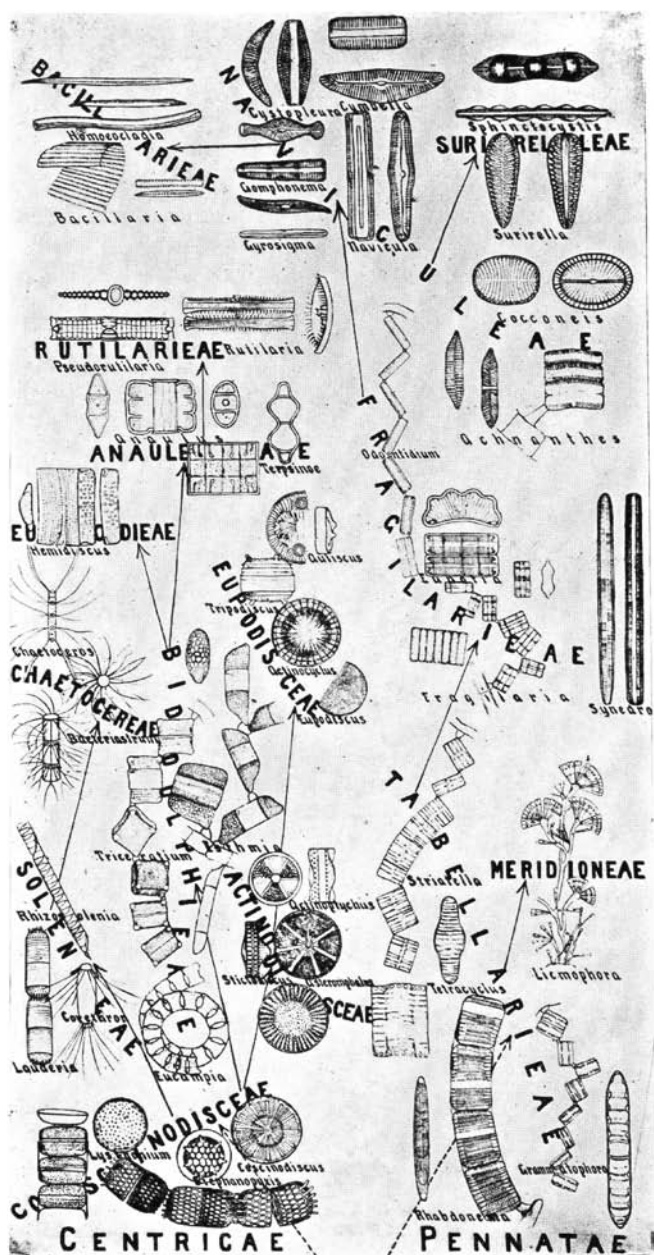


Figure 2 – Reproduction of a figure from Bessey (1900: plate V). Bessey adopted a modified version of Schütt’s classification to construct his phylogenetic tree.

that is, characters from specimens (see below). While it may be that the stratigraphic record is of limited use, it does not mean that the fossils themselves are of limited value. Fossils are indeed only just specimens, albeit from extinct organisms, their value is in providing character data as evidence for their relationships – a view that many diatomists have understood intuitively (for example, see Andrews 1974, on *Raphidodiscus marylandicus* Christian).

2. Assume all data so far examined are poor or of insufficient value and choose another source

This has been the usual way of dealing with character conflict. If all the data do not speak directly to one solution the data must be wrong or poor in some way – and this is the most common way of dealing with inconsistencies in morphological data. The solution has (usually) been: find a new data source. This ‘data du jour’ approach has run through the history of diatom systematics and is evident today. Currently, that new data source is DNA. The promises of palaeontology were never fulfilled and, in the search for more relevant data, “Today the magic of paleontology is supplanted by the magic of DNA” (Nelson & Ladiges 2009: 40). Here another notion (story) is attached to its justification: every living thing is made of DNA; these are the building blocks of life; so it must be the source of data, the crucible of truth (Nelson et al. 1987: 543). Inspection of some trees derived from DNA sequence reveals problems (see, for example, Rampen et al. on *Attheya* where, after a thorough analysis of their data, they conclude: “...we were unable to determine the exact phylogenetic position of the investigated *Attheya* species within diatoms”, Rampen et al. 2009: 444, abstract). The argument proceeds in the usual fashion: DNA must be right, morphology must be wrong, an argument similar in many respects to that made by Pfitzer for plastids as opposed to symmetry and Schütt, for symmetry as opposed to plastids.

With respect to (most) fossil data, there is no DNA, and simply mapping their characters onto a DNA derived phylogenetic tree does not account for which characters are or are not homologues, in spite of those who might speak on their behalf. When DNA sequence trees do not agree with morphology, it suggests a problem requiring solution rather than argument for or against any particular data source. This leads to the third, and our preferred, solution.

3. Dealing with conflict is a problem of interpretation

Examination of the characters might be profitable. For example, one might understand Pfitzer's plastid characters as two different 'things' (we acknowledge that plastid characters are more complex than the way in which we deal with them here): they may be considered either as 'many small plastids' or a 'few large ones'. Treated as such they may be seen as alternatives, each suggesting a sub-group of diatoms. This is how Pfitzer understood matters and derived two sub-groups: Coccochromaticae and Placochromaticae. When characters are viewed this way, as unit characters, it reflects an approach (an assumption) that was once part of a method of classification called phenetics (Sneath & Sokal 1973). The assumptions, as far as they went, were that if something could be described or, better, measured, then it qualified as a taxonomic character, and with enough of these kinds of characters, truth would emerge (a different, earlier, crucible of truth). Phenetics was found to be false. Why and how, for example, could measuring the number of striae in 10 µm on diatom valves be related to some historical change? What relationship does the character '10 striae in 10 µm' have with '30 striae in 10 µm'? One might concoct some scenario but it seems almost pointless (we acknowledge that '10 striae in 10 µm' versus '30 striae in 10 µm' can help *identify and distinguish* two species but we are not addressing the problem of identification). These measurements are not characters in the sense used in phylogenetic reasoning.

It might be supposed that various kinds of plastids are, in fact, related to each other in some fundamental (phylogenetic) way. That is, 'small plastids' (coccochrome) and 'large ones' (placochrome) are different aspects of the same thing, modifications of each other, if you like. What does that mean? It can represent a simple hypothesis: that, for example, the coccochrome condition is primitive, the placochrome condition derived. What does that mean? It means that the coccochrome condition *on its own* is not a character (it is a description), just part of a character; as a modified character (the placochrome condition) must be a modification of something else (the coccochrome condition). Thus, general data on plastids might yield evidence for two groups:

1. Diatoms (as a whole): coccochromes + placochromes
2. Raphid diatoms plus some 'araphid' diatoms: placochromes

All characters can be subject to such understanding and interpretation. These characters are not propositions concerning truth but propositions about characters and the groups they might represent. Consider valve symmetry. One way is to consider symmetry as a property of the valve sternum. If a valve has a sternum, it exhibits bilateral symmetry. But to consider the sternum as two characters – presence of a sternum, absence of a sternum – leads to problems. First, mere observation suggests that all raphid diatoms have a sternum, even if in some cases it is obscured or integrated closely with the raphe. Thus, presence of a sternum picks out all pennate diatoms, raphid and 'araphid' alike; the sternum *on its own* is not a character (it is a description), just part of one. Thus, the characters are 'sternum plus raphe' and 'the raphe' itself (again, we acknowledge that the raphe encompasses a number of different structures, themselves ripe for interpre-

tation). Thus, general data on the valve sternum also yields evidence for two groups:

1. Pennate diatoms (raphid + 'araphid' diatoms)
2. Raphid diatoms

If we consider symmetry alone, rather than just the possession of a sternum, another group emerges:

3. Pennate diatoms + some non centric 'centric' diatoms (Schütt's Hemicycliaca).

On their own, these are simply hypotheses about the relationships of characters and how they support the relationships of groups. Combined, the separate character hypotheses yield further hypotheses: That there is support for pennate diatoms; there is support for raphid diatoms; there is support for a group composed of raphid diatoms plus some 'araphid' diatoms; and there is support for a group of all pennate diatoms plus some non centric 'centric' diatoms. Viewed this way, the characters do not now conflict. They complement one another and lend support to the various groupings, each at a different hierarchical level: this exercise demonstrates that these data discussed above and used by Smith and Pfitzer, are, after all, useful. Thus conflict is resolved by *understanding* data rather than ignoring it or by explaining it away, resolving Tansley's general complaint.

The interpretation above is a simplified form of cladistic analysis (phylogenetic systematics), a view that is now accepted as commonplace in almost all other areas of systematic endeavour (see, for example, the papers in Cracraft & Donoghue 2004). Such analyses understand characters as dynamic, the parts relating to one another, rather than the static unit characters of phenetics. Cladistic analysis also attempts to integrate all kinds of data, such as the studies of Kociolek & Stoermer using sexual reproduction, chloroplasts and valve features for the cymbelloid and gomphonemoid diatoms (Kociolek & Stoermer 1988).

CONCLUSIONS

There is a need to synthesise data available for understanding taxon relationships, rather than proposing more and more elaborate scenarios to explain character evolution. Thus, cladistic analysis allows a direct connection of phylogeny with classification.

For a scientific – as opposed to a narrative based – approach to phylogeny, one needs to consider a number of items: classification can be viewed as directly equivalent to phylogeny. The sources of data are morphology, palaeontology and molecules. Morphology and palaeontology, when understood as derived from specimens (character-bearers), are the same thing. No data source has any unique privilege over any another. Creating stories to explain away character conflict has no place in scientific studies.

REFERENCES

- Andrews G.W. (1974) Systematic position and stratigraphic significance of the marine Miocene diatom *Raphidodiscus marylandicus* Christian. *Nova Hedwigia Beihefte* 45: 231–243.
- Anonymous (1919) The reconstruction of elementary botanical teaching: with a discussion of the problems raised. Reprinted

- from the *New Phytologist* 1917, 1918 and 1919. London, William Wesley and Co.
- Bessey C.E. (1900) The modern conception of the structure and classification of diatoms, with a revision of the tribes and a rearrangement of the North American genera. *Transactions of the American Microscopical Society* 21: 61–86.
- Blackman F.F., Blackman V.H., Keeble F., Oliver F.W., Tansley A.G. (1917) The reconstruction of elementary botanical teaching. *New Phytologist* 16: 241–252.
- Boney A.D. (1991) The ‘Tansley Manifesto’ affair. *New Phytologist* 118: 3–21.
- Chin T.G. [Jin Dexiang] (1978) On the phylogenesis of diatoms. *Journal of Xiamen University (Natural Science)* 2: 1–8.
- Chin T.G. [Jin Dexiang] (1991) Marine diatomology (Haiyang Guizaoxue). Xiamen, Xiamen University Press [In Chinese].
- Cracraft J., Donoghue M.J. (2004) *Assembling the tree of life*. New York, N.Y., Oxford University Press.
- Cuerrier A., Kiger R.W., Stevens P.F. (1996) Charles Bessey, evolution, classification, and the New Botany. *Huntia* 9: 179–213.
- Godwin H. (1957) Arthur George Tansley, 1871–1955. *Biographical Memoirs of Fellows of the Royal Society* 3: 227–246.
- Godwin H. (1985) Early development of the *New Phytologist*. *New Phytologist* 100: 1–4.
- Kocielek J.P., Stoermer E.F. (1988) A preliminary investigation of the phylogenetic relationships of the freshwater, apical pore field-bearing cymbelloid and gomphonemoid diatoms (Bacillariophyceae). *Journal of Phycology* 24: 377–385.
- Nelson G.J., Atz W., Kallman K.D., Smith C.L. (1987) Donn Eric Rosen, 1929–1986. *Copeia* 1987: 541–547.
- Nelson G.J., Ladiges P.Y. (2009) Biogeography and the molecular dating game: a futile revival of phenetics? *Bulletin de la Société Géologique de France* 180: 39–43.
- Overfield R.A. (1993) *Science with practice: Charles E. Bessey and the maturing of American botany*. Ames, Iowa, Iowa State University Press.
- Patterson C. (1987) Introduction. In Patterson C. (ed.) *Molecules and morphology in evolution: conflict or compromise?*: 1–22. Cambridge, Cambridge University Press.
- Pool R.J. (1915) A brief sketch of the life and work of Charles Edwin Bessey. *Transactions of the American Microscopical Society* 34: 299–305.
- Pfitzer E. (1871) *Untersuchungen über Bau und Entwicklung der Bacillariaceen (Diatomaceen)*. Bonn, Adolf Marcus [as Heft 2 of J. Hanstein, *Botanische Abhandlungen aus dem Gebiet der Morphologie und Physiologie*. 1872].
- Rampen S.W., Schouten S., Panoto F.E., Brink M., Andersen R.A., Muyzer G., Abbas B., Sinninghe Damsté J.S. (2009) Phylogenetic position of *Attheya longicornis* and *Attheya septentrionalis* (Bacillariophyta). *Journal of Phycology* 45: 444–453.
- Raven J.A., Waite A.M. (2004) The evolution of silicification in diatoms: inescapable sinking and sinking as escape? *New Phytologist* 162: 45–61.
- Richards R.J. (2008) *Tragic sense of life: Ernst Haeckel and the struggle over evolutionary thought*. Chicago, University of Chicago Press.
- Schütt F. (1896) Bacillariales. In: Engler A., Prantl K. (eds) *Die Natürlichen Pflanzenfamilien*, 1(1b): 31–153. Leipzig, W. Englemann.
- Smith H.L. (1872) *Conspectus of the families and genera of the Diatomaceae*. *The Lens* 1: 1–19, 72–93, 154–157.
- Sneath P.H.A., Sokal R.R. (1973) *Numerical taxonomy*. San Francisco, Freeman.
- Van Heurck H.F. (1878) *Le microscope: la construction, son manie- ment et son application à l’anatomie végétale et aux diatomées*. 3^e Ed. Bruxelles, E. Ramlot.
- Van Heurck H.F. (1880–1885) *Synopsis des diatomées de Belgique*. Anvers, Brouwers & co.
- Van Heurck H.F. (1896) *A treatise on the Diatomaceae, containing introductory remarks on the structure, life history, collection, cultivation and preparation of Diatoms*. London, W. Wesley & Son.
- Paper based on a keynote presented during the Symposium “Diatom Taxonomy in the 21st Century” (Meise 2009). Manuscript received 17 Nov. 2009; accepted in revised version 23 Jun. 2010.

Communicating Editor: Bart Van de Vijver.