Full Paper

New Phytologist

Comparative genomics reveals dynamic genome evolution in host specialist ectomycorrhizal fungi

Lotus A. Lofgren[1,2], Nhu H. Nguyen[3], Rytas Vilgalys[4], Joske Ruytinx[5], Hui-Ling Liao[6], Sara Branco[7], Alan Kuo[8], Kurt LaButti[8], Anna Lipzen[8], William Andreopoulos[8], Jasmyn Pangilinan[8], Robert Riley[8], Hope Hundley[8], Hyunsoo Na[8], Kerrie Barry[8], Igor V. Grigoriev[8,9], Jason E. Stajich[1] and Peter G. Kennedy[2]

[1]Department of Microbiology and Plant Pathology, University of California Riverside, Riverside, CA 92507, USA; [2]Department of Plant and Microbial Biology, University of Minnesota, St Paul, MN 55108, USA; [3]Department of Tropical Plant and Soil Science, University of Hawaii, Manoa, HI 96822, USA; [4]Department of Biology, Duke University, Durham, NC 27708, USA; [5]Research group Microbiology, Department of Bio-engineering Sciences, Vrije Universiteit Brussel, Brussel BE1500, Belgium; [6]Department of Soil Microbial Ecology, University of Florida, Quincy, FL 32351, USA; [7]Department of Integrative Biology, University of Colorado Denver, Denver, CO 80204, USA [Correction added after online publication 6 February 2021: the preceding affiliation has been updated]; [8]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; [9]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

Author for correspondence:
Lotus A. Lofgren
Email: lotuslofgren@gmail.com

Received: *19 October 2020*
Accepted: *16 December 2020*

New Phytologist (2021) **230:** 774–792
doi: 10.1111/nph.17160

Key words: genomics, G-protein coupled receptors, host specificity, secondary metabolites, small secreted proteins, *Suillus*.

Summary

- While there has been significant progress characterizing the 'symbiotic toolkit' of ectomycorrhizal (ECM) fungi, how host specificity may be encoded into ECM fungal genomes remains poorly understood.
- We conducted a comparative genomic analysis of ECM fungal host specialists and generalists, focusing on the specialist genus *Suillus*. Global analyses of genome dynamics across 46 species were assessed, along with targeted analyses of three classes of molecules previously identified as important determinants of host specificity: small secreted proteins (SSPs), secondary metabolites (SMs) and G-protein coupled receptors (GPCRs).
- Relative to other ECM fungi, including other host specialists, *Suillus* had highly dynamic genomes including numerous rapidly evolving gene families and many domain expansions and contractions. Targeted analyses supported a role for SMs but not SSPs or GPCRs in *Suillus* host specificity. Phylogenomic-based ancestral state reconstruction identified *Larix* as the ancestral host of *Suillus*, with multiple independent switches between white and red pine hosts.
- These results suggest that like other defining characteristics of the ECM lifestyle, host specificity is a dynamic process at the genome level. In the case of *Suillus*, both SMs and pathways involved in the deactivation of reactive oxygen species appear to be strongly associated with enhanced host specificity.

Introduction

]Fungi play critical and diverse roles in ecosystems as pathogens, saprotrophs and symbionts of both plants and animals (Peay *et al.*, 2016). With the rapid rise of available genome sequences for fungi, there has been growing interest in linking fungal ecological lifestyle with genome content (Kohler *et al.*, 2015; Martino *et al.*, 2018; Knapp *et al.*, 2018; Lofgren *et al.*, 2019; Haridas *et al.*, 2020). Ectomycorrhizal (ECM) fungi form mutualistic associations with a wide range of woody plants, representing *c.* 60% of all trees in the Earth's forest ecosystems (Steidinger *et al.*, 2019). Comparative genomic analyses have revealed multiple insights into the ECM fungal lifestyle, including losses of plant cell wall-degrading enzymes, presumably as an adaptation from free-living to plant-associated symbioses (Kohler *et al.*, 2015; Peter *et al.*, 2016; Miyauchi *et al.*, 2020) as well as an abundance of lineage-specific genes involved in the degradation of organic matter (Floudas *et al.*, 2012; Kohler *et al.*, 2015; Sipos *et al.*, 2017).

A distinguishing feature of the ECM fungal lifestyle relative to other types of mycorrhizal interactions is the presence of highly host-specific associations (Molina *et al.*, 1992; Bruns *et al.*, 2002). One of the best documented examples is the ECM fungal genus *Suillus*, which forms nearly exclusive associations with trees in the family Pineaceae (Kretzer *et al.*, 1996; Lofgren *et al.*, 2018). These associations are tightly coupled, with a given *Suillus* species forming specialized associations with a single host genus (particularly the genera *Pinus*, *Larix* and *Pseudotsuga*), and even

with different subgenera within the genus *Pinus* (Kretzer *et al.*, 1996; Liao *et al.*, 2016). The evolution of fungal–host associations within *Suillus*, however, is dynamic and involves multiple independent host-switching events (Nguyen *et al.*, 2016). In addition to their emerging use as a model of fungal host specificity (Liao *et al.*, 2014, 2016; Nguyen *et al.*, 2016), *Suillus* fungi have well-demonstrated ecological importance, particularly as facilitators of tree establishment in native and exotic ranges (Dickie *et al.*, 2010; Policelli *et al.*, 2019) and as producers of prolific, long-distance extramatrical mycelium, which acts as a major below-ground carbon sink (Agerer, 2001; Bidartondo *et al.*, 2001).

Current understanding of the drivers of fungal host specificity is heavily influenced by the field of plant pathology. Seminal work on pathogen host switching, host range expansions/contractions and context-dependent compatibility have helped to elucidate both the genetic underpinnings and ecological pressures selecting for the diverse range of specificity relationships observed across the fungal phylogeny (Gilbert & Webb, 2007; Schulze-Lefert & Panstruga, 2011; Lo Presti *et al.*, 2015). Despite these advances, the processes facilitating host specificity in fungal mutualisms are less well understood. For example, in pathogen systems, restricted host range is accompanied by gene losses, presumably correlated to the loss of traits needed to colonize diverse hosts (Spanu *et al.*, 2010; Visser *et al.*, 2010; Baroncelli *et al.*, 2016). Whether this pattern also holds for fungal mutualists is not yet clear, as the ecological pressures structuring genome evolution in mutualisms may be different from those of antagonisms (Gladieux *et al.*, 2014; McLaughlin & Malik, 2017; Stajich, 2017).

Multiple classes of molecules have garnered repeated attention in relation to fungal host specificity, including three that display differential expression during the establishment of mycorrhizas in compatible host interactions in *Suillus*: small secreted proteins (SSPs), secondary metabolites (SMs), and G-protein coupled receptors (GPCRs) (Liao *et al.*, 2016). SSPs are often species-specific (termed SSSPs) and have been shown to play critical roles during the process of ECM mycorrhization generally (Plett *et al.*, 2014), with mycorrhizal-induced SSPs (MiSSPs) constituting 8–28% of the genes upregulated during symbiosis (Martin *et al.*, 2008; Kohler *et al.*, 2015). Although the majority of fungal SSPs display little sequence conservation with known proteins and most are functionally uncharacterized, it has been hypothesized that many function as effectors (Kim *et al.*, 2016; Plett *et al.*, 2020). Fungal effectors are secreted, cysteine-rich molecules that play a key role in host susceptibility to colonization and are well defined in fungal pathogen systems (Uhse & Djamei, 2018). In pathogens, effectors structure host specificity at multiple scales, spanning kingdoms to individual tissue types (Skibbe *et al.*, 2010; Irieda *et al.*, 2018). This range implies that not all effector targets are present in all host species, and specific suites of effectors are required to interact with specific host genotypes. Indeed, it has been suggested that host specificity in fungal pathogens may be directly regulated by the specific complement of effectors produced, where mutation, loss, or gain of effectors, modulate colonization success and resultant host ranges (Pritchard & Birch, 2011; Dong *et al.*, 2014; Sharma *et al.*, 2014; vanDam *et al.*, 2016).

Secondary metabolites were among the first molecular factors identified to play a role in fungal host specificity, as host-specific toxins (HSTs) associated with fungal pathogenesis (Walton & Panaccione, 1993). The genes responsible for secondary metabolite production are generally clustered in fungal genomes, allowing for the coordinated transcription of multistep reactions leading to the biosynthesis of complex molecules (Keller & Hohn, 1997). These molecules represent a large number of bioactive compounds synthesized by a limited number of core biosynthetic enzymes, primarily nonribosomal peptide synthases (NRPS), polyketide synthases (PKS) and terpene synthases or cyclases (terpenes). The products of SM clusters have multiple functions, including virulence (Collemare & Lebrun, 2011), antibacterial activity (de Weert *et al.*, 2007), communication (Brakhage, 2013) and host metabolic changes, such as the induction of growth factors and genes related to nutrient acquisition (Contreras-Cornejo *et al.*, 2016). Diverse SMs have also been associated with fungal host specificity, although the mechanisms differ widely across fungal lifestyle and phylogeny (Dunkle *et al.*, 1991; Walton, 2006; Tsuge *et al.*, 2016; Liao *et al.*, 2016).

G-protein coupled receptors are the largest class of signal transduction molecules in eukaryotes, functioning in the sensing of numerous external stimuli (Kochman, 2014). Although GPCRs exhibit low sequence similarity, they share a common architecture, including the presence of seven transmembrane domains, an extracellular N-terminus and an intracellular C-terminus. The role of GPCRs in the transduction of environmental signals has also been shown to extend to host recognition in fungi. For example, Pth11-like GPCRs are involved in host species recognition in the entomopathogenic fungal genus *Metarhizium* (Gao *et al.*, 2011) and are differentially expressed among fungal and insect hosts in the fungal genus *Tolypocladium* (Quandt *et al.*, 2016). Further, in the ECM fungi *Laccaria bicolor* and *Tuber melanosporum*, GPCR and G-protein related transcripts were found to be the most highly upregulated signaling genes transcribed during ECM colonization (Voiblet *et al.*, 2001; Martin *et al.*, 2010; Plett *et al.*, 2012).

Although *Suillus* species have previously displayed differential regulation of SSPs, SM clusters and GPCRs depending on host compatibility (Liao *et al.*, 2016), it is unknown whether these expression differences are directly related to host specificity or associated with the process of ECM colonization more generally. To better understand how host specificity may be directly encoded onto the genomes of ECM fungi, we conducted a global analysis of gene family dynamics along with a targeted analysis of SSPs, SMs and GPCRs for 23 *Suillus* species (including 22 newly sequenced genomes) and 23 non-*Suillus* ECM species (hereafter referred to as 'Other ECM'). The Other ECM group included five representatives from the genus *Rhizopogon*, the sister genus to *Suillus* (Kretzer *et al.*, 1996), as well as 18 species representing a broader phylogenetic range (Table 1). Although mostly comprising host generalists, the Other ECM fungal group also included several species with high and moderate host specificity (see later for definitions of specificity categories), which were chosen based on genome availability (Table 1). We also made the same global and targeted comparisons among white pine-, red pine-, and

Table 1 Ectomycorrhizal fungal species included in this study, identifying information by JGI project code, host specificity information with citations, and citation for each genome sequencing project.

Species	JGI project code	Specificity	Specificity citation	Genome citation
Amanita muscaria	Amamu1	Very low	*Northofagus* (Dunk *et al*., 2012), *Quercus* (Vargas *et al*., 2019), *Pinus*, (Sawyer *et al*., 2001)	Kohler *et al*. (2015)
Cantherellus anzutake	Cananz1	Low	*Pinus*, *Quercus* (Ogawa *et al*., 2019)	Miyauchi *et al*. (2020)
Gautieria morchelliformis	Gaumor1_1	Low	*Pinus*, *Quercus* (Nedelin *et al*., 2016)	Miyauchi *et al*. (2020)
Gyrodon lividus	Gyrli1	High	*Alnus* (Hayward & Thiers, 1984; Henrici, 2006)	Miyauchi *et al*. (2020)
Hebeloma cylindrosporum	Hebcy2	Very Low	*Pinus*, *Cistus*, (Marmeisse *et al*., 2004), *Larix (*Wong & Fortin, 1989*)*, *Quercus* (Oh *et al*., 1995), *Dryas* (Melville *et al*., 1987)	Kohler *et al*. (2015)
Hydnum rufescens	Hydru2	Low	*Picea* (Grebenc *et al*., 2009), *Abies* (Ważny, 2014), *Pinus* (Feng *et al*., 2016)	Miyauchi *et al*. (2020)
Hysterangium stoloniferum	Hyssto1	Low	*Quercus* (Castellano, 1999), *Picea* (Raidl & Agerer, 1998)	Miyauchi *et al*. (2020)
Laccaria amethystina	Lacam2	Very Low	*Fagus*, *Abies*, *Carpinus*, *Quercus* (Roy *et al*., 2008) *Pinus* (Ter-amoto *et al*., 2012)	Kohler *et al*. (2015)
Laccaria bicolor	Lacbi2	Very Low	*Populus*, *Pseudotsuga* (Plett *et al*., 2015), *Pinus* (Reininger & Sieber, 2012; Hazard *et al*., 2017)	Martin *et al*. (2008)
Paxillus involutus	Paxin1	Low	*Alnus* (Murphy & Miller, 1994), *Betula*, *Picea* (Hedh *et al*., 2009)	Kohler *et al*. (2015)
Piloderma olivaceum	Pilcr1	Low	*Pinus* (Heinonsalo *et al*., 2015), *Pseudotsuga* (Kranabetter *et al*., 2012)	Kohler *et al*. (2015)
Pisolithus microcarpus	Pismi1	Low	*Eucalypus*, *Acacia* (Martin *et al*., 2002)	Kohler *et al*. (2015)
Pisolithus tinctorius	Pisti1	Low	*Pinus*, *Quercus* (Oh *et al*., 1995; Martin *et al*., 2002)	Kohler *et al*. (2015)
Rhizopogon salebrosus	Rhisa1	Moderate	*Pinus* (red) (Kennedy & Bruns, 2005), *Pinus* (white) (Kohout *et al*., 2011)	Unpublished – used with permission
Rhizopogon truncatus	Rhitru1	Low	*Pinus* (red) (Massicotte *et al*., 1999), *Tsuga* (Trappe, 2009)	Unpublished – used with permission
Rhizopogon vesiculosus	Rhives1	High	*Pseudotsuga* (Massicotte *et al*., 1994)	Mujic *et al*. (2017)
Rhizopogon vinicolor	Rhivi1	High	*Pseudotsuga* (Grubisha *et al*., 2002)	Mujic *et al*. (2017)
Rhizopogon vulgaris	Rhivul1	Moderate	*Pinus* (red and white) (Mujic *et al*., 2017)	Unpublished – used with permission
Russula brevipes	Rusbre1	Low	*Pinus*, *Picea* (Bergemann & Miller, 2002)	Unpublished – used with permission
Russula compacta	Ruscom1	Very Low	*Anthonotha*, *Cryptosepalum*, *Paramacrolobium*, *Uapaca* (Diédhiou *et al*., 2010) *Quercus*, *Rhododendron*, *Myrica*, *Pinus*, *Cedrus*, *Cupressus* (Bhatt *et al*., 2014)	Unpublished - used with permission
Scleroderma citrinum	Sclci1	Low	*Larix* (Richter & Bruhn, 1990), *Pinus* (Mohan *et al*., 1993), *Picea* (Brunner *et al*., 1992)	Kohler *et al*. (2015)
Suillus americanus	Suiame1	High	*Pinus* (white) (Smith & Thiers, 1964)	This paper
Suillus ampliporus	Suiamp1	High	*Larix* (Nguyen *et al*., 2016)	This paper
Suillus bovinus	Suibov1	High	*Pinus* (red) (Dahlberg & Stenlid, 1994)	This paper
Suillus brevipes	Suibr2	High	*Pinus* (red) (Smith & Thiers, 1964)	This paper
Suillus clintonianus	Suicli1	High	*Larix* (Nguyen *et al*., 2016)	This paper
Suillus cothurnatus	Suicot1	High	*Pinus* (red) (Nguyen *et al*., 2016)	This paper
Suillus decipiens	Suidec1	High	*Pinus* (red) (Nguyen *et al*., 2016)	This paper
Suillus fuscotomentosus	Suifus1	High	*Pinus* (red) (Siegel & Schwarz, 2016)	This paper
Suillus weaverae (granulatus)	Suigr1	Unknown	Unknown (white in Smith & Thiers, 1964 but white or red in Kuo & Methven, 2010)	This paper
Suillus hirtellus	Suihi1	High	*Pinus* (red) (Smith & Thiers, 1964)	This paper
Suillus lakei	Suilak1	High	*Pseudotsuga* (Smith & Thiers, 1964)	This paper
Suillus luteus	Suilu4	High	*Pinus* (red) (Smith & Thiers, 1964)	Kohler *et al*. (2015)
Suillus occidentalis	Suiocc1	High	*Pinus* (red) (Nguyen *et al*., 2016)	This paper
Suillus paluster	Suipal1	High	*Larix* (Nguyen *et al*., 2016)	This paper
Suillus spraguei (pictus)	Suipic1	High	*Pinus* (white) (Smith & Thiers, 1964)	This paper
Suillus placidus	Suipla1	High	*Pinus* (white) (Smith & Thiers, 1964)	This paper
Suillus plorans	Suiplo1	High	*Pinus* (white) (Nguyen *et al*., 2016)	This paper
Suillus subalutaceus	Suisu1	High	*Pinus* (red) (Nguyen *et al*., 2016)	This paper
Suillus subaureus	Suisub1	Low	*Pinus* (white), *Larix*, *Quercus* (Lofgren *et al*., 2018)	This paper
Suillus cf. *subluteus*	Suisubl1	High	*Pinus* (white) (Nguyen *et al*., 2016)	This paper
Suillus tomentosus	Suitom1	High	*Pinus* (red) (Smith & Thiers, 1964)	This paper
Suillus variegatus	Suivar1	High	*Pinus* (red) (Nguyen *et al*., 2016)	This paper

Table 1 (Continued)

Species	JGI project code	Specificity	Specificity citation	Genome citation
Suillus discolor	Suidis1	High	Pinus (white) (Nguyen *et al*., 2016)	This paper
Thelephora terrestris	Theter1	Very low	*Alnus*, *Pseudotsuga* (Miller *et al*., 1992), *Eucalyptus* (Ingleby & Mason, 1996), *Pinus* (Pera & Alvarez, 1995)	Miyauchi *et al*. (2020)
Thelephora ganbajun	Thega1	Low	*Pinus*, *Keteleeria*, *Cunninghamia* (Mortimer *et al*., 2012)	Miyauchi *et al*. (2020)

larch-associated *Suillus* species to assess genomic differences specific to a given host genus association. Further, we used phylogenomic-based ancestral state reconstruction to identify the ancestral host of *Suillus* as well as major host-switching and speciation events in the evolutionary history of the genus. Based on previous studies of fungal host specificity conducted in pathogen systems, we hypothesized that *Suillus* would possess significantly more gene family contractions than expansions, consistent with reduced host range, and significantly more contractions than Other ECM fungal species, while displaying greater molecular diversity in *Suillus* lineages which have recently switched host groups. We further hypothesized that gene losses in *Suillus* would be reflected in the targeted analysis, with less diversity of SSPs, SM clusters and GPCRs in *Suillus* over Other ECM fungal species.

Materials and Methods

Fungal strains, extraction and genome preparation

Cultures from 22 *Suillus* species were isolated from sporocarps (dikaryons) associated with different host genera: *Pinus* subgenus *Pinus* (hereafter 'red pine'), $n = 10$; *Pinus* subgenus *Strobus* (hereafter 'white pine'), $n = 7$; *Larix* (hereafter 'larch'), $n = 3$; *Pseudotsuga*, $n = 1$; or *Quercus*, $n = 1$ (Table 1). The genome for *S. luteus* (red pine) was previously sequenced and is described in Kohler *et al.,* (2015). *Suillus* genomes were coded by host association, as noted earlier, and only those with more than three representatives and associated with a single host genus were used for the intrageneric analyses. All ECM fungal species were encoded on a host specificity scale based on reported associations in the primary literature (Table 1), with 'very low' indicating colonization of hosts from three or more genera from distantly related lineages (e.g. gymnosperms and angiosperms), 'low' indicating colonization of hosts from two desperate genera, 'moderate' indicating colonization of multiple species in closely related genera (e.g. *Quercus* and *Fagus*), or different subgenera of a single genus, and 'high' indicating the ability to colonize species within a single subgenus. Complete metadata for all 46 species are publicly available on JGI's MycoCosm Portal at https://mycocosm.jgi.doe.gov (Grigoriev *et al*., 2014) and the *Suillus* genomes can be accessed directly at the JGI *Suillus* Portal at https://mycocosm.jgi.doe.gov/mycocosm/home/releases?flt=suillus. Cultured *Suillus* isolates were grown on shakers in liquid Modified Melin-Norkrans media at room temperature. For all species except *Suillus bovinus* and *Suillus variegatus*, DNA and RNA were coextracted using CTAB/chloroform and LiCl precipitation as described in Liao *et al.* (2014). DNA for *S. bovinus* and *S. variegatus* were prepared as earlier, but RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Germany), eliminating genomic DNA with the Turbo 365 DNA-free kit (Applied Biosystems, Waltham, MA, USA). A detailed account of genome sequencing and assembly, transcriptome sequencing and assembly, and genome annotation for newly sequenced *Suillus* genomes is presented in the Supporting Information Methods S1 (see Table S1 for genome sequencing and assembly statistics).

Phylogenomic analysis

We used two different approaches for the phylogenomic analyses. The first approach used a gene tree method capable of taking advantage of multicopy orthologs (Emms & Kelly, 2019), while the second was based on an in-house multilocus single-copy orthologs approach (https://doi.org/10.5281/zenodo.1257002). To take advantage of multicopy orthologs, we ran OrthoFinder 2.0 with Diamond and inferred gene trees using the Blast-based hierarchical clustering algorithm DendroBlast (Emms & Kelly, 2015). Using the full set of all unrooted gene trees, the species tree was inferred using the algorithm Stag (Species Tree Inference from All Genes) and rooted by Stride (Emms & Kelly, 2018). To resolve the backbone of the *Suillus* phylogeny, we ran PHYling (https://doi.org/10.5281/zenodo.1257002) using 434 single-copy markers (https://doi.org/10.5281/zfenodo.3630031) and the maximum likelihood (ML) algorithm IQ-Tree (Nguyen *et al*., 2015). The best-fit model according to Bayesian information criterion score was determined using the ModelFinder function in IQ-Tree and was determined to be JTT + F+I + G4, which was run with 1000 rapid bootstrap iterations. Individual branch support values were assessed using a Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) in IQ-Tree over 1000 iterations. Both trees were constructed using only species in the genera *Suillus* and *Rhizopogon* ($n = 28$ genomes). Additionally, the PHYling method was used to construct a third tree consisting of all 46 species used in this study, for use in the Café analysis.

Ancestral state reconstruction was performed using the R packages phytools (Revell, 2012) and ape (Paradis *et al*., 2004). Model selection was preformed using the 'fitMk' function in phytools, and weighted Akaike information criterion (AIC) calculated with the 'aic.w' function to choose between ER (equal rates), SYM (symmetric backward and forward rates), ARD (all-rates-different) models. For both trees, the ER model had the highest AIC (Stag: ER = 0.853 896 36, SYM = 0.146 059 68, ARD = 0.000 043 97; PHYling: ER = 0.845 568 52, SYM = 0.154 335 21, ARD = 0.000 096 26) and was used for

calculating the ancestral state probability at each node using the 'ace' function in APE. Vertical node ordering and tree comparison were set using the 'cophylo' function in PHYTOOLS.

Comparative analysis

All genomes were assessed for completeness using Benchmarking of Universal Single-Copy Orthologs (BUSCO) v.4.0.5 based on the BUSCO model set for basidiomycota_odb10 (Simão *et al.*, 2015). To calculate gene family expansions and contractions for all 46 species, we clustered the protein sequences into families using the Markov Cluster Algorithm (MCL) (Enright *et al.*, 2002), and the per-family counts were recorded. A time-calibrated ultrametric tree (see later) was generated using R8S v.1.81 (Sanderson, 2003), with outgroup calibration estimated using TIMETREE (http://www.timetree.org). Rapidly evolving gene families were identified using Computational Analysis of gene Family Evolution (CAFÉ v.4.2.1) (De Bie *et al.*, 2006), on the MCL generated family counts with an inflation value of 1.5. Functions for each gene family were assigned by compiling all available annotations including gene ontology (GO) terms (Ashburner *et al.*, 2000), InterPro (Mitchell *et al.*, 2019) UniProt (UniProt Consortium, 2018) and Pfam domains (El-Gebali *et al.*, 2019) for each gene in a given family, with consensus annotation chosen on the basis of term frequency. In the case of ties, both annotations are reported. Global distribution of InterPro domains for all proteins was assessed between *Suillus* and Other ECM fungi using one-sided *t*-tests at $P < 0.001$. For domains significantly over- or underrepresented between *Suillus* and Other ECM fungi, GO terms were assigned using INTERPRO2GO (Camon *et al.*, 2005). GO enrichment was assessed using a Fisher's exact test and the algorithm = 'weight01' to take GO hierarchy into consideration. For the most significant term ('oxidation-reduction process', GO:0055114), all significant InterPro domains falling under the parent term were retrieved from the INTERPRO2GO database using a custom R script and visualized using the R package PHEATMAP v.1.0.12 (https://CRAN.R-project.org/package=pheatmap). Auxiliary Activity enzymes were annotated as in Ruiz-Dueñas *et al.* (2020) and retrieved from the CAZy database via the MycoCosm portal.

To predict SSPs, we applied signalP5 (Armenteros *et al.*, 2019) to screen proteins containing a secretion signal peptide (eukaryote option with default settings). The resultant dataset was then filtered to include only proteins lacking predicted transmembrane helices using TMHMM (Krogh *et al.*, 2001). A custom script was used to filter proteins to those consisting of < 300 amino acids. Ortholog prediction of SSPs and SSSPs was carried out using ORTHOFINDER2 (Emms & Kelly, 2019). Secondary metabolite clusters were identified using ANTISMASH FUNGI 5.0 (Blin *et al.*, 2019). Orthologous SM clusters were assigned using BIGSCAPE (Navarro-Muñoz *et al.*, 2020). GPCRs were identified using a custom pipeline employing Hidden Markov Model (HMM) queries. First, we constructed a catalog of fungal GPCRs that had been previously characterized to class (Table S2). These GPCR sequences were curated from published studies that included either experimental validation or high sequence homology to well-defined fungal GPCR classes. GPCRs representing each class ($n = 2$–9 per class, average 5) were used to construct a database for each class by running all representatives of a given class though PSI-BLAST (Altschul *et al.*, 1997), with a standard hit cut-off of $P < 0.005$, and five search iterations. From the final search iteration for each class, 800 sequences were retained and used to construct HMM models using HMMER (http://hmmer.org). The total proteomes of each species were filtered to contain only proteins with six to eight transmembrane domains using PHOBIUS (Krogh *et al.*, 2001), and the resultant dataset for each species was searched against the HMM for each class. Proteins with matches < $1.0e^{-5}$ were considered a match to that class. If a single protein had a significant match to more than one HMM class, the protein was classified as class 'unknown'.

Statistical analysis

To compare differences in genome size, predicted proteome size, Auxiliary Activity enzymes, SSP and SSSP richness, SM clusters and GPCRs between *Suillus* and Other ECM fungi, we used phylogenetic generalized least squares (PGLS) analysis, which accounts for the phylogenetic structure in our dataset. Differences were tested for significance using 'pgls' in the R package CAPER (Orme *et al.*, 2014). To account for variable phylogenetic signal in the model residuals, λ was optimized for each model using ML (Revell, 2010). PGLS analyses were run twice for each comparison, once comparing *Suillus* with the full set of Other ECM fungal species, and once comparing *Suillus* with the Other ECM set excluding species with high (*Gyrodon lividus*, *Rhizopogon vesiculosis*, *Rhizopogon vinicolor*) and moderate (*Rhizopogon salebrosus*, *Rhizopogon vulgaris*) host specificity. Phylogenetic autocorrelation was assessed with Bloomberg's K using the R package PHYLOSIGNAL (Keck *et al.*, 2016). Phylogenetic signal at internal nodes separating *Suillus* from Other ECM fungi and Suillaceae (*Suillus* and *Rhizopogon*) from Other ECM fungi were assessed using 'phyloSignalINT' with Bloomberg's K, and evidence for local phylogenetic autocorrelation was assessed with Moran's *i* using 'lipaMoran' in phylosignal. Differences in genome size, predicted proteome size, SSP richness and SSSP richness among *Suillus* species associating with different host groups were evaluated using type 1 sum-of-squares ANOVAs, with boxCox testing and log transformations applied when necessary to meet variance assumptions. When significant, differences among group means were determined using Tukey's honestly significant difference at $\alpha < 0.05$. To account for unequal sample sizes among groups, a parallel set of analyses was run using a series of randomization tests. For intrageneric *Suillus* comparison, multifactor randomization tests were implemented using the COIN package in R (Zeileis *et al.*, 2008), at $\alpha = 0.05$. Significant differences between groups were determined using pairwise permutation tests, implemented with the package RCOMPANION with a Benjamini–Hochberg correction for multiple comparisons. Significant differences in SM cluster abundance were assessed for each SM cluster type using multiple *post hoc t*-tests, with a Holm

adjustment for multiple comparisons. All data analysis was carried out in R (R Core Team, 2017). All programing scripts associated with this project are available at: https://github.com/MycoPunk/Suillus_comp_genomics

Results

The *Suillus* genome assemblies had high coverage and were nearly gap-free, with an average depth of coverage of 145, an average of 1254 scaffolds per genome, an average of one gap per genome, an average scaffold N50 of 66.4 and an average L50 of 0.33 Mbp (Table S1). Genome size ranged from 42.34 to 114.21 Mbp (mean = 63.73). The number of gene models ranged from 13 537 to 22 673 (mean = 17 340). Repeat content varied widely from 1.59 to 30.23 Mbp, and was significantly related to genome size, even when accounting for the size contribution of repetitive elements ($R^2 = 0.45$, $P = 1.0e^{-04}$). Neither genome size nor proteome size was found to be significantly different between *Suillus* and Other ECM fungi or among *Suillus* associating with different host groups (Table 2). All genomes had high degrees of genome completeness (Table S3), despite having been sequenced on a variety of sequencing platforms, with BUSCO not significantly different between *Suillus* and Other ECM fungi (Welch's two-sample *t*-test with a mean BUSCO of 95.8% in *Suillus* and 93.8% in Other ECM fungi).

Compared with Other ECM fungi, *Suillus* had significantly higher numbers of both rapidly expanding gene families and rapidly contracting gene families (CAFÉ analysis expansions: $t = 3.913$, df = 44, $P = 3.129e^{-4}$; contractions: $t = -5.1108$, df = 39.831, $P = 8.397e^{-06}$). These results held in the full species dataset as well as when the Other ECM fungal species with high and moderate host specificity were removed (expansions: $t = 3.282$, df = 39, $P = 2.176e^{-3}$, contractions: $t = -4.501$, df = 28.853, $P = 1.022e^{-4}$). *Suillus* had means of 59 expansions and 23 contractions, respectively, vs 46 and seven for the Other ECM fungi (Fig. 1a). For both gene family expansions and contractions among *Suillus* species, differences among host groups were not significant (Fig. 1b). Analysis of rapidly evolving gene families exclusive to *Suillus* (i.e. not rapidly evolving in Other ECM fungi) revealed six rapidly expanding gene families (Fig. 1c) and 14 rapidly contracting gene families represented in at least six species (Fig. 1d). Functional annotations were assigned to 15 of these 20 rapidly evolving gene families. Investigation of InterPro domain abundances between *Suillus* and Other ECM fungi showed a total of 1616 domains that were significantly overrepresented (Table S4) and 769 that were significantly underrepresented ($P < 0.001$; Table S5). Mapping differentially represented domains to GO annotations identified 18 overrepresented GO terms ($P < 0.01$; Table 3). This number was reduced to nine when excluding the Other ECM fungi with high and moderate specificities. In both cases, the most significantly overrepresented GO term in *Suillus* was 'oxidation-reduction process' (GO:0055114) ($P = 7.1e^{-13}$, as compared with all Other ECM fungi; and $P = 6.1e^{-14}$ when excluding Other ECM fungi with high and moderate specificities). There were also three significantly underrepresented GO terms in *Suillus* ('translation' GO:0006412, 'sulfate assimilation' GO:0000103, and 'cell redox homeostasis' GO:0045454), but the only term that remained significant when high- and moderate-specificity Other ECM fungi were removed was cell redox homeostasis (Table 3). Reverse targeting of all InterPro domains under 'oxidation-reduction process' revealed 91 domains that were significantly more abundant in *Suillus* ($P < 0.001$), including multiple domains associated with the detoxification of reactive oxygen species (ROS) such as thioredoxin reductase, pyridine nucleotide-disulfide oxidoreductase, aldehyde dehydrogenases, glutathione, superoxide dismutase and multiple catalases (Aguirre *et al.*, 2006; Morel *et al.*, 2008) (Fig. 2; Table S6). Investigation into the distribution of genes encoding Auxiliary Activity enzymes yielded only one enzyme group that was

Table 2 Comparisons of genomic features of *Suillus* vs Other ECM (ectomycorrhizal) fungi as well as among *Suillus* species with differing host associations.

	Genome size	Proteins	SSPs	SSSPs	Terpenes	NRPS-like	GPCRs
Suillus	62.55 ± 3.68	17307 ± 431.74	394 ± 14.18	112 ± 7.52	23 ± 1.50	12 ± 0.62	28.09 ± 1.06
Other ECM	59.40 ± 5.47	16672 ± 796.65	364 ± 23.98	172 ± 20.59	13 ± 1.29	6 ± 0.87	44.87 ± 3.15
PGLS	ns	ns	ns	ns	df = 44, $F = 6.15$, $P = 0.017$	df = 44, $F = 5.96$, $P = 0.019$	ns
PGLS w/o H/M	ns	ns	ns	ns	df = 39, $F = 29.12$, $P = 3.57e^{-06}$	df = 39, $F = 11.27$, $P = 0.002$	ns
Red pine	61.08 ± 4.51	17388 ± 615.80	370 ± 8.45	95 ± 5.44	24 ± 2.40	11 ± 0.81	25.51 ± 1.11
White pine	69.98 ± 10.40	17870 ± 1115.15	420 ± 33.65	121 ± 15.40	25 ± 2.86	14 ± 1.43	28.32 ± 1.90
Larch	55.72 ± 4.56	16214 ± 342.41	417 ± 46.38	140 ± 11.14	17 ± 3.28	11 ± 2.03	33.83 ± 2.17
ANOVA	ns	ns	ns	df = 2, $F = 4.468$, $P = 0.028$	ns	ns	ns
Randomization	ns	ns	ns	$Z = 2.549$, $P = 0.011$	ns	ns	ns

GPCR, G-protein coupled receptor; NRPS, nonribosomal peptide synthase; ns, not significant; PGLS, phylogenetic generalized least-squares analysis; PGLS w/o H/M, PGLS analysis excluding Other ECM with high and moderate host specificities; SSP, small secreted protein; SSSP, species-specific small secreted proteins.
Values represent means ± 1 SE.

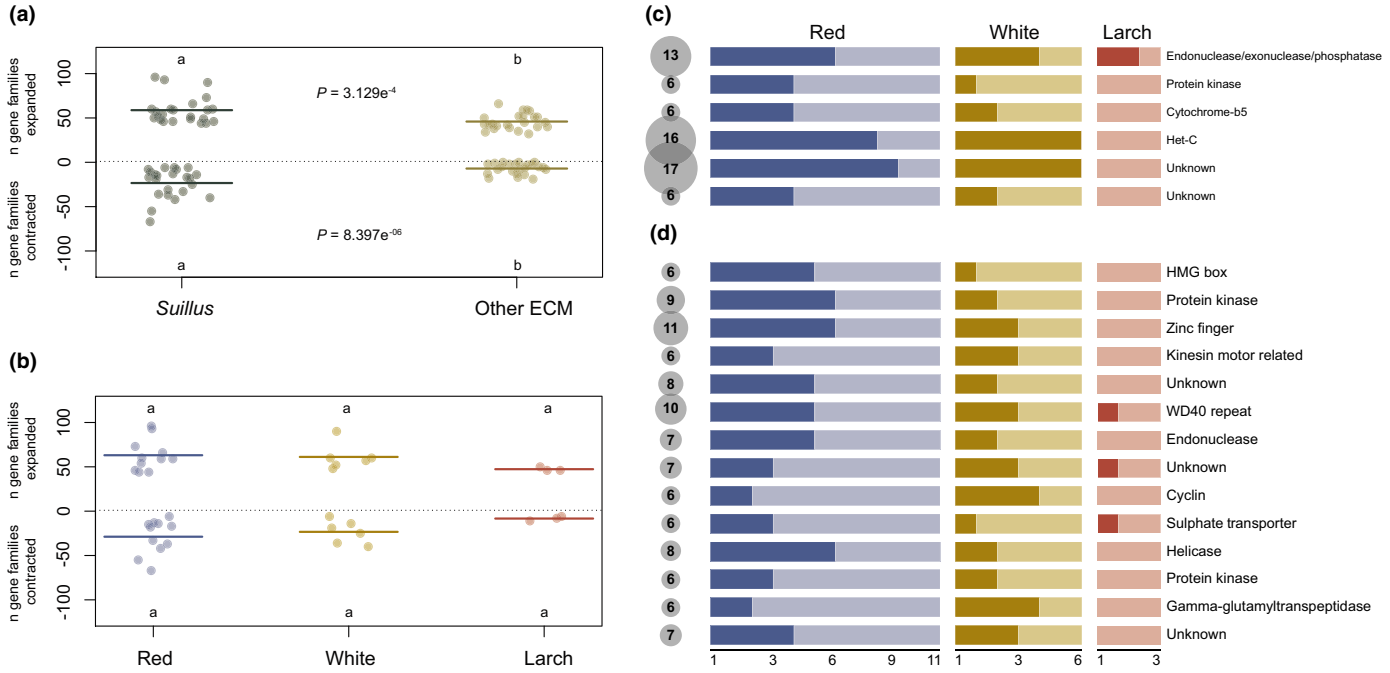

Fig. 1 Gene family expansions and contractions. *Suillus* fungi encode significantly more rapidly evolving gene families than Other ectomycorrhizal (ECM) fungi, and exhibit both expansions and contractions. (a) Rapid gene family expansions and contractions for *Suillus* and Other ECM ('Other ECM' in the figure) fungi computed by CAFÉ (expansions: $t = 3.91$, df = 44, $P = 3.129e^{-4}$; contractions: $t = -5.11$, df = 40, $P = 8.397e^{-06}$). Different letters above (expansions) or below (contractions) the groups indicate significant differences between group means. (b) Rapid gene family expansions and contractions among *Suillus* species associating with different host groups. (c, d) Rapidly expanding (c) and rapidly contracting (d) gene families in *Suillus* that were not rapidly expanding in any of the Other ECM fungi. Only gene families expanding or contracting in more than six species were considered. Shaded bars indicate the number of species that family was expanding or contracting for red pine, white pine and larch hosts, respectively, with the *x*-axis specifying the total number of species associated with that host. Gray circles indicate the number of *Suillus* species in which that gene family was rapidly evolving.

significantly different between *Suillus* and Other ECM fungi: AA3_dist, which encodes GMC oxidoreductase ($P < 0.001$) and was expanded in *Suillus* (Fig. S1).

There were no significant differences in SSP richness or SSSP richness between *Suillus* and Other ECM fungi (Fig. 3a,b), regardless of whether or not high- and moderate-specificity Other ECM fungi were included. SSP richness also did not significantly differ among *Suillus* species (Fig. 3c), but red pine-associated *Suillus* species had significantly lower SSSP abundance than larch-associated *Suillus* species, with white pine-associated *Suillus* species being intermediate (Fig 3d) (ANOVA: df = 2, $F = 4.468$, $P = 0.028$, with Tukey red pine vs larch P-value = 0.039; randomization: $Z = 2.255$, $P = 0.011$, with red pine vs larch P-value = 0.024). Significant phylogenetic signal existed across the phylogeny for both SSPs and SSSPs for both the explanatory variable (SSP: $K = 0.249$, $P < 0.01$; SSSPs: $K = 1.202$, $P < 0.001$) and the residuals (SSP: $K = 0.439$, $P < 0.001$; SSSPs: $K = 1.299$, $P < 0.001$). For SSPs, the phylogenetic signal was significantly associated with the split between Suillaceae and Other ECM fungi ($P < 0.01$), while for SSSPs the signal was between *Suillus* and Other ECM fungi ($P < 0.01$). Local Moran's *i* showed no notable phylogenetic signal between *Suillus* and Other ECM fungi, or within *Suillus*.

The SM cluster analysis categorized core biosynthetic enzymes as belonging to five groups: terpenes, t1pks, NRPS-like terpenes, NRPS-like, NRPS or indole. *Suillus* had significantly higher abundance of terpene SM clusters, with an average of 23 in *Suillus* and 13 in Other ECM fungi (all Other ECM: $P < 0.05$; Other ECM with high- and moderate-specificity species excluded: $P < 0.001$). *Suillus* also had a significantly higher abundance of NRPS-like SM clusters than Other ECM fungi, with an average of 12 in *Suillus* and six in Other ECM fungi (all non-*Suillus* ECM: $P < 0.05$; Other ECM with high- and moderate-specificity species excluded: $P < 0.001$) (Fig. 4a,b). No significant differences were found among *Suillus* fungi associating with different hosts (Fig. 4c,d). Significant phylogenetic signal existed across the phylogeny for both NRPS-like SM clusters and terpene SM clusters in the explanatory variable (NRPS-like: $K = 0.182$, $P < 0.01$; terpenes: $K = 0.221$, $P < 0.01$) and in the residuals for NRPS, but not in the residuals for terpenes (NRPS-like: $K = 1.89$, $P < 0.05$; terpenes: $K = 0.12$, $P > 0.05$) For NRPS-like SM clusters, phylogenetic signal was associated with the split between Suillaceae and Other ECM fungi ($P < 0.01$). For terpenes, the phylogenetic signal was equally associated with the split between Suillaceae and Other ECM fungi and the split between *Suillus* and other EMC fungi (both at $P < 0.01$) (Fig. 4e). Local Moran's *i* showed notable phylogenetic signal between *Suillus* and Other ECM fungi for both NRPS-like SM clusters (Fig 4f) and terpene SM clusters (Fig 4g). The distribution of orthologous SM clusters revealed that terpene clusters were predominantly shared across

Table 3 Significant gene ontology (GO) terms overrepresented and underrepresented in *Suillus* vs other ectomycorrhizal (ECM) fungi for both the full complement of Other ECM fungi and excluding Other ECM fungi with high and moderate host specificities.

GO.ID	Overrepresented: all Other ECM	*P*-value
GO:0055114	Oxidation-reduction process	7.10E-13
GO:0006556	S-adenosylmethionine biosynthetic process	1.00E-07
GO:0006412	Translation	4.20E-07
GO:0009113	Purine nucleobase biosynthetic process	9.00E-06
GO:0006419	Alanyl-tRNA aminoacylation	8.90E-05
GO:0009058	Biosynthetic process	0.00042
GO:0000256	Allantoin catabolic process	0.00042
GO:0000154	rRNA modification	0.0008
GO:0006481	C-terminal protein methylation	0.0008
GO:0006099	Tricarboxylic acid cycle	0.00086
GO:0055085	Transmembrane transport	0.00124
GO:0005975	Carbohydrate metabolic process	0.00133
GO:0006006	Glucose metabolic process	0.00233
GO:0019551	Glutamate catabolic process to 2-oxoglutarate	0.00236
GO:0001522	Pseudouridine synthesis	0.0026
GO:0019915	Lipid storage	0.00464
GO:0006097	Glyoxylate cycle	0.00758
GO:0006891	Intra-Golgi vesicle-mediated transport	0.00758

GO.ID	Overrepresented: Other ECM without high and moderate	*P*-value
GO:0055114	Oxidation-reduction process	6.10E-14
GO:0009058	Biosynthetic process	1.00E-04
GO:0000256	Allantoin catabolic process	0.00016
GO:0000154	rRNA modification	0.00041
GO:0055085	Transmembrane transport	0.00078
GO:0019551	Glutamate catabolic process to 2-oxoglutarate	0.00123
GO:0006099	Tricarboxylic acid cycle	0.00206
GO:0006097	Glyoxylate cycle	0.004
GO:0001522	Pseudouridine synthesis	0.00855

GO.ID	Underrepresented: all Other ECM	*P*-value
GO:0006412	Translation	1.40E-06
GO:0000103	Sulfate assimilation	0.003
GO:0045454	Cell redox homeostasis	0.0047

GO.ID	Underrepresented: Other ECM without high and moderate	*P*-value
GO:0045454	Cell redox homeostasis	0.0012

host groups, with only one unique cluster in white pine-associated *Suillus* species and two in larch-associated *Suillus* species (Fig. 5a). Conversely, five unique NRPS clusters were found in white pine-associated *Suillus* species, with no unique clusters in the red pine- or larch-associated *Suillus* species (Fig. 5b).

G-protein coupled receptor abundance was not significantly different between *Suillus* fungi (mean = 15) and Other ECM fungi (mean = 19), regardless of whether or not Other ECM fungi with high and moderate specificity were excluded (Fig. 6a). GPCRs were identified with significant similarity to 12 of the 14 currently proposed classes (including 19 with similarity to Pth11-like class 14) and 13 GPCRs were classified into more than one class. No significant differences in GPCR abundance were found among red pine- (16), white pine- (15) and larch-associated (16) *Suillus* species. Among *Suillus*, GPCRs had significant similarity to 10 classes, including eight with similarity to class 14, and one classified into more than one class (Fig. 6b). Significant phylogenetic signal existed across the phylogeny for GPCRs in both the explanatory variable ($K = 0.424$, $P < 0.001$) and in the residuals ($K = 0.482$, $P < 0.001$); however, specific significant signal at either the split between Suillaceae and Other ECM fungi or the split between *Suillus* and Other ECM fungi was not detected. Similarly, local Moran's *i* showed no evidence of clustering of phylogenetic signal.

In the STAG analyses, a total of 12 717 protein-based phylogenetic trees were constructed, of which 4728 contained representatives in all species and were used in consensus tree determination (Fig. 7). STAG and PHYLING generated consistent branch topology, with the exception of a single bipartition which grouped *S. brevipes*, *S. occidentalis* and *S. luteus* with *S. placedus* and *S. weaverae* in the STAG tree but separated *S. placedus* and *S. weaverae* in the PHYLING tree. Notably, this was the only poorly supported bipartition in the PHYLING tree (ML = 20.1/SH-aLRT = 56) whereas all other ML bipartitions had greater than 90% branch support (Fig 7.) It should be noted that the branch support values generated for the STAG and PHYLING trees, respectively, represent fundamentally different metrics and are therefore not directly comparable. Branch support on the PHYLING tree represents ML bootstrap values, whereas branch support on the STAG tree represents the percentage of the 12 717 independent gene trees that support that bipartition. The three larch (*Larix*)-associated *Suillus* species (*S. clintonianus*, *S. ampliporus* and *S. paluster*) clustered on low nodes of the tree, but were not monophyletic. Ancestral state reconstruction supported larch as the ancestral host of the genus *Suillus*. Larch-associated ancestors gave rise to a single independent origin for red pine-associated *Suillus* species, while white pine-associated *Suillus* species evolved via three independent host-switching events from red pine onto white pine, with support for one reversion (in the *S. tomentosus* clade) from white pine back to red pine.

Discussion

Contrary to our expectation that reduced host range would result in consistent gene losses, we found that the ECM fungal host specialist genus *Suillus* has numerous rapidly evolving gene families, representing both gene family expansions and contractions, as well as an unexpected number of protein-coding domain expansions. Targeted analysis of molecular classes previously identified as important in the structuring of host specificity relationships for fungi supported significant enrichment of terpene- and NRPS-like SM clusters. Conversely, we found no evidence of significant differences in the abundance of SSPs, or the abundance or class distribution for GPCRs between *Suillus* and Other ECM fungi. Additionally, the comparisons among *Suillus* species specializing on different host groups identified no genomic differences among the three classes of molecules investigated, with the possible exception of SSSPs, which showed a loss of richness in red pine-associated species relative to larch-associated species.

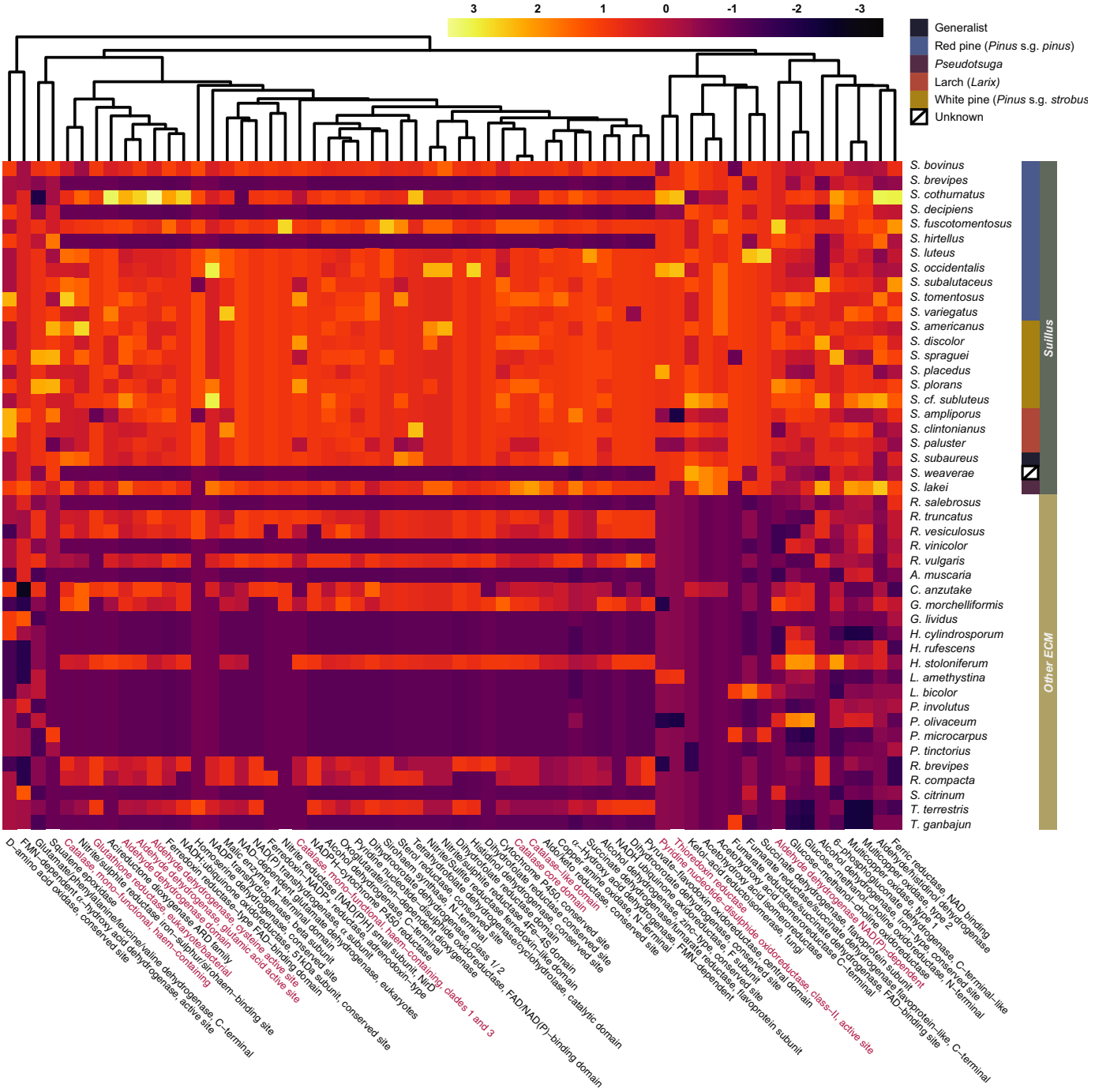

Fig. 2 Redox-related InterPro domain counts for *Suillus* and Othe rectomycorrhizal (ECM) fungi. Differences in domain abundance highlight reactive oxygen species (ROS)-related pathways in host-specific *Suillus*. The most significantly different InterPro domain annotations were determined by *t*-tests at $P > 0.001$. GO enrichment of these domains identified the most significant parent term as 'oxidation-reduction process' (GO:0055114, $P = 7.1e^{-13}$). Reverse mapping the parent term back onto the significantly different domains identified 91 overrepresented redox-related domains in *Suillus*. Those with $P < 0.0005$ are displayed ($n = 62$). Count data normalized to zero. Nonnormalized counts for all species for all 91 domains can be found in Supporting Information Table S6. Domain names highlighted in pink are of particular interest and are addressed in the Discussion section.

Interestingly, the most significant GO term (GO:0055114) associated with the overrepresented domains in *Suillus* was related to redox processes. Further investigation of the domains under this parent term revealed a diversity of domains associated with oxidative stress, such as thioredoxin reductase, glutathione, superoxide dismutase, catalases, pyridine nucleotide-disulfide oxidoreductase and multiple aldehyde dehydrogenases which are coupled to the generation and detoxification of ROS primarily via lipid peroxidation (Sato *et al.*, 2009; Singh *et al.*, 2013; Xiong *et al.*, 2013). It is known that plant-produced ROS compounds

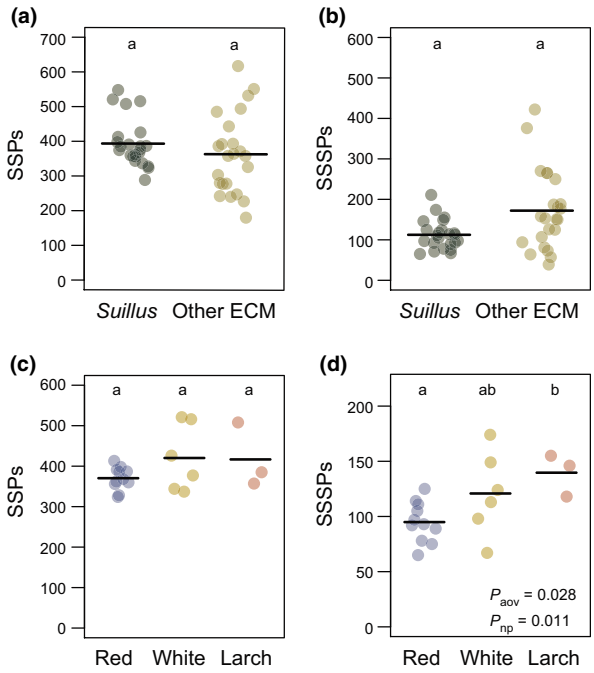

Fig. 3 Small secreted protein (SSP) distributions for *Suillus* ectomycorrhizal (ECM) and Other ECM ('Other ECM' in the figure) fungi. *Suillus* do not encode a significantly different diversity of SSPs compared with Other ECM fungi. However, species-specific small secreted proteins (SSSPs) are less diverse in red pine-associated *Suillus* than in larch-associated species. Different letters indicate significant differences between groups. (a, b) *Suillus* vs Other ECM fungi. (c, d) *Suillus* compared by host association. Paov, *P*-value ANOVA test; Pnp, *P*-value nonparametic (randomization) test.

play an important role in plant defense during colonization by fungal pathogens (Segal & Wilson, 2018) and many of the domains identified here are associated with ROS deactivation in the plant apoplast (Nogueira-Lopez *et al.*, 2018). It is also known that plants produce ROS when exposed to capable ECM strains, but do not produce ROS when exposed to incompetent ECM isolates that are incapable of undergoing mycorrhization (Gafur *et al.*, 2004). Additionally, ROS pathways are known to alter auxin signaling in plants (Zhao *et al.*, 2012), and these processes may be interconnected in ECM fungi. For example, in addition to their role in ROS deactivation, aldehyde dehydrogenases also catalyze the last step in the production of fungal-derived IAA, an auxin phytohormone that initiates essential changes to root architecture and Hartig net formation during ECM colonization (Vayssières *et al.*, 2015; Krause *et al.*, 2015). Further, IAA is upregulated during ECM associations with compatible, but not incompatible, hosts in the host specialist *Tricholoma vaccinum* (Krause *et al.*, 2015). Interestingly, the only significant GO term for domain contractions in *Suillus* was GO:0045454 'cell redox homeostasis', implying that redox evolution in *Suillus* may also include adaptive losses. Although ECM fungi have largely lost the genes for plant cell wall-degrading enzymes found in decay fungi (Kohler *et al.*, 2015; Miyauchi *et al.*, 2020), the ability to decompose soil organic matter via nonenzymatic oxidation is thought to be widespread (Shah *et al.*, 2016; Hess *et al.*, 2018). This capacity allows for the acquisition of nitrogen via Fenton chemistry and has been characterized in other members of the order Boletales, such as *Paxillus involutus* (Op De Beeck *et al.*, 2018; Nicolás *et al.*, 2019). To investigate the potential for the redox-related gene expansions in *Suillus* to play a role in decomposition, we looked at the distribution of Auxiliary Activity enzymes, which function in the oxidative breakdown of organic matter (Levasseur *et al.*, 2013). We found that out of 20 annotated Auxiliary Activity enzymes, only GMC oxidoreductase was significantly different between *Suillus* and Other ECM fungi, appearing between one and six times in all but four species of *Suillus*, and absent in all Other ECM fungi except the host specialist *G. lividus* (Fig. S1). These results suggest that the redox-related genes expanded in *Suillus* may have functions apart from those involved in decomposition, and hint at targeted ROS deactivation as a possible mechanism mediating the enhanced host specificity present in this fungal genus.

We found a lower (but not significant) number of SSSPs in *Suillus* than in Other ECM fungi. Given the expected sequence similarity between more closely related species, however, we expected that *Suillus* would have a smaller number of SSSPs than a more diverse group of ECM fungi simply as a result of phylogenetic conservatism. A true measure of relative SSSP abundance for a given genus will require calling SSSPs relative to a dataset of multiple genera with well-represented species in each genus and where each genus represents comparable total intergenic patristic distance. SSSPs were significantly more abundant in larch-associated species than in red pine-associated *Suillus* species. Although this trend could be influenced by the comparatively low number of larch-associated species in our dataset, it may also be reflective of the relaxed host switching observed between red and white pine-, but not larch-associated *Suillus*.

The SM cluster enrichment encountered in *Suillus* relative to Other ECM fungi included both genes encoding terpenes and NRPS-like enzymatic cores. Terpene enrichment is consistent with host–fungal communication and may relate to host specificity, as terpenes have been found to play critical roles in the process of recognition and response among fungi, bacteria, plants and insects (Zhao *et al.*, 2012). Basidiomycete fungi primarily produce sesqui-, di- and triterpenes (Quin *et al.*, 2014), while many plant hosts, notably pines, produce a large number of monoterpenes that can inhibit fungal growth (Melin & Krupa, 1971; Huber & Bohlmann, 2006). ECM-derived sesquiterpenes are associated with lateral root development and increased substrate availability and may be of particular importance to the process of mycorrhization (Ditengou *et al.*, 2015). Future work on the identification and classification of ECM terpene genes would benefit from a high-fidelity method specific to this class of enzymes, such as that used by Quin *et al.* (2013) to identify sesquiterpene-encoding SM clusters. Unlike terpenes, the functions of fungal NRPS-like gene products are largely unknown. The most well-studied fungal NRPS genes are recognized as virulence factors (such as HC toxin in *Cochliobolus*) and are found at higher abundance in pathogens than in other fungal lifestyles (Dunkle *et al.*, 1991; Yoder & Turgeon, 2001). However, NRPS

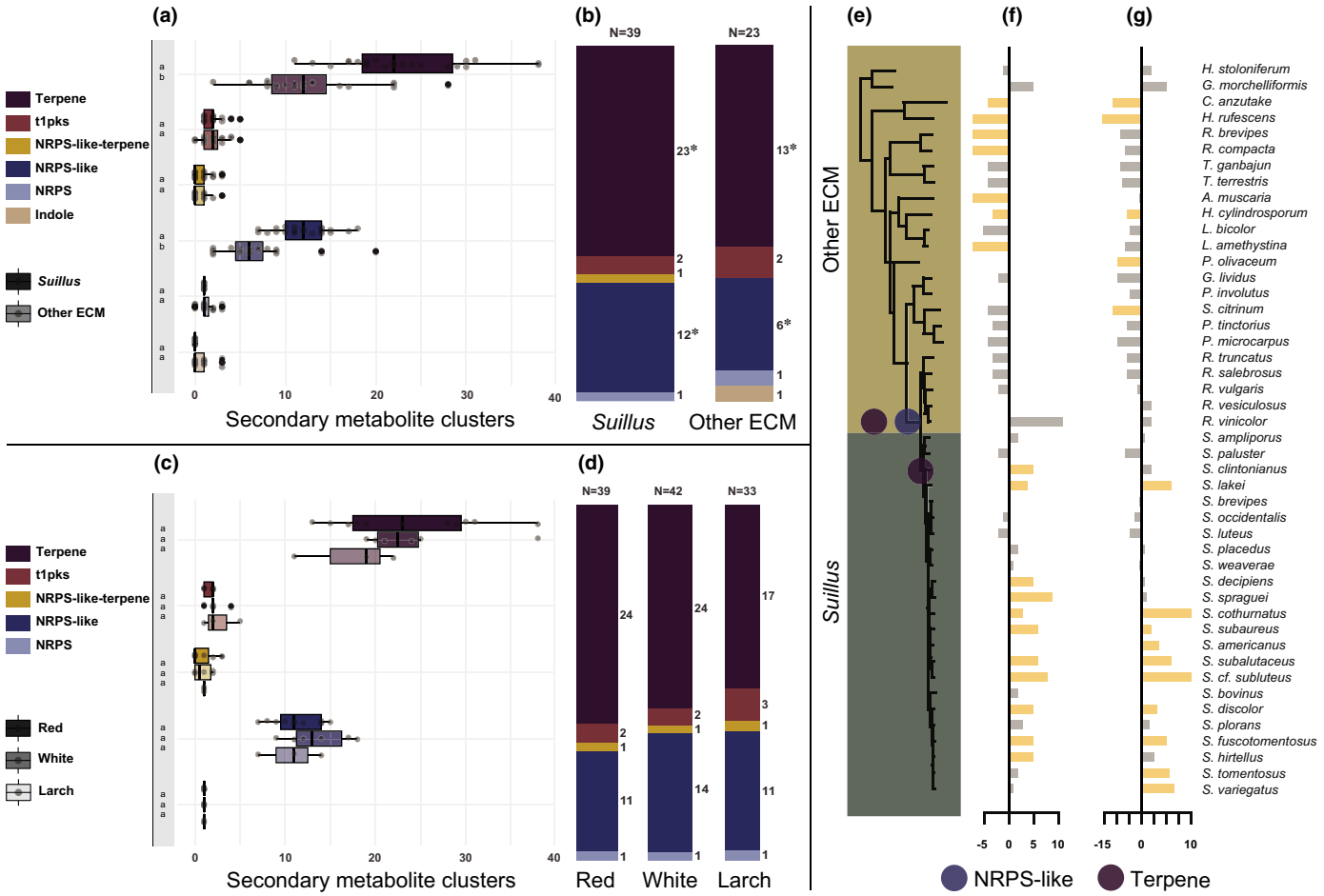

Fig. 4 Secondary metabolite cluster identity and distributions for *Suillus* ectomycorrhizal (ECM) and Other ECM ('Other ECM' in the figure) fungi. *Suillus* have significant enrichment of terpene and nonribosomal peptide synthase (NRPS)-like secondary metabolite clusters compared with Other ECM fungi. (a, b) *Suillus* vs Other ECM fungi; (c, d) *Suillus* species associated with different host groups. (a, c) Box plots of secondary metabolite (SM) cluster distributions represent the interquartile region (IQR) intersected by the median. Whiskers extend to the most extreme data point, no more than 1.5 × IQR. (b, d) Spine plots displaying mean abundance of all (*N*) and each (to the right of the graph) SM cluster group between *Suillus* and Other ECM fungi. Significantly different groups are highlighted with an asterisk. (e) For NRPS-like SM clusters, significant phylogenetic autocorrelation was detected at the node delineating the separation between Suillaceae (*Suillus* and *Rhizopogon*) and Other ECM fungi. For terpenes, significant phylogenetic autocorrelation was detected at both the node delineating the separation Suillaceae and the node delineating *Suillus* and Other ECM fungi. (f, g) Local Moran's *i* showed clustered phylogenetic signal for both NRPS-like (f) and terpene (g) SM clusters, with *Suillus* generally displaying positive signal, and Other ECM fungi generally displaying negative signal.

and NRPS-like genes have also been shown to act as anti-herbivory agents in the symbiotic fungus *Epichloë* (Tanaka *et al.*, 2005) and are involved in fungal siderophore production, and iron acquisition from host tissues (Oide & Turgeon, 2020). Although a specific role for NRPS-like genes in host specificity has not been identified, our observation of higher diversity in host-specialized *Suillus* compared with Other ECM fungal species suggests that SMs may play a role in structuring host specificity, and represents an intriguing future research direction.

Although the precise role of GPCRs in ECM mycorrhization remains unclear, G-protein signaling is well established as a primary system for communication both between microbes and between microbes and hosts (Hughes & Sperandio, 2008; Brown *et al.*, 2018; Dierking & Pita, 2020). Lacking a bioinformatic tool capable of classifying newly characterized fungal-specific GPCRs, we employed a novel identification approach based on protein similarity to one of 14 HMM models, each built on a literature-curated database of fungal-specific GPCRs. This method was tested using leave-one-out cross-validation with a classification accuracy of > 98%, for all classes except classes 3 and 5, which were conflated with one another and called incorrectly *c.* 4-% of the time. This result, coupled with the high number of hits to class 2 GPCRs, and the high sequence similarity of a limited number of GPCRs in class 11 (thought to be exclusive to the fungal subphylum Pezizomycotina (Brown *et al.*, 2018)), highlight the necessity for a systematic review of GPCR diversity in basidiomycete fungi, and a universal classification scheme for fungal GPCR diversity more generally.

Given previous results linking differential regulation of SSPs and GPCRs with successful mycorrhization in *Suillus* (Liao *et al.*, 2016), it is likely that these classes play a role in successful mycorrhization in general rather than acting in a host-specific manner.

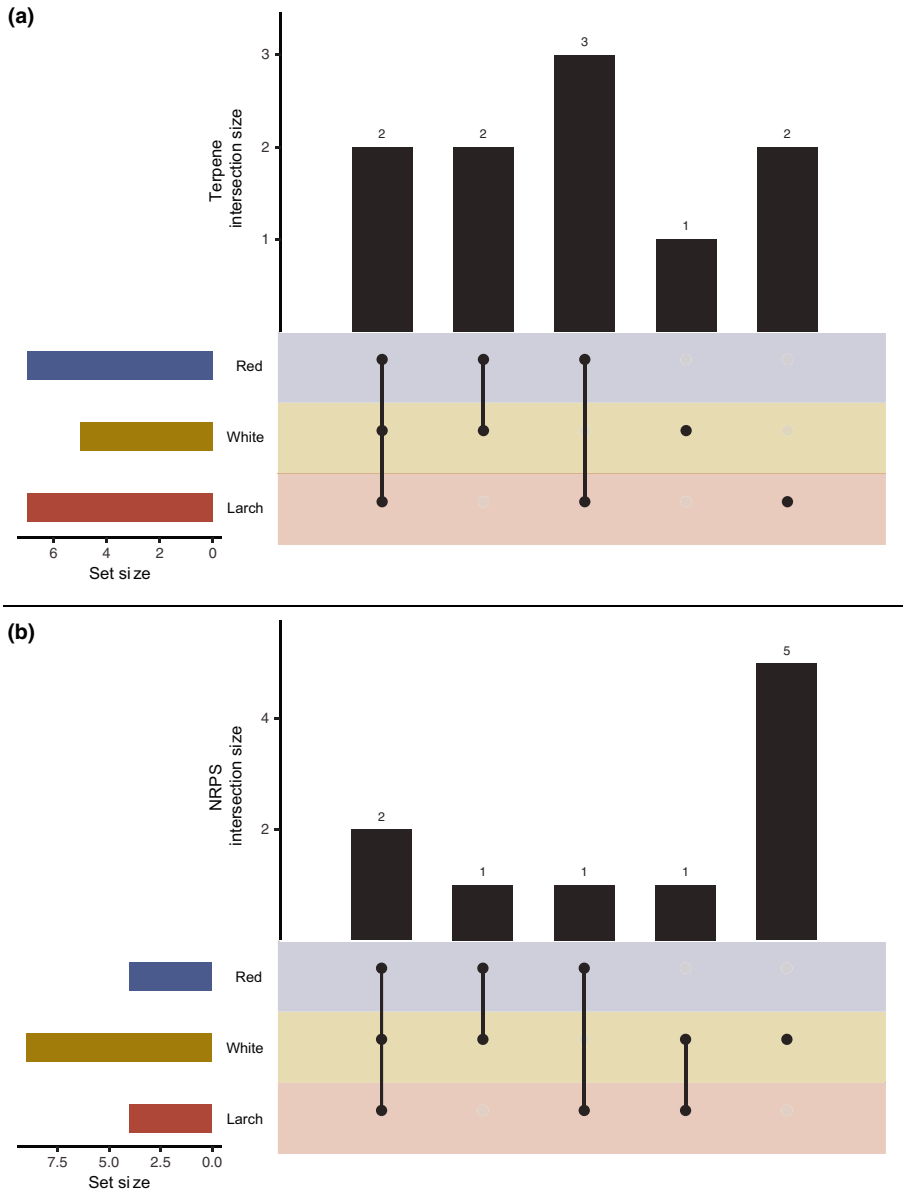

Fig. 5 Distributions of orthologous secondary metabolite clusters with significant differences among *Suillus* species associated with red pine, white pine, and larch hosts. Orthologous cluster analysis highlights a lack of unique terpene-related clusters in larch-associated *Suillus*, and a diversification in nonribosomal peptide synthase (NRPS)-related clusters for white pine-associated *Suillus*. Overlap of unique orthologous clusters by host group for (a) terpene-related clusters and (b) NRPS-related clusters.

Alternatively, if these molecular classes are capable of mediating successful mycorrhization in a host-specific manner, they appear to do so as a result of modified transcriptional regulation rather than gene diversification. For the SM clusters and the oxidative stress pathways identified in the global analysis, follow-up studies are needed to determine the precise role that these classes play. Along with the unique host specificity observed in *Suillus*, some species in the genus also exhibit unique profiles of metal tolerance (Ruytinx *et al.*, 2013), where copy number variation and single nucleotide polymorphisms in oxidative stress pathways have been linked to metal-adapted populations (Bazzicalupo *et al.*, 2020). Similarly, the functional role of terpenes and NRPS-like secondary metabolites varies widely across the fungal phylogeny (Keller *et al.*, 2005). However, given the strong phylogenetic signal of these classes within host-specialist *Suillus*, and the documented importance of these classes in fungal–host signaling pathways (Xue *et al.*, 2008; Rohlfs & Churchill, 2011; Kües *et al.*, 2018), future work should prioritize the potential of SMs and oxidative stress responses in structuring host-specific mycorrhization with a focus on functional analysis.

It has long been thought that high host specificity is driven by ecological trade-offs connected to resource specialization (MacArthur & Levins, 1964; Whittaker & Feeny, 1971). This argument assumes that maintaining access to diverse resources can only be accomplished at the sacrifice of performance, which in turn selects for an optimized state of derived host specialization

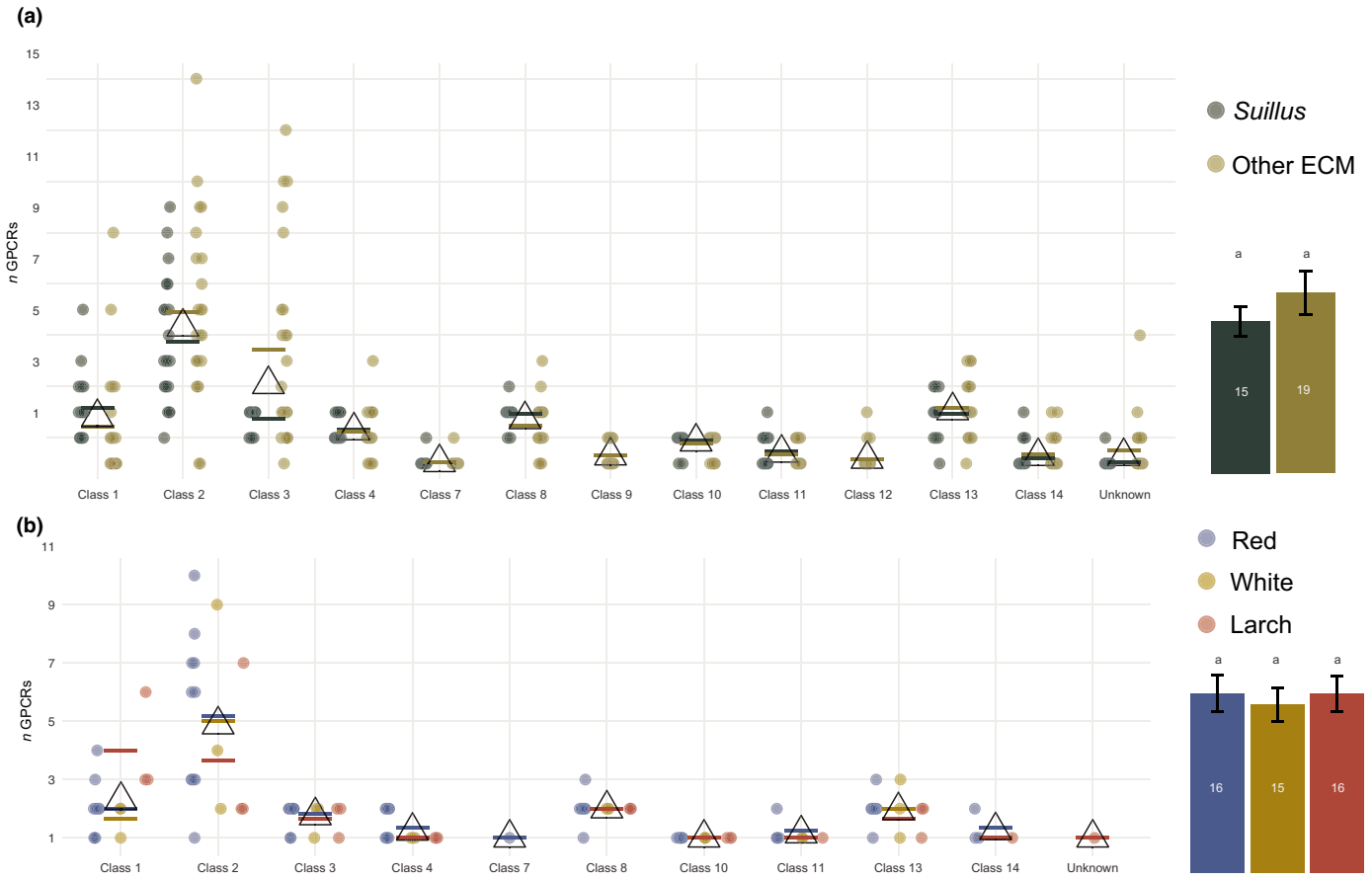

Fig. 6 G-protein coupled receptor (GPCR) distributions for *Suillus* ectomycorrhizal (ECM) and Other ECM ('Other ECM' in the figure) fungi. GPCR abundance or identity was not significantly different for either inter- or intrageneric comparisons. (a, b) Distribution of GPCRs by class for *Suillus* vs Other ECM fungi (a) and among *Suillus* species with different host associations (b). Horizontal lines indicate the mean of each group for a given class, and triangles represent the total mean for each class. Bar graphs represent the mean, and error bars represent the mean ± 1 SD.

(Huey, 1984; Bruns *et al.*, 2002). Because gene loss is assumed to be permanent and host range contractions are often associated with genetic losses (Spanu *et al.*, 2010; Visser *et al.*, 2010; Baroncelli *et al.*, 2016), specialization was long considered to be an evolutionary irreversible state (Simpson, 1953; Moran, 2002). However, numerous examples have indicated bidirectional transitions between generalist and specialist lineages, and it is now accepted that high host specificity is neither universally derived nor an innately optimized resource acquisition strategy (Appel & Martin, 1992; Desdevises *et al.*, 2002; Stireman, 2005; Hardy & Otto, 2014; Ouvrard *et al.*, 2015). Our analyses included one example of a *Suillus* species capable of colonizing multiple host genera: *Suillus subaureus* is known to colonize white pine, larch and oak (Lofgren *et al.*, 2018). Our ancestral state reconstruction supports an evolutionary trajectory where *S. subaureus* was derived from white pine-associated ancestors, making its ability to colonize multiple host groups an example of host expansion. Taken together, our results support the conclusion that host range contractions are not obligately associated with genetic losses, and that specialization on a given host is not an evolutionary irreversible state, as evidenced by many host-switching events, one host expansion event, and one reversion between host groups in the *Suillus* phylogeny.

In our dataset, which was based largely on public genome availability, Other ECM fungi exhibiting high or moderate host specificity were limited to *G. lividus*, and four species of *Rhizopogon* (closely related to *Suillus*). This small number of Other ECM species with enhanced host specificity notably limits our ability to detect global genomic correlates of specificity. That said, our results did not change when species with high and moderate host specificities were excluded from the Other ECM group. This suggests that *Suillus* fungi may display unique patterns of dynamic genome evolution relative to Other ECM fungal host specialists. Better resolution of the extent to which there are common pathways structuring host specificity across diverse ECM lineages will require additional genome sequencing projects targeting other host specialist groups, such as the phylogenetically diverse range of ECM fungal host specialists associated with *Alnus* (Kennedy & Hill, 2010) as well as a targeting of ECM fungal genera that contain a more even balance of host specialists and generalists (such as *Lactarius*). Despite this limitation, our work suggests that ECM fungi, like their pathogenic counterparts, may have a number of specificity-related traits imprinted on their genomes. As such, this study adds to a rapidly growing body of work linking fungal genomic architecture and ecological lifestyle (Floudas *et al.*, 2012; Kohler *et al.*, 2015; Lofgren *et al.*,

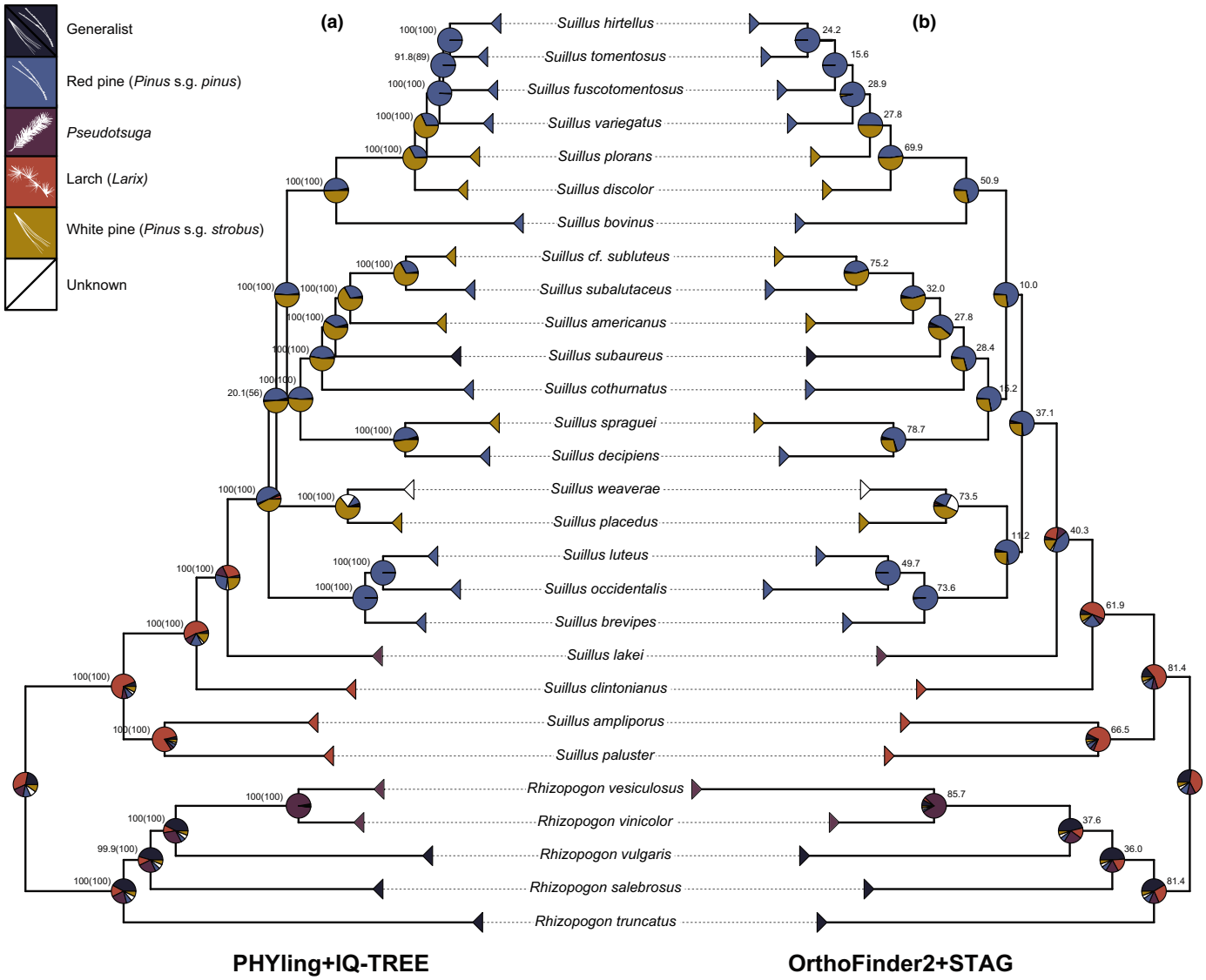

Fig. 7 Phylogenetic reconstruction of host ancestral state and switching within the ectomycorrhizal (ECM) fungal genus *Suillus*. Ancestral state reconstruction supports larch (*Larix*) as the ancestral host within the genus and multiple independent host-switching events between red and white pine hosts. Pie charts represent the posterior probabilities of ancestral host association at each internal node. (a) PHYling + IQ-Tree tree. Branch support values represent maximum likelihood bootstrap support values over 1000 iterations, support values in parenthesis represent Shimodaira–Hasegawa approximate likelihood ratio test values over 1000 iterations. (b) OrthoFinder2 + Stag tree. Branch support values represent the percentage of bipartitions of individual gene trees that support that bipartition.

2019; Miyauchi *et al.*, 2020). In order to assess the generality of these findings more fully, it will also be important to look at similar patterns in other groups of plant-associated fungi, such as endophytes (Knapp *et al.*, 2018). Fortunately, the rate of fungal genome generation and public release continues to increase rapidly (Grigoriev *et al.*, 2014), making knowledge of the genetic mechanisms defining fungal compatibility and host range across diverse ecological lifestyles readily obtainable.

Acknowledgements

We thank all members of the *Suillus* Consortium for use of genome sequences, encouragement and project feedback, staff at the Minnesota Super Computing Institute for computational assistance, H. Nielsen and the Signal P team, B. Stielow for use of genome data for the species *G. morchelliformis* and *H. stonoliferum*, J. Xu for use of genome data for the species *T. ganbajun*, F. Martin for use of genome data for the species *T. terrestris*, *C. anzutake*, *H. rufescens*, *G. lividus*, and B. Looney for use of genome data for the species *R. compacta* and *R. brevipies*, B. Henrissat for the annotation and curation of Auxiliary Activity enzymes via the CAZy database. Special thanks to A. Certano and K-H. Chen for extracting genomic DNA and RNA for several of the species used in this study. Comparative analyses were performed on the High-Performance Computing Cluster at the University of California, Riverside,

in the Institute of Integrative Genome Biology, supported by NSF DBI-1429826 and NIH S10-OD016290. JES is a CIFAR fellow in the Fungal Kingdom: Threats and Opportunities program. The work conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the US Department of Energy under contract no. DE-AC02-05CH11231. Project support was provided by a National Science Foundation Division of Environmental Biology grant (no. 1554375) to PGK and RV.

Authors' contributions

LAL and PGK planned the project and wrote the manuscript. AK wrote the supplementary methods for genome sequencing and assembly. RV, NHN, JR, H-LL, SB, IVG and JES provided editorial input for the manuscript. LAL carried out comparative genomics and statistical analysis. LAL and JES carried out phylogenomics. HH and HN sequenced genomes and transcriptomes. AL assembled transcriptomes. KL, WA, JP and RR assembled genomes. AK annotated genomes. KB and IVG coordinated sequencing and analysis at JGI.

ORCID

William Andreopoulos https://orcid.org/0000-0001-9097-1123
Kerrie Barry https://orcid.org/0000-0002-8999-6785
Sara Branco https://orcid.org/0000-0002-9890-933X
Igor V. Grigoriev https://orcid.org/0000-0002-3136-8903
Hope Hundley https://orcid.org/0000-0001-5071-4587
Peter G. Kennedy https://orcid.org/0000-0003-2615-3892
Alan Kuo https://orcid.org/0000-0003-3514-3530
Kurt LaButti https://orcid.org/0000-0002-5838-1972
Hui-Ling Liao https://orcid.org/0000-0002-1648-3444
Anna Lipzen https://orcid.org/0000-0003-2293-9329
Lotus A. Lofgren https://orcid.org/0000-0002-0632-102X
Hyunsoo Na https://orcid.org/0000-0002-4386-017X
Nhu H. Nguyen https://orcid.org/0000-0001-8276-7042
Robert Riley https://orcid.org/0000-0003-0224-0975
Joske Ruytinx https://orcid.org/0000-0001-5835-3704
Jason E. Stajich https://orcid.org/0000-0002-7591-0020
Rytas Vilgalys https://orcid.org/0000-0001-8299-3605

Data availability

Complete metadata for all species are publicly available on JGI's Mycocosm Portal at https://mycocosm.jgi.doe.gov. *Suillus* genomes can be accessed directly at the JGI *Suillus* portal at https://mycocosm.jgi.doe.gov/mycocosm/home/releases?flt=suillus. All newly sequenced genomes have been deposited into GenBank: accession numbers for each sequencing project can be found in Table S1. All programing scripts associated with this project are available at: https://github.com/MycoPunk/Suillus_comp_genomics (https://doi.org/10.5281/zenodo.4321603).

References

Agerer R. 2001. Exploration types of ectomycorrhizae A proposal to classify ectomycorrhizal mycelial systems according to their patterns of differentiation and putative ecological importance. *Mycorrhiza* **11**: 107–114.

Aguirre J, Hansberg W, Navarro R. 2006. Fungal responses to reactive oxygen species. *Medical Mycology* **44**: S101–S107.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389–3402.

Appel HM, Martin MM. 1992. Significance of metabolic load in the evolution of host specificity of *Manduca sexta*. *Ecology* **73**: 216–228.

Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**: 420–423.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al.* 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**: 25–29.

Baroncelli R, Amby DB, Zapparata A, Sarrocco S, Vannacci G, Le Floch G, Harrison RJ, Holub E, Sukno SA, Sreenivasaprasad S *et al.* 2016. Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. *BMC Genomics* **17**: 1–17.

Bazzicalupo AL, Ruytinx J, Ke Y-H, Coninx L, Colpaert JV, Nguyen NH, Vilgalys R, Branco S. 2020. Fungal heavy metal adaptation through single nucleotide polymorphisms and copy-number variation. *Molecular Ecology* **29**: 4157–4169.

Bergemann SE, Miller SL. 2002. Size, distribution, and persistence of genets in local populations of the late-stage ectomycorrhizal basidiomycete, *Russula brevipes*. *New Phytologist* **156**: 313–320.

Bhatt R, Vishwakarma M, Singh U, Joshi S. 2014. Macrofungal diversity in Adwani forest of Garhwal Himalaya, Uttarakhand. **44**: 16.

Bidartondo MI, Ek H, Wallander H, Soderstrom B. 2001. Do nutrient additions alter carbon sink strength of ectomycorrhizal fungi? *New Phytologist* **151**: 543–550.

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research* **47**: W81–W87.

Brakhage AA. 2013. Regulation of fungal secondary metabolism. *Nature Reviews Microbiology* **11**: 21–32.

Brown NA, Schrevens S, Van Dijck P, Goldman GH. 2018. Fungal G-protein-coupled receptors: Mediators of pathogenesis and targets for disease control. *Nature Microbiology* **3**: 402–414.

Brunner I, Amiet R, Zollinger M, Egli S. 1992. Ectomycorrhizal syntheses with *Picea abies* and three fungal species: a case study on the use of an in vitro technique to identify naturally occurring ectomycorrhizae. *Mycorrhiza* **2**: 89–96.

Bruns TD, Bidartondo MI, Taylor DL. 2002. Host specificity in ectomycorrhizal communities: what do the exceptions tell us? *Integrative and Comparative Biology* **42**: 352–359.

Camon EB, Barrell DG, Dimmer EC, Lee V, Magrane M, Maslen J, Binns D, Apweiler R. 2005. An evaluation of GO annotation retrieval for BioCreAtIvE and GOA. *BMC Bioinformatics* **6**: S17.

Castellano MA. 1999. Hysterangium. In: Cairney JWG, Chambers SM, eds. *Ectomycorrhizal fungi key genera in profile*. Berlin/Heidelberg, Germany: Springer, 311–323.

Collemare J, Lebrun M-H. 2011. Fungal secondary metabolites: ancient toxins and novel effectors in plant-microbe interactions. In: Martin F, Kamoun S, eds. *Effectors in plant-microbe interactions*. Hoboken, NJ, USA: Wiley-Blackwell, 377–400.

Contreras-Cornejo HA, Macías-Rodríguez L, del-Val E, Larsen J. 2016. Ecological functions of *Trichoderma* spp. and their secondary metabolites in the rhizosphere: interactions with plants. *FEMS Microbiology Ecology* **92**: fiw036.

Dahlberg A, Stenlid J. 1994. Size, distribution and biomass of genets in populations of *Suillus bovinus* (L.: Fr.) Roussel revealed by somatic incompatibility. *New Phytologist* **128**: 225–234.

De Bie T, Cristianini N, Demuth JP, Hahn MW. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269–1271.

Desdevises Y, Morand S, Legendre P. 2002. Evolution and determinants of host specificity in the genus *Lamellodiscus* (Monogenea). *Biological Journal of the Linnean Society* **77**: 431–443.

Dickie IA, Bolstridge N, Cooper JA, Peltzer DA. 2010. Co-invasion by Pinus and its mycorrhizal fungi. *New Phytologist* **187**: 475–484.

Diédhiou AG, Selosse M-A, Galiana A, Diabaté M, Dreyfus B, Bâ AM, Faria SMD, Béna G. 2010. Multi-host ectomycorrhizal fungi are predominant in a Guinean tropical rainforest and shared between canopy trees and seedlings. *Environmental Microbiology* **12**: 2219–2232.

Dierking K, Pita L. 2020. Receptors mediating host-microbiota communication in the metaorganism: the invertebrate perspective. *Frontiers in Immunology* **11**: 1251.

Ditengou FA, Müller A, Rosenkranz M, Felten J, Lasok H, van Doorn MM, Legué V, Palme K, Schnitzler J-P, Polle A. 2015. Volatile signalling by sesquiterpenes from ectomycorrhizal fungi reprogrammes root architecture. *Nature Communications* **6**: 6279.

Dong S, Stam R, Cano LM, Song J, Sklenar J, Yoshida K, Bozkurt TO, Oliva R, Liu Z, Tian M *et al.* 2014. Effector specialization in a lineage of the Irish potato famine pathogen. *Science* **343**: 552–555.

Dunk CW, Lebel T, Keane PJ. 2012. Characterisation of ectomycorrhizal formation by the exotic fungus *Amanita muscaria* with *Nothofagus cunninghamii* in Victoria, Australia. *Mycorrhiza* **22**: 135–147.

Dunkle LD, Cantone FA, Ciuffetti LM. 1991. Accumulation of host-specific toxin produced by *Cochliobolus carbonum* during pathogenesis of maize. *Physiological and Molecular Plant Pathology* **38**: 265–273.

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A *et al.* 2019. The Pfam protein families database in 2019. *Nucleic Acids Research* **47**: D427–D432.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* **16**: 157.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**: 238.

Emms EM, Kelly S. 2018. STAG: Species Tree Inference from All Genes. bioRxiv: 267914. doi: 10.1101/267914

Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**: 1575–1584.

Feng B, Wang X-H, Ratkowsky D, Gates G, Lee SS, Grebenc T, Yang ZL. 2016. Multilocus phylogenetic analyses reveal unexpected abundant diversity and significant disjunct distribution pattern of the Hedgehog Mushrooms (*Hydnum* L.). *Scientific Reports* **6**: 25586.

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS *et al.* 2012. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**: 1715–1719.

Gafur A, Schützendübel A, Langenfeld-Heyser R, Fritz E, Polle A. 2004. Compatible and incompetent *Paxillus involutus* isolates for ectomycorrhiza formation in vitro with poplar (*Populus* x *canescens*) differ in H2O2 production. *Plant Biology* **6**: 91–99.

Gao Q, Jin K, Ying SH, Zhang Y, Xiao G, Shang Y, Duan Z, Hu X, Xie XQ, Zhou G *et al.* 2011. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genetics* **7**: e1001264.

Gilbert GS, Webb CO. 2007. Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Sciences, USA* **104**: 4979–4983.

Gladieux P, Ropars J, Badouin H, Branca A, Aguileta G, de Vienne DM, de la Vega RCR, Branco S, Giraud T. 2014. Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Molecular Ecology* **23**: 753–773.

Grebenc T, Martín MP, Kraigher H. 2009. Ribosomal ITS diversity among the European species of the genus *Hydnum* (Hydnaceae). *Anales del Jardín Botánico de Madrid* **66**: 121–132.

Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F *et al.* 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic acids research* **42**: D699–D704.

Grubisha LC, Trappe JM, Molina R, Spatafora JW. 2002. Biology of the ectomycorrhizal genus *Rhizopogon*. VI. Re-examination of infrageneric relationships inferred from phylogenetic analyses of ITS sequences. *Mycologia* **94**: 607–619.

Hardy NB, Otto SP. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society B: Biological Sciences* **281**: 20132960.

Haridas S, Albert R, Binder M, Bloem J, LaButti K, Salamov A, Andreopoulos B, Baker SE, Barry K, Bills G *et al.* 2020. 101 Dothideomycetes genomes: A test case for predicting lifestyles and emergence of pathogens. *Studies in Mycology* **96**: 141–153.

Hayward D, Thiers HD. 1984. *Gyrodon Lividus* in California. *Mycologia* **76**: 573–575.

Hazard C, Kruitbos L, Davidson H, Mbow FT, Taylor AFS, Johnson D. 2017. Strain identity of the ectomycorrhizal fungus *Laccaria bicolor* is more important than richness in regulating plant and fungal performance under nutrient rich conditions. *Frontiers in Microbiology* **8**: 1874.

Hedh J, Johansson T, Tunlid A. 2009. Variation in host specificity and gene content in strains from genetically isolated lineages of the ectomycorrhizal fungus *Paxillus involutus* s. lat. *Mycorrhiza* **19**: 549–558.

Heinonsalo J, Sun H, Santalahti M, Bäcklund K, Hari P, Pumpanen J. 2015. Evidences on the ability of mycorrhizal genus *Piloderma* to use organic nitrogen and deliver it to scots pine. *PLoS ONE* **10**: e0131561.

Henrici A. 2006. Notes and records. *Field Mycology* **7**: 31–34.

Hess J, Skrede I, Chaib De Mares M, Hainaut M, Henrissat B, Pringle A. 2018. Rapid divergence of genome architectures following the origin of an ectomycorrhizal symbiosis in the genus *Amanita*. *Molecular Biology and Evolution* **35**: 2786–2804.

Huber DPW, Bohlmann J. 2006. The role of terpene synthases in the direct and indirect defense of conifers against insect herbivory and fungal pathogens. In: *Multigenic and induced systemic resistance in plants*. New York, NY, USA: Springer Science, 296–313.

Huey RB. 1984. Is a jack-of-all temperatures a master of none? *Evolution* **38**: 441–444.

Hughes DT, Sperandio V. 2008. Inter-kingdom signaling: communication between bacteria and their hosts. *Nature Reviews Microbiology* **6**: 111–120.

Ingleby K, Mason PA. 1996. Ectomycorrhizas of *Thelephora terrestris* formed with Eucalyptus globulus. *Mycologia* **88**: 548–553.

Irieda H, Inoue Y, Mori M, Yamada K, Oshikawa Y, Saitoh H, Uemura A, Terauchi R, Kitakura S, Kosaka A *et al.* 2018. Conserved fungal effector suppresses PAMP-triggered immunity by targeting plant immune kinases. *Proceedings of the National Academy of Sciences, USA* **116**: 496–505.

Keck F, Rimet F, Bouchez A, Franc A. 2016. Phylosignal: An R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* **6**: 2774–2780.

Keller NP, Hohn TM. 1997. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology* **21**: 17–29.

Keller NP, Turner G, Bennett JW. 2005. Fungal secondary metabolism - from biochemistry to genomics. *Nature Reviews. Microbiology* **3**: 937–947.

Kennedy PG, Bruns TD. 2005. Priority effects determine the outcome of ectomycorrhizal competition between two *Rhizopogon* species colonizing *Pinus muricata* seedlings. *New Phytologist* **166**: 631–638.

Kennedy PG, Hill LT. 2010. A molecular and phylogenetic analysis of the structure and specificity of *Alnus rubra* ectomycorrhizal assemblages. *Fungal Ecology* **3**: 195–204.

Kim K-T, Jeon J, Choi J, Cheong K, Song H, Choi G, Kang S, Lee Y-H. 2016. Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Frontiers in Plant Science* **7**: 186.

Knapp DG, Németh JB, Barry K, Hainaut M, Henrissat B, Johnson J, Kuo A, Lim JHP, Lipzen A, Nolan M *et al.* 2018. Comparative genomics provides insights into the lifestyle and reveals functional heterogeneity of dark septate endophytic fungi. *Scientific Reports* **8**: 1–13.

Kochman K. 2014. Superfamily of G-protein coupled receptors (GPCRs) - extraordinary and outstanding success of evolution. *Advances in Hygiene & Experimental Medicine* **68**: 1225–1237.

Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canbäck B, Choi C, Cichocki N, Clum A *et al.* 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics* **47**: 410–415.

Kohout P, Sýkorová Z, Bahram M, Hadincová V, Albrechtová J, Tedersoo L, Vohník M. 2011. Ericaceous dwarf shrubs affect ectomycorrhizal fungal community of the invasive *Pinus* strobus and native *Pinus sylvestris* in a pot experiment. *Mycorrhiza* **21**: 403–412.

Kranabetter JM, Stoehr MU, O'Neill GA. 2012. Divergence in ectomycorrhizal communities with foreign Douglas-fir populations and implications for assisted migration. *Ecological Applications* **22**: 550–560.

Krause K, Henke C, Asiimwe T, Ulbricht A, Klemmer S, Schachtschabel D, Boland W, Kothe E. 2015. Biosynthesis and secretion of indole-3-acetic acid and its morphological effects on *Tricholoma vaccinum*-spruce ectomycorrhiza. *Applied and Environmental Microbiology* **81**: 7003–7011.

Kretzer A, Li Y, Szaro T, Bruns TD. 1996. Internal transcribed spacer sequences from 38 recognized species of Suillus sensu lato: phylogenetic and taxonomic implications. *Mycologia* **88**: 776–785.

Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**: 567–580.

Kües U, Khonsuntia W, Subba S, Dörnte B. 2018. Volatiles in communication of agaricomycetes. In: Anke T, Schüffler A, eds. *The mycota. Physiology and genetics: selected basic and applied aspects*. Cham, Switzerland: Springer International, 149–212.

Kuo M, Methven A. 2010. *100 Cool mushrooms*. Ann Arbor, MI, USA: University of Michigan Press.

Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels* **6**: 41.

Liao H-L, Chen Y, Bruns TD, Peay KG, Taylor JW, Branco S, Talbot JM, Vilgalys R. 2014. Metatranscriptomic analysis of ectomycorrhizal roots reveals genes associated with *Piloderma-Pinus* symbiosis: improved methodologies for assessing gene expression in situ. *Environmental Microbiology* **16**: 3730–3742.

Liao H-L, Chen Y, Vilgalys R. 2016. Metatranscriptomic study of common and host-specific patterns of gene expression between pines and their symbiotic ectomycorrhizal fungi in the genus *Suillus*. *PLoS Genetics* **12**: e1006348.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. 2015. Fungal effectors and plant susceptibility. *Annual Review of Plant Biology* **66**: 513–545.

Lofgren L, Nguyen NH, Kennedy PG. 2018. Ectomycorrhizal host specificity in a changing world: can legacy effects explain anomalous current associations? *New Phytologist* **220**: 1273–1284.

Lofgren LA, Uehling JK, Branco S, Bruns TD, Martin F, Kennedy PG. 2019. Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology* **28**: 721–730.

MacArthur R, Levins R. 1964. Competition, habitat selection, and character displacement in a patchy environment. *Proceedings of the National Academy of Sciences, USA* **51**: 1207–1210.

Marmeisse R, Guidot A, Gay G, Lambilliotte R, Sentenac H, Combier J-P, Melayah D, Fraissinet-Tachet L, Debaud JC. 2004. *Hebeloma cylindrosporum*-a model species to study ectomycorrhizal symbiosis from gene to ecosystem. *New Phytologist* **163**: 481–498.

Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V *et al.* 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**: 88–92.

Martin F, Díez J, Dell B, Delaruelle C. 2002. Phylogeography of the ectomycorrhizal *Pisolithus* species as inferred from nuclear ribosomal DNA ITS sequences. *New Phytologist* **153**: 345–357.

Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R *et al.* 2010. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**: 1033–1038.

Martino E, Morin E, Grelet GA, Kuo A, Kohler A, Daghino S, Barry KW, Cichocki N, Clum A, Dockter RB *et al.* 2018. Comparative genomics and transcriptomics depict ericoid mycorrhizal fungi as versatile saprotrophs and plant mutualists. *New Phytologist* **217**: 1213–1229.

Massicotte HB, Melville LH, Peterson RL, Molina R. 1999. Biology of the ectomycorrhizal fungal genus, *Rhizopogon*. *New Phytologist* **142**: 355–370.

Massicotte HB, Molina R, Luoma DL, Smith JE. 1994. Biology of the ectomycorrhizal genus, *Rhizopogon*. II. Patterns of host-fungus specificity following spore inoculation of diverse hosts grown in monoculture and dual culture. *New Phytologist* **126**: 677–690.

McLaughlin RN, Malik HS. 2017. Genetic conflicts: the usual suspects and beyond. *The Journal of Experimental Biology* **220**: 6–17.

Melin E, Krupa S. 1971. Studies on ectomycorrhizae of Pine II. Growth inhibition of mycorrhizal fungi by volatile organic constituents of *Pinus silvestris* (Scots Pine) roots. *Physiologia Plantarum* **25**: 337–340.

Melville LH, Massicotte HB, Peterson RL. 1987. Ontogeny of early stages of ectomycorrhizae synthesized between *Dryas integrifolia* and *Hebeloma cylindrosporum*. *Botanical Gazette* **148**: 332–341.

Miller SL, Koo CD, Molina R. 1992. Early colonization of red alder and Douglas fir by ectomycorrhizal fungi and *Frankia* in soils from the Oregon coast range. *Mycorrhiza* **2**: 53–61.

Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI *et al.* 2019. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research* **47**: D351–D360.

Miyauchi S, Kiss E, Kuo A, Drula E, Kohler A, Sánchez-García M, Morin E, Andreopoulos B, Barry KW, Bonito G *et al.* 2020. Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nature Communications* **11**: 5125.

Mohan V, Natarajan K, Ingleby K. 1993. Anatomical studies on ectomycorrhizas. *Mycorrhiza* **3**: 51–56.

Molina R, Massicotte H, Trappe JM. 1992. *Specificity phenomena in mycorrhizal symbioses: community-ecological consequences and practical implications*. London, UK: Chapman & Hall.

Moran NA. 2002. The evolution of host-plant alternation in aphids: evidence for specialization as a dead end. *American Naturalist* **132**: 681–706.

Morel M, Kohler A, Martin F, Gelhaye E, Rouhier N. 2008. Comparison of the thiol-dependent antioxidant systems in the ectomycorrhizal *Laccaria bicolor* and the saprotrophic *Phanerochaete chrysosporium*. *New Phytologist* **180**: 391–407.

Mortimer PE, Karunarathna SC, Li Q, Gui H, Yang X, Yang X, He J, Ye L, Guo J, Li H *et al.* 2012. Prized edible Asian mushrooms: ecology, conservation and sustainability. *Fungal Diversity* **56**: 31–47.

Mujic AB, Kuo A, Tritt A, Lipzen A, Chen C, Johnson J, Sharma A, Barry K, Grigoriev IV, Spatafora JW. 2017. Comparative Genomics of the ectomycorrhizal sister species *Rhizopogon vinicolor* and *Rhizopogon vesiculosus* (Basidiomycota: Boletales) reveals a divergence of the mating Type B Locus. *G3: Genes. Genomes, Genetics* **7**: 1775–1789.

Murphy JF, Miller OK. 1994. Mycorrhizal syntheses with *Alnus serrulata* (Ait.) Willd. *Castanea* **59**: 156–166.

Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S *et al.* 2020. A computational framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology* **16**: 60–68.

Nedelin T, Gyosheva M, Kostov K, Savev S. 2016. New Records and data on hypogeous ectomycorrhizal fungi in Bulgaria. *Forestry Ideas* **22**: 113–126.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**: 268–274.

Nguyen N, Vellinga EC, Bruns TD, Kennedy P. 2016. Phylogenetic assessment of global *Suillus* ITS sequences supports morphologically defined species and reveals synonymous and undescribed taxa. *Mycologia* **108**: 1216–1228.

Nicolás C, Martin-Bertelsen T, Floudas D, Bentzer J, Smits M, Johansson T, Troein C, Persson P, Tunlid A. 2019. The soil organic matter decomposition mechanisms in ectomycorrhizal fungi are tuned for liberating soil organic nitrogen. *ISME Journal* **13**: 977–988.

Nogueira-Lopez G, Greenwood DR, Middleditch M, Winefield C, Eaton C, Steyaert JM, Mendoza-Mendoza A. 2018. The apoplastic secretome of *Trichoderma virens* during interaction with maize roots shows an inhibition of

plant defence and scavenging oxidative stress secreted proteins. *Frontiers in Plant Science* **9**: 409.

Ogawa W, Takeda Y, Endo N, Yamashita S, Takayama T, Fukuda M, Yamada A. 2019. Repeated fruiting of Japanese golden chanterelle in pot culture with host seedlings. *Mycorrhiza* **29**: 519–530.

Oh KI, Melville LH, Peterson RL. 1995. Comparative structural study of *Quercus serrata* and *Q. acutissima* formed by *Pisolithus tinctorius* and *Hebeloma cylindrosporum*. *Trees* **9**: 171.

Oide S, Turgeon BG. 2020. Natural roles of nonribosomal peptide metabolites in fungi. *Mycoscience* **61**: 101–110.

Op De Beeck M, Troein C, Peterson C, Persson P, Tunlid A. 2018. Fenton reaction facilitates organic nitrogen acquisition by an ectomycorrhizal fungus. *New Phytologist* **218**: 335–343.

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W. 2014. *Caper: comparative analyses of phylogenetics and evolution in R.* [WWW document] URL https://cran.r-project.org/web/packages/caper/index.html [accessed 1 July 2020].

Ouvrard D, Chalise P, Percy DM. 2015. Host-plant leaps versus host-plant shuffle: a global survey reveals contrasting patterns in an oligophagous insect group (Hemiptera, Psylloidea). *Systematics and Biodiversity* **13**: 434–454.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Peay KG, Kennedy PG, Talbot JM. 2016. Dimensions of biodiversity in the earth mycobiome. *Nature Reviews Microbiology* **14**: 434–447.

Pera J, Alvarez IF. 1995. Ectomycorrhizal fungi of *Pinus pinaster*. *Mycorrhiza* **5**: 193–200.

Peter M, Kohler A, Ohm RA, Kuo A, Krützmann J, Morin E, Arend M, Barry KW, Binder M, Choi C *et al.* 2016. Ectomycorrhizal ecology is imprinted in the genome of the dominant symbiotic fungus *Cenococcum geophilum*. *Nature Communications* **7**: 1–15.

Plett JM, Daguerre Y, Wittulsky S, Vayssieres A, Deveau A, Melton SJ, Kohler A, Morrell-Falvey JL, Brun A, Veneault-Fourrey C *et al.* 2014. Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the *Populus* JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proceedings of the National Academy of Sciences, USA* **111**: 8299–8304.

Plett JM, Kohler A, Martin F. 2012. De-constructing a mutualist: How the molecular blueprints of model symbiotic fungi are changing our understanding of mutualism. In: Essered K, ed. *Fungal associations, 2nd edn.* Berlin/Heidelberg, Germany: Springer, 93–117.

Plett JM, Plett KL, Wong-Bajracharya J, de Pereira MdF, Costa MD, Kohler A, Martin F, Anderson IC. 2020. Mycorrhizal effector PaMiSSP10b alters polyamine biosynthesis in *Eucalyptus* root cells and promotes root colonization. *New Phytologist* **228**: 712–727.

Plett JM, Tisserant E, Brun A, Morin E, Grigoriev IV, Kuo A, Martin F, Kohler A. 2015. The Mutualist *Laccaria bicolor* expresses a core gene regulon during the colonization of diverse host plants and a variable regulon to counteract host-specific defenses. *Molecular Plant–Microbe Interactions* **28**: 261–273.

Policelli N, Bruns TD, Vilgalys R, Nuñez MA. 2019. Suilloid fungi as global drivers of pine invasions. *New Phytologist* **222**: 714–725.

Pritchard L, Birch P. 2011. A systems biology perspective on plant-microbe interactions: Biochemical and structural targets of pathogen effectors. *Plant Science* **180**: 584–603.

Quandt CA, Di Y, Jaiswal P, Spatafora JW, Elser J. 2016. Differential expression of genes involved in host recognition, attachment, and degradation in the Mycoparasite *Tolypocladium ophioglossoides*. *G3: Genes, Genomes, Genetics* **6**: 731–741.

Quin MB, Flynn CM, Wawrzyn GT, Choudhary S, Schmidt-Dannert C. 2013. Mushroom hunting by using bioinformatics: application of a predictive framework facilitates the selective identification of sesquiterpene synthases in basidiomycota. *Chembiochem: A European Journal of Chemical Biology* **14**: 2480–2491.

Quin MB, Flynn CM, Schmidt-Dannert C. 2014. Traversing the fungal terpenome. *Natural Product Reports* **31**: 1449–1473.

R Core Team. 2017. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. [WWW document] URL http://www.R-project.org/ [accessed 1 July 2020]. **55**: 275–286.

Raidl S, Agerer R. 1998. Hysterangium stoloniferum Tul & Tul.+ Picea abies (L) Karst. *Descriptions of Ectomycorrhizae* **3**: 31–35.

Reininger V, Sieber TN. 2012. Mycorrhiza reduces adverse effects of dark septate endophytes (DSE) on growth of conifers. *PLoS ONE* **7**: e42865.

Revell LJ. 2010. Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution* **1**: 319–329.

Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.

Richter DL, Michigan SU, Bruhn JN. 1990. *Scleroderma citrinum* (Gasteromycetes, Sclerodermatales) and *Larix decidua* form ectomycorrhizae in pure culture. *Nova Hedwigia* **50**: 355–360.

Rohlfs M, Churchill ACL. 2011. Fungal secondary metabolites as modulators of interactions with insects and other arthropods. *Fungal Genetics and Biology* **48**: 23–34.

Roy M, Dubois M-P, Proffit M, Vincenot L, Desmarais E, Selosse M-A. 2008. Evidence from population genetics that the ectomycorrhizal basidiomycete *Laccaria amethystina* is an actual multihost symbiont. *Molecular Ecology* **17**: 2825–2838.

Ruiz-Dueñas FJ, Barrasa JM, Sánchez-García M, Camarero S, Miyauchi S, Serrano A, Linde D, Babiker R, Drula E, Ayuso-Fernández I *et al.* 2020. Genomic analysis enlightens agaricales lifestyle evolution and increasing peroxidase diversity. *Molecular Biology and Evolution.* msaa301. doi: https://doi.org/10.1093/molbev/msaa301

Ruytinx J, Nguyen H, Hees MV, Beeck MOD, Vangronsveld J, Carleer R, Colpaert JV, Adriaensen K. 2013. Zinc export results in adaptive zinc tolerance in the ectomycorrhizal basidiomycete *Suillus bovinus*. *Metallomics* **5**: 1225–1233.

Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**: 301–302.

Sato I, Shimizu M, Hoshino T, Takaya N. 2009. The Glutathione System of *Aspergillus nidulans* Involves a Fungus-specific Glutathione S-Transferase. *Journal of Biological Chemistry* **284**: 8042–8053.

Sawyer NA, Chambers SM, Cairney JWG. 2001. Distribution and persistence of *Amanita muscaria* genotypes in Australian Pinus radiata plantations. *Mycological Research* **105**: 966–970.

Schulze-Lefert P, Panstruga R. 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends in Plant Science* **16**: 117–125.

Segal LM, Wilson RA. 2018. Reactive oxygen species metabolism and plant-fungal interactions. *Fungal Genetics and Biology* **110**: 1–9.

Shah F, Nicolás C, Bentzer J, Ellström M, Smits M, Rineau F, Canbäck B, Floudas D, Carleer R, Lackner G *et al.* 2016. Ectomycorrhizal fungi decompose soil organic matter using oxidative mechanisms adapted from saprotrophic ancestors. *New Phytologist* **209**: 1705–1719.

Sharma R, Mishra B, Runge F, Thines M. 2014. Gene Loss rather than gene gain is associated with a host jump from monocots to dicots in the smut fungus *Melanopsichium pennsylvanicum*. *Genome Biology and Evolution* **6**: 2034–2049.

Siegel N, Schwarz C. 2016. *Mushrooms of the redwood coast: a comprehensive guide to the fungi of coastal Northern California.* Berkeley, CA, USA: Ten Speed Press.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212.

Simpson GG. 1953. *The major features of evolution.* New York, NY, USA: Columbia University Press.

Singh S, Brocker C, Koppaka V, Ying C, Jackson B, Matsumoto A, Thompson DC, Vasiliou V. 2013. Aldehyde Dehydrogenases in cellular responses to oxidative/electrophilic stress. *Free Radical Biology & Medicine* **56**: 89–101.

Sipos G, Prasanna AN, Walter MC, O'Connor E, Bálint B, Krizsán K, Kiss B, Hess J, Varga T, Slot J *et al.* 2017. Genome expansion and lineage-specific genetic innovations in the forest pathogenic fungi *Armillaria*. *Nature Ecology & Evolution* **1**: 1931–1941.

Skibbe SD, Doehlemann G, Fernandes J, Walbot V. 2010. Maize tumors caused by Ustilago genes in host and pathogen. *Science* **67**: 89–93.

Smith AH, Thiers HD. 1964. *A Contribution Toward a Monograph of North American Species of Suillus.* Ann Arbor, MI, USA: Lubrecht & Cramer.

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ *et al.* 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**: 1543–1546.

Stajich JE. 2017. Fungal genomes and insights into the evolution of the kingdom. *Microbiology Spectrum* **5**: 619–633.

Steidinger BS, Crowther TW, Liang J, Van Nuland ME, Werner GDA, Reich PB, Nabuurs GJ, de-Miguel S, Zhou M, Picard N *et al.* 2019. Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature* **569**: 404–408.

Stireman JO. 2005. The evolution of generalization? Parasitoid flies and the perils of inferring host range evolution from phylogenies. *Journal of Evolutionary Biology* **18**: 325–336.

Tanaka A, Tapper BA, Popay A, Parker EJ, Scott B. 2005. A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. *Molecular Microbiology* **57**: 1036–1050.

Teramoto M, Wu B, Hogetsu T. 2012. Transfer of 14C-photosynthate to the sporocarp of an ectomycorrhizal fungus *Laccaria amethystina*. *Mycorrhiza* **22**: 219–225.

Trappe JM. 2009. Diversity, ecology, and conservation of truffle fungi in forests of the Pacific Northwest. In: US Department of Agriculture, Forest Service, Pacific Northwest Research Station, vol 772.

Tsuge T, Harimoto Y, Hanada K, Akagi Y, Kodama M, Akimitsu K, Yamamoto M. 2016. Evolution of pathogenicity controlled by small, dispensable chromosomes in Alternaria alternata pathogens. *Physiological and Molecular Plant Pathology* **95**: 27–31.

Uhse S, Djamei A. 2018. Effectors of plant-colonizing fungi and beyond. *PLoS Pathogens* **14**: e1006992.

UniProt Consortium T. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **46**: 2699.

vanDam P, Fokkens L, Schmidt SM, Linmans JHJ, Corby Kistler H, Ma LJ, Rep M. 2016. Effector profiles distinguish formae speciales of *Fusarium oxysporum*. *Environmental Microbiology* **18**: 4087–4102.

Vargas N, Gonçalves SC, Franco-Molano AE, Restrepo S, Pringle A. 2019. In Colombia the Eurasian fungus *Amanita muscaria* is expanding its range into native, tropical *Quercus humboldtii* forests. *Mycologia* **111**: 758–771.

Vayssières A, Pěnčík A, Felten J, Kohler A, Ljung K, Martin F, Legué V. 2015. Development of the Poplar-*Laccaria bicolor* ectomycorrhiza modifies root auxin metabolism, signaling, and response. *Plant Physiology* **169**: 890–902.

Visser B, Le Lann C, den Blanken FJ, Harvey JA, van Alphen JJM, Ellers J. 2010. Loss of lipid synthesis as an evolutionary consequence of a parasitic lifestyle. *Proceedings of the National Academy of Sciences, USA* **107**: 8677–8682.

Voiblet C, Duplessis S, Encelot N, Martin F. 2001. Identification of symbiosis-regulated genes in *Eucalyptus globulus - Pisolithus tinctorius* ectomycorrhiza by differential hybridization of arrayed cDNAs. *The Plant Journal* **25**: 181–191.

Walton JD. 2006. HC-toxin. *Phytochemistry* **67**: 1406–1413.

Walton JD, Panaccione DG. 1993. Host-Selective toxins and disease specificity: perspectives and progress. *Annual Review of Phytopathology* **31**: 275–303.

Ważny R. 2014. Ectomycorrhizal communities associated with silver fir seedlings (*Abies alba* Mill.) differ largely in mature silver fir stands and in Scots pine forecrops. *Annals of Forest Science* **71**: 801–810.

de Weert S, Kuiper I, Lagendijk EL, Lamers GEM, Lugtenberg BJJ. 2007. Role of Chemotaxis Toward Fusaric Acid in Colonization of Hyphae of *Fusarium oxysporum* f. sp. radicis-lycopersici by *Pseudomonas fluorescens* WCS365. *Molecular Plant–Microbe Interactions* **17**: 1185–1191.

Whittaker RH, Feeny PP. 1971. Allelochemics: chemical interactions between species. *Science* **171**: 757–770.

Wong KKY, Fortin JA. 1989. A Petri dish technique for the aseptic synthesis of ectomycorrhizae. *Canadian Journal of Botany* **67**: 1713–1716.

Xiong C, Xia Y, Zheng P, Wang C. 2013. Increasing oxidative stress tolerance and subculturing stability of *Cordyceps militaris* by overexpression of a glutathione peroxidase gene. *Applied Microbiology and Biotechnology* **97**: 2009–2015.

Xue C, Hsueh Y-P, Heitman J. 2008. Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS Microbiology Reviews* **32**: 1010–1032.

Yoder OC, Turgeon BG. 2001. Fungal genomics and pathogenicity. *Current Opinion in Plant Biology* **4**: 315–321.

Zeileis A, van de Wiel MA, Hornik K, Hothorn T. 2008. Implementing a class of permutation tests: the coin package. *Journal of Statistical Software* **28**: 1–23.

Zhao F-Y, Han M-M, Zhang S-Y, Wang K, Zhang C-R, Liu T, Liu W. 2012. Hydrogen peroxide-mediated growth of the root system occurs via auxin signaling modification and variations in the expression of cell-cycle genes in rice seedlings exposed to cadmium stress. *Journal of Integrative Plant Biology* **54**: 991–1006.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Distribution of Auxiliary Activity enzymes with count information.

Methods S1 Detailed methods for genome sequencing, assembly and annotation.

Tables S1 Sequencing and assembly statistics for the 22 newly sequenced *Suillus* genomes used in this study.

Table S2 Lookup information for the literature-curated GPCRs used to construct the GPCR database for each family used in this study.

Table S3 Genome completeness scores.

Table S4 InterPro domains overrepresented in *Suillus* with count information.

Table S5 InterPro domains underrepresented in *Suillus* with count information.

Table S6 'Oxidation-reduction process' InterPro domains over-represented in *Suillus* with count information.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.