ELSEVIER

British Mycological Society promoting fungal science

journal homepage: www.elsevier.com/locate/funbio

Hawksworthiomyces gen. nov. (*Ophiostomatales*), illustrates the urgency for a decision on how to name novel taxa known only from environmental nucleic acid sequences (ENAS)

Z. Wilhelm DE BEER[a,*], *Seonju MARINCOWITZ*[a], *Tuan A. DUONG*[b], *Jae-Jin KIM*[c], *Andre RODRIGUES*[d], *Michael J. WINGFIELD*[a]

[a]*Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria 0002, South Africa*
[b]*Department of Genetics, Forestry and Agricultural Research Institute (FABI), University of Pretoria, Pretoria 0002, South Africa*
[c]*Division of Environmental Science and Ecological Engineering, Korea University, 5-1 Anam-dong, Seongbuk-gu, Seoul 136-701, South Korea*
[d]*Department of Biochemistry and Microbiology, UNESP – São Paulo State University, 13506-900, Rio Claro, SP, Brazil*

ARTICLE INFO

Article history:
Received 14 February 2016
Received in revised form
2 July 2016
Accepted 12 July 2016
Available online 29 July 2016
Corresponding Editor:
Geoffrey Gadd

Keywords:
Melbourne Code
Nomenclature
Operational taxonomic units
Ophiostoma
Sporothrix

ABSTRACT

There have been many recent studies using environmental nucleic acid sequences (ENAS) to assess fungal diversity. As a result, more than a third of all fungal sequences in GenBank are of environmental origin. But inconsistent annotation of the thousands of undescribed taxa represented by these sequences limits access to these data. Consequently, these ENAS and the taxa they represent are rarely considered in other studies, and especially not in taxonomic treatments. This problem is confounded by the fact that the current version of the International Code of Nomenclature for Algae, Fungi, and Plants (Melbourne Code) prohibits the description of novel taxa known only from ENAS. There have been suggestions to emend the Code to allow a systematic nomenclatural treatment of these currently 'orphan' taxa but this has yet to occur. In this study, we considered the feasibility of using sequences from environmental studies to resolve the generic status of *Sporothrix lignivora*. This species forms a single lineage distinct from *Sporothrix* and other genera in the *Ophiostomatales*. BLAST searches in GenBank using LSU and ITS sequences of *S. lignivora* corresponded with several sequences from environmental studies. This also led to the discovery of isolates collected in diversity studies based on culturable fungi, with high similarity to *S. lignivora*. Phylogenetic analyses including taxa representing all major genera and lineages in the *Ophiostomatales* revealed a distinct, well-supported lineage that included *S. lignivora* and the ENAS. This confirmed the presence of a new genus in the *Ophiostomatales* described here as *Hawksworthiomyces* gen. nov., with *S. lignivora* as type species. Whereas only one described species was known in the so-called *S. lignivora* complex, our analyses revealed nine additional lineages in what is now *Hawksworthiomyces*. For three of these lineages, we were able to obtain isolates and these are described as *Hawksworthiomyces taylorii* sp. nov., *Hawksworthiomyces crousii* sp. nov., and *Hawksworthiomyces hibbettii*

* *Corresponding author*. Tel.: +27 12 4203938; fax: +27 12 4203960.
E-mail address: wilhelm.debeer@fabi.up.ac.za (Z. Wilhelm de Beer).
http://dx.doi.org/10.1016/j.funbio.2016.07.004

sp. nov. Five of the lineages each included one or more sequences from single studies, and thus remain unnamed. The remaining lineage included two sequences from separate studies of fungi inhabiting conifer wood. One of these sequences was an uncultured fungus clone from a spruce log in Sweden. The other sequence was for an isolate from a western red cedar fencepole in British Columbia, Canada, that was subsequently lost. These two ITS sequences differ in only two nucleotide positions. We are confident that they represent the same taxon and meet the criteria for an ENAS species, for which we provide the name, *Hawksworthiomyces sequentia* sp. nov. ENAS, and designate a DNA sequence as type in the absence of a type specimen. This case study makes it clear that environmental sequences and those from lost isolates can be extremely valuable in phylogeny-based taxonomic studies. It emphasises the fact that the Code should be emended to enable the naming of such taxa in a manner that will facilitate their incorporation in other studies.

Taxonomic novelties

Hawksworthiomyces Z.W. de Beer, Marinc., M.J. Wingf. **gen. nov.**; ***Hawksworthiomyces lignivorus*** (De Meyer, Z.W. de Beer, M.J. Wingf.) Z.W. de Beer, Marinc., M.J. Wingf. **comb. nov.**; ***Hawksworthiomyces taylorii*** Z.W. de Beer, Marinc., M.J. Wingf. **sp. nov.**; ***Hawksworthiomyces crousii*** Z.W. de Beer, Marinc., M.J. Wingf. **sp. nov.**; ***Hawksworthiomyces hibbettii*** Z.W. de Beer, Marinc., M.J. Wingf. **sp. nov.**; ***Hawksworthiomyces sequentia*** Z.W. de Beer, T.A. Duong, M.J. Wingf. **sp. nov. ENAS**.

Introduction

> *"… If one does not name the organisms, one does not know of them. Such organisms would just be … spots in a jungle or, as is the case of millions of organisms even today, they would not exist for us at all …"* --- Jouni Issakainen (Issakainen 1999)

The impact of DNA sequencing on fungal taxonomy and nomenclature during the past 25 years has been dramatic (Taylor 2011; Hibbett & Taylor 2013). In 1992, Berbee & Taylor (1992) showed that it is possible to link sexual and asexual fungi in molecular phylogenies based on DNA sequences. Within a decade, most major taxonomic treatments included phylogenies based on sequences, and in 2011 the practice of dual nomenclature was abandoned and the necessary changes were made to the Code to enforce 'one fungus one name' principles (Hawksworth *et al.* 2011; McNeill *et al.* 2012).

Not only were genera redefined, but DNA sequences have rapidly led to the formulation of new species concepts in fungi. Of these, Genealogical Concordance Phylogenetic Species Recognition (GCPSR) based on unlinked sets of characters (Taylor *et al.* 2000, 2006a,b), has become the standard for all the major contemporary taxonomic works. In phylogenetic studies applying the GCPSR, the name assigned to a lineage representing a species, is typically determined by DNA sequences of the type specimen or ex-type isolate of that species. To facilitate DNA-based species recognition, the ITS region (ribosomal internal transcribed spacer regions 1 and 2, including the 5.8S subunit) was accepted as universal barcode for fungi with the intention that every fungal species would be represented by a sequence in online databases (Schoch *et al.* 2012). Following the selection of the barcode, GenBank has established the curated, RefSeq Targeted Loci (RTL) database, where public sequence accessions were confirmed to be linked to valid species names and correctly annotated type specimens (Schoch *et al.* 2014). ITS has also become the gene region used most often in DNA-based surveys considering fungal diversity in different environments (Kõljalg *et al.* 2013). Since it is understood that the ITS does not always distinguish between closely related species, a more sensitive barcoding region (translation elongation factor 1-α) was recently recommended to be used in addition to ITS for fungal barcoding (Stielow *et al.* 2015).

The estimated number of fungal species on earth remains a matter of considerable debate (Hawksworth 2001; O'Brien *et al.* 2005; Mueller & Schmit 2007; Blackwell 2011). But it is clear that DNA sequences have contributed substantially to the rate of taxon discovery and that millions of species remain undescribed. A small part of this increase in numbers of newly discovered taxa is due to an improved ability to distinguish between cryptic species, but it is fungal molecular ecologists that are now at the forefront of species discovery (Hibbett *et al.* 2009, 2011; Hibbett & Glotzer 2011). By 2011 more than a third of all fungal sequences in GenBank were of environmental origin (e.g. soil, wood, leaf litter, etc.). This is in contrast to sequences for the remaining taxa that were primarily from specimens and mostly living cultures (Hibbett *et al.* 2011). Yet, the number of 'species' added annually to GenBank had already 'tipped' by 2009 in favour of environmental nucleic acid sequences (ENAS), as opposed to specimen-based sequences (Hibbett *et al.* 2011; Hibbett & Glotzer 2011; Taylor 2011).

In most environmental studies, sequences are grouped based on similarity into molecular operational taxonomic units (MOTUs), species hypotheses (SHs) or virtual taxa (VT) (Ryberg *et al.* 2008; Kõljalg *et al.* 2013; Öpik *et al.* 2014; Ryberg 2015). Different studies have applied different criteria in designating these 'taxa', and the naming or coding of MOTUs is generally applied informally and this varies from one study to another (Taylor & Hibbett 2013). These practices cause major confusion because there is no centralized database compiling all these names in a systematic manner. This makes the