Fungal Biology 123 (2019) 517–527

Contents lists available at ScienceDirect

Fungal Biology

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

Evaluation of ITS2 molecular morphometrics effectiveness in species delimitation of Ascomycota – A pilot study

Natesan Sundaresan, Amit Kumar Sahu, Enthai Ganeshan Jagan, Mohan Pandi*

Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 15 June 2018
Received in revised form
1 April 2019
Accepted 2 May 2019
Available online 8 May 2019

Corresponding Editor: Gabor M. Kovacs

Keywords:
CBC
Fungal barcoding
INDELs
Sequence-secondary structure
Species specific variation

ABSTRACT

Exploring the secondary structure information of nuclear ribosomal internal transcribed spacer 2 (ITS2) has been a promising approach in species delimitation. However, Compensatory base changes (CBC) concept employed in this approach turns futile when CBC is absent. This prompted us to investigate the utility of insertion/deletion (INDELs) and substitutions in fungal delineation at species level. Upon this rationale, 116 strains representing 97 species, belonging to 6 genera (*Colletotrichum*, *Boeremia*, *Leptosphaeria*, *Peyronellaea*, *Plenodomus* and *Stagonosporopsis*) of Ascomycota were retrieved from Q-bank for molecular morphometric analysis. CBC, INDELs and substitutions between the species of their respective genus were recorded. Most species combinations lacked CBC. Among the substitution events, transitions were predominant. INDELs were less frequent than the substitutions. These evolutionary events were mapped upon the helices to discern species specific variation sites. In 68 species unique variation sites were recognised. The remaining 29 species shared absolute similarity with distinctly named species. The variation sites catalogued in them overlapped with other distinct species and resulted in the blurring of species boundaries. Species specific variation sites recognized in this study are the preliminary results and they could be discerned with absolute confidence when larger datasets encompassing all described species of genera were investigated. They could be of potential use in barcoding fungi at species level. This study also concludes that the ITS2 molecular morphometric analysis is an efficient third dimensional study of the fungal species delimitation. This may help to avoid the false positives in species delimitations and to alleviate the challenges in molecular characterization.

1. Introduction

Fungi represent a wide group of heterotrophic organisms with diverse ecological functions. A conservative estimate of their number exceeds 1.5 million although highly debated and continues to raise (Hawksworth, 1991, 2001). About 98 128 species had been catalogued (Bass and Richards, 2011), while the major chunk remains undescribed to lower taxonomic level. Because of their pleomorphic nature, erroneously named entries in public domains and lack of reference sequences, most isolates could not be distinguished/named at the species level. These obstacles have rendered uncertainty over the years in delimitating species boundaries. Conventionally, morphological description of spores, sporangial appendages, hyphal structures and colony characteristics had been employed in delimiting the species boundaries (Cai et al., 2009; Diaz et al., 2012). These descriptions are highly subjective in nature and are compounded by their pleomorphism and plasticity in different culture conditions (Nilsson et al., 2006; Cai et al., 2009; Kusari et al., 2014). Hence, they lead to assigning multiple names for a single organism. Nevertheless, morphological descriptions were fairly enough to distinguish most fungus at the genus level and would fall short in species delimitations. Such instances in the mycological research are well documented (Begerow et al., 2010; Money, 2013).

Mycologists and researchers resorted to molecular approaches along with morphological features to reduce the conflicts and resolve the confusions in the delimitation of fungal species (Arnold et al., 2000; Davis, 2010; Diaz et al., 2012). This has been proven as a quick and reliable method for delimiting the morphologically similar but genetically distinct fungal species and vice versa (Athokpam and Tandon, 2014). The internal transcribed spacer (ITS) region is commonly used in the species delimitation of fungi and diverse eukaryotes. The ITS region composed of ITS1 and ITS2 lies

* Corresponding author.
E-mail address: an_pandi@rediffmail.com (M. Pandi).

https://doi.org/10.1016/j.funbio.2019.05.002
1878-6146/

between the 18S–5.8S and 5.8S–28S rDNA cistrons respectively. The complete ITS spanning about 500–700 bp includes both ITS segments and highly conserved 5.8S region. Their features such as homogenized high copy number, ease of amplification even from small quantities of DNA and phylogenetic informativeness, renders them an ideal marker for species delimitation (Coleman, 2003; Bargues et al., 2014; Rampersad, 2014).

While several other molecular markers are being contended, ITS was proposed as the universal barcode for fungal species delimitation and widely employed (Begerow et al., 2010; Ihrmark et al., 2012; Bazzicalupo et al., 2013; Blaalid et al., 2013). Cytochrome C oxidase subunit I (COI), Actin (ACT), β – tubulin (TUB2), Calmodulin (CAL), Glutamine synthase (GS), RNA polymerase II subunits (RPB1 and RPB2), transcription elongation factor 1 alpha (TEF), Mini-chromosome maintenance complex component 7 (MCM7) gene, and Glyceraldehyde 3 – phosphate dehydrogenase (GAPDH) were the other markers employed to varying degree in cataloguing different taxons. Multigene based phylogenetic analysis was also proposed to circumvent the problems in identification of the isolates at the species level. Except ITS, other marker genes lack reference datasets and universal primers (Cai et al., 2009; Porras-Alfaro and Bayman, 2011; Réblová and Réblová, 2013; Rampersad, 2014). Also, the widespread adoption of this method is unlikely owing to the cost implied (Money, 2013).

In about three decades of molecular (predominantly ITS based) phylogenetic studies to delimit the fungi and describe the fungal diversity tentatively, many strides had been achieved (Taylor, 2011; Ferro et al., 2014). The most important of them has been the rendition of phylogenetic evidence to suggest that sexual (teleomorph) and asexual (anamorph) stages of a fungus had the same gene sequence. This implored scientists to discontinue the double naming system that tolerated separate names for the anamorph and teleomorph phases of the same fungus, a significant step in reducing the confusion in naming a fungus (Hawksworth et al., 2011; Taylor, 2011). Incorporation of one fungi one name concept would eliminate the synonymous entries and the number of catalogued fungal species might dwindle to range between 43 271 (Mora et al., 2011) and ~59 000 (Hawksworth, 2004, 1992). Nevertheless, flooding of public databases with several uncharacterized and erroneously named entries continues to plague fungal species delimitation (Nilsson et al., 2006; Ozerskaya et al., 2010).

Recently, efforts to curate the ITS data flooded in the global databases, have been taken up to reduce the misidentification of fungal species. Also, in an effort to enhance the robustness in phylogenetic delimitation using ITS sequences, the ITS1 and ITS2 subunits were explored further. ITS2 has been proven to be phylogenetically more informative than the ITS1. In addition, incorporation of secondary structural information of ITS2 in phylogenetic analysis has been shown to increase the accuracy and robustness of the study (Ahvenniemi et al., 2009; Yan et al., 2018). ITS2 secondary structure has a common core and genus/species specific structural features, which is much more conserved on the structure level than on the sequence level. The method of utilizing the secondary structure information of nucleic acids to assess the phylogenetic relationship among the species is referred to as molecular morphometrics (Billoud et al., 2000). ITS2 sequence-secondary structure based delimitation at both genus and species levels of animals, insects and plants were already documented and proven as a useful marker by several authors (Banerjee et al., 2008; Barik et al., 2011; Chen et al. 2004, 2010; Gao et al., 2010; Padhi et al., 2016). However, ITS2 molecular morphometrics analysis based fungal delimitation studies are yet to be explored in detail and it has been the rationale of this study.

ITS2 based molecular morphometrics analysis has rendered a CBC concept of species delimitation which corroborates well with the biological species concept based on the genetic incompatibility between organisms (Wolf et al., 2013). In a pair-wise comparison of secondary structures of two species, CBCs are observed in the paired regions (loop/helix). Both the nucleotides of a pairing mutates to retain the base pairing (e.g., GC to AU and vice versa) and such changes are known as CBCs. Presence of one or more CBCs in a pair-wise comparison, has been shown as a diagnostic feature of the distinctness of the species (Muller et al., 2007). However, the diagnostic credibility turns void when CBC is 0. In such cases, the pair under consideration could either represent same or distinct species and this remains a bottleneck for wider application of CBC concept based species delimitation. In this perspective CBCs, INDELs and substitution (transition and transversion) events were recorded and we intended to explore their influence in fungal barcoding. Evolutionary events such as INDELs and substitution have been shown to be phylogenetically informative in many markers (Fritz et al., 1994; Salvi and Mariottini, 2012), however, they have not been investigated in ITS2 molecular morphometric studies. We postulate them to be useful, especially when CBCs fail to provide unequivocal evidence for the distinctness of species. Based upon this background, in this study, we intended to investigate the utility of ITS2 molecular morphometrics in fungal species delimitation. Representatives from six common genera of Ascomycetes were included in this investigation. In this perspective, CBCs, INDELs and substitution events were recorded and their utility in fungal barcoding was explored.

2. Materials and methods

2.1. Dataset

Internal transcribed spacer sequences of fungal species characterized by morphological features and multi-loci phylogeny available from Q-bank database (www.Q-bank.eu) were retrieved. Among the resources available at Q-bank, 6 genera were chosen for this pilot study. They include *Colletotrichum*, *Boeremia*, *Leptosphaeria*, *Peyronellaea*, *Plenodomus* and *Stagonosporopsis*. In total, 346 sequences of six different genera belonging to Ascomycota were retrieved, of which, 314 sequences had a complete ITS2 region and they were used for this study. The inclusion of sequences from six different genera provides an opportunity to investigate the utility of this study in a wider group of fungi. They represented 97 species, among which 76 species had more than one strains and the remaining species had single strain. We have analysed the similarities among the multiple strains of a species and a representative sequence was chosen in case of 100% similarity. Conversely, when variations were encountered among the multiple strains of a species, representatives for each variant were included in this study. The accession number and binomial name of the sequences used in the dataset were listed in Supplementary Table S1.

2.2. Molecular phylogenetic analysis

The Fungal ITS extractor was employed to extract the ITS2 region (Nilsson et al., 2010). Subsequently, the full length of ITS and ITS2 sequence based phylogenetic analysis were performed using Maximum likelihood with the program RAxML 7.4.2 (Stamatakis et al., 2008) and a Bayesian approach with the MrBayes 3.2.2 program (Ronquist and Huelsenbeck, 2003). RAxML run with Maximum likelihood + rapid bootstrap for 1000 replicates using GTR model. Bayesian analyses run with 1 000 000 generations starting with a random tree and employing 2 simultaneous chains were executed. Every 100th tree was saved into a file. The first 100 000 generations were deleted as the “burn in” of the chain. The program Tracer 1.6 was applied to compare splits frequencies in the

different runs and to plot cumulative split frequencies to ensure that stationarity was reached. Of the remaining trees, a majority rule consensus tree with average branch lengths was calculated using the sumt option of MrBayes. Bootstrap (BP) values ≥50 and posterior probability (PP) values ≥0.5 were represented above the branches of ML tree (BP/PP). Phylogenetic trees were visualized using the program TreeGraph 2 (http://treegraph.bioinfweb.info/Download).

2.3. Secondary structure prediction

Preceding the structure prediction, the 5′ and 3′ end of ITS2 region was added with 20 bp of the 5.8S and 28S rRNA sequences respectively. The added sequences aid in canonical base pairing and folding of the ITS2 region (Chen et al., 2004; Bridge et al., 2008). Secondary structures were modelled using Mfold software version 3.2 (http://mfold.rna.albany.edu/?q=mfold/RNAFolding-Form) using the default settings for a linear molecule folded at 37 °C in 1 M NaCl with no bivalent ions. Maximum interior loop size and asymmetry were set at 30, and suboptimality was set between 3 and 5% to give a minimum of three possible structures for each sequence. In almost every taxon, multiple alternative folding patterns per sequence with Minimal Free Energy (MFE) values were observed. The structure with minimum free energy and the common core was chosen for analysis (Bridge et al., 2008; Caisová et al., 2013; Athokpam and Tandon, 2014). For those structures which do not follow the common core, Homology Modelling (HM) was performed using the ITS2 database (http://its2.bioapps.biozentrum.uni-wuerzburg.de/) with the default settings ITS2 PAM 50 matrix, 75% threshold for the helixes transfer similarity, gap open penalty of 20 and gap extension penalty of 4 for the transfer helixes with the custom modelling option available (Selig et al., 2008; Ahvenniemi et al., 2009). Structural information's were downloaded in Vienna format for further analysis and jpeg format to visualize the secondary structure models. Inferred individual structures were examined for the following parameters: common core, length, base composition, GC content, and structural energy.

2.4. Molecular morphometric analysis

The ITS2 sequences were simultaneously aligned with their secondary structures in 4SALE Ver. 1.7 using clustalW binary file. The alignment of sequence with secondary structure information and single letter amino acid codes of sequence secondary structure information were exported for neighbour-joining and maximum likelihood phylogenetic analyses respectively. The consensus secondary structures were also visualized in 4SALE. The CBCs were counted using the CBC Analyzer as implemented in 4SALE (Seibel et al., 2006; Ahvenniemi et al., 2009; Ruhl et al., 2010; Buchheim et al., 2011). In addition, evolutionary events of INDELs and substitution events were also manually recorded.

The sequence − secondary structure based phylogenetic trees were reconstructed by neighbour-joining (NJ) through the use of an ITS2 sequence-structure specific and general time reversible substitution model in ProfDistS Ver.0.9.9 (http://profdist.bioapps.biozentrum.uni-wuerzburg.de/cgi-bin/index.php?section=download) (Wolf et al., 2008; Buchheim et al., 2011). The maximum likelihood (ML) tree was calculated using Phangorn as implemented in the statistical framework R. The R script is available from the 4SALE homepage (http://4sale.bioapps.biozentrum.uni-wuerzburg.de) (Wolf et al., 2014; Heeg and Wolf, 2015). Bootstrap support was estimated based on 1000 pseudo-replicates for both methods and the resulting tree was visualized with Tree Graph 2 (http://treegraph.bioinfweb.info/Download).

2.5. Cataloguing the variation sites

Pair-wise comparison of sequence and secondary structure of all species (included in this study) combinations from respective genera were investigated to document the variable sites. In species pair-wise comparisons of sequence-secondary structure, when a site is non-homologous to its counterpart, then it is denoted as a variable site. Across all the possible species pair-wise comparisons in a genus, occurrences of variable sites and the evolutionary events responsible for the variation were enlisted. The variation sites have been categorised as follows: If a site appears to be variable in more than 10 pair-wise comparisons, then they have been denoted as highly varied site (HVs). Similarly, when a site is observed to vary in 5–9 species pair-wise comparisons and <5 species pair-wise comparisons, they are termed as moderately varied site (MVs) and less varied site (LVs) respectively. The threshold levels set were specific for this dataset. However, similar cut-off levels were earlier reported by Caisová et al. (2011) in discerning the role of evolutionary events in speciation of Ulvales. Catalogued variation sites and evolutionary events were inferred to discern probable species specific variation sites (barcode). Graphical representations of the data were generated using GraphPad PRISM 6.

3. Result

The ITS2 sequences of 116 strains representing 97 species were analysed in our dataset and they belong to the following genera: *Colletotrichum* (18 strains), *Boeremia* (23 strains), *Leptosphaeria* (13 strains), *Peyronellaea* (26 strains), *Plenodomus* (14 strains) and *Stagonosporopsis* (22 strains).

3.1. Nucleotide information of ITS2 region

The length, base composition, and GC% were summarized in the Supplementary Table S1. The length of the ITS2 region ranged from 149 to 162 nucleotides. The shortest ITS2 region was found in *Plenodomus lingam* CBS 260.94 while the longest ITS2 region was found in *Plenodomus influorescens* CBS 143.84 and *Plenodomus congestus* CBS 244.64. The GC % of all these strains ranged above 50% except *Boeremia* spp., *Plenodomus lindquistii* CBS 381.67 and *Plenodomus pimpinellae* CBS 101637. The average GC% of *Colletotrichum*, *Stagonosporopsis, Peyronellaea*, *Leptosphaeria*, *Boeremia,* and *Plenodomus* were 54.85 ± 1.53, 51.05 ± 1.14, 50.85 ± 0.53, 53.23 ± 1.48, 48.90 ± 0.40 and 52.37 ± 2.04 respectively.

3.2. Sequence divergence across species in respective genera

Evolutionary events such as transition (Ts), transversion (Tv) and INDELs, contributes to the sequence divergence and eventually leads to speciation. Distribution of the sequence variability in ITS2 across the species combinations of genera was represented in Fig. 1. In this dataset, *Boeremia* spp. and *Peyronellaea* spp. possessed the least number of variations, contrastingly *Leptosphaeria* spp. and *Plenodomus* spp. featured the highest number of observed variations. About 17 species in this dataset had been represented by more than one strain due to the presence of variations in their sequences. Interestingly we have found ≤2 variations between the strains of a species.

3.3. Secondary structural features

The secondary structures predicted for these sequences were following the common core rule and their minimum free energy (kcal) was presented in Supplementary Table S1. The minimum free energy of structure varies between the species. The average

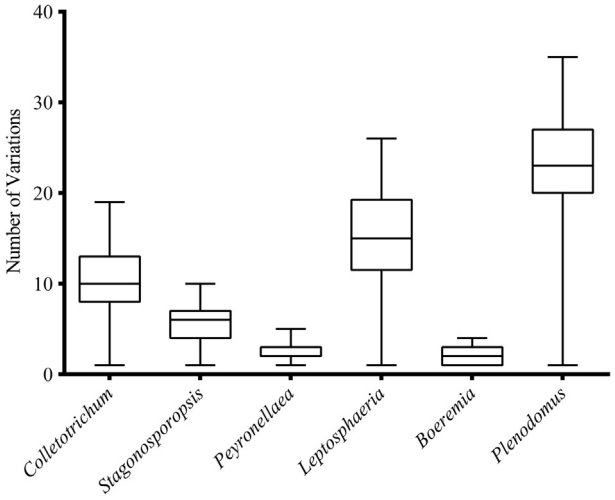

Fig. 1. Summarizes the number of variations in ITS2 sequences accounted between the species of their respective genus.

minimum free energy (kcal/mol) of *Colletotrichum*, *Boeremia*, *Leptosphaeria*, *Peyronellaea*, *Plenodomus,* and *Stagonosporopsis* were 72.42 ± 1.99, 58.66 ± 1.72, 72.96 ± 2.70, 62.62 ± 2.14, 73.07 ± 3.95 and 62.16 ± 1.88 respectively. In this study, two distinct types of secondary structures were observed, however, helix III had been the longest in both types. In type 1, 3 helices were bound around the central bulge and 4th helix does not occur. In another type, helix 3 and 4 were bound around the central bulge and the helix 1 and 2 were bound at the extended subdomain. Out of these 116 strains, 96 strains (82.76%) possessed type 1 structure and the remaining 20 strains had type 2 structure. Comparison and phylogenetic analyses would be effective with homogenous structural pattern and hence the strains with 4 helices (type 2) were remodelled into the type 1 structure using phylogenetically closely related strains possessing three helix structures as a template.

The general features of the ITS2 secondary structure, included the presence of UGGC motif in the 5′ apex of helix 3. Additionally, purine − purine and pyrimidine − pyrimidine mismatches had been observed and were depicted in Fig. 2. Among *Colletotrichum*, one UC mismatch in helix 1 and one or 2 UU mismatches in helix 3 had been observed. In *Staganosporopsis*, most species featured one UU and UC mismatches in helix 2, also one UU and AA mismatches in helix 3. Also, GA mismatch was found in helix 3 of all species. Similarly, in *Peyronellaea*, one UU, UC, GA mismatch in helix 2, one AA and GA mismatch in helix 3 was observed in all species. Also, one UU mismatch in helix 3 was found in most species. In *Leptosphaeria*, one UU mismatch was observed in helix 3 of most species. Among *Boeremia*, one UU, GA and AA/AG mismatches were observed in the 3rd helix of all species. Likewise, one UU mismatch was found in the 3rd helix of *Plenodomus* spp.

3.4. Cataloguing the evolutionary events

Out of 97 species, 29 species shared absolute similarities with other species recognized erstwhile. Similarities in other molecular markers reported erstwhile for these 29 species were also analysed (data not shown). Most marker genes possessed high similarity (>97%) and very few marker genes had divergence accounting for phylogenetic information to delineate the species. For example, in *Peyronellaea*, ACT, CAL and TEF markers had considerable divergence and all other markers had higher similarity. *P. protuberans* CBS 381.96 and *P. obtusa* CBS 377.93 had been named distinctly, even though they had absolute similarity in all marker genes except TEF (98.96%). Likewise, *P. alectorolophi* CBS 132.96 and *P. protuberans* CBS 381.96 possessed absolute similarity in all marker genes except ACT (99.62%). Also, in many cases, all marker genes were lacking for most of the species. For example, *P. pomorum* CBS 115.67 and *P. pomorum* var. *circinata* CBS 285.76 shared an absolute similarity in ITS2 and ITS. No other markers information was available. Such instances were observed throughout the dataset. The remaining 68 species were investigated for species specific evolutionary events.

Evolutionary events were manually analysed and catalogued through the sequence and secondary structure comparisons of all the species in the genera, to access their utility in species delimitation. Among them, transitions were more frequent than the transversion. INDELs were less frequent than the substitution events in all genus included in this study. The distribution percentage of INDELs and substitution events were depicted in Fig. 3.

The catalogued variations (Table 1) were majorly distributed in helix II, III and non-helical regions. In the majority of the investigated genera, helix III possessed the maximum variations. Exceptionally, maximum variations were observed in non-helical regions of *Boeremia* and helix II of *Plenodomus*. Least proportion of variations were observed in helix I of all genera investigated in this study, while absolute conservation in this region was evidenced in *Peyronellaea* and *Boeremia*.

3.5. ITS2 molecular morphometric analysis

Prior to the ITS2 molecular morphometrics analysis, full length of ITS and ITS2 sequence based phylogenetic trees were generated using ML and Bayesian methods. They exhibited concordant tree topologies and are provided in Supplementary Fig. S4a–c. These results prove that the ITS2 region possess comparable phylogenetic significance as that of the full length of ITS region. As a part of the molecular morphometric analysis, the ITS2 sequence secondary structure based phylogenetic tree were generated in NJ and ML methods and they also showed similar clade formations. Hence ML tree is presented with bootstrap support, deduced from both the methods (Figs. 4–6). Single letter amino acid codes of nuclear ribosomal ITS2 sequence-secondary structure alignment information were submitted in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S23745). The phylogenetic tree generated based on full length of ITS sequence, ITS2 sequence and ITS2 sequence − secondary structure showed similar clade formations. Further, CBC was observed among few species and in the majority of species comparisons they were lacking (Supplementary Table S2a–c). Species specific variation sites were catalogued and depicted in Supplementary Table S3. Helix-wise distribution of species specific HVs and MVs sites were represented at respective leaves of the ML tree. These results concur that the molecular morphometrics analysis of ITS2 marker could be eventually used in species delimitations.

3.5.1. Colletotrichum *and* Stagonosporopsis

In *Colletotrichum*, 15 species (18 strains) employed in the dataset were resolved into 5 distinct clades (Fig. 4). *C. lineola* CBS 109228 showed absolute similarity with *C. dematium* CBS 123728 and differed with *C. dematium* CBS 115524 by one variation. *C. dematium* CBS 123728 and CBS 115524 possessed 9 species specific variation sites and 1 strain specific variation sites. Similarly, *C. spinaciae* CBS 108.40 and *C. circinans* CBS 111.21 had absolute similarity and resolved together in the same clade. Three *C. truncatum* strains included in this study were resolved together in the same clade. However, strain specific variation sites were observed between them. *C. truncatum* CBS 136.30, CBS 125327 and CBS 112998

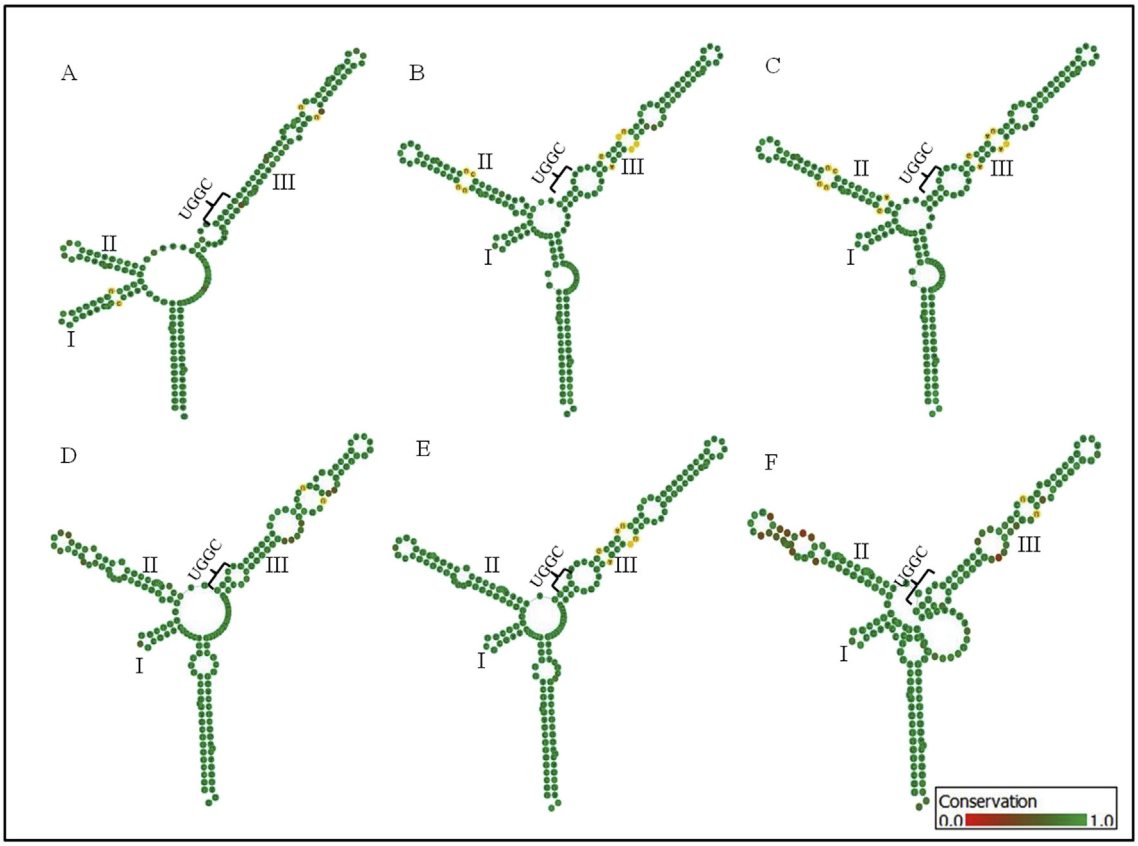

Fig. 2. Consensus secondary structures of ITS2 sequences (A) *Colletotrichum*, (B) *Stagonosporopsis*, (C) *Peyronellaea*, (D) *Leptosphaeria*, (E) *Boeremia* and (F) *Plenodomus*. The helices were numbered I–III from 5′ to 3′ direction. The conserved parts indicated in green (>51%; most conserved) to red (least conserved). Purine–purine and pyrimidine–pyrimidine mismatches are highlighted in yellow colour. UGGC motif in helix III was also depicted. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 3. Distribution frequencies of INDELs and substitution (transversion and transition) across genera.

possessed 7 species specific variation sites and 5 strain specific variation sites.

In *Stagonosporopsis,* 19 species (22 strains) employed in the dataset were resolved into 6 distinct clades (Fig. 4). *S. andigena, S. hortensis, S. trachelii, S. ajacis* were represented by two strains in this study. Strain specific variation sites were documented in them. *S. andigena* CBS 101.80 and CBS 269.80 possessed 7 species specific variation sites and 1 strain specific variation sites. *S. hortensis* CBS 104.42 and 572.85 featured 6 species specific variation sites and 1 strain specific variation sites. *S. trachelii* CBS 384.68 and CBS 379.91 featured 7 species specific variation sites and lacked strain specific variation sites. *S. ajacis* CBS 177.93 and CBS 176.93 featured 8 species specific variation sites and 1 strain specific variation sites. *S. caricae* CBS 248.90 and *S. dennisii* CBS 135.96 – showed absolute similarity and resolved together in sub-clade of clade 8. Likewise, *S. chrysanthemi* ATCC 10748, *S. rudbeckiae* CBS 109180, *S. ligulicola* var. *inoxydabilis* CBS 425.90 shared absolute similarity with one another and resolved together in clade 10.

3.5.2. Peyronellaea and Leptosphaeria

In *Peyronellaea*, 22 species (26 strains) employed in the dataset were resolved into 4 distinct clades (Fig. 5). *P. pomorum, P. pinodella, P. americana* and *P. crutisii* were represented by 2 strains each. However, species specific and strain specific variations sites documented in them shared similarities with other species. *P. pomorum* CBS 115.67 and *P. pomorum* var. *circinata* CBS 285.76 shared an absolute similarity and resolved together in clade 1. *P. australis* CBS 444.81, *P. arachidicola* CBS 315.90, *P. alectorolophi* CBS 132.96, *P. pinodella* CBS 100580, *P. protuberans* CBS 381.96, *P. obtusa* CBS 377.93, shared an absolute similarity and resolved together in clade 2. *P. glomerata* CBS 112448, and *P. aurea* CBS 269.93 shared an absolute similarity and resolved together as a subclade in clade 2. In clade 3, *P. coffeae-arabica* CBS 123380, *P. americana* CBS 112525, *P. zeaemaydis* CBS 588.69 and *P. subglomerata* CBS 110.92 shared absolute similarity with one another. They showed one variation with P. americana PD 80.1143. *P. sancta* CBS 281.83, *P. pomorum* var. *cyanea* CBS 388.80, *P. crutisii* PD 71.6, *P. pomorum* var. *pomorum* CBS 539.66 shared absolute similarity with one another and hence resolved together as clade 4.

In *Leptosphaeria*, 11 species (13 strains) employed in the dataset were resolved into 2 distinct clades (Fig. 5). *L. veronicae* possessed 7 species specific variation sites and 4 strain specific variation sites. *L. slovacica* possessed 11 species specific variation sites and 3 strain specific variation sites. *L. macrocapsa* CBS 640.93 and *L. errabunda*

Table 1
Documentation of helices wise ITS2 sequence variations across the species of their respective genus.

		Helix I		Helix II		Helix III		Non helical region	
Colletotrichum	INDELs	–	(02.45)	02.53	(25.91)	07.26	(42.53)	12.83	(29.11)
	Ts	01.18		18.48		21.77		05.81	
	Tv	01.27		04.89		13.50		10.46	
Stagonosporopsis	INDELs	–	(07.03)	–	(06.33)	15.86	(58.84)	12.55	(27.80)
	Ts	07.03		06.33		24.30		12.65	
	Tv	–		–		18.67		02.60	
Peyronellaea	INDELs	–	-	–	(37.07)	–	(41.46)	16.59	(21.47)
	Ts	–		37.07		26.83		–	
	Tv	–		–		14.63		04.88	
Leptosphaeria	INDELs	–	(03.35)	16.84	(40.17)	09.62	(45.08)	–	(11.40)
	Ts	03.35		14.43		18.0.41		06.49	
	Tv	–		08.89		17.05		04.92	
Boeremia	INDELs	–	–	–	(27.85)	–	(27.85)	26.58	(44.30)
	Ts	–		27.85		11.39		–	
	Tv	–		–		16.46		17.72	
Plenodomus	INDELs	–	(02.24)	25.21	(54.02)	04.35	(30.14)	05.02	(13.60)
	Ts	01.61		17.36		17.78		05.51	
	Tv	0.63		11.46		08.01		03.07	

Ts: transition, Tv: transversion and INDELs: insertion and deletion.

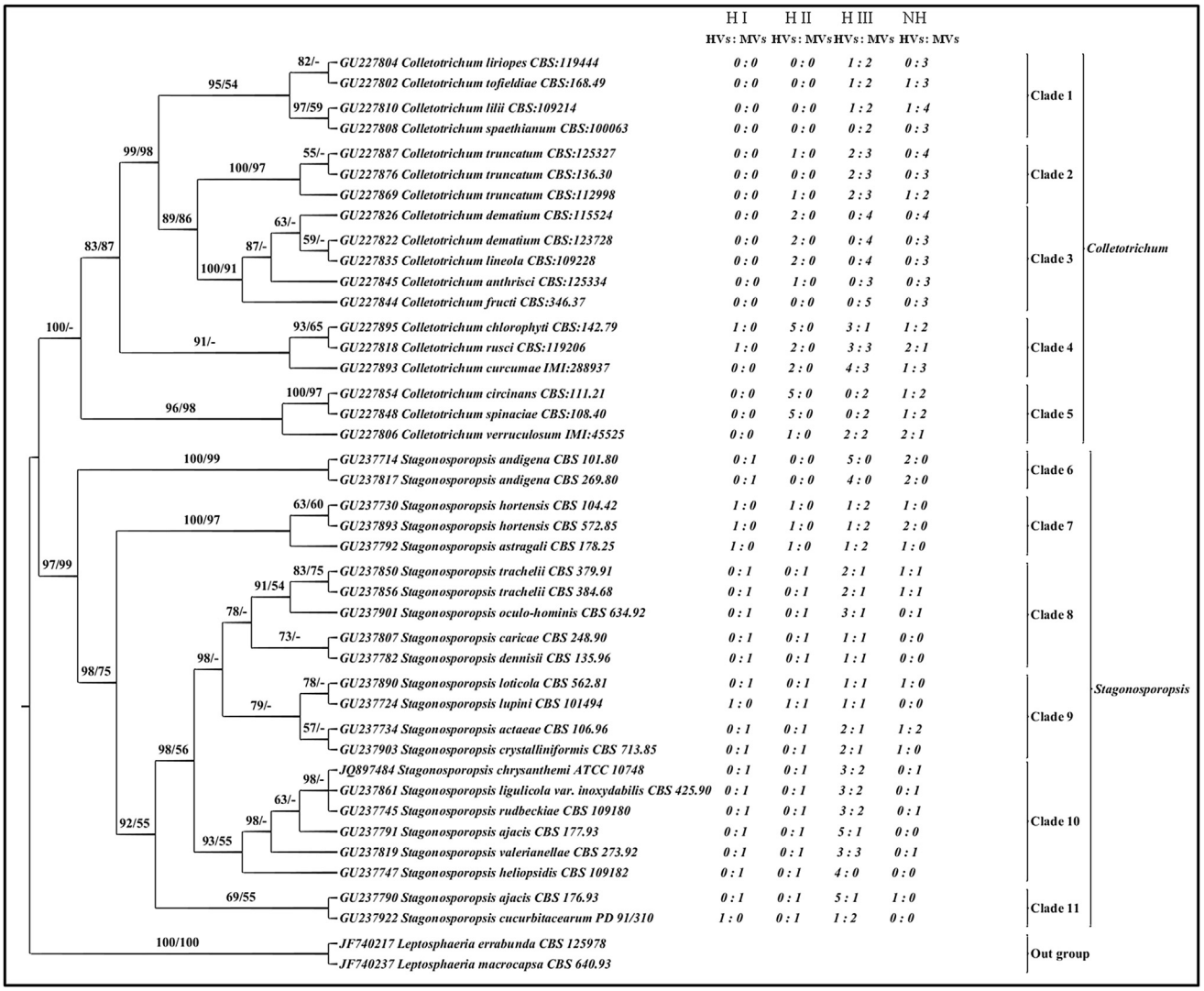

Fig. 4. Phylogram generated from maximum likelihood analysis based on ITS2 sequence – secondary structure of *Colletotrichum* spp. and *Stagnosporopsis* spp. Bootstrap (BP) values ≥50% were represented above the branches as ML BP/NJ BP (ML – maximum likelihood and NJ – neighbour joining). Number of highly variable sites (HVs) and moderately variable sites (MVs) of respective leaves were represented helices wise as HVs:MVs. (HI: Helix I, HII: Helix II, HIII: Helix III and NH: non-helical region).

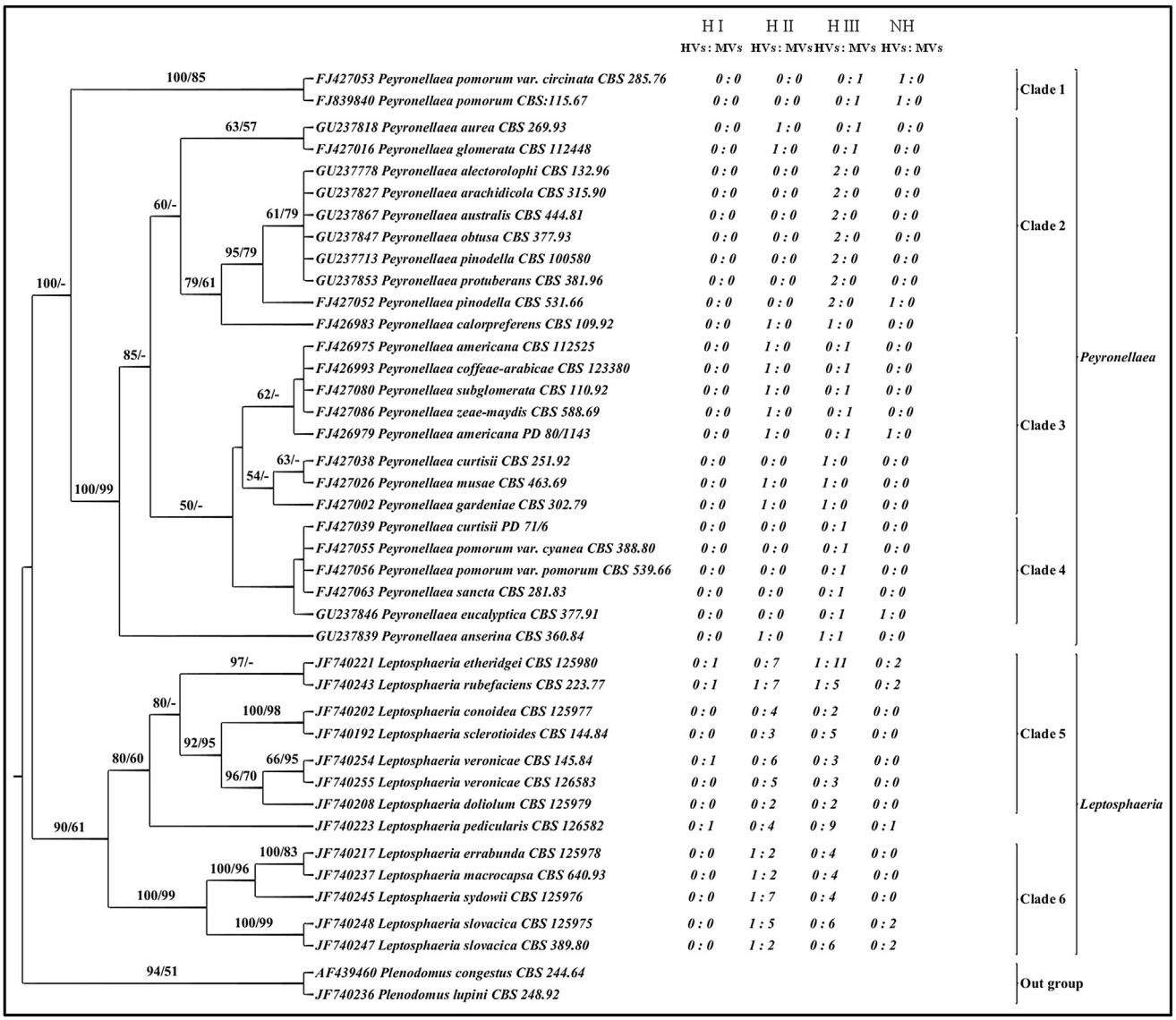

Fig. 5. Phylogram generated from maximum likelihood analysis based on ITS2 sequence – secondary structure of *Peyronellaea* spp. and *Leptosphaeria* spp. Bootstrap (BP) values ≥50% were represented above the branches as ML BP/NJ BP (ML – maximum likelihood and NJ – neighbour joining). Number of highly variable sites (HVs) and moderately variable sites (MVs) of respective leaves were represented helices wise as HVs:MVs. (HI: Helix I, HII: Helix II, HIII: Helix III and NH: non-helical region).

CBS 125978 shared an absolute similarity and resolved in a sub-clade of clade 6.

3.5.3. *Boeremia and Plenodomus*

In Boeremia, 19 species (23 strains) employed in the dataset were resolved into 3 distinct clades (Fig. 6). Clade 1 encompassed 13 strains. *B.exigua* var. *gilvescens* CBS 101156, *B. telephii* CBS 109175, *B. foveata* CBS 109176, *B. exigua* var. *viburni* CBS 100354, *B. exigua* var. *pseudolilacis* CBS 101207, *B. exigua* var. *populi* CBS 100167, *B. exigua* var. *heteromorpha* CBS 101196, *B. exigua* var. *forsythia* CBS.101197, *B. exigua* var. *exigua* CBS 101150 featured absolute similarity. Similarly, *B. exigua* var. *forsythia* CBS 101212 and *B. lycopersici* CBS 101199 possessed no variations and resolved together in the same clade. *B. sambuci-nigrae* CBS 109170, *B. strasseri* CBS 126.93, *B. hedericola* CBS 366.91, *B. exigua* var. *lilacis* CBS 101206, *B. diversispora* CBS 109.79 and *B. crinicola* CBS 109.79 shared an absolute similarity and resolved together in clade 2. *B. exigua* var. *forsythia* CBS.101197 and CBS 101212, *B. lycopersici* 101199 strains resolved with other species and also their species specific variations sites shared similarities with other species. *B. lycopersici* CBS 101200 and CBS 109172 possessed 1 species specific variation sites and 1 strain specific variation sites. Likewise, *B. exigua* var. *linicola* CBS 109.49 and CBS 112.28 possessed 1 species specific variation sites and 1 strain specific variation sites.

In *Plenodomus,* 12 species (14 strains) employed in the dataset were resolved into 4 distinct clades (Fig. 6). *P. agnitus* CBS 121.89 and CBS 126584 possessed 16 species specific variation sites and 1 strain specific variation sites. Similarly, *P. hendersoniae* possessed 11 species specific variation sites and 1 strain specific variation sites.

4. Discussion

Fungal identification and taxonomical regrouping had been constantly evolving albeit slowly, to address the discrepancies and inadequacies in species delimitations (Fig. 7), which inadvertently affects most aspects of mycological studies. The utility of ITS2 as a phylogenetic marker to delineate the species boundaries had been well documented in animals, plants, and protistans (Coleman,

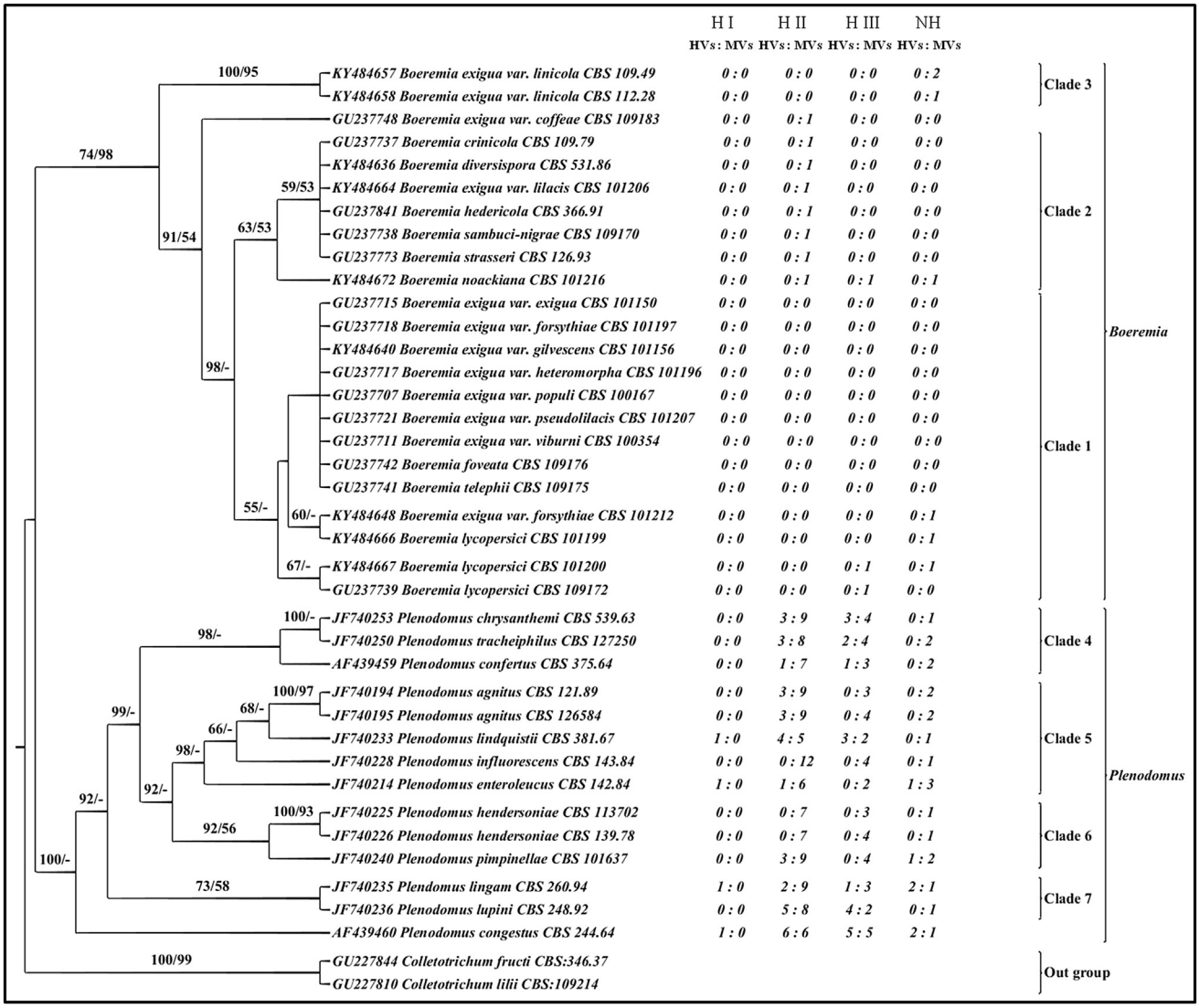

Fig. 6. Phylogram generated from maximum likelihood analysis based on ITS2 sequence – secondary structure of *Boeremia* spp. and *Plenodomus* spp. Bootstrap (BP) values ≥50% were represented above the branches as ML BP/NJ BP (ML – maximum likelihood and NJ – neighbour joining). Number of highly variable sites (HVs) and moderately variable sites (MVs) of respective leaves were represented helices wise as HVs:MVs. (HI: Helix I, HII: Helix II, HIII: Helix III and NH: non-helical region).

2007). In this study, full length of ITS and ITS2 sequence based phylogenetic analysis were performed preceding the ITS2 sequence-secondary structure based molecular morphometrics analysis. ITS2 region possess comparable phylogenetic significance as that of the full length of ITS region. Erstwhile, Yang et al. (2018) showed that the taxonomic preferences of ITS and ITS2 were similar in fungi and concluded that ITS2 alone might be a more suitable marker to delineate the species of fungal community. Similar observations were documented in basidiomycetes (Badotti et al., 2017) and plants (Wang et al., 2016; Sheth et al., 2018). Hence, we progressed to investigate the ITS2 molecular morphometrics in detail. Molecular morphometrics has been the recent revolution in species delimitation, demonstrated to be more informative than the erstwhile methods. CBC, a hallmark of molecular morphometric analysis, has been recognized as species delimiter. However, lack of CBCs in species comparisons are futile in delimiting species boundaries. Henceforth, in this study, we explored the utilities of substitution and INDEL events in molecular morphometric based species delimitation.

ITS2 database has been a stride in ITS2 based comparative phylogenetics research. The implementation of iterative HM for structure prediction in ITS2 database renders 4 helical patterns universally (Koetschan et al., 2010). Three helix pattern structure had been reported in beetles (Ruhl et al., 2010), nematodes of subfamily Anguininae (Ma et al., 2008). Similarly, Coleman (2007) showed that fungal ITS2 secondary structures had 2–4 helical patterns using MFE based structure prediction. Similarly, Landis and Gargas (2007), reported 3–4 helical structures for fungal ITS2. Padhi et al. (2016) also reported 3 helices in *Fusarium* spp. Hence, we intended to investigate the MFE based structure prediction and extend it for molecular morphometrics analysis. Secondary structures predicted based on MFE, yielded structures with 3 helical patterns for most species in the dataset and only in few cases 4 helical patterns were observed. Consensus structure predicted in both methods had the 3 helical pattern. This proves that the 4th helix lacked conservation and corroborates with previous observations that 4th helix may not be always present (Miao et al., 2008; Ruhl et al., 2010). Common structural features were found to

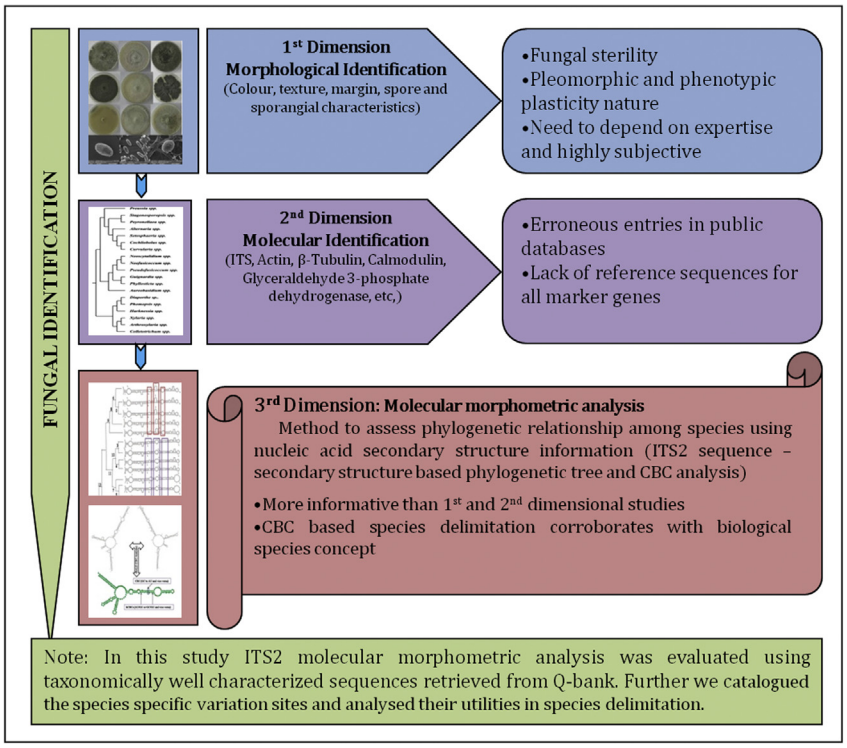

Fig. 7. Schematic representation of the progress in fungal identification approaches.

be mostly congruent in both cases except for the 4th helix. Further, ITS2 sequence–structure based phylogenetic resolution and CBCs among the strains were mostly congruent in both HM and MFE folded structures based analysis (data not shown). Henceforth, structures predicted based on MFE, were taken up for molecular morphometric analysis.

CBC based species delimitations had been shown to concur with biological species concept and predicted with 93% confidence level. However, its utility is tarnished, due to the fact that there is 76% probability for compared species to be same when CBC is 0 (Muller et al., 2007). In this dataset, most species combinations lacked CBCs due to the formation of hemi-CBCs or bulges as a result of evolutionary events or lack of structural compatibility. For instance, in our dataset, *Boeremia* were represented by closely related species and they lacked CBCs despite documentation of evolutionary events. Similarly, *Plenodomus* spp. represented in this dataset were highly divergent which had impacted the structural similarity between the compared species pairs, hence no CBCs were found. Hence, documenting the variations (evolutionary events) in the ITS2 sequences and their distribution in the secondary structure was attempted to define species boundaries.

Evolutionary events had been documented to be phylogenetically informative (Fritz et al., 1994; Caisová et al., 2011; Nagy et al., 2012; Salvi and Mariottini, 2012). In our dataset, among substitution events, the frequency of transition was observed to be higher than the frequency of transversion. These observations corroborate well with erstwhile reports in several markers (Fitch, 1967; Lyons and Lauring, 2017). Also, INDELs were less frequent than the substitutional events in all genus included in this study. Nagy et al. (2012) had observed the same pattern in ITS1 and ITS2 loci. These evolutionary events render both Watson and non-Watson base pairing thereby inducing bulges or mismatches in the secondary structures, additional information in structure based phylogenetic studies (Zhang et al., 2015). The evolutionary events catalogued were majorly distributed in helix II, III and non-helical regions. Least proportion of variations was observed in helix I of all genera investigated in this study, while absolute conservation in this region was evidenced in *Peyronellaea* and *Boeremia*. Miao et al. (2008), similarly observed distribution pattern of several CBCs and hemi-CBCs in helix II and III.

Nevertheless, defining the cut-off range for accounting the variation sites within a species and identifying the speciation event are mandatory for effective use of these data. In this dataset, variation sites accounted in 10 or more species combinations were regarded as HVs sites and variation sites documented in 5–9 species combinations were regarded as MVs sites. Variation sites accounted in fewer species combinations (<5) were regarded as LVs sites and were not included in the discerning species specific variation sites. The number and distribution of HVs and MVs sites were observed to be species specific (Supplementary Table S3). Erstwhile, Caisová et al. (2011) had documented evolutionary events, based on which conservation and variation sites were discerned to evaluate their role in ITS2 evolution and speciation events in Ulvales.

Species specific variation sites were recognized for 68 species included in this study, while the remaining 29 species shared similarities with other species. For example, *C. lineola* CBS 109228 showed absolute similarity with *C. dematium* CBS 123728 and differed with *C. dematium* CBS 115524 by one variation. *C. dematium* CBS 123728 possessed 9 species specific variation sites, which inadvertently overlaps with *C. lineola*. *C. dematium* and *C. lineola* were considered synonym and were recently re-established as separate species by Damm et al. (2009). This had perpetuated misnaming over the years. Similarly, *P. americana* PD 80.1143 possessed one distinct variation site than *P. americana* CBS 112525 which had absolute similarity with *P. coffeae-arabica* CBS 123380, *P. zeaemaydis* CBS 588.69 and *P. subglomerata* CBS 110.92. They were observed to resolve closely in multi-loci phylogeny based on ITS,

LSU and TUB markers. However, these markers failed to resolve the *P. protuberans*, *P. alectorolophi* and *P. obtuse* because of the absolute similarity (Aveskamp et al., 2010).

Interestingly, in *S. trachelii* strains the sequence divergence had not affected the species specific sites (HVs and MVs). The variation may be accounted in LVs site. Contrastingly, in *C. truncatum*, *S. andigena, S. hortensis, L. veronicae, L. slovacica, B. exigua* var. *linicola, P. hendersoniae* and *P. agnitus*, apart from the species specific variation sites, additional variation sites (HVs and MVs) were observed which may provide insight at the strain level. Species specific variation sites recognized in this study is a preliminary result and they could be discerned with absolute confidence when larger datasets encompassing all described species of genera were investigated.

Multi-marker based studies are economically less feasible, inherit inconsistencies in fixing similarity threshold and are prone to phylogenetic incongruence with individual gene-based trees (Begerow et al., 2010; Gazis et al., 2011). Concatenating many such multi-markers with their inherent inconsistencies might bring in unrealistic phylogenetic lineation and rise of a new taxon. In addition, not all reported species possess the complete set of marker data, which would limit its inclusion in the phylogenetic analysis and species delineation of fungal isolates. It is wise to curate and tap into the enormous ITS data; improvise the phylogenetic utility by incorporating the secondary structural data. On the contrary, implications of Intragenomic variations (IaGVs) in ITS phylogeny need to be understood. IaGVs are distributed throughout the length of ITS, however, more pronounced in ITS1 and relatively less in ITS2. Despite the observance of multiple haplotypes as a result of IaGVs, the dominant haplotypes corresponded to the sequence obtained by Sanger sequencing method using PCR primers (Alanagreh et al., 2017; Wolf et al., 2013), implying they have no impact on ITS2 bases phylogenetic analysis. Few contradictions to the use of ITS2 due to intragenomic variations were also reported in insect groups (*Agrodiaetus* of *Lepidoptera* (Shapoval and Lukhtanov, 2015), *Pediculus humanus* of Phthiraptera (Leo and Barker, 2002)). However, the detailed investigation on intragenomic variations in fungal ITS2 and its influence in phylogenetic analysis is lacking and could be a potential scope of future research.

Implementation and exploration of ITS2 sequence secondary structure information provide a strong impetus for fungal species delimitation. The result of the current study also validates the use of ITS2 molecular morphometric analysis and reveals an insight of species specific variations in species delineation. They could be of potential use in barcoding fungi at the species level. Extending the study with larger datasets could help in redefining the species boundaries of contentious species and regrouping the inaccurately named sequences, thereby alleviating the problems in fungal taxonomy and identification. More concerted efforts are mandatory to resolve the pertinent issues and problems in identification otherwise the sequence based identification would turn obsolete like the morphological identification at the species level in the near future.

Author contributions

MP and NS designed the experiments and revised the manuscript. NS performed the experiments and the data were analysed with the help of AKS. NS and EGJ wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest regarding the present study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.funbio.2019.05.002.

References

Ahvenniemi, P., Wolf, M., Lehtonen, M.J., Wilson, P., German-Kinnari, M., Valkonen, J.P.T., 2009. Evolutionary diversification indicated by compensatory base changes in ITS2 secondary structures in a complex fungal species, *Rhizoctonia solani*. J. Mol. Evol. 69, 150–163. https://doi.org/10.1007/s00239-009-9260-3.

Alanagreh, L., Pegg, C., Harikumar, A., Buchheim, M., 2017. Assessing intragenomic variation of the internal transcribed spacer two: adapting the Illumina metagenomics protocol. PLoS One 12, e0181491.

Arnold, A., Maynard, Z., Gilbert, G., Coley, P., Kursar, T., 2000. Are tropical fungal endoyphytes hyperdiverse? Ecol. Lett. 3, 267–274.

Athokpam, V.D., Tandon, V., 2014. Morphological and molecular characterization of *Posthodiplostomum* sp. (Digenea: Diplostomidae) metacercaria in the muscles of snakeheads (*Channa punctata*) from Manipur, India. Helminthologia 51, 141–152. https://doi.org/10.2478/s11687-014-0221-z.

Aveskamp, M.M., de Gruyter, J., Woudenberg, J.H.C., Verkley, G.J.M., Crous, P.W., 2010. Highlights of the Didymellaceae: a polyphasic approach to characterise *Phoma* and related pleosporalean genera. Stud. Mycol. 65, 1–60. https://doi.org/10.3114/sim.2010.65.01.

Badotti, F., Oliveira, F.S., Garcia, C.F., Vaz, A.B.M., Fonseca, P.L.C., Nahum, L.A., Oliceira, G., Goes-Neto, A., 2017. Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). BMC Microbiol. 17, 42. https://doi.org/10.1186/s12866-017-0958-x.

Banerjee, A.K., Arora, N., Murthy, U.S., 2008. Stability of ITS2 secondary structure in *Anopheles*: what lies Beneath ? Int. J. Integr. Biol. 1, 232–238.

Bargues, M.D., Zuriaga, M.A., Mas-Coma, S., 2014. Nuclear rDNA pseudogenes in chagas disease vectors: evolutionary implications of a new 5.8S+ITS-2 paralogous sequence marker in triatomines of North, Central and northern South America. Infect. Genet. Evol. 21, 134–156. https://doi.org/10.1016/j.meegid.2013.10.028.

Barik, B.P., Tayung, K., Jagadev, P.N., 2011. Molecular phylogeny and RNA secondary structure of Fusarium species with different lifestyles. Plant Pathol. Quar. 1, 205–219. https://doi.org/10.5943/ppq/1/2/5.

Bass, D., Richards, T.A., 2011. Three reasons to re-evaluate fungal diversity “on Earth and in the ocean. Fungal Biol. Rev. 25, 159–164. https://doi.org/10.1016/j.fbr.2011.10.003.

Bazzicalupo, A.L., Bálint, M., Schmitt, I., 2013. Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. Fungal. Ecol. 6, 102–109. https://doi.org/10.1016/j.funeco.2012.09.003.

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. https://doi.org/10.1007/s00253-010-2585-4.

Billoud, B., Guerrucci, M.A., Masselot, M., Deutsch, J.S., 2000. Cirrepede phylogeny using an novel approach: molecular morphometrics. Mol. Biol. Evol. 17, 1435–1445.

Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M., Kauserud, H., 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. Mol. Ecol. Resour. 13, 218–224. https://doi.org/10.1111/1755-0998.12065.

Bridge, P.D., Schlitt, T., Cannon, P.F., Buddie, A.G., Baker, M., Borman, A.M., 2008. Domain II hairpin structure in ITS1 sequences as an aid in differentiating recently evolved animal and plant pathogenic fungi. Mycopathologia 166, 1–16. https://doi.org/10.1007/s11046-008-9094-3.

Buchheim, M.A., Keller, A., Koetschan, C., Förster, F., Merget, B., Wolf, M., 2011. Internal transcribed spacer 2 (nu ITS2 rRNA) sequence-structure phylogenetics: towards an automated reconstruction of the green algal tree of life. PLoS One 6, e16931. https://doi.org/10.1371/journal.pone.0016931.

Cai, L., Hyde, K., Taylor, P., Weir, B., Waller, J., Abang, M., Zhang, J., Yang, Y., Phoulivong, S., Liu, Z., Prihastuti, H., Shivas, R., McKenzie, E., Johnston, P., 2009. A polyphasic approach for studying *Colletotrichum*. Fungal Divers. 39, 183–204.

Caisová, L., Marin, B., Melkonian, M., 2013. A consensus secondary structure of ITS2 in the Chlorophyta identified by phylogenetic reconstruction. Protist 164, 482–496. https://doi.org/10.1016/j.protis.2013.04.005.

Caisová, L., Marin, B., Melkonian, M., 2011. A close-up view on ITS2 evolution and speciation – a case study in the Ulvophyceae (Chlorophyta, Viridiplantae). BMC Evol. Biol. 11, 262. https://doi.org/10.1186/1471-2148-11-262.

Chen, C.A., Chang, C.C., Wei, N.V., Chen, C.H., Lein, Y.T., Lin, H.E., Dai, C.F., Wallace, C.C., 2004. Secondary structure and phylogenetic utility of the ribosomal internal transcribed Spacer 2 (ITS2) in Scleractinian corals. Zool. Stud. 43, 759–771.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., Leon, C., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5, e8613. https://doi.org/10.1371/journal.pone.0008613.

Coleman, A.W., 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res. 35, 3322–3329. https://doi.org/10.1093/nar/gkm233.

Coleman, A.W., 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19, 370–375. https://doi.org/10.1016/S0168-9525(03)00118-5.
Damm, U., Woudenberg, J.H.C., Cannon, P.F., Crous, P.W., 2009. *Colletotrichum* species with curved conidia from herbaceous hosts. Fungal Divers. 39, 45–87.
Davis, S., 2010. DNA sequencing of fungi in a microbiology laboratory. Microbiol. Aust. 142–144.
Diaz, P.L., Hennell, J.R., Sucher, N.J., 2012. Genomic DNA extraction and barcoding of endophytic fungi. In: Sucher, N.J., Hennell, J.R., Carlef, M.C. (Eds.), Plant DNA Fingerprinting and Barcoding: Methods and Protocols, Methods in Molecular Biology. Science + Business Media, Springer, pp. 171–179. https://doi.org/10.1007/978-1-61779-609-8.
Ferro, M., Antonio, E.A., Souza, W., Bacci, M., 2014. ITScan: a web-based analysis tool for internal transcribed spacer (ITS) sequences. BMC Res. Notes 7, 1–4. https://doi.org/10.1186/1756-0500-7-857.
Fitch, W.M., 1967. Evidence suggesting a non-random character to nucleotide replacements in naturally occurring mutations. J. Mol. Biol. 26, 499–507. https://doi.org/10.1016/0022-2836(67)90317-8.
Fritz, G.N., Conn, J., Cockburn, A., Seawright, J., 1994. Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae). Mol. Biol. Evol. 11, 406–416. https://doi.org/10.1093/oxfordjournals.molbev.a040122.
Gao, T., Yao, H., Song, J., Zhu, Y., Liu, C., Chen, S., 2010. Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evol. Biol. 10, 324. https://doi.org/10.1186/1471-2148-10-324.
Gazis, R., Rehner, S., Chaverri, P., 2011. Species delimitation in fungal endophyte diversity studies and its implications in ecological and biogeographic inferences. Mol. Ecol. 20, 3001–3013. https://doi.org/10.1111/j.1365-294X.2011.05110.x.
Hawksworth, D.L., 2004. Fungal diversity and its implications for genetic resource collections. Stud. Mycol 50, 9–18. https://doi.org/10.5598/imafungus.2011.02.01.14.
Hawksworth, D.L., 2001. The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol. Res. 105, 1422–1432. https://doi.org/10.1017/S0953756201004725.
Hawksworth, D.L., 1992. The need for a more effective biological nomenclature for the 21st century. Bot. J. Linn. Soc. 109, 543–567.
Hawksworth, D.L., 1991. The fungal dimension of biodiversity: magnitude, significance and conservation. Mycol. Res. 95, 641–655.
Hawksworth, D.L., Crous, P.W., Redhead, S.A., Reynolds, D.R., Samson, R.A., Seifert, K.A., Taylor, J.W., Al, E., 2011. The Amesterdam declaration of fungal nomenclature. IMA Fungus 2, 105–112. https://doi.org/10.5598/imafungus.2011.02.01.14.
Heeg, J.S., Wolf, M., 2015. ITS2 and 18S rDNA sequence-structure phylogeny of Chlorella and allies (Chlorophyta, Trebouxiophyceae, Chlorellaceae). Plant Gene 4, 20–28. https://doi.org/10.1016/j.plgene.2015.08.001.
Ihrmark, K., Bödeker, I.T.M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandström-Durling, M., Clemmensen, K.E., Lindahl, B.D., 2012. New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol. Ecol. 82, 666–677. https://doi.org/10.1111/j.1574-6941.2012.01437.x.
Koetschan, C., Förster, F., Keller, A., Schleicher, T., Ruderisch, B., Schwarz, R., Müller, T., Wolf, M., Schultz, J., 2010. The ITS2 database III sequences and structures for phylogeny. Nucleic Acids Res. 38, D275–D279. https://doi.org/10.1093/nar/gkp966.
Kusari, S., Singh, S., Jayabaskaran, C., 2014. Biotechnological potential of plant-associated endophytic fungi: hope versus hype. Trends Biotechnol. 32, 297–303.
Landis, F.C., Gargas, A., 2007. Using ITS2 secondary structure to create species-specific oligonucleotide probes for fungi. Mycologia 99, 681–692. https://doi.org/10.3852/mycologia.99.5.681.
Leo, N.P., Barker, S.C., 2002. Intragenomic variation in ITS2 rDNA in the louse of humans, *Pediculus humanus*: ITS2 is not a suitable marker for population studies in this species. Insect Mol. Biol. 11, 651–657.
Lyons, D.M., Lauring, A.S., 2017. Evidence for the selective basis of transition-to-transversion substitution bias in two RNA viruses. Mol. Biol. Evol. 34, 3205–3215. https://doi.org/10.1093/molbev/msx251.
Ma, H., Overstreet, R.M., Subbotin, S.A., 2008. ITS2 secondary structure and phylogeny of cyst-forming nematodes of the genus *Heterodera* (Tylenchida: Heteroderidae). Org. Divers. Evol. 8, 182–193. https://doi.org/10.1016/j.ode.2007.09.001.
Miao, M., Warren, A., Song, W., Wang, S., Shang, H., Chen, Z., 2008. Analysis of the internal transcribed spacer 2 (ITS2) region of Scuticociliates and related Taxa (Ciliophora, Oligohymenophorea) to infer their evolution and phylogeny. Protist 159, 519–533. https://doi.org/10.1016/j.protis.2008.05.002.
Money, N.P., 2013. Against the naming of fungi. Fungal Biol 117, 463–465. https://doi.org/10.1016/j.funbio.2013.05.007.
Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How many species are there on earth and in the ocean? PLoS Biol. 9, e1001127. https://doi.org/10.1371/journal.pbio.1001127.
Muller, T., Philippi, N., Dandekar, T., Schultz, J., Wolf, M., 2007. Distinguishing species. RNA 13, 1469–1472. https://doi.org/10.1261/rna.617107.
Nagy, L.G., Kocsubé, S., Csanádi, Z., Kovács, G.M., Petkovits, T., Vágvölgyi, C., Papp, T., 2012. Re-Mind the gap! insertion - deletion data reveal neglected phylogenetic potential of the nuclear ribosomal internal transcribed spacer (ITS) of fungi. PLoS One 7, e49794. https://doi.org/10.1371/journal.pone.0049794.
Nilsson, R.H., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K.H., Köljalg, U., 2006. Taxonomic reliability of DNA sequences in public sequences databases: a fungal perspective. PLoS One 1, e59. https://doi.org/10.1371/journal.pone.0000059.
Nilsson, R.H., Veldre, V., Hartmann, M., Unterseher, M., Amend, A., Bergsten, J., Kristiansson, E., Ryberg, M., Jumpponen, A., Abarenkov, K., 2010. An open source software package for automated extraction of ITS1 and ITS2 from fungal ITS sequences for use in high-throughput community assays and molecular ecology. Fungal Ecol 3, 284–287. https://doi.org/10.1016/j.funeco.2010.05.002.
Ozerskaya, S.M., Kochkina, G.A., Ivanushkina, N.E., 2010. Fungal diversity in GenBank: problems and possible solutions. Inoculum 61, 1–4.
Padhi, S., Panda, M.K., Das, D., Tayung, K., 2016. ITS2 RNA secondary structure analysis reveals close affinity between endophytic and pathogenic fungi: a case study in *Fusarium* species. Ann. Microbiol. 66, 625–633. https://doi.org/10.1007/s13213-015-1142-x.
Porras-Alfaro, A., Bayman, P., 2011. Hidden fungi, emergent properties: endophytes and microbiomes. Annu. Rev. Phytopathol. 49, 291–315. https://doi.org/10.1146/annurev-phyto-080508-081831.
Rampersad, S.N., 2014. ITS1, 5.8S and ITS2 secondary structure modelling for intraspecific differentiation among species of the *Colletotrichum gloeosporioides* sensu lato species complex. SpringerPlus 3, 684. https://doi.org/10.1186/2193-1801-3-684.
Réblová, M., Réblová, K., 2013. RNA secondary structure, an important bioinformatics tool to enhance multiple sequence alignment: a case study (Sordariomycetes, Fungi). Mycol. Prog. 12, 305–319. https://doi.org/10.1007/s11557-012-0836-8.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180.
Ruhl, M.W., Wolf, M., Jenkins, T.M., 2010. Compensatory base changes illuminate morphologically difficult taxonomy. Mol. Phylogenet. Evol. 54, 664–669. https://doi.org/10.1016/j.ympev.2009.07.036.
Salvi, D., Mariottini, P., 2012. Molecular phylogenetics in 2D: ITS2 rRNA evolution and sequence-structure barcode from Veneridae to Bivalvia. Mol. Phylogenet. Evol. 65, 792–798. https://doi.org/10.1016/j.ympev.2012.07.017.
Seibel, P.N., Müller, T., Dandekar, T., Schultz, J., Wolf, M., 2006. 4SALE – a tool for synchronous RNA sequence and secondary structure alignment and editing. BMC Bioinf. 7, 498. https://doi.org/10.1186/1471-2105-7-498.
Selig, C., Wolf, M., Müller, T., Dandekar, T., Schultz, J., 2008. The ITS2 database II: homology modelling RNA structure for molecular systematics. Nucleic Acids Res. 36, D377–D380. https://doi.org/10.1093/nar/gkm827.
Shapoval, N., Lukhtanov, V., 2015. Intrgenomic variations of multicopy ITS2 marker in *Agrodiaetus* blue butterflies (Lepidoptera, Lycaenidae). Comp. Cytogenet. 9, 483–497.
Sheth, B.P., Punia, S., Dheer, M., Rakhashiya, P.M., Patel, P.P., Thaker, V.S., 2018. Phylogenetic implications and secondary structure analyses of *Vigna mungo* (L.) Hepper genotypes based on nrDNA ITS2 sequences. Comput. Biol. Chem. 78, 389–397. https://doi.org/10.1016/j.compbiolchem.2018.12.018.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771. https://doi.org/10.1080/10635150802429642.
Taylor, J.W., 2011. One fungus = One Name: DNA and fungal nomenclature twenty years after PCR. IMA Fungus 2, 113–120. https://doi.org/10.5598/imafungus.2011.02.02.01.
Wang, X.Y., Zheng, S.H., Liu, Y., Han, J.P., 2016. ITS2, a better DNA barcode than ITS in identification of species in Artemisia L. Chinese herbal medicines, vol. 8, pp. 352–358.
Wolf, M., Chen, S., Song, J., Ankenbrand, M., Müller, T., 2013. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences – a proof of concept. PLoS One 8, e66726. https://doi.org/10.1371/journal.pone.0066726.
Wolf, M., Koetschan, C., Müller, T., 2014. ITS2, 18S, 16S or any other RNA – simply aligning sequences and their individual secondary structures simultaneously by an automatic approach. Gene 546, 145–149. https://doi.org/10.1016/j.gene.2014.05.065.
Wolf, M., Ruderisch, B., Dandekar, T., Schultz, J., Müller, T., 2008. ProfDistS: (profile-) distance based phylogeny on sequence – structure alignments. Bioinformatics 24, 2401–2402. https://doi.org/10.1093/bioinformatics/btn453.
Yan, D.-H., Gao, Q., Sun, X., Song, X., Li, H., 2018. ITS2 sequence–structure phylogeny reveals diverse endophytic *Pseudocercospora* fungi on poplars. Genetica 146, 187–198. https://doi.org/10.1007/s10709-018-0011-5.
Yang, R.-H., Su, J.-H., Shang, J.-J., Wu, Y.-Y., Li, Y., Bao, D.-P., Yao, Y.-J., 2018. Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. PLoS One 13, e0206428.
Zhang, W., Yuan, Y., Yang, S., Huang, J., Huang, L., 2015. ITS2 secondary structure improves discrimination between medicinal "Mu tong" species when using DNA barcoding. PLoS One 10, e0131185. https://doi.org/10.1371/journal.pone.0131185.