J. Phycol. 49, 215–225 (2013)

DOI: 10.1111/jpy.12020

MINIREVIEW

ALGAL TAXONOMY: A ROAD TO NOWHERE?[1]

Olivier De Clerck[2]
Phycology Research Group, Ghent University, Krijgslaan 281-S8, 9000 Ghent, Belgium
Michael D. Guiry
AlgaeBase and Irish Seaweed Research Group, Ryan Institute, National University of Ireland, University Road, Galway, Ireland
Frederik Leliaert
Phycology Research Group, Ghent University, Krijgslaan 281-S8, 9000 Ghent, Belgium
Yves Samyn
Belgian National Focal Point to the Global Taxonomy Initiative, Royal Belgian Institute of Natural Sciences, Vautierstraat 29, 1000 Brussels, Belgium
and Heroen Verbruggen
School of Botany, University of Melbourne, Melbourne, Victoria 3010 Australia

The widespread view of taxonomy as an essentially retrogressive and outmoded science unable to cope with the current biodiversity crisis stimulated us to analyze the current status of cataloguing global algal diversity. Contrary to this largely pessimistic belief, species description rates of algae through time and trends in the number of active taxonomists, as revealed by the web resource AlgaeBase, show a much more positive picture. More species than ever before are being described by a large community of algal taxonomists. The lack of any decline in the rate at which new species and genera are described, however, is indicative of the large proportion of undiscovered diversity and bears heavily on any prediction of global algal species diversity and the time needed to catalogue it. The saturation of accumulation curves of higher taxa (family, order, and classes) on the other hand suggest that at these taxonomic levels most diversity has been discovered. This reasonably positive picture does not imply that algal taxonomy does not face serious challenges in the near future. The observed levels of cryptic diversity in algae, combined with the shift in methods used to characterize them, have resulted in a rampant uncertainty about the status of many older species. As a consequence, there is a tendency in phycology to move gradually away from traditional names to a more informal system whereby clade-, specimen- or strain-based identifiers are used to communicate biological information. Whether these informal names for species-level clades represent a temporary situation stimulated by the lag between species discovery and formal description, or an incipient alternative or parallel taxonomy, will be largely determined by how well we manage to integrate historical collections into modern taxonomic research. Additionally, there is a pressing need for a consensus about the organizational framework to manage the information about algal species names. An eventual strategy should preferably come out of an international working group that includes the various databases as well as the various phycological societies. In this strategy, phycologists should link up to major international initiatives that are currently being developed, such as the compulsory registration of taxonomic and nomenclatural acts and the introduction of Life Science Identifiers.

Key index words: **algae; AlgaeBase; barcoding; dark species; DNA taxonomy; species discovery; taxonomy**

> *… if our rates of discovery continue at the present pace, we should have a definitive total for insects in a little over fifteen thousand years. The rest … may take a little longer* – B. Bryson (2003).

Long before Carl von Linné (1707–1778), the instigator and popularizer of modern nomenclature and taxonomy, mankind showed a propensity to document, describe, and catalogue all the living things on our planet (Godfray and Knapp 2004). The practice of giving names to organisms, probably born out of necessity among hunter-gatherer socie-

---

[1] Received 1 September 2012. Accepted 15 October 2012.
[2] Author for correspondence: e-mail Olivier.declerck@ugent.be.

ties and later a privileged pastime of the rich, has transformed itself over the years into a modern science embracing the principles of evolutionary biology as well as revolutionary innovations on multiple technological fronts (Goldstein and DeSalle 2011). The importance of describing and naming species (alpha taxonomy) transcends the mere cataloguing of diversity. Clearly, it is an enabling discipline, delivering basic and indispensable knowledge for many fields of human interest and underpinning a host of scientific research, including biodiversity conservation, understanding of ecosystem services, documenting climate change, and ecological modeling. As a central tenet, and virtually unchallenged for over 250 years, taxonomy makes use of a rank-based classification system whereby Linnaean binomials serve as anchors for biological information about a species, including its taxonomic affinities, morphology, distribution, and ecological role.

The task of discovering and documenting biodiversity has been given an increased sense of urgency now that anthropogenic impacts are dramatically altering the biota of the Earth (Cardinale et al. 2012), with current extinction rates now as high as 27,000 of the known species per annum (Wilson 1992, Maddison et al. 2012). However, the general perception persists that taxonomy is a science in crisis, and a discipline that suffers from a lack of prestige and resources (Wheeler 2004, Boero 2010). The magnitude of biological diversity seems so overwhelming and the discovery and description of new species proceeds at such a slow pace (6,000–18,000 species of eukaryotes per year; Mora et al. 2011) that millennia of basic descriptive taxonomy seem to await us before alpha taxonomy reaches its goal. The inadequate pace of taxonomy has resulted in several proposals (and the inevitably accompanying contentious debates) aimed at making the process of describing species more efficient (e.g., Tautz et al. 2003, Godfray et al. 2007, Wheeler 2008, Clark et al. 2009, Smith et al. 2009, Patterson et al. 2010, Deans et al. 2012, Maddison et al. 2012).

Contrary to the widespread pessimistic view of taxonomy as an essentially retrogressive and outmoded science unable to cope with the current biodiversity crisis (Goldstein and DeSalle 2011), a much more optimistic view was presented by Joppa et al. (2011). These authors analyzed rates of species descriptions for a number of animal and plant taxa and concluded that, contrary to the generally accepted view, both the rates of species description and the number of taxonomists have increased exponentially since the 1950s. Interestingly, they also noted a marked decline in the number of species described per taxonomist, something that they attribute to the increasing difficulty of finding new species in an ever-declining "missing-species pool." This perceived decline has been attributed to the choice of test cases (flowering plants, cone snails, spiders, amphibians, birds, and mammals), representing well studied and taxonomist-rich groups (Bacher 2012, Samyn and De Clerck 2012). Applying a much broader taxon sampling, Costello et al. (2012) confirmed the existence of a greater taxonomic effort as reported by Joppa et al. (2011); however, the rate of species description was shown to be virtually linear from the mid-19th century onward with no evidence of decrease or, indeed, increase.

In this context, the questions of how we are doing as algal taxonomists and what the status of descriptive algal taxonomy is, are relevant to ask. Does a daunting task await phycologists with the prospect of several hundreds or even thousands of years of alpha taxonomy at the present pace or could the remaining algal diversity perhaps be described at the current rate within the next century? Answering this question requires accurate estimates of the total number of algal species, the number of species already described and the rate at which species are currently being described. Until recently, a lack of primary data on algal species descriptions, the number of currently accepted names and synonyms largely prevented any quantitative analysis of algal diversity and trends in descriptive taxonomy. Today, information on algal diversity and taxonomy is available in a number of online databases, each with their own focus and limitations. The *Index Nominum Algarum* (*INA*; Silva and Moe 1999) is a nomenclator, a high-quality compendium of all algal names, their original places of publication, and type information; but *INA* generally offers no information on accepted names and taxonomic synonyms. AlgaTerra (Jahn and Kusber 2012) and the on-line Catalogue of Diatom Names at the California Academy of Sciences (Fourtanier and Kociolek 2012) integrate nomenclatural information on terrestrial and freshwater microalgae, and diatoms, respectively, but are limited in their taxonomic information. AlgaeBase (Guiry and Guiry 2012) databases taxonomic, nomenclatural, bibliographic, and biogeographical information on micro-algae and macroalgae from marine as well as freshwater and terrestrial environments on a worldwide basis. Even though AlgaeBase is a work in progress (Guiry 2012), it does not have the taxonomic limitations of the other algal databases. We have therefore used AlgaeBase as a primary source of data to generate a set of metrics enabling us to evaluate the current state of algal taxonomy.

Here, we analyze the rates of species description through time as well as trends in the number of algal taxonomists and evaluate the progress of algal taxonomy in the light of recent estimates of global algal diversity (Appendix S1, see Supporting Information for technical details). In addition, we discuss general trends in species discovery, contrast these with species description rates, highlight important challenges that algal taxonomy is facing and make recommendations toward future taxonomic practices.

ALGAL DIVERSITY AND TAXONOMIC EFFORT

The rate at which new algal species are described provides a direct measure of taxonomic effort in the phycological community. Additionally, growth in the cumulative numbers of described species may provide an indication of the global species diversity. A cumulative curve of algal species descriptions through time (Fig. 1), based on more than 32,000 species names in AlgaeBase of the estimated 44,000 described to date (Guiry 2012), shows no sign of leveling off. Description rates of currently recognized species show an initial lag phase, during which relatively few species were described (from 1753 to about 1850), followed by an era of high taxonomic effort that persists until the present day. Unlike well-known groups in which rates decrease as fewer species remain to be described such as the birds, mammals, and higher plants (Costello and Wilson 2011, Joppa et al. 2011), there is no evidence for a decrease in the description rates of algal species. Although the yearly number of descriptions shows considerable variation through time, there is a gradual overall increase of the description rate over time (Fig. 2). Temporal variation in description rates reveals an erratic pattern that can best be explained as a combination of global events (e.g., World Wars I & II) combined with a disproportionate taxonomic output by a few individuals (Table 1). Taxonomic effort was relatively low during most of the 1960s, 1970s, and 1980s, but picked up again at the end of the century. This increase of species descriptions coincides with and might be the result of an increased use of molecular sequence data in systematics. However, this trend is not persisting as evidenced by the lower description rates since the turn of the millennium.

The number of taxa that are currently regarded as synonyms is about half the number of currently recognized species and an additional 20,000 species names are flagged as uncertain in AlgaeBase (Fig. 1). Although some of these names are undoubtedly invalid or illegitimate, or published in obscure or poorly known sources, most of these names represent taxa the current status of which is difficult to assess. Note that we use the terms invalid and illegitimate in the sense of the *International Code of Nomenclature for Algae, Fungi and Plants* (Mcneill et al. 2012). In zoology the term “valid name” is often used to denote a currently recognized name, as opposed to a synonym (see for example, Appeltans et al. 2012). The sheer number of these names, however, is daunting. From the mid-19th until the mid-20th century the number of species with uncertain status is of the same order of magnitude and often surpasses the currently accepted species as well as the synonyms (Fig. 2). In other words, there is considerable uncertainty regarding the taxonomic status of a large number of algal species names. This uncertainty on the status of a large part of the

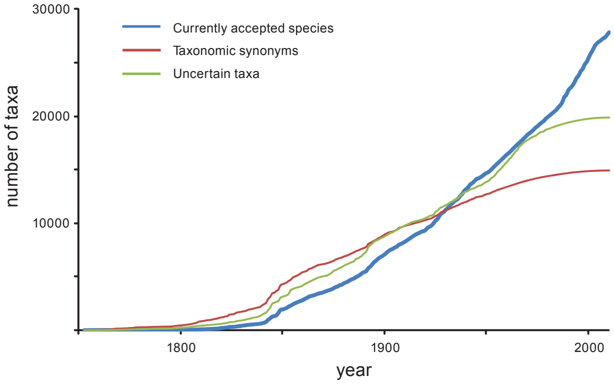

FIG. 1. Cumulative number of currently recognized algal species, taxonomic synonyms and taxa of uncertain taxonomic status described over time. Extracted from AlgaeBase.

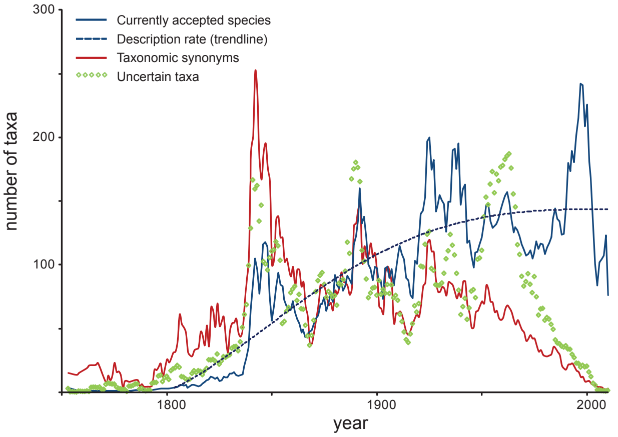

FIG. 2. Description of algal species per year. Currently recognized species, taxonomic synonyms and species flagged as “uncertain” are presented as a 5-year moving average. A trend line depicts the gradual increase in description rate.

TABLE 1. A preliminary list, extracted from AlgaeBase, of algal taxonomists that have described more than 1,000 species.

Author	Nationality	Main group	No. species
Friedrich Traugott Kützing (1807–1893)	German	All	2,636
Christian Gottfried Ehrenberg (1795–1876)	German	Microalgae	2,055
Albert Grunow [Grunov] (1826–1914)	German	All	1,251
Friedrich Hustedt (1886–1968)	German	Diatoms	1,219
Horst Lange-Bertalot (1936–present)	German	Diatoms	1,145
Jacob Georg Agardh (1813–1901)	Swedish	Seaweeds	1,144
Boris Vassilievich Skvortsov [*Скворцов*] (1890–1980)	Russian	Microalgae	1,073
William Henry Harvey (1811–1866)	Irish	Seaweeds	1,061

described species was fittingly illustrated in a recent study by Thessen et al. (2012) on the dinoflagellate genus *Gymnodinium* in which 38% of 236 extant species was only known from the original description. As the authors pointed out, these singletons seriously limit our current understanding of biodiversity. They may be validly described species that are members of a rare biosphere. Alternatively, they may reflect that taxonomic practices are too poor to allow the organism to be re-identified or that the descriptions are not known to other researchers. In many groups, particularly in green microalgae and diatoms, the existence of large numbers of varieties and forms may indicate many undescribed species, or it may mean that the characters upon which these infraspecific taxa are based are relatively minor variants not worthy of consideration at the species level.

The temporal accumulation of taxa of higher taxonomic levels is depicted in Figure 3A–E. For these graphs, the description of the earliest species belonging to the higher taxon was used as the date of discovery of that particular taxon. The discovery date of a taxon may therefore predate the actual description of the taxon by several decades. For example, the class Mamiellophyceae was described as recently as 2010 (Marin and Melkonian 2010), but the oldest species assigned to it, *Monomastix opisthostigma* Scherfell, dated from 1912 (Scherffel 1912). Taxa of all but one of the currently recognized phyla, the Chlorarachniophyta being an exception, were discovered prior to 1850. The accumulation of classes, orders, and families reveals a remarkably similar pattern, with a more or less linear discovery rate in the 19th century, gradually leveling off in the 20th century. It is interesting to note that the accumulation curve of classes lags behind those of orders and families since 1900 (Fig. S1, in the Supporting Information). In our opinion, this illustrates appropriately how the deep divergences that separate many algal groups have only become appreciated very slowly, coinciding with advances in culture techniques, analytical biochemistry (e.g., pigment and storage-compound

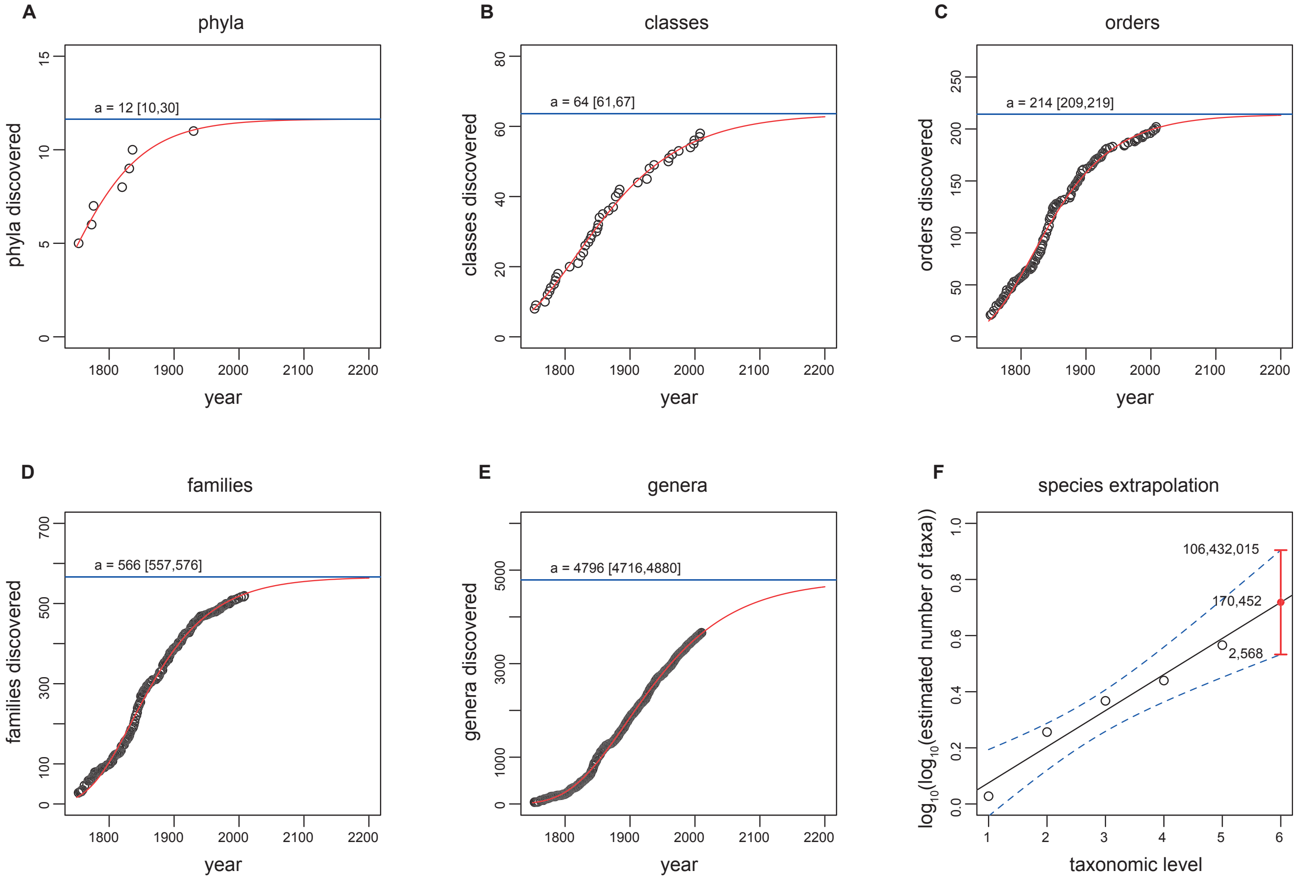

FIG. 3. Temporal accumulation of taxa based on the description of the earliest described element for each taxonomic level (A: phyla, B: classes, C: orders, D: families, E: genera) with indication of the estimated asymptote (with 95% CI). (F) Linear regression between the asymptotic numbers of higher taxa and the numerical hierarchy of each taxonomic rank, with indication of the 95% confidence interval (dashed line) and the estimated global algal diversity (Appendix S1, see Supporting Information).

composition), microscopy (TEM and SEM to visualize ultrastructural details of the cell) and gene sequencing technology (molecular phylogenetics, DNA taxonomy). The importance of the latter is demonstrated by the sudden rise in the number of algal classes from 1990 onward (Fig. S1). The description of new genera follows a persistent linear trend, similar to the species accumulation curve, which indicates that many more algal genera (as well as species) remain to be discovered de novo.

As a general trend, the number of practicing taxonomists increased 4-fold over the ~110 year period from 1900 to 2010 (Fig. 4). Whereas in 1900, barely 50 different phycologists were annually involved in the description of new species, this number rose sharply to nearly 200 by the year 2000. The number of species described per taxonomist shows a number of peaks in the 19th century which result from the taxonomic output of a few extremely productive individuals that each described more than 1,000 species in a very short time, and two that described over 2,000 species (Table 1). Taxonomic output *per capita* decreased slightly during the 20th century, but seems to pick up again in the last two decades.

Apart from the current rates of species description and the size of the phycological community involved in describing new species, an informed estimate of the time needed to describe all species, is critically dependent on the overall magnitude of the algal diversity. The variety of approaches that has been proposed to predict extant species numbers serves as evidence for the complexity of the problem (May 1988, Wilson and Costello 2005, Bouchet 2006, Bebber et al. 2007, Costello and Wilson 2011, Mora et al. 2011). Estimates more often than not vary widely depending on the rationale and the source of data used. Estimates on the magnitude of the global and marine algal diversity were recently published by Guiry (2012) and Appeltans et al. (2012), respectively. These estimates range from 72,500 to 115,000, the estimate by Guiry being the more conservative one. Our own regression analyses of the species and higher taxon data from AlgaeBase serve to illustrate the difficulty in predicting species numbers from discovery curves. A two-step procedure (See Mora et al. 2011) first estimates the number of higher taxa with nonlinear regression (Fig. 3, A–E) and subsequently estimates the number of species by fitting a regression model of the number of taxa against the taxonomic level (Fig. 3G). This yielded an estimate of ~170,000 species, but the uncertainty about the asymptotes in Figure 3, A–F result in a massive 95% confidence interval, ranging from a few thousand to more than a hundred million of species. These statistical confidence limits are obviously unrealistic because 44,000 species have already been described. If the interval is narrowed to incorporate a more realistic lower limit of 60,000, the upper limit comes out at half a million species. A direct extrapolation of the species-level data predicted a total of 68,348 species (Fig. 5). While not too far from Guiry's estimate of 72,500, our result needs to be interpreted with caution because fitting growth curves without information about the decelerating part of the trend is highly unreliable. Furthermore, the current estimate was based on 32,000 of a presumed 44,000 described species. The effect of an additional 12,000 species on the shape of the curve and therefore also on the asymptote is virtually impossible to predict.

A moderately positive picture of algal taxonomy emerges under the assumption of a global estimate of 72,500 species of which 44,000 have already been described. Under this scenario it will take less than 200 years to document the missing species pool at the current description rate of 150 species per year. In reality, of course, it will take longer because description rates will decrease as it becomes more

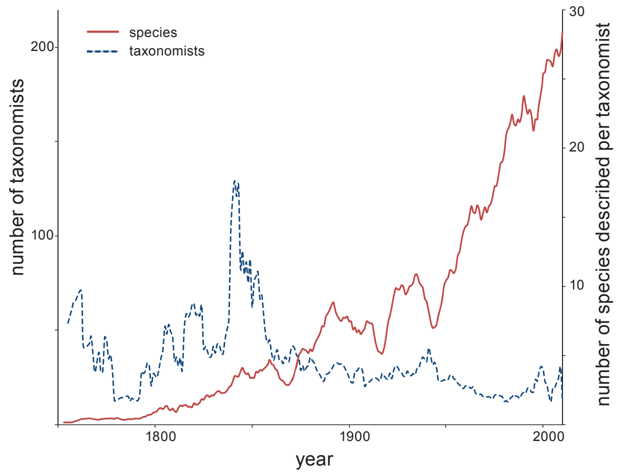

FIG. 4. Temporal trend of the number of algal taxonomists and the number of species described per taxonomist. Data are 5-year moving averages to show trends.

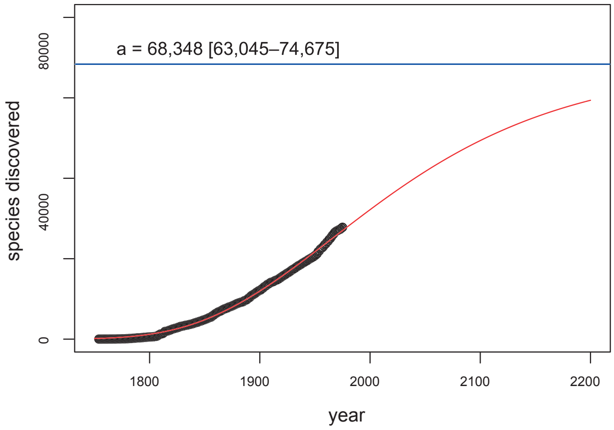

FIG. 5. Estimate of global algal diversity based on the rate species descriptions.

difficult to discover new species. If an estimate of 115,000 or 170,000 algal species is closer to the truth, many more centuries of basic descriptive taxonomic work are ahead of us. Despite the fact that there is considerable uncertainty on the number of algal species, it is also clear that hyper-estimates predicting more than one million algal species (see for example Andersen 1992, John 1994) are unlikely to hold up under the assumptions of our analyses. This reasonably positive picture does not imply, however, that algal taxonomy does not face serious challenges in the near future. Below we discuss some of these issues that urgently need our attention.

SPECIES DISCOVERY

Sequencing of target genes in individual organisms or more recently by environmental sequencing (e.g., Rusch et al. 2007) has revolutionized algal systematics at every taxonomic level. The dramatic increase in genetic data generation has propelled the birth of DNA taxonomy, in which species are delimited based on sequence data using evolutionary species concepts (Vogler and Monaghan 2007) as well as DNA barcoding, which identifies specimens based on sequence similarity against a database of sequences of a priori defined species (Hebert et al. 2003). Both DNA taxonomy and barcoding have been enthusiastically adopted by phycologists (e.g., Saunders 2005, Robba et al. 2006, Lane et al. 2007, Guillemin et al. 2008, Sherwood et al. 2008, McDevit and Saunders 2009, Gile et al. 2010, Le Gall and Saunders 2010, Moniz and Kaczmarska 2010, Trobajo et al. 2010, Hamsher et al. 2011) and the use of molecular markers has become the rule rather than the exception in assessments of algal species-level diversity (Medlin et al. 2007, Cianciola et al. 2010). The enthusiasm with which DNA sequence data have been embraced has resulted in an explosion of sequences in repositories such as GenBank and BOLD Systems (www.barcodinglife.com). An increasing number of these sequences, however, are not linked to a full and proper species name (i.e., a Latin binomial consisting of a genus name and specific epithet). Examining the sequence data deposited in Genbank, we found that the percentage of sequences of properly named algae decreased gradually from ~90% in 1993 to <20% in 2010 (Fig. 6; see also Verbruggen 2012).

The spectacular increase in the number of sequenced specimens that have not been linked to existing species or that have not been formally described as new species, coined "dark taxa" by Page (2012), is the result of two interlinked factors. First, describing species has remained an inherently time-consuming process in stark contrast to the dizzying speed at which sequences can be generated nowadays and hence diversity can be discovered. Second, genetic data have confronted biologists with the fact that morphology does not always adequately reflect species boundaries. This mismatch between genetic diversity and morphology, although observed in all groups of organisms (Bickford et al. 2007), becomes more of an issue when organisms decrease in morphological complexity and size (e.g., Whittaker et al. 2005, Verbruggen et al. 2009). Indeed, levels of cryptic diversity reach such proportions in some algal groups that researchers become deterred from formally describing the observed diversity (e.g., several diatom and green algal genera; (Denboh et al. 2003, Beszteri et al. 2007, McManus and Lewis 2011, Souffreau et al. 2012); *Symbiodinium* coral endosymbionts (LaJeunesse 2001); the red algal genus *Portieria*, (Payo et al. in press). Cryptic diversity, however, has another side effect that reflects on the growing proportion of dark taxa. In situations where cryptic diversity is rife and morphology is therefore by definition inadequate to properly identify specimens, it is virtually impossible to link specimens to existing names with something that represents a little more than educated guesswork. It is the uncertainty that coincides with linking gene sequences to published names, that constrains taxonomists from doing so.

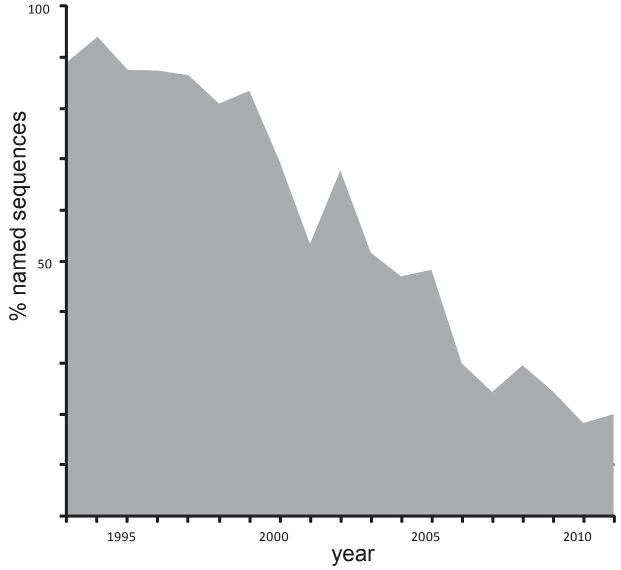

FIG. 6. Temporal trend of the percentage of algal sequences with Latin binomials in GenBank

TOWARD A FUTURE WITHOUT NAMES?

> *Ultimately scientific names are only human readable strings for referring to biological taxa* – R. Hyam (2009)

The increasing number of articles that fail to link the discovered diversity to formal taxonomic names, results in an interesting, but highly controversial discussion that even divides the authors of this article up to a certain level: should all species have a name and should it be a Latin binomial? Page (2010) argues that abandoning the traditional name system does not necessarily mean the end of biology. Indeed, microbiology seems to be doing fine as a discipline even though only a mere 5,000 species have been formally described (Pace 1997) and <1% of all prokaryotic sequences submitted in GenBank is linked to a formal name (Page 2012). Strictly speaking, formal Latin binomials are not necessary for a lot of science; however, political conventions such as the Convention on International Trade in Endangered Species of Wild Fauna and Flora and the Red lists of the International Union for the Conservation of Nature use scientific names as currency in framing their conservation efforts.

Nevertheless, we do not argue here for or against Latin binomials, but we seek merely to emphasize their functionality. Traditional binomials are formalized identifiers to which biological or other information is attached. The importance of the identifier is that it unequivocally refers to the correct species and despite the traditional attachment to Latin, the form of the identifier is subordinate to its function. The observed levels of cryptic diversity in algae combined with the shift in methods used to characterize them has resulted in the rampant uncertainty about the nature of many older species, and the names of such old species are becoming increasingly ineffective at fulfilling their "unique identifier" function. As a consequence, and similar to what happened in bacteriology several decades ago, there is a tendency in phycology to gradually move away from species-based identifiers to a more informal system whereby clade-, specimen- or strain-based identifiers are used to communicate biological information. Examples are rife in studies focusing not only on microbial diversity (Rodriguez et al. 2005, Viprey et al. 2008, Worden et al. 2009, Blanc et al. 2010) and zooxanthellae (LaJeunesse 2001, Pinzón and LaJeunesse 2010, LaJeunesse and Thornhill 2011), but also in several seaweed genera informal clade names have been used for more than a decade (Van Der Strate et al. 2002, Zuccarello and West 2003, Leliaert et al. 2009, Gutner-Hoch and Fine 2011, Kamiya et al. 2011). Whether these informal names for species-level clades present a temporary situation stimulated by the discrepancy in the rates of species discovery and formal description, or an incipient alternative or parallel taxonomy, will be largely determined by how well we manage the past.

THE DEAD WEIGHT OF THE PAST

> *In some taxonomic groups the past acts as a dead weight on the subject, the complex synonymy and scattered type material deterring anyone from attempting a modern revision* – Godfray (2002)

The major challenge lies not with future descriptions of taxa: there is no doubt the field of taxonomy is versatile enough to be able to incorporate genomic or other integrative data. An increasing number of diagnoses of new species already explicitly refer to GenBank accession numbers of representative gene sequences and formal protocols describing how to integrate barcodes in diagnoses or descriptions of new species have been proposed (Evans and Mann 2009). We are of the opinion that all species descriptions must be accompanied by representative DNA reference sequences of the type specimen, and recommend storing a DNA extract of the type specimen in DNA banks, specialized cryogenic storage facilities (Hodkinson et al. 2007). The requirement for DNA sequences in the diagnosis or description of new species and other taxa is already being informally implemented by many phycological journals. However, many species are still described based on morphology alone, often on the basis of a single collection, especially in diatoms.

The major challenge of algal taxonomy in the 21st century is presented by the interpretation of what has been described in the past, in attempting to match Linnaean names with DNA sequences (Tautz et al. 2003). Ideally, one should sequence the type of every single species, currently accepted species as well as the synonyms and species of uncertain status. It is, however, wishful thinking to assume that DNA information can be obtained from most type specimens, either because types are inadequate, or unavailable, or because they cannot be used for DNA extraction (e.g., formalin-preserved specimens, microscope slides, oxidized diatom frustules, certain fumigated herbaria, drawings; Saunders and McDevit 2012). For such cases Tautz et al. (2003) proposed a modus operandi whereby experienced taxonomists identify newly collected specimens that could be used for DNA extraction. The policy of the ICN (Art. 9.7) to select epitypes where the holotype, lectotype or even a previously designated neotype is "…demonstrably ambiguous and cannot be critically identified for purposes of the precise application of the name of a taxon" is entirely fit for such a purpose. The magnitude of the problem is, of course, exacerbated in phycology by the sheer volume of synonyms and names of uncertain affinity, which outnumber the currently accepted species by a factor of 3–5 to 1, and the fact that many algal types simply have been destroyed, or have not been located or even designated (Silva 2008). The seriousness of the problem of heterotypic synonyms has been recognized by higher plant taxonomists who, in a huge collaborative effort, clarified synonymies for approximately 110,000 flowering plants (Scheffers et al. 2012). There is no doubt sorting out the past represents a daunting task, but with piles of unlocated or uninformative types, heaps of DNA clusters that need a name and a limited amount of time available, phycologists will need to make a number of important decisions.

There is a crucial role for collections and herbaria in bridging this gap between species discovery and Linnaean binomials (and infraspecific names) by locating type material and making it available online and for sequencing. A preliminary web search shows that depressingly little information on algal types can be retrieved online (with a few notable exceptions such as the National Herbarium Nederland, the Linnaean herbarium in London, the Smithsonian Institution, and the type specimens from the Galapagos Islands at the University Herbarium of Berkeley). Additionally, technological progress to increase both the speed and the success rate to generate representative DNA sequences from type specimens is much needed (Särkinen et al. 2012, Saunders and McDevit 2012). Target enrichment methods in combination with high-throughput sequencing (e.g., Stiller et al. 2009) offer the potential to boost the integration of historical collections in present-day modern systematic research. Although these techniques are increasingly applied to generate genomic information from paleontological samples (Rizzi et al. 2012), to date they have not seen wide application in taxonomic research (see Mason et al. 2011). In addition, herbaria are an extraordinary resource for the discovery of new species. Half of the species of flowering plants, estimated still to be described, have probably been collected and stored in herbaria (Bebber et al. 2010) and there is no reason to assume this would be different for algae.

FUTURE MANAGEMENT OF ALGAL DIVERSITY

Regardless of whether the Linnaean classification system will hold up in the future or whether it will gradually fade away and will become replaced by some alternative system, there is a crucial role for an organizational framework that is able to manage the information about algal species. Databases are critical to the future of algal diversity management, but the level of investment in them has been very low. The maintenance of nomenclatural and taxonomic databases, each with their own merits, has been without exception the task of a handful of dedicated individuals. Repositories may not survive the demise of these people without considerable international effort on the part of the scientific community and those that fund biodiversity research. There are considerable difficulties in integrating the various databases, not least the politics, both national and international, but one may consider this an open invitation to reflect on how we will manage our algal diversity information in the future. An eventual strategy should preferably come out of an international working group that includes the databases as well as representatives of the various phycological societies. In this we should link up to major international initiatives that are currently being developed. Mycologists, for example, as an additional criterion for the valid publication of new taxa and taxonomic changes will need to cite a unique identifier, issued by one of three registration databases, MycoBank, *Index Fungorum* and Fungal Name (a Chinese repository) in the protologue (Crous et al. 2004). Importantly, each taxon will be assigned a unique identifier which serves to link databases of individual specimens with critical metadata including geographical distribution, DNA, ecology, and morphology. AlgaeBase is the only algal database currently providing Life Science Identifiers (LSIDs), unique identifiers for names at species level (and above), through the LSID Web Resolver using Taxonomic Database Working Group standards (http://www.tdwg.org). A Special Committee on the Registration of Algal and Plant names (including fossils) set up by the International Botanical Congress is presently looking at extending registration to all plants, an initiative that would be much valued in phycology.

Although the compulsory registration of names of new taxa and other nomenclatural innovations represents only one facet to improve the efficiency of taxonomy, it provides a mechanism that would help solve problems encountered in determining which names are effectively published, reducing the publication of invalid taxa, and assuring that data are available online to the entire scientific community immediately on publication. Registration should obviously be extended to the large number of previously published names, the current status of which is often uncertain. These efforts, in combination with a better integration of historical collections in systematic research, should result in a coherent and stable taxonomic framework in which names can fulfill their "unique identifier" function again.

We are very grateful to Ward Appeltans, Wendy and Caoilte Guiry, Paul Kirk, Lennert Tyberghein, Bart Vanhoorne, and Pieter Vanormelingen for assistance, valuable information, and useful discussions. Likewise, we are grateful to Richard Moe and Paul Silva for the *INA* resource which was critical to provide description dates for taxa of uncertain status. AlgaeBase was funded by the Programme for Research in Third-level Institutions (PRTLI 2, 3 and 4 to the Ryan Institute and the Environmental Change Institute) funded by the Department of Education and Science in Ireland, by Atlantic Philanthropies, and by the European Union. HV is funded by the Australian Research Council (FT110100585). FL is indebted to the FWO-Vlaanderen for a grant as postdoctoral researcher. Many thanks to Craig Schneider and an anonymous, but enthusiastic referee for their reviews.

Andersen, R. A. 1992. Diversity of eukaryotic algae. *Biodivers. Conserv.* 1:267–92.

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., Bamber, R. et al. 2012. The magnitude of global marine species diversity. *Curr. Biol.* 23:2189–202.

Bacher, S. 2012. Still not enough taxonomists: reply to Joppa et al. *Trends Ecol. Evol.* 27:65–6.

Bebber, D. P., Carine, M. A., Wood, J. R. I., Wortley, A. H., Harris, D. J., Prance, G. T., Davidse, G., Paige, J., Pennington, T. D., Robson, N. K. B. & Scotland, R. W. 2010. Herbaria are a major frontier for species discovery. *Proc. Natl Acad. Sci. USA* 107:22169–71.

Bebber, D. P., Marriott, F. H. C., Gaston, K. J., Harris, S. A. & Scotland, R. W. 2007. Predicting unknown species numbers using discovery curves. *Proc. R. Soc. Biol. Sci. Ser. B* 274:1651–8.

Beszteri, B., John, U. & Medlin, L. K. 2007. An assessment of cryptic genetic diversity within the *Cyclotella meneghiniana* species complex (Bacillariophyta) based on nuclear and plastid genes, and amplified fragment length polymorphisms. *Eur. J. Phycol.* 42:47–60.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K. & Das, I. 2007. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* 22:148–55.

Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., Lindquist, E. et al. 2010. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 22:2943–55.

Boero, F. 2010. The study of species in the era of biodiversity: a tale of stupidity. *Diversity* 2:115–26.

Bouchet, P. 2006. The magnitude of marine biodiversity. *In* Duarte, C. M. [Ed.] *The Exploration of Marine Biodiversity: Scientific and Technological Challenges*. Fundación BBVA, Bilbao, Spain, pp. 31–64.

Bryson, B. 2003. *A Short History of Nearly Everything*. Broadway Books, New York, 298 pp.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A. et al. 2012. Biodiversity loss and its impact on humanity. *Nature* 486:59–67.

Cianciola, E. N., Popolizio, T., Schneider, C. W. & Lane, C. E. 2010. Using molecular-assisted alpha taxonomy to better understand red algal biodiversity in Bermuda. *Diversity* 2:946–58.

Clark, B. R., Godfray, H. C. J., Kitching, I. J., Mayo, S. J. & Scoble, M. J. 2009. Taxonomy as an eScience. *Phil. Trans. R. Soc. Ser. A* 367:953–66.

Costello, M. J. & Wilson, S. P. 2011. Predicting the number of known and unknown species in European seas using rates of description. *Glob. Ecol. Biogeogr.* 20:319–30.

Costello, M. J., Wilson, S. & Houlding, B. 2012. Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Syst. Biol.* 61:871–83.

Crous, P. W., Gams, W., Stalpers, J. A., Robert, V. & Stegehuis, G. 2004. MycoBank: an online initiative to launch mycology into the 21st century. *Stud. Mycol.* 50:19–22.

Deans, A. R., Yoder, M. J. & Balhoff, J. P. 2012. Time to change how we describe biodiversity. *Trends Ecol. Evol.* 27:78–84.

Denboh, T., Ichimura, T., Hendrayanti, D. & Coleman, A. W. 2003. *Closterium moniliferum-ehrenbergii* (Charophyceae, Chlorophyta) species complex viewed from the 1506 group I intron and ITS2 of nuclear rDNA. *J. Phycol.* 39:960–77.

Evans, K. M. & Mann, D. G. 2009. A proposed protocol for nomenclaturally effective DNA barcoding of microalgae. *Phycologia* 48:70–4.

Fourtanier, E. & Kociolek, J. P. 2012. Catalogue of diatom names. California Academy of Sciences. Available at http://research.calacademy.org/research/diatoms/names/index.asp (last accessed 20 August 2012).

Gile, G. H., Stern, R. F., James, E. R. & Keeling, P. J. 2010. DNA barcoding of chlorarachniophytes using nucleomorph ITS sequences. *J. Phycol.* 46:743–50.

Godfray, H. C. J. 2002. Challenges for taxonomy - The discipline will have to reinvent itself if it is to survive and flourish. *Nature* 417:17–19.

Godfray, H. C. J., Clark, B. R., Kitching, I. J., Mayo, S. J. & Scoble, M. J. 2007. The Web and the structure of taxonomy. *Syst. Biol.* 56:943–55.

Godfray, H. C. J. & Knapp, S. 2004. Taxonomy for the twenty-first century – introduction. *Phil. Trans. R. Soc. Ser. B* 359:559–69.

Goldstein, P. Z. & DeSalle, R. 2011. Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *BioEssays* 33:135–47.

Guillemin, M. L., Akki, S. A., Givernaud, T., Mouradi, A., Valero, M. & Destombe, C. 2008. Molecular characterisation and development of rapid molecular methods to identify species of Gracilariaceae from the Atlantic coast of Morocco. *Aquat. Bot.* 89:324–30.

Guiry, M. D. 2012. How many species of algae are there? *J. Phycol.* 48:1057–63.

Guiry, M. D. & Guiry, G. M. 2012. AlgaeBase. World-wide electronic publication. Available at http://www.algaebase.org (last accessed 30 June 2012).

Gutner-Hoch, E. & Fine, M. 2011. Genotypic diversity and distribution of *Ostreobium quekettii* within scleractinian corals. *Coral Reefs* 30:643–50.

Hamsher, S. E., Evans, K. M., Mann, D. G., Poulickova, A. & Saunders, G. W. 2011. Barcoding diatoms: exploring alternatives to COI-5P. *Protist* 162:405–22.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. Biol. Sci. Ser. B* 270:313–21.

Hodkinson, T. R., Waldren, S., Parnell, J. A. N., Kelleher, C. T., Salamin, K. & Salamin, N. 2007. DNA banking for plant breeding, biotechnology and biodiversity evaluation. *J. Plant. Res.* 120:17–29.

Hyam, R. 2009 Calling time on biological nomenclature. Available at http://www.hyam.net/blog/archives/509 (last accessed 4 August 2009).

Jahn, R. & Kusber, W.-H. 2012 AlgaTerra information system. Available at http://www.algaterra.org (last accessed 15 August 2012).

John, D. M. 1994. Biodiversity and conservation: algal perspective. *The Phycologist* 38:3–21.

Joppa, L. N., Roberts, D. L. & Pimm, S. L. 2011. The population ecology and social behaviour of taxonomists. *Trends Ecol. Evol.* 26:551–3.

Kamiya, M., West, J. A. & Hara, Y. 2011. Induction of apomixis by outcrossing between genetically divergent entities of *Caloglossa leprieurii* (Ceramiales, Rhodophyta) and evidence of hybrid apomicts in nature. *J. Phycol.* 47:753–62.

LaJeunesse, T. C. 2001. Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the its region: In search of a "species" level marker. *J. Phycol.* 37:866–80.

LaJeunesse, T. C. & Thornhill, D. J. 2011. Improved resolution of reef-coral endosymbiont (*Symbiodinium*) species diversity, ecology, and evolution through *psbA* non-coding region genotyping. *PLoS ONE* 6:e29013.

Lane, C. E., Lindstrom, S. C. & Saunders, G. W. 2007. A molecular assessment of northeast Pacific *Alaria* species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding. *Mol. Phylogenet. Evol.* 44:634–48.

Le Gall, L. & Saunders, G. W. 2010. DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora. *J. Phycol.* 46:374–89.

Leliaert, F., Verbruggen, H., Wysor, B. & De Clerck, O. 2009. DNA taxonomy in morphologically plastic taxa: Algorithmic species delimitation in the *Boodlea* complex (Chlorophyta: Cladophorales). *Mol. Phylogenet. Evol.* 53:122–33.

Maddison, D. R., Guralnick, R., Hill, A., Reysenbach, A. L. & McDade, L. A. 2012. Ramping up biodiversity discovery via online quantum contributions. *Trends Ecol. Evol.* 27:72–7.

Marin, B. & Melkonian, M. 2010. Molecular phylogeny and classification of the mamiellophyceae class. nov. (Chlorophyta) based on sequence comparisons of the nuclear- and plastid-encoded rRNA operons. *Protist* 161:304–36.

Mason, V. C., Li, G., Helgen, K. M. & Murphy, W. J. 2011. Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Res.* 21:1695–704.

May, R. 1988. How many species are there on earth. *Science* 241:1441–9.

McDevit, D. C. & Saunders, G. W. 2009. On the utility of DNA barcoding for species differentiation among brown macroalgae (Phaeophyceae) including a novel extraction protocol. *Phycol. Res.* 57:131–41.

McManus, H. A. & Lewis, L. A. 2011. Molecular phylogenetic relationships in the freshwater family Hydrodictyaceae (Sphaeropleales, Chlorophyceae), with an emphasis on *Pediastrum duplex*. *J. Phycol.* 47:152–63.

Mcneill, J., Barrie, F. R., Buck, W. R., Demoulin, V., Greuter, W., Hawkworth, D. L., Herendeen, P. S. et al. 2012. *International Code of Nomenclature for Algae, Fungi and Plants (Melbourne Code) Adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011.* Koeltz, Koenigstein, Germany, 232 pp.

Medlin, L. K., Metfies, K., John, U. & Olsen, J. L. 2007. Algal molecular systematics: a review of the past and prospects for the future. *In* Brodie, J. & Lewis, J. [Eds.] *Unravelling the Algae: The Past, Present, and Future of Algal Systematics.* Taylor and Francis, Boca Raton, Florida, pp. 341–53.

Moniz, M. B. J. & Kaczmarska, I. 2010. Barcoding of diatoms: nuclear encoded ITS revisited. *Protist* 161:7–34.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B. & Worm, B. 2011. How many species are there on earth and in the ocean? *PLoS Biol.* 9:e1001127.

Pace, N. R. 1997. A molecular view of microbial diversity and the biosphere. *Science* 276:734–40.

Page, R. D. M. 2010. Are names really the key to the big new biology? Available at http://iphylo.blogspot.be/2010/10/are-names-really-key-to-big-new-biology.html (last accessed April 12).

Page, R. D. M. 2012. Dark taxa: GenBank in a post-taxonomic world. Available at http://iphylo.blogspot.be/2011/04/dark-taxa-genbank-in-post-taxonomic.html (last accessed April 12).

Patterson, D. J., Cooper, J., Kirk, P. M., Pyle, R. L. & Remsen, D. P. 2010. Names are key to the big new biology. *Trends Ecol. Evol.* 25:686–91.

Payo, D. A., Colo, J., Calumpong, H. & De Clerck, O. 2011. Variability of non-polar secondary metabolites in the red alga *Portieria*. *Marine Drugs* 9:2438–68.

Payo, D. A., Leliaert, F., Verbruggen, H., D'Hondt, S., Calumpong, H. & De Clerck, O. in press. Extensive cryptic diversity and fine scale endemism of a marine red alga in the Philippines. *Proc. R. Soc. Biol. Sci. Ser. B.* doi:10.1098/rspb.2012.2660

Pinzón, J. H. & LaJeunesse, T. C. 2010. Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. *Mol. Ecol.* 20:311–25.

Rizzi, E., Lari, M., Gigli, E., De Bellis, G. & Caramelli, D. 2012. Ancient DNA studies: new perspectives on old samples. *Gen. Sel. Evol.* 44:19.

Robba, L., Russell, S. J., Barker, G. L. & Brodie, J. 2006. Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). *Am. J. Bot.* 93:1101–8.

Rodriguez, F., Derelle, E., Guillou, L., Le Gall, F., Vaulot, D. & Moreau, H. 2005. Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). *Environ. Microbiol.* 7:853–9.

Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., Wu, D. Y. et al. 2007. The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol.* 5:398–431.

Samyn, Y. & De Clerck, O. 2012. No name, no game. *Eur. J. Taxonomy* 10:1–3.

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S. & Bakker, F. T. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE* 7:e43808.

Saunders, G. W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Phil. Trans. R. Soc. Ser. B* 360:1879–88.

Saunders, G. W. & McDevit, D. C. 2012. Acquiring DNA sequence data from dried archival red algae (Florideophyceae) for the purpose of applying available names to contemporary genetic species: a critical assessment. *Botany-Botanique* 90:191–203.

Scheffers, B. R., Joppa, L. N., Pimm, S. L. & Laurance, W. F. 2012. What we know and don't know about Earth's missing biodiversity. *Trends Ecol. Evol.* 9:501–10.

Scherffel, A. 1912. Zwei neue, trichocystenartige Bildungen führende Flagellaten. *Arch. Protistenk.* 27:94–128.

Sherwood, A. R., Vis, M. L., Entwisle, T. J., Necchi, O. J. & Presting, G. G. 2008. Contrasting intra versus interspecies DNA sequence variation for representatives of the Batrachospermales (Rhodophyta): Insights from a DNA barcoding approach. *Phycol. Res.* 56:269–79.

Silva, P. C. 2008. Historical review of attempts to decrease subjectivity in species identification, with particular regard to algae. *Protist* 159:153–61.

Silva, P. C. & Moe, R. L. 1999. The index nominum algarum. *Taxon* 48:351–3.

Smith, V. S., Rycroft, S. D., Harman, K. T., Scott, B. & Roberts, D. 2009. Scratchpads: a data-publishing framework to build, share and manage information on the diversity of life. *BMC Bioinformatics* 10: S6.

Souffreau, C., Vanormelingen, P., Van de Vijver, B., Isheva, T., Verleyen, E., Sabbe, K. & Vyverman, W. 2012. Molecular evidence for distinct antarctic lineages in the cosmopolitan terrestrial diatoms *Pinnularia borealis* and *Hantzschia amphioxys*. *Protist* doi: http://dx.doi.org/10.1016/j.protis.2012.04.001. in press.

Stiller, M., Knapp, M., Stenzel, U., Hofreiter, M. & Meyer, M. 2009. Direct multiplex sequencing (DMPS)-a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Res.* 19:1843–8.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003. A plea for DNA taxonomy. *Trends Ecol. Evol.* 18:70–4.

Thessen, A. E., Patterson, D. J. & Murray, S. A. 2012. The taxonomic significance of species that have only been observed once: the genus *Gymnodinium* (Dinoflagellata) as an example. *PLoS ONE* 7:e44015.

Trobajo, R., Mann, D. G., Clavero, E., Evans, K. M., Vanormelingen, P. & McGregor, R. C. 2010. The use of partial *cox*1, *rbc*L and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). *Eur. J. Phycol.* 45:413–25.

Van Der Strate, H. J., Boele-Bos, S. A., Olsen, J. L., Van De Zande, L. & Stam, W. T. 2002. Phylogeographic studies in the tropical seaweed *Cladophoropsis membranacea* (Chlorophyta, Ulvophyceae) reveal a cryptic species complex. *J. Phycol.* 38:572–82.

Verbruggen, H. 2012. No name, new game. Available at http://phycoweb.wordpress.com/2012/04/10/no-name-new-game/ (last accessed 10 April 2012).

Verbruggen, H., Vlaeminck, C., Sauvage, T., Sherwood, A. R., Leliaert, F. & De Clerck, O. 2009. Phylogenetic analysis of *Pseudochlorodesmis* strains reveals cryptic diversity above the family level in the siphonous green algae (Bryopsidales, Chlorophyta). *J. Phycol.* 45:726–31.

Viprey, M., Guillou, L., Ferreol, M. & Vaulot, D. 2008. Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environ. Microbiol.* 10:1804–22.

Vogler, A. P. & Monaghan, M. T. 2007. Recent advances in DNA taxonomy. *J. Zool. Syst. Evol. Res.* 45:1–10.

Wheeler, Q. D. 2004. Taxonomic triage and the poverty of phylogeny. *Phil. Trans. R. Soc. Ser. B* 359:571–83.

Wheeler, Q. D. 2008. *The new taxonomy.* CRC Press, Boca Raton.

Whittaker, R. J., Araujo, M. B., Paul, J., Ladle, R. J., Watson, J. E. M. & Willis, K. J. 2005. Conservation Biogeography: assessment and prospect. *Divers. Distrib.* 11:3–23.

Wilson, E. O. 1992. *The diversity of life.* Penguin Books, London.

Wilson, S. P. & Costello, M. J. 2005. Predicting future discoveries of European marine species by using a non-homogeneous renewal process. *J. R. Statist. Soc. Ser. C.* 54:897–918.

Worden, A. Z., Lee, J. H., Mock, T., Rouze, P., Simmons, M. P., Aerts, A. L., Allen, A. E. et al. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas. *Science* 324:268–72.

Zuccarello, G. C. & West, J. A. 2003. Multiple cryptic species: molecular diversity and reproductive isolation in the *Bostrychia radicans/B. moritziana* complex (Rhodomelaceae, Rhodophyta) with focus on North American isolates. *J. Phycol.* 39:948–59.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. Rescaled temporal accumulation of taxa based on the description of the earliest described element for each taxonomic level.

Appendix S1. Data sources and methodological details.