WHEN SHOULD WE DESCRIBE SPECIES?

Two intense debates are underway in fungal taxonomy, one more public than the other, with little consideration that one debate may inform the other. The more public debate, which continues in this issue of *IMA Fungus* (see p. 213), concerns possible mechanisms for typification of, and the attendant wisdom of accepting, DNA-only species. The quieter debate, occurring more among editors, often to the frustration of authors of rejected papers, concerns the acceptability of papers describing single new species, especially when based on single specimens or single cultures. Proponents of DNA-typified species presently focus on often semantic concerns to define minimum requirements, sometimes nodding towards best practices, while practitioners of traditional taxonomy propose species that meet the requirements of the *International Code of Nomenclature for algae, fungi, and plants* (ICN), but fail to meet the editorial policies of an increasing number of journals. To my eyes, these are actually the same debate. Both concern our concepts of discovery, and the inconsistent logic separating the two reveals a fault line between what the para-legal ICN and the editor/reviewer-enforced policies that determine what actually is published.

What is discovery?

In the provocative book *Reinventing Discovery* (Nielsen 2011), a quantum physicist argues that Big Data has fundamentally changed the nature of discovery in science. The image of a tortured genius experiencing eureka moments alone in a laboratory among books and arcane instruments is being replaced by teams of analysts organizing masses of data to generate empirical, statistically-supported new knowledge – the human role is as a fact-checker for software and computers, perhaps supported by citizen scientists. In the old world, the solitary taxonomist focused for a life-time on one group, and in the brave new world is replaced by a far more efficient Illumina sequencer that feeds millions of sequences into a bioinformatic pipeline and identifies thousands of OTUs that don't match any known species in the databases. As stated by Hibbett *et al.* (2011), "...molecular ecology is clearly the major arena of contemporary species discovery..." rather than conventional taxonomy. My question is whether this process actually discovers new species, or simply indicates that there are new species to be found? In modern ecology, when you have a substrate in your hand that contains DNA sequences of a thousand species, half of them unknown, have you discovered 500 new species or have you picked up a handful of dirt?

In conventional taxonomy, a specimen in hand allows measurement and descriptions of characters and character states, elucidation of ecology and observation of behaviour. Reproduced observations and a progressive chain of experiments are core elements of science. A DNA sequence is an observation that can be reproduced as long as there is a physical specimen that can be re-sampled and extended with existing or future technologies. A printed DNA sequence clearly does not meet this condition. The Law of Conservation of Information is controversial, but suggests that no new information is derived by rearrangement of existing data (Medawar 1988). Does the act of naming a sequence provide new information that is not already inherent in the sequence itself? I would say not.

Whether it is one specimen or a hundred, with a specimen in hand it seems clear that you have made a discovery. Does the knowledge that someone else has detected the same DNA in a different handful of dirt really change the picture? There are no characters other than nucleotides, there is no differential ecology or behaviour attributable to the specific unknowns, unless they can be inferred in some way by information inherent in the genetic sequences.

Is there any conceptual similarity between a species based on one specimen and a species based on a few DNA sequences? Does a double standard exist, where our historical practise allows (but is now actively discouraging) what some perceive as low quality species descriptions with an old technology, while preventing what some would consider a higher quality of species description using a new technology?

Single specimen species: Four examples

It is easier to offer my own work up for criticism than to question the decisions of others. Here are four species known originally from single specimens, two published, one published long ago by someone else, and one unpublished, with some rationalization and *post-facto* analysis of their present status.

(1) The bamboo spathes that yielded *Charomyces amphimelas* (Seifert 1987) were collected by a fellow MSc student during his holiday in Hawai'i. My MSc supervisor, R.J. Bandoni, had stacks of damp chambers all over his office. After six months, a black, a wiry growth filled one dish, but it did not belong to Bandoni's beloved heterobasidiomycetes, so the voluminous specimen cycled around the students in the lab until it ended up with me. The fungus was stunning (Fig. 1A), but didn't grow when I tried to culture it, the specimen having been stealth-colonized by *Trichoderma*. In 1987, the process was to show the fungus to everyone who might have seen it before, so I sent hunks of the fungus to the experts. None knew it; all were intrigued. The morphology was so distinctive that I had no hesitation describing it as a new genus, an approach that was not questioned by the reviewers of the paper. But I wonder if I would describe it today. Outside of compilations, and this editorial, this paper has never been cited (and self-citations don't count!), not even when a second species was added to the genus (although the accidental overlap of the generic name *Charomyces* with *Saccharomyces* fools some search engines!). As far as I am aware, the fungus has never been re-collected (no records in data.GBIF.org or mycoportal.org). I can speculate about its phylogenetic relationships based on conidium ontogeny, but it remains identifiable only for those skilled in the art of micromorphology, not to those skilled in the art of DNA sequencing.

(2) *Hirsutella uncinata* grew in a damp chamber from the nut-like follicles of a *Hakea* sp. that I picked off the ground in the Mount Tomah Botanical Garden outside Sydney, Australia, in 1999 (Seifert & Boulay 2004), while waiting for Pedro Crous and Brett Summerell to tire of looking at leaf spots. I will never forget my perplexity back in Ottawa when I removed the conidiophores of the *Hirsutella* with the mounting needle and found them to be stiff as wire instead of floppy and flaccid like the *Verticillium* I was expecting to see; then the delight as I looked at the preparation through the compound microscope, captured forever in the micrograph taken at that second (Fig. 1B). It was a true eureka moment, a powerful feeling of discovering something new. The generic assignment was clear, and once again I consulted with the experts. This time, there was a culture and some DNA sequences. Unlike *Charomyces,* this fungus is identifiable by morphologists and DNA sequencers alike. No sexual morph is known and the insect host, buried in the rock-like substrate, is uncharacterized. The paper has very modest citations, but has had some sequence traction as the closest relative of *Ophiocordyceps sinensis* of oriental medicine fame. To my knowledge this fungus also has not been seen again.

(3) *Harpagomyces lomnikii.* A few years ago, photographs of chains of pitted, doliiform cells with ornamenting hooks were circulated on the internet (Fig. 1C) by palaeontologists who found them in an archaeological dig in Argentina (Fernández *et al.* 2010). I was probably the only person alive who recognized them as something described in Poland a hundred years earlier as *Harpagomyces*. This genus was compiled among the hyphomycetes by Carmichael *et al.* (1980), but excluded from our 2011 compilation (Seifert *et al.* 2011). These odd cells had blown through a window in Warsaw, were given a name, and were then forgotten. Honestly, I felt I had given the Argentinians a name with no information attached; we did not even know what kingdom it belonged to. This fungus is named but of unknown classification, not really identifiable by morphology except by accident, and unidentifiable with DNA sequencing. How useful is this name?

(4) *The Brain Fungus*. My last example is undescribed and unnamed as far as I know; for convenience, we will call it the Brain Fungus. I found it once in my backyard, growing in a pile of old pine logs. It looked like an *Acremonium* or *Mortierella* through the dissecting microscope, but I practically fell off my chair when I looked through the compound. What was this? Spores always in pairs, produced inside some kind of sporangium (Fig. 1D) . . . but the spores never seemed to separate. I immediately made some single spore isolations (or single brain isolations) and wondered whether this might be a bacterium and if the antibacterial antibiotics in the medium would be a problem. Of course, nothing grew. We tried direct sequencing — no sequences. Is this an ascomycete, a zygomycete, a "hyphomycete" or something else? The hyphae suggest that it is probably a fungus, but I've shown pictures to colleagues and in conference presentations, hoping someone will recognize it and at least tell me what phylum it belongs to. Here we have an organism of ambiguous affinities, easily identifiable by morphology but unnamed, with no correlating DNA sequences, and the log pile is now gone. How best to catalogue this organism for future generations when there are only a few photographs and one dried specimen to provide evidence? My present strategy is to keep showing the pictures to people, like you, right now . . . looking for illumination, something that feels like knowledge rather than a chance observation.

What makes a species worth naming?

As one of the founders of *Fungal Planet* (Crous *et al.* undated), I am distressed at the response by many journals who consider that this is now the only viable avenue for description of single species, whether based on one gathering or many. In fact, my descriptions of *Charomyces* and *Hirsutella uncinata* would have fitted fairly well into *Fungal Planet*, and then might have actually received some sideways citations! If I were to describe the Brain Fungus now, I could do so validly but I would be creating an analogue of *Harpagomyces*. There would be no retrieval mechanism for future scientists to locate the description of this organism, and the use of the name would only occur if some future devotee of the historical literature of the selected taxonomic group happened to stumble on the publication.

Names are supposed to mean something; their purpose is to convey information. At a more basic level, naming something ensures that it is catalogued and forces future taxonomists to consider it. In my opinion, the alphanumeric serial codes introduced for species hypotheses by UNITE (Kõljalg *et al.* 2013) seem to be the solution for unnamed taxon retrieval for DNA data. I wish we had something like that for morphology!

Arbitrary rules, mindlessly applied, devalue the importance of competent, state-of-the-art work in systematics, too often in pursuit of citations. Are the limitations of what we can determine about a species from a DNA sequence more severe than what we can determine about a species when we have only one specimen? If not, why are so many journals reluctant to allow single species descriptions based on morphology, but lining up to publish controversial papers on DNA defines taxa that test the limits of the ICN?

Taxonomists worry about maintaining quality, so let us talk about quality. Our debates over whether to allow formal naming of sequence-only species or single species descriptions need to synchronized. Quality may not need to be legislated in the ICN, but it still needs to be enforced; there is a strong tendency among mycologists to use the ICN as a quality assurance mechanism. The framers of the ICN have to accept this.

We should not confuse data, whatever its technical components, with understanding and knowledge. What does it take to raise species description above banality, above trivia that could be extracted by any child or by a machine? Do we want machine taxonomy in fungal biology? From one perspective this seems like a paranoid question and, from another, prescient. If DNA sequences comprise both the description and the type, it is a short step to a pipeline that automatically describes and names the OTUs as species. The question of machine-automated species description is staring us in the face. Surely we should be discussing it?

Acknowledgement: I am grateful to M. Virginia Bianchinotti for the photograph used as Fig. 1C.

Carmichael JW, Kendrick WB, Conners IL, Sigler L (1980) *Genera of Hyphomycetes*. Edmonton: University of Alberta Press.

Crous PW, Seifert KA, Samson RA, Hawksworth DL (eds) [Undated] *Fungal Planet: a peer-reviewed global initiative to promote the study of fungal biodiversity.* http:/www.fungalplanet.org

Fernández AL, Grill S, Seifert KA, Bianchinotti V, Martínez G (2010) First record of

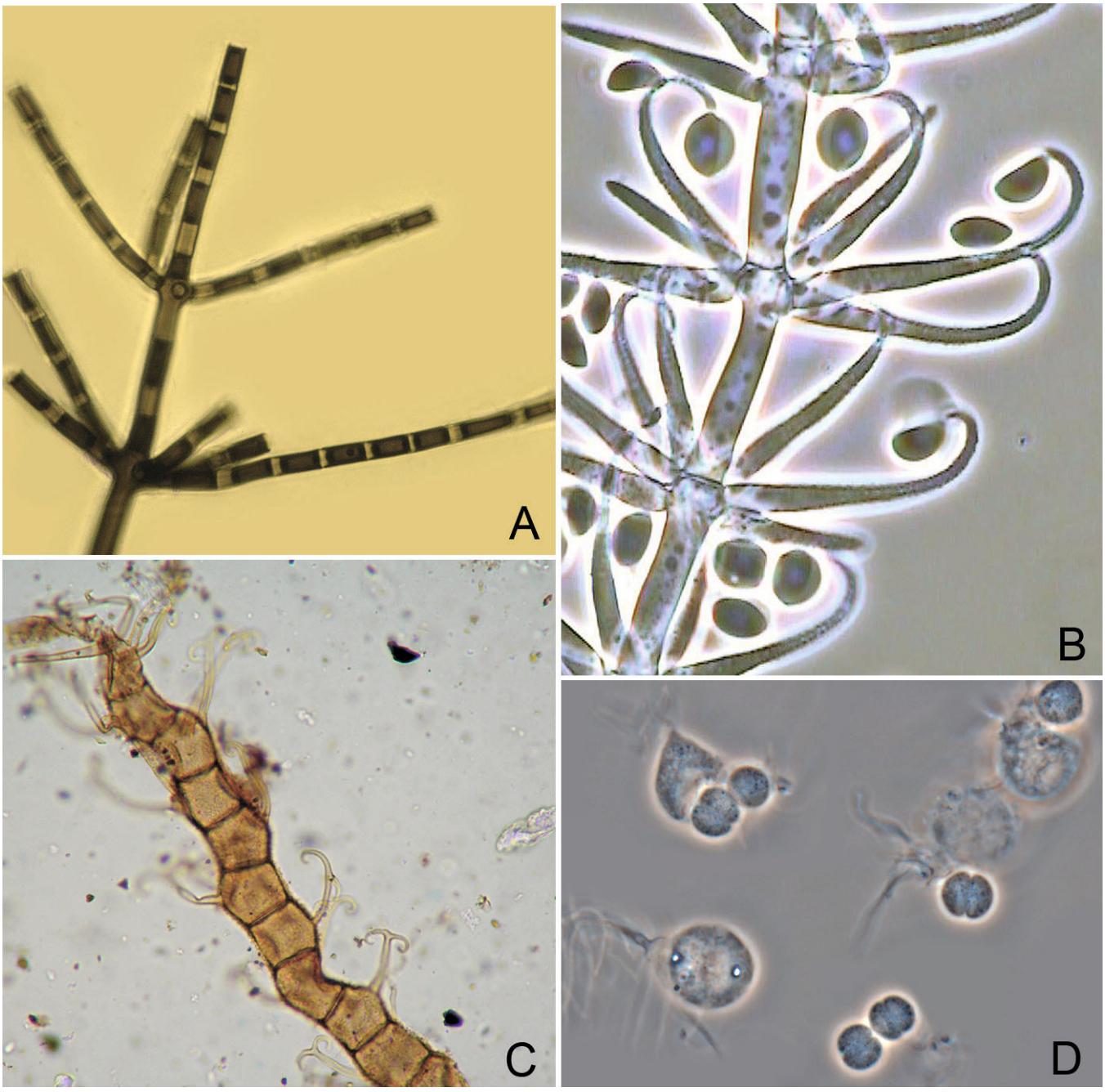

Fig. 1. A few of my favorite things. **A.** *Charomyces amphimelas* (from Seifert *et al.* 2011). **B.** *Hirsutella uncinata* (Seifert & Boulay 2004). **C.** *Harpagomyces* sp. (Fernández *et al.* 2010). **D.** The brain fungus.

Harpagomyces (*Eumycota, incertae sedis*) in Cenozoic sediments, lower Colorado Basin, Argentina. In: *X Congreso Argentino de Paleontología y Bioestratigrafía. VII Congreso Latinoamericano de Paleontología*. La Plata: Lugar.

Hibbett DS, Ohman A, Glotzer D, Nuhn M, Kirk P, Nilsson RH (2011) Progress in molecular and morphological taxon discovery in Fungi and options for formal classification of environmental sequences. *Fungal Biology Reviews* **25**: 38–47.

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, *et al.* (2013) Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology* **22**: 5271–5277.

Medawar PB (1988) *The Limits of Science*. Oxford: Oxford University Press.

Nielsen M (2011) *Reinventing Discovery: the new era of networked science*. Princeton: Princeton University Press.

Seifert KA (1987) *Charomyces*, a new genus of hyphomycetes with thallic-arthric conidia. *Canadian Journal of Botany* **65**: 230–232.

Seifert KA, Boulay H (2004) *Hirsutella uncinata*, a new hyphomycete from Australia. *Mycologia* **96**: 929–934.

Seifert K, Morgan-Jones G, Gams W, Kendrick B. (2011) *The Genera of Hyphomycetes*. [CBS Biodiversity Series vol. 9.] Utrecht: CBS/KNAW Fungal Biodiversity Centre.

Keith A. Seifert
President, IMA
(keith.seifert@agr.gc.ca)