

The species concept in *Aspergillus*: recommendations of an international panel

R.A. Samson^{1*}, J. Varga^{1,2}, S.M. Witiak³ and D.M. Geiser³

¹CBS Fungal Biodiversity Centre, Uppsalalaan 8, NL-3584 CT Utrecht, The Netherlands; ²Department of Microbiology, Faculty of Science and Informatics, University of Szeged, Közép fasor 52, H-6726 Szeged, Hungary; ³Department of Plant Pathology, Pennsylvania State University, University Park 16802, U.S.A.

*Correspondence: Robert A. Samson, r.samson@cbs.knaw.nl

INTRODUCTION

During an International Workshop held from 12–14 April 2007 in Utrecht, The Netherlands, participants discussed what an *Aspergillus* species is and how we delimit a species. In several sessions the species concept was presented by researchers in *Aspergillus* covering traditional and modern taxonomy, genetics, clinical, industrial and applied microbiology, biochemistry and genomics. Several points were discussed including the following:

- What and how many genes should be used to delimit an *Aspergillus* taxon?
- How does the phylogenetic species concept translate to practical and routine diagnoses?
- What is the impact of *Aspergillus* taxonomy in terms of epidemiology, case definitions and biological understanding of disease?
- What are the roles of *Aspergillus* databases for species identification?
- What is the value and impact of polyphasic approaches for species identification?
- What genes/methods can be used to design kits for rapid identification?
- How should new species be proposed?

The following issues are described below in more detail:

1. Which and how many genes are required to delimit a species?

During the discussion, several points were dealt with. It was suggested that genealogical concordance principles should be tested before we can answer the question of how many genes are needed to define a species. It was agreed that ITS sequences should be determined for later quick identification purposes, although sequence data are not required by the International Code of Botanical Nomenclature. For species delimitation, the polyphasic approach was suggested as the “gold standard” using a combination of multilocus sequence data, morphological, physiological characteristics and ecological data. However, no one character could be used as a “gold standard” to test the null hypothesis that this is not a new species, only applying the

whole set of characters will enable us to define new species. For species descriptions it is recommended to examine several (2–3) gene sequences (e.g. ITS, calmodulin, β -tubulin, actin) and submit them to recognised sequence databases. If the description of the new species meets the current conditions of the Botanical Code, it is not possible to reject a species. However, such an action can be recommended in the review of manuscripts describing new *Aspergillus* species. Regarding extrolites, it was suggested that a set of 4–8 compounds to be used, rather than a single molecule – no recommendations for which molecule. Additional points: it would be good to have a database (DNA and other types of characters) for quick identification; characters that are used for identification versus characters that are used to delimit a species should be separated.

Summary: A polyphasic approach is the gold standard, but it is difficult to define a standard set of criteria that must be met. Morphological characteristics are sometimes variable and need to be combined with other characters. Some suggestions for other characters included DNA sequences, physiological and ecological data and extrolite analyses. Regarding sequence data, ITS has been suggested because it has been used widely, but ITS sequences frequently show little or no variation between otherwise easily recognised closely related species. Solid phylogenetic species recognition generally requires multiple, more variable loci such as β -tubulin, calmodulin, actin, and other intron-rich protein coding genes. Difficulties regarding sequence data include the fact that not everyone has access to PCR and sequencing facilities, it is still unknown if other genes are more informative, and sequence data are not required by the Botanical Code. Consequently the absence of sequence data is presently not ground for rejection of new species.

2. What do we do about dual nomenclature?

It was suggested to use the same species name for both the anamorph and teleomorph, although it is not invalid to use another name according to the Botanical Code. A single culture should have only one name, but again it is not invalid to use a second name. Several participants suggested using *Aspergillus* as the primary name and the teleomorph as a secondary name, but experts of the International Botanical Code of Nomenclature disagreed as the second name will be invalid. Another suggestion was to epitypify all

new species to prepare for a single name, although again, it is not in agreement with the Botanical Code. Regarding clinical researchers who depend on one of the names, it was suggested to use the name “group” (or complex) if identification is based on morphology, which cannot distinguish between closely related species. Most, although not all participants, agreed to give preference to the teleomorph name. Researchers working on *Aspergillus* genetics mentioned that in a specific field, a name has a certain meaning, and preferred the name *Aspergillus* for phylogeny and genetics. Another suggestion was to give the name as follows: teleomorph [anamorph genus] if both exist, but it was rejected by most participants because names should not become too long by combining anamorph and teleomorph names. In the final vote, it was accepted by most participants to use dual names where necessary, single name in normal use, depending on the state that is seen, and treat it as a recommendation. Regarding Latin description of new species, it was accepted to have a short Latin diagnosis, followed by a more detailed English description. The Botanical Code accepts a valid description or a diagnosis in Latin. The majority of the workshop saw the need for a separate fungal nomenclatural code such as the code which the bacteriologists use.

Summary: Dual nomenclature is not an ideal system, but there are several difficulties with changing this. The majority voted to keep two names, but there were dissenting opinions. According to the Botanical Code, two names are legal and changing the code is difficult. However, different research communities sometimes use different names for a single species, causing confusion. The different names signal different morphological and physiological characteristics that people are looking for. Using a terminology such as complex or group might be helpful, but not precise enough for some applications. Names should not become too long by combining anamorph and teleomorph names. Single names would simplify things now that DNA features are available and identification does not necessarily depend only on morphology. However, it is difficult to have users switch to different names, and there is no consensus on which (teleomorph or anamorph) name to use.

3. What are the standards for describing and storing type cultures?

It was proposed that ex type cultures of new species be deposited in 2–3 different culture collections, preferably located in different parts of the world (Asia, Europe, US) to deal with import/export issues. If there is no available type culture, the *Aspergillus* community has the option of declaring it invalid. It was also suggested that at least a single locus DNA sequence must be provided for publication. There should be a limit to the time (e.g. six months) between publishing and depositing in collections. These suggestions are not accommodated by the Botanical Code, but were accepted as recommendations for good practice. New taxa should always be compared with ex type cultures of related species.

Summary: There should be 2–3 different open sources of a type culture of a newly proposed species in order to be scientifically valid. These should be deposited in a timely manner. New descriptions that are not deposited will be considered invalid by the *Aspergillus* community.

4. What sorts of databases do we need?

It was agreed that different specialised databases are needed for key identification purposes, with good links between databases. However, each database requires funding and curation. The use of a Wikipedia approach for the databases was questioned because of a lack of quality control, although it is a strong possibility as a clearinghouse for general information regarding protocols and media. It was suggested to use the *Aspergillus* website as a clearing house by linking to other sites. It was accepted that links should show the focus of each individual database, and there is a need for links to other communities as well. The databases that are linked to the *Aspergillus* website should be of high quality. Some overlap between databases is not a real problem. Regarding a simple database for species identification, it was suggested to include basic sequences for identification, pictures and links to media protocols. Although such a database already exists, it is in a more complicated form. Finally, it was agreed that the Wikipedia approach, not suitable for databases, was a good idea for media and protocols.

Summary: Databases are critical for identification and biology. In general, focused databases that are linked to other related databases were encouraged. The links should give an idea about the content of the database. A simple database for identification was proposed which would include basic sequences, photos, links to media/growth protocols, or the possibility to make a phylogenetic tree to get around the nomenclature problem.

RECOMMENDATIONS

- A polyphasic approach is recommended for describing new *Aspergillus* species, including molecular, morphological, physiological and ecological data.
- Any new species should be compared to type strains of presumed relatives.
- Any proposed new species should show evidence for evolutionary divergence from other taxa, particularly unique DNA characters at multiple loci, in addition to any distinctive extrolites and morphological characters.
- Latin descriptions can be short diagnoses, indicating differences from related taxa.
- Detailed morphological and physiological descriptions should be provided.
- Media used for the description should be based on the use of media recommended by the International Commission of *Penicillium* and *Aspergillus*: Malt Extract Agar and Czapek's Agar, with referenced formulas.
- Type cultures of new *Aspergillus* species should be deposited in at least two international recognized culture collections.
- If type cultures are not made available for the scientific community, the species will be considered invalid.
- New species names should be registered at MycoBank (www.MycoBank.org).
- For the description of new taxa, multiple, independent

loci are recommended for use, particularly loci for which large datasets already exist, such as ITS, β -tubulin, calmodulin, actin, RNA polymerase.

- Sequences must be deposited in recognized genetic databases.
 - Use dual names where necessary, and a single name in normal use, depending on the state (teleomorph or anamorph) that is observed.
 - Focused databases that are linked to other related databases are encouraged, with links giving an idea about the content of the database.
- A simple database for identification was proposed which would include basic sequences, photos, links to media/growth protocols, or the possibility to make a phylogenetic tree.

This protocol will be endorsed by the IUMS International Commission of *Penicillium* and *Aspergillus*.