

## A Return to Linnaeus's Focus on Diagnosis, Not Description: The Use of DNA Characters in the Formal Naming of Species

SUSANNE S. RENNER\*

*Systematic Botany and Mycology, University of Munich (LMU), Menzinger Street 67, 80638 Munich, Germany*

\*Correspondence to be sent to: *Systematic Botany and Mycology, University of Munich (LMU), Menzinger Street 67, 80638 Munich, Germany;*  
E-mail: renner@lrz.uni-muenchen.de

Received 12 September 2015; reviews returned 2 April 2016; accepted 5 April 2016  
Associate Editor: Tiffani Williams

**Abstract.**—Descriptions and diagnoses are alternative choices in all *Codes of Nomenclature* because Linnaeus relied on diagnoses, not descriptions, to name ca. 13,400 animals, plants, and fungi. A diagnosis names characters in which a new taxon differs from the most similar known taxon; a description mixes taxonomically informative and uninformative features, usually without indicating which is which. The first formal diagnoses of new taxa that included DNA-based characters came out in 2001, and by November 2015, at least 98 names of species of acoels, lichens, angiosperms, annelids, alveolates, arachnids, centipedes, turtles, fishes, butterflies, mollusks, nematodes, and pathogenic fungi have been published based on diagnostic mitochondrial, plastid, or nuclear DNA substitutions, indels, or rarely genetic distances, with or without additional morphological features. Authors have found diverse ways to specify the diagnostic traits (all published studies are here tabulated). While descriptions try to “cover” within-species variation, a goal rarely accomplished because of (i) the stochastic nature of specimen availability (thousands of species are known from single collections) and (ii) the subjective circumscription of species, the purpose of diagnoses was and is speedy identification. Linnaeus tried to achieve this by citing images, geographic occurrence, and previous literature. The renewed attention to sharp diagnosis now coincides with worldwide barcoding efforts, may speed up formal naming, and matches the increasing reliance on DNA for both classification and identification. I argue for DNA-based diagnoses of new species becoming a recommendation in all *Codes*, not just the bacterial code. [Codes of Nomenclature; description; diagnosis; DNA-based diagnosis; naming new species; nomenclature.]

“Descriptions cannot be made full enough and accurate enough to satisfy later workers. Each generation of taxonomists must see the actual specimens used by earlier generations, and I think the tendency now is, or should be, to make descriptions short, but of course explicit and carefully calculated, and to make specimens widely available.”

P.J. Darlington, Jr., 1971, p. 146

The naming of organisms following standardized conventions is the basis for linking new information to existing knowledge. It is also the basis for biological classification, effective communication, and extrapolation of findings about organisms. The mere accession numbers of DNA sequences (or other strings of numbers lacking an agreed system of the numbers' innate significance) do not permit extrapolation of information about morphological traits, biogeographic ranges, or sharing of published knowledge across disciplines, all of which is possible with a widely used naming convention. Most researchers are using the conventions of the Linnaean system, with the fixed starting points being Linnaeus's treatments of plants and animals (1753, 1758; Persoon and Fries for certain fungi and the names of lichens).

Since about 2000, taxonomists have increasingly tried to combine morphological and molecular data for detecting and delimiting species (reviewed in Wheeler 2008; Begerow et al. 2010; Hibbett and Taylor 2013; Vences et al. 2013), and since 2003, DNA barcoding has become the method of choice for reliable identification, at least for insects, certain fungi, tropical trees, and many aquatic organisms (Hebert et al. 2003; Koljalg et al. 2013; Hausmann et al. 2013; Kress et al. 2015). Surprisingly, however, DNA characters have rarely been used in the formal description of species (Cook et al. 2010). Of 310 barcoding publications surveyed by Kress et al. (2015) that led to the discovery of new species, only one (Félix et al. 2014; Table 1) used DNA traits in species protologs.

To date, two papers have discussed DNA-based formal diagnoses (Cook et al. 2010; Tripp and Lendemer 2014). Both overlooked that the practice began 15 years ago (Westheide and Hass-Cordes 2001), and they either focused on a hypothetical example (Cook et al. 2010) or examples from 2012 and 2013 (Tripp and Lendemer 2014). Tripp and Lendemer (2012) also raised a potential problem with one type of DNA diagnosis, namely genetic distances, which I take up in the “Discussion” section. No previous paper has surveyed the conceptual and factual history of DNA-based formal naming, and the absence of a review of how taxonomists have

TABLE 1. All studies published so far that have included DNA-based diagnoses in species protocols, that is, for the formal diagnosis of type material

Taxon	Number and geographic region of new species	Type of molecular diagnosis	Morphological description (yes/no); Morphological diagnosis (yes/no)	Deposition of type material (not always the specimen from which DNA was isolated)	Deposition of sequences or alignment(s)	References
Acoelomorpha (Platyhelminthes) Nemertodermatida, Nemertodermatidae	9 species of <i>Nemeritoides</i> , worldwide in marine mesopsammon	Specified substitutions in LSU and SSU rRNA and histone 3	Yes Yes	Type material in a public collection	Sequences in GenBank	Meyer-Wachsmuth et al. 2014
Alveolata Dinophyceae	5 species of <i>Alexandrium</i>	"The combined nucleotide sequences of the holotype strain .... The complete list of diagnostic D1-D2 LSU, ITS (Internal Transcribed Spacer)/5.8S and SSU Gen-Bank sequences which can be used as a genetic type for this species are reported in Supplementary Material."	Yes Yes	Holo- and isotypes on SEM stubs and in formol in Senckenberg herbarium (FR)	Sequences in GenBank	John et al. 2014
Angiospermae Solanaceae	1 species of <i>Bryntofelia</i> from Brazil	Specified substitutions in the plastid <i>rnh1F</i> gene and in nuclear ITS	Yes No	Type material in various herbaria	Sequences in GenBank	Filipowicz et al. 2012
Angiospermae Buxaceae	3 species of <i>Buxus</i> from Cuba	Specified substitutions in the plastid <i>matK</i> gene and in coding region of <i>trnL</i> spacer	Yes Yes	Type material in various herbaria	Sequences in GenBank	González Gutiérrez et al. 2013
Angiospermae Ehretiaceae	1 species of <i>Rochefortia</i> from Lesser Antilles	Specified substitutions in the nuclear ITS1 region	Yes Yes	Type material in various herbaria	Sequences in GenBank	Irimia and Gottschling 2016

(continued)

TABLE 1. Continued

Annelida, Syllidae Bristle worms	1 species of <i>Petitita</i> Seychelles	Specified substitutions in nuclear ITS and 8 diagnostic RFLP (Restriction Fragment Length Polymorphism) fragments with specified primers	Yes No	Holotype "used up," three syntypes in Senckenberg Museum	Several nuclear ITS seqs. in GenBank	Westheide and Hass-Cordes 2001
Annelida, Serpulidae Tube worms	1 species of <i>Galeolaria</i> , Australia	Specified substitutions in mtDNA cytochrome <i>b</i> gene and nuclear ITS	Yes No (Description is called diagnosis)	Type material in public collection	Sequences in GenBank	Halt et al. 2009
Arachnida, Opiliones, Cyphophthalmi: Neogoveidae	2 species of <i>Metasiro</i> , South Carolina, USA	Specified substitutions in mt COI gene	Yes No	Type material in a public collection	Sequences in GenBank	Clouse and Wheeler 2014
Arthropoda, Chilipoda, Craterostigmidae Centipedes	1 species of <i>Craterostigmus</i> , New Zealand	Specified substitutions in nuclear 18S and 28S rRNA, mitochondrial (mt) 16S rRNA, and the protein-encoding cytochrome <i>c</i> oxidase subunit I (COI)	Yes Yes	Type material in a public collections (including DNA voucher from holotype)	Sequences in GenBank	Edgecombe and Giribet 2008
Ascomycota, Lecanoraceae Lichens	5 species of <i>Rhizoplaca</i>	Unspecified substitutions in nuclear ITS <i>"Rhizoplaca</i> <i>polymorpha</i> consists of specimens recovered within 'clade IVc' in Leavitt et al. (2011a), which is supported as a lineage distinct from all other populations according to coalescent-based genetic analysis of multiple genetic loci."	No! No	Type material in public collection and MycoBank	Sequences in GenBank	Leavitt et al. 2013
Ascomycota, Lecanoraceae Lichens	1 species of <i>Leparia</i>	Specified substitutions in nuclear ITS (in Latin)	Yes No	Type material in public collections	Sequences in GenBank	Lendemer 2011
Ascomycota, Parmeliaceae Lichens	1 species of <i>Parmelia</i> Europe	Specified substitutions in a group I intron and nuclear ITS (in Latin)	No No	Type material in public collections	Sequences in GenBank	Molina et al. 2011

(continued)

TABLE 1. Continued

Chordata, Pisces Ictaluridae	1 species of <i>Noturus</i>	Specified substitutions in cyt b and RAG2 genes (listed in a table)	Yes No (Description is called diagnosis)	Type material in various collections	Sequences in GenBank	Egge and Simons 2006
Chordata, Reptilia, Pelomedusidae Turtles	6 species of <i>Pelomedusa</i>	Specified substitutions and indels in nuclear 12S rDNA	Yes No (Description is called diagnosis) No No	Type material in various collections	12S rDNA sequences in GenBank	Petzold et al. 2014
Fungi Onygenales, Coccidioidomycosis A dimorphic pathogenic fungus	1 species of <i>Coccidioides</i>	Specified substitutions in nuclear Chitin synthase gene; two microsatellite differences (in Latin)	No No	American- type culture collection #28868; killed sample in Jepson Herbarium, CA	Not cited, unclear	Fisher et al. 2002
Fungi Neocallimastigales	1 species of <i>Pirromyces</i>	"The least inclusive clade containing organisms with nuclear rRNA ITS sequences with GenBank accessions GQ850318, GQ850355 & GQ850368..."	No No	Holotype K(M) 173535.	Sequences in GenBank	Kirk 2012
Fungi Mortierellales	1 species of <i>Mortierella</i>	"With an ITS sequence (GenBank JQ693160) that is distinct from other members of the gamsii/ elongata clade, deviating in the ITS1 region from other species in the clade; with a 94–97% similarity. With a sister group relationship to a possibly polyphyletic clade containing <i>Mortierella</i> <i>sclerotiiella</i> (basal; GenBank HQ63031, ex type), <i>M. cogitans</i> (GenBank HQ630281, ex type), and <i>M. acrotoma</i> (GenBank HQ630328, ex type)."	No No	Holotype IMI 398111	Sequences in GenBank	Bridge and Hughes 2012

(continued)

TABLE 1. Continued

			No	Type material in public collection	Sequences from GenBank (none new)	
Lepidoptera: Hesperiidae: Eudamiae	10 species of <i>Astraptes</i>	Specified substitutions in mitochondrial cytochrome c oxidase subunit I gene fragment (COI)	No			Brower 2010
Mollusca, Gastropoda, Abyssochrysidae Hydrothermal-vent snails	5 species of <i>Alviniconcha</i> , deep sea, hydrothermal vents, Western Pacific and Indian oceans	Specified substitutions in mt COI, 12S mt RNA, 16S rRNA, nuclear 28S, and 18S rRNA	Yes	Type material in a public collection	Sequences in GenBank	Johnson et al. 2015
Mollusca, Gastropoda, Coralliophilidae Coral snails	14 species of <i>Leptoconchus</i> , parasitic snails living in corals in Indo-West Pacific of Egypt, the Maldives, Thailand, Palau, and Indonesia	Specified substitutions in mt COI and nuclear ITS	Yes	685 snails collected from 327 hosts, incl. type material, deposited in various collections	Sequences in GenBank; alignment said to be in TreeBase, but not there	Gittenberger and Gittenberger 2011
Mollusca, Gastropoda, Glaucidae Nudibranchia	3 species of <i>Glaucus</i> , pelagic, Pacific Ocean	Specified substitutions in mt COI and 16S rRNA	Yes	Type material in a public collection	Sequences in GenBank	Churchill et al. 2014
Gastropoda, Microhedyliidae Micro slugs	9 species of <i>Ponitohedyle</i> from the mesopsammon in oceans worldwide	Specified substitutions in mt COI, 16S rRNA, nuclear 28S, and 18S rRNA	Yes	Type material in various collections	Sequences in GenBank	Jörger and Schrödl 2013
Nematoda, Rhabditidae	15 species of <i>Caenorhabditis</i>	This species differs by SSU, LSU and ITS2 DNA sequences (JN636069) from all other species, listed in Tables 1 and 2. "Note that these ribosomal DNA sequences may vary within the species"	No	The type culture specimens are deposited at the <i>Caenorhabditis</i> Genetics Center	Sequences in GenBank	Félix et al. 2014

Note: Full references are in the main text's reference list. Several studies rely entirely on DNA and provide no morphological diagnoses or descriptions.

incorporated molecular characters into protologs has led to uncertainty and reinventing of the wheel (Cook et al. 2010; Jörger and Schrödl 2013).

Species are always delimited against already known species (Linnaeus 1753, 1758; Mayr 1992; Naciri and Linder 2015). This holds true regardless of whether they are conceived as created (Linnaeus 1753, 1758) or as the result of evolution (Mayr 1992). Huge numbers of “cryptic” species—a term only meaningful relative to the particular technology used for studying organisms—can be distinguished with genomic data, and taxonomists are facing the challenge of naming at least some of this organismal diversity as it may be relevant for their research interests. It is useful then to consider how earlier taxonomists facing large numbers of new species mastered the task.

Foremost among taxonomists naming species is Linnaeus, who named ca. 6000 species of plants and 4400 species of animals (Müller-Wille 2006; Jarvis 2007). To do so, Linnaeus focused on diagnostic features in which a species differs from closest relatives. He was rightly proud of this idea and devoted much thought to the drafting of his diagnostic phrases, which were for him the true names of species. “Linnaeus held that these diagnosis should not exceed 12 words in length, and he and Jacquin even managed on occasion to reduce them to one word” (Stearn 1992, p. 144). In addition, he cited previous literature, available illustrations, and species’ ranges where known (Jarvis 2007). Longer descriptions based on multiple specimens and indicating the range of trait variation became widespread with the *Prodromus* project of the two De Candolles (I. Friis, personal communication, May 2015). Alphonse de Candolle was president of the International Botanical Congress in 1866 in London and wielded an immense influence (Nicolson 1991), and, obviously, a focus on within-species variation fit with Darwinian views on descent with modification.

While descriptions that mix taxonomically informative and uninformative traits became customary after about 1850, none of the *Codes of Nomenclature* stipulates that a new taxon must be described because such a requirement would have made Linnaeus’s names unavailable (under the zoological code) or invalid (under the botanical and mycological code). Instead, all Codes leave a choice between either a description or a diagnosis. The *Code of Nomenclature for algae, fungi, and plants* (McNeill et al. 2012, article 32.2) defines a diagnosis as “a statement of that which in the opinion of its author distinguishes the taxon from others.” The *International Code of Zoological Nomenclature* (ICZN 1999, article 13.1.3) defines it thus “When describing a new taxon, an author should make clear his or her purpose to differentiate the taxon by including with it a diagnosis, that is to say, a summary of the characters that differentiate the new nominal taxon from related or similar taxa,” and the *Bacteriological Code* (Lapage et al. 1992) states that any name proposal “must contain a brief diagnosis, i.e., a statement or list of those features that led the author to conclude that the proposed taxon is sufficiently different from other recognized taxa...”

Besides a diagnosis or description, a type specimen must be clearly indicated, and it is the type material that provides the objective standard of reference for the application of the name it bears.

The combination of the type method (i.e., name-bearing specimens deposited in one or more collections) and the discrete nature of nucleotide characters (substitutions or insertions/deletions of codons) begs reconsideration of Linnaeus’s focus of diagnosing species by features that distinguish them from their known closest relatives, instead of describing mixed sets of traits that vary at different hierarchical levels. Here, I consider the ways in which taxonomists have incorporated DNA characters directly into the publication of new species names, and I also review the history of DNA-based formal naming. I conclude with recommendations about best practice DNA-based diagnosis.

## MATERIALS AND METHODS

### *Literature Search and Data Documentation*

I compiled published molecular diagnoses through internet searches, surveying relevant journals and corresponding with colleagues. A molecular diagnosis involves the formal naming of a taxon by listing the DNA or protein characters in which it differs from its closest relative(s) in the protolog, thus associating it with a binomial Latinized name and the type material with its place of deposition. I checked that the molecular data indicated in the diagnosis were accessible in the cited database, usually the National Center for Biotechnology Information (NCBI: <http://www.ncbi.nlm.nih.gov/>, accessed 28 April 2016).

My survey focused on species names. An example of a molecular diagnosis of a higher taxon is that of the family Ambuchananiaceae Seppelt & H.A. Crum ex A.J. Shaw, “fam. nov. Plantae heterogeneae in morphologia, synapomorphis molecularibus in DNA nuclei mitochondri et plasti unitae. Type: *Ambuchanania*” (Shaw et al. 2010, p. 1523). Of course, this was before botanists abolished the Latin requirement on 1 January 2012. The baselines for bacterial names are *Approved Lists*, with a starting point of 1980, and new bacterial names are reviewed by a nomenclature committee and published in the *IJSEM* (Lapage et al. 1992). As mentioned in the “Introduction,” the proposal of a new bacterial name must contain a type designation and a brief diagnosis, that is, a statement or list of those features that led the author to conclude that the proposed taxon is sufficiently different from other recognized taxa to justify its naming (see also Stackebrandt and Goebel 1994). For the present review, I focus on eukaryotes.

## RESULTS

### *The Use of DNA Characters in Species Diagnoses since 2001*

The first to discuss how DNA characters might be used in species diagnosis were Don Reynolds and



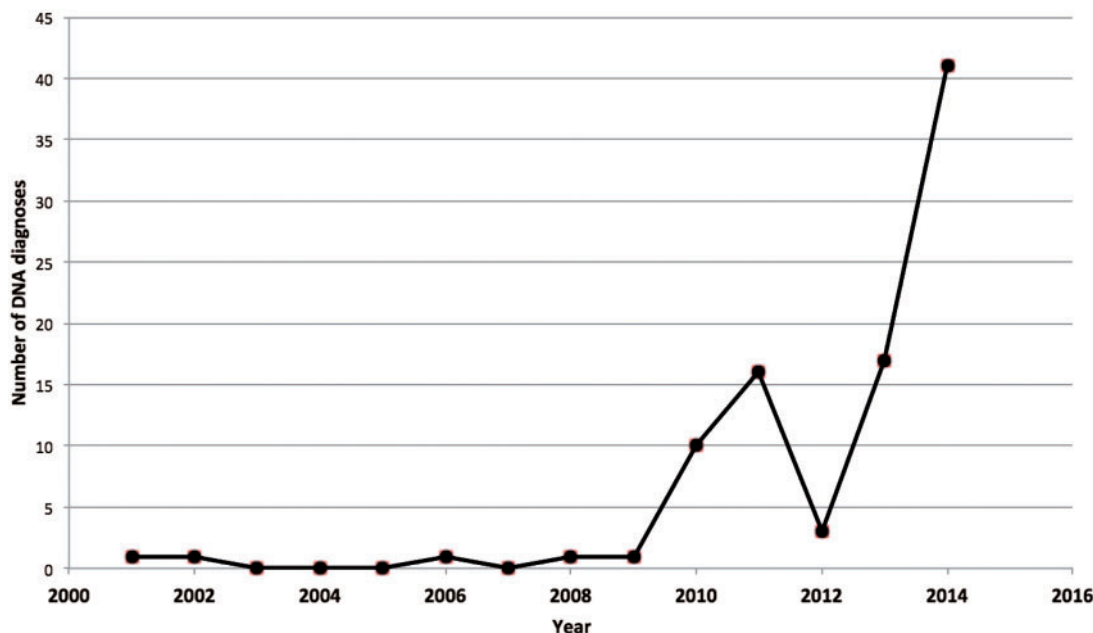


FIGURE 1. DNA-based formal diagnoses of new species of eukaryotes (mainly fungi, animals, and plants) since 2000 (based on data in Table 1). I am aware of only one molecular diagnosis published in 2015 (Irimia and Gottschling 2016).

John Taylor (1991) who clarified that the existing rules of the *International Code of Botanical Nomenclature* (as it was then still called) allowed DNA-based species naming and that DNA itself could serve as the type element. They provide two hypothetical examples of new fungal species names, one with a mix of DNA and morphological type materials and the other with DNA type material only, and call on herbaria to prepare for storing DNA material as types. Almost a quarter of a century has passed since this prescient article, but taxonomists are still feeling the need to defend the use of DNA characters in protologs (Cook et al. 2010; Jörger and Schrödl 2013; Tripp and Lendemer 2014), and the approach is only slowly becoming more common (Fig. 1).

Perhaps, surprisingly, Reynolds and Taylor (1991) devoted more discussion to the idea of using genomic material as type material than to the utility of nucleic acid characters as diagnostic tools because they thought it “unavoidable that DNA will serve as character source for contemporary taxonomic descriptions” (Reynolds and Taylor 1991, p. 311). Their hypothetical diagnosis, for a species collected on “a health food candy bar”, takes the shape, “5′-3′, ATGCCTAATAACTACCTAGC, AACT GATACTAATACC, Nucleotide positions 116-136, 1200-1216, Small Nuclear rDNA (1800 BP); 5′-3′, TATAGCCGCTAATCG CTAGATAA, Nucleotide positions 100-123, Mitochondrial Small rDNA (1648 BP).” (l.c., p. 314). The first molecular diagnosis of a real, not hypothetical, taxon is of a polychaete annelid from the Seychelles, the protolog of which differentiates it from morphologically similar individuals from Rhodos and Tenerife (Westheide and Hass-Cordes 2001). Twenty specimens were available for microscopy, and 13 others were used for RAPD fingerprinting or sequencing of the

nuclear internal transcribed species region of ribosomal DNA. The diagnostic DNA characters consist of eight RAPD bands obtained with specified primers and of characteristic substitutions in the ITS2.

By November 2015, 98 molecular diagnoses of species of Acoelomorpha, Alveolata, Angiospermae, Annelida, Arachnida, Arthropoda, Ascomycota, Chordata (Reptilia and Pisces), Fungi, Lepidoptera, Mollusca, and Nematoda have been published (Table 1). Relatively few protologs refrain from also providing a morphological description (Brower 2010; Molina et al. 2011; Leavitt et al. 2013).

#### *Names of Species of Fungi with DNA-Based Diagnoses*

Mycologists were at the forefront of DNA-based formal species naming, probably because their organisms pose particular challenges, as pointed out by Reynolds and Taylor (1991, p. 315), “recognition of DNA as at least part of the type element is certain to diminish the reliance on sexual characters for classification and undermine the maintenance of a separate form-classification for fungi lacking sex.” Most mycologists see no problem in diagnosing species by specific DNA substitutions (Taylor 2011; Schoch et al. 2012; Koljalg et al. 2013), and mycologists have also developed some of the most creative DNA-based diagnoses (Table 1). Thus, Hobbett et al. (2011, p. 45) proposed this form, “The least inclusive group containing organisms with nuclear rRNA ITS sequences with GenBank accessions AB244041 and DQ054545.” This exact form was used by Kirk (2012) in the protolog of *Piromyces cryptodigmaticus* Fliegerová, K. Voigt &

P.M. Kirk, diagnosed as “The least inclusive clade containing organisms with nuclear rRNA ITS sequences with GenBank accessions GQ850318, GQ850355 & GQ850368; with a sister group relationship to the clade containing the proposed epitype of *Piromyces communis* with a nuclear rRNA ITS sequence with GenBank accession AY429665; the closest named common ancestor, *Cyllumyces aberensis*, with a nuclear rRNA ITS sequence with GenBank accession FJ483845. Holotype K(M) 173535.” A similar form has been used for five species of lichens (Leavitt et al. 2013, p. 11), “*Rhizoplaca polymorpha* consists of specimens recovered within ‘clade IVc’ in Leavitt et al. (2011), which is supported as a lineage distinct from all other populations according to coalescent-based genetic analysis of multiple genetic loci.”

This form of clade-based diagnosis (“The least inclusive clade containing...”) has been challenged by Tripp and Lendemer (2012), who have requested the Committee on the application of the *Code of Nomenclature for algae, fungi, and plants* to decide on the validity of this form, which in their view goes against the requirement in Article 32.2(d) that a diagnosis cannot describe properties such as purely aesthetic features, economic, medicinal or culinary usage, cultural significance, cultivation techniques, geographical origin, or geological age. This matter is currently unsolved, and I have not found examples from outside fungi and lichens of this form of diagnosis (Table 1).

#### *Names of Animal Species with DNA-Based Diagnoses*

While nuclear 18S and 28S rRNA, mitochondrial 16S rRNA, and protein-encoding cytochrome *c* oxidase subunit I (COI or *cox1*) sequences have all been used in the diagnoses of new animal species names, the barcoding region, *cox1*, which pinpoints the correct species in many groups of insects (Hausmann et al. 2013), has been used especially often (Table 1). In most studies, DNA diagnostic features serve to corroborate morphological differences. For example, diagnostic COI substitutions that agree with shell characters clearly diagnose species of parasitic snails, but “impoverished anatomical details [alone] do not allow identification” (Gittenberger and Gittenberger 2011; p. 21). Several large-bodied species, such as turtles, have also been diagnosed with molecular *cum* morphological traits (Petzold et al. 2014). DNA-derived traits, mixed with morphology, have also been used in a key to 205 described braconid Hymenoptera *Apanteles* from Mesoamerica (Fernández-Triana et al. 2014), but Fernández-Triana and colleagues decided not to use DNA barcoding traits as species diagnoses, instead using the form “sequences in BOLD: 2, barcode compliant sequences: 2.”

#### *Names of Plant Species with DNA-Based Diagnoses*

Between January 1935 and 2012, botanists (and mycologists) had to write any diagnosis in Latin

(Table lists three such molecular diagnoses). Since 2012, however, a few plant species have been diagnosed with nucleotide substitutions described in English (Table 1), and one study even provides both molecular and morphological diagnoses and molecular and morphological descriptions (González et al. 2013).

## DISCUSSION

### *Advantages of the Sharper Diagnosis of Type Material*

A key advantage of molecular diagnoses is their utility for more precisely characterizing type material than is possible with morphological traits. The better a type collection (including syntypes and paratypes) is characterized, the more reliable the identification of future specimens. This does not mean that unidentified specimens in the future will need to be sequenced for identification. Instead, identification may continue to rely on morphological matching of preserved specimens or, increasingly, of images using machine learning. Having stringent diagnoses that specify DNA differences among closely related species (or subspecific taxa) can facilitate identification in those cases where the correct identification of a specimen is crucial, for example, for parasites of crops or of animals, especially us, but also for specimens that are incomplete, poorly preserved, or immature, so that diagnostic features are missing. Also, as pointed out by Cook et al. (2010), it is often quicker and cheaper to use diagnostic DNA features than to rely on the traditional expert-centered paradigm of identification.

The many studies that have clarified erroneous application of names or relationships among living and extinct species by sequencing DNA from type material attest to the importance of DNA diagnosis, now and in the future (Stuart and Fritz 2008; Hausmann et al. 2009; Sebastian et al. 2010; Stuckas and Fritz 2011; Stuckas et al. 2013; Fritz et al. 2014; Petzold et al. 2014; Heupink et al. 2014; Cappellini et al. 2014; Renner et al. 2014; Speidel et al. 2015; Erpenbeck et al. 2016).

### *Easy Accessibility, Interpretability, and Utility in Automated Keys*

Several taxonomic journals have hypertext markup language that allows direct linkages between new species names and sequences in GenBank or other sequence databanks (Penev et al. 2010). Sequences mentioned in diagnoses will serve as a standard for future reference, as pointed out by Reynolds and Taylor (1991) and Tautz et al. (2003), together with the type material deposited in one or, better, more museum collections (cf. the Darlington quote at the top of this paper). “DNA sequence information is digital and is not influenced by subjective assessments. It would be reproducible at any time and by any person, speaking any language. Hence, it would be a universal communication tool and resource for taxonomy, which



can be linked to any kind of biological or biodiversity information. Even if a query sequence does not produce an exact match, it will be possible to link an organism to closely related ones" (Tautz et al. 2003, p. 71). These authors, therefore, proposed that an attempt be made to provide a DNA sequence alongside all future taxonomic samples and species descriptions. In my view, this should become a recommendation in all Codes. Taxonomists, however, have begun to go further by including DNA characters directly in the diagnosis of nominal new taxa. This makes the type material more valuable and is safer for the future than if sequences come from other specimens that may be less well-preserved than type material typically is (or should be). Most importantly, sharp diagnosis of the types of species names will help avoid the publication of unnecessary names (new synonyms).

Last, DNA sequence databases with automated matching can replace identification keys. The functionality of such species-naming pipelines has been demonstrated in fungi (Koljalg et al. 2013). For animals, the concept of a Barcode Index Number (BIN) has been proposed (Ratnasingham and Hebert 2013), namely a persistent, species-level taxonomic registry using patterns of nucleotide variation in the barcode region of the cytochrome c oxidase I (COI) gene. The system begins by examining the correspondence between groups of specimens identified to species through prior taxonomic work and those inferred from the analysis of COI sequence variation using several algorithms.

#### *Differences Between Barcoding and DNA-Based Diagnosis, and How the Two Approaches Will Increasingly Reinforce Each Other*

There are three differences between barcoding and using DNA features in the protocols of new species. First, barcoding relies on a few universally agreed markers; DNA-based diagnosis does not, but can instead use a mix of other DNA traits, even indels (cf. Table 1). Second, barcoding is about identifying unknown material by matching sequences to named sequences in a database. This is not the purpose of DNA-based diagnoses, which serve to better describe a new species' type collection(s). For barcoding, one does not need to study type material or deposit a type in a designated public collection, as one does to name a species. Third, barcoding one's material is not a requirement or recommendation in any of the *Codes of Nomenclature*, while diagnosis is a recommendation in all of them, providing the foundation for the view advocated here, that two or three examples of DNA-based diagnoses (perhaps from Table 1) be added to encourage the use of DNA-based diagnoses.

One of the early criticisms of DNA barcoding (identifying species with DNA sequence markers) originated from the misconception that it was equivalent to DNA taxonomy, and as pointed out by a reviewer of this Point of View, it may be important to stress that I am

arguing here for a (continued) modification in how we diagnose types, hence, an aspect of *nomenclatural* work, not in how we circumscribe species, which is a matter of taxonomy, not nomenclature.

#### *Genetic Distance Less Suitable than Diagnostic Substitutions?*

Tripp and Lendemer (2012) have raised the question whether node-based diagnoses (Hibbett et al. 2011; Kirk 2012), rather than diagnostic substitutions, are valid and have submitted a request to the *Nomenclature Commission (for plants and fungi)* for clarification of two examples involving fungal names published without reference to specific characters distinguishing them from their closest relatives (see above, "Results" section for a specific example, *Rhizoplaca polymorpha*). Based on my reading of 98 molecular diagnoses, I agree with Tripp and Lendemer that discrete DNA features of type specimens are more useful than node-based diagnoses, which focus on phylogenetic context, not specimens. At least one study, however, has combined genetic distances and discrete trait states (Meyer-Wachsmuth et al. 2014), and the naming of bacteria has long relied on distances (see section "Materials and Methods").

#### CONCLUSIONS

DNA-based diagnoses along with (generously loaned) museum specimens and stably archived specimen images to my mind are more important today than attempts to "cover" morphological variation in populations (which obviously can and will continue). Such attempts are always limited by the availability of material, which causes taxonomists to wait, often for years, until that perfect second or third collection shows up. Even where several specimens are available, which and how much variation to document and describe remains subjective (Darlington 1971) and is a matter of idiosyncratic taxonomic practice. Reducing the time spent on long descriptions and instead focusing on sharp diagnoses might lead to faster naming of new species (also Riedel et al. 2013).

The process of species discovery (their delimitation from already known species) is a question of human interest and available technologies, and this implies that we will never know all species, even if "knowing" is defined as having lists of agreed upon names as done by bacteriologists and increasingly mycologists. My recommendation would be to include examples of DNA-based diagnoses in the *Codes* to help practitioners (see section "Differences Between Barcoding and DNA-Based Diagnosis, and How the Two Approaches Will Increasingly Reinforce Each Other"). Matching with existing names, or naming as new, the thousands of unnamed entities waiting in collections (each more or less incompletely represented) will become easier with both ongoing barcoding efforts and the inclusion of DNA traits in the diagnoses of types.

## ACKNOWLEDGMENTS

I thank Frank Anderson for critical comments; K. Jörger, Zoological Natural History Collections, Munich, and D. Begerow, Systematic Mycology, University of Bielefeld, for pointing me to DNA diagnoses in their fields of expertise; Ib Friis and Henrik Balslev for the opportunity to present this work at a meeting of the Royal Danish Academy of Sciences and Letters in May 2015; and Michael Ohl and Staffan Müller-Witte for the opportunity to discuss the topic with philosophers, historians, and sociologists at a meeting in Hannover in June 2015.

## REFERENCES

- Begerow D., Nilsson H., Unterseher M., Maier W. 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Appl. Microbiol. Biotechnol.* 87:99–108.
- Bridge P.D., Hughes K.A. 2012. *Index Fungorum* 7 (1).
- Brower A.V.Z. 2010. Alleviating the taxonomic impediment of DNA barcoding and setting a bad precedent: names for ten species of "*Astraptes fulgerator*" (Lepidoptera: Hesperidae: Eudaminae) with DNA-based diagnoses. *Syst. Biodivers.* 8:485–491.
- Cappellini E., Gentry A., Palkopoulou E., Ishida Y., Cram D., Roos A.M., Watson M., Johansson U.S., Fernholm B., Agnelli P., Barbagli F., Littlewood D.T.J., Kelstrup C.D., Olsen J.V., Lister A.M., Roca A.L., Dalén L., Gilbert M.T.P. 2014. Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae). *Zool. J. Linn. Soc.* 170:222–232.
- Churchill C.K.C., Ángel V., Foighil Diarmaid O. 2014. Molecular and morphological systematics of neustonic nudibranchs (Mollusca: Gastropoda: Glaucidae: *Glaucus*), with descriptions of three new cryptic species. *Invertebr. Syst.* 28:174–195.
- Clouse R.M., Wheeler W.C. 2014. Descriptions of two new, cryptic species of *Metasiro* (Arachnida: Opiliones: Cyphophthalmi: Neogoveidae) from South Carolina, USA, including a discussion of mitochondrial mutation rates. *Zootaxa* 3814:177–201.
- Cook L.G., Edwards R.D., Crisp M.D., Hardy N.B. 2010. Need morphology always be required for new species descriptions? *Invertebr. Syst.* 24:322–326.
- Darlington P.J. Jr. 1971. The carabid beetles of New Guinea. Part IV. General consideration; analysis and history of fauna; Taxonomic supplement. *Bull. Mus. Comp. Zool.* 142:129–337.
- Edgecombe D.E., Giribet G. 2008. A New Zealand species of the trans-Tasman centipede order Craterostigmomorpha (Arthropoda: Chilopoda) corroborated by molecular evidence. *Invertebr. Syst.* 22:1–15.
- Egge J.J.D., Simons A.M. 2006. The challenge of truly cryptic diversity: diagnosis and description of a new madtom catfish (Ictaluridae: *Noturus*). *Zool. Scr.* 35:581–595.
- Erpenbeck D., Ekins M., Enghuber N., Hooper J.N.A., Lehnert H., Polisenio A., Schuster A., Setiawan E., Voogd N.J. de, Wörheide G., Soest R.W.M. van. 2016. Nothing in (sponge) biology makes sense – except when based on holotypes. *J. Mar. Biol. Assoc. U.K.* 96: 305–311.
- Feilix M.-A., Braendle C., Cutter A.D. 2014. A streamlined system for species diagnosis in *Caenorhabditis* (Nematoda: Rhabditidae) with name designations for 15 distinct biological species. *PLoS One* 9(4): e94723. doi: 10.1371/journal.pone.0094723
- Fernández-Triana J.L., Whitfield J.B., Rodríguez J.J., Smith M.A., Janzen D.H., Hallwachs W.D., Hajibabaei M., Burns J.M., Solis M.A., Brown J., Cardinal S., Goulet H., Hebert P.D.N. 2014. Review of *Apanteles sensu stricto* (Hymenoptera, Braconidae, Microgastrinae) from Area de Conservación Guanacaste, northwestern Costa Rica, with keys to all described species from Mesoamerica. *ZooKeys* 383:1–565.
- Filipowicz N., Nee M.H., Renner S.S. 2012. Description and molecular diagnosis of a new species of *Brunfelsia* (Solanaceae) from the Bolivian and Argentinean Andes. *PhytoKeys* 10:83–94.
- Fisher M.C., Koenig G.L., White T.J., Taylor J.W. 2002. Molecular and phenotypic description of *Coccidioides posadasii* sp. nov., previously recognized as the non-California population of *Coccidioides immitis*. *Mycologia* 94:73–84.
- Fritz U., Petzold A., Kehlmaier C., Kindler C., Campbell P., Hofmeyr M.D., Branch W.R. 2014. Disentangling the *Pelomedusa* complex using type specimens and historical DNA (Testudines: Pelomedusidae). *Zootaxa* 3795:501–522.
- Gittenberger A., Gittenberger E. 2011. Cryptic, adaptive radiation of endoparasitic snails: sibling species of *Leptoconchus* (Gastropoda: Coralliophilidae) in corals. *Org. Divers. Evol.* 21:21–41.
- González Gutiérrez P.A., Köhler E., Borsch T. 2013. A new species of *Buxus* (Buxaceae) from northeastern Cuba based on morphological and molecular character, including some comments on molecular diagnosis. *Willdenowia* 43:125–137.
- Halt M.N., Kupriyanova E.K., Cooper S.J.B., Rouse G.W. 2009. Naming species with no morphological indicators: species status of *Galeolaria caespitosa* (Annelida: Serpulidae) inferred from nuclear and mitochondrial gene sequences and morphology. *Invertebr. Syst.* 23:205–222.
- Hausmann A., Godfray H.C.J., Huemer P., Mutanen M., Rougerie R., van Nieukerken E.J., Ratnasingham S., Hebert P.D.N. 2013. Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS One* 8(12): doi: 10.1371/journal.pone.0084518.
- Hausmann A., Hebert P., Mitchell A., Rougerie R., Sommerer M., Edwards T., Young C.J. 2009. Revision of the Australian *Oenochroma vinaria* Guenée, 1858 species-complex (Lepidoptera, Geometridae, Oenochrominae): DNA barcoding reveals cryptic diversity and assesses status of type specimen without dissection. *Zootaxa* 2239: 1–21.
- Hawksworth D.L., Kalin-Arroyo M.T. 1995. Magnitude and distribution of biodiversity. In: Heywood V., editor. *Global Biodiversity Assessment*. Cambridge, UK: Cambridge University Press. pp. 107–191.
- Hebert P.D.N., Cywinska A., Ball S.L., deWaard J.R. 2003. Biological identifications through DNA barcodes. *Proc. Roy. Soc. Lond. B* 270:313–321.
- Heupink T.H., van Grouw H., Lambert D.M. 2014. The mysterious Spotted Green Pigeon and its relation to the Dodo and its kindred. *BMC Evol. Biol.* 14:136.
- Hibbett D.S., Taylor J.W. 2013. Fungal systematics: is a new age of enlightenment at hand? *Nat. Rev. Microbiol.* 11:129–133.
- Hibbett D.S., Ohman A., Glotzer D., Nuhn M., Kirk P., Nilsson R.H. 2011. Progress in molecular and morphological taxon discovery in fungi and options for formal classification of environmental sequences. *Fungal Biol. Rev.* 25:38–47.
- ICZN. 1999. *International Code of Zoological Nomenclature*. 4th ed. London, UK: The International Trust for Zoological Nomenclature. 306 pp. Available from: <http://iczn.org/iczn/index.jsp> (accessed 21 June 2015).
- Irimia R.-E., Gottschling M. 2016. A new species of *Rochefortia* (Ehretiaceae, Boraginales) from the Lesser Antilles. *Phytotaxa* 236:62–70.
- Jarvis C.E. 2007. *Order out of chaos: Linnaean plant names and their types*. London: Linnean Society of London in association with the Natural History Museum.
- John U., Litaker R.W., Montresor M., Murray S., Brosnahan M. L., Anderson D.M. 2014. Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: the introduction of five species with emphasis on molecular-based (rDNA) classification. *Protist* 165:779–804.
- Johnson S.B., Warén A., Tunnicliffe V., Van Dover C., Wheat G., Schultz T.F., Vrijenhoek R.C. 2015. Molecular taxonomy and naming of five cryptic species of *Alviniconcha* snails (Gastropoda: Abyssochrysoidea) from hydrothermal vents. *Syst. Biodivers.* 13:278–295.
- Jörger K.M., Schrödl G.M. 2013. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front. Zool.* 10:59.
- Kirk P.M. 2012. *Index Fungorum* 1 (1).
- Koljalg U., Nilsson R.H., Abarenkov K., Tedersoo L., Taylor A.F.S. and 37 others. 2013. Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22:5271–5277.

- Kress W.J., García-Robledo C., Uriarte M., Erickson D.L. 2015. DNA barcodes for ecology, evolution, and conservation. *Trends Ecol. Evol.* 30:25–35.
- Lapage S.P., Sneath P.H.A., Lessel E.F., et al. editors. 1992. International Code of Nomenclature of Bacteria: Bacteriological Code, 1990 Revision. Washington, DC: ASM Press.
- Leavitt S.D., Fankhauser J.D., Leavitt D.H., Porter L.D., Johnson L.A., St. Clair L.L. 2011. Complex patterns of speciation in cosmopolitan “rock posy” lichens – Discovering and delimiting cryptic fungal species in the lichen-forming *Rhizoplaca melanophthalma* species-complex (Lecanoraceae, Ascomycota). *Mol. Phyl. Evol.* 59: 587–602.
- Leavitt S.D., Fernández-Mendozà F., Pérez-Ortega S., Sohrabi M., Divakar P.K., Lumbsch H.T., St. Clair L.L. 2013. DNA barcode identification of lichen-forming fungal species in the *Rhizocarpon melanophthalma* species-complex (Lecanorales, Lecanoraceae), including five new species. *MycKeys* 7:1–22.
- Lendemer J.C. 2011. A taxonomic revision of the North American species of *Lepraria* s.l. that produce divaricatic acid, with notes on the type species of the genus *L. incana*. *Mycologia* 103:1216–1229.
- Linnaeus C. 1753. *Species plantarum*. Stockholm: Salvius.
- Linnaeus C. 1758. *Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. Stockholm: Salvius
- Mayr E. 1992. A local flora and the biological species concept. *Am. J. Bot.* 79:222–238
- McNeill J. Chairman of the Editorial Committee. 2012. International Code of Nomenclature for algae, fungi, and plants (Melbourne Code), Adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011 (electronic ed.). Bratislava: International Association for Plant Taxonomy.
- Meyer-Wachsmuth I., Curini Galletti M., Jondelius U. 2014. Hyper-cryptic marine meiofauna: species complexes in Nemertodermatida. *PLoS One* 9: e107688.
- Molina M.C., Divakar P.K., Millanes A.M., Sánchez E., Del-Prado R., Hawksworth D.L., Crespo A. 2011. *Parmelia sulcata* (Ascomycota: Parmeliaceae), a sympatric monophyletic species complex. *Lichenologist* 43:585–601.
- Müller-Wille S. 2006. Linnaeus’ herbarium cabinet: a piece of furniture and its function. *Endeavour* 30:60–64.
- Naciri Y., Linder P. 2015. Species delimitation and relationships: the dance of the seven veils. *Taxon* 64:3–16.
- Nicolson D.H. 1991. A history of botanical nomenclature. *Ann. Missouri Bot Garden.* 78:33–56.
- Penev L., Kress W.J., Knapp S., Li, D.-Z., Renner S.S. 2010. Fast, linked, and open – the future of taxonomic publishing for plants: launching the journal *PhytoKeys* 1:1–14.
- Petzold A., Vargas-Ramírez M., Kehlmaier C., Vamberger M., Branch W.R., du Preez L., Hofmeyr M.D., Meyer L., Schleicher A., Siroky P., Fritz U. 2014. A revision of African helmeted terrapins (Testudines: Pelomedusidae: *Pelomedusa*), with descriptions of six new species. *Zootaxa* 3795:523–548.
- Ratnasingham S. and Hebert P.D.N. 2013. A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS One* 8(7): e66213.
- Renner S.S., Chomicki G., Greuter W. 2014. Proposal to conserve the name *Momordica lanata* (*Citrullus lanatus*) (watermelon, Cucurbitaceae), with a conserved type, against *Citrullus battich*. *Taxon* 63:941–942.
- Reynolds D.R., Taylor J.W. 1991. DNA specimens and the “International code of botanical nomenclature”. *Taxon* 40:311–315.
- Riedel A., Sagata K., Suhardjono Y.R., Tänzler R., Balke M. 2013. Integrative taxonomy on the fast track – towards more sustainability in biodiversity research. *Front. Zool.* 10:15.
- Schoch C.L., Seifert K.A., Huhndorf S., Robert V., Spouge J.L., Levesque C.A., Chen W., and Fungal Barcoding Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. USA* 109: 6241–6246.
- Sebastian P., Schaefer H., Renner S.S. 2010. Darwin’s Galapagos gourd: providing new insights 175 years after his visit. *J. Biogeogr.* 37: 975–980.
- Shaw A.J., Cox C.J., Buck W.R., Devos N., Buchanan A.M., Cave L., Seppelt R., Shaw B., Larrain J., Andrus R., Greilhuber J., Temsch E.M. 2010. Newly resolved relationships in an early land plant lineage: Bryophyta class Sphagnopsida (peat mosses) *Am. J. Bot.* 97: 1511–1531.
- Speidel W., Hausmann A., Müller G.C., Kravchenko V., Mooser J., Witt T.J., Prosser S., Hebert P.D.N. 2015. Taxonomy 2.0: Next-Generation-Sequencing of old type specimens supports the description of two new species of the *Lasiocampa decolorata* group from Morocco (Lepidoptera, Lasiocampidae). *Zootaxa* 3999: <http://dx.doi.org/10.11646/zootaxa.3999.3.5>, accessed 28 April 2016.
- Stackebrandt E., Goebel B.M. 1994. Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* 44:846–849.
- Stearn W.T. 1992, 4th ed. *Botanical Latin: history, grammar, syntax, terminology and vocabulary*. Portland, OR: Timber Press.
- Stuart B.L., Fritz U. 2008. Historical DNA from type museum specimens clarifies diversity of Asian leaf turtles (*Cyclemys*). *Biol. J. Linn. Soc.* 94:131–141.
- Stuckas H., Fritz U. 2011. Identity of *Pelodiscus sinensis* revealed by DNA sequences of an approximately 180-year-old type specimen and a taxonomic reappraisal of *Pelodiscus* species (Testudines: Trionychidae). *J. Zool. System. Evol. Res.* 49:335–339.
- Stuckas H., Gemel R., Fritz U. 2013. One extinct turtle species less: *Pelusius seychellensis* is not extinct, it never existed. *PLoS One* 8: e57116.
- Tautz D., Arctander P., Minelli A., Thomas R.H., Vogler A.P. 2003. A plea for DNA taxonomy. *Trends Ecol. Evol.* 18:70–74.
- Taylor J.W. 2011. One Fungus = one Name: DNA and fungal nomenclature twenty years after PCR. *IMA Fungus* 2: 113–120.
- Tripp E.A., Lendemer J.C. 2012. (4–5) Request for binding decisions on the descriptive statements associated with *Mortierella sigyensis* (fungi: *Mortierellaceae*) and *Piromyces cryptodignaticus* (fungi: *Neocallimastigaceae*). *Taxon* 61:886–888.
- Tripp E.A., Lendemer J.C. 2014. Sleepless nights: When you can’t find anything to use but molecules to describe new taxa. *Taxon* 63: 969–971.
- Vences M., Guayasamin J.M., Miralles A., De La Riva I. 2013. To name or not to name: Criteria to promote economy of change in Linnaean classification schemes. *Zootaxa* 3636:201–244.
- Westheide W., Hass-Cordes E. 2001. Molecular taxonomy: Description of a cryptic *Petititia* species (Polychaeta: Syllidae) from the island of Mahé (Seychelles, Indian Ocean) using RAPD markers and ITS2 sequences. *J. Zool. Syst. Evol. Res.* 39:103–111.
- Wheeler Q.D., editor. 2008. *The new taxonomy. The Systematics Association Special Volume Series 76*, Boca Raton, FL: CRC Press, Taylor & Francis Group.