environmental microbiology reports

Environmental Microbiology Reports (2018) **00**(00), 00–00 doi:10.1111/1758-2229.12612

The genome and microbiome of a dikaryotic fungus (*Inocybe terrigena*, Inocybaceae) revealed by metagenomics

Mohammad Bahram,[1,2*†] Dan Vanderpool,[3] Mari Pent,[2] Markus Hiltunen[1] and Martin Ryberg[1]

[1]*Department of Organismal Biology, Evolutionary Biology Centre Uppsala University, Norbyvägen 18D, Uppsala, 75236 Sweden.*

[2]*Department of Botany, Institute of Ecology and Earth Sciences, University of Tartu, 40 Lai St, Tartu, 51005 Estonia.*

[3]*Division of Biological Sciences, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA.*

Summary

Recent advances in molecular methods have increased our understanding of various fungal symbioses. However, little is known about genomic and microbiome features of most uncultured symbiotic fungal clades. Here, we analysed the genome and microbiome of Inocybaceae (Agaricales, Basidiomycota), a largely uncultured ectomycorrhizal clade known to form symbiotic associations with a wide variety of plant species. We used metagenomic sequencing and assembly of dikaryotic fruiting-body tissues from *Inocybe terrigena* (Fr.) Kuyper, to classify fungal and bacterial genomic sequences, and obtained a nearly complete fungal genome containing 93% of core eukaryotic genes. Comparative genomics reveals that *I. terrigena* is more similar to ectomycorrhizal and brown rot fungi than to white rot fungi. The reduction in lignin degradation capacity has been independent from and significantly faster than in closely related ectomycorrhizal clades supporting that ectomycorrhizal symbiosis evolved independently in *Inocybe*. The microbiome of *I. terrigena* fruiting-bodies includes bacteria with known symbiotic functions in other fungal and non-fungal host environments, suggesting potential symbiotic functions of these bacteria in fungal tissues regardless of habitat conditions. Our study demonstrates the usefulness of direct metagenomics analysis of fruiting-body tissues for characterizing fungal genomes and microbiome.

Received 13 November, 2017; revised 19 December, 2017; accepted 4 January, 2018. *For correspondence. E-mail bahram@ut.ee; Tel. +372 737 6222; Fax +372 737 6222. [†]Present address: Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

Introduction

Ectomycorrhizal fungi constitute a major component of fungal communities in terrestrial ecosystems, functioning to facilitate nutrient uptake and carbon storage by plants (Smith and Read, 2010). Evidence from phylogenetic studies, particularly ribosomal RNA genes, suggests that ectomycorrhizal fungi have evolved independently in several fungal lineages (Hibbett *et al*., 2000; Tedersoo *et al*., 2010; Veldre *et al*., 2013) but the genomic adaptations associated with these functional shifts are only beginning to be characterized (Kohler *et al*., 2015).

Recent genomic studies focusing on fungi and using high-throughput sequencing (HTS) technologies have improved our understanding of the evolution of mycorrhizal symbiosis. These studies provide evidence that while mycorrhizal fungi have retained much of their enzymatic capacity to release nutrients from complex organic compounds, certain carbohydrate active enzyme (CAZyme) families have contracted during the evolution of mycorrhizal fungi compared to their saprotrophic ancestors (Martin *et al*., 2010; Kohler *et al*., 2015). In addition, metabolite pathways and secondary metabolites vary between obligate biotrophs and saprotrophs. This adaptation reflects a shift in life history, contributing to variation in genome size, characteristic reductions in gene families such as transporters and plant cell wall degradation enzymes (Martin *et al*., 2010; Spanu *et al*., 2010; Floudas *et al*., 2012; Kohler *et al*., 2015). Ectomycorrhizal fungi possess a repertoire of genes encoding cellulose degrading enzymes to help release simple organic compounds that are available for uptake by plants, which substantially accelerates soil nutrient cycling (Martin *et al*., 2010). So far, most fungal genome sequencing studies have employed culture-dependent techniques, which omit major ectomycorrhizal lineages that are difficult to culture.

Similar to animals and plants, fungal tissues can harbour a diverse array of prokaryote associates. In root and soil fungi, bacteria may contribute to the formation

and regulation of mycorrhizal associations (Torres-Cortés *et al*., 2015). A few studies on fungal fruiting-bodies suggest that bacteria may have important ecological roles in fungal spore dispersal (Splivallo *et al*., 2015), gene expression (Riedlinger *et al*., 2006; Deveau *et al*., 2007) and mycotoxin production (Lackner *et al*., 2009). In addition, bacteria may affect fungal growth (Chen *et al*., 2013) and mycorrhization (Frey-Klett *et al*., 2007; Aspray *et al*., 2013), yet we know little about the associated bacterial taxa and functions in epigeous fruiting-bodies.

We performed metagenomic sequencing on DNA extracted from fruiting-body tissues to characterize the genome and microbiome of a dikaryotic fungus *Inocybe terrigena* (Inocybaceae, Agaricales). Inocybaceae is a diverse ectomycorrhizal fungal lineage that is thought to have evolved independently from other ectomycorrhizal lineages (Matheny *et al*., 2009). Nevertheless, there is no published work on the evolution of mycorrhizal status in Inocybaceae using comparative genomics. We compared the genome of *I. terrigena* with 58 other genomes from the Agaricomycetidae to look for significant expansion or contraction in CAZyme gene families across multiple, independent evolutionary transitions to mycorrhizal symbiosis (Matheny *et al*., 2009). Additionally, we characterized the taxonomic composition and potential functions of the associated bacteria in fruiting-body tissues.

Results

Genomic features of Inocybe terrigena

After quality filtering and trimming, 29,244,774 paired end reads (96%) were used for assembly. Detailed descriptive statistics for sequencing and assembly are available in Supporting Information Table S1. BLAST searches were performed with the resulting 31 471 assembled contigs against reference genomes resulted in 2262 contigs (N50: 30 623, total length: 26 127 495 bp) and 11 596 contigs (N50: 8777, total length 34 853 629 bp) initially identified as fungal and bacterial respectively. Mitochondrial contigs were determined based on coverage ($>$1200X) and GC content ($<$ 0.40), resulting in 5 contigs, with a combined length of 67 kb. In total, 85.8% of non mitochondrial fungal sequences had a GC content between 0.4 and 0.5 (Supporting Information Table S2). Most bacterial contigs are outside this interval (97.2%). Among non-fungal genera, *Pedobacter*, *Pseudomonas* and *Burkholderia* had the highest coverages of 40.24 ± 20.64, 5.73 ± 3.24 and 4.23 ± 1.99 respectively (Supporting Information Fig. S2). In total, 56.5% and 10.6% of total bacterial assembly length belonged to *Pseudomonas* and *Pedobacter*, with GC content of 0.608 ± 0.032 and 0.396 ± 0.039 respectively (Fig. 1; Supporting Information Fig. S2). The remaining contigs of bacterial origin had GC content of 0.500 ± 0.115 and coverage of 18.86 ± 108.85 (Fig. 1; Supporting Information Fig. S2). Based on this we classified previously unclassified contigs with median coverage of $>$1 and GC content $>$ 0.35 and $<$ 0.56 (Fig. 1) as fungal, but made no further attempts classifying contigs as bacteria. The final assembly included 5409 fungal contigs (N50: 19 600, total length: 37 436 744 bp) and 12,040 bacterial contigs (N50: 9 606, total length 39 035 878 bp; Supporting Information Table S2).

The MGRAST analysis revealed that 77.42% of contigs were of bacterial origin followed by 22.47% fungal, 0.04% viral and 0.03% archaeal contigs. Mapping all reads to the assembly resulted in a total of 28 632 347 mapped reads (93% of all reads), with a median coverage of 35X in the fungal contigs (22 274 636 reads; 78% of mapped reads) and 5X for bacterial contigs (5 072 696 reads; 18%). Variant calling revealed 3.4 variant sites per kb in the fungal contigs.

BLAST analyses of extracted ITS sequences against UNITE (Kõljalg *et al*., 2013) resulted in one ITS sequence matching *Inocybe terrigena* (accession number: JF908091; identity = 99%, *e*-value = 1 e–137, median coverage = 2335, contig length = 8222), and five additional ITS sequences matching other Eukaryotes. All of the additional eukaryotic contigs including ITS had very low coverage: *Hypomyces odoratus* (identity = 100%, *e*-value = 1 e −86, median coverage = 2, contig length = 745), *Selaginella deflexa* (identity = 98%, *e*-value = 6 e −102, median coverage = 2, contig length = 1175), *Drosophila subauraria* (identity = 84%, *e*-value = 3 e −23, median coverage = 1, contig length = 438), *Cucumis melo* (identity = 100%, *e*-value = 3 e −04, median coverage = 3, contig length = 1041). These results indicate a negligible contribution of other Eukaryotes to the resulting *I. terrigena* assembly.

The Maker2 gene annotation of fungal ($n = 2261$), unclassified (contigs matching neither fungi nor bacteria; $n = 17196$) and ambivalent (contigs matching both fungi and bacteria; $n = 417$) contigs resulted in 11 918 fungal genes identified (coverage = 50.58 ± 142.49, median $\pm$ SD; GC content = 0.467 ± 0.0295). The Prodigal annotation of bacterial contigs resulted in 63 328 genes. Of these, 3289 additional fungal genes were identified based on BLAST searches (median coverage = 57.33 ± 152.85; GC content = 0.471 ± 0.043), indicating some fungal contigs were misidentified as bacterial contigs. Thus, we confirmed the identification of all classified contigs based on BLAST searches using annotated genes against the NCBI protein database, as implemented in the pipeline (Supporting Information Table S2; Supporting Information 1). Based on genes annotated by Prodigal and genes identified from unclassified contigs using Maker2 ($n = 8190$, median coverage= 10.70 ± 11.77; GC content = 0.599 ± 0.065), 62 023 genes were identified as bacterial (11 188 and 50 835 by Maker2 and Prodigal respectively; median

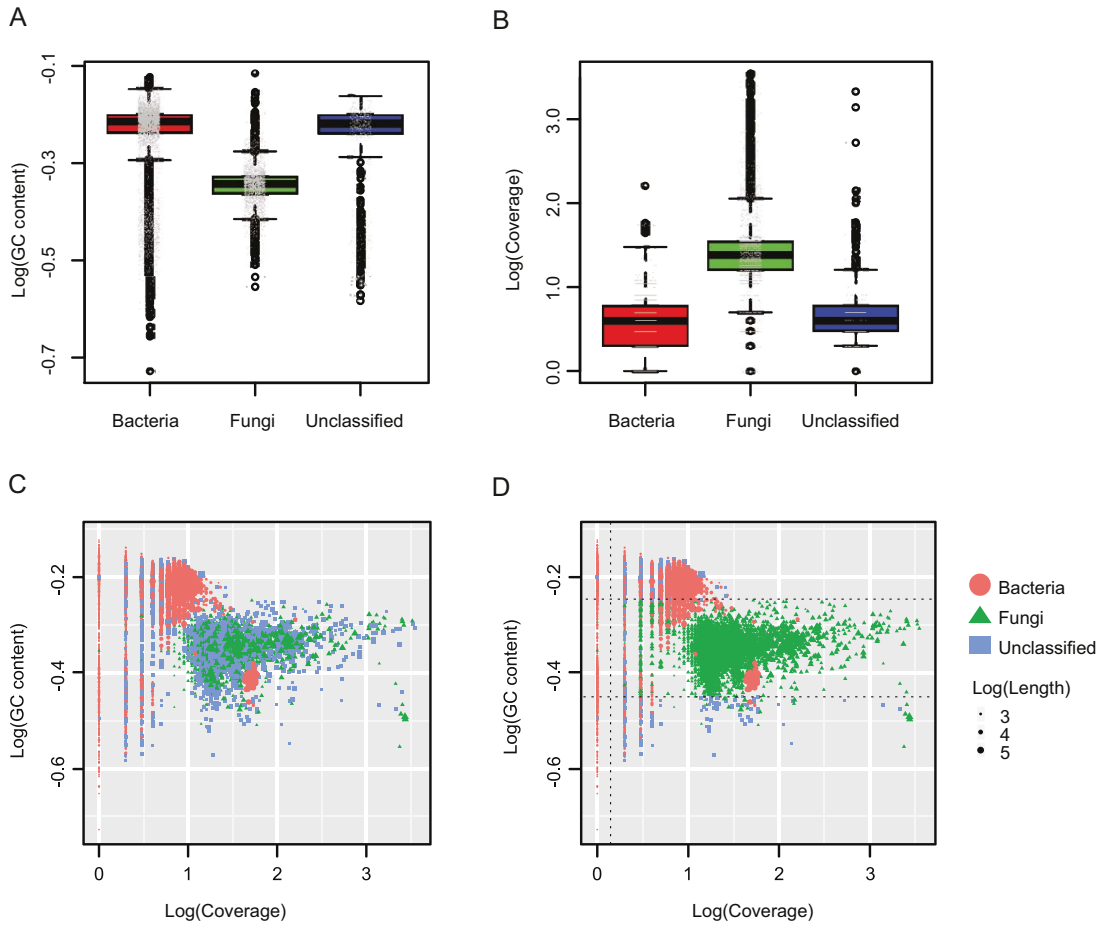

Fig. 1. Bacterial and fungal assemblies can be separated based on GC content and coverage.
A–B. The boxplots of GC content (A) and median coverage (B) for bacterial, fungal contigs identified based on Blast searches.
C. The scatterplot of GC content as a function of median coverage of contigs.
D. Same as C, however, unclassified contigs with median coverage > 1 and GC content > 0.35 and < 0.56 were classified as fungi. In C and D, unclassified contigs of < 1000 bp length have been excluded.

coverage = 10.07 ± 13.24; GC content = 0.588 ± 0.080; Supporting Information Table S3). After this additional round of filtering, 2855 genes remain unclassified (median coverage = 6.36 ± 37.47; GC content = 0.598 ± 0.073).

In total, the *I. terrigena* genome included 15 207 genes, which is lower than those reported for most ectomycorrhizal fungi (Supporting Information Fig. S1; Supporting Information Table S5). Of 15 207 genes, 24.6% genes were with Pfam domains based on InterPro domain assignments. Using KEGGmapper, 4467 out of 15 207 fungal genes were functionally annotated (29.4%; Supporting Information Fig. S3 and Table S3). Using CAZY pipeline, we identified 396 Pfam domains from CAZy Families, including 46 auxiliary activities (AA, 8 families), 33 carbohydrate-binding modules (CBM, 11 families), 47 carbohydrate esterases (CE, 7 families), 168 glycoside hydrolases (GH, 31 families), 94 glycosyltransferases (GT, 34 families) and 8 polysaccharide lyases (PL, 4 families).

The CEGMA analysis indicated 91% of 242 full-length, core eukaryotic genes were present in our fungal assembly and 93% were present when partial and full-length alignments were considered. (Supporting Information Table S4).

Comparative genomics

OrthoMCL clustering of 59 Agaricomycetes whole proteomes resulted in 870 single-copy genes present in >75% of taxa, and a concatenated alignment of 148 316 amino acid positions. The resulting Maximum Likelihood phylogenetic tree (Fig. 2) is largely consistent with previous phylogenies (e.g., Kohler *et al.*, 2015) and the ASTRAL phylogeny (Supporting Information 2) had well supported parts. Our CAFE analysis of CAZYmes indicated that 6 of ∼ 220 CAZY families found were significantly expanded or contracted in the 59 genomes analysed here (Fig. 3; Supporting Information Table S6). Cluster analysis revealed two major groups of fungi based on CAZyme profiles, and it placed *I. terrigena* in the same

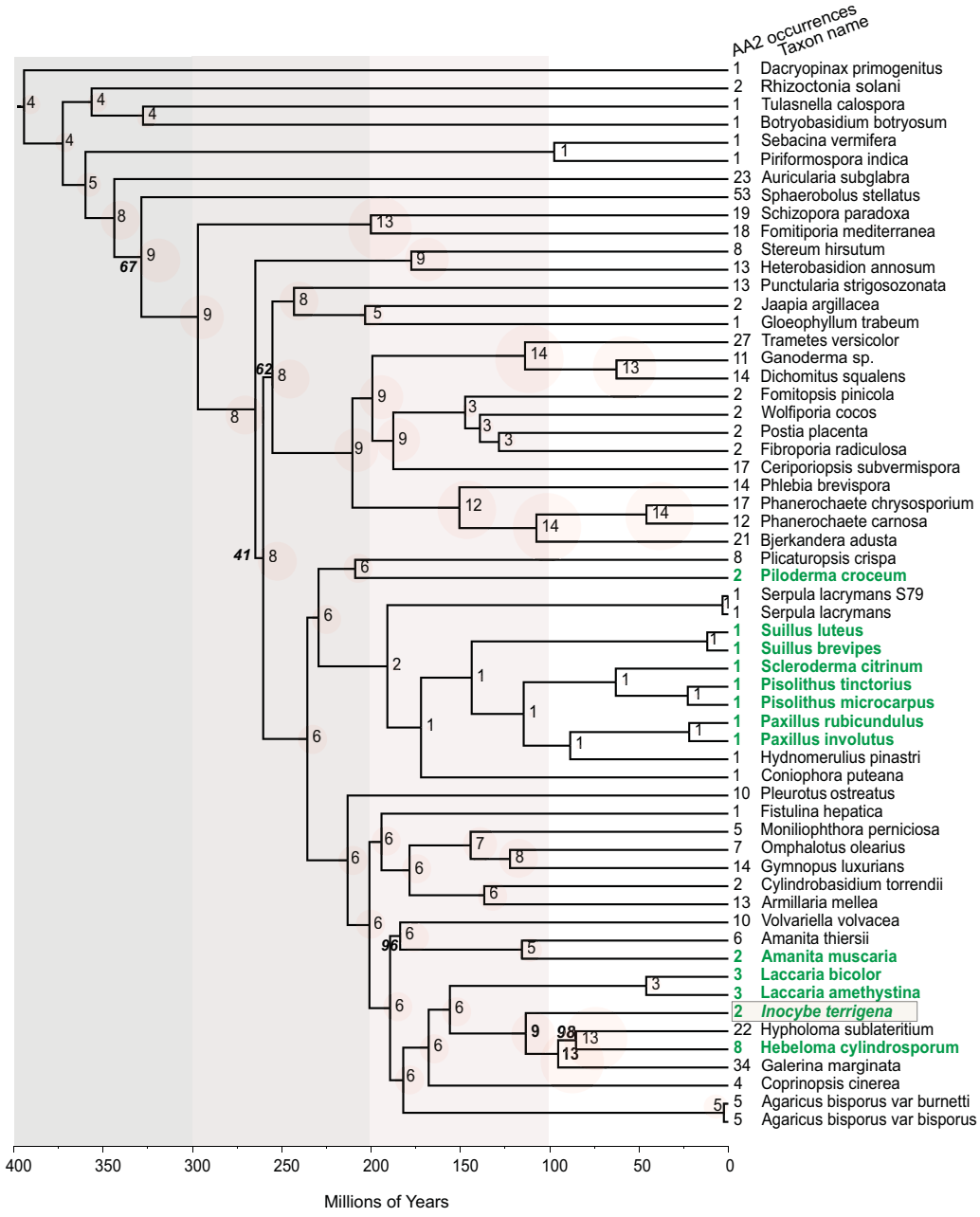

Fig. 2. Phylogenetic tree showing the phylogenetic relationship of *Inocybe terrigena* and other published ectomycorrhizal and saprotrophic fungi. All clades have a support value of 100% except those that are indicted in italic text. Non-italic node labels and numbers next to taxon names represent the number of AA2 gene occurrences for each parent node and taxon, indicating gain/loss of AA2 genes in each taxon. Note that *I. terrigena* belongs to a clade which has gained three AA2 genes from its parent node, but it has lost seven AA2 genes compared to other members of this clade. Numbers on the tree nodes and edges are estimated (based on CAFE) and observed copy numbers of AA2. Ectomycorrhizal fungi are indicated with green text.

cluster as other ectomycorrhizal taxa (Fig. 2). Compared to other genomes included in our comparative genomics analysis, *I. terrigena* has no significantly expanded CAZyme families; however, it is significantly reduced in an important lignin degrading CAZyme family (AA2), which contains class II lignin-modifying peroxidases. While *I. terrigena* belongs to a clade which has gained 3 AA2 from its parent node, it has lost 7 AA2 compared to other members of this clade.

Microbiome of Inocybe terrigena

Proteobacteria and Bacteroidetes comprised 80.9% and 16.6% of the taxonomic composition in bacterial

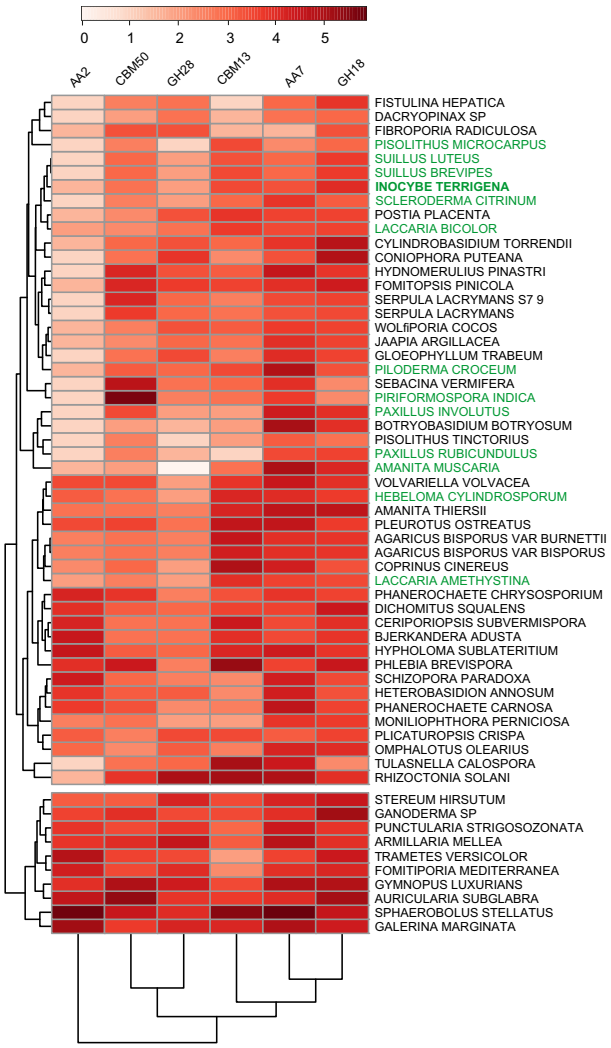

Fig. 3. Heatmap of six CAZyme families that showed significant expansion or contraction across 59 analysed genomes in this study. The scale shows the copy numbers (Log) of CAZyme genes in each family. Ectomycorrhizal fungi are indicated with green text.

metagenomics data (Fig. 4A). Gammaproteobacteria (49.9%), Betaproteobacteria (26.7%), Flavobacteria (10.3%), Sphingobacteria (4.6%) and Alphaproteobacteria (3.7%) were the dominant bacterial classes (Fig. 4B). *Pseudomonas*, *Chryseobacterium,* *Herbaspirillum*, *Burkholderia* and *Pedobacter* were the dominant genera (Fig. 4D). In total, the *I. terrigena* metagenome included 63 bacterial species based on MG-RAST pipeline, with three dominant species *Pseudomonas fluorescens*, *P. syringae* and *Herbaspirillum seropedicae* contributing 61% to the total bacterial genes. The bacterial contigs contained 28 288 predicted proteins that mostly matched genes with known functions including *metabolism* (55.3%), *information storage and processing* (22.1%), and *cellular processes and signalling* (8.2%) (Supporting Information Fig. S4).

Discussion

Genomic features of Inocybe terrigena

Our study is the first to analyse the genome and microbiome of a dikaryotic fungus. Using one tenth of an Illumina lane, we obtained a genome with a comparable completeness with those reported in previous studies (Kohler *et al*., 2015; Quandt *et al*., 2015), indicating the acceptable performance of this approach to investigate the genomes of unculturable fungi based on epigeous fruiting-body tissues. Such analysis is facilitated by significant differences in length, GC content and median coverage between bacterial and fungal genomes, which enable classification of assemblies into fungal or bacterial origin. The GC content of the *I. terrigena* genome (0.47) is in the same range as other ectomycorrhizal species in the Agaricales such as *Laccaria bicolor* (0.47) and *L. amethystina* (0.47), *Hypholoma sublateritum* (0.51), *Galerina marginata* (0.48) and *Hebeloma cylindrosporum* (0.48). The GC content of all major bacterial groups except three Bacteroidetes (comprising 5.4% of contigs) were outside this range (Supporting Information Table S2). Together, the contrasting GC content and coverage can be used to accurately remove possible contamination (Laetsch and Blaxter, 2017) and separate bacterial and fungal contigs, which can reduce the computation cost and database biases, particularly when specific genes are targeted.

Our analysis revealed that while *I. terrigena* has a smaller proteome based on the number of gene models compared to the previously sequenced ectomycorrhizal Basidiomycetes (Supporting Information Fig. S1 and Table S5, Martin *et al*., 2010; Kohler *et al*., 2015), it has a more similar CAZyme profile to those of brown rot and ectomycorrhizal than to white rot lineages (Figs. 2, 3). Given the significant reduction in certain CAZyme families as a common pattern in genome evolution of ectomycorrhizal fungi (Kohler *et al*., 2015), the increase in the AA2 family in the lineage leading to the most recent common ancestor to *Inocybe* and its sister clade indicates that this line was not ectomycorrhizal; therefore, the significant reduction in AA2 in the lineage leading to *Inocybe* supports a separate origin of the ectomycorrhizal nutritional mode in this lineage (Matheny *et al*., 2006).

Structure and function of associated bacteria of Inocybe terrigena

Proteobacteria and Bacteroidetes were the most abundant phyla in *I. terrigena*, constituting the largest fraction of bacterial assembly (Supporting Information Fig. S3). Similarly to our study, Proteobacteria has been identified as a very common bacterial group in several

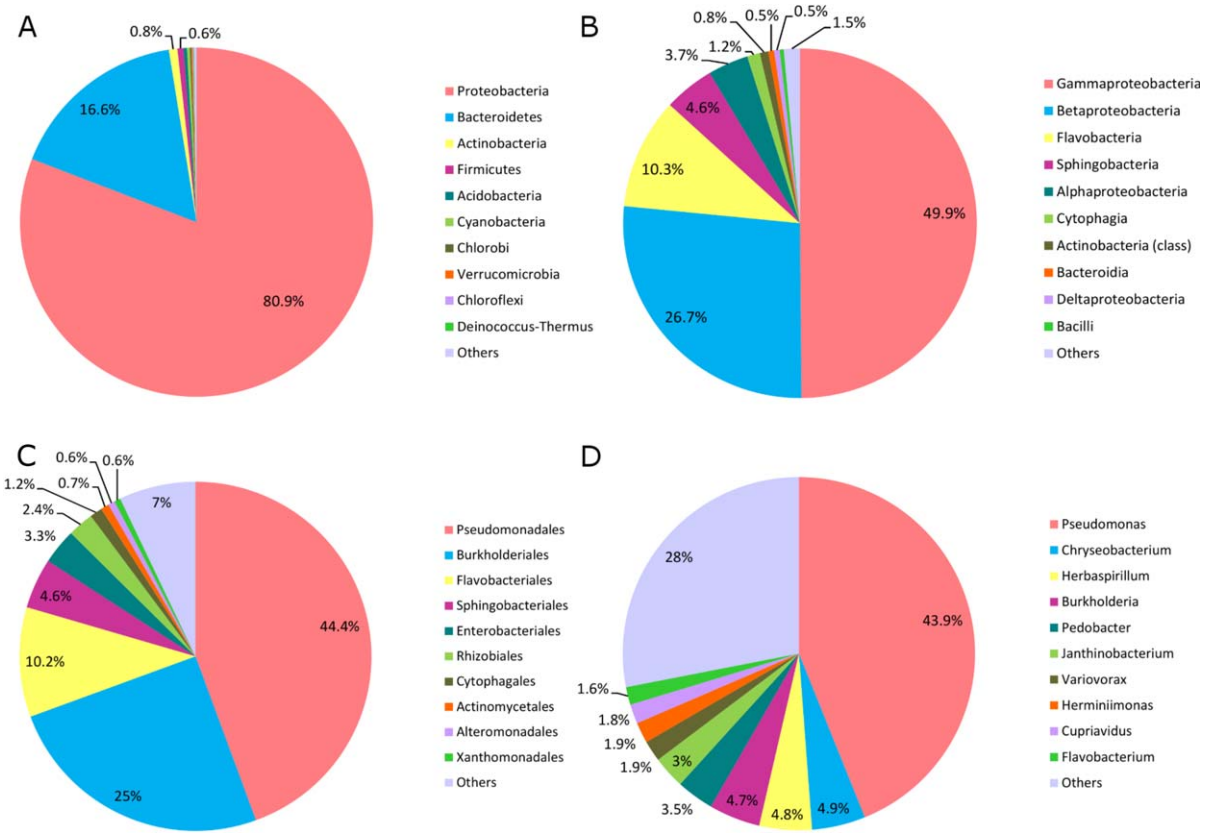

Fig. 4. Pie chart showing the relative abundance of 10 most common bacterial taxa at phylum (A), class (B), order (C) and genus (D) level in *Inocybe terrigena* fruitbody based on representative hits of RefSeq database (at *e*-value $< 1 \times 10^{-5}$ and % identity > 60) using MG-RAST. Bacterial taxa with abundance $\geq 0.5\%$ are presented. All fungal, bacterial and unclassified contigs were included.

ascomycetous (Barbieri *et al*., 2005; 2007; Quandt *et al*., 2015; Benucci and Bonito, 2016) as well as basidiomycetous (Kumari *et al*., 2013; Pent *et al*., 2017) fruiting-bodies. There is also evidence that the relative abundance of Proteobacteria is higher in the mycosphere (Warmink *et al*., 2009) and ectomycorhizosphere (Uroz *et al*., 2012; Antony-Babu *et al*., 2013) compared to the surrounding soil, suggesting the tendency of Proteobacteria for colonizing fungus-related habitats. A high relative abundance of Bacteroidetes, the second-largest bacterial phylum in *I. terrigena*, is also often found in ectomycorhizosphere, mycosphere and fruiting-bodies of ascomycetous as well as basidiomycetous fungi (Uroz *et al*., 2012; Antony-Babu *et al*., 2013; Benucci and Bonito, 2016; Halsey *et al*., 2016; Pent *et al*., 2017). In particular, the relative abundance of Sphingobacteria in *I. terrigena* was similar to that in *Elaphomyces granulatus* (Quandt *et al*., 2015). Similarly to *I. terrigena* fruiting-body Acidobacteria, Actinobacteria, Verrucomicrobia, Firmicutes and Cyanobacteria form a small fraction of the bacterial community in basidiomycetous (Pent *et al*., 2017) as in ascomycetous (Antony-Babu *et al*., 2013) fruiting-bodies, whereas they are highly represented in soil (Eilers *et al*., 2010; Bergmann *et al*., 2011). The observed level of specificity of fungal associated bacteria may be related to close fungal-bacterial interactions or fungal-driven changes in habitat conditions, such as change in pH or nutrient conditions in soil (Danell *et al*., 1993; Nazir *et al*., 2010a,b).

The most abundant bacterial genera associated with *I. terrigena* are known to have symbiotic functions in fungal tissues. Although major bacterial taxa found in *I. terrigena* (Supporting Information Fig. S3) are common in soil (Baldani *et al*., 1986; Janssen, 2006), the genus *Pseudomonas* is also one of the most abundant bacterial groups in basidiomycetous fruiting-bodies and ectomycorrhizas (Deveau *et al*., 2007; Kumari *et al*., 2013). While *Pseudomonas* strains are mainly saprotrophic, they are also known to alleviate detrimental effect of pathogens on plant roots and leaves (Haas and Defago, 2005) and facilitate mycorrhizal establishment (Dominguez *et al*., 2012). Some strains of *P. fluorescens* can promote the growth of mycelium, ascus opening or other morphological changes in fungi (Citterio *et al*., 2001; Cho *et al*., 2003; Deveau *et al*., 2007). Certain *Pseudomonas* species have also been reported as bacterial

'fungiphiles' in the mycospheres, pointing to their close relationship with fungi (Warmink *et al*., 2009). Furthermore, *Pedobacter* has been identified as tolaasin detoxifying bacteria from Agaricales (Tsukamoto *et al*., 2002), and *Chryseobacterium* has been identified in mycospheres of several fungi (Warmink *et al*., 2009). *Herbaspirillum seropedicae* has been reported as nitrogen fixing root associated bacterium (Baldani *et al*., 1986). Another dominant group, *Burkholderia* is also known to have beneficial interactions with fungi improving the formation of mycorrhiza (Aspray *et al*., 2006; Frey-Klett *et al*., 2007) and providing the fungal partner with nutrients in stress conditions (Stopnisek *et al*., 2016). It is possible that these bacteria play similar roles in epigeous fungal fruiting-bodies; however, more replicated studies are needed to understand the specific functions of these bacteria and to exclude the possibility that they passively colonize fungal fruiting-bodies. Taken together, these data indicate the dominance of symbiotic bacteria in fungal epigeous fruiting-bodies.

Despite many similarities, there were some differences between the microbiome of basidiomycetous *I. terrigena* and hypogeous ascomycetous fruiting-bodies (Barbieri *et al*., 2005; 2007; Quandt *et al*., 2015; Benucci and Bonito, 2016). Particularly compared to our dataset where Gamma and Betaproteobacteria were the dominant classes, Alphaproteobacteria and Actinobacteria were the dominant classes in ascomycetous fruiting-bodies. This may be explained by the higher relative abundance of Alphaproteobacteria in soil (Janssen, 2006; Fierer *et al*., 2012; Pent *et al*., 2017) as well as more intimate association between soil and hypogeous fruiting-bodies, compared to epigeous fruiting-bodies. The differences between bacterial communities of epi- and hypogeous fruiting-bodies may also be ascribed to different environmental conditions below- and aboveground. In contrast to the microbiome of hypogeous species, which are typically dominated by *Bradyrhizobium* species (Barbieri *et al*., 2005; 2007; Antony-Babu *et al*., 2013; Quandt *et al*., 2015), both *Pseudomonas and Pedobacter* dominated the *I. terrigena microbiome*. Several potato-associated *Pseudomonas* species are able to counteract both plant-pathogenic fungi and plant-parasitic nematode (Krechel *et al*., 2002), and some *Pedobacter* are associated with soil or plant-pathogenic nematodes (Tian *et al*., 2011; Baquiran *et al*., 2013). Thus, it is tempting to suggest that the high abundance of *Pseudomonas* and *Pedobacter* may reflect their protective functions against pathogens and parasites in epigeous fruiting-bodies.

Comparing the relative abundance of functional gene categories in *I. terrigena* and non-desert soil microbiome (Fierer *et al*., 2012) reveals a similar functional composition between the two distinct environments. This, together with remarkable similarity in their taxonomic composition, suggests that soil microbes act as a major species source for fungal associated bacterial communities (Pent *et al*., 2017). Nonetheless, the high relative abundance of genes functionally related to environmental and genetic information processing (Supporting Information Figs. S2 and S4) may facilitate processing a large amount of information from their host environment, to enhance mycorrhizal colonization and reduce the impact of harmful environmental conditions and pathogens (Frey-Klett *et al*., 2007).

Conclusions

This study demonstrates that with appropriate filtering, metagenomic sequencing of fungal fruiting-body tissues enables near complete genome sequencing from dikaryotic fruiting-body tissues, which is comparable to those from cultured fungal isolates. With further advances in HTS technology, e.g., overcoming length limitation and improving assembly algorithms and hence genome assembly quality, metagenomics will also be useful to study the population genomics of uncultured fungi. In addition, metagenomics enabled us to characterize the associated bacterial taxa and functions in fruiting-body tissues. Certain groups of these bacteria appear to have symbiotic functions with the fungal host.

Experimental procedures

DNA was extracted from a dried collection of *Inocybe terrigena* (field accession no: MR00339; coordinates: N59°0′7.52, E17°42′2.42"; date: 2013-09-20; herbarium: UPS), collected and identified by MR, using Plant mix DNeasy DNA Isolation kit. To obtain DNA of sufficient quantity, four DNA extractions from lamellae of the fruiting-body were pooled. Lamellae are internal spore bearing layered structures in fruiting-bodies and are, therefore, expected to provide higher DNA per unit of material. In addition, as lamellae are protected during development they have a lower risk of contamination than other exposed parts of the fruiting-body. We took extra care to minimize contamination during sample collecting, storing and handling. PCR free, paired-end, 300bp insert libraries were constructed from total DNA and sequenced on 1/10 of an Illumina HiSeq 2500 lane at Sci-Life laboratory (Stockholm, Sweden). Raw Illumina reads, available at the Sequence Read Archive (SRA) under accession number SRP066410, were quality filtered and trimmed using Fastx Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/2) using the following settings: -q 25 -p 90 and -t 25 -l 20 respectively. Genome assembly was performed using Spades de novo (Bankevich *et al*., 2012) with default settings. To sort assembled contigs into fungi and bacteria, we used

BLAST queries to compare contigs to previously published whole genomes of bacteria and fungi downloaded from GenBank and the Joint Genome Institute (JGI; www.jgi.doe.gov) website respectively. To evaluate the accuracy and coverage of the resulting assemblies, we mapped post-filtered reads to the assembled contigs using Bowtie 2 (Langmead and Salzberg, 2012). We used median coverage in combination with GC content to determine whether contigs not classified by BLAST were of bacterial or fungal origin, by comparison to the values of contigs classified by BLAST. Unclassified contigs shorter than 1000 bp were excluded from the analysis, as GC content and median coverage varied greatly for these and taxonomic assignment was deemed unreliable. Bacterial and eukaryotic species present in the final assemblies were identified using (previously constructed) HMMer profiles to recognize and trim ribosomal operons and ITS region respectively (Lagesen *et al*., 2007; Bengtsson-Palme *et al*., 2013). Additionally, we used the MG-RAST pipeline (Meyer *et al*., 2008) to infer both taxonomic and functional annotations based on assemblies.

Protein annotation of fungal and bacterial assemblies was performed using the Maker2.3.36 pipeline (Holt and Yandell, 2011) and Prodigal (Hyatt *et al*., 2010) with default parameters respectively. Core Eukaryotic Mapping Genes Approach (CEGMA) (Parra *et al*., 2009) was used to evaluate genome completeness and to generate preliminary annotations as training sets in Maker2. RepeatModeler (Smit and Hubley, 2011) was used to generate a classified repeat library for the metagenome assembly. This repetitive sequence library was combined with a Maker2 provided a transposable element library and was used in RepeatMasker 3.0 (Smit *et al*., 2010) for masking within the Maker2 pipeline. Additionally, the proteomes of *Laccaria bicolor*, *Coprinopsis cinerea* and the UniProt reference proteomes (The UniProt Consortium, 2014) were used as protein evidence to generate hints for *ab-initio* predictors. Three rounds of training were used for the *ab-initio* gene predictors SNAP and Augustus while self-training was used for GeneMark-ES. Overcalling genes is common for *ab-initio* gene predictors (Larsen and Krogh, 2003) so 'keep_preds = 0' was set in Maker2 to only call gene models which were supported by protein evidence (AED < 1). Predicted proteins lacking functional evidence were scanned for protein family domains (PFAM) using InterProScan5 (IPRscan5) (Jones *et al*., 2014), those containing PFAM domains were promoted to full gene models using scripts included with Maker2. Protein models were assigned to multigene families using the OrthoMCL software (v2.0.9) (Li *et al*., 2003). BLASTX was used to compare and sort bacterial and fungal protein models against the whole NCBI protein database. Analyses were facilitated with a custom perl pipeline, which is available in the Supporting Information. Fungal protein models were compared to non-redundant genomes using BlastKOALA (Kanehisa *et al*., 2016) and eggNOG-mapper (Huerta-Cepas *et al*., 2017) for functional annotation with additional screening for carbohydrate-active enzymes (CAZYmes) using the dbCAN pipeline (Yin *et al*., 2012).

Using the mapped read dataset from Bowtie, duplicate reads were marked using MarkDuplicates function of Picard Tools were realigned around indels using RealignerTargetCreator and IndelRealigner of Genome Analysis Toolkit (McKenna *et al*., 2010). Variant calling was performed on the realigned reads using Platypus (Rimmer *et al*., 2014). The level of heterozygosity was quantified from the filtered variant data set.

Phylogenetic analysis was performed based on complementary supermatrix and supertree approaches with 74 publicly available Agaricomycetes genomes (retrieved from the JGI database). Single-copy genes were identified based on MCL clusters on BLAST e-values with the inflation parameter set to 2.0, in an additional quality check they were clustered using OrthoMCL with only those single copy genes present in 75% of taxa used for phylogenetic analysis. Amino acid sequences were aligned using MAFFT 7 (Katoh and Standley, 2013) with highly variable regions removed using Gblocks (Talavera *et al*., 2007). The resulting concatenated alignment was used to estimate a maximum likelihood phylogeny with 1000 UltraFast (UF) bootstrap replicates (Minh *et al*., 2013) using IQ-Tree 1.5.5 (Nguyen *et al*., 2015). Sequence evolution was modelled using the Posterior Mean Site Frequency (PSMF) model (Wang *et al*., 2017) and 60 mixture classes. PhyloBayes 3.3 (Lartillot *et al*., 2013) with site-specific evolutionary rates modelled using non-parametric infinite mixtures (CAT-GTR) was used to generate a fossil calibrated ultrametric tree from our whole genome concatenated tree. A diffuse gamma prior with the mean equal to the standard deviation was used for the root of the tree. A log-normal auto-correlated relaxed clock model (Thorne *et al*., 1998) with a Dirichlet prior on divergence times was used for dating the tree. Two fossils were employed to calibrate node times to geological time. The first corresponds to a minimum age of 90My for the Agaricales (Hibbett *et al*., 1997). The second corresponds to a minimum age of 360My (Stubblefield *et al*., 1985) for the Basidiomycota. Trees for individual genes were constructed using the PROTGAMMAWAG model and 100 bootstraps in RAxML 8.2.4 (Stamatakis, 2014). The gene trees were used to estimate to species trees using ASTRAL; support values were calculated using the bootstrapped trees (Mirarab and Warnow, 2015). Significant expansion or contraction of CAZyme gene families was determined

for a reduced phylogeny of the 59 taxa belonging to Agaricomycotina using CAFE 3.1 (Han *et al*., 2013).

Acknowledgements

We thank Sandra Lorena Ament-Velásquez and Fabien Burki for advice on bioinformatics, and three anonymous reviewers for constructive comments on earlier version of this article. The authors would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure. Funding was provided by Uppsala University (Department of Organismal Biology), Kaptain Carl Stenholm and Estonian Science Foundation Agency (PUT1317).

References

Antony-Babu, S., Deveau, A., Nostrand, J.D.V., Zhou, J., Le Tacon, F., Robin, C., *et al*. (2014) Black truffle-associated bacterial communities during the development and maturation of *Tuber melanosporum* ascocarps and putative functional roles. *Environ Microbiol* **16**: 2831–2847.

Aspray, T.J., Jones, E.E., Whipps, J.M., and Bending, G.D. (2006) Importance of mycorrhization helper bacteria cell density and metabolite localization for the *Pinus sylvestris-Lactarius rufus* symbiosis. *FEMS Microbiol Ecol* **56**: 25–33.

Aspray, T.J., Jones, E.E., Davies, M.W., Shipman, M., and Bending, G.D. (2013) Increased hyphal branching and growth of ectomycorrhizal fungus *Lactarius rufus* by the helper bacterium *Paenibacillus* sp. *Mycorrhiza* **23**: 403–410.

Baldani, J.I., Baldani, V.L.D., Seldin, L., and Döbereiner, J. (1986) Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. *Int J Syst Evol Microbiol* **36**: 86–93.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Baquiran, J.-P., Thater, B., Sedky, S., De Ley, P., Crowley, D., Orwin, P.M., and Badger, J.H. (2013) Culture-Independent investigation of the microbiome associated with the nematode *Acrobeloides maximus*. *PLoS One* **8**: e67425.

Barbieri, E., Bertini, L., Rossi, I., Ceccaroli, P., Saltarelli, R., Guidi, C., *et al*. (2005) New evidence for bacterial diversity in the ascoma of the ectomycorrhizal fungus *Tuber borchii* Vittad. *FEMS Microbiol Lett* **247**: 23–35.

Barbieri, E., Guidi, C., Bertaux, J., Frey-Klett, P., Garbaye, J., Ceccaroli, P., *et al*. (2007) Occurrence and diversity of bacterial communities in Tuber magnatum during truffle maturation. *Environ Microbiol* **9**: 2234–2246.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., *et al*. (2013) Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol Evol* **4**: 914–919.

Benucci, G.M.N., and Bonito, G.M. (2016) The truffle microbiome: species and geography effects on bacteria associated with fruiting bodies of hypogeous Pezizales. *Microbiol Ecol* **72**: 4–8.

Bergmann, G.T., Bates, S.T., Eilers, K.G., Lauber, C.L., Caporaso, J.G., Walters, W.A., *et al*. (2011) The under-recognized dominance of Verrucomicrobia in soil bacterial communities. *Soil Biol Biochem* **43**: 1450–1455.

Chen, S., Qiu, C., Huang, T., Zhou, W., Qi, Y., Gao, Y., *et al*. (2013) Effect of 1-aminocyclopropane-1-carboxylic acid deaminase producing bacteria on the hyphal growth and primordium initation of Agaricus bisporus. *Fung Ecol* **6**: 110–118.

Cho, Y.S., Kim, J.S., Crowley, D.E., and Cho, B.G. (2003) Growth promotion of the edible fungus *Pleurotus ostreatus* by fluorescent pseudomonads. *FEMS Microbio Lett* **218**: 271–276.

Citterio, B., Malatesta, M., Battistelli, S., Marcheggiani, F., Baffone, W., Saltarelli, R., *et al*. (2001) Possible involvement of *Pseudomonas fluorescens* and Bacillaceae in structural modifications of Tuber borchii fruit bodies. *Can J Microbiol* **47**: 264–268.

Dahm, H., Wrótniak, W., Strzelczyk, E., and Bednarska, E. (2005) Diversity of culturable bacteria associated with fruiting bodies of ectomycorrhizal fungi. *Phytopathol Pol* **38**: 51–62.

Danell, E., Alström, S., and Ternström, A. (1993) *Pseudomonas fluorescens* in association with fruitbodies of the ectomycorrhizal mushroom *Cantharellus cibarius*. *Mycol Res* **97**: 1148–1152.

Deveau, A., Palin, B., Delaruelle, C., Peter, M., Kohler, A., Pierrat, J.C., *et al*. (2007) The mycorrhiza helper Pseudomonas fluorescens BBc6R8 has a specific priming effect on the growth, morphology and gene expression of the ectomycorrhizal fungus *Laccaria bicolor* S238N. *New Phytol* **175**: 743–755.

Dominguez, J.A., Martin, A., Anriquez, A., and Albanesi, A. (2012) The combined effects of *Pseudomonas fluorescens* and *Tuber melanosporum* on the quality of *Pinus halepensis* seedlings. *Mycorrhiza* **22**: 429–436.

Doré, J., Perraud, M., Dieryckx, C., Kohler, A., Morin, E., Henrissat, B., *et al*. (2015) Comparative genomics, proteomics and transcriptomics give new insight into the exoproteome of the basidiomycete *Hebeloma cylindrosporum* and its involvement in ectomycorrhizal symbiosis. *New Phytol* **208**: 1169–1187.

Eilers, K.G., Lauber, C.L., Knight, R., and Fierer, N. (2010) Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. *Soil Biol Biochem* **42**: 896–903.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., *et al*. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci USA* **109**: 21390–21395.

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R.A., Henrissat, B., *et al*. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**: 1715–1719.

Frey-Klett, P., Garbaye, J., and Tarkka, M. (2007) The mycorrhiza helper bacteria revisited. *New Phytol* **176**: 22–36.

Haas, D., and Defago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nat Rev Microbiol* **3**: 307–319.

Halsey, J.A., de Cássia Pereira e Silva, M., and Andreote, F.D. (2016) Bacterial selection by mycospheres of Atlantic Rainforest mushrooms. *Antonie Leeuwenhoek* **109**: 1353–1365.

Han, M.V., Thomas, G.W., Lugo-Martinez, J., and Hahn, M.W. (2013) Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol Biol Evol* **30**: 1987–1997.

Hibbett, D., Grimaldi, D., and Donoghue, M. (1997) Fossil mushrooms from Miocene and Cretaceous ambers and the evolution of Homobasidiomycetes. *Am J Bot* **84**: 981.

Hibbett, D.S., Gilbert, L.-B., and Donoghue, M.J. (2000) Evolutionary instability of ectomycorrhizal symbioses in basidiomycetes. *Nature* **407**: 506–508.

Holt, C., and Yandell, M. (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform* **12**: 491.

Huerta-Cepas, J., Forslund, K., Pedro Coelho, L., Szklarczyk, D., Juhl Jensen, L., von Mering, C., *et al*. (2017) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Bio Evol* **34**: 2115–2122.

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site *identification*. *BMC Bioinform* **11**: 119.

Janssen, P.H. (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA Genes. *Appl Environ Microbiol* **72**: 1719–1728.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., *et al*. (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**: 726–731.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., *et al*. (2015) Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat Genet* **47**: 410–415.

Krechel, A., Faupel, A., Hallmann, J., Ulrich, A., and Berg, G. (2002) Potato-associated bacteria and their antagonistic potential towards plant-pathogenic fungi and the plant-parasitic nematode *Meloidogyne incognita* (Kofoid & White) Chitwood. *Can J Microbiol* **48**: 772–786.

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F.S., Bahram, M., *et al*. (2013) Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* **22**: 5271–5277.

Kumari, D., Reddy, M.S., and Upadhyay, R.C. (2013) Diversity of cultivable bacteria associated with fruiting bodies of wild Himalayan *Cantharellus* spp. *Ann Microbiol* **63**: 845–853.

Lackner, G., Partida-Martinez, L.P., and Hertweck, C. (2009) Endofungal bacteria as producers of mycotoxins. *Trends Microbiol* **17**: 570–576.

Laetsch, D.R., and Blaxter, M.L. (2017) BlobTools: Interrogation of genome assemblies. *F1000Research* **6**: 1287.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucl Acid Res* **35**: 3100–3108.

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lartillot, N., Rodrigue, N., Stubbs, D., and Richer, J. (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst Biol* **62**: 611–615.

Larsen, T.S., and Krogh, A. (2003) EasyGene–a prokaryotic gene finder that ranks ORFs by statistical significance. *BMC Bioinform* **4**: 21.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**: 2178–2189.

Martin, F., Kohler, A., Murat, C., Balestrini, R., Coutinho, P.M., Jaillon, O., *et al*. (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**: 1033–1038.

Matheny, P.B., Curtis, J.M., Hofstetter, V., Aime, M.C., Moncalvo, J.M., Ge, Z.W., *et al*. (2006) Major clades of Agaricales: a multilocus phylogenetic overview. *Mycologia* **98**: 982–995.

Matheny, P.B., Aime, M.C., Bougher, N.L., Buyck, B., Desjardin, D.E., Horak, E., *et al*. (2009) Out of the Palaeotropics? historical biogeography and diversification of the cosmopolitan ectomycorrhizal mushroom family Inocybaceae. *J Biogeogr* **36**: 577–592.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., *et al*. (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., *et al*. (2008) The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform* **9**: 386.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**: 1188–1195.

Mirarab, S., and Warnow, T. (2015) ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* **31**: i44–i52.

Mu, L.L., Yun, Y.B., Park, S.J., Cha, J.S., and Kim, Y.K. (2015) Various pathogenic Pseudomonas strains that cause brown blotch disease in cultivated mushrooms. *J Appl Biol Chem* **58**: 349–354.

Munsch, P., Alatossava, T., Marttinen, N., Meyer, J.M., Christen, R., and Garden, L. (2002) *Pseudomonas costantinii* sp. nov., another causal agent of brown blotch disease, isolated from cultivated mushroom sporophores in Finland. *Int J Syst Evol Microbiol* **52**: 1973–1983.

Nazir, R., Warmink, J.A., Boersma, H., and van Elsas, J.D. (2010a) Mechanisms that promote bacterial fitness in

fungal-affected soil microhabitats. *FEMS Microbiol Ecol* **71**: 169–185.

Nazir, R., Boersma, F.G.H., Warmink, J.A., and van Elsas, J.D. (2010b) *Lyphyllum* sp. strain Karsten alleviates pH pressure in acid soil and enhances the survival of *Variovorax paradoxus* HB44 and other bacteria in the mycosphere. *Soil Biol Biochem* **42**: 2146–2152.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.

Olson, Å., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., *et al*. (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytol* **194**: 1001–1013.

Parra, G., Bradnam, K., Ning, Z., Keane, T., and Korf, I. (2009) Assessing the gene space in draft genomes. *Nucleic Acids Res* **37**: 289–297.

Quandt, C.A., Kohler, A., Hesse, C., Sharpton, T., Martin, F., and Spatafora, J.W. (2015) Metagenome sequence of Elaphomycesgranulatus from sporocarp tissue reveals Ascomycota ectomycorrhizal fingerprints of genome expansion and a Proteobacteria rich microbiome. *Environ Microbiol* **17**: 2952–2968.

Pent, M., Põldmaa, K., and Bahram, M. (2017) Bacterial communities in boreal forest mushrooms are shaped both by soil parameters and host identity. *Front Microbiol* **8**: 836.

Riedlinger, J., Schrey, S.D., Tarkka, M.T., Hampp, R., Kapur, M., and Fiedler, H.-P. (2006) Auxofuran, a novel metabolite that stimulates the growth of fly agaric, is produced by the mycorrhiza helper bacterium Streptomyces strain AcH 505. *Appl Environ Microbiol* **72**: 3550–3557.

Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S.R., Wilkie, A.O., WGS500 Consortium, *et al*. (2014) Integrating mapping-, assembly-and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* **46**: 912–918.

Smit, A., and Hubley, R. (2011) *RepeatModeler-1.0. 5*. WA: Institute of Systems Biology.

Smit, A.F., Hubley, R., and Green, P. (2010). *RepeatMasker Open-3.0*. URL http://www.repeatmasker.org/

Smith, S.E., and Read, D.J. (2010) *Mycorrhizal Symbiosis*. 3rd ed. London, UK: Academic Press.

Spanu, P.D., Abbott, J.C., Amselem, J., Burgis, T.A., Soanes, D.M., Stuber, K., *et al*. (2010) Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**: 1543–1546.

Splivallo, R., Deveau, A., Valdez, N., Kirchhoff, N., Frey-Klett, P., and Karlovsky, P. (2015) Bacteria associated with truffle-fruiting bodies contribute to truffle aroma. *Environ Microbiol* **17**: 2647–2660.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stopnisek, N., Zühlke, D., Carlier, A., Barberán, A., Fierer, N., Becher, D., *et al*. (2016) Molecular mechanisms underlying the close association between soil Burkholderia and fungi. *Isme* **10**: 253–264.

Stubblefield, S.P., Taylor, T.N., and Beck, C.B. (1985) Studies of paleozoic fungi. IV. Wood-decaying fungi in Callixylon newberryi from the upper Devonian. *Am J Bot* **72**: 1765–1774.

Talavera, G., Castresana, J., Kjer, K., Page, R., and Sullivan, J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Boil* **56**: 564–577.

Tedersoo, L., May, T.W., and Smith, M.E. (2010) Ectomycorrhizal lifestyle in fungi: global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza* **20**: 217–263.

Thorne, J.L., Kishino, H., and Painter, I.S. (1998) Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol* **15**: 1647–1657.

Tian, X., Cheng, X., Mao, Z., Chen, G., Yang, J., and Xie, B. (2011) Composition of Bacterial communities associated with plant-parasitic nematode *Bursaphelenchus mucronatus*. *Curr Microbiol* **62**: 117–125.

Torres-Cortés, G., Ghignone, S., Bonfante, P., and Schüßler, A. (2015) Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: Transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proc Natl Acad Sci USA* **112**: 7785–7790.

Tsukamoto, T., Murata, H., and Shirata, A. (2002) Identification of non-pseudomonad bacteria from fruit bodies of wild agaricales fungi that detoxify tolaasin produced by *Pseudomonas tolaasii*. *Biosci Biotech Biochem* **66**: 2201–2208.

UniProt Consortium (2014) UniProt: a hub for protein information. *Nucl Acid Res* **43**: D204-12.

Uroz, S., Oger, P., Morin, E., and Frey-Klett, P. (2012) Distinct ectomycorrhizospheres share similar bacterial communities as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Appl Environ Microbiol* **78**: 3020–3024.

Veldre, V., Abarenkov, K., Bahram, M., Martos, F., Selosse, M.-A., Tamm, H., *et al*. (2013) Evolution of nutritional modes of Ceratobasidiaceae (Cantharellales, Basidiomycota) as revealed from publicly available ITS sequences. *Fung Ecol* **6**: 256–268.

Vicente, C.S.L., Nascimento, F., Espada, M., Barbosa, P., Mota, M., Glick, B.R., *et al*. (2012) Characterization of bacteria associated with pinewood nematode *Bursaphelenchus xylophilus*. *PLoS One* **7**: e46661.

Wang, H.C., Minh, B.Q., Susko, E., and Roger, A.J. (2017) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst Biol* doi:10.1093/sysbio/syx068 (in press).

Warmink, J.A., Nazir, R., and van Elsas, J.D. (2009) Universal and species-specific bacterial “fungiphiles” in the mycospheres of different basidiomycetous fungi. *Environ Microbiol* **11**: 300–312.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **40**: W445–W451.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Fig. S1. The *I. terrigena* genome includes a smaller number of genes compared to other ecotmycorrhizal genomes. Boxplot showing the number of gene models reported

previously for different functional guilds of fungi. Dash line corresponds to the number of gene models uncovered for *Inocybe terrigena* in this study.

Fig. S2. Relationship of contig coverage and contig length of dominant bacterial genera in *I. terrigena*.

Fig. S3. The relative abundance of fungal functional gene categories in *Inocybe terrigena* genome.

Fig. S4. Pie chart showing the relative abundance of bacterial gene functional categories in *Inocybe terrigena* fruitbody based on KEGG Orthology groups (A) and Subsystems (B). Functional categories with relative abundance of $\geq$ 0.5% are presented. Only contigs that were initially identified as bacteria based on BLAST searches were included.

Table S1. Major features of the genome of *Inocybe terrigena*.

Table S2. Features and taxonomic identification of contigs assembled based on Inocybe terrigena metagenome.

Table S3. Genes identified from fungal contigs based on Inocybe terrigena metagenome. Tabs present basic features, KEGG id and KEGG pathways retrieved from GhostKOALA.

Table S4. Results of CEGMA (Statistics of the completeness of the genome based on 248 CEGs).

Table S5. Genomes used for comparative genomics.

Table S6. CAZyme gene profiles of 59 Agaricomycetes genomes used in comparative genomics analysis.