Shallow Genome Sequencing for Phylogenomics of Mycorrhizal Fungi from Endangered Orchids

Sarah A. Unruh[1], J. Chris Pires[1], Lawrence Zettler[2], Luigi Erba[2], Igor Grigoriev[3], Kerrie Barry[3],

Christopher Daum[3], Anna Lipzen[3], and Jason E. Stajich[4]

1. Division of Biological Sciences, University of Missouri, Columbia, MO, United States

Sarah Unruh Roles: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review and Editing

J. Chris Pires Roles: Conceptualization, Funding Acquisition, Project Administration, Resources, Supervision, Writing – Review and Editing

2. Department of Biology, Illinois College, Jacksonville, IL, United States

Lawrence W. Zettler Roles: Funding acquisition, Resources

Luigi Erba Roles: Investigation

3. United States of America Department of Energy Joint Genome Institute, Walnut Creek, CA, United States

Igor Grigoriev Roles: Project Administration, Resources

Kerrie Barry Roles: Project Administration

Christopher Daum: Resources

Anna Lipzen: Investigation

4. Microbiology and Plant Pathology, University of California-Riverside, Riverside, CA, United States

Jason E. Stajich Roles: Data Curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – Review and Editing

ABSTRACT

41

42

43 Most plant species form symbioses with mycorrhizal fungi and this relationship is especially

44 important for orchids. Fungi in the genera *Tulasnella, Ceratobasidium,* and *Serendipita* are

45 critically important for orchid germination, growth and development. The goals of this study are

46 to understand the phylogenetic relationships of mycorrhizal fungi and to improve the taxonomic

47 resources for these groups. We identified 32 fungal isolates with the internal transcribed spacer

48 region and used shallow genome sequencing to functionally annotate these isolates. We

49 constructed phylogenetic trees from 408 orthologous nuclear genes for 50 taxa representing 14

50 genera, 11 families, and five orders in Agaricomycotina. While confirming relationships among

51 the orders Cantharellales, Sebacinales, and Auriculariales, our results suggest novel relationships

52 between families in the Cantharellales. Consistent with previous studies, we found the genera

53 *Ceratobasidium* and *Thanatephorus* of Cerabotasidiaceae to not be monophyletic. Within the

54 monophyletic genus *Tulasnella*, we found strong phylogenetic signals that suggest a potentially

55 new species and a revision of current species boundaries (e.g. *Tulasnella calospora*); however it

56 is premature to make taxonomic revisions without further sampling and morphological

57 descriptions. There is low resolution of *Serendipita* isolates collected . More sampling is needed

58 from areas around the world before making evolutionary-informed changes in taxonomy. Our

59 study adds value to an important living collection of fungi isolated from endangered orchid

60 species, but also informs future investigations of the evolution of orchid mycorrhizal fungi.

61

INTRODUCTION

Fungi are more than mere decomposers, they form symbioses with every other group of organisms on Earth. Fungal interactions span the entire symbiotic spectrum, from parasitism to mutualism. Most intertwined with plants, may have even enabled development/existence of land plants (Lutzoni et al., 2018). As a result of this long-term association, fungi are essential symbionts to almost every plant species on Earth. The fungi live in plant roots are called mycorrhizal fungi and associate with more than 85% of plant species (Smith and Read, 2008). Mycorrhizal fungi are critical for plant health and function by helping obtain and retain water, mediating defense responses, participating in signaling between roots, and facilitating the exchange of nutrients like carbon, phosphorus, and nitrogen (Barto et al., 2012; Jung et al., 2012; Peterson and Massicotte, 2004; Wang et al., 2017; Yoder et al., 2010). The plant group that relies the most on their mycorrhizal fungi are orchids.

Orchids rely on their mycorrhizal symbionts to stimulate plant development during seed germination by providing carbon resources (Kuga et al., 2014). Orchid mycorrhizal fungi (ORM) form hyphal coils termed pelotons inside the cells of orchid embryos and in the adult roots, tubers, or rhizomes (McCormick et al., 2016; Rasmussen et al., 2015). These pelotons are the sites of nutrient exchange and the molecular nature of this marketplace remains poorly understood though exciting new research shedding light (Fochi et al., 2017a; Fochi et al., 2017b; Kuga et al., 2014). Most orchids associate with mycobionts belonging to the basidiomycete groups Sebacinales, Ceratobasidiaceae and Tulasnellaceae. In addition to orchid mycorrhizal fungi, these groups contain saprotrophs, plant pathogens, and ectomycorrhizal representing a wide array of metabolic capabilities (Kohler et al., 2015; Nagy et al., 2016). Furthermore, molecular studies have revealed simultaneous root colonization by multiple fungal partners in

86 both photosynthetic terrestrial and epiphytic orchids (Martos et al., 2012). Concluding sentence

87 that makes the argument that there are many dynamics we need to better understand so we need

88 to characterize the diversity of these fungi to untangle their interactions and mechanisms.

89   Although fungi play critical roles, they are rarely visible on the landscape. The number of

90 extant fungal species on Earth ranges from 2-5 million (Blackwell, 2011; Hawksworth and

91 Lücking, 2017) up to 166 million species (Larsen et al., 2017). Most species are microscopic and

92 over the last few decades species identification has relied on molecular methods. Historically,

93 these methods often have used a single molecular marker such at ITS (Nilsson et al., 2014).

94 However modern genome sequencing methods are important tools to discover and describe

95 taxonomic, phylogenetic and functional diversity. The use of different, new analytical tools has

96 also greatly benefited our knowledge of the below-ground ecology of orchids and orchid

97 mycorrhizal fungi. On the right track with multiple markers and Bayesian species delimitations

98 (Ruibal et al., 2014; Ruibal et al., 2013; Whitehead et al., 2017). New species of Tulasnella and

99 relatives are constantly being identified (Linde et al., 2017). Continue to combine sequencing

100 with taxonomic knowledge to provide a comprehensive description of the species that associate

101 with orchids.

102   The genera of orchid fungi we have sampled belong to two orders, Cantharellales and

103 Sebacinales, in the Agaricomycetes. Cantharellales is sister to the rest of class Agaricomycetes

104 and comprises seven families total (Ceratobasidiaceae, Tulasnellaceae, Botryobasidiaceae,

105 Cantharellaceae, Clavulinaceae, Hydnaceae, and Aphelariaceae), though Hibbett et al., (2014),

106 define Cantharellaceae and Clavulinaceae as synonymous with Hydnaceae and the status of

107 Aphelariaceae is unknown (Kirk et al., 2008; Leacock, 2018). Ceratobasidiaceae has two genera

108 (*Ceratobasidium* and *Rhizoctonia*/*Thanatephorus*) that have been demonstrated to be

109　polyphyletic (Veldre et al., 2013). In fact, the type specimen for *Ceratobasidium* has since been

110　reclassified as a member of the order Auriculariales based on the characters like the shape of the

111　basidia and the dolipore (specialized hyphal septa) ultrastructure , leading Oberwinkler et al.,

112　(2013a) to restrict *Ceratobasdium* and Ceratobasidiaceae to the type specimen and reclassifying

113　*Ceratobasidium* spp. as *Rhizoctonia* (Kirk et al., 2008). Tulasnellaceae contains 3 genera and c.

114　50 sp (Kirk et al., 2008). In addition to these described families, the genus *Sisotrema* is known to

115　be polyphyletic with members in Auriculariales as well as Cantharellales. Successively sister to

116　the rest of the Agaricomycetes is the order Sebacinales which includes two families – the

117　Sebacinaceae and Serendipitaceae (Weiss et al., 2016). Though this order comprises a wide

118　swath of diversity, it remains difficult to adequately describe species due to a high volume of

119　environmental sequence data without information about morphological characters (Oberwinkler

120　et al., 2013b; Weiss et al., 2016).

121　　　In this study, our primary goal is to shallowly sequence a rich living collection of fungi

122　isolated from orchid roots and seedlings to provide a phylogenetic framework for future genome-

123　enabled evolutionary and functional studies. Our secondary goal, with the addition of key

124　outgroups, is to answer a series of nested phylogenetic questions about the relationships among

125　the orders, families and genera of Agaricomycetes, with a focus on Ceratobasicaceae,

126　Tulasnellaceae, and Sebacineaceae. We screened taxa using ITS sequencing, and after

127　contaminants were removed we chose 32 taxa for shallow genome sequencing. A total of 50 taxa

128　were analyzed and we extracted 408 orthologous genes. Two highly-supported phylogenetic

129　trees were constructed with RAxML and ASTRAL-III that were overall highly congruent. We

130　discuss how our study provides new insight into the relationships of these orchid mycorrhizal

131　fungi, highlights areas for taxonomic attention and we suggest future research directions.

2. MATERIALS AND METHODS

2.1 Taxonomic sampling

32 environmental samples were isolated from endangered orchids. These samples span three genera in three families in two orders. Outgroup genomes were chosen from the repository in Mycocosm to capture the breadth of taxonomic diversity (Grigoriev et al., 2014). Two super outgroups (*Kockovaella* sp and *Calocera sp)* were chosen from the successively sister classes outside the ingroup class Agaricomycetes [Tremellomycetes, [Dacrymycetes, [Agaricomycetes]]]. In the Cantharellales we sampled the three genomes in Ceratobasidiaceae (*Rhizoctonia solani*, *Thanatephorus cucumeris*, and *Ceratobasidium sp AG1)*, the two genomes in Tulasnellaceae (*Tulasnella calospora* AL13/4D*,* and *Tulasnella calospora* UAMH9824*)*, and one genome each from 4 of the remaining 5 families *Botryobasidium botryosum* (Botryobasidiaceae), *Clavulina sp* (Clavulinaceae), *Cantharellus anzutake* (Cantharellaceae), and *Hydnum rufescens* (Hydnaceae) . We also included three genomes in Serendipitaceae (Sebacinales) *Sebacina vermifera* (syn. *Serendipita vermifera)*, *Piriformospora indica* (syn. *Serendipita indica*), and *Serendipita sp.* 407. We sampled representatives from the order Auriculariales to capture the entire diversity of these sequences (*Oliveonia pauxilla, Auricularia subglabra, Aporpium caryae, and Exidia glandulosa*).

2.2 Fungal Isolates

The 32 fungal samples used in this study were isolated from roots or protocorms (the seedling stage) of endangered orchid species in areas spanning from Hawaii to Florida, with a focus on the Midwest and the Florida Panther National Wildlife Refuge (Table 1). For the full description of the isolation techniques used, see Zettler and Corey (2018). Briefly, root tissue

6

156 was surface-sterilized then placed in a petri dish with sterile water and finely diced with a

157 scalpel. Fungal Isolation Media (Clements et al., 1986) was poured on the diced root tissue and

158 left at ambient temperature. After 24-48 hours, the plates were examined with a dissecting

159 microscope to identify fungal growth. Mycelia were excised and placed on Difco Potato

160 Dextrose Agar (PDA; Becton Dickinson and Co., Sparks, MD, Mfr # BD 213400). Those fungi

161 with morphological characteristics consistent with fungi in the form genus *Rhizoctonia* as

162 identified in Currah et al., (1997) were retained for identification with ITS sequencing (Figure

163 1).

164   Fungi were grown in flasks with 75ml of full strength Difco Potato Dextrose Broth

165 (Difco Becton Dickinson and Co., Sparks, MD, Mfr # BD 254920) on a shaker table until there

166 was enough tissue for extraction. Depending on the isolate this took 2-6 weeks. Often multiple

167 flasks of each isolate were grown at one time to speed up this process. For extraction, the entire

168 contents of each flask was poured into a 150mL Polystyrene Bottle Top Filter 0.45um (Corning

169 Incorporated, Corning, USA, Cat # 430627) and washed with DI water. These samples were

170 weighed to determine how many samples could be processed from each sample (minimum of 0.2

171 grams filtered weight/tube). Fungi were isolated with either the Bacterial/Fungal DNA extraction

172 kit (Zymo Research, Irvine, USA, Cat # D6005, Lot # ZRC201856) according to manufacturer

173 protocol or a CTAB, phenol chloroform isoamyl procedure (Supplemental Figure S1). When the

174 Zymo kit was used, fungi were added to lysis tubes and put on bead beater for two rounds of four

175 minutes. If the CTAB extraction was employed, fungal tissue was ground with liquid Nitrogen in

176 ceramic mortar and pestle. Extracted DNA was assayed on a NanoDrop 2000 (ThermoFisher

177 Scientific, USA, cat # ND-2000) and on a Qubit 2.0 Fluorometer (ThermoFisher Scientific,

178 USA, cat # Q32866) with the Qubit double-stranded DNA High Sensitivity Assay kit

179 (ThermoFisher Scientific, USA, cat # Q32851). We followed JGI instruction for sample

180 submission by submitting approximately 500 ng of each sample in a total volume of 25-35 uL in

181 one 96-well plate provided by JGI.

182 **2.3 ITS sequencing for Species Identification**

183 To determine species identity, we sequenced the internal transcribed spacer (ITS) region

184 of the rDNA. We used the same DNA extraction methods referenced above. We used the primer

185 pairs ITS1/ITS4-Tul or ITS1-OF/ITS4-OF for isolates presumed to be *Tulasnella* as the ITS

186 sequences in this genus are highly divergent and not captured well with other primers (Taylor

187 and McCormick, 2008). For the genera *Ceratobasidium* and *Serendipita,* the general primers

188 ITS1/ITS4 or ITS1-OF/ITS4-OF were used and if these did not successfully amplify the ITS

189 region of *Serendipita* isolates the primer pair ITS3Seb/NL4 (Bellemain et al., 2010; Ray et al.,

190 2015; White et al., 1990). The amplified DNA was cleaned with the DNA Clean and

191 Concentrator-25 kit (Zymo Research, Irvine, USA, cat # D4033). These PCR products were

192 assessed on a 1.5% agarose gel and Sanger sequencing was performed at the University of

193 Missouri DNA Core Facility. These sequences were evaluated for confidence in base calling and

194 edited by trimming low quality bases from the beginning and end of each sequence in Geneious

195 9.1.8 (http://www.geneious.com/). These trimmed sequences were queried against NCBI's

196 default nucleotide-nucleotide database as well as the UNITE database for species identification

197 (Nilsson et al., 2019). These sequences were generated for the purpose of accurate species ID

198 before sending DNA samples for shallow genome sequencing.

199
200 **2.4 Shallow Genome Sequencing and Quality Control**

201 Shallow genome sequencing of 32 samples, quality control, and filtering were performed

202 at the Joint Genome Institute (JGI) under a Community Sequencing Proposal (#2000). Samples

203   were run on an Illumina NovaSeq with 2x151 base pair (bp) reads. The quality control and

204   filtering at the JGI use BBmap to remove contamination and remove low quality reads (Bushnell

205   B., BBMap. http://sourceforge.net/projects/bbmap/). Three samples were sequenced at the

206   University of Missouri's DNA Core Facility which were run on an Illumina NextSeq 500

207   machine on one lane with 45 other samples generating 2x150 bp reads.

208   **2.5 Shallow Genome Assembly and Annotation**

209   All cleaned and filtered sequences from the Joint Genome Institute and the University of

210   Missouri were assembled with the AAFTF pipeline for read assembly, remove vector

211   contamination and duplicate contigs, contig sequence polishing and sorting the contigs by length

212   (Stajich, JE., Automatic Assembly For the Fungi. https://github.com/stajichlab/AAFTF). The

213   pipeline performs assembly with Spades 3.10.0 using default parameters which consider 3 kmer

214   values and select the optimal assembly based on summary statistics (Nurk et al., 2013). As a

215   measure to assess genome completeness, all samples were run through BUSCO 3.0.2 using the

216   Basidiomycota database (Simao et al., 2015). For most samples, RNA sequence data was used to

217   facilitate annotation. When samples were too distantly related to map efficiently to the RNA

218   sequencing reads, these taxa were annotated without aligning to the RNA sequences (Table 5).

219   The RNA sequences used for reference were also generated from JGI CSP #2000 and will be

220   published as part of a separate study.

221   All samples were then prepared for gene prediction using Funannotate 1.6.0 (Palmer JP,

222   Stajich JE. 2018, https://github.com/nextgenusfs/funannotate), which performs all the steps

223   necessary for genome annotation from gene prediction training to final gene consensus model,

224   functional prediction, and dataset preparation for deposition into GenBank. The tool first runs

225   RepeatMasker 4.0.7 (http://www.repeatmasker.org). This "softmasks" the genome by converting

226   repetitive elements into lowercase letters in the assembly files. This step is necessary for the gene

227   prediction steps that follow. After masking, each assembly is run through a training step to

228   provide the initial models for the *ab initio* gene prediction programs AUGUSTUS 3.3.0 (Keller

229   et al., 2011; Stanke and Waack, 2003), SNAP (Korf, 2004), CodingQuarry (Testa et al.,

230   2015), and GeneMark-ES/ET 4.38.0 (Lomsadze et al., 2014). Protein sequences are also aligned

231   with diamond (Buchfink et al., 2015) and gene models polished with exonerate (Slater and

232   Birney, 2005). When RNASeq reads were available for a strain, these were applied as part of a

233   training step which first aligned short RNASeq reads, followed by assembly of these reads into

234   contig with Trinity. Finally these assembled transcripts were aligned to the genome to produce

235   gene models which were used for gene predictor training. Table 5 has the strains which were

236   able to use the RNASeq data as support for gene model training and prediction. These combined

237   evidence of these gene predictions, both *ab initio* and protein and transcript sequence based,

238   were combined with EvidenceModeler to use combined evidence to predict a final set of protein

239   coding genes. In addition tRNA gene predictions were performed with tRNAScan-SE (Lowe and

240   Eddy, 1997). The resulting predicted protein files were then used for the phylogenetic analyses.

241   **2.6 Phylogenomic analysis**

242   We used the pipeline PHYling 1.0 (https://doi.org/10.5281/zenodo.1257001) developed

243   by the Stajich lab, to extract orthologous genes from the predicted proteins of our taxa (Spatafora

244   et al., 2016). PHYling uses Hmmer3 (v3.2.1) to compare our predicted proteins to a list of

245   Profile-Hidden-Markov models of phylogenetically informative markers. The list we used is the

246   434 orthologous gene set (https://doi.org/10.5281/zenodo.1251476) constructed by the 1000

247   Fungal Genomes Project and identified as single-copy in orthologous gene clusters available

248   from the Joint Genome Institute's MycoCosm repository (Grigoriev et al., 2014). We used

249 hmmsearch to compare each sample's proteome to the 434 gene list. The protein sequence

250 homologs we identified were aligned to the marker-profile HMM with hmmalign. These

251 alignments were concatenated to run a phylogenetic analysis with RAxML 8.2.12 (Stamatakis,

252 2006; Stamatakis et al., 2008). The model of evolution was determined automatically and

253 bootstrapped with 100 replicates. The gene trees generated from RAxML were used to construct

254 a consensus tree with ASTRAL-III 5.6.3 (Mirarab et al., 2014; Zhang et al., 2017).

255 **2.7 Data accessibility**

256 Isolates with UAMH numbers are stored in the UAMH Centre for Global Microfungal

257 Biodiversity repository. Raw DNA sequence data have been deposited in SRA and are associated

258 with BioProjects listed in Table 3. Scripts used for these analyses and all alignments, trees, and

259 intermediate files will be made available in a Dryad repository upon publication. BioProject IDs

260 and JGI Mycocosm repositories are summarized in Table 3.

261
262 <div align="center">**3. RESULTS**</div>
263 **3.1 ITS identifications**

264 For the 35 isolates studied, ITS identifications, primers used and the length of each

265 sequence are summarized in Table 2. One sample sent to the Joint Genome Institute was not

266 sequenced due to poor DNA quality. Two isolates were identified as contaminants (isolates 420

267 and 422) and were excluded from further analysis (Table 2). Only four out of 35 isolates were

268 identified to species.

269 **3.2 Shallow genome sequencing and annotation**

270 Shallow genome sequencing of 32 fungal isolates resulted in a wide range in the number

271 of genes annotated in each individual genome. The isolate *Serendipita sp* 396 has the least

272 number of annotated genes at 8,285 and *Ceratobasidium sp* 428 has the most at 25,099. The

273 BUSCO completeness scores ranged from 54.2% to 96.6% of the 1335 orthologues in the

274 BUSCO dataset. For assembly statistics see Table 4 and for BUSCO completeness scores see

275 Table 5. Out of 435 orthologous genes, 429 had enough significant hits for further analysis. The

276 number of genes present for each taxa ranged from 299 in *Tulasnella sp* 408 to 425 in the

277 outgroups *Auricularia subglabra* and *Botryobasidium botryosum*. For full matrix occupancy see

278 Table 6. The outgroup *Kockovaella imperatae* contained 408 of the 429 genes so those 408

279 sequences were included in the phylogenetic analyses. The concatenated alignment has 128,774

280 distinct alignment patterns and is 14.31% gaps.

281 **3.3 Phylogenetic analysis**

282   The best concatenated tree likelihood is -3406977.36. The bootstrap (BS) support is

283 overall very high with the majority of branches at 100 (Figure 2). Eight branches have bootstrap

284 values below 100, and, of those, only three are below 75. The ASTRAL-III tree shows high

285 congruence with the concatenated tree and all but five branches are supported with 0.7 local

286 posterior probability or higher (Figure 3). The two phylogenies have the exact same topology on

287 the class, order, and family level and recapitulate with high support previously published

288 relationships between orders in the Agaricomycetes [Cantharellales, [Sebacinales,

289 [Auriculariales]]]. The phylogenies are highly congruent within Cantharellales, however, the

290 relationships between *Serendipita* isolates are quite different as discussed below.

291   Within the Cantharellales, we have strong support (94 BS, 0.99 posterior probability) for

292 Ceratobasidiaceae as sister to the rest of the order. Within Ceratobasidiaceae, the Ceratobasidium

293 isolates cluster together with very strong support with the exception of *Ceratobasidium sp* 423,

294 which is nested within *Rhizoctonia solani* and *Thanatephorus cucumeris*. The only difference

295 between the ML and ASTRAL-III in the family is the placement of *Ceratobasidium sp* 370. In

296 the ML tree, 370 is sister to a clade of [414, [394+UAMH11750]] and in the ASTRAL-III tree,

297 370 is sister with isolate 414 and equally related to 394+UAMH11750. There is no phylogenetic

298 signal based on orchid source, geographic location (Figure 3, Table 1). Both trees show

299 Tulasnellaceae as sister to the clade [Botryobasidium, [Clavulina, [Cantherellus + Hydnum]]]

300 with 100BS and 1.0 pp. The relationships in *Tulasnella* are highly supported with all but one

301 branch with 100 BS values and all but two branches with pps less than 1.0. Notably, the genome

302 sequence and the shallow genome sequence data for *Tulasnella calospora* UAMH 9824 are sister

303 to each other in the tree, though two other isolates are included in a clade with *Tulasnella*

304 *calospora* AL13.

305   The samples in the Sebacinales are not as well-resolved. The *Serendipia* isolates have the

306 least support overall due to the short branches of all isolates aside from *Serendipita* 399, which is

307 sister to the rest. All *Serendipita spp* in this study are most closely related to *Serendipita*

308 (=*Piriformospora) indica* with 100 BS/1.0. It is important to note our inclusion of the reference

309 genome *Serendipita 407* (Serendipita sp._407_v1.0) and a shallow genome sequence of the same

310 isolate (Serendipita_sp_407.Orchid). In our dataset these two samples are not sister to each other.

311 In the quartet-based ASTRAL-III tree, *Serendipita* 400 and 411 are sister to each other with 0.77

312 posterior probability, whereas in the concatenated tree, the genome of isolate 407 was sister to

313 the rest of the Serendipita isolates aside from 399. The short branches in this group indicate a

314 small number of changes in the alignment in the ML tree and a high degree of discordance in the

315 ASTRAL-III tree. All of the Serendipita isolates are from epiphytic orchids in the Florida

316 Panther National Wildlife Refuge (Figure 3, Table 1). In the ASTRAL tree, Serendipita spp tend

317 to cluster with orchid source compared to the ML tree.

4. DISCUSSION

4.1 Overview

The primary goal of this study was to use shallow genome sequencing and phylogenetic methods to uncover the evolutionary relationships in a collection of fungal isolates that interact with endangered orchid species. The secondary goal was to leverage current genomic resources to investigate relationships among the orders, families and genera of Agaricomycetes, with a focus on Ceratobasicaceae, Tulasnellaceae, and Sebacineaceae. Understanding of species in the fungal genera that facilitate orchid germination is extremely poor, as the number of formally described species is much lower than the diversity of fungi revealed from metagenomic or environmental sequencing. The results of this study add to our understanding of the genetic diversity of these fungal taxa and provide an example of how sequence data can be incorporated with taxonomic expertise to better describe fungal species.

The fungi that help germinate orchids were first categorized under one "form genus" called *Rhizoctonia* (Currah et al., 1997). This classification is not phylogenetically informative and today we know many orchid symbionts come from two orders (Cantharellales and Sebacinales) in the class Agaricomycetes (Hibbett, 2006). However, the taxonomy remains to be fully resolved. One reason classification can be difficult in these taxa is that these isolates do not sporulate or make sexual structures in laboratory conditions. Another is that traditionally, fungi were classified under two different names – the sexual stage (teleomorph) or vegetative state (anamorph). This policy ended during the 2011 International Botanical Congress when the Nomenclature Section voted to eliminate this dual nomenclature system (Hibbett and Taylor, 2013). Many of the names published in literature are no longer considered the correct taxonomy though in many cases these changes are not strongly reinforced. This study examines the

342   phylogenetic relationships of a collection of isolates so that the genetic distance of these strains

343   is known and to provide a framework for future evolutionary questions. Data from these

344   phylogenies can also provide evidence for new species or to revise current species concepts.

345   Understanding of taxonomy and species relationships is critical for testing evolutionary

346   hypotheses. Increased sampling within taxonomic groups and from sites around the globe is

347   necessary for future studies.

348

349   **4.2 Relationships among Orders and Families**

350   We used shallow genome sequencing for phylogenomics to describe the evolutionary

351   relationships among a collection of orchid mycorrhizal fungi. We also included numerous

352   outgroups to span the amount of biodiversity represented by these fungi. The large number of

353   coding genes allowed us to provide strong evidence for relationships between orders and a novel

354   result within the families of Cantharellales. Our results show strong support for the relationships

355   [Cantharellales, [Sebacinales, [Auriculariales]]]. This is consistent with previously reported

356   studies (Nagy et al,. 2016). Within Cantharellales, the taxonomy is less certain and is still

357   undergoing changes. For example, Dictionary of the Fungi lists seven families while Hibbett et

358   al., (2014) claim four by defining Clavulinaceae and Cantharellaceae as synonymous with

359   Hydnaceae. This decision seems to be based on the authors' interpretations as the data in the

360   papers they cite don't support this conclusion (Leacock 2018). Gónzalez et al., (2016), found

361   some support for the relationships [Tulasnellaceae, [Ceratobasidiaceae +Botryobasidiaceae,

362   [Hydnaceae]]] based on the markers ITS-LSU, rpb2, tef1, and atp6. They did state that multiple

363   coding genes would be necessary to see if their result was robust (Gónzalez et al., 2016). Our

364   results show strong support (99 BS and .94 posterior probability) for Ceratobasidiaceae as the

365   sister family to [Tulasnellaceae, [Botryobasidiaceae + rest of Cantharellales]]. We did only

366   include one sample from the four groups besides Ceratobasidiaceae and Tulasnellaceae so more

367   sampling is needed in this group of fungi to produce a robust and consistent phylogenetic

368   inference.

369   **4.3 Relationships in Ceratobasidiaceae**

370   The *Ceratobasidium* samples are closely related with the exception of isolate

371   *Ceratobasidium sp* 423 that is nested within *Rhizoctonia solani* and *Thanatephorus cucumeris*

372   (Figures 2, 3). These results are consistent with Veldre et al., (2013), who found that the genera

373   Ceratobasididum and Thanatephorus are polyphyletic. Given the type specimen for

374   Ceratobasidium has since been placed in the Auriculariales, Oberwinkler et al., (2013a)

375   recommended Ceratobasidium should be renamed Rhizoctonia. Given these taxonomic

376   conundrums, attention is needed to make a robust classification system. Something we found

377   affirming was the close relationship of isolates Ceratobasidium 11750 and Ceratobasidium 394.

378   Based on a nearly identical ITS sequence alignment, these isolates were assumed to be very

379   closely related. This result is noteworthy because they have differential abilities to germinate

380   seeds from the endangered Ghost orchid, *Dendrophylax lindenii*. 394 can germinate seeds but

381   379 does not. More sampling is needed to compare how the isolates included in our study are

382   related to other *Ceratobasidium spp*. that are in defined Anastomosis Groups.

383

384   **4.4 Relationships in Tulasnellaceae**

385   Our *Tulasnella* isolates show a well-supported monophyletic clade in both phylogenetic

386   trees (Figures 2 and 3). Without further targeted sampling, it is premature to delimit species

387   boundaries; however, one species that could use revision is *Tulasnella calospora*. In both the

16

388   concatenated and coalescent phylogenies, the two *T. calospora* genomes are not sister to each

389   other but include the isolates 408 and 417, which were not identified as *T. calospora* based on

390   the ITS sequence. This result could be a function of the relatively low number of orthologous

391   genes that we recovered from 408 and 417, 291 and 330 out of 434, respectively (Tables 5 and

392   6). However, others have voiced concern over the species concept (Melissa McCormick, pers.

393   comm.).

394   Three isolates in this analysis are from the Hawaiian island of Molokai (330, 331, and

395   332; Table 1). These isolates cluster very closely in both phylogenies and are sister to three

396   isolates of *Tulasnella inquilina*. These isolates turn pink when exposed to light and have highly

397   divergent ITS sequences from the other Tulasnella isolates in this analysis. The strong support

398   for the monophyly of these Hawaiian samples, and their placement in the tree, suggest a

399   potentially new species. With increased sampling, more robust methods to delineate species

400   boundaries such as those used in (Whitehead et al., 2017) and we will have the power to better

401   describe the diversity of orchid mycorrhizal fungi.

402   **4.5 Relationships in Sebacinales**

403   All of the *Serenipita* isolates in this analysis are from the Florida National Wildlife

404   Panther Refuge (NWPR) in Florida and they are associated with three different epiphytic orchid

405   species (Table 1). In both phylogenetic analyses, *Serendipita* 399 is sister to the rest of our

406   samples. Growing on PDA, 399 looks morphologically distinct from the other *Serendipita sp* due

407   to a darker orange pigment and a crustose layer on the surface of the agar. This isolate also

408   grows much more slowly than other Serendipita taxa, it would take longer than four weeks for

409   the fungus to grow to the edge of a standard petri dish. For the remaining samples, it could be,

410   that there is one main species or population of *Serendipita* that grows in orchid roots in the

411   NWPR as their relationships are poorly resolved in the RAxML phylogeny and highly

412   incongruent between the two phylogenies. However, in the ASTRAL analysis, the *Serendipita*

413   isolates cluster somewhat closely by the orchid species from which they were isolated though

414   this is not a strong signal (Figure 3). A more thorough and targeted analysis is required to

415   determine the number of distinct populations of these fungi in the Florida National Panther

416   Wildlife Refuge similar to that conducted by Ruibal et al., (2017) to describe the population

417   structure of *Tulasnella prima* in Australia. It would be interesting to survey the fungi growing in

418   the roots of all plant species in the NWPR to determine the genetic diversity of *Serendipita*

419   across the landscape. Such an experiment would show whether orchids are using a narrow

420   distribution of fungi or if the plants are less discerning but the genetic diversity of the fungi is

421   simply very low.

422   Another result from our analysis shows that these fungal strains are most closely related

423   to *Piriformospora indica*, a known ectomycorrhizal fungus species (Varma et al., 2001). Many

424   fungi in the order Sebacinales are ecologically characterized as ectomycorrhizal fungi and

425   interact with a wide diversity of plant species (Kohler et al., 2015). Indeed, researchers are

426   isolating fungi in the Sebacinales from plants like switchgrass (*Panicum virgatum*) to determine

427   the benefit of these fungi for applications in agriculture (Craven and Ray, 2019). Orchids might

428   contribute to this effort, as it took more than one year for the Craven lab to isolate one strain of

429   *Sebacina vermifera ssp. bescii* from switchgrass; similar fungi are much more easy to isolate

430   from orchid roots (Prasun Ray, pers. comm.). Orchids could be environmental filters for fungi

431   that could be beneficial in many plant-fungal interactions.

432

433 **4.6 Future directions**

434   The next steps stemming from this study are to combine the phylogenetic relationships

435 with taxonomic expertise to name new species or to revisit problematic species concepts like

436 *Tulasnella calospora.* Additionally, it would be beneficial to sequence the genome of the type

437 specimens for many of these genera and species. Being able to compare the genetic sequences of

438 the type specimens would be extremely beneficial for fungal species that do not present sexual

439 characteristics in the lab. A set of fifteen isolates from the collection have been sequenced on the

440 PacBio platform and will be assembled into reference genomes as part of another aim of the

441 Community Sequencing Proposal (Table 3).

442

443   **ACKNOWLEDGEMENTS**

444

456   Science Foundation, IOS 1339156, and the Department of Energy (DOE), Defense Threat

457   Reduction Agency (DTRA).

458
459

460 LITERATURE CITED
461

462 Barto, E.K., Weidenhamer, J.D., Cipollini, D., Rillig, M.C., 2012. Fungal superhighways: do
463   common mycorrhizal networks enhance below ground communication? Trends in Plant
464   Science 17, 633-637.
465 Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010. ITS as
466   an environmental DNA barcode for fungi: an in silico approach reveals potential PCR
467   biases. BMC microbiology 10, 189.
468 Blackwell, M., 2011. The Fungi: 1, 2, 3 … 5.1 million species? American Journal of Botany 98,
469   426-438.
470 Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using
471   DIAMOND. Nature Methods 12, 59-60.
472 Bushnell, B., 2014. BBmap.
473 Clements, M., Muir, H., Cribb, P., 1986. A preliminary report on the symbiotic germination of
474   European terrestrial orchids. Kew Bulletin, 437-445.
475 Craven, K.D., Ray, P., 2019. More than Serendipity: The Potential to Manage Soil Carbon and
476   Emissions While Promoting Low-Input Agriculture with Serendipitoid Mycorrhizae.
477   Phytobiomes Journal 3, 161-164.
478 Currah, R., Zelmer, C., Hambleton, S., Richardson, K., 1997. Fungi from orchid mycorrhizas.
479   Orchid Biology. Springer, pp. 117-170.
480 Fochi, V., Chitarra, W., Kohler, A., Voyron, S., Singan, V.R., Lindquist, E.A., Barry, K.W.,
481   Girlanda, M., Grigoriev, I.V., Martin, F., Balestrini, R., Perotto, S., 2017a. Fungal and
482   plant gene expression in the Tulasnella calospora-Serapias vomeracea symbiosis provides
483   clues about nitrogen pathways in orchid mycorrhizas. New Phytol 213, 365-379.
484 Fochi, V., Falla, N., Girlanda, M., Perotto, S., Balestrini, R., 2017b. Cell-specific expression of
485   plant nutrient transporter genes in orchid mycorrhizae. Plant Sci 263, 39-45.
486 Gónzalez, D., Rodriguez-Carres, M., Boekhout, T., Stalpers, J., Kuramae, E.E., Nakatani, A.K.,
487   Vilgalys, R., Cubeta, M.A., 2016. Phylogenetic relationships of Rhizoctonia fungi within
488   the Cantharellales. Fungal Biol 120, 603-619.
489 Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A.,
490   Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014.
491   MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Research 42,
492   D699-D704.
493 Hawksworth, D.L., Lücking, R., 2017. Fungal Diversity Revisited: 2.2 to 3.8 Million Species.
494   The Fungal Kingdom. American Society of Microbiology.
495 Hibbett, D.S., 2006. A phylogenetic overview of the Agaricomycotina. Mycologia 98, 917-925.
496 Hibbett, D.S., Bauer, R., Binder, M., Giachini, A.J., Hosaka, K., Justo, A., Larsson, E., Larsson,
497   K.H., Lawrey, J.D., Miettinen, O., Nagy, L.G., Nilsson, R.H., Weiss, M., Thorn, R.G.,
498   2014. 14 Agaricomycetes. In: McLaughlin, D.J., Spatafora, J.W. (Eds.), Systematics and
499   Evolution: Part A. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 373-429.
500 Hibbett, D.S., Taylor, J.W., 2013. Fungal systematics: is a new age of enlightenment at hand?
501   Nat Rev Microbiol 11, 129-133.
502 Jung, S.C., Martinez-Medina, A., Lopez-Raez, J.A., Pozo, M.J., 2012. Mycorrhiza-Induced
503   Resistance and Priming of Plant Defenses. Journal of Chemical Ecology 38, 651-664.
504 Keller, O., Kollmar, M., Stanke, M., Waack, S., 2011. A novel hybrid gene prediction method
505   employing protein multiple sequence alignments. Bioinformatics 27, 757-763.

506 Kirk, P.M., Ainsworth, G.C., Bisby, G.R., 2008. Ainsworth & Bisby's dictionary of the fungi.
507   CABI.
508 Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., Canback, B., Choi, C.,
509   Cichocki, N., Clum, A., Colpaert, J., Copeland, A., Costa, M.D., Dore, J., Floudas, D.,
510   Gay, G., Girlanda, M., Henrissat, B., Herrmann, S., Hess, J., Hogberg, N., Johansson, T.,
511   Khouja, H.R., LaButti, K., Lahrmann, U., Levasseur, A., Lindquist, E.A., Lipzen, A.,
512   Marmeisse, R., Martino, E., Murat, C., Ngan, C.Y., Nehls, U., Plett, J.M., Pringle, A.,
513   Ohm, R.A., Perotto, S., Peter, M., Riley, R., Rineau, F., Ruytinx, J., Salamov, A., Shah,
514   F., Sun, H., Tarkka, M., Tritt, A., Veneault-Fourrey, C., Zuccaro, A., Mycorrhizal
515   Genomics Initiative, C., Tunlid, A., Grigoriev, I.V., Hibbett, D.S., Martin, F., 2015.
516   Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in
517   mycorrhizal mutualists. Nat Genet 47, 410-415.
518 Korf, I., 2004. Gene finding in novel genomes. BMC Bioinformatics 5, 59.
519 Kuga, Y., Sakamoto, N., Yurimoto, H., 2014. Stable isotope cellular imaging reveals that both
520   live and degenerating fungal pelotons transfer carbon and nitrogen to orchid protocorms.
521   New Phytol 202, 594-605.
522 Larsen, B.B., Miller, E.C., Rhodes, M.K., Wiens, J.J., 2017. Inordinate Fondness Multiplied and
523   Redistributed: the Number of Species on Earth and the New Pie of Life. The Quarterly
524   Review of Biology 92, 229-265.
525 Leacock, P.R., 2018. Cantharellales - MycoGuide.
526 Linde, C.C., May, T.W., Phillips, R.D., Ruibal, M., Smith, L.M., Peakall, R., 2017. New species
527   of Tulasnella associated with terrestrial orchids in Australia. IMA Fungus 8, 28-47.
528 Lomsadze, A., Burns, P.D., Borodovsky, M., 2014. Integration of mapped RNA-Seq reads into
529   automatic training of eukaryotic gene finding algorithm. Nucleic Acids Research 42,
530   e119-e119.
531 Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: A Program for Improved Detection of Transfer
532   RNA Genes in Genomic Sequence. Nucleic Acids Research 25, 955-964.
533 Lutzoni, F., Nowak, M.D., Alfaro, M.E., Reeb, V., Miadlikowska, J., Krug, M., Arnold, A.E.,
534   Lewis, L.A., Swofford, D.L., Hibbett, D., Hilu, K., James, T.Y., Quandt, D., Magallón,
535   S., 2018. Contemporaneous radiations of fungi and plants linked to symbiosis. Nature
536   Communications 9.
537 Martos, F., Munoz, F., Pailler, T., Kottke, I., Gonneau, C., Selosse, M.-A., 2012. The role of
538   epiphytism in architecture and evolutionary constraint within mycorrhizal networks of
539   tropical orchids. Molecular Ecology 21, 5098-5109.
540 McCormick, M.K., Taylor, D.L., Whigham, D.F., Burnett, R.K., 2016. Germination patterns in
541   three terrestrial orchids relate to abundance of mycorrhizal fungi. Journal of Ecology 104,
542   744-754.
543 Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014.
544   ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30,
545   i541-i548.
546 Nagy, L.G., Riley, R., Tritt, A., Adam, C., Daum, C., Floudas, D., Sun, H., Yadav, J.S.,
547   Pangilinan, J., Larsson, K.-H., Matsuura, K., Barry, K., Labutti, K., Kuo, R., Ohm, R.A.,
548   Bhattacharya, S.S., Shirouzu, T., Yoshinaga, Y., Martin, F.M., Grigoriev, I.V., Hibbett,
549   D.S., 2016. Comparative Genomics of Early-Diverging Mushroom-Forming Fungi
550   Provides Insights into the Origins of Lignocellulose Decay Capabilities. Molecular
551   Biology and Evolution 33, 959-970.

552   Nilsson, R.H., Hyde, K.D., Pawłowska, J., Ryberg, M., Tedersoo, L., Aas, A.B., Alias, S.A.,
553   Alves, A., Anderson, C.L., Antonelli, A., Arnold, A.E., Bahnmann, B., Bahram, M.,
554   Bengtsson-Palme, J., Berlin, A., Branco, S., Chomnunti, P., Dissanayake, A., Drenkhan,
555   R., Friberg, H., Frøslev, T.G., Halwachs, B., Hartmann, M., Henricot, B., Jayawardena,
556   R., Jumpponen, A., Kauserud, H., Koskela, S., Kulik, T., Liimatainen, K., Lindahl, B.D.,
557   Lindner, D., Liu, J.-K., Maharachchikumbura, S., Manamgoda, D., Martinsson, S.,
558   Neves, M.A., Niskanen, T., Nylinder, S., Pereira, O.L., Pinho, D.B., Porter, T.M.,
559   Queloz, V., Riit, T., Sánchez-García, M., De Sousa, F., Stefańczyk, E., Tadych, M.,
560   Takamatsu, S., Tian, Q., Udayanga, D., Unterseher, M., Wang, Z., Wikee, S., Yan, J.,
561   Larsson, E., Larsson, K.-H., Kõljalg, U., Abarenkov, K., 2014. Improving ITS sequence
562   data for identification of plant pathogenic fungi. 67, 11-19.
563   Nilsson, R.H., Larsson, K.-H., Taylor, A.F S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D.,
564   Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., Saar, I., Kõljalg, U., Abarenkov,
565   K., 2019. The UNITE database for molecular identification of fungi: handling dark taxa
566   and parallel taxonomic classifications. Nucleic Acids Research 47, D259-D264.
567   Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky,
568   A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R.,
569   Clingenpeel, S.R., Woyke, T., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2013.
570   Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. Springer
571   Berlin Heidelberg, Berlin, Heidelberg, pp. 158-170.
572   Oberwinkler, F., Riess, K., Bauer, R., Kirschner, R., Garnica, S., 2013a. Taxonomic re-
573   evaluation of the Ceratobasidium-Rhizoctonia complex and Rhizoctonia butinii, a new
574   species attacking spruce. Mycological Progress 12, 763-776.
575   Oberwinkler, F., Riess, K., Bauer, R., Selosse, M.-A., Weiß, M., Garnica, S., Zuccaro, A.,
576   2013b. Enigmatic Sebacinales. Mycological Progress 12, 1-27.
577   Palmer, J., Stajich, J.E., 2018. Funannotate.
578   Peterson, R.L., Massicotte, H.B., 2004. Exploring structural definitions of mycorrhizas, with
579   emphasis on nutrient-exchange interfaces. Canadian Journal of Botany 82, 1074-1088.
580   Rasmussen, H.N., Dixon, K.W., Jersakova, J., Tesitelova, T., 2015. Germination and seedling
581   establishment in orchids: a complex of requirements. Ann Bot 116, 391-402.
582   Ray, P., Ishiga, T., Decker, S.R., Turner, G.B., Craven, K.D., 2015. A Novel Delivery System
583   for the Root Symbiotic Fungus, Sebacina vermifera, and Consequent Biomass
584   Enhancement of Low Lignin COMT Switchgrass Lines. BioEnergy Research 8, 922-933.
585   Ruibal, M.P., Peakall, R., Foret, S., Linde, C.C., 2014. Development of Phylogenetic Markers for
586   Sebacina (Sebacinaceae) Mycorrhizal Fungi Associated with Australian Orchids.
587   Applications in Plant Sciences 2, 1400015.
588   Ruibal, M.P., Peakall, R., Smith, L.M., Linde, C.C., 2013. Phylogenetic and microsatellite
589   markers for Tulasnella (Tulasnellaceae) mycorrhizal fungi associated with Australian
590   orchids. Appl Plant Sci 1.
591   Ruibal, M.P., Triponez, Y., Smith, L.M., Peakall, R., Linde, C.C., 2017. Population structure of
592   an orchid mycorrhizal fungus with genus-wide specificity. Scientific Reports 7.
593   Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015.
594   BUSCO: assessing genome assembly and annotation completeness with single-copy
595   orthologs. Bioinformatics 31, 3210-3212.
596   Slater, G.S.C., Birney, E., 2005. Automated generation of heuristics for biological sequence
597   comparison. BMC Bioinformatics 6, 31.

598   Smith, S.E., Read, D.J., 2008. Mycorrhizal symbiosis. Academic press.
599   Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., Bonito, G.,
600   Corradi, N., Grigoriev, I., Gryganskyi, A., James, T.Y., O'Donnell, K., Roberson, R.W.,
601   Taylor, T.N., Uehling, J., Vilgalys, R., White, M.M., Stajich, J.E., 2016. A phylum-level
602   phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia
603   108, 1028-1046.
604   Stajich, J.E., 2019. Automatic Assembly For the Fungi.
605   Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with
606   thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.
607   Stamatakis, A., Hoover, P., Rougemont, J., 2008. A Rapid Bootstrap Algorithm for the RAxML
608   Web Servers. Systematic Biology 57, 758-771.
609   Stanke, M., Waack, S., 2003. Gene prediction with a hidden Markov model and a new intron
610   submodel. Bioinformatics 19, ii215-ii225.
611   Taylor, D.L., McCormick, M.K., 2008. Internal transcribed spacer primers and sequences for
612   improved characterization of basidiomycetous orchid mycorrhizas. New Phytol 177,
613   1020-1033.
614   Testa, A.C., Hane, J.K., Ellwood, S.R., Oliver, R.P., 2015. CodingQuarry: highly accurate
615   hidden Markov model gene prediction in fungal genomes using RNA-seq transcripts.
616   BMC Genomics 16, 170.
617   Varma, A., Singh, A., Sudha, Sahay, N.S., Sharma, J., Roy, A., Kumari, M., Rana, D., Thakran,
618   S., Deka, D., Bharti, K., Hurek, T., Blechert, O., Rexer, K.H., Kost, G., Hahn, A., Maier,
619   W., Walter, M., Strack, D., Kranner, I., 2001. Piriformospora indica: An Axenically
620   Culturable Mycorrhiza-Like Endosymbiotic Fungus. In: Hock, B. (Ed.), Fungal
621   Associations. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 125-150.
622   Veldre, V., Abarenkov, K., Bahram, M., Martos, F., Selosse, M.-A., Tamm, H., Kõljalg, U.,
623   Tedersoo, L., 2013. Evolution of nutritional modes of Ceratobasidiaceae (Cantharellales,
624   Basidiomycota) as revealed from publicly available ITS sequences. Fungal Ecology 6,
625   256-268.
626   Wang, W., Shi, J., Xie, Q., Jiang, Y., Yu, N., Wang, E., 2017. Nutrient Exchange and Regulation
627   in Arbuscular Mycorrhizal Symbiosis. Molecular Plant 10, 1147-1158.
628   Weiss, M., Waller, F., Zuccaro, A., Selosse, M.A., 2016. Sebacinales - one thousand and one
629   interactions with land plants. New Phytol 211, 20-40.
630   White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal
631   ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and
632   applications 18, 315-322.
633   Whitehead, M.R., Catullo, R.A., Ruibal, M., Dixon, K.W., Peakall, R., Linde, C.C., 2017.
634   Evaluating multilocus Bayesian species delimitation for discovery of cryptic mycorrhizal
635   diversity. Fungal Ecology 26, 74-84.
636   Yoder, J.A., Imfeld, S.M., Heydinger, D.J., Hart, C.E., Collier, M.H., Gribbins, K.M., Zettler,
637   L.W., 2010. Comparative water balance profiles of Orchidaceae seeds for epiphytic and
638   terrestrial taxa endemic to North America. Plant Ecology 211, 7-17.
639   Zettler, L.W., Corey, L.L., 2018. Orchid Mycorrhizal Fungi: Isolation and Identification
640   Techniques. Orchid Propagation: From Laboratories to Greenhouses—Methods and
641   Protocols. Springer, pp. 27-59.
642   Zhang, C., Sayyari, E., Mirarab, S., 2017. ASTRAL-III: Increased Scalability and Impacts of
643   Contracting Low Support Branches. Springer International Publishing, Cham, pp. 53-75.

644
645

Figure 1. Morphological examples of Tulasnella, Ceratobasidium, and Serendipita.
One representative from each genus from the Zettler collection. All three isolates started growing on Potato Dextrose Agar on the same day as indicated by the date on the petri dish (25 November 2015). Photographs: Sarah Unruh.
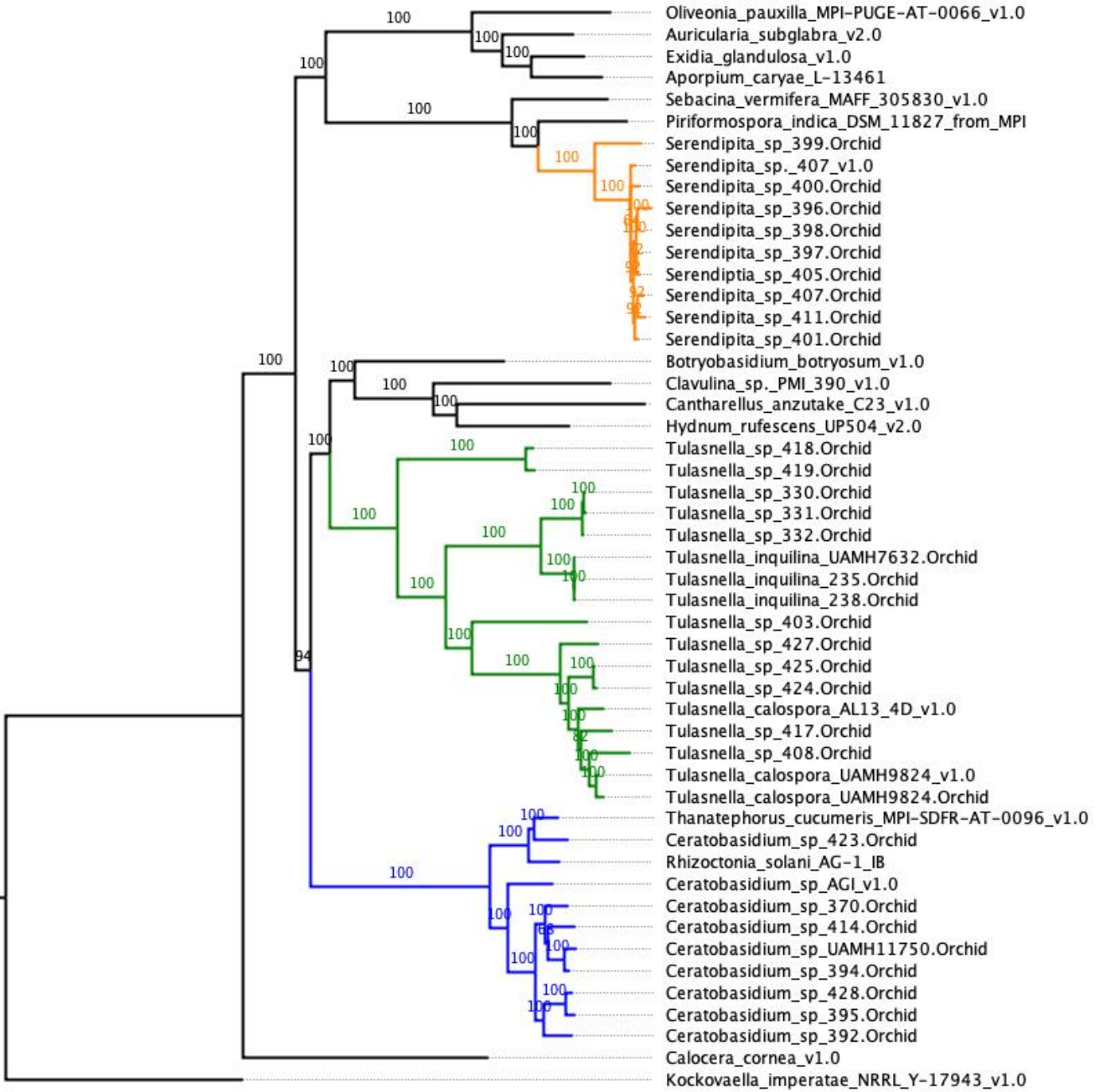
Figure 2. Concatenation-based phylogeny of orchid mycorrhizal fungi.

Phylogenetic tree of the orchid mycorrhizal fungi in the Zettler collection with outgroups from the MycoCosm repository (genome.jgi.doe.gov/mycocosm/home). Alignments were made with the Phyling pipeline and the phylogeny was built with RAxML.
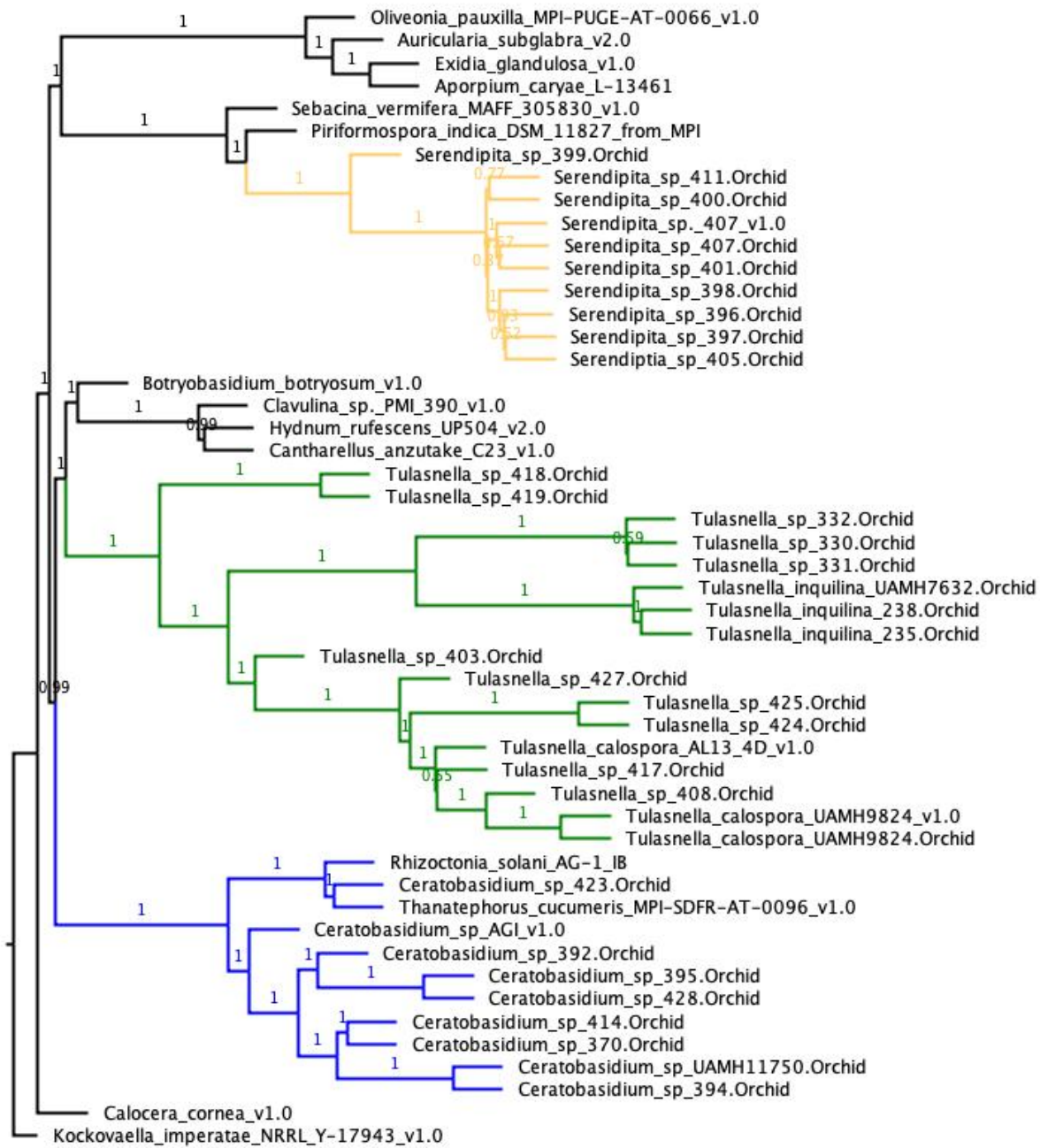
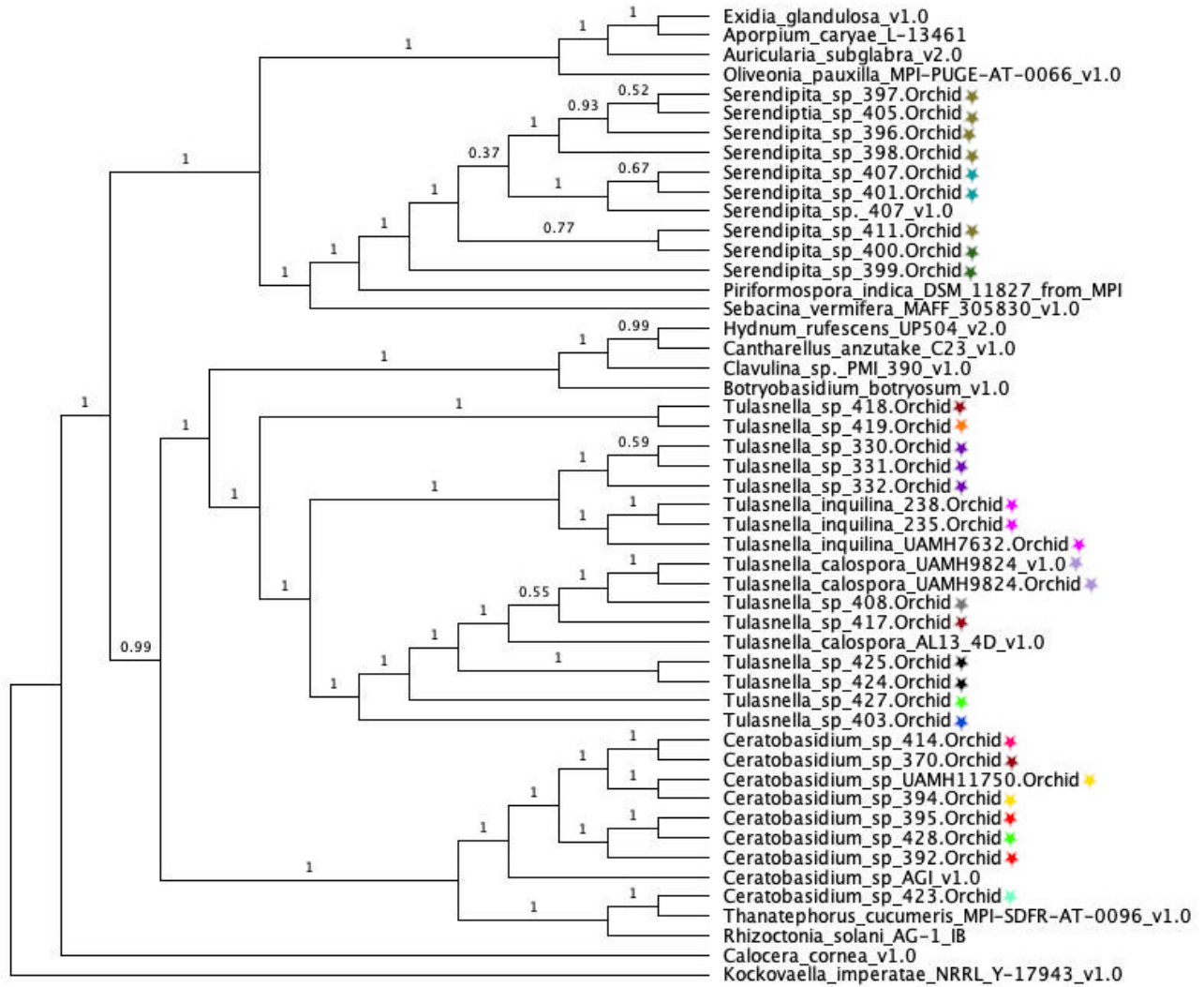**Figure 3. Quartet-based phylogeny of orchid mycorrhizal fungi.**

Phylogenetic tree of the 32 orchid mycorrhizal fungi in the Zettler collection with 16 outgroups from the MycoCosm repository (genome.jgi.doe.gov/mycocosm/home). Alignments were made with the Phyling pipeline the gene trees were produced with RAxML and the tree was inferred using ASTRAL-III. All posterior probabilities are reported on the tree.

Figure 4. Annotated Quartet-based phylogeny.

Phylogenetic tree of the 32 orchid mycorrhizal fungi in the Zettler collection with 18 genomes from the MycoCosm repository (genome.jgi.doe.gov/mycocosm/home). Branches were transformed in FigTree and annotated with colored stars indicating the origin they were isolated from.

Sample ID	Species	Strain	Orchid source	Tissue source	Location
Cerato11750	Ceratobasidium sp	UAMH11750	*Dendrophylax lindenii* (Lindl.) Benth. ex Rolfe	root	Florida Panther National Wildlife Refuge (NWR)
Cerato370	Ceratobasidium sp	370	*Platanthera leucophaea* (Nutt.) Lindl.	root	Tuscola Co., MI
Cerato392	Ceratobasidium sp	392	*Campylocentrum pachyrrhizum* (Rchb.f.) Rolfe	root	Florida Panther National Wildlife Refuge (NWR)
Cerato394	Ceratobasidium sp	394	*Dendrophylax lindenii* (Lindl.) Benth. ex Rolfe	root	Florida Panther National Wildlife Refuge (NWR)
Cerato395	Ceratobasidium sp	395	*Campylocentrum pachyrrhizum* (Rchb.f.) Rolfe	root	Florida Panther National Wildlife Refuge (NWR)
Cerato414	Ceratobasidium sp	414	*Platanthera lacera* (Michx.) G.Don	root	Fayette Co., IL
Cerato423	Ceratobasidium sp	423	*Spiranthes vernalis* Engelm. & A.Gray	root	Madison Co., IL
Cerato428	Ceratobasidium sp	428	*Dendrophylax porrectus* (Rchb.f.) Carlsward & Whitten	root	Florida Panther National Wildlife Refuge (NWR)
Serend396	Serendipita sp	396	*Prosthechea cochleata* (L.) W.E.Higgens	root	Florida Panther National Wildlife Refuge (NWR)
Serend397	Serendipita sp	397	*Prosthechea cochleata* (L.) W.E.Higgens	root	Florida Panther National Wildlife Refuge (NWR)
Serend398	Serendipita sp	398	*Prosthechea cochleata* (L.) W.E.Higgens	root	Florida Panther National Wildlife Refuge (NWR)
Serend399	Serendipita sp	399	*Encyclia tampensis* Small	root	Florida Panther National Wildlife Refuge (NWR)
Serend400	Serendipita sp	400	*Encyclia tampensis* Small	root	Florida Panther National Wildlife Refuge (NWR)
Serend401	Serendipita sp	401	*Epidendrum amphistomum* A.Rich	root	Florida Panther National Wildlife Refuge (NWR)
Serend405	Serendiptia sp	405	*Prosthechea cochleata* (L.) W.E.Higgens	root	Florida Panther National Wildlife Refuge (NWR)
Serend407	Serendipita sp	407	*Epidendrum amphistomum* A.Rich	root	Florida Panther National Wildlife Refuge (NWR)
Serend411	Serendipita sp	411	*Prosthechea cochleata* (L.) W.E.Higgens	root	Florida Panther National Wildlife Refuge (NWR)
Tulasn330	Tulasnella sp	330	*Platanthera holochila* (Hillebr.) Kraenzl.	peloton	Molokai, HI
Tulasn331	Tulasnella sp	331	*Platanthera holochila* (Hillebr.) Kraenzl.	peloton	Molokai, HI
Tulasn332	Tulasnella sp	332	*Platanthera holochila* (Hillebr.) Kraenzl.	peloton	Molokai, HI
Tulasn403	Tulasnella sp	403	*Oeceoclades maculata* Lindl.	root	Florida Panther National Wildlife Refuge (NWR)
Tulasn408	Tulasnella sp	408	*Polystachya concreta* (Jacq.) Garay & H.R.Sweet	root	Florida Panther National Wildlife Refuge (NWR)
Tulasn417	Tulasnella sp	417	*Platanthera leucophaea* (Nutt.) Lindl.	root	McHenry Co., IL
Tulasn418	Tulasnella sp	418	*Platanthera leucophaea* (Nutt.) Lindl.	root	McHenry Co., IL
Tulasn419	Tulasnella sp	419	*Cypripedium candidum* Muhl. ex Willd.	protocorm/seedling	McHenry Co., IL
Tulasn424	Tulasnella sp	424	*Platanthera paramoena* A.Gray	root	Fayette Co., IL
Tulasn425	Tulasnella sp	425	*Platanthera paramoena* A.Gray	root	Fayette Co., IL
Tulasn427	Tulasnella sp	427	*Dendrophylax porrectus* (Rchb.f.) Carlsward & Whitten	root	Florida Panther National Wildlife Refuge (NWR)
Tulasn9824	Tulasnella calospora	UAMH9824	*Spiranthes brevilabris* Lindl.	root	Levy Co., FL
Tulinq235	Tulasnella inquilina	235	*Platanthera integrilabia* (Correll) Luer	root	McMinn Co., TN
Tulinq238	Tulasnella inquilina	238	*Platanthera integrilabia* (Correll) Luer	root	McMinn Co., TN
Tulinq7632	Tulasnella inquilina	UAMH7632	*Platanthera integrilabia* (Correll) Luer	root	Greenville, SC

Table 1. Description of fungal isolates.

UAMH numbers refer to the repository number for isolates deposited in the UAMH Centre for Global Microfungal Diversity

SampleID	Morphological ID	Top hit UNITE	Top hit GenBank	Primer sequenced	Length edited in base pairs
Cerato11750	*Ceratobasidium*	Ceratobasidiaceae	Uncultured *Ceratobasidium* clone LP8-Cer1	ITS1	641
Cerato370	*Ceratobasidium*	*Ceratobasidium*	*Ceratobasidium* UAMH 9847	ITS4	538
Cerato392	*Ceratobasidium*	Basidiomycota (same as ncbi	orchid mycorrhizae KH4-8	ITS4	550
Cerato394	*Ceratobasidium*	*Ceratobasidium*	*Ceratobasidium*	ITS1	586
Cerato395	*Ceratobasidium*	Ceratobasidiaceae	*Ceratobasidium sp* JTO161	ITS1	548
Cerato414	*Ceratobasidium*	Ceratobasidiaceae	*Ceratobasidium sp*	ITS1	100
Cerato423	*Ceratobasidium*	*Ceratobasidium*	Uncultured Ceratobasidiaceae clone 207	ITS4	390
Cerato428	*Ceratobasidium*	Ceratobasidiaceae	*Ceratobasidium sp* JTO161	ITS1	420
Serend396	*Serendipita*	Sebacinales (orchid fungus)	uncultured Sebacinales clone	ITS1	189
Serend397	*Serendipita*	Sebacinales	uncultured Sebacinales clone	NL4	680
Serend398	*Serendipita*	Sebacinales	uncultured Sebacinales clone	NL4	567
Serend399	*Serendipita*	Sebacinales	uncultured Sebacinales clone	NL4	740
Serend400	*Serendipita*	Sebacinales	*Serendipita sp* MAFF 305831	NL4	780
Serend401	*Serendipita*	Sebacinales	Uncultured Sebacinales clone LP49-23S	ITS4	614
Serend405	*Serendipita*	*Serendipita*	*Serendipita sp* MAFF 305831	NL4	380
Serend407	*Serendipita*	Sebacinales	Uncultured *Sebacina* mycobiont of Riccardia palmata	*	2316
Serend411	*Serendipita*	Sebacinales	Uncultured Sebacinales clone LP49-23S	ITS3Seb	880

Tulasn330	*Tulasnella*	Tulasnellaceae	Uncultured Tulasnellaceae isolate 55P-Leu13	ITS4	570
Tulasn331	*Tulasnella*	Tulasnellaceae	Uncultured Tulasnellaceae	ITS1	620
Tulasn332	*Tulasnella*	Tulasnellaceae	Uncultured Tulasnellaceae	ITS4-OF	810
Tulasn403	*Tulasnella*	*Tulasnella*	*Tulasnella sp* CH01	ITS1	490
Tulasn408	*Tulasnella*	Tulasnellaceae	Uncultured Tulasnellaceae clone DOf-YC9	ITS4	622
Tulasn417	*Tulasnella*	*Tulasnella*	*Tulasnella sp* 9 MM-2012	ITS1	870
Tulasn418	*Tulasnella*	Tulasnellaceae	Uncultured Tulasnellaceae P94	ITS1	350
Tulasn419	*Tulasnella*	*Tulasnella*	Tulasnellaceae sp Pch 253	ITS4	368
Tulasn424	*Tulasnella*	Tulasnellaceae	Tulasnellaceae	ITS4	605
Tulasn425	*Tulasnella*	*Tulasnella*	*Tulasnella sp* 149	ITS1	570
Tulasn427	*Tulasnella*	Tulasnellaceae	Uncultured *Tulasnella* clone 998OF	ITS4	380
Tulasn9824	*Tulasnella calospora*	*Tulasnella calospora*	*Tulasnella calospora* isolate Pch-QS-0-1	ITS4-Tul	148
Tulinq235	*Epulorhiza inquilina*	*Tulasnella*	*Tulasnella sp* 3MV-2011 PA 053A	ITS4	650
Tulinq238	*Epulorhiza inquilina*	Tulasnellaceae	*Tulasnella sp* 3MV-2011 PA 053A	ITS4	758
Tulinq7632	*Epulorhiza inquilina*	*Tulasnella*	*Tulasnella sp* 3MV-2011 PA 053A	ITS1	790
Isolate 420*	*Tulasnella*	*Phanerochaete australis*	*Phanerochaete australis*	ITS1	350
Isolate 422*	*Tulasnella*	*Trichoderma petersenii*	*Trichoderma sp* isolate ARMI-23	ITS4	390

Table 2. Identifications of fungal isolates based on the internal transcribed spacer (ITS)

SampleID	BioProject or JGI web portal	BioSample
*Ceratobasidium_sp_UAMH11750.Orchid	PRJNA558776	SAMN12498506
Ceratobasidium_sp_370.Orchid	PRJNA557749	SAMN12427929
Ceratobasidium_sp_392.Orchid	PRJNA557750	SAMN12427914
*Ceratobasidium_sp_394.Orchid	PRJNA557751	SAMN12427897
*Ceratobasidium_sp_395.Orchid	PRJNA557752	SAMN12427926
Ceratobasidium_sp_414.Orchid	PRJNA557753	SAMN12427925
*Ceratobasidium_sp_423.Orchid	PRJNA557754	SAMN12427910
Ceratobasidium_sp_428.Orchid	PRJNA557755	SAMN12427923
Serendipita_sp_396.Orchid	PRJNA557757	SAMN12427894
Serendipita_sp_397.Orchid	PRJNA557758	SAMN12427928
Serendipita_sp_398.Orchid	PRJNA557759	SAMN12427900
Serendipita_sp_399.Orchid	PRJNA557760	SAMN12427906
*Serendipita_sp_400.Orchid	PRJNA557761	SAMN12427895
Serendipita_sp_401.Orchid	PRJNA557762	SAMN12427903
*Serendipita_sp_405.Orchid	PRJNA557763	SAMN12427917
*Serendipita_sp_407.Orchid	PRJNA558790	SAMN12498938
*Serendipita_sp_411.Orchid	PRJNA557734	SAMN12427911
Tulasnella_sp_330.Orchid	PRJNA557739	SAMN12427924
Tulasnella_sp_331.Orchid	PRJNA557740	SAMN12427908
*Tulasnella_sp_332.Orchid	PRJNA557741	SAMN12427902
Tulasnella_sp_403.Orchid	PRJNA557742	SAMN12427920
Tulasnella_sp_408.Orchid	PRJNA557743	SAMN12427916
Tulasnella_sp_417.Orchid	PRJNA557744	SAMN12427919
Tulasnella_sp_418.Orchid	PRJNA557745	SAMN12427921
*Tulasnella_sp_419.Orchid	PRJNA557746	SAMN12427922
Tulasnella_sp_424.Orchid	PRJNA557747	SAMN12427899
*Tulasnella_sp_425.Orchid	PRJNA557748	SAMN12427912
*Tulasnella_sp_427.Orchid	PRJNA557733	SAMN12427904
*Tulasnella_calospora_UAMH9824.Orchid	PRJNA558788	SAMN12498837
Tulasnella_inquilina_235.Orchid	PRJNA557736	SAMN12427891

**Tulasnella_inquilina*_238.Orchid*	PRJNA557737	SAMN12427893
**Tulasnella_inquilina*_UAMH7632.Orchid*	PRJNA557738	SAMN12427890
*Aporpium_caryae*_L-13461.	https://mycocosm.jgi.doe.gov/Elmca1	
*Auricularia_subglabra*_v2.0	https://mycocosm.jgi.doe.gov/Aurde3_1	
*Botryobasidium_botryosum*_v1.0	https://mycocosm.jgi.doe.gov/Botbo1	
*Calocera_cornea*_v1.0	https://mycocosm.jgi.doe.gov/Calco1	
*Cantharellus_anzutake*_C23_v1.0	https://mycocosm.jgi.doe.gov/Cananz1	
*Clavulina_sp.*_PMI_390_v1.0	https://mycocosm.jgi.doe.gov/ClaPMI390	
*Exidia_glandulosa*_v1.0	https://mycocosm.jgi.doe.gov/Exigl1	
*Hydnum_rufescens*_UP504_v2.0	https://mycocosm.jgi.doe.gov/Hydru2	
*Kockovaella_imperatae*_NRRL_Y-17943_v1.0	https://mycocosm.jgi.doe.gov/Kocim1	
*Oliveonia_pauxilla*_MPI-PUGE-AT-0066_v1.0	https://mycocosm.jgi.doe.gov/Olipa1	
*Piriformospora_indica*_DSM_11827_from_MPI	https://mycocosm.jgi.doe.gov/Pirin1	
*Rhizoctonia_solani*_AG-1_IB	https://mycocosm.jgi.doe.gov/Rhiso1	
*Sebacina_vermifera*_MAFF_305830_v1.0	https://mycocosm.jgi.doe.gov/Sebve1	
*Serendipita_sp.*_407_v1.0	https://mycocosm.jgi.doe.gov/Serend1	
*Thanatephorus_cucumeris*_MPI-SDFR-AT-0096_v1.0	https://mycocosm.jgi.doe.gov/Thacu1	
*Tulasnella_calospora*_AL13_4D_v1.0	https://mycocosm.jgi.doe.gov/Tulca1	

Table 3. List of taxa and data availability.

Asterisks*indicate isolates selected for reference genome sequencing.

SampleID	CONTIG COUNT	TOTAL LENGTH	MIN	MAX	MEDIAN	MEAN	L50	N50	L90	N90
Ceratobasidium_sp_UAMH11750.Orchid	7239	47766782	2000	131072	4320	6598.53	1452	8784	5266	2947
Ceratobasidium_sp_370.Orchid	8028	47284605	2000	145775	3838	5889.96	1613	7304	5991	2715
Ceratobasidium_sp_392.Orchid	8904	52465834	1500	100768	3806	5892.39	1693	8316	6228	2516
Ceratobasidium_sp_394.Orchid	8562	53454716	1500	94250	3864	6243.25	1547	9440	5859	2601
Ceratobasidium_sp_395.Orchid	8769	67010313	1500	125675	4575	7641.73	1478	12298	5716	3127
Ceratobasidium_sp_414.Orchid	4161	50425407	1500	342430	4339	12118.58	349	34639	2143	4179
Ceratobasidium_sp_423.Orchid	15380	66434938	1500	107431	2946	4319.57	3306	5259	11578	2042
Ceratobasidium_sp_428.Orchid	8999	69097995	1500	121965	4552	7678.41	1493	12317	5876	3139
Serendipita_sp_396.Orchid	1431	20638744	2072	267195	9279	14422.6	279	17663	1083	6736
Serendipita_sp_397.Orchid	4253	28775535	1500	270096	3582	6765.94	580	10829	2793	2618
Serendipita_sp_398.Orchid	4302	28851264	1500	267525	3546	6706.48	574	11054	2835	2583
Serendipita_sp_399.Orchid	5013	31004825	1500	127831	3844	6184.88	906	8927	3450	2622
Serendipita_sp_400.Orchid	4392	28560853	1502	143991	3564	6502.93	635	10584	2927	2562
Serendipita_sp_401.Orchid	3823	28571286	1500	662216	3626	7473.52	429	13497	2433	2804
Serendipita_sp_405.Orchid	4254	28724145	1500	280457	3539	6752.27	566	10892	2796	2594
Serendipita_sp_407.Orchid	4028	27230538	1500	297010	3154	6760.31	426	12923	2590	2442
Serendipita_sp_411.Orchid	4211	28400685	1500	296955	3574	6744.4	574	10730	2767	2605
Tulasnella_sp_330.Orchid	1013	42302809	1500	967722	8117	41759.93	66	175815	323	19296
Tulasnella_sp_331.Orchid	3446	44512311	1500	298887	4965	12917.1	311	35916	1764	4762
Tulasnella_sp_332.Orchid	3329	44576393	1501	416704	5260	13390.33	297	36313	1699	5096

Tulasnella_sp_403.Orchid	2963	29964626	1502	501591	4201	10112.93	303	23340	1647	3550
Tulasnella_sp_408.Orchid	10591	63626598	1500	92931	3595	6007.61	1850	9042	7284	2490
Tulasnella_sp_417.Orchid	9866	61487333	1500	224139	3694	6232.25	1709	9469	6703	2528
Tulasnella_sp_418.Orchid	3880	32841821	1501	268652	3496	8464.39	411	18771	2277	2844
Tulasnella_sp_419.Orchid	3865	33665047	1500	229676	3923	8710.23	419	18466	2308	3120
Tulasnella_sp_424.Orchid	781	48431701	1507	1054277	13975	62012.42	71	195410	275	36329
Tulasnella_sp_425.Orchid	769	48399368	1507	1089567	16223	62938.06	74	189857	290	37534
Tulasnella_sp_427.Orchid	6018	39289859	1500	134256	3802	6528.72	956	9789	4051	2663
Tulasnella_calospora_UAMH9824.Orchid	4164	49802335	1500	345740	5668	11960.21	481	25557	2334	4747
Tulasnella_inquilina_235.Orchid	1742	44191844	1503	570014	3948	25368.45	119	102565	508	11931
Tulasnella_inquilina_238.Orchid	2874	45488573	1500	444022	4355	15827.62	240	54071	1242	5477
Tulasnella_inquilina_UAMH7632.Orchid	2898	46174977	1501	520136	4010	15933.39	209	56439	1211	5431

Table 4. Assembly statistics.

SampleID	RNASeq	Gene Count	BUSCO Complete %	BUSCO Single	BUSCO Fragmented	BUSCO Missing	BUSCO # Genes
Ceratobasidium_sp_UAMH11750.Orchid	Cerato379	16971	65.8	60.3	11.6	22.6	1335
Ceratobasidium_sp_370.Orchid	CeratoAll	11343	78.4	77.7	8.5	13.1	1335
Ceratobasidium_sp_392.Orchid	CeratoAll	14816	72.3	66.3	13.4	14.3	1335
Ceratobasidium_sp_394.Orchid	Cerato394	18818	71.1	62.6	13.6	15.3	1335
Ceratobasidium_sp_395.Orchid	Cerato395	19777	71.1	42.9	13.2	15.7	1335
Ceratobasidium_sp_414.Orchid	CeratoAll	12172	96.6	95.5	1.6	1.8	1335
Ceratobasidium_sp_423.Orchid	CeratoAll	13213	77.6	76.6	11.7	10.7	1335
Ceratobasidium_sp_428.Orchid	CeratoAll	25061	71.4	39.4	11.9	16.7	1335
Serendipita_sp_396.Orchid	Serend400	8272	65.5	64.9	5.8	28.7	1335
Serendipita_sp_397.Orchid	Serend400	12078	74.7	73.8	11.9	13.4	1335
Serendipita_sp_398.Orchid	Serend400	12311	74.6	73.6	11.4	14	1335
Serendipita_sp_399.Orchid	Serend400	11252	72.2	64.6	11.6	16.2	1335
Serendipita_sp_400.Orchid	Serend400	12369	72	70.9	11.3	16.7	1335
Serendipita_sp_401.Orchid	Serend400	11951	76.4	75.4	10	13.6	1335
Serendipita_sp_405.Orchid	Serend400	11992	73.3	72.6	10.7	16	1335
Serendipita_sp_407.Orchid	Serend400	11442	69	67.8	13.3	17.7	1335
Serendipita_sp_411.Orchid	Serend400	11996	74.5	73.2	10.6	14.9	1335
Tulasnella_sp_330.Orchid		9146	95.4	94.5	1.8	2.8	1335
Tulasnella_sp_331.Orchid		9039	92.9	91.7	2.8	4.3	1335
Tulasnella_sp_332.Orchid		9025	93	91.7	2.7	4.3	1335

Tulasnella_sp_403.Orchid	Tulinq7632	8272	77.7	77.2	10.3	12	1335
Tulasnella_sp_408.Orchid	Tulinq7632	15407	54.2	48.4	17.5	28.3	1335
Tulasnella_sp_417.Orchid		13362	62.4	56.7	15.3	22.3	1335
Tulasnella_sp_418.Orchid	Tulasn419	12415	86.3	85.5	6.1	7.6	1335
Tulasnella_sp_419.Orchid	Tulasn419	12897	88	87.3	6	6	1335
Tulasnella_sp_424.Orchid		11832	95.6	94.3	1.6	2.8	1335
Tulasnella_sp_425.Orchid		10834	96	94.4	1.4	2.6	1335
Tulasnella_sp_427.Orchid		11876	71.4	62	14.6	14	1335
Tulasnella_calospora_UAMH9824.Orchid	Tulinq7632	12307	88.9	87.3	5.4	5.7	1335
Tulasnella_inquilina_235.Orchid	Tulinq7632	13948	95.4	94.3	1.9	2.7	1335
Tulasnella_inquilina_238.Orchid	Tulinq7632	14664	94.1	92.5	2.6	3.3	1335
Tulasnella_inquilina_UAMH7632.Orchid	Tulinq7632	14741	94.2	92.5	2.3	3.5	1335

Table 5. Annotation and BUSCO completeness metrics.
Taxa without an RNA sequence listed did not sufficiently map to the Tulinq7632 RNA sequences and were annotated without expression data. The colors in the BUSCO complete % column range from blue-green (lowest percentage) to dark red (highest percentage).

Sample ID	Number best hit genes (429 total)
Ceratobasidium_sp_UAMH11750.Orchid	354
Ceratobasidium_sp_370.Orchid	356
Ceratobasidium_sp_392.Orchid	368
Ceratobasidium_sp_394.Orchid	376
Ceratobasidium_sp_395.Orchid	369
Ceratobasidium_sp_414.Orchid	387
Ceratobasidium_sp_423.Orchid	376
Ceratobasidium_sp_428.Orchid	382
Serendipita_sp_396.Orchid	311
Serendipita_sp_397.Orchid	391
Serendipita_sp_398.Orchid	371
Serendipita_sp_399.Orchid	358
Serendipita_sp_400.Orchid	375
Serendipita_sp_401.Orchid	382
Serendipita_sp_405.Orchid	379
Serendipita_sp_407.Orchid	364
Serendipita_sp_411.Orchid	382
Tulasnella_sp_330.Orchid	401
Tulasnella_sp_331.Orchid	401
Tulasnella_sp_332.Orchid	398
Tulasnella_sp_403.Orchid	376
Tulasnella_sp_408.Orchid	291
Tulasnella_sp_417.Orchid	330
Tulasnella_sp_418.Orchid	400
Tulasnella_sp_419.Orchid	408
Tulasnella_sp_424.Orchid	408
Tulasnella_sp_425.Orchid	403
Tulasnella_sp_427.Orchid	376
Tulasnella_calospora_UAMH9824.Orchid	398
Tulasnella_inquilina_235.Orchid	416
Tulasnella_inquilina_238.Orchid	410
Tulasnella_inquilina_UAMH7632.Orchid	417
Aporpium_caryae_L-13461.	423
Auricularia_subglabra_v2.0	425
Botryobasidium_botryosum_v1.0	425
Calocera_cornea_v1.0	410
Cantharellus_anzutake_C23_v1.0	412
Ceratobasidium_sp_AGI_v1.0	422

Clavulina_sp._PMI_390_v1.0	423
Exidia_glandulosa_v1.0	422
Hydnum_rufescens_UP504_v2.0	413
Kockovaella_imperatae_NRRL_Y-17943_v1.0	408
Oliveonia_pauxilla_MPI-PUGE-AT-0066_v1.0	419
Piriformospora_indica_DSM_11827_from_MPI	417
Rhizoctonia_solani_AG-1_IB	410
Sebacina_vermifera_MAFF_305830_v1.0	420
Serendipita_sp._407_v1.0	415
Thanatephorus_cucumeris_MPI-SDFR-AT-0096_v1.0	424
Tulasnella_calospora_AL13_4D_v1.0	405
Tulasnella_calospora_UAMH9824_v1.0	426

Table 6. Matrix Occupany.

Stajich lab CTAB Protocol:

1. Add Lysis Buffer (650 µL Buffer A, 650 µL Buffer B, 260 µL Buffer C, 175 µL .1% PVP, 10 µL Proteinase K) to 2 mL microcentrifuge tube, mix, and split equally into two 2 mL tubes.
2. Place in hot plate and heat to 65º C.
3. Grind young fungal tissue in liquid nitrogen, add 50-100 mg of tissue to each tube.
4. Incubate 30 min at 65º mixing by inversion frequently (2-5 min).
5. Add 280 µL KAc to each tube, mix by inversion, incubate on ice for 5 min.
6. Add 500-700 (the more the better) µL PCI, mix by inversion (>5 min) or vortex briefly then incubate for 2 min at room temp (RT).
7. Spin at 6,000 g for 10 min
8. Take supernatant, add equal volume CI (usually about 1000ul).
9. Mix by inversion (>5 min) then incubate for 2 min.
10. Spin at 6,000 g for 10 min
11. Take supernatant (usually 700uL):
    a. RNAse treatment (2.5 µL RNAse, 37º, 90-120 min)*
    b. Optional additional CI washes
12. Add 1/10 vol NaAc, mix, add 1 vol Isopropanol.
13. Incubate at RT 5 min, should start to see lots of DNA threads.
14. Spin at 3,000 g for 2 min, pour out the supernatant.
15. Wash with 1 mL freshly prepared, cold 70% ethanol.
16. Spin at 3,000 g for 2 min, pipette out the EtOH. Remove as much EtOH as possible before drying.
17. Dry pellet at RT for 10-15 min and/or 65º for <2 min to dry any leftover ethanol
    a. Resuspend in 50-100 µL TE (adjusted to pH9) at 65º Optional CI wash (add 600-800 TE buffer at 65°, resuspend DNA, add equal volume CI, mix as directed in step 9, carry on protocol from there minus the RNAse and CI steps, I usually take 500-600 supernatant if added 800 uL CI).
18. Nanodrop, 260/280 is indicative of nucleic acid and 260/230 indicative of protein
19. Qubit
20. Run on Gel
21. Check ITS and 16S by PCR

Supplemental Figure S1. CTAB DNA extraction protocol from Stajich lab