FUNGAL BIOLOGY 119 (2015) 940–945

British Mycological Society promoting fungal science

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

How many DNA markers are needed to reveal cryptic fungal species?

*Sudhagar V. BALASUNDARAM, Ingeborg B. ENGH, Inger SKREDE, Håvard KAUSERUD**

Section for Genetics and Evolutionary Biology (EVOGENE), Department of Biosciences, University of Oslo, P.O. Box 1066 Blindern, 0316 Oslo, Norway

ARTICLE INFO

Article history:
Received 2 February 2015
Received in revised form
17 June 2015
Accepted 14 July 2015
Available online 26 July 2015
Corresponding Editor:
Kerry O'Donnell

Keywords:
Cryptic species
DNA barcode
Genetic markers
Phylogenetic informativeness
Serpula

ABSTRACT

In the fungal kingdom there is a high prevalence of morphologically defined species that includes closely related 'cryptic' biological species with similar phenotypes. Due to evolutionary processes like incomplete lineage sorting and introgression through hybridization, several independent DNA markers are essential to resolve closely related fungal species. In this study we wanted to analyze how many independent loci are necessary to reveal the cryptic species, using the genus *Serpula* as a model system. DNA sequences from ten different DNA loci, eight nuclear and two mitochondrial DNA markers, were obtained from various cryptic species within *Serpula*. The inclusion of five loci gave a highly confident separation of the cryptic species. Several other loci performed better than the standard DNA barcoding marker ITS in separating the cryptic species. The DNA loci *tub*, *hsp*, *rpb2* and *tef* gave, on average, best support for the different cryptic species in single gene trees. We conclude that the analyses of a few but informative independent DNA loci, such as *tub*, *hsp*, *rpb2* and *tef* in addition to the standard DNA barcode ITS, may give a good indication about the existence of cryptic species in fungi.

Introduction

The fungal kingdom comprises high richness of species occupying diverse niches, the latest estimate being six million species (Taylor *et al.* 2014). Traditionally, fungal species have been separated based on their fruit body morphology. However, the fruiting structures represent only a short phase in the fungal life cycle. Moreover, many fruit bodies have a very simple structure and include limited number of characters, e.g. corticoid fungi with resupinate fruit bodies. This means that a limited number of characters often have been available for delimiting fungal taxa, as compared to e.g. animals and plants, where a larger part of the phenotype is available for characterization and species delimitations. This is may be one important reason why the prevalence of cryptic species, i.e. morphologically indiscernible biological/phylogenetic units present within taxonomic species, seems to be highly widespread in the fungal kingdom.

The adoption of molecular DNA based genetic analyses has revealed that cryptic species are highly prevalent in fungal morphospecies. Early in the 1990's, Vilgalys & Sun (1994) demonstrated that the oyster mushroom *Pleurotus ostreatus* includes high levels of phylogenetic divergence, where eight

* *Corresponding author*. Tel.: +47 22854832; fax: +47 22854664.
E-mail addresses: sudhagar@ibv.uio.no (S. V. Balasundaram), Ingeborg.Bjorvand.Engh@mycoteam.no (I. B. Engh), inger.skrede@ibv.uio.no (I. Skrede), havard.kauserud@ibv.uio.no (H. Kauserud).
http://dx.doi.org/10.1016/j.funbio.2015.07.006

phylogenetic groups were detected. Eight cryptic species were identified in the genus *Neurospora* (Dettman *et al.* 2003) and in *Saccharomyces*, several cryptic species were identified by phylogenetic analysis of twenty genes across eight yeast taxa (Rokas *et al.* 2003). In the root rot fungus *Heterobasidion annosum* (sensu lato) a number of phylogenetic groups have been detected in Europe and in North America (Garbelotto *et al.* 1998; Johannesson & Stenlid 2003). The widespread wood decay fungi *Coniophora arida*, *Coniophora olivacea* (Kauserud *et al.* 2007a) and *Coniophora puteana* (Kauserud *et al.* 2007b) comprise a significant number of cryptic species. In the fungal kingdom, identification of cryptic species is essential for a better understanding of species richness and for proper conservation and management plans.

Phylogenetic species recognition relies upon the analyses of multiple un-linked DNA markers, where congruence in tree topology across markers is evaluated (Taylor *et al.* 2000). Groups that are stable and can be recognized across several markers likely represent evolutionary independent lineages that recombine within but not across groups. However, how many independent markers are needed to differentiate those groups? This question certainly depends upon factors like lineage age, demographic history and population size and, hence, vary from case to case.

In this study, we are focusing on cryptic species found in the family *Serpulaceae* (*Basidiomycota, Boletales*). The genus *Serpula* belongs to this family and consists of saprotrophic fungi, which are involved in wood degradation, resulting in brown rot (Binder & Hibbett 2006). In addition, the *ectomycorrhizal* genera *Austropaxillus* and *Gymnopaxillus* are grouping inside *Serpula* (Skrede *et al.* 2011). The dry rot fungus *Serpula lacrymans* is divided into two groups that represent genetically well-separated lineages (Kauserud *et al.* 2007c). One of them is *S. lacrymans* var. *shastensis*, residing naturally in North America, while the other is *S. lacrymans* var. *lacrymans*, which is colonizing buildings on all continents (Kauserud *et al.* 2007c). The sister species *Serpula himantioides* appears to include five cryptic species, some of them with a worldwide distribution (Carlsen *et al.* 2011).

Our aim in this study is to evaluate how many genetic markers are needed to resolve cryptic species, using *Serpula* as a model group. To address this question, ten different DNA markers were sequenced, two representing mitochondrial DNA (COX and mtSSU) and eight representing nuclear DNA (*gpd*, *hsp*, ITS, LSU, SSU, *rpb2*, *tef* and *tub*).

Materials and methods

Materials

In this study we analyzed 29 collections and isolates of *S. lacrymans* var. *lacrymans*, *S. lacrymans* var. *shastensis*, *S. himantioides*, *Austropaxillus* spp., and *S. incrassata* strains from the family *Serpulaceae* of the order *Boletales* (Table S1).

Molecular work

A small amount of fungal tissue was homogenized in a Mixer Mill (MM301, Retsch GmbH & Co., Haan, Germany) before we extracted total DNA following the 2 % CTAB miniprep method described by Murray & Thompson (1980) with minor modifications from Gardes & Bruns (1993). We dissolved the dried DNA pellet in 100-μL of sterile milli-Q H_2O, and used further dilutions for PCR amplification. Ten different DNA loci were PCR amplified using the primers and PCR programs listed in Table S2. Cycle sequencing was performed using the ABI BigDye Terminator sequencing buffer and v3.1 Cycle Sequencing kit (Life Technologies, Carlsbad, CA). Sequences were processed on an ABI 3730 DNA analyser (Life Technologies). Sequences were assembled and edited using BioEdit 7 (Hall 1999). All sequences have been deposited in GenBank, and accession numbers are given in Table S1.

Phylogenetic analyses

Single gene alignments were constructed for the ten different genetic markers. These were combined into a concatenated alignment using AIR-Appender (Kumar *et al.* 2009) with default parameters. Further, randomized concatenated alignments were constructed that included two to nine gene markers. For each number of loci (2–9) 20 replicate alignments were obtained. In total, 150 randomized gene alignments were constructed in this manner. All the sequences were aligned with ClustalW (Larkin *et al.* 2007) with default parameters and then manually corrected. In total, ten single gene trees, a single concatenated tree and 150 alternative trees were constructed. All phylogenetic trees were constructed using maximum likelihood analysis as implemented in the GTRCAT approximation in RaxML (Stamatakis 2006) running 1000 bootstraps. Phylogenetic trees were viewed in FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/) and were manually inspected for tree topology and branch support for selected clades. All phylogenetic analyses were done on the Abel computing cluster from The University of Oslo (http://www.uio.no/english/services/it/research/hpc/abel/).

We analyzed phylogenetic informativeness (PI) to calculate an informativeness profile for each gene to resolve branching order against a particular epoch in a phylogenetic tree (chronogram). The online program PhyDesign (Lopez-Giraldez & Townsend 2011) was used to measure the PI of the ten loci. We calculated net and per site phylogenetic informativeness. The net PI resolved particular nodes in the phylogenetic tree to assess the utility of molecular markers, whereas the PI per site is relevant because it compares relative informativeness of genes without the confounding influence of sequence length (Townsend 2007; Fong & Fujita 2011). PI finds genes with an optimal rate of evolution during a given period of time on the phylogenetic tree.

Results and discussion

Multi-gene phylogeny

The best maximum likelihood tree based on the ten combined genetic markers is shown in Fig 1. The tree supports the topology observed in previous phylogenetic studies of *Serpula* (Carlsen *et al.* 2011; Skrede *et al.* 2011). Not surprisingly, the current tree, based on additional loci, provides higher support for all nodes. The ectomycorrhiza forming genus *Austropaxillus*