ORIGINAL ARTICLE

Unraveling intragenomic polymorphisms in the high-quality genome of *Hypoxylaceae*: a comprehensive study of the rDNA cistron

Marjorie Cedeño-Sanchez[1,2] · Tian Cheng[1] · Christopher Lambert[1,2,3] · Miroslav Kolarík[4] · Eric Kuhnert[5] · Russell J. Cox[5] · Jörn Kalinowski[6] · Bart Verwaaijen[6,7] · Marc Stadler[1,2]

Abstract

Sequencing ribosomal DNA loci, such as ITS and LSU, are among the first steps to characterize and identify newly isolated fungal strains. Substantial doubt has recently been raised concerning the unequivocal suitability of using this sequence information for inferring inter- and intraspecific relationships, as especially the ITS locus was repeatedly shown to harbor high intragenomic plasticity. The evolutionary mechanisms leading to this phenomenon are poorly understood, but this is about to change, owing to the availability of third generation sequencing techniques. The current study attempts a comprehensive analysis of rDNA cistrons in the genomes of over 40 selected species of *Hypoxylaceae* (*Xylariales*), using a combination of 2nd generation (Illumina) and 3rd generation sequencing technologies (Oxford Nanopore), along with extensive usage of bioinformatic tools to obtain high-quality genomes. Our findings revealed high intragenomic polymorphisms in distinct *Hypoxylaceae* species, indicating the presence of deep rDNA paralogs. Phylogenetic inference of the ITS rDNA showed low support for the primary clades within *Hypoxylaceae*, while inference using the protein encoding gene *RPB2* showed better support. We also found that proteinogenic genes are not prone to intraspecific and intragenomic polymorphisms as much as the rDNA loci. Our study highlights the limitations of solely relying on the rDNA cistron as a barcode for fungi. We propose *TUB2* as a new primary barcoding marker for *Hypoxylaceae* and other *Xylariales*. Moreover, we demonstrate the feasibility to retrieve some DNA loci that had hitherto not been used for classical phylogenetic studies (such as *TEF1*) in *Xylariales* directly from genomes without the necessity to work with primers.

Keywords Barcoding · Genomics · Phylogeny · Sanger sequencing · *Xylariales*

Introduction

The internal transcribed spacer (ITS) rDNA has been widely used as the primary barcode marker for fungi due to its high variability and being easily amplifiable by PCR.

Additionally, the LSU region has been used to complement ITS-based identifications (Klaubauf et al. 2010; Schoch et al. 2012; Heeger et al. 2018; Ceballos-Escalera et al. 2022). ITS-driven barcoding and even species delimitation has a long history of being problematic in the order *Xylariales* (*Ascomycota*). For instance, the genus *Annulohypoxylon*,

Section Editor: Marco Thines

✉ Marc Stadler
marc.stadler@helmholtz-hzi.de

[1] Department Microbial Drugs, Helmholtz-Zentrum für Infektionsforschung GmbH, Inhoffenstrasse 7, 38124 Braunschweig, Germany

[2] Institute of Microbiology, Technische Universität Braunschweig, Spielmannstraße 7, 38106 Braunschweig, Germany

[3] Division Molecular Cell Biology, Zoological Institute, Technische Universität Braunschweig, Spielmannstrasse 7, 38106 Braunschweig, Germany

[4] Institute of Microbiology of the ASCR, v.v.i., Vídeňská 1083, 14220 Praha, Czech Republic

[5] Centre of Biomolecular Drug Research (BMWZ), Institute for Organic Chemistry, Leibniz University Hannover, Schneiderberg 38, 30167 Hannover, Germany

[6] Center for Biotechnology (CeBiTec), Bielefeld University, Universitätsstraße 27, 33615 Bielefeld, Germany

[7] Department of Genetics, Martin-Luther-University-Halle-Wittenberg, Halle (Saale), Germany

which was rendered paraphyletic in older ITS-based phylogenetic analyses (Sánchez-Ballesteros et al. 2000). Proteinogenic sequences derived from *TUB2* and *α-actin* unequivocally resolved into a monophyletic clade comprising *Annulohypoxylon* (Hsieh et al. 2005), which is identical with *Hypoxylon* sect. *Annulata sensu* Ju and Rogers. The genus was later further subdivided, following a multi-genealogy and incorporating chemotaxonomic and morphological evidence, and *Jackrogersella* was split from *Annulohypoxylon* (Wendt et al. 2018). The concurrent resurrection of the *Hypoxylaceae* and the division of the *Xylariaceae s. lat.* into different families would not have been feasible based on rDNA data alone.

A lack of specificity of ITS data not only affects generic segregation, but also species recognition. Members of the genus *Daldinia*, as well as the *Hypoxylon rubiginosum* and *H. fuscum* complexes, are known to have nearly identical ITS sequences and can only be segregated by polyphasic approaches (Stadler et al. 2013; Kuhnert et al 2014; Lambert et al. 2021). Beneath homologous ITS sequences between well-resolved species, Stadler et al. (2020) even reported on finding heterogeneous sequences among the multiple copies of distinct ITS loci found in high-quality genomes generated with 3rd generation sequencing technology (Oxford Nanopore), polished with 2nd generation sequencing technology (Illumina). How these findings of the pilot study featuring 14 genome sequences extend to other members of the *Hypoxylaceae* have so far not been followed up upon. Here, we attempt to fill this gap by providing genome sequences for an additional 44 strains of *Xylariales*, including several ex-type cultures.

Materials and methods

A total of 44 genome sequences from *Xylariales* were newly generated and subsequently used for the current study. All details are provided in Table 1.

Genomic DNA preparation and extraction

All fungal strains were grown in 500-mL Erlenmeyer flasks with 100 mL YMG medium) (10 g malt extract, 4 g glucose, 4 g yeast extract ad 1 L deionized water; pH 6,3) and then placed in a shaking incubator at 220 rpm and 25 °C, for 3 to 5 days depending on growth speed of the fungus. The mycelium was harvested using a Büchner funnel with filter paper (MN 640 w, Macherey-Nagel, Düren, Germany) by vacuum filtration, then frozen with liquid nitrogen and ground to fine powder in a mortar. DNA extraction was performed with the GenElute® Plant Genomic DNA Miniprep Kit (Sigma Aldrich, St. Louis, MO, USA), following manufacturer's instructions.

Instructions for Nanopore and Illumina library preparations were followed as reported in Stadler et al. (2020).

Assembly

Basecalling of Nanopore reads was done live with the Guppy algorithm embedded in the MinKNOW platform on GridION (Oxford Nanopore Technologies). Adapters were trimmed using Porechop v0.2.4 (https://github.com/rrwick/Porechop). Genome assemblies were performed using canu v2.1.1 (Koren et al. 2017). This assembly was then polished with Illumina short read data using Pilon v1.24 (Walker et al. 2014), applying four cycles with bwa-mem v0.7.17 (Li 2013) and four further rounds with Bowtie2 v2.4.4 (Langmead and Salzberg 2012). Mitochondrial contigs and assembly artifacts were removed based on coverage and a BLAST (Camacho et al. 2009) search against published mitochondrial sequences. The remaining nuclear genome contigs were further curated and sorted with the "clean" option included in funannotate v1.8.16 (Palmer and Stajich 2020).

Identification of ITS, LSU, *TUB2*, *RPB2*, *TEF1*, and *ACT1* sequence copies

Genome sequences were used to create individual BLAST databases in Geneious prime 2023.4 (https://www.geneious.com). Previously published sequences for ITS, LSU, *TUB2*, *RPB2*, *ACT1*, and *TEF1* were used as homology search templates to locate each region in the genomes and mapped in the query genome. All sequences obtained were aligned using the ClustalW algorithm with standard settings provided by Geneious prime 2023.4 (https://www.geneious.com) and edited manually, if necessary. The ITS (ITS1, 5.8S, and ITS2) sequence of *Hypoxylon sporistriatatunicum* DSM 115550 (GenBank Acc. No. MN056426) was employed as a reference for obtaining the ITS region. The minimum and maximum percentages of homology, as well as maximum bp deviation, were obtained after sequence alignment.

Molecular phylogenetic inference

Extracted ITS, LSU sequences were aligned by using MAFFT online (http://mafft.cbrc.jp/ alignment/server/, Katoh et al. 2019). A maximum-likelihood phylogenetic tree was constructed using IQ-TREE v. 2.1.3 [-b 1000 -abayes -m MFP] (Minh et al. 2020), with selection of the appropriate nucleotide exchange model by ModelFinder (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017) based on Bayesian information criterion. Branch support was calculated with non-parametric bootstrap (Felsenstein 1985) and approximate Bayes test (Anisimova et al. 2011). In total, 1000

Table 1 Details of the genome sequences selected from *Xylariales*, including strain IDs, sequencing methods, and references where original sequence data have previously been described. O/I, Oxford Nanopore/Illumina; PB, PacBio. Type specimens are labeled with T (holotype), IT (isotype), PT (paratype), and ET (epitype)

Species	Strain	Sequencing method	Genome sequenced reference
Annulohypoxylon michelianum	CBS 119993	O/I	This study
A. truncatum (ET)	CBS 140778	O/I	Wibberg et al. 2021
Creosphaeria sassafras	CBS 127876	O/I	This study
Daldinia childae (T)	CBS 122881	O/I	This study
D. concentrica	CBS 113277	O/I	Wibberg et al. 2021
Hypomontagnella. monticulosa	MUCL 54604	O/I	Wibberg et al. 2021
Hyp. spongiphila (T)	CLL-205	O/I	Wibberg et al. 2021
Hyp. submonticulosa	DAOMC 242471; CBS 115280	O/I	Wibberg et al. 2021
Hypoxylon addis (T)	MUCL 52797	O/I	This study
H. duranii (T)	ATCC 58730	O/I	This study
H. dussii (T)	MUCL 53766	O/I	This study
H. erythrostoma	MUCL 53759	O/I	This study
H. fragiforme (ET)	MUCL 51264	PB	Wibberg et al. 2021
H. fuscum (ET)	CBS 113049	O/I	This study
H. gibriacense (T)	MUCL 52698	O/I	This study
H. guilanense (T)	MUCL 57726	O/I	This study
H. haematostroma (ET)	MUCL 53301	O/I	This study
H. hinnuleum (T)	MUCL 3621	O/I	This study
H. hypomiltum	MUCL 53306	O/I	This study
H. invadens (T)	MUCL 54175	O/I	This study
H. investiens	MUCL 53316	O/I	This study
H. lateripigmentum (T)	MUCL 53304	O/I	This study
H. lechatii (T)	MUCL 54609	O/I	This study
H. lienhwacheense	MFLUCC 14-1231	O/I	Wibberg et al. 2021
H. macrocarpum	CBS 119012	O/I	This study
H. olivaceopigmentum (T)	DSM 107924	O/I	This study
H. perforatum	DSM 107930	O/I	This study
H. petriniae (T)	CBS 114746	O/I	This study
H. porphyreum	CBS 119022	O/I	This study
H. pulicicidum	CBS 122622, MUCL 49879	O/I	Wibberg et al. 2021
H. rickii (ET)	MUCL 53309	O/I	Wibberg et al. 2021
H. rubiginosum (ET)	MUCL 52887	PB	Wibberg et al. 2021
H. rubiginosum	DSM 106870	O/I	This study
H. sporistriatatunicum (T)	DSM 115550	O/I	This study
H. texense (T)	DSM 107933	O/I	This study
H. ticinense	CBS 115271	O/I	This study
H. trugodes (ET)	MUCL 54794	O/I	This study
Jackrogersella cohaerens	CBS 119126	O/I	This study
J. minutella	CBS 135445	O/I	This study
J. multiformis (ET)	CBS 119016	O/I	Wibberg et al. 2021
Parahypoxylon papillatum (T)	ATCC 58729	O/I	This study
Pa. ruwenzoriense (T)	MUCL 51392	O/I	This study
Pyrenopolyporus hunteri (ET)	MUCL 49339	O/I	Wibberg et al. 2021
Xylaria hypoxylon (ET)	CBS 122620	O/I	Wibberg et al. 2021

bootstrap replicates were mapped onto the ML tree with the best (highest) ML score. Single locus trees were calculated following the same methodology and checked for congruence with the multigene phylogenetic tree.

GC content and ITS secondary structure prediction

Intragenomic polymorphism of two strains was remarkably high (*Pa. papillatum* ATCC 58729, *Hyp. monticulosa* MUCL 54604). We investigated the notion if this is due to the polymorphism actually representing pseudogenes by following the workflow of Kolařík et al. (2021) and Prahl et al. (2021). Briefly, similarity in *p*-distance was calculated in MEGA11 software (Kumar et al. 2018). The GC content of the ITS region and its components (ITS1, 5.8S and ITS2) was calculated in DaMBE version 6.4.42 (Xia and Xie 2001). Modeling of hybridization of a proximal stem containing the 5.8S and 28S motifs that delimitate the secondary structure of the ITS2 was done in the web interface Internal Transcribed Spacer 2 Ribosomal RNA Database, ITS2-DB, (http://its2.bioapps.biozentrum.uni-wuerzburg.de/, Ankenbrand et al. 2015) using its "Annotate" tool, which is based on hidden Markov models (HMMs) (Keller et al. 2009). In order to predict the folding of the ITS2 secondary structure, we used an expected value for detection of significant hits below 0.001 (*E*-value < 0.001) and HMMs for fungal organisms with a minimum size of 150 nucleotides. We selected only those secondary structures that were obtained by direct folding and preferred over other modeled secondary structures. Free energy structures of the ITS2 were obtained using the RNAfold online tool, which were visualized using a force directed graph layout (forna) (Kerpedjiev et al. 2015) via the web application provided by ViennaRNA Web Services (http://rna.tbi.univie.ac.at/).

Results

A comprehensive analysis was conducted on the genomes of 44 strains of *Xylariales*. We report the generation of 30 new genome sequences derived from *Hypoxylaceae* and one from *Lopadostomataceae*, enabling taxonomic marker studies of unprecedented detail in this group. Hence, we sought to evaluate the suitability of different frequently used protein-encoding genes (*TUB2*, *RPB2*) and we explored *TEF1* and *ACT1* for phylogenetic reconstruction as well for their suitability as potential barcodes and to encompass the extent of observable intragenomic variability of rDNA. With the exception of the here reported *H. dussii* genome harboring two copies of *TEF1*, all proteinogenic genes were found to contain only a single copy, while we encountered multiple copies and intragenomic polymorphisms in the ITS (ITS1, 5.8S, ITS2) and LSU region. As previously

discussed (Stadler et al. 2020), we cannot exclude that the actual number of copies is higher in the genomes as alignment or sequencing errors can still not be excluded even if the current, 3rd generation technologies are used. Therefore, the numbers given below reflect at least those that are actually present in the assembled genomes, but possibly, more paralogs could exist in reality.

Intragenomic variation in the ITS and LSU regions

The analysis of the strains revealed interesting findings regarding identity and the number of copies found in the corresponding genome. The majority of the strains studied showed varying degrees of identity among ITS and LSU, i.e., sequences displaying 100% identity in the ITS region but not in the LSU region, as well as vice versa. Nevertheless, for 12 strains, sequences were identical (100%) or showed minor deviances for 11 strains 98–99.9% for both regions (ITS and LSU). The number of copies of ITS and LSU located in one genome was shown to diverge as well. In our data set, approximately 22% of the strains analyzed displayed different numbers of ITS and LSU copies. For instance, *J. cohaerens* showed the highest variety (14 ITS and 10 LSU copies), followed by *Pa. papillatum* with 11 ITS and nine LSU copies. The genomes of *H. addis*, *H. dussii*, *H. guilanense*, *H. sporistriatatunicum*, and *Hyp. monticulosa* were observed to harbor only one extra copy of ITS. For *P. hunteri* and *X. hypoxylon*, one extra copy of LSU was found. Even though we cannot be sure if these phenomena are due to sequencing or assembly artifacts, we would like to point them out here. As already discussed by Stadler et al. (2020), the real numbers of these loci in the genomes could actually even be higher.

Notably, *H. petriniae* exhibited the highest number of copies in its genome, with 40 ITS and 38 LSU copies, showing identity of 99.1 to 100% and 98.5 to 100%, respectively, followed by *H. guilanense* with 21 ITS and 20 LSU copies with an identity of 99.3 to 100% and 99.4 to 100%, respectively. These findings add to the record of observed genetic complexity and diversity within the studied strains. The results are summarized in Table 2.

Two different genomes of *H. rubiginosum* were analyzed. The strain DSM 106870 is an endophyte isolated from *Fraxinus excelsior* as antagonist of *Hymenoscyphus fraxineus* (Halecker et al. 2020), and its genome was sequenced by Oxford Nanopore technology in the current study. The ex-epitype strain MUCL 52887 was isolated from the wood of *Fagus sylvatica* (Wendt et al. 2018), and its genome had been sequenced using PACBIO (Stadler et al 2020). We were unable to retrieve neither ITS nor LSU regions from MUCL 52887, whereas the genome sequence of strain DSM 10687 exhibited five copies for both. Nevertheless, the ITS obtained from strain MUCL 52887 through Sanger

Table 2 Details of the numbers of copies found in ITS and LSU and % identity divergence for all genome sequences studied

Species	Strain	# of ITS copies	Maximum deviation (# of bp)	% identity		# of LSU copies	Maximum deviation (# of bp)	% identity	
				Minimal	Maximal			Minimal	Maximal
Annulohypoxylon michelianum	CBS 135445	8	0	100	100	8	2	100	100
A. truncatum (ET)	CBS 140778	3	3	99.4	100	3	69	98.7	99.7
Creosphaeria sassafras	CBS 127876	6	0	100	100	6	1	100	100
Daldinia childiae (T)	CBS 122881	6	0	100	100	6	1	100	100
D. concentrica	CBS 113277	3	1	99.8	100	3	37	98.9	99.9
Hypomontagnella monticulosa	MUCL 54604	4	96	88.2	99.9	3	42	99.1	99.6
Hyp. spongiphila (T)	CLL-205	5	0	100	100	5	12	99.7	100
Hyp. submonticulosa	DAOMC 242471	1	NA	NA	NA	1	NA	NA	NA
Hypoxylon addis (T)	MUCL 52797	8	1	99.8	100	7	5	99.9	100
H. duranii (T)	ATCC 58730	7	0	100	100	7	0	100	100
H. dussii (T)	MUCL 53766	8	0	100	100	7	1	100	100
H. erythrostoma	MUCL 53759	5	0	100	100	5	3	99.9	100
H. fragiforme (ET)	MUCL 51264	19	4	99.3	100	19	109	96.9	100
H. fuscum (ET)	CBS 113049	8	0	100	100	8	3	99.9	100
H. gibriacense (T)	MUCL 52698	15	0	100	100	15	49	98.9	100
H. guilanense (T)	MUCL 57726	21	4	99.3	100	20	25	99.4	100
H. haematostroma (ET)	MUCL 53301	8	0	100	100	8	8	99.8	100
H. hinnuleum (T)	MUCL 3621	8	0	100	100	8	0	100	100
H. hypomiltum	MUCL 53306	11	3	99.6	100	11	39	98.9	100
H. invadens (T)	MUCL 54175	7	0	100	100	7	0	100	100
H. investiens	MUCL 53316	6	0	100	100	6	6	99	100
H. lateripigmentum (T)	MUCL 53304	21	2	99.6	100	21	10	99.8	100
H. lechatii (T)	MUCL 54609	6	0	100	100	6	0	100	100
H. lienhwacheense	MFLUCC 14-1231	4	23	96.4	97.8	4	98	96.9	99.8
H. macrocarpum	CBS 119012	8	19	96.3	100	8	59	98.3	100
H. olivaceopigmentum (T)	DSM 107924	10	0	100	100	10	1	100	100
H. perforatum	DSM 107930	11	0	100	100	11	11	99.8	100
H. petriniae (T)	CBS 114746	41	5	99.1	100	38	69	98.5	100
H. porphyreum	CBS 119022	6	1	99.8	100	6	2	100	100
H. pulicicidum	ATCC 74245	4	13	98.2	100	4	85	98.1	99.6
H. rickii (ET)	MUCL 53309	4	6	98.9	99.8	4	25	99.3	99.8
H. rubiginosum (ET)	MUCL 52887	0	NA	NA	NA	1	NA	NA	NA
H. rubiginosum	DSM 106870	5	0	100	100	5	32	99.4	100
H. sporistriatatunicum (T)	DSM 115550	5	0	100	100	4	2	100	100
H. texense (T)	DSM 107933	6	0	100	100	6	0	100	100
H. ticinense	CBS 115271	7	1	99.8	100	7	1	100	100
H. trugodes (ET)	MUCL 54794	3	1	99.8	100	3	0	100	100
Jackrogersella cohaerens	CBS 119126	14	1	99.8	100	10	5	99.8	100
J. minutella	CBS 135445	8	0	100	100	8	8	99.8	100
J. multiformis (ET)	CBS 119016	3	1	99.8	100	3	19	99.4	99.8
Parahypoxylon papillatum (T)	ATCC 58729	11	57	89.2	100	9	4	99.9	100
Pa. ruwenzoriense (T)	MUCL 51392	6	0	100	100	6	1	100	100
Pyrenopolyporus hunteri (ET)	MUCL 49339	5	28	96.7	99.8	6	106	99.6	100
Xylaria hypoxylon (ET)	CBS 122620	13	28	94.3	100	12	127	96.2	99.9

sequencing (*H. rubiginosum* MUCL 52887, GenBank acc. number KC477232 in Stadler et al. 2013) was found to be 100% identical to the ITS copies obtained from *H. rubiginosum* DSM106870 (Supplementary Information Fig. 7B). The LSU (*H. rubiginosum* MUCL 52887, GenBank acc. number KY610469 in Wendt et al. 2018) sequence contains a deletion, but otherwise shares 100% sequence identity with four out of five retrieved LSU sequences obtained from the *H. rubiginosum* DSM 106870 genome (Supplementary Information Fig. 7C).

Furthermore, we analyzed the four ITS and LSU copies of the *H. rickii* genome sequence in more detail, which displayed identity from 98.9 to 99.8%. Six additional ITS sequences were downloaded from GenBank, including two sequences acquired from the ex-type strain MUCL 53309 (acc. number KC968932 in Kuhnert et al. 2014, KY610416 in Wendt et al. 2018) from Martinique; the strain YMJ 25 (acc. number JQ009313 in Hsieh et al. 2005) from Mexico; strain H19R (acc. number AJ390408 in Sánchez-Ballesteros et al. 2000) from Mexico; the strain "Hyporck" (acc. number MN490062, "Unpublished") from Malaysia; and the strain CFE-152 (acc. number MN653266, "Unpublished") from India (Supplementary Information Fig. 7A). Remarkably, five sequences of these strains exhibited 100% identity with the ITS-2 sequence obtained from the genome of *H. rickii* (Supplementary Information Table 1). However, strain CFE-152 (acc. number MN653266) from India displayed identities ranging from 93.7 to 94.3 % against the sequences of *H. rickii* from the genome. A BLAST comparison against the ITS sequences deposited in GenBank revealed a closest hit with 95.99%, representing a *Hypoxylon* sp. (acc. number KR155059.1 and KR155047.1, "Unpublished") followed by *H. rickii* strain MUCL 53309 with 95.13%, confirming the dubious identification of this strain. For the LSU region, a unique available sequence in GenBank (acc. number KY610416 in Wendt et al. 2018) was retrieved that shared 100% identity with the LSU 2 of the here reported genome sequence of *H. rickii*.

Phylogenetic tree reconstruction

The analysis featured representatives for each molecularly well-established genus within the *Hypoxylaceae*, specifically *Annulohypoxylon* (2 strains), *Daldinia* (2 strains), *Hypomontagnella* (3 strains), *Hypoxylon* (29 strains), *Jackrogersella* (3 strains), *Pyrenopolyporus* (1 strain), and *Parahypoxylon* (2 strains). Additionally, one representative from both *Xylariaceae* and *Lopadostomataceae* (*Xylaria hypoxylon* and *Creosphaeria sassafras*) were included as outgroups.

In our phylogenetic analyses, 44 sequences (or 45 when considering *TEF1*) were employed. The final MAFFT alignment consisted of 4884 nucleotides for *RPB2*, 2351

nucleotides for *TUB2*, 2534 nucleotides for *TEF1*, and 1946 for *ACT1*. The alignment of each locus is available in the Supplementary Information Tables 4-7.

Out of the individual protein coding phylogenetic trees, the phylogram of *RPB2* showed the highest support in the core clades of *Hypoxylaceae* consisting of the genera *Annulohypoxylon*, *Daldinia*, *Hypomontagnella*, *Hypoxylon*, *Jackrogersella*, *Pyrenopolyporus*, and *Parahypoxylon* BS and PP support (1/100), as well as for the outgroup (Fig. 1). Similar results were obtained for the combined rooted phylogenetic tree (Fig. 2).

In contrast, the inferred trees obtained from *ACT1*, *TUB2*, and *TEF1* loci showed a low to moderate support for principal clades (or groups) when compared to previously inferred phylogenies described for the family *Hypoxylaceae* by Wendt et al. (2018), Lambert et al. (2019), Becker et al. (2020), and Cedeno-Sanchez et al. (2023). The phylogenetic tree of each locus is available in the Supplementary Information Figs. 2-4.

Following the idea presented by Stadler et al. (2020), we inferred a phylogenetic tree for all extracted ITS rDNA sequences. The final data matrix comprised 145 sequences. Sequence multiples were reduced to one unique representative per ITS copy to improve readability. The final MAFFT alignment consisted of 1482 nucleotides. The resulting phylogenetic tree displayed low to moderate support for the primary clades within *Hypoxylaceae*. Sequences extracted from the same strain were consistently resolving in the same clade with high support (Supplementary Information Fig. 1).

Intragenomic polymorphisms—deep paralogues

The intragenomic polymorphisms recovered from two strains (*Pa. papillatum* ATCC 58729, *Hyp. monticulosa* MUCL 54604) were remarkably large; hence, we investigated whether they could be regarded as pseudogenes. In both strains, we detected a prominent haplotype (nine and three copies in the first and second strain, respectively) that fully matched the reported sequences for the particular strain obtained through Sanger sequencing. Additionally, we found minor haplotypes (always represented by a single copy) that exhibited a divergence of 89–92% compared to the major haplotype and displayed significantly lower GC content (Supplementary Information Table 2). Notably, the minor haplotypes of both strains exhibited highly divergent sequences in the conservative 5.8S rDNA when compared to all other sequences in the dataset. The alignment of 5.8S rDNA of all strains involved 165 positions and 26 unique haplotypes. Excluding the minor haplotypes of *Pa. papillatum* and *Hyp. monticulosa*, there were 20 variable positions each. When including *Pa. papillatum* minor haplotypes, an additional 20 variable positions were found. Similarly, incorporating the minor copy of *Hyp. monticulosa* resulted in an additional 31 positions, making a total of 51 variable positions. This deep divergence found in
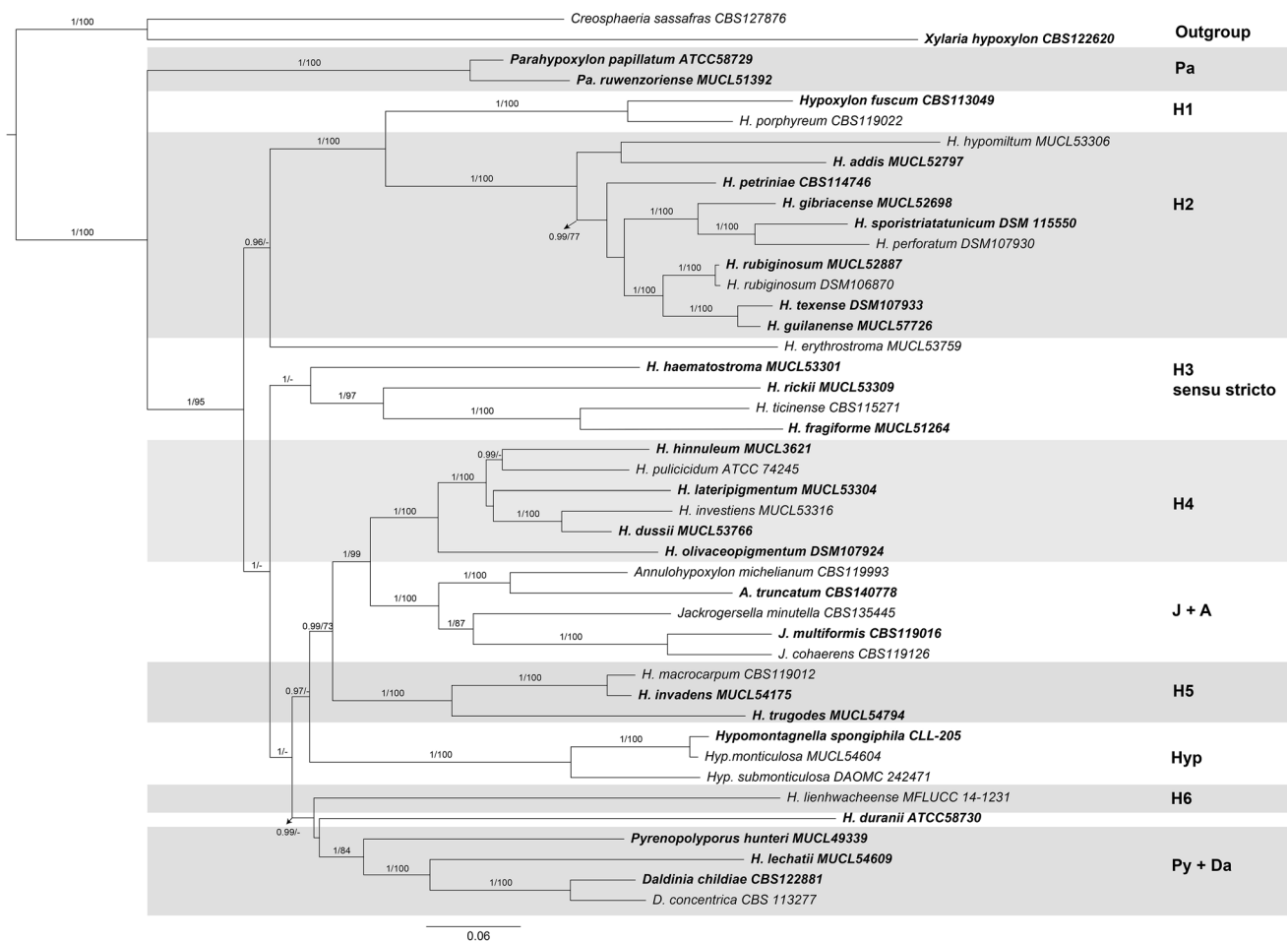
Fig. 1 Molecular phylogenetic maximum likelihood (lLn = −67265.272) tree inferred from the whole *RPB2* sequence using IQ-TREE. Support valu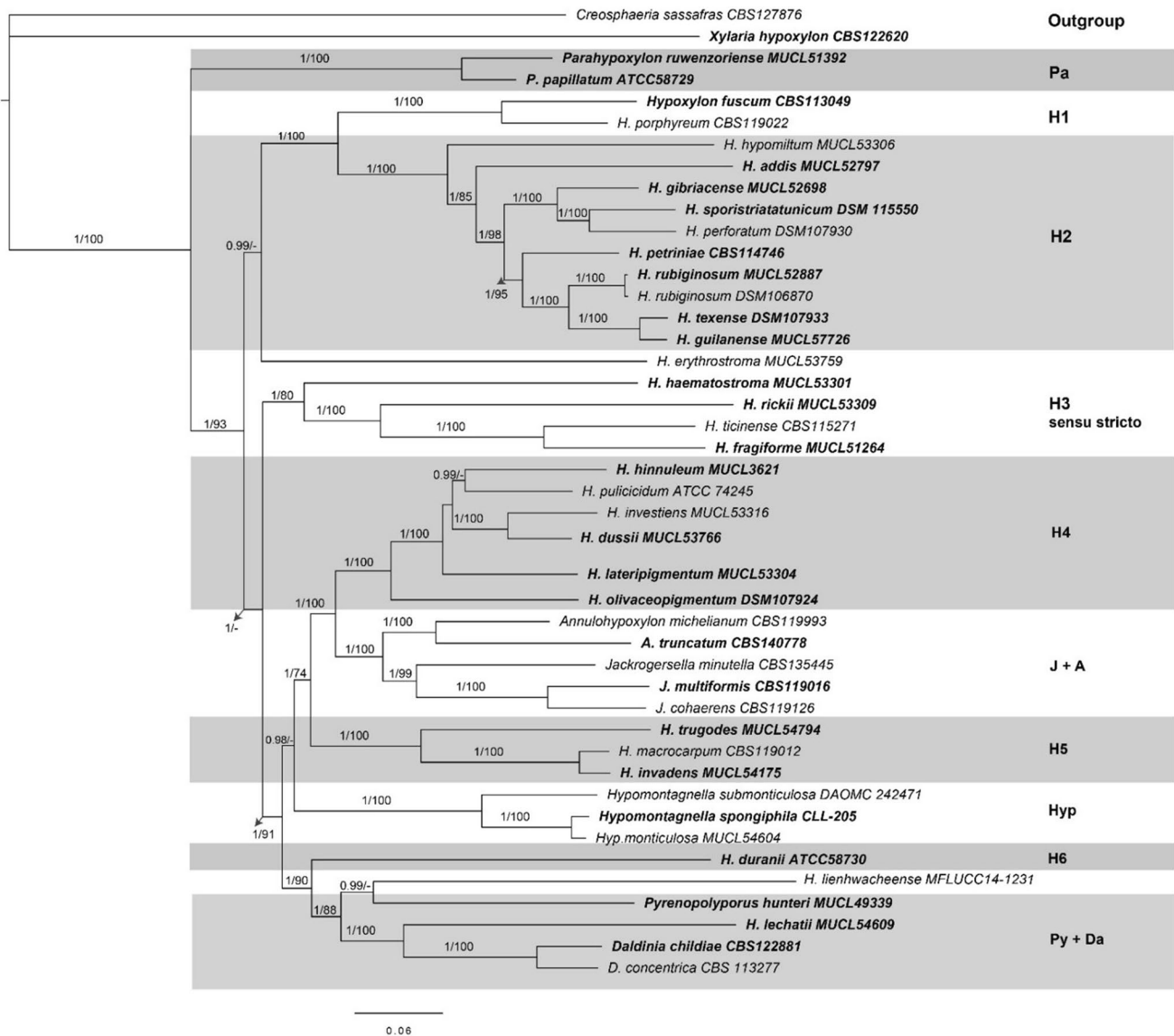es were calculated following Bayesian inference methodology and from 1000 bootstrap replicates. Bayesian posterior probability scores ≥ 0.95/bootstrap support values ≥ 70 are indicated along branches. Type material is highlighted in bold letters

ITS is also reflected in the adjacent LSU sequences. Specifically, in *Pa. papillatum*, the minor ITS haplotypes (2nd, 3rd) continued with sequences that differed from the major copy by 11.5% (2nd, first 1000 bp used) and 12.6% (3rd, first 1000 bp used), and they differed from each other by 8.7%. For *Hyp. monticulosa*, the LSU haplotypes presented in the genome differed in 1% (1000 bp used).

To investigate further deviations in rDNA, we studied hybridization of the proximal stem containing 5.8S and 28S (LSU) motifs. In the genome of *Pa. papillatum* ATCC58729, the dominant (i.e., 1st) haplotype exhibited the typical hybridization of the proximal stem with a Gibbs free energy (ΔG) of −17 and enthalpy (ΔH) of −140.2 for the ensemble. Its structure is formed by an imperfect stem that harbors one free nucleotide on the 5.8S strand and one free nucleotide on the LSU strand, forming typical bulge loops. However, in other two minor haplotypes, variations in the ΔG and ΔH were detected when compared with the major sequence, and assembly of the typical proximal stem was highly modified (Fig. 3).

As an additional indicator of potential rDNA pseudogenes, we studied the secondary structure of ITS2 sequences. In almost all haplotypes, the standard pattern with four helices (I–IV) was observed, with helix III being the longest (Fig. 3). However, for the minor 3rd *Pa. papillatum* ATCC 58729 haplotype, the prediction failed due to the absence of similar models in the database (see http://its2.bioapps.biozentrum.uni-wuerzburg.de/). On the other hand, the direct modeling of another minor 2nd *Pa. papillatum* haplotype resulted in a structure with a much higher free energy (−28.49 kcal/mol) compared to the major haplotype (−44.78). In *Hyp. monticulosa*, both haplotype models exhibited specific free energy values (Supplementary Information Table 2).

Springer

Fig. 2 Molecular phylogenetic maximum likelihood (lLn = −148775.562) tree inferred from a multigene alignment featuring proteinogenic nucleotide sequences derived from whole *ACT1*, *TUB2*, *TEF1*, and *RPB2* genes. Support values were calculated following Bayesian inference methodology and from 1000 bootstrap replicates. Bayesian posterior probability scores ≥ 0.95/bootstrap support values ≥ 70 are indicated along branches. Type material is highlighted in bold letters

Discussion

Recent genome sequencing projects showed that relying solely on rDNA marker for species identification and phylogenetic analysis can drive problematic species delimitations and incorrect identifications, especially in species complex, due to the presence of intragenomic polymorphisms (Stadler et al. 2020; Paloi et al. 2022; Bradshaw et al. 2023). To further explore this notion, we strived to obtain high-quality genomes for rDNA analysis, combining 2nd generation sequencing technologies such as Illumina with 3rd generation sequencing technologies like Oxford Nanopore, along with extensive usage of bioinformatic tools, which significantly increased the accuracy of the genome assembly (Stadler et al. 2020; Paloi et al. 2022; Hoang et al. 2022).

The results obtained here are well-aligned with the previous results by Stadler et al. (2020), which revealed high intragenomic polymorphisms in a pilot study featuring a smaller subset of strains accommodated within the *Hypoxylaceae*. Additional sequence analysis revealed the presence of deep rDNA paralogs. Intragenomic variation in the rDNA cistron can likely be traced back to be caused by nucleotide deletions, insertions, and substitutions within the genome (Bradshaw et al. 2023; Paloi et al. 2022).
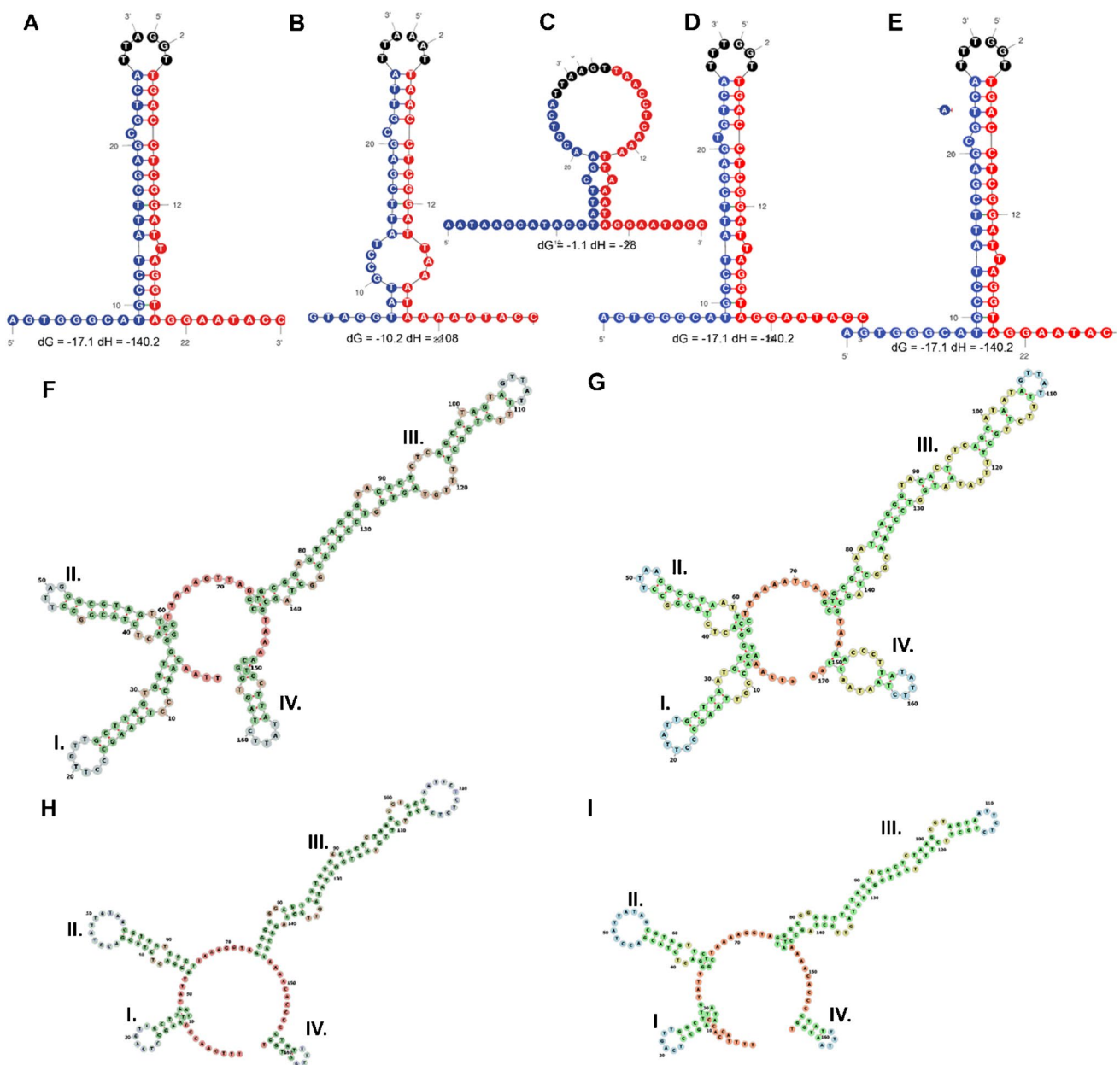
Fig. 3 A–E Hybridization model of the proximal stem region and **F–I** ITS2 secondary structure prediction in *Hyp. monticulosa* and *Pa. papillatum*. **A** Major haplotype (i.e., 1st haplotype) of *H. papillatum*, **B** minor haplotype (i.e., 2nd haplotype) of *Pa. papillatum*, **C** minor haplotype (i.e., 3rd haplotype) of *Pa. papillatum*, **D** major haplotype of *Hyp. monticulosa*, **E** minor haplotype of *Hyp. monticulosa*, **F** major haplotype (i.e., 1st haplotype) of *Pa. papillatum*, **G** minor haplotype (i.e., 2nd haplotype) of *Pa. papillatum*, **H** major haplotype of *Hyp. monticulosa*, and **I** minor haplotype of *Hyp. monticulosa*. The helices are indicated by Roman numerals (I–IV)

Although the number of LSU copies did not differ significantly in most of the studied strains, probably due to the rDNA cistron being arranged in tandem throughout the genome (Torres-Machorro et al. 2010), considerable variations were observed in genomes derived from, e.g., *J. cohaerens*, *Pa. papillatum*, *H. addis*, *H. dussii*, *H. guialense*, *H. sporistriatatunicum*, *Hyp. monticulosa*, *Py. hunteri*, and *X. hypoxylon*, where more copies of ITS were found compared to LSU, and in some cases, only a partial LSU

was recovered. Additionally, we were not able to retrieve ITS sequences from *H. rubiginosum* MUCL 52887 and found only one copy in *Hyp. submonticulosa*. The quality of third-generation sequencing technologies and their impending implications for studying phenomena such as polymorphisms of the rDNA cistron, including their multiplicity (Bradshaw et al. 2023), are widely discussed for different groups of fungi (Paloi et al. 2022; Stadler et al. 2020). The apparent absence, or at least failure to detect ITS

sequences inside a genome sequence for technical reasons, is not an isolated case, as was recently discussed by Bradshaw et al. (2023). Here, the authors were unable to locate the ITS sequences for a quarter of all taxa evaluated and only a single ITS copy for half of the taxa studied.

The study by Stadler et al. (2020) reported polymorphisms in the ITS region for the species *H. lienhwacheense*, *H. rickii*, *Hyp. monticulosa, and Py. hunteri*. In our analysis, apart from *Py. hunteri*, all these species also exhibited such variations in the LSU region. On the other hand, the genomes of *A. truncatum*, *D. concentrica*, *H. pulicicidum*, and *J. multiformis* were found to possess polymorphisms in the LSU region but not in ITS (Table 2). Despite this, for the species studied here, the ITS rDNA showed to have a lower rate of polymorphisms than the LSU region. We expected the opposite, taking into consideration that the ITS region (~ 400–900 bp) contains two introns (ITS1 and ITS2) that are highly variable and a well-conserved small non-coding RNA (5.8S). Usually, the LSU, with ~ 3000–5000 bp, is highly conserved within species because of the crucial function it plays (ribosome function and protein synthesis; cf. Gregory et al. 2019), and many variations in its sequence could cause disruption in these processes.

After in-depth analyses of sequences retrieved from the genome of *H. rickii*, we concluded that only one of the four sections (where the ITS and LSU can be located in the genome) were amplified using Sanger sequencing methods (Supplementary Information Fig. 6) in the past. This is interesting because the target sites of commonly used primers (ITS1, ITS4, ITS5) located in the genome did not diverge, and thus should be amplified by PCR stochastically. We encountered a similar phenomenon when studying the *Py. hunteri* genome sequence, where only one of the five sections can be found in the literature (Supplementary Information Fig. 5). For both regions (ITS and LSU), however, one of the sequences retrieved from the genome presents mismatches with the primers ITS1 and ITS5, hence offering an explanation on why this region is not targeted for amplification. A second explanation would be a highly condensed rDNA cistron, with the consequence that these sections are not actively transcribed when the cell needs to produce ribosomes for protein synthesis.

Phylogenetic analysis of the ITS showed low to medium support for the main clades of the *Hypoxylaceae* reported in previous studies (Wendt et al. 2018; Lambert et al. 2019; Becker et al. 2020; Cedeño-Sanchez et al. 2023). This result is congruent with older phylogenetic studies solely relying on ITS rDNA data alone but does not reflect taxonomic advances of the last decade (compare with Sánchez-Ballesteros et al. 2000; Hsieh et al. 2005). Strikingly, sequences from the same strain consistently formed well-supported clades, confirming the reliability of our results. This result suggests that ITS can rather be used to estimate taxonomic affinities towards the different species complexes described for the *Hypoxylaceae* in a "quick-and-dirty"-fashion. Due to LSU showing a similar pattern of intragenomic polymorphism, we would consider it risky to apply it as complementary barcoding marker to ITS for identification to species level. On the contrary, all the protein-coding regions studied here clearly displayed better phylogenetic resolution and hence can be regarded as much more suited for barcoding for fungi due to being highly conserved and not displaying intragenomic polymorphisms, at least in the here studied strains.

Of note, we pioneered the retrieval of *TEF1* data from all here investigated genomes, marking its inaugural inclusion in a *Xylariales* phylogeny. This locus has been well-established for other taxonomic groups, such as *Amphisphaeriaceae*, *Cainiaceae*, *Cladosporiaceae*, *Clypeosphaeriaceae*, *Diatrypaceae*, and *Hyponectriaceae* (Jaklitsch and Voglmayr 2012; Dai et al. 2014; Vicente et al. 2021, Samarakoon et al. 2022). However, applying this locus alone did not show any advantage in direct comparison to the others, at least for *Hypoxylaceae*. Other retrieved genes (*TUB2* and *RPB2*) have already been successfully applied to infer well-resolved phylogenies (Hsieh et al. 2005; Wendt et al. 2018; Lambert et al. 2019; Becker et al. 2020; Cedeño-Sanchez et al. 2023). Comparing the results obtained in the study of Hsieh et al. (2005) with a partial *ACT1* and in this study using a whole *ACT1* gene, it is clear that the *ACT1* gene alone is insufficient for solving the relationships among species in the *Hypoxylaceae*.

Our investigation revealed high intragenomic polymorphism of rDNA in distinct strains within the *Hypoxylaceae*, indicating the presence of paralogs. Specifically, we found multiplets of rDNA sequences in *Hyp. monticulosa* and *Pa. papillatum* genomes, which we propose to call deep rDNA paralogs. They uniformly exhibited a significantly lower GC content compared to the major copy and showed highly variable 5.8S rDNA sequences, which were otherwise found to be highly conserved across species. Some of the paralogs retained the necessary motifs for maintaining the required secondary structure of ITS2, while two paralogs exhibited thermodynamically less stable structure (i.e., it has a high free energy) which could not be modeled anymore for technical reasons (Supplementary Information Table 2). Specifically, in two minor haplotypes of *Pa. papillatum*, conservative motifs in 5.8S and LSU sequences were disrupted, resulting in a corresponding proximal stem structure with a highly atypical form. Low GC content and mutations in otherwise well-conserved sequence segments are typical signs for a pseudogene (see Kolařík and Vohník 2018 and Stadler et al. 2020 for review). In eukaryotes, the hybridized 5.8S and LSU rRNA parts, forming so-called proximal stems, have a free nucleotide on each side with approximately six base pairs in between. The structural pattern of this proximal stem is necessary for successful detection of its associated

processing machinery (see Keller et al. 2009) and has been proposed as a diagnostic character for pseudogene detection (Harpke and Peterson 2007, 2008). In general, the 3rd ITS haplotype of *Hyp. monticulosa* deviates the most, as it exhibits the lowest GC content, an unpredictable ITS2 secondary sequence, and a highly disrupted proximal stem structure. Furthermore, the LSU sequence adjacent to this haplotype shows the highest observed divergence (11.5%) from the major haplotype. To conclude, it is highly likely that the captured deep paralogs represent pseudogenes (at least in the case of *Hyp. monticulosa*) or sequences in an advanced stage of pseudogene formation. Variations in rDNA identity spotted for the other explored genomes can be explained with nucleotide deletions, insertions, and substitutions in the genome.

Lastly, we want to stress that the major conclusion of this study is that it reinforces our opinion that the rDNA cistron alone is insufficient as a universal barcode marker for fungi (Paloi et al. 2022; Bradshaw et al. 2023), especially in the *Hypoxylaceae* (Stadler et al. 2020). We propose to evaluate the *TUB2* gene as a new primary barcoding marker for *Hypoxylaceae* as a substitute, as ample reference sequences have already been obtained for *Hypoxylon* and many other genera of *Xylariales* in the past. A phylogeny based on *RPB2* sequences resolved the hitherto accepted topology for *Hypoxylaceae* and helped to further stabilize the phylogeny, but the number of available sequences derived from type and reliably identified vouchers is far lower than in case of *TUB2*. The suitability of *TUB2* for many taxa needs to be examined by the inclusion of additional vouchers belonging to the same species (complex) to assess interspecific variability. The reason is that there are no molecular data at all or only a single sequence dataset available for more than 50% of the described *Hypoxylaceae* taxa. Nevertheless, it has already been shown that the few *Hypoxylon* taxa for which multiple specimens were sequenced that the *TUB2* locus shows little variability in, e.g., *H. fragiforme*, *H. rubiginosum* (according to data from GenBank arising from multiple independent studies), and *H. fuscum* (cf. Lambert et al. 2021). Our study contributes to a better understanding of the genetic diversity and evolutionary dynamics within *Xylariales* and emphasizes the need for holistic approaches, such as multi-gene sequencing in fungal barcoding endeavors and phylogenetic analyses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11557-023-01940-2.

Data availability All sequence data generated in this study are available in GenBank (https://www.ncbi.nlm.nih.gov/genbank/).

Declarations

Conflict of interest The authors declare no competing interests.

References

Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol 60(5):685–699. https://doi.org/10.1093/sysbio/syr041

Ankenbrand MJ, Keller A, Wolf M, Schultz J, Förster F (2015) ITS2 database V: twice as much. Mol Biol Evol 32:3030–3032. https://doi.org/10.1093/molbev/msv174

Becker K, Lambert C, Wieschhaus J, Stadler M (2020) Phylogenetic assignment of the fungicolous *Hypoxylon invadens* (*Ascomycota*, *Xylariales*) and investigation of its secondary metabolites. Microorganisms 8:1397. https://doi.org/10.3390/microorganisms8091397

Bradshaw MJ, Aime MC, Rokas A, Maust A, Moparthi S, Jellings K, Pane AM, Hendricks D, Pandey B, Li Y, Pfister D (2023) Extensive intragenomic variation in the internal transcribed spacer region of fungi. iScience 26:107317. https://doi.org/10.1016/j.isci.2023.107317

Camacho C, Coulouris G, Avagyan Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinform. 10:1–9. https://doi.org/10.1186/1471-2105-10-421

Ceballos-Escalera A, Richards J, Arias MB, Inward DJG, Vogler AP (2022) Metabarcoding of insect-associated fungal communities: a comparison of internal transcribed spacer (ITS) and large-subunit (LSU) rRNA markers. MycoKeys 88:1–33. https://doi.org/10.3897/mycokeys.88.77106

Cedeño-Sanchez M, Charria-Girón E, Lambert C, Luangsa-ard JJ, Decock C, Franke R, Brönstrup M, Stadler M (2023) Segregation of the genus *Parahypoxylon* (*Hypoxylaceae*) from *Hypoxylon* by a polyphasic taxonomic approach. MycoKeys 95:131–162

Chernomor O, Von Haeseler A, Minh BQ (2016) Terrace aware data structure for phylogenomic inference from supermatrices. Syst Biol 65(6):997–1008. https://doi.org/10.1093/sysbio/syw037

Dai DQ, Wijayawardene NN, Bhat DJ, Chukeatirote E, Bahkali AH, Zhao RL, Xu JC, Hyde KD (2014) *Pustulomyces* gen. nov. accommodated in *Diaporthaceae*, *Diaporthales*, as revealed by morphology and molecular analyses. Cryptog Mycol 35(1):63–72

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.2307/2408678

Gregory B, Rahman N, Bommakanti A et al (2019) The small and large ribosomal subunits depend on each other for stability and accumulation. Life Sci Alliance 2(2):e201800150. https://doi.org/10.26508/lsa.201800150

Halecker S, Wennrich JP, Rodrigo S, Andrée N, Rabsch L, Baschien C, Steinert M, Stadler M, Surop F, Schulz B (2020) Fungal endophytes for biocontrol of ash dieback: the antagonistic potential of *Hypoxylon rubiginosum*. Fungal Ecol 45:100918. https://doi.org/10.1016/j.funeco.2020.100918

Harpke D, Peterson A (2007) 5.8S motifs for the identification of pseudogenic ITS regions. Botany 86:300–305. https://doi.org/10.1139/B07-134

Harpke D, Peterson A (2008) Extensive 5.8S nrDNA polymorphism in *Mammillaria (Cactaceae)* with special reference to the identification of pseudogenic internal transcribed spacer regions. J Plant Res 121:261–270. https://doi.org/10.1007/s10265-008-0156-x

Heeger F, Bourne EC, Baschien C, Yurkov A, Bunk B, Spröer C, Overmann J, Mazzoni C, Monaghan MT (2018) Long-read DNA metabarcoding of ribosomal RNA in the analysis of fungi from aquatic environments. Mol Ecol Res 18:1500–1514. https://doi.org/10.1111/1755-0998.12937

Hoang MTV, Irinyi L, Meyer W (2022) Long-read sequencing in fungal identification. Microbiol Aust 43:14–18

Hsieh HM, Ju YM, Rogers JD (2005) Molecular phylogeny of *Hypoxylon* and closely related genera. Mycologia 97(4):844–865. https://doi.org/10.1080/15572536.2006.11832776

Jaklitsch WM (2012) Voglmayr H (2012) Phylogenetic relationships of five genera of *Xylariales* and *Rosasphaeria* gen. nov. (*Hypocreales*). Fungal Divers 52:75–98. https://doi.org/10.1007/s13225-011-0104-2

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166. https://doi.org/10.1093/bib/bbx108

Keller A, Schleicher T, Schultz J, Müller T, Dandekar T, Wolf M (2009) 5.8S-28S rRNA interaction and HMMbased ITS2 annotation. Gene 430:50–57. https://doi.org/10.1016/j.gene.2008.10.012

Kerpedjiev P, Hammer S, Hofacker I (2015) Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. Brief Bioinform 31:3377–3379. https://doi.org/10.1093/bioinformatics/btv372

Klaubauf S, Inselsbacher E, Zechmeister-Boltenstern S, Wanek W, Gottsberger R, Strauss J, Gorfer M (2010) Molecular diversity of fungal communities in agricultural soils from Lower Austria. Fungal Divers 44:65–75. https://doi.org/10.1007/s13225-010-0053-1

Kolařík M, Vohník M (2018) When the ribosomal DNA does not tell the truth: the case of the taxonomic position of *Kurtia argillacea*, an ericoid mycorrhizal fungus residing among *Hymenochaetales*. Fungal Biol 122:1–18. https://doi.org/10.1016/j.funbio.2017.09.006

Kolařík M, Wei IC, Hsieh SY, Piepenbring M, Kirschner R (2021) Nucleotide composition bias of rDNA sequences as a source of phylogenetic artifacts in *Basidiomycota* a case of a new lineage of a uredinicolous *Ramularia*-like anamorph with affinities to *Ustilaginomycotina*. Mycol Prog 20:1553–1571. https://doi.org/10.1007/s11557-021-01749-x

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–236. https://doi.org/10.1101/gr.215087.116

Kuhnert E, Fournier J, Peršoh D, Luangsa-ard JJD, Stadler M (2014) New *Hypoxylon* species from Martinique and new evidence on the molecular phylogeny of *Hypoxylon* based on ITS rDNA and β-tubulin data. Fungal Divers 64:181–203. https://doi.org/10.1007/s13225-013-0264-3

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Lambert C, Wendt L, Hladki AI, Stadler M, Sir EB (2019) *Hypomontagnella* (*Hypoxylaceae*): a new genus segregated from *Hypoxylon* by a polyphasic taxonomic approach. Mycol Prog 18:187–201. https://doi.org/10.1007/s11557-018-1452-z

Lambert C, Pourmoghaddam MJ, Cedeño-Sanchez M, Surup F, Khodaparast SA, Krisai-Greilhuber I, Voglmayr H, Stradal TEB, Stadler M (2021) Resolution of the *Hypoxylon fuscum* complex (*Hypoxylaceae, Xylariales*) and discovery and biological characterization of two of its prominent secondary metabolites. J Fungi 7:131. https://doi.org/10.3390/jof7020131

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

Li H (2013) Aligning sequence reads, clone sequences and assemblycontigs with BWA-MEM. arXiv Preprint. arXiv:1303.3997 [q-bio.GN]

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Haeseler AV, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. https://doi.org/10.1093/molbev/msaa015

Palmer JM, Stajich J (2020) Funannotate v1.8.1: eukaryotic genome annotation. 10.5281/zenodo.4054262

Paloi S, Luangsa-ard JJ, Mhuantong W, Stadler M, Kobmoo N (2022) Intragenomic variation in nuclear ribosomal markers and its implication in species delimitation, identification and barcoding in fungi. Fungal Biol Rev 42:1–33. https://doi.org/10.1016/j.fbr.2022.04.002

Prahl RE, Khan S, Deo RC (2021) The role of internal transcribed spacer 2 secondary structures in classifying mycoparasitic *Ampelomyces*. PloS One 16:e0253772. https://doi.org/10.1371/journal.pone.0253772

Samarakoon MC, Hyde KD, Maharachchikumbura SSN, Stadler M, Jones EBG, Promputtha I, Suwannarach N, Camporesi E, Bulgakov TS, Liu JK (2022) Taxonomy, phylogeny, molecular dating and ancestral state reconstruction of *Xylariomycetidae* (*Sordariomycetes*). Fungal Divers 112:1–88. https://doi.org/10.1007/s13225-021-00495-5

Sánchez-Ballesteros J, González V, Salazar O, Acero J, Portal MA, Julián M, Rubio V, Bills GF, Polishook JD, Platas G, Mochales S, Peláez F (2000) Phylogenetic study of *Hypoxylon* and related genera based on ribosomal ITS sequences. Mycologia 92:964–977. https://doi.org/10.1080/00275514.2000.12061240

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W et al (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. PNAS 109:6241–6246. https://doi.org/10.1073/pnas.1117018109

Stadler M, Kuhnert E, Peršoh D, Fournier J (2013) The *Xylariaceae* as model example for a unified nomenclature following the "One Fungus-One Name" (1F1N) concept. Mycol Int J Fungal Biol 4:5–21. https://doi.org/10.1080/21501203.2013.782478

Stadler M, Lambert C, Wibberg D, Kalinowski J, Cox RJ, Kolařík M, Kuhnert E (2020) Intragenomic polymorphisms in the ITS region of high-quality genomes of the *Hypoxylaceae (Xylariales, Ascomycota)*. Mycol Prog 19:235–245. https://doi.org/10.1007/s11557-019-01552-9

Torres-Machorro AL, Hernández R, Cevallos AM, López-Villaseñor I (2010) Ribosomal RNA genes in eukaryotic microorganisms: witnesses of phylogeny? FEMS Microbiol Rev 34:59–86. https://doi.org/10.1111/j.1574-6976.2009.00196.x

Vicente TF, Gonçalves MF, Brandão C, Fidalgo C, Alves A (2021) Diversity of fungi associated with macroalgae from an estuarine environment and description of *Cladosporium rubrum* sp. nov. and *Hypoxylon aveirense* sp. nov. Int J Syst Evol Microbiol 71:004630. https://doi.org/10.1099/ijsem.0.004630

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963

Wendt L, Sir EB, Kuhnert E, Heitkämper S, Lambert C, Hladki AI, Romero AI, Luangsa-ard JJ, Srikitikulchai P, Peršoh D, Stadler M (2018) Resurrection and emendation of the *Hypoxylaceae*, recognised from a multigene phylogeny of the *Xylariales*. Mycol Prog 17:115–154. https://doi.org/10.1007/s11557-017-1311-3

Wibberg D, Stadler M, Lambert C, Bunk B, Spröer C, Rückert C, Kalinowski J, Cox RJ, Kuhnert E (2021) High quality genome sequences of thirteen *Hypoxylaceae (Ascomycota)* strengthen the phylogenetic family backbone and enable the discovery of new taxa. Fungal Divers 106:7–28

Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. J Hered 92:371–373. https://doi.org/10.1093/jhered/92.4.371

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.