Systematics, diversification, and functional diversity of Russulaceae (Russulales)

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Brian Patrick Looney
May 2017

Copyright © 2017 by Brian P. Looney
All rights reserved.

DEDICATION

This is for my mom, my dad, my wife, and my beautiful, beloved kitty Azuki.

ACKNOWLEDGEMENTS

Though specific acknowledgements are given for those that contributed to particular Chapters, I would like to acknowledge those individuals and organizations that have contributed to my training, thusfar, as a mycologist and research scientist. First, I would like to acknowledge my advisor, Brandon Matheny, and committee member, Karen Hughes, for mentoring me as a technician and independent researcher in their labs prior to starting graduate school. Without either of these people I would certainly be a worse scientist or even not one at all. I would like to recognize the Great Smoky Mountains mushroom crew, specifically Steve Trudell, Mike Wood, Jay Justice, and Else Vellinga, who first trained me how to photograph, collect, describe, and identify a wide breadth of mushroom-forming fungi. I greatly value this experience as my introduction to mycological research. I want to thank my labmates Joshua Birkebak, Marisol Sánchez-García, and Martin Ryberg who trained me in molecular techniques, phylogenetics, and good practices in science. Together with my subsequent labmates, Emma Harrower, Christine Braaten, Hailee Korotkin, and Rachel Swenie, I was the recipient of countless considerations, criticisms, ideas, and encouragement to stay focused and accomplish my research goals. I would like to recognize the professional and expert mycologists that assisted in my exploration and training in *Russula* systematics, including Steve Miller, Bart Buyck, David Lewis, Jorinde Nuytinck, and especially Felix Hampe and Slavomír Adamčík. I would also like to thank my committee member, Jessy Labbé, for mentoring me in advanced molecular biology skills and challenging me with developing research in the field of genomics, a field that is brand new to me. I would like to thank the Cumberland Mycological Society and the Gulf States Mycological Society for providing opportunities for collecting specimens and communicating my research to a broad audience. Finally, I would like to thank Ron Petersen for being a role model and for many stimulating conversations on mycology and life as a researcher.

ABSTRACT

The family Russulaceae is an iconic family of mushroom-forming Basidiomycetes both because of their importance as edible mushrooms in many parts of the world and their species richness in both temperate and tropical forested biomes. While much mycological research has been focused on this group, recent systematic and ecological research has failed to develop a comprehensive or cohesive organization by which to understand the evolutionary relationships, patterns of diversification, or functional importance of the group. Recently, interest in ectomycorrhizal fungi (EmF), of which Russulaceae is a key lineage, has greatly increased due to the recognition of the importance of EmF in carbon sequestration in the face of global climate change. By specifically taking a lineage-based approach to the study of Russulaceae, this work is an attempt to elucidate the biological importance of this group as a model for understanding important biological patterns in EmF. To this end, this dissertation work seeks to address five key questions: 1) What are the major systematic relationships in the *Russula*, Russulaceae, and their placement within Russulales? 2) What are the biogeographic and host patterns in *Russula*? 3) What factors have contributed to the high diversification of *Russula*? 4) What are the functional differences between major groups within Russulaceae? 5) To what extent have members of Russulaceae retained the ability to decompose soil organic matter?

To address these main questions, my research has combined the collection and study of sporocarps with molecular phylogenetics and contemporary evolutionary analytics. These efforts have led to the first multi-gene phylogeny of the genus *Russula* with a clade-based classification system proposed. By applying ancestral area reconstruction methodologies and diversification analyses using state speciation-extinction (SSE) models, I have inferred a temperate origin associated with angiosperms for *Russula*. I have provided support for a higher net diversification rate in temperate species of *Russula* that is not a result of migration. Here I present a molecular systematic revision of the Roseinae clade and provide support for at least 5 new species. Finally, in a comparative genomic analysis I show that Russulaceae are widely diverse in gene content, indicating diverse functional roles.

TABLE OF CONTENTS

INTRODUCTION 1
CHAPTER I Systematics, biogeography, and global diversification patterns in the genus *Russula* 4
Abstract 5
Introduction 6
Materials and Methods 8
Taxon sampling, DNA sequencing and phylogenetic analyses of the core data set 8
Clustering analyses of environmental sequences and metadata aquisition 9
Mega-phylogeny and BEAST analyses 10
Ancestral state reconstructions and diversification analyses 11
Results 12
Analysis of clustered MOTUs suggests both biogeographic distribution and plant association are phylogenetically overdispersed 12
Ancestral reconstruction and molecular clock methods suggest an extratropical origin of Russula associated with angiosperms during the Palaeogene 15
State speciation-extinction models suggest higher rates of diversification for Russula in the extratropics and in association with Pinaceae 18
Discussion 22
Into and out of the tropics 22
Higher diversification rates in the extratropics explain a reversal of the latitudinal diversity gradient (LDG) 23
Host switching is an important driver of diversification in Russula 25
Potential for additional drivers of diversification 26
Sampling and methodological considerations 27
Conclusions 29
Acknowledgments 30
References 31
Appendix 39
CHAPTER II Systematics, biogeography, and character evolution of *Russula* subsection *Roseinae* 66
Abstract 67
Introduction 67
Materials and Methods 70
Taxon Sampling 70
Molecular Sampling 72
Phylogenetic Inference and Time Calibration 73
Coalescent species delimitation approaches 74
Species Tree Estimation 74
Phylogeographic and host analyses 75
Ancestral State Reconstruction 75
Results 76

Phylogeographic reconstruction, dating, and phylogenetic informativeness of genes 76
Species delimitation in the Roseinae clade 76
Ancestral range and host reconstruction of the Roseinae clade 79
Ancestral reconstruction of traits in the Roseinae Clade 81
Discussion 81
Systematics of the Incrustatula clade 81
Species delimitation in the Roseinae clade 85
Phylogeography and host association of the Roseinae clade 85
Morphological trait evolution in the Roseinae clade 87
Conclusion 89
Acknowledgments 89
References 90
Appendix 97
CHAPTER III Russulaceae as a model system to recognize functional diversity in an ectomycorrhizal lineage 98
Abstract 99
Introduction 99
Systematics and Sampling Strategy 101
Diversification, Biogeographic, and Host Association Patterns 107
Ecology, Life History, and Ecosystem Function 110
Growth Characteristics *in vitro* 112
Gene Content and Molecular Composition 113
Conclusions and Future Prospective 120
References 122
CHAPTER IV Phylogenomics and comparative genomics of Russulaceae 132
Abstract 133
Introduction 133
Materials and Methods 135
Taxon sampling and nucleic acid extraction 135
Genome Sequencing and Annotation 137
Phylogenomics and gene family reconstructions 138
Pan-genome assembly and network analyses 139
Transposable elements and microsatellite analysis 140
Results 140
Phylogenomic reconstruction shows trends in genome size 140
Transposable elements and simple sequence repeats analyses 141
Pangenome and network analyses 145
Evolution of lignin degrading enzymes 150
Discussion 150
Acknowledgments 157
References 158
CONCLUSION 164
VITA 165

LIST OF TABLES

Table 1. Crown ages and ancestral character states reconstructed for *Russula* and major clades. 16
Table 2. Model comparisons for BiSSE, BiSSE-ness, and GeoSSE analyses. 19
Table 3. Maximum Likelihood estimates of parameters for the best model for BiSSE and GeoSSE analyses. 21
Table 4. Taxon sampling and Genbank accession numbers° for constraint phylogeny. 39
Table 5. Final MOTU clusters with associated metadata for mycoheterotrophic parasitism, gasteroid morphology, plant association and geographic distribution. 48
Table 6. Russulaceae sampled with stats (location collected, date collected, GenBank #'s) 106
Table 7. Genome assembly and annotation statistics. Green number indicate smallest and red is largest. 142

LIST OF FIGURES

Figure 1. Maximum parsimony ancestral reconstructions of geography (left) and host association (right) along an ultrametric mega-phylogeny of environmental *Russula* MOTUs inferred using r8s. Major clades are designated by colour. Geographical and host tree metadata associated with MOTU clusters are designated with coloured lines at the tips, with equivocal tips inferred from the analysis. Areas are coded dark blue for extratropical distribution, green for tropical distribution and orange for cosmopolitan distribution. Plant association data are coded red for angiosperm association, aqua for Pinaceae association and black for generalist. Net relatedness indices produced in phylocom indicate phylogenetic overdispersion for both character sets. 13

Figure 2. Map projection of the global distribution of *Russula* MOTUs. The areas of circles are scaled by the number of MOTUs relative to total MOTUs recovered. Top numbers represent number of endemic MOTUs. Bottom numbers indicate total number of MOTUs. Lines represent number of overlapping distributions for widespread MOTUs. 14

Figure 3. Boxplot comparing average taxon age based on terminal branch lengths of taxa from a secondarily time-calibrated mega-phylogeny with ranges in the tropics or extratropics (Top) and host association with Pinaceae or angiosperms (Bottom). P-values resulted from nonparametric Mann–Whitney *U*-tests...... 17

Figure 4. Posterior probability density means and standard error representing relative diversification (k-l) and dispersal rates for geographic and host state-specific models for *Russula*. (A) Estimates for extratropical MOTUs (light blue) and tropical MOTUs (orange) with differential transition rates from tropics to extratropics (grey) for the best-supported equal transition GeoSSE model. (B) Estimates for angiosperm-associated MOTUs (red) and Pinaceae-associated MOTUs (blue) with differential transition rates from Pinaceae association to angiosperm association (grey) for the best-supported full-parameter GeoSSE model. (C) Estimates for endemic MOTUs (green) and widespread MOTUs (purple) with differential transition rates of contraction from widespread to endemic (grey) for the best-supported full-parameter BiSSE model. (D) Estimates for host-specific MOTUs (yellow) and host generalist MOTUs (pink) with differential transition rates of restriction from host generalist to specialist (grey) for the best-supported full-parameter BiSSE model 20

Figure 5. Phylogenetic relationships of *Russula* inferred from nuclear ribosomal and single-copy (ITS, nrLSU, *rpb1* and *rpb2*) sequences derived from a maximum-likelihood analysis. 41

Figure 6. Chronogram of Russulaceae inferred in BEAST 2. The 95% HPD posterior probabilities >0.9 are reported. 42

Figure 7. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for GeoSSE tropical vs. extratropical distribution. 43
Figure 8. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for GeoSSE angiosperm vs. Pinaceae association. 44
Figure 9. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for BiSSE binary tropical/temperate endemicity vs. widespread. 45
Figure 10. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for BiSSE binary angiosperm/Pinaceae specificity vs. generalist association. 46
Figure 11. BAMM analysis showing (A) the top nine shift configurations of the most credible shift configuration set. Red circles represent increases in diversification rate, while blue circles represent slowdowns in diversification. The size of the circle indicates how significant the shift is. The f values indicate what proportion of the confidence can be assigned to that particular scenario; and (B) a circle phylogeny of the best shift configuration. 47
Figure 12. North American and European members of *Russula* subsect. *Roseinae*: A) *Russula albida* (photo Per Marstad); B) *Russula leporina* nom. prov.; C) *Russula magnorosea* nom. prov. (photo Per Marstad); D) *Russula* cf. *magnorosea*; E) *Russula peckii*; F) *Russula minutula* (photo Per Marstad); G) *Russula peckii* (photo Per Marstad); H) *Russula niveopersonata* nom. prov.; I) *Russula austrorubellipes* nom. prov. (photo Steve Trudell); J) *Russula velutipes* 71
Figure 13. Ultrametric chronogram of the Incrustatula clade inferred in BEAST. Bootstrap support is reported along branches (black) from ML reconstruction in RAxML and is followed by posterior probabilities. Mean estimations of ages in million years are reported at nodes (red). Hyphens are used when bootstrap support is below 50%, posterior probability is below 0.5, or if the node was not recovered by either method. Inset shows phylogenetic informativeness inferred in PhyDesign of the five nuclear markers used for inferring the phylogeny. 77
Figure 14. Maximum Likelihood phylogeny of the Roseinae clade inferred in RaXML with bootstrap support reported along branches. Collections sampled are listed with collector number and locality. Best supported species models for SpedeSTEM and BP&P are given in colored bars with the full species model (FULL) given first. Species epithets are given for well-resolved clades recovered from both coalescent models, with questionable taxa given *confer* status. 78
Figure 15. Ancestral area and host reconstruction of the best-supported model estimated in the R package 'BioGeoBEARS' 80

Figure 16. Microscopic characters used in ancestral trait reconstruction of the Roseinae clade: A) Pleurocystidia with dispersed elements (*R. rubellipes*, NY00253510, holotype); B) Pleurocystidia with partly dispersed elements and a refringent body (*R. pseudopeckii*, NY00253511, holotype); C) Pleurocystidia with large refringent bodies only in apex (*R. peckii*, NYS f3630.7, holotype); D) Dense trichoderm pileipellis composed of narrow, cylindrical, frequently branched hyphal terminations and narrow primordial hyphae (*R. pseudopeckii*, NY00253511, holotype); E) Loose trichoderm pileipellis composed of moderately wide, attenuated, occasionally branched hyphal terminations and wide, short-celled primordial hyphae (*R.* sp., SAV F-3576); F) Epithelium pileipellis of inflated, ellipsoid, rarely branched elements and moderately wide primordial haphae (*R. rubellipes*, NY00253510, holotype) ; G) Spores exhibiting short width and ornamentation height (*R. pseudopeckii*, NY00253511, holotype); H) Spores exhibiting medium width and ornamentation height (*R. rubellipes*, NY00253510, holotype); and I) Spores exhibiting tall width and ornamentation height (*R. praeumbonata*, NY00760516, syntype). Scale bar equals 10 µm but only 5 µm for spores...... 82

Figure 17. Ancestral character reconstruction using ML in Mesquite for both macro- and microcharacters. Characters that were not measured for given species are marked as an X. Reconstructions where no character was resolved as less than 25% significant were excluded... 83

Figure 18. Gene tree reconstructions of Incrustatula clade for five loci (ITS, *rpb1*, *rpb2*, *tef1*, *mcm7*) used in multi-gene analyses showing major clade relationships (Red-Core Roseinae; Pink-Residual Roseinae; Orange-Albida; and Purple-Lilaceinae). Asterisks and green stars mark well-supported clade. 97

Figure 19. Maximum Likelihood phylogenetic reconstruction of Russulaceae with sampling from Buyck et al. (2008) and Looney et al. (2016). Clades are collapsed based on clades from Looney et al. (2016), the molecular subgeneric classification of *Lactarius* from (Verbeken & Nuytinck, 2013), and the subgeneric classification of *Lactifluus* from De Crop et al. (2017). Taxon sampling for the Mycorrhizal Genome Initiative is highlighted in blue and RGI samples are highlighted in red. Samples marked with * are still being prepared for submission to sequence...... 103

Figure 20. Maximum Likelihood phylogenetic reconstruction of Russulales dataset from Miller et al. (2006) with a backbone constraint topology of well-supported relationships from the Mycocosm Portal phylogeny (Grigoriev *et al.*, 2014). Clades are collapsed based on clades recovered from Miller et al. (2006) and bootstrap values ≥ 50 are reported. Representative samples of species and/or genera for which genomes are currently available are highlighted in red...... 108

Figure 21.Growth and colony morphology of *M. och* (*Multifurca ochricompacta*), *L.del* (*Lactarius* cf. *deliciosus*), and *L.psa* (*Lactarius psammicola* with response to a temperature gradient in MMN. 114
Figure 22.Growth of *L.del* (*Lactarius* cf. *deliciosus, M.och* (*Multifurca ochricompacta*), and *L.psa* (*Lactarius psammicola*) in terms of colony diameter (mm) in MMN in response to different temperature treatments. Each data point is a mean of three replicates with a standard deviation of the mean. 115
Figure 23. Growth and colony morphology of *M. och* (*Multifurca ochricompacta*), *L.del* (*Lactarius* cf. *deliciosus*), and *L.psa* (*Lactarius psammicola* with response to a temperature gradient in MMN. 116
Figure 24. Growth of *L.del* (*Lactarius* cf. *deliciosus, M.och* (*Multifurca ochricompacta*), and *L.psa* (*Lactarius psammicola*) in terms of colony diameter (mm) in MMN in response to different pH treatments. Each data point is a mean of three replicates with a standard deviation of the mean. 117
Figure 25. Phylogenetic reconstruction of Russulales genomes using 2,518 single-copy genes in RaXML with 1,000 bootstrap iterations. Taxon labels correspond to JGI identifiers (Ruscom1–*Russula compacta*; Rusbre1–*Russula brevipes*; Rusdis1–*Russula dissimulans*; Rusrug1–*Russula rugulosa*; Lacsub1–*Lactifluus subvellereus*; Lacvol1–*Lactifluus* cf. *volemus*; Muloch1–*Multifurca ochricompacta*; Lacqui1–*Lactarius quietus*; Glocon1–*Gloeopeniophorella convolvens*; Aurvu1–*Auriscalpium vulgare*; Lenvul1–*Lentinellus vulpinus*; Clapy1–*Artomyces pyxidata*; Stehi1–*Stereum hirsutum*; Hetan2–*Heterobasidion irregulare*; Densp1–*Dentipellis sp.*; Lopni1–*Peniophora sp.*; Ricme1–*Peniophora cinereus*; Varmin1–*Vararia minispora*). Genomes are coded for nutritional mode as Ectomycorrhizal (ECM), Saprotrophic (Sap), and Pathogen (Path) with genome size in Megabases (Mb) graphed as bars. 143
Figure 26. Comparison of the proportion of Russulales genomes that are comprised of transposable elements and other repeated elements... 144
Figure 27. Comparison of the number and different classes of microsatellite repeats found in Russulales genomes. Colors correspond to different classes (dark blue–mononucleotide; orange–dinucleotide; gray–trinucleotide; yellow–tetranucleotide; light blue–pentanucleotide; green–hexanucleotide microsatellite). 146
Figure 28. Heat-map reconstruction of the pan-genome of Russulales from a sequence homology-based protein families approach highlighting presence and absence of Pfam gene families based on Markov Clustering Algorithm. Core proteins shared by all members are clustered at the top with variability in presence and absence found throughout the rest. Abbreviations for genomes are JGI identifiers (see Fig. 25). 147

Figure 29. Heat-map reconstruction of unique Pfam gene families recovered in pan-genome analysis. Abbreviations for genomes are JGI identifiers (see Fig. 25). 148
Figure 30. Proportional Similarity networks of the gene family content profiles of 18 Russulales species at different Proportional Similarity thresholds. Each node represents a species, and edges represent the similarity between the gene family profiles of species. 149
Figure 31. Network of Pfam domains that are positively enriched for nutritional mode within Russulales. Squares represent different species, while dots indicated enriched domains. 151
Figure 32. A) Phylogenetic reconstruction of the lignin peroxidase gene family from Russulales that codes for lignin peroxidase. Well-supported clades (bootstrap ≥70) are named based on included taxa. Clades including or made up of members of Russulaceae are colored based on nutritional mode (green–ectomycorrhizal; red–saprotrophic). B) Phylogenetic reconstruction of the lignin peroxidase gene family only from Russulaceae. Species with multiple gene copies are highlighted in color, including two ectomycorrhizal species (*M. ochricompacta* & *R. brevipes*) and the saprotrophic representative (*G. convolvens*). Abbreviations for genomes are JGI identifiers (see Fig. 25). 152
Figure 33. Phylogenetic reconstruction of the multi-copper oxidase (MCO) gene family from Russulaceae that codes for laccases. Gene copy numbers are numbered with JGI identifiers as abbreviations for genomes (see Fig. 1). Members of Russulaceae are colored based on nutritional mode (green–ectomycorrhizal; red–saprotrophic). 153

INTRODUCTION

The family Russulaceae is an iconic family of mushroom-forming Basidiomycetes that have attracted the attention of mycologists throughout the past few centuries, including a "founding father" of mycology Elias Magnus Fries. Rolf Singer wrote of the genus *Russula* in his 1986 edition of *Agaricales in Modern Taxonomy* that "there is no genus in the Agaricales where more species have been studied anatomically and chemically… where more type specimens and authentic material has been critically revised in the light of modern methods… [where] more effort has been spent by local and traveling specialists." Despite all of this classical work, knowledge of how to recognize North American species of *Russula* is severely lacking, even leading a taxonomic expert of the group, Bart Buyck, to call for a new initiative towards the study of *Russula* in the eastern USA and write that "local expertise on the genus has completely vanished." As a speciose group of fungi, it is imperative that biologists are able to recognize species so we can make comparisons between studies and better understand species distribution patterns and the ecological roles individual species or groups of species play in their environment. This is especially important for groups of ectomycorrhizal fungi, like Russulaceae, which are obligate mutualists with dominant trees and shrubs of many forested ecosystems. It is for these reasons that I have undertaken this project to elucidate the systematics, biogeography, and functional diversity of Russulaceae.

For Chapter 1 I present a global meta-analysis of the genus *Russula*, the most speciose genus in Russulaceae, to test for the evolutionary pattern that has driven a reversal in the latitudinal diversity gradient pattern in ectomycorrhizal fungi. To accomplish this goal, I erected a multi-gene phylogeny for the genus *Russula* using vouchered specimens of species representative of morphologically defined infrageneric groupings in order to identify the major clades of the genus and use this tree as a topology constraint for a global phylogeny of *Russula*. To reconstruct the global phylogeny of *Russula*, sequences of the ribosomal internal transcribed spacer of *Russula* deposited in GenBank were clustered into near species-level molecular taxonomic units (MOTUs) and the constraint topology was enforced in RaXML using a combined Maximum likelihood and Maximum parsimony approach. MOTUs were then coded by geographic state (Tropical vs. Extratropical) and host association (Angiosperm vs. Pinaceae) for ancestral state reconstruction and state speciation and extinction (SSE) modeling.

In Chapter 2 I present a systematic revision and biogeographic analysis of a charismatic group of mostly red-capped *Russula* species that have traditionally been placed in *Russula* subsect. *Roseinae* Singer ex Sarnari. Specimens were collected on field trips conducted between 2012 and 2016 focused on the eastern United States. Combined with historical specimens, including type material, all putative members of this subsection and its sister group, *Russula* subsect. *Lilaceinae* (Melzer & Zvára) Jul. Schäff., were sampled for multiple molecular markers to reconstruct a subsection-level and species-level multi-gene

phylogeny. Gene markers were tested for phylogenetic informativeness at these different taxonomic levels. I used a combined approach for species delimitation encompassing morphology, geography, phylogeny, and evolution. For evolution-based delimitation of species, I applied the multispecies coalescent model using Bayesian phylogenetics and phylogeography (BP&P), a Bayesian approach that takes a species tree and concatenated sequence alignment as inputs, and Speedy Species Tree Estimation Using Maximum Likelihood (SpedeSTEM), a likelihood approach that takes gene trees and gene alignments as inputs. The biogeographic and host reconstruction were performed in the R package 'BioGeoBEARS', which compares different biogeographical models and incorporates the *j* parameter, which simulates founder-event jump dispersal or long-distant dispersal. Morphological characters were reconstructed using ancestral state reconstruction in Mesquite to test for conservation of diagnostic characters at the species-level. Finally, species determinations are made based on multiple lines of evidence and proposals for the erection of new and amendment of existing taxonomic groups is made.

For Chapter 3 I offer a literature review and growth study of Russulaceae as an overview of the current knowledge of the family and as a preview of the whole-genome dataset of Russulaceae I have produced through the Russulaceae Genome Initiative (RGI) in collaboration with the Joint Genome Institute of the U.S. Department of Energy. I begin this overview by placing the RGI in the broader context of the current state of mycological genomic studies to illustrate why this project is an appropriate outcome of what has already been accomplished. Next, I attempt to elucidate the current state of systematics within Russulaceae, its placement in relation to other groups in Russulales for which genomes have been sampled, and justification for the sampling strategy of the RGI. A section is given on the evolutionary history, biogeography, and host relationships of Russulaceae, contrasting my own research with other studies on *Russula* and the potentially larger body of literature on the genera *Lactarius* and *Lactifluus*. Following this, I present an overview of studies where Russulaceae has been highlighted as an ecologically unique or important group, including studies on its life history. For samples where I was able to obtain an axenic tissue culture and sequence a genome, a growth study was designed to illustrate the potential experimental applications of these cultures in conjunction with an annotated genome. For future comparative genomic studies, a section is given on the current body of knowledge on the biochemical and genetic diversity that has been elucidated using relevant species of Russulaceae. Hypotheses based on this current state of knowledge are proposed throughout. Finally, a conclusion section summarizes some key points and proposes future analyses for the RGI dataset.

Chapter 4 represents a comparative genomic study through the RGI of the order Russulales with the addition of dense genome sampling of the family Russulaceae. Sampling from Russulaceae is necessary for such a study as they represent about two-thirds of the species diversity of the order. To begin the study, annotated Pfam (http://pfam.xfam.org/) gene families from 18

representative genomes of Russulales, including 8 newly sequenced genomes of Russulaceae, were compared across the order to infer the pan-genome of Russulales. Gene families for which the entire order shared exactly one orthologous copy were extracted and aligned for a phylogenomic reconstruction of relationships within the Russulales. Genomes were analyzed for transposable element (TE) and simple sequence repeat (SSR) content, and genome size and repeated element content was compared across the phylogeny between saprotrophic and ectomycorrhizal members. To look at overall functional similarity between the different members of Russulaceae, a similarity analysis was implemented using gene family profiles from the pan-genome assembly and calculating similarity for all potential species pairings using proportional similarity as the similarity metric and visualized as a network. To look for specific functional similarities between different trophic modes a gene network analysis of enrichment on Pfam domains was performed using Fisher's exact test for both positively and negatively enriched domains by trophic mode. Finally, I reconstructed the phylogenies of the lignin peroxidase and multi-copper oxidase gene families to test whether Russulaceae has retained the ability to degrade lignin from a white-rot saprotrophic ancestor and whether there have been expansions in these gene families indicating neofunctionalization.

CHAPTER I
SYSTEMATICS, BIOGEOGRAPHY, AND GLOBAL DIVERSIFICATION PATTERNS IN THE GENUS *RUSSULA*

A version of this chapter was originally published as:
Looney, B.P., M. Ryberg, F. Hampe, M. Sánchez-García and P.B. Matheny (2016) Into and out of the tropics: global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi. *Molecular Ecology* 25: 630-647.

The doctoral student collected many specimens for this study, performed all of the lab work for generating sequences, constructed both phylogenetic datasets, inferred all of the phylogenies, performed ancestral area and host reconstructions, performed diversification analyses, and was the primary author of the manuscript. M. Ryberg assisted with clustering analyses and provided comments on the manuscript. F. Hampe provided European samples for sequencing and provided comments on the manuscript. M. Sánchez-García produced a useful Perl script for extracting metadata, assisted with diversification analyses, and provided comments on the manuscript. P.B. Matheny coordinated research efforts, provided lab space and training, and significantly contributed to editing the manuscript.

Abstract

Ectomycorrhizal (ECM) fungi, symbiotic mutualists of many dominant tree and shrub species, exhibit a biogeographic pattern counter to the established latitudinal diversity gradient of most macroflora and fauna. However, an evolutionary basis for this pattern has not been explicitly tested in a diverse lineage. In this study, we reconstructed a mega-phylogeny of a cosmopolitan and hyperdiverse genus of ECM fungi, *Russula*, sampling from annotated collections and utilizing publicly available sequences deposited in GenBank. Metadata from molecular operational taxonomic unit cluster sets were examined to infer the distribution and plant association of the genus. This allowed us to test for differences in patterns of diversification between tropical and extratropical taxa, as well as how their associations with different plant lineages may be a driver of diversification. Results show that *Russula* is most species-rich at temperate latitudes and ancestral state reconstruction shows that the genus initially diversified in temperate areas. Migration into and out of the tropics characterizes the early evolution of the genus, and these transitions have been frequent since this time. We propose the ‘generalized diversification rate’ hypothesis to explain the reversed latitudinal diversity gradient pattern in *Russula* as we detect a higher net diversification rate in extratropical lineages. Patterns of diversification with plant associates support host switching and host expansion as driving diversification, with a higher diversification rate in lineages associated with Pinaceae and frequent transitions to association with angiosperms.

Introduction

A long established global pattern of biogeography proposed for macroorganisms is the latitudinal diversity gradient (LDG), observed by early naturalists and corroborated over several centuries in numerous studies (Von Humboldt 1807; Hillebrand 2004). This pattern has been supported for all major groups of macroflora and fauna including plants, amphibians, mammals, birds, reptiles, and marine and terrestrial invertebrates (Hillebrand 2004). Alternatively, microbes have traditionally been considered to follow the 'everything is everywhere, but the environment selects' model, although some heterogeneity has been shown for certain groups (Baas-Becking 1934; Fontaneto et al. 2008). At the interface of these two global distribution patterns are fungi, which have traditionally been considered to follow the microbial model but more recently been found to be highly geographically segregated (Taylor et al. 2006). Due to this intermediate position that fungi have traditionally held, biogeographic patterns of fungi have been poorly understood and have received less attention (Lumbsch et al. 2008; Tedersoo et al. 2012). Given recent advances in molecular methods for detecting species from environmental samples, it is now much more feasible to investigate patterns in their global distribution (Tedersoo et al. 2014a).

Recent studies have demonstrated that ectomycorrhizal (ECM) fungi exhibit a biogeographic pattern counter to the LDG, where ECM fungal diversity increases away from the tropics and towards the temperate/ boreal interface (Tedersoo et al. 2012, 2014a). ECM fungi are obligate symbionts with plant roots of primarily tree and shrub species, whereby the fungus provides water and nutrients (viz, nitrogen and phosphorus) to the plant in exchange for photosynthates (Alexopoulos et al. 1996). This symbiosis is necessary for these fungi to complete their life cycle, and it is also critical for their plant partners as this symbiosis provides a competitive advantage (Perry et al. 1989). There are an estimated 25 000 ECM species worldwide, and this biotrophic association has evolved independently some 80 times primarily in the Ascomycota and Basidiomycota (Rinaldi et al. 2008; Tedersoo & Smith 2013). ECM symbiosis with only those plant lineages that allow ECM colonization makes ECM fungi an ideal guild to investigate global biogeographic patterns, as their biogeography and diversification patterns are probably heavily influenced by the distribution, dispersion and diversification patterns of their plant partners (Hoeksema 2010).

An initial meta-analysis by Tedersoo et al. (2012) showed a reversal of the LDG in ECM fungi by analysis of metadata from numerous fungal communities, and following studies have highlighted potential ecological drivers of this pattern. This work demonstrated that ECM fungal richness peaks between 4000 and 4500 km from the equator (36°–40.5° N/S). Edaphic, climatic and biotic factors were tested in a multivariate model as predictors for ECM species richness, of which several were significant, including mean annual temperature, mean annual precipitation, anthropogenic disturbance, soil texture and ECM plant family. Soil volume had a positive correlation with ECM species richness, which has been proposed as a possible ecological driver for a reversed LDG by allowing more

stratification for niche space (Peay et al. 2010; Smith et al. 2011; Kennedy et al. 2012). For a lineage level analysis, ECM plant family explained 34% of the variation in ECM fungal communities, indicating that species tend to be segregated by ECM plant partner (Tedersoo et al. 2012). A follow-up study utilized a standardized sampling approach while collecting environmental soil samples from 365 sites around the globe (Tedersoo et al. 2014a). This latter study confirmed the reversed LDG trend in ECM fungi but upheld the standard LDG pattern for saprotrophic, parasitic and pathogenic fungi (Tedersoo et al. 2014a). The strongest predictors explaining global ECM richness in this analysis included the ratio of ECM plant abundance relative to non-ECM plant abundance in a community, total ECM plant species richness and soil pH. While these studies have focused on identifying ecological factors that predict ECM species richness, evolutionary mechanisms that might help explain how these factors contribute to the reversed LDG pattern have been largely overlooked (Kennedy et al. 2012).

The ability to model the evolutionary dynamics underlying the LDG has resulted in a number of testable hypotheses that could be applied to ECM fungi. The ‘tropical conservatism hypothesis’ has been proposed as a general explanation of the LDG, where lineages have a tropical origin, and the conserved environmental niches of these organisms restrict dispersal to the extratropics, most likely due to climatic restraints (Latham & Ricklefs 1993; Wiens & Donoghue 2004). The ‘out of the tropics’ hypothesis proposes that the tropics can act simultaneously as a museum and a cradle for these lineages, where dispersal events to the extratropics are frequent yet the lineages will still concurrently persist and diversify in the tropics (Jablonski et al. 2006). The ‘diversification rate hypothesis’ proposes that a higher net diversification rate in the tropics is driving the LDG, whether due to a higher rate of molecular evolution, stable climatic conditions over evolutionary time, or periods of tropical expansion in the evolutionary past (Rohde 1992; Jansson et al. 2013). These three hypotheses have been proposed as a nested hierarchy, with the ‘tropical conservatism hypothesis’ being the most restrictive (Kerkhoff et al. 2014). Sánchez-Ramírez et al. (2015a) recently tested for an evolutionary pattern to explain the reversed LDG in a clade of the ECM genus Amanita and found that temperate lineages have a higher speciation rate. This study seeks to further test for these patterns in a hyperdiverse genus of ECM fungi.

Russula is the largest genus in the order Russulales comprising some 750–900 described species (Kirk et al. 2008; Buyck & Atri 2011). *Russula* can, therefore, be considered the second most taxonomically diverse genus of ECM fungi after the genus Cortinarius (Kirk et al. 2008). The genus is a dominant ECM lineage in tropical, temperate, boreal and tundra ECM communities (Singer 1986; Buyck et al. 1996; Geml et al. 2009). Russulaceae have also been hypothesized to have a tropical origin (Buyck et al. 1996), which according to established biogeographic hypotheses (Wiens & Donoghue 2004; Jablonski et al. 2006; Jansson et al. 2013), would suggest the family should be most diverse in the tropics. Indeed, Tedersoo & Nara (2010) found the/russulalactarius lineage to be

more diverse in tropical forests; however, this conclusion was tentative as statistical support was lacking. Members of the genus *Russula* are ecologically diverse as they associate with every major ECM plant lineage (Singer 1986), are host to mycoheterotrophic members of Ericaceae and Orchidaceae (Kennedy et al. 2011a), and occasionally have gasteroid fruit body morphology, which makes up a significant a proportion of the diet of many small mammals (Lebel & Tonkin 2007). Phylogenetic relationships within the genus have been proposed (Eberhardt 2002; Miller & Buyck 2002; Buyck et al. 2008), but taxon sampling and gene sampling have been sparse to date. The first major phylogenetic treatment of the genus identified six major clades using a single molecular marker. Not unexpectedly, internodal support was lacking for most higher-level relationships (Miller & Buyck 2002). A later multigene analysis of the family Russulaceae resolved four genera, but taxon sampling was not adequate to resolve major clades within *Russula* (Buyck et al. 2008). Because a multigene treatment with sufficient taxon sampling is unavailable, a more robust phylogenetic framework for the group is necessary to investigate the history of their diversification.

The objectives of this study are to: (i) produce a robust phylogeny of the genus *Russula* as a basis to investigate its patterns of diversification; (ii) utilize clustering of global sampling and metadata associated with DNA sequences of *Russula* to resolve its global distribution and ECM plant associations; (iii) use ancestral state reconstruction methods to infer the evolutionary history of its biogeography and plant association; and (iv) compare biogeographic models to infer rates of diversification and transitions in biogeographic states and plant associations. By examining the history of diversification of a large genus of ECM fungi, we seek to understand what general evolutionary patterns exist and whether co-evolution or host switching might be driving this pattern at a large scale.

Materials and Methods

Taxon sampling, DNA sequencing and phylogenetic analyses of the core data set

Vouchered specimens from North America and Europe were sequenced to infer a multigene phylogeny of *Russula* (Table 4, Appendix). To ensure sampling of wide phylogenetic diversity, type species of major infrageneric groups were targeted from three of the most relevant infrageneric classification systems proposed for *Russula* (Romagnesi 1967; Singer 1986; Sarnari 1998). Full morphological descriptions with color notes (Kornerup & Wanscher 1967) were made for identification of all specimens. Specimens were dehydrated and deposited at the TENN and GENT herbaria [herbarium abbreviations per Thiers (continuously updated)]. DNA extraction and PCR protocols followed that of Birkebak et al. (2013). Four loci were targeted for Infrageneric clade-level resolution including two nrDNA regions (nuclear ribosomal large subunit (LSU)

and internal transcribed spacers (ITS) and two single-copy genes (*rpb1* and *rpb2*, which encode the largest and second largest subunits of RNA polymerase II, respectively). We refer to this alignment as the 'core data set'. The following primer pairs were used for amplification: ITS using ITS1F–ITS4 (White et al. 1990; Gardes & Bruns 1993); LSU using LR0R–LR5 (Vilgalys & Hester 1990); *rpb2* using b6F–b7.1R (Matheny 2005); and *rpb1* using gAf–fCr (Matheny et al. 2002) with int2F and int2.1R as internal sequencing primers. Sequences were assembled using SEQUENCHER 4.9 (Gene Codes, Ann Arbor, MI, USA). Alignments incorporating multilocus data from previous systematic studies (Buyck et al. 2008; Van de Putte et al. 2012; Looney 2015) were constructed separately for each gene region using MAFFT 6.717 (Katoh & Toh 2008) using the L-INS-i algorithm and manually adjusted in MACCLADE 4.08 (Maddison & Maddison 2005). Intergene conflict was investigated by inferring phylogenies for each locus using RAXMLGUI 1.2 (Stamatakis et al. 2008; Silvestro & Michalak 2012) and manually inspecting topologies to ensure that the same major groupings were recovered. Data sets were then concatenated in SEAVIEW 4.3.0 (Gouy et al. 2010) to construct a supermatrix alignment. Regions of the ITS data set with ambiguous site alignments were excluded (sites 100–112, 269–284, 302–319, 550–562, 816–904). PARTITIONFINDER 1.0.1 (Lanfear et al. 2012) determined the optimal evolutionary models and partition scheme for a partitioned analysis for both the core data set and mega-phylogeny. The alignment for the core data set is available online at Dryad Digital Depository (http://dx.doi.org/10.5061/dryad.gn4p4).

A multigene phylogeny was inferred using RAXMLGUI 1.2 (Stamatakis et al. 2008; Silvestro & Michalak 2012) executing 1000 rapid ML bootstraps replicates (Figure 5, Appendix). For further assessing clade support, MRBAYES 3.2 (Ronquist et al. 2012) was used for a Bayesian analysis of 1 000 000 generations using default priors until the standard deviation of split frequencies reached below 0.01. Outgroups were selected from the remaining three genera in the family Russulaceae: *Lactifluus deceptivus*, *Lactarius lignyotus* and *Multifurca zonaria*. Bootstrap values >70% and posterior probabilities >0.95 are considered as evidence for strongly supported relationships.

Clustering analyses of environmental sequences and metadata aquisition

All putative ITS sequences of the Russula clade, including the genus *Russula* and associated sequestrate genera *Macowanites*, *Cystangium*, *Gymnomyces* and *Martellia*, were extracted from GenBank using the bioinformatics program *emerencia* (Ryberg et al. 2009). To ensure adequate statistical power for diversification analyses and minimize the effects of low taxon sample size and high character state bias, we assembled a data set including >300 species using traits representing a minimum of 10% of the sampling, as suggested by Davis et al. (2013). Sequences were screened for chimeric assembly using a chimera checker (Nilsson et al. 2010) and manually pruned if sequence quality was low,

indicated by either long strings of ambiguous nucleotides or having >50% missing data. This data set is hereafter referred to as the 'GenBank data set.'

Two rounds of clustering analyses were performed on the GenBank data set to define molecular operational taxonomic units (MOTUs): one using cd-hit (Li & Godzik 2006) with a 99% identity and 80% coverage threshold and a second using CLUSTERTREE 1.0 (deposited in Dryad) with a 0.02 branch length cut-off using phylogenies inferred in FASTTREE (Price et al. 2009). FASTTREE was also used to visualize alignment quality and identify dubious sequences for exclusion based on extremely long branches. Representative sequences from each cluster were selected based on greatest sequence coverage, lowest number of polymorphic sites and whether they were identified to species. Representative sequences were aligned in MAFFT and then manually edited in MACCLADE. Due to the size of the data set and the variability of the region across the Russula clade, ClustalW was used to automatically align specific regions using SEAVIEW.

Biogeographic and ECM plant associate data were extracted from GenBank using a custom Perl script and by manually reading through primary literature. Biogeographic coding for tropical vs. extratropical used the latitudinal cut-off of the 23.5° parallels, and regional coding was performed by continent with the Middle East partitioned as the Eurasian territories from the Arabian Peninsula north through Turkey and east through Iran due to this region's intermediate position between Europe and the majority of Asia. ECM plant associates were inferred if the plant associate was reported in GenBank, a sequence was derived from a known root tip, or if the sample was reported from a monodominant forest (i.e. oak forest, well-described hardwood forest with no potential Pinaceae hosts, pine plot, etc.). The plant associates for clusters were used to determine the maximum level of host specificity of MOTUs supported by global sampling then coded as Pinaceae, angiosperm or generalist (i.e. associating with both Pinaceae and angiosperms) associates for the general data set. For a more refined analysis, MOTUs were also coded by ECM plant family, with generalist MOTU clusters coded as angiosperm or generalist as necessary.

Mega-phylogeny and BEAST analyses

The core data set was used as a backbone topology in RAXML to preserve higher-level relationships after merging the multilocus date with the GenBank data set. Additional gene sampling from clustered GenBank data (i.e. LSU, *rpb1* and *rpb2*) was incorporated into the supermatrix to estimate a mega-phylogeny. This was accomplished by aligning and concatenating associated sequences of LSU, *rpb1* and *rpb2* from any sequence of the same cluster set/MOTU (Smith et al. 2009). Using a backbone topology in RAXML allowed environmental MOTUs to be added to the starting tree using a maximum parsimony (MP) criterion. The tree was then optimized under normal ML parameters. The constrained mega-phylogeny was then ultrametricized using the Powell algorithm for nonparametric rate smoothing implemented in R8S 1.7 (Sanderson 2003). The core data set

was then excluded from the analysis to prevent taxon redundancies using the drop.tip function in the 'ape' package in R (Paradis et al. 2004).

To infer the crown ages of *Russula* and its major clades, the core data set was aligned with previously published multigene data sets of Russulaceae (Buyck et al. 2008; Van de Putte et al. 2012; Looney 2015) and outgroups through the AFTOL project (aftol.org). A chronogram of Russulaceae was inferred from three independent runs in BEAST 2 with 50 000 000 generations and a burn-in of the first 10% of trees generated so that all ESS values exceeded 200 (Figure 6, Appendix). Secondary calibrations were taken from Floudas et al. (2012) using normally distributed mean age estimates of Russulales, Boletales, Agaricales, Agaricomycetidae and the ancestral node of all three orders.

Ancestral state reconstructions and diversification analyses

Ancestral state reconstruction was performed using MP and ML approaches in MESQUITE 2.75 (Maddison & Maddison 2001) and Bayesian estimation in BAYESTRAITS V2 (Pagel et al. 2004). Significance in the ML and Bayesian analyses was determined by comparison of the negative log-likelihood of the character states with a difference threshold of 2. To test whether biogeographical range or plant association p is conserved in clades, the distributions of the traits on the phylogeny were tested for phylogenetic conservatism using PHYLOCOM 4.2 (Webb et al. 2008). Mann–Whitney U-tests were performed under different assumed sampling biases by incrementally reducing the biased state ages by 10% for geography and plant association character sets to test for differences in mean ages using the 'STATS' package version 3.2.1 in R. Diversification rates associated with geography and plant association were analyzed using the binary state speciation and extinction (BiSSE) model, the BiSSE–node enhanced state shift (BiSSE-ness) model for detecting cladogenetic shifts associated with character states, and the geographic state speciation and extinction (GeoSSE) model, a variant of the BiSSE model that allows species to occupy both binary states simultaneously (i.e. widespread or generalist). SSE analyses were implemented in the R package 'DIVERSITREE' (Maddison et al. 2007; FitzJohn et al. 2009; Goldberg et al. 2011; Magnuson-Ford & Otto 2012). Maximum-likelihood outputs from the models were tested and compared using the 'anova' function in R, and parameter estimates were found using a Markov chain Monte Carlo (mcmc) method using 1000 steps. To test for the effects of sampling bias for character states, 10 iterations of the mcmc analysis were performed with assumed sampling biases at 10% increments for 1000 steps implemented in the 'DIVERSITREE' package in R (Figure 7, 8, 9, & 10, Appendix). Finally, a BAMM approach (Rabosky et al. 2014) for trait independent analysis of diversification rate shifts was employed using the BAMMtools package in R to minimize the problem highlighted by Rabosky & Goldberg (2015) in which a single shift in diversification rate in a single diverse clade can bias estimates for that trait throughout the entire tree (Figure 11, Appendix).

Results

Analysis of clustered MOTUs suggests both biogeographic distribution and plant association are phylogenetically overdispersed

A total of 3510 ITS sequences of *Russula* were extracted from GenBank, with 3337 sequences resulting from a search for 'Russula' and 173 from searches for 'Macowanites', 'Cystangium', 'Gymnomyces' and 'Martellia'. A total of 162 sequences were excluded from the analyses due to low sequence quality, low coverage or as chimeric sequences. From the initial total, 21.6% of the sequences were already identified to species while 78.4% represented unidentified or environmental sequences from soil or root sampling.

Clustering analyses resulted in a phylogenetic tree with 1064 MOTUs (Figure 1). Of these, 202 were unique to the tropics, 844 were extratropical and 18 were found in both areas. An analysis of phylogenetic conservatism of geographic states across the mega-phylogeny showed that biogeographic states are phylogenetically overdispersed with a net relatedness index (NRI) of –2.1046 (P = 0.017). At a continental scale, North America had the greatest number of represented MOTUs at 441, with Europe and Asia also having a high number of MOTUs at 295 and 225, respectively (Figure 2). Most tropical MOTUs (105 or 51%) were sampled from Africa. Over 93% of MOTUs were recovered as endemic to a single continent, with the most range overlap detected between North America and Europe, which shared 62 MOTUs.

For ECM plant associate data, 158 of the MOTUs were recovered as associates of Pinaceae, 443 MOTUs as angiosperm associates, 60 as generalists and 403 were equivocal with no metadata available. An analysis of phylogenetic conservatism of plant associate states across the mega-phylogeny shows these states are also phylogenetically overdispersed with a NRI of –2.141 (P = 0.016), indicating that plant association is highly labile within clades. *Russula* MOTUs were recovered from 16 different plant families, with 25% of MOTUs associated with only Fagaceae, 24% with Pinaceae, 12% as angiosperm generalists and 9% as generalists (angiosperm and Pinaceae). Other notable ECM plant families include the Fabaceae and Dipterocarpaceae, which make up a large proportion of tropical plant associates, and the Myrtaceae that comprise many of the south temperate associates in Australia. Ancillary ecological roles were investigated, and 17 MOTUs (17%) were recovered as hosts for orchids or achlorophyllous members of the Ericaceae (Table 5, Appendix). Fifty-three MOTUs (5%) were recovered with a gasteroid morphology, and not all of the members in these clusters shared this morphology.

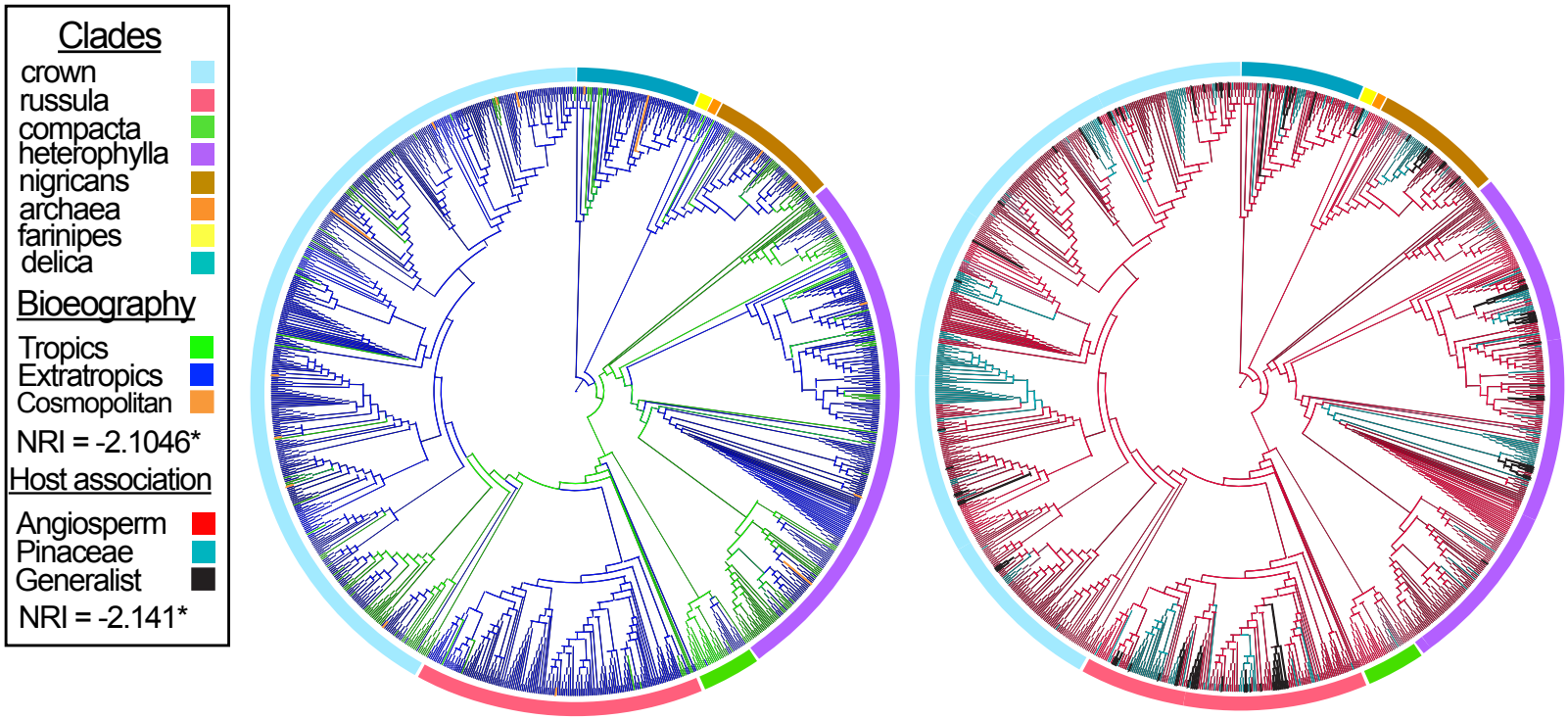


Figure 1. Maximum parsimony ancestral reconstructions of geography (left) and host association (right) along an ultrametric mega-phylogeny of environmental *Russula* MOTUs inferred using r8s. Major clades are designated by colour. Geographical and host tree metadata associated with MOTU clusters are designated with coloured lines at the tips, with equivocal tips inferred from the analysis. Areas are coded dark blue for extratropical distribution, green for tropical distribution and orange for cosmopolitan distribution. Plant association data are coded red for angiosperm association, aqua for Pinaceae association and black for generalist. Net relatedness indices produced in phylocom indicate phylogenetic overdispersion for both character sets.

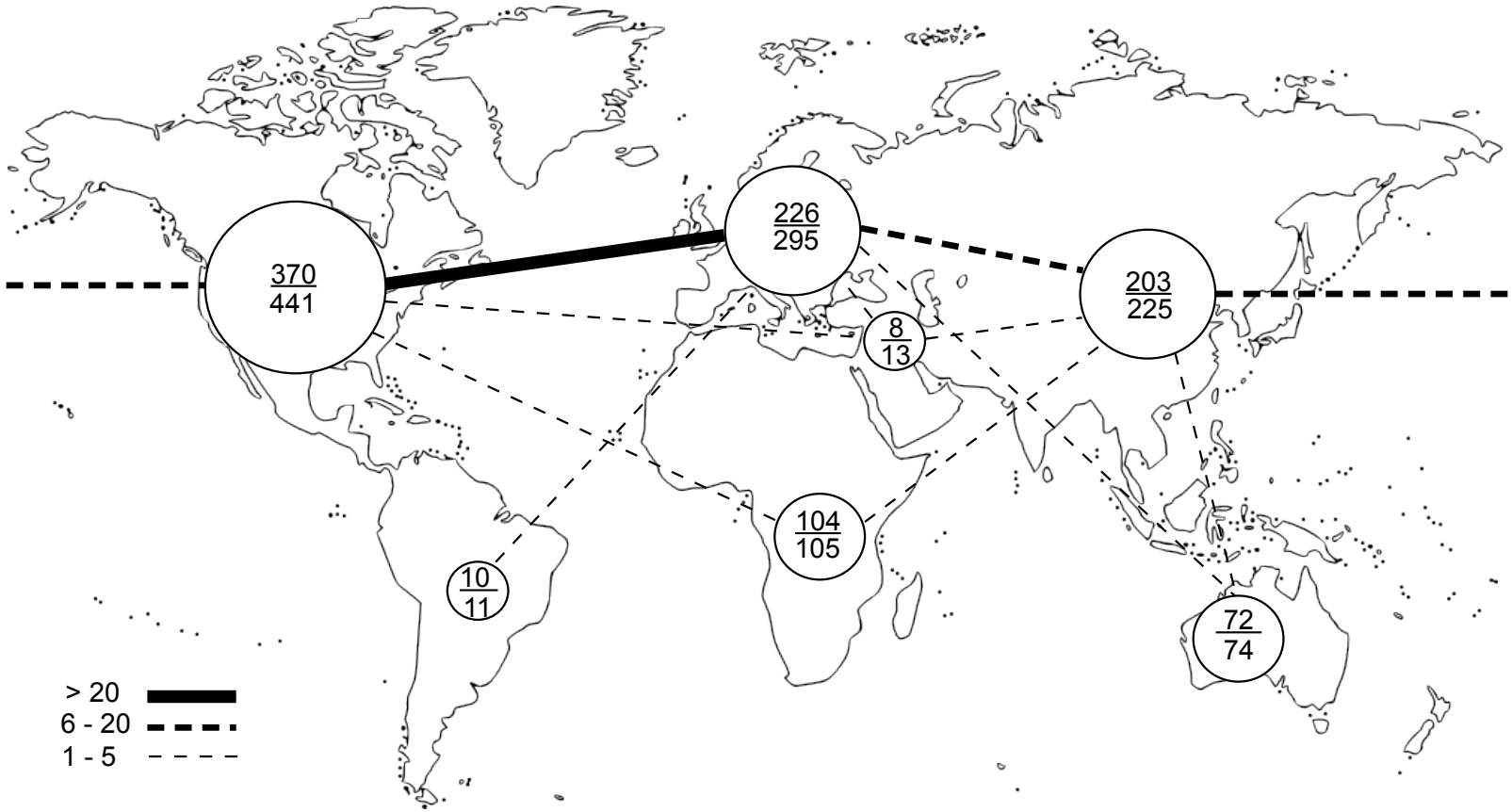


Figure 2. Map projection of the global distribution of *Russula* MOTUs. The areas of circles are scaled by the number of MOTUs relative to total MOTUs recovered. Top numbers represent number of endemic MOTUs. Bottom numbers indicate total number of MOTUs. Lines represent number of overlapping distributions for widespread MOTUs.

Ancestral reconstruction and molecular clock methods suggest an extratropical origin of Russula associated with angiosperms during the Palaeogene

The ancestral range of *Russula* was resolved with statistical support from ML and Bayesian inference as extratropical (Table 1). The delica, nigricans, archaea and farinipes clades were all resolved by ML as most likely having an extratropical origin with statistical support. The ancestor of heterophylla, russula, compacta and crown clade was ambiguous according to ML analysis, but it and all subtending major clades except the russula clade were resolved as tropical by MP analysis. The compacta clade had the highest likelihood support for a tropical ancestry. In a multistate reconstruction separating north and south temperate, as well as Neotropical and palaeotropical MOTUs, we could not reject a palaeotropical origin for *Russula*. However, support was much higher for a north temperate origin. This was true for most of the major clades except for delica, farinipes, archaea and russula clades, which were all significantly supported as having a north temperate origin. The four-state parsimony reconstruction agrees with the binary model, where tropical groups originated in the palaeotropics.

Ancestral plant association was reconstructed as ambiguous between angiosperm and Pinaceae under ML and Bayesian analyses, but all major clades were inferred as having an angiosperm association according to MP. A multistate reconstruction of major ECM plant families refuted an ancestral association with Myrtaceae for *Russula* and some individual major clades, yet the plant association reconstruction was ambiguous for all other families. MP reconstruction of plant family association supported either an ancestral association between Pinaceae or Fagaceae for all temperate clades except the russula clade, inferred as Fagaceae or Fabaceae. Fagaceae was inferred as the ancestral association for tropical clades under parsimony.

Using secondary time calibrations, *Russula* split from Lactarius and Multifurca ca. 55 (41–60) million years (MY) ago with a crown age of 44 (33–55) MY (Figure 6, Appendix). Of the eight major clades, heterophylla was inferred as the oldest group at 42 MY, with compacta second oldest at 37 MY old (Table 1). The youngest major clades inferred were the delica, farinipes, russula and crown clades, all around 30 MY old. Comparisons using a Mann–Whitney *U*-test of the taxon age for biogeographic ranges showed that tropical taxa, on average, are significantly older, with an average age of 7.8 MY, compared to extratropical taxa with an average age of 3.3 MY (Figure 3). Accounting for potential taxon sampling biases from the north temperate zone, this effect holds true if our sampling misses <2 tropical species for every one extratropical species (50% bias). Angiosperm associates, with a mean age of 5.2 MY old, were found to be, on average, significantly older than Pinaceae associates with an average of 2.5 MY. This effect holds true if sampling our sampling misses <1.25 angiosperm associates for every Pinaceae associate (20% bias).

Table 1. Crown ages and ancestral character states reconstructed for *Russula* and major clades.

Clades	Age	Geography binary					Host binary					Geography 4-state			Host family 6-state		
	MY	Geog PP	PP state	Geog ML	ML state	Geog MP	Host PP	PP state	Host ML	ML state	Host MP	ML	ML state	MP	ML	ML state	MP
root	43.96	0.88*	Temp	0.89*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.78/0.21*	**Ntem**/Ptro	Ntem	0.55	Fag	Pin,Fab
all except delica	43.52	0.79	Temp	0.89*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.78/0.22*	**Ntem**/Ptro	Ntem	0.54	Fag	Pin,Fab
het/rus/com/cro	43.18	0.64	Trop	0.88	Temp	Trop	0.50	Equi	0.50	Equi	Angi	0.48/0.52*	Ntem/**Ptro**	Ptro	0.53	Fag	Fab
nigricans/arc/far	42.78	0.88*	Temp	0.90*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.88/0.12*	**Ntem**/Ptro	Ntem	0.49	Fag	Pin,Fab
heterophylla	42.17	0.50	Trop	0.86	Temp	Trop	0.50	Equi	0.50	Equi	Angi	0.37/0.64*	Ntem/**Ptro**	Ptro	0.5	Fag	Fab
rus/com/cro	42.14	0.65	Trop	0.84	Temp	Trop	0.50	Equi	0.51	Angi	Angi	0.45/0.55*	Ntem/**Ptro**	Ptro	0.49	Fag	Fab
nigricans & arc	37.88	0.87	Temp	0.93*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.95*	Ntem	Ntem	0.5	Equi	Pin,Fab
compacta	37.03	0.82	Trop	0.53	Temp	Trop	0.50	Equi	0.50	Equi	Angi	0.12/0.87*	Ntem/**Ptro**	Ptro	0.48	Fab	Fab
nigricans	36.65	0.80	Temp	0.93*	Temp	Temp	0.50	Equi	0.51	Pina	Angi	0.98*	Ntem	Ntem	0.5	Equi	Pin,Fab
archaea	33.73	0.69	Temp	0.90*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.97*	Ntem	Ntem	0.5	Equi	Pin,Fab
russula/crown	33.25	0.72	Temp	0.82	Temp	Trop	0.50	Equi	0.50	Equi	Angi	0.50/0.50*	Ntem/Ptro	Ptro	0.5	Equi	Fab
delica	31.44	0.80	Temp	0.89*	Temp	Temp	0.50	Equi	0.50	Equi	Angi	0.95*	Ntem	Ntem	0.5	Equi	Pin,Fab
farinipes	30.57	0.83	Temp	0.91*	Temp	Temp	0.50	Equi	0.52	Angi	Angi	0.97*	Ntem	Ntem	0.32	Fag	Pin,Fab
russula	30.29	0.85	Temp	0.86	Temp	Temp	0.50	Equi	0.51	Angi	Angi	0.91*	Ntem	Ntem	0.39	Fag	Fab,Fag
crown	29.94	0.64	Trop	0.72	Temp	Trop	0.50	Equi	0.54	Angi	Angi	0.44/0.55*	Ntem/**Ptro**	Ptro	0.57	Fag	Fab

rus = russula clade
com = compacta clade
het = heterophylla clade
far = farinipes clade
arc = archaea clade
cro = crown clade
Temp = Temperate
Trop = Tropical
Angi = Angiosperm associate
Pina = Pinaceae associate
Equi = Equivocal
Geography 4-state = north temperate (Ntem), neotropics (Ntro), paleotropics (Ptro), and south temperate (Stem)
Host family 6-state = Pinaceae (Pin), Betulaceae (Bet), Dipterocarpaceae (Dip), Fabaceae (Fab), Fagaceae (Fag), and Myrtaceae (Myr)

* indicates significance based on the difference of –LnLik greater than 2
bold indicates states that have a higher likelihood when multiple states are found significant

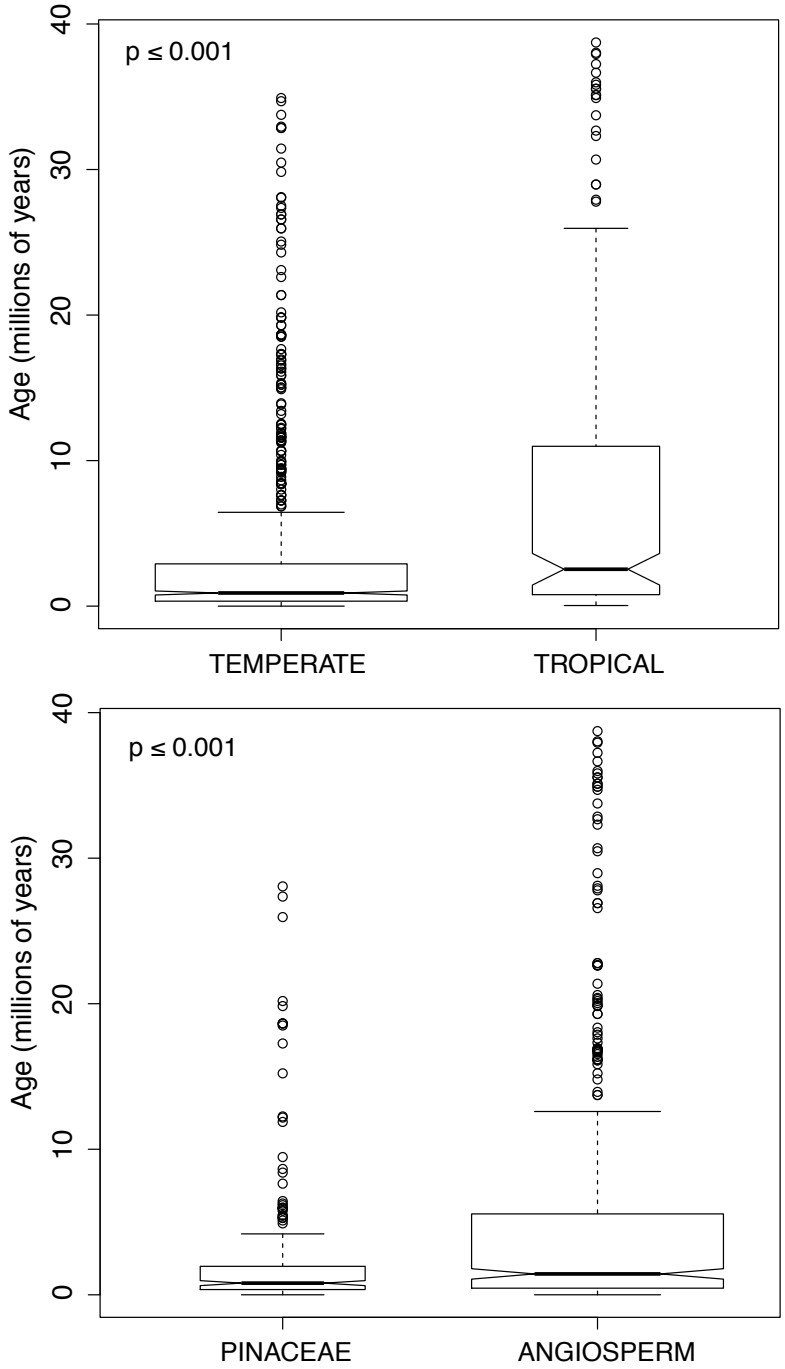


Figure 3. Boxplot comparing average taxon age based on terminal branch lengths of taxa from a secondarily time-calibrated mega-phylogeny with ranges in the tropics or extratropics (Top) and host association with Pinaceae or angiosperms (Bottom). P-values resulted from nonparametric Mann–Whitney *U*-tests.

State speciation-extinction models suggest higher rates of diversification for Russula *in the extratropics and in association with Pinaceae*

Model testing for GeoSSE, BiSSE and BiSSE-ness in an ANOVA framework showed significant support for the full model in four of the five data sets (Table 2). The model best supported for the GeoSSE geography data set was one that constrained transition rates between character states to be equal, indicating that dispersal between the tropics and extratropics is bidirectional. The full models for all other analyses were supported over models constraining speciation and extinction rates to be equal and pure-birth models, demonstrating that for all data sets diversification patterns differ between character states and extinction rates should be estimated. For both GeoSSE data sets, the best models were supported over models that constrained combined states to zero, indicating that speciation rates for widespread and host generalist taxa should be estimated. For BiSSE analyses, the full model was supported over models constraining transition rates as equal, demonstrating that rates of biogeographical and plant associate expansion or restriction are unidirectional. Finally, the BiSSE-ness analysis for plant association found the full cladogenic model supported over an anagenesis model of diversification, indicating that host switches are driving cladogenic events.

ML estimates of the best model were used as starting values for Bayesian inference of model parameters (Table 3). Rates of diversification were found to be significantly higher in extratropical lineages than tropical lineages, with extratropical lineages having a positive diversification rate and tropical lineages having a mean estimate of a negative rate, although we cannot reject a neutral diversification rate (Figure 4A). Diversification rate estimates for host specificity support a higher diversification rate with Pinaceae-associated taxa over angiosperm- associated taxa, with angiosperm-associated taxa having a negative diversification rate. However, we were not able to reject Pinaceae-associated MOTUs with a neutral diversification rate (Figure 4B). Transitions from Pinaceae association to angiosperm association are estimated to occur at rates 15.3 times higher than from angiosperm to Pinaceae. Diversification rate estimates for biogeographic range indicates that widespread taxa are diversifying at the same rate as those restricted to either the tropics or extratropics (Figure 4C). Transition rates, however, are more biased towards range contraction at rates 3.5 times higher than range expansions. Diversification rate estimates for host specificity indicate that host generalists are diversifying much faster than host specialists, with host specialists having a negative diversification rate (Figure 4D). Transition rates, however, are much more biased towards host specialization with rates being 5.6 times higher than range expansion events. These findings hold true under moderate taxon sampling biases (Figures 7, 8, 9, and 10, Appendix).

Table 2. Model comparisons for BiSSE, BiSSE-ness, and GeoSSE analyses.

	Df	lnLik	AIC	ChiSq	Pr(>|Chi|)	
GeoSSE geography models (Tropical vs. Extratropical)						
full	7	1000.7	-1987.3	NA	NA	
no.sAB	6	978.8	-1945.5	43.8	0	***
eq.div	5	934.2	-1858.4	133.0	0	***
no.mu	5	742.0	-1474.1	517.3	0	***
eq.trans	6	1000.3	-1988.7	0.7	0.4	
GeoSSE plant association models (Angiosperm vs. Pinaceae)						
full	7	278.1	-542.2	NA	NA	
no.sAB	6	276.1	-540.1	4.0	0	*
eq.div	5	248.0	-485.9	60.3	0	***
no.mu	5	66.3	-122.5	423.6	0	***
BiSSE geography models (Endemic vs. Widespread)						
full	6	1121.9	-2231.9	NA	NA	
eq.trans	5	1081.4	-2152.9	81.0	0	***
eq.div	4	1008.4	-2008.8	227.0	0	***
no.mu	4	957.7	-1907.5	328.4	0	***
BiSSE plant association models (Specific vs. Generalist)						
full	6	561.0	-1109.9	NA	NA	
eq.trans	5	480.7	-951.5	160.5	0	***
eq.div	4	397.2	-786.5	327.4	0	***
no.mu	4	512.9	-1017.8	96.2	0	***
BiSSE-ness plant association models (Angiosperm vs. Pinaceae)						
full	10	623.9	-1227.7	NA	NA	
no.trans	9	622.8	-1227.5	2.2	0.1	***
eq.div	8	432.8	-849.6	382.1	0	***
no.mu	9	618.1	-1218.2	11.5	0	***
no.pc	8	581.6	-1147.2	84.6	0	***

full = model with all parameters
eq.trans = model with transition rates constrained as equal
no.sAB = model with no dual-state speciation
eq.div = model with diversification constrained as equal
no.mu = model with extinction constrained to 0
no.pc = model with no cladogenic diversification

* = $P \leq 0.05$; ** = $P \leq 0.005$; *** = $P \leq 0.0005$

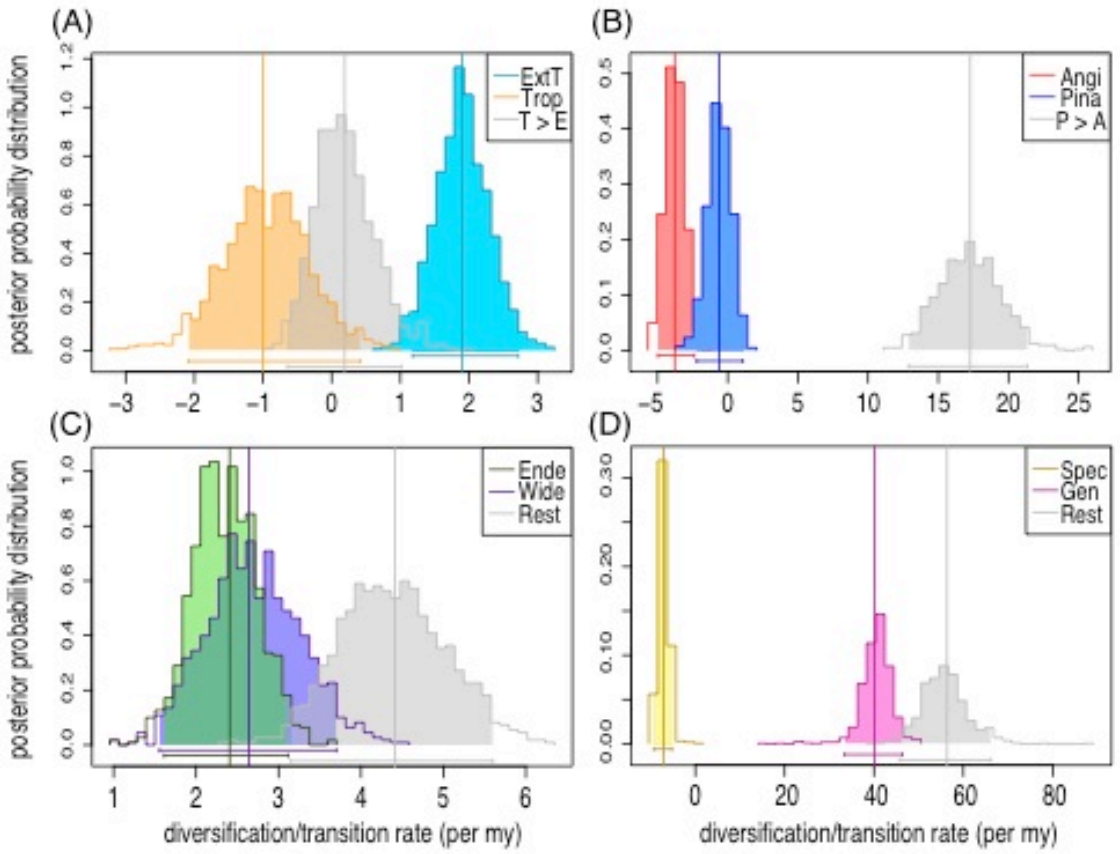


Figure 4. Posterior probability density means and standard error representing relative diversification (k-l) and dispersal rates for geographic and host state-specific models for *Russula*. (A) Estimates for extratropical MOTUs (light blue) and tropical MOTUs (orange) with differential transition rates from tropics to extratropics (grey) for the best-supported equal transition GeoSSE model. (B) Estimates for angiosperm-associated MOTUs (red) and Pinaceae-associated MOTUs (blue) with differential transition rates from Pinaceae association to angiosperm association (grey) for the best-supported full-parameter GeoSSE model. (C) Estimates for endemic MOTUs (green) and widespread MOTUs (purple) with differential transition rates of contraction from widespread to endemic (grey) for the best-supported full-parameter BiSSE model. (D) Estimates for host-specific MOTUs (yellow) and host generalist MOTUs (pink) with differential transition rates of restriction from host generalist to specialist (grey) for the best-supported full-parameter BiSSE model

Table 3. Maximum Likelihood estimates of parameters for the best model for BiSSE and GeoSSE analyses.

	sA	sB	sAB	xA	xB	dAB	dBA	DA	DB	TB->A	T ratio
GeoSSE Geog	28.9	6.6	63.8	27.2	8.2	0.9	0.9	1.7	-1.6	0.4	1
GeoSSE Plant	11.0	33.3	14.2	16.1	35.0	1.4	20.8	-5.1	-1.7	19.5	15.3
BiSSE Geog	36.3	2.7	N/A	34.1	0	2.1	7.2	2.2	2.7	5.1	3.5
BiSSE Host	0.5	74.9	N/A	7.2	40.2	19.4	109	-6.7	34.7	89.6	5.6

sA = speciation rate A
sB = speciation rate B
sAB = speciation rate for dual-state
xA = extinction rate A
GeoSSE Geog = A for Extratropical; B for Tropical
GeoSSE Plant = A for Angiosperm; B for Pinaceae
BiSSE Geog = A for Endemic; B for Widespread
BiSSE Plant = A for Host Specific; B for General
xB = extinction rate B
dAB = dispersal rate from A to B
dBA = dispersal rate from B to A
DA = net diversification rate A
DB = net diversification rate B
TB->A = transition rate from B to A
T ratio = transition rate B to A divided by A to B

Discussion

Into and out of the tropics

Russula is among the most taxonomically diverse ECM lineages in the tropics (Buyck et al. 1996; Tedersoo & Nara 2010). However, ancestral area analyses support that the genus is an ancestrally temperate group. In addition, diversification rate analyses support a higher net rate of diversification among taxa in extratropical regions. This suggests a complex biogeographic history for *Russula*, and likely most other ECM lineages, which falls counter to the predictions of established biogeographic hypotheses (Wiens & Donoghue 2004; Jablonski et al. 2006; Jansson et al. 2013).

The most recent common ancestor of *Russula* was probably an angiosperm associate that began to diversify ca. 40 MY ago during the Eocene in North temperate regions of Eurasia and/or North America. The late Eocene marked the beginning of transition to icehouse Earth conditions where, despite large fluctuations in CO2 levels, Antarctic ice began to form and global climates began the period of cooling leading to modern conditions (Lear et al. 2008).

It has been suggested that diversification of ECM fungi was facilitated by an expanded niche space caused by cooling climates (Bruns et al. 1998; Ryberg & Matheny 2012), and *Russula* is a group that appears to be well adapted to temperate climates and able to occupy these novel niches. The early history of the group shows the divergence of the delica, farinipes, archaea and nigricans clades occur in the north temperate zone. There is evidence for switches to the tropics in the ancestors of the heterophylla, compacta, and crown clades with a major reversal back to the extratropics in the most recent common ancestor of the russula clade. Since this early history, transitions between the tropics and extratropics have been frequent events in the evolutionary history of *Russula* with at least 47 independent shifts to the tropics and a comparable number of shifts to the extratropics. Only the compacta clade is composed of more tropical taxa than extratropical taxa, thus representing the only major tropical clade in *Russula* based on current sampling.

Several ECM clades have been hypothesized as tropical in origin (Matheny et al. 2009; Kennedy et al. 2012; Wilson et al. 2012; Sánchez-Ramírez et al. 2015b). The ancestral origin of Inocybaceae, *Amanita* sect. *Caesareae*, and most ECM clades of Sclerodermatineae (Matheny et al. 2009; Wilson et al. 2012; Sánchez-Ramírezet al. 2015b). The ECM Sebacinaceae is the only major ECM lineage that has been shown explicitly to have a north temperate origin (Tedersoo et al. 2014b). No ECM lineages have yet been found endemic to the Neotropics, nor have any groups been reconstructed with a Neotropical origin. South temperate taxa in the family Inocybaceae are largely derived from north temperate progenitors, and Neotropical taxa have been shown to have immigrated from elsewhere (Matheny et al. 2009). Two lineages, *Austropaxillus*

and ECM Hysterangiales, have been inferred as having south temperate origins (Hosaka et al. 2008; Skrede et al. 2011).

While *Russula* has been inferred as having a north temperate origin, the family Russulaceae may have its origins in the tropics, given that *Lactifluus*, an ECM genus of over 120 species that has been hypothesized as the sister clade to the rest of Russulaceae, is largely a tropical clade (Verbeken et al. 2011). In this case, *Russula* would represent a major clade that diversified outside of its ancestral range to a greater extent than the other major clades of the ECM lineage (i.e. Russulaceae), similar to what has been found in some other ECM lineages (Matheny et al. 2009; Kennedy et al. 2012).

Although ECM clades vary greatly in age, significant diversification episodes have coincided with specific geologic periods during the evolution of these groups. The oldest ECM lineage (Tuberaceae, Ascomycota) is a cosmopolitan group ca. 160 million years old (Bonito et al. 2013) originating in the late Jurassic, while *Austropaxillus* (order Boletales) has been identified as a young ECM lineage with a mean age of 22 million years (Skrede et al. 2011). We recovered the ECM lineage Russulaceae to have a mean age of 76 million years, originating during the late Cretaceous, which is consistent with ages of several ECM clades of Agaricales (Ryberg & Matheny 2012). The crown age of *Russula* (44 MY) during the Eocene corresponds with ages of many of the major clades of Tuberaceae (30–54 MY) as well as major clades within the ECM Sebacinaceae (30–45 MY) (Bonito et al. 2013; Tedersoo et al. 2014b). *Russula*, therefore, conforms with an emerging pattern in which the origin of ECM association is ancient, in this case ECM evolving in the ancestor of Russulaceae during the late Cretaceous, but diversification of the major extant clades has occurred much more recently in the Eocene, during which the global climate began cooling and temperate conditions expanded.

Higher diversification rates in the extratropics explain a reversal of the latitudinal diversity gradient (LDG)

For many of the proposed explanations of the LDG pattern, biological justifications could also apply to groups originating outside the tropics. The ‘biogeographical conservatism hypothesis’ has been proposed as an alternative to the ‘tropical conservatism hypothesis’, which suggests that thermal or climatic tolerances may restrict groups to certain environmental niches regardless of whether they originate in the tropics (Pyron & Burbrink 2009). As an alternative to the ‘out of the tropics’ model, an ‘into the tropics’ model would suggest that lineages outside the tropics are not dispersal limited in regard to the tropics but those lineages can continue to diversify alongside endemic extratropical lineages for an overall greater accumulation of species. Some processes proposed for the ‘diversification rate hypothesis’ could also apply to groups with an extratropical ancestry. This includes an accelerated rate of molecular evolution, relatively stable climatic conditions, or an expanded niche space due to biotic and abiotic factors. However, the extratropics cannot be said to have seen gross expansions

compared with modern conditions considering the relatively constant cooling trend of global climates. For an explanation of this pattern applied to nontropical groups as well, we propose the 'generalized diversification rate' hypothesis, which states that patterns of diversity can be explained by regional abiotic or biotic factors that promote an increased diversification rate regardless of the biogeographic origin of a group or dispersability into or out of the region.

Diversification patterns in *Russula* support the 'generalized diversification rate' hypothesis as an explanation of the reversed LDG. Lineages of *Russula* in the extratropics exhibit a higher rate of net diversification as they transition into and out of the tropics at relatively equal rates. A pattern of phylogenetic niche conservatism has been proposed as good support for the 'tropical conservatism hypothesis', where we should expect tropical lineages to disperse infrequently into the extratropics, thus allowing tropical clades to diversify or persist as long branches (Crisp & Cook 2012). We do see the tropics acting as a museum with tropical taxa having a much lower extinction rate and higher average species ages (Figure 3). However, we found the distribution of tropical MOTUs to be phylogenetically overdispersed, indicating that transitions have occurred into and out of the tropics frequently during the last 40 MY. Additionally, an 'into the tropics' model can be rejected as transition rates between the tropics and extratropics were found to be equal in *Russula*. Diversification patterns in *Russula* agree with the findings of Sánchez-Ramírez et al. (2015a) that extratropical ECM taxa have a higher speciation rate than tropical taxa; however, extinction was indicated as a significant variable in our models for *Russula* diversification. Given these trends, extinction in tropical environments may be driven by an unavailability of abundant niche space from fewer soil horizons, more fragmented host distributions, and a lack of community partitioning due to a lower host lineage diversity (Tedersoo & Nara 2010).

Tedersoo et al. (2012) suggested that clade age might explain why ECM fungi are more species rich at temperate latitudes than in the tropics. If this is correct, then temperate lineages should be older and more diverse than tropical lineages. Kennedy et al. (2012) found no support for the 'clade age' hypothesis in the ECM genus *Clavulina*, which was found to be tropical in origin and containing several derived temperate lineages. One of these temperate lineages was found to be diversifying at nearly 2.5 times the rate elsewhere in the tree. With a north temperate origin, *Russula* provides a good test for the 'clade age' hypothesis. Diversification patterns in *Russula* reject the 'clade age' hypothesis and support an overall higher diversification rate for extratropical taxa as a generalized pattern, even when major clades are not restricted to the tropics or extratropics. The 'clade age' hypothesis is also confounded as a generalizable pattern for ECM fungi by the paucity of evidence for temperate origins for a majority of diverse ECM lineages.

In the extratropics, *Russula* is characterized by high speciation and extinction rates, indicating a high species turnover evident by the low average age of extratropical taxa. This finding is consistent with the prediction of Buyck et al. (1996) that temperate ECM fungi may experience higher competition due to

exposure to ‘foreign invaders’, whereas the tropics act like a museum because of the relative isolation from competition. We find some evidence for latitudinal optima described by Sánchez-Ramírez et al. (2015a) for Amanita sect. Caesareae. In *Russula*, there is a much higher rate of transition to either the tropics or extratropics rather than range expansion to both. This could indicate that the subtropics represent a barrier for dispersal and that different adaptations are required for surviving in tropical vs. extratropical habitats and mycorrhizal communities.

Host switching is an important driver of diversification in Russula

To explain the reversal of the LDG in ECM fungi, increased ECM plant diversity in temperate regions was proposed as a driving evolutionary force, but neither codiversification nor host switching has been investigated in this context (Kennedy et al. 2012; Tedersoo et al. 2012; Põlme et al. 2013). The diversification of major clades in *Russula* corresponds to the time of diversification for major ECM plant lineages, including Fagaceae, Betulaceae, Salicaceae, Malvaceae, Cistaceae and Dipterocarpaceae (Bell et al. 2010). This is consistent with the hypothesis that codiversification with hosts or host switching may have been an important driver of diversification for ECM fungi. Evolution of ECM plant diversity makes sense as a driver for the reversed LDG pattern in ECM fungi as several diverse ECM plant lineages (e.g. Myrtaceae, Fagaceae and Pinaceae) have their diversity centres outside the tropics (Pryor 1959; Richardson 2000; Nixon 2006). ECM plant lineage association was found to be conserved in major clades in the Agaricales, such as Cortinarius, Hygrophorus and Inocybaceae, but conservation was not found in others (Ryberg & Matheny 2012). If codiversification is an important driver for ECM fungal diversity, then we should expect *Russula* clades to be host-restricted to particular plant lineages genera or families. We find that plant association in *Russula* is not conserved by plant lineage at the family level, evidenced by phylogenetic overdispersion and the lack of signal for inferring ancestral plant associations. We also find support for a model showing that host switching is driving cladogenic events over an anagenic model of host diversification. With these analyses combined, there is strong evidence that host switching is an important driver for diversification in *Russula* and is more plausible than a codiversification scenario of diversification.

Although it has not been found to be an important driver for diversification of *Russula*, there is some evidence that codiversification may be an important process for select ECM fungi and for ECM plant lineages in general. ECM plants comprise select lineages of Gnetaceae, Pinaceae and numerous lineages of angiosperms, including members of Betulaceae, Dipterocarpaceae, Fabaceae, Fagaceae, Juglandaceae, Myrtaceae, Nothofagaceae and Salicaceae (Brundrett 2009) among others. Many fungal lineages containing ECM fungi, including Russulales, have been found to be younger than the diversification of angiosperms (Hibbett & Matheny 2009). Consistent with these findings, ancestral

plant associates for a number of ECM lineages, now including *Russula*, have been inferred as angiosperm (Matheny et al. 2009; Ryberg & Matheny 2011; Wilson et al. 2012; Bonito et al. 2013). Only the ECM Sebacinaceae has been recovered as having an ancestral association with Pinaceae (Tedersoo et al. 2014b). Studies of the ECM plant genus *Alnus* have found historical distributions consistent with their associates, giving strong support for codispersal for this plant lineage with their associates (Kennedy et al. 2011b; Põlmeet al. 2013). Another ECM plant group that shows a strong signal of association and, potentially, codiversification with its fungal partners is *Pinus*, whose species are nearly ubiquitous with the ECM genera *Suillus* and *Rhizopogon* (Bruns et al. 2002). Studies looking at codiversification from the perspective of species-rich ECM plant lineages, such as *Quercus* or *Eucalyptus*, have not been attempted. Nonetheless, if ECM fungi are codiversifying with their plant associates, this may be an important process for diversification of ECM plants as particular host-specific fungal associates may be necessary partners for those plant lineages, whereas host switching may be a primary process by which most ECM fungal lineages diversify.

A surprising result from the GeoSSE model comparison of plant association was that MOTUs associated with Pinaceae have higher speciation rates than the ones associated with angiosperms. In this case, we can see an evolutionary source–sink dynamics, where a majority of species initially evolve as associates with Pinaceae but preferentially switch to angiosperm hosts where they either expand their host range or go extinct. A potential mechanism to explain this pattern would be orogenesis events that can act like a species pump similar to glacial refugia (Sedano & Burns 2010; Wang et al. 2012). Many Pinaceae species are montane and will probably track elevational gradients as mountains are uplifted. These events are ideal for populations not able to track this migration due to dispersal limitation or thermal tolerances to become isolated and either speciate or switch to an angiosperm host. Populations that are able to track Pinaceae associates may have opportunities to host switch and speciate with other members of Pinaceae in different life zones or community types (i.e. pine to spruce dominant community) (Tang & Ohsawa 1997). It is also probably that climate fluctuations over geological time create this effect at the temperate–boreal interface (Sandel et al. 2011). *Russula* generalists that associated with both Pinaceae and angiosperms have a higher diversification rate than more host-specific species, which also indicates that host switching or expansion may be more important drivers than co-evolution with the plant associate. Again, host specificity is characterized by an evolutionary source–sink dynamic, where speciation occurs with generalist species, but their host ranges are frequently restricted, which may increase extinction rates.

Potential for additional drivers of diversification

An important criticism of trait-based diversification analyses broached by Rabosky & Goldberg (2015) is that a hidden trait or traits may be driving

diversification patterns that, by chance, may be correlated with the trait being tested. This criticism is not a concern for our latitudinal assessment, as we are interested in analysing a pattern explicitly to discover the evolutionary process. This criticism is relevant when considering whether ECM plant associate lineage is driving ECM fungal diversification, but this issue is more a problem of interpretation than any flaw in the models. ECM plant associations have been proposed as potential drivers of the reversed LDG pattern, and our results are consistent with this. However, this is not to say that other associated factors may not be more important at other spatial scales, including root stratification (Kalliokoski et al. 2010), mycorrhizal root signaling (Felten et al. 2009), ability to associate with arbuscular mycorrhizae (Kennedy et al. 2011a,b), or even something external to the associate such as community type (McGuire et al. 2013) or stratification of the soil (Rosling et al. 2003). A final possibility is that key adaptations of the fungi may be playing a role in diversification with different plant host lineages, as adaptive radiations in fungi have been shown to be driven by a combination of environmental opportunity and phonological adaptations to take advantage of that opportunity (Gaya et al. 2015). For *Russula*, this may include adaptations to labile characters such as changes in spore morphology in response to changing environments such as temperature and moisture for differential dispersability and germination, pigmentation of the pileus cuticle as protection against radiation or to attract animal dispersal vectors (Eberhardt 2002), different suites of oxidative enzymes for accessing nutrients in recalcitrant plant matter or expansions in small secreted proteins used in root colonization (Kohler et al. 2015). By identifying traits that support a pattern of diversification, we can develop additional hypotheses to test for a ‘smoking gun’ trait, if one exists.

Sampling and methodological considerations

Using a total data approach, we were able to achieve maximal global sampling of *Russula*; however, there are some caveats and biases inherent to this approach. The total number of recovered *Russula* MOTUs (1064) exceeds the number of currently accepted species in the genus (750–900 spp.) indicating that numerous novel species of *Russula* have not been formally described. The majority of the GenBank studies evaluated here originated in North America or Europe, which have the highest number of MOTUs. We recovered 441 MOTUs in North America, which is near the total number of species reported from both the USA (419 spp.) and Mexico (66 spp.) (Kong et al. 2002; Buyck 2007). A high number of MOTUs (62) are shared between North America and Europe, which closely agrees with the number of species described from Europe that are also reported in North America (87 spp.) (Buyck 2007). Although sampling bias towards the extratropics was anticipated and accounted for in our diversification analyses (Figures 7, 8, 9, & 10, Appendix), this bias may not be as pronounced given 1) the smaller land mass with available ECM habitat; 2) the lack of ECM plant richness in the tropics; and 3) the number of tropical *Russula* taxa described

compared to MOTUs recovered from molecular sampling efforts. The recovered number of MOTUs for tropical Africa (105) closely approximates the number of described species (129–_165 spp.) (Buyck et al. 1996; Verbeken & Buyck 2002). The total MOTUs recovered for the Neotropics (45) also matches well the number of described taxa from the region (42 spp.) (Buyck et al. 1996). Species estimates for tropical Asia are more difficult to obtain due to the application of traditional European names to species from this area, but any bias towards the extratropics in this region is probably offset by a lack of sampling from the temperate Himalayan region of south China where we should expect a high diversity coinciding with a high number of ECM plant lineages (Das et al. 2010). An assessment of *Russula* diversity for tropical Asia should be an objective for future studies. For south temperate sampling, we recovered 74 MOTUs from Australia and New Zealand, which exceeds the number of species described from this region, given that the largest study in the genus from this region describes 33 species (McNabb 1973). A few disjunct distributions of MOTUs are probably explained by local introductions from pine plantations (Dickie et al. 2010). Six MOTUs were recovered as having a holarctic distribution throughout North America, Europe and Asia, three of which were independently sampled by fourteen different GenBank studies (Table 5, Appendix).

Accounting for almost half of all of the MOTUs for which it was possible to retrieve ECM plant associate data, host preference for *Russula* strongly favours the Fagaceae (165) and Pinaceae (157). This is not surprising given that these families are the most species diverse ECM plant lineages in north temperate regions (Pryor 1959; Nixon 2006). *Russula* sequences were detected from 16 different plant families. Plant families where *Russula* was not detected but where we might expect to find *Russula* include Gnetaceae, Casuarinaceae and Cistaceae, which are mostly south temperate or tropical lineages (Brundrett 2009; Tedersoo & Põlme2012). Tedersoo et al. (2014b) hypothesized groups that associate with more plant lineages should be older, but this is not the case with *Russula*, a relatively young group that associates with nearly all known ECM plant lineages.

We used a robust, multigene phylogeny as a guide tree for the Genbank data set mega-phylogeny due to the variability of the ITS region from which most of the environmental data were based, which allowed the conservation of higher-level relationships. The final ultrametric topology was therefore dependent on relationships inferred based on the phylogeny of the core data set, where some nodes were not supported by bootstrapping or posterior probability. However, these clades were resolved in both maximum likelihood and Bayesian inference, giving some confidence for the topology. Taxon sampling for the core data set was also biased towards North American and European taxa due to reliance on the major classification systems, which are based on those regions. Cluster sets were considered regardless of cluster size, as excluding singleton cluster sets would reduce sampling beyond the necessary limits for SSE models. The calculated average of sequences per cluster set was 2.6, with 79% of sequences coded as extratropical, 12% as tropical and 9% as widespread. Given these

sample sizes, we can certainly infer presence data for all geography and some hosts, but we cannot be certain that the full geographic range or host is being captured for MOTUs. There is a stronger bias for tropical samples being undersampled, where fewer studies have been conducted and many MOTUs were only detected once. There must also be a bias towards recovering more host specialists, as there must be at least two sequences in the cluster set with conflicting hosts to be considered a generalist. It is probably that some MOTUs that could be considered generalists were not coded as such because none of their other hosts were sampled within their geographic range. Given these limitations, only potential geographic dispersal and host switches can be tested. Also, as there is no consistent sampling strategy for GenBank sequences, there may be biases in our ability to detect rare taxa from locations that have only had sampling done from fruit body or root collections. Our approach, however, was able to achieve much greater sampling than would be possible without a worldwide network of sampling researchers and sampling sites, and we propose that efforts should continue to report metadata for sequence data submitted to online data repositories and support databases for global sampling data such as UNITE (https://unite.ut.ee/), GBIF (http://www.gbif.org/) and fungimap (http://fungimap.org.au/).

Conclusions

Investigation of diversification patterns in fungi is challenging given the immense diversity of these groups, their cosmopolitan distributions and the necessity to approximate complete global taxon sampling across a phylogeny. Utilizing available sequence data from various environmental sources can help mitigate these challenges by allowing for a more complete assessment of global diversity and more accurate estimation of evolutionary patterns. Using state-specific diversification models, we found strong support for the ‘generalized diversification rate’ hypothesis as an evolutionary process accounting for high extratropical diversity in *Russula*. Application of these models to other lineages of fungi may confirm our findings as a generalizable pattern. We also found evidence that host switching is an important driver in *Russula* diversification, allowing us to generate new hypotheses about trait-driven diversification in fungi. For example, a finer-scale analysis comparing diversification between taxa from lowland tropical forests and montane tropical forests or between specific ECM plant lineages may indicate, counter to most other guilds of fungi (Tedersoo et al. 2014a), that climate effects are less important than host effects for ECM fungi. Also, while this and other studies have focused on evolutionary dynamics at the tropical interface, the boreal–temperate interface has also been highlighted as an important biogeographic boundary, where ECM begins to drop off northward as part of a unimodal distribution (Tedersoo et al. 2012, 2014a;). Future studies in *Russula* should examine this relationship to determine whether the same evolutionary or ecological forces are governing this pattern, especially considering that boreal forests are composed of a higher density of ECM plants

than most temperate forests, while temperate systems can contain a higher ECM plant species richness.

Acknowledgments

This work was financially supported by National Science Foundation awards DEB-0949517 to PBM and a Doctoral Dissertation Improvement Grant DEB-1501293 to BPL and PBM. Research funding was also received from the University of Tennessee through the Chancellor's Award Fund. The Mycological Society of America's Mentor Travel grant is gratefully acknowledged to finance travel to present and receive feedback concerning this research. The authors would also like to thank the Gulf States Mycological Society and Cumberland Mycological Society for the opportunity to collect specimens for this study. We would like to thank Slavomír Adamčík (Slovak Academy of Sciences) and Steven Miller (University of Wyoming) for technical training in identification of *Russula* species. We thank the curators and staff from herbaria at the University of Gent (GENT) and the University of Tennessee (TENN) for access to collections.

References

Alexopoulos C, Mims C, Blackwell M (1996) Introductory Mycology. John Wiley & Sons, New York City, New York.

Baas-Becking LGM (1934) Geobiologie of Inleiding tot de Milieukunde. WP Van Stockum & Zoon, The Hague, The Netherlands.

Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. American Journal of Botany, 97, 1296–1303.

Birkebak JM, Mayor JR, Ryberg M, Matheny PB (2013) A systematic, morphological and ecological overview of the Clavariaceae (Agaricales). Mycologia, 105, 896–911.

Bonito G, Smith ME, Nowak M et al. (2013) Historical biogeography and diversification of truffles in the Tuberaceae and their newly identified southern hemisphere sister lineage. PLoS ONE, 8, e52765.

Brundrett MC (2009) Mycorrhizal associations and other means of nutrition of vascular plants: understanding the global diversity of host plants by resolving conflicting information and developing reliable means of diagnosis. Plant and Soil, 320, 37–77.

Bruns TD, Szaro TM, Gardes M et al. (1998) A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. Molecular Ecology, 7, 257–272.

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

Buyck B (2007) A new initiative towards the study of *Russula* in the eastern USA. Pagine Di Micologia, 27, 81–86.

Buyck B, Atri NS (2011) A *Russula* (Basidiomycota, Russulales) with an unprecedented hymenophore configuration from northwest Himalaya (India). Cryptogamie, Mycologie, 32, 185–190.

Buyck B, Thoen D, Watling R (1996) Ectomycorrhizal fungi of the Guinea-Congo Region. Proceedings of the Royal Society of Edinburgh, 104B, 313–333.

Buyck B, Hofstetter V, Eberhardt U, Verbeken A, Kauff F (2008) Walking the thin line between *Russula* and *Lactarius*: the dilemma of *Russula* subsect. *Ochricompactae*. Fungal Diversity, 28, 15–40.

Crisp MD, Cook LG (2012) Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? New Phytologist, 196, 681–694.

Das K, Putte K, Buyck B (2010) New or interesting *Russula* from Sikkim Himalaya (India). Cryptogamie, Mycologie, 31, 373–387.

Davis MP, Midford PE, Maddison W (2013) Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. BMC Evolutionary Biology, 13, 1–11.

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

Eberhardt U (2002) Molecular kinship analyses of the agaricoid Russulaceae: correspondence with mycorrhizal anatomy and sporocarp features in the genus *Russula*. Mycological Progress, 1, 201–223.

Felten J, Kohler A, Morin E et al. (2009) The ectomycorrhizal fungus *Laccaria bicolor* stimulates lateral root formation in poplar and *Arabidopsis* through auxin transport and signaling. Plant Physiology, 151, 1991–2005.

FitzJohn RG, Maddison WP, Otto SP (2009) Estimating trait dependent speciation and extinction rates from incompletely resolved phylogenies. Systematic Biology, 58, 595–611.

Floudas D, Binder M, Riley R et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science, 336, 1715–1719.

Fontaneto D, Barraclough TG, Chen K, Ricci C, Herniou EA (2008) Molecular evidence for broad-scale distributions in bdelloid rotifers: everything is not everywhere but most things are very widespread. Molecular Ecology, 17, 3136–3146.

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes-application to the identification of mycorrhizae and rusts. Molecular Ecology, 2, 113–118.

Gaya E, Fern_andez-Brime S, Vargas R et al. (2015) The adaptive radiation of lichen-forming Teloschistaceae is associated with sunscreening pigments and a bark-to-rock substrate shift. Proceedings of the National Academy of Sciences USA, 122, 11600–11605.

Geml J, Laursen GA, Timling I et al. (2009) Molecular phylogenetic biodiversity assessment of arctic and boreal ectomycorrhizal Lactarius Pers. (Russulales; Basidiomycota) in Alaska, based on soil and sporocarp DNA. Molecular Ecology, 18, 2213–2227.

Goldberg EE, Lancaster LT, Ree RH (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. Systematic Biology, 60, 451–465.

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and
Evolution, 27, 221–224.

Hibbett DS, Matheny PB (2009) The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. BMC Biology, 7, 1–13.

Hillebrand H (2004) On the generality of the latitudinal diversity gradient. The American Naturalist, 163, 192–211.

Hoeksema JD (2010) Ongoing coevolution in mycorrhizal interactions. New Phytologist, 187, 286–300.

Hosaka K, Castellano MA, Spatafora JW (2008) Biogeography of Hysterangiales (Phallomycetidae, Basidiomycota). Mycological Research, 112, 448–462.

Jablonski D, Roy K, Valentine JW (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. Science, 314, 102–106.

Jansson R, Rodríguez-Castañeda G, Harding LE (2013) What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. Evolution, 67, 1741–1755.

Kalliokoski T, Pennanen T, Nygren P, Sievänen R, Helmisaari R-S (2010) Belowground interspecific competition in mixed boreal forests: fine root and ectomycorrhiza characteristics along stand developmental stage and soil fertility gradients. Plant and Soil, 330, 73–89.

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Briefings in Bioinformatics, 9, 286–298.

Kennedy AH, Taylor DL, Watson LE (2011a) Mycorrhizal specificity in the fully mycoheterotrophic *Hexalectris* Raf. (Orchidaceae: Epidendroideae). Molecular Ecology, 20, 1303–1316.

Kennedy PG, Garibay-Orijel R, Higgins LM, Angeles-Arguiz R (2011b) Ectomycorrhizal fungi in Mexican *Alnus* forests support the host co-migration hypothesis and continental-scale patterns in phylogeography. Mycorrhiza, 21, 559–568.

Kennedy PG, Matheny P, Ryberg K et al. (2012) Scaling up: examining the macroecology of ectomycorrhizal fungi. Molecular Ecology, 21, 4151–4154.

Kerkhoff AJ, Moriarty PE, Weiser MD (2014) The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. Proceedings of the National Academy of Sciences USA, 111, 8125–8130.

Kirk PM, Cannon PF, Stalpers JA (2008) The Dictionary of the Fungi. CABI Publishing, Wallingford, UK.

Kohler A, Kuo A, Nagy LG et al. (2015) Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nature Genetics, 47, 410–415.

Kong A, Montoya A, Estrada-Torres A (2002) *Russula herrerae*, a new species with marginal veil from Mexico. Mycologia, 94, 290–296.

Kornerup A, Wanscher JH (1967) Methuen Handbook of Colour, 2nd edn., Methuen & Co Ltd, London.

Lanfear R, Calcott B, Ho SY, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution, 29, 1695–1701.

Latham RE, Ricklefs RE (1993) Continental comparisons of temperate-zone tree species diversity. In: Species Diversity in Ecological Communities: Historical and Geographical Perspectives (eds Ricklefs RE, Schluter D), pp. 294–314. University of Chicago Press, Chicago, Illinois.

Lear CH, Bailey TR, Pearson PN, Coxall HK, Rosenthal Y (2008) Cooling and ice growth across the Eocene-Oligocene transition. Geology, 36, 251–254.

Lebel T, Tonkin JE (2007) Australasian species of *Macowanites* are sequestrate species of *Russula* (Russulaceae, Basidiomycota). Australian Systematic Botany, 20, 355–381.

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22, 1658–1659.

Looney BP (2015) Molecular annotation of type specimens of *Russula* species described by W.A. Murrill from the southeast United States. Mycotaxon, 129, 255–268.

Lumbsch HT, Buchanan PK, May TW, Mueller GM (2008) Phylogeography and biogeography of fungi. Mycological Research, 112, 423–424.

Maddison WP, Maddison D (2001) Mesquite: a modular system for evolutionary analysis. Version 2.74. Available from http:// mesquiteproject.org.

Maddison D, Maddison WP (2005) MacClade 4: Analysis of Phylogeny and Character Evolution. Version 4.08a. Sinauer Associates Inc., Sunderland, Connecticut.

Maddison WP, Midford PE, Otto SP (2007) Estimating a binary character's effect on speciation and extinction. Systematic Biology, 56, 701–710.

Magnuson-Ford K, Otto SP (2012) Linking the investigations of character evolution and species diversification. The American Naturalist, 180, 225–245.

Matheny PB (2005) Improving phylogenetic inference of mushrooms with RPB1 and RPB2 nucleotide sequences (*Inocybe*; Agaricales). Molecular Phylogenetics and Evolution, 35, 1–20.

Matheny PB, Liu YJ, Ammirati JF, Hall BD (2002) Using RPB1 sequences to improve phylogenetic inference among mushrooms (*Inocybe*, Agaricales). American Journal of Botany, 89, 688–698.

Matheny PB, Aime MC, Bougher NL et al. (2009) Out of the Palaeotropics? Historical biogeography and diversification of the cosmopolitan ectomycorrhizal mushroom family Inocybaceae. Journal of Biogeography, 36, 577–592.

McGuire KL, Allison SD, Fierer N, Treseder KK (2013) Ectomycorrhizal-dominated boreal and tropical forests have distinct fungal communities, but analogous spatial patterns across soil horizons. PLoS ONE, 8, e68278.

McNabb R (1973) Russulaceae of New Zealand 2. *Russula* Pers. ex SF Gray. New Zealand Journal of Botany, 11, 673–730.

Miller SL, Buyck B (2002) Molecular phylogeny of the genus *Russula* in Europe with a comparison of modern Infrageneric classifications. Mycological Research, 106, 259–276.

Nilsson RH, Abarenkov K, Veldre V et al. (2010) An open source chimera checker for the fungal ITS region. Molecular Ecology Resources, 10, 1076–1081.

Nixon K (2006) Global and neotropical distribution and diversity of oak (genus *Quercus*) and oak forests. In: Ecology and Conservation of Neotropical Montane Oak Forests (ed. Kappelle M), pp. 3–13. Springer, Heidelberg, Germany.

Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. Systematic Biology, 53, 673–684.

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

Peay KG, Kennedy PG, Davies SJ, Tan S, Bruns TD (2010) Potential link between plant and fungal distributions in a dipterocarp rainforest: community and phylogenetic structure of tropical ectomycorrhizal fungi across a plant and soil ecotone. New Phytologist, 185, 529–542.

Perry D, Margolis H, Choquette C, Molina R, Trappe J (1989) Ectomycorrhizal mediation of competition between coniferous tree species. New Phytologist, 112, 501–511.

Põlme S, Bahram M, Yamanaka T et al. (2013) Biogeography of ectomycorrhizal fungi associated with alders (*Alnus* spp.) in relation to biotic and abiotic variables at the global scale. New Phytologist, 198, 1239–1249.

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution, 26, 1641–1650.

Pryor L (1959) Species distribution and association in *Eucalyptus*. In: Biogeography and Ecology in Australia (eds Keast A, Crocker RL, Christian CS), pp. 461–471. Springer Science, Dordrecht, The Netherlands.

Pyron RH, Burbrink FT (2009) Can the tropical conservatism hypothesis explain temperate species richness patterns? An inverse latitudinal biodiversity gradient in the New World snake tribe Lampropeltini. Global Ecology and Biogeography, 18, 406–415.

Rabosky DL, Goldberg EE (2015) Model inadequacy and mistaken inferences of trait-dependent speciation. Systematic Biology, 64, 340–355.

Rabosky DL, Grundler M, Anderson C et al. (2014) BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. Methods in Ecology and Evolution, 5, 701–707.

Richardson DM (2000) Ecology and Biogeography of *Pinus*. Cambridge University Press, Cambridge, UK. Rinaldi A, Comandini O, Kuyper TW (2008) Ectomycorrhizal fungal diversity: separating the wheat from the chaff. Fungal Diversity, 33, 1–45.

Rohde K (1992) Latitudinal gradients in species diversity: the search for the primary cause. Oikos, 65, 514–527.

Romagnesi H (1967) Les Russules D'Europe et D'Afrique du Nord. Bordas, Paris, France.

Ronquist F, Teslenko M, van der Mark P et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61, 539–542.

Rosling A, Landeweert R, Lindahl B et al. (2003) Vertical distribution of ectomycorrhizal fungal taxa in a podzol soil profile. New Phytologist, 159, 775–783.

Ryberg M, Matheny PB (2011) Dealing with incomplete taxon sampling and diversification of a large clade of mushroom-forming fungi. Evolution, 65, 1862–1878.

Ryberg M, Matheny PB (2012) Asynchronous origins of ectomycorrhizal clades of Agaricales. Proceedings of the Royal Society London Series B: Biological Sciences, 279, 2003–2011.

Ryberg M, Kristiansson E, Sjökvist E, Nilsson RH (2009) An outlook on the fungal internal transcribed spacer sequences in GenBank and the introduction of a web-based tool for the exploration of fungal diversity. New Phytologist, 181, 471–477.

Sánchez-Ramírez S, Etienne RS, Moncalvo JM (2015a) High speciation rate at temperate latitudes explains unusual diversity gradients in a clade of ectomycorrhizal fungi. Evolution, 69, 2196–2209.

Sánchez-Ramírez S, Tulloss RE, Amalfi M, Moncalvo JM (2015b) Palaeotropical origins, boreotropical distribution and increased rates of diversification in a clade of edible ectomycorrhizal mushrooms (*Amanita* section *Caesareae*). Journal of Biogeography, 42, 351–363.

Sandel B, Arge L, Dalsgaard B et al. (2011) The influence of Late Quaternary climate-change velocity on species endemism. Science, 334, 660–664.

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

Sarnari M (1998) Monographia Illustrata del Genere *Russula* in Europa. Associazioni Micologica Bresadola, Trento, Italy.

Sedano RE, Burns KJ (2010) Are the Northern Andes a species pump for Neotropical birds? Phylogenetics and biogeography of a clade of Neotropical tanagers (Aves: Thraupini). Journal of Biogeography, 37, 325–343.

Silvestro D, Michalak I (2012) raxmlGUI: a graphical front-end for RAxML. Organisms Diversity & Evolution, 12, 335–337.

Singer R (1986) The Agaricales in Modern Taxonomy. Koeltz Scientific Books, Koenigstein, Germany.

Skrede I, Engh IB, Binder M et al. (2011) Evolutionary history of Serpulaceae (Basidiomycota): molecular phylogeny, historical biogeography and evidence for a single transition of nutritional mode. BMC Evolutionary Biology, 11, 1–13.

Smith SA, Beaulieu JM, Donoghue MJ (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. BMC Evolutionary Biology, 9, 1–12.

Smith ME, Henkel TW, Aime MC et al. (2011) Ectomycorrhizal fungal diversity and community structure on three co-occurring leguminous canopy tree species in a Neotropical rainforest. New Phytologist, 192, 699–712.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. Systematic Biology, 57, 758–771.

Tang CQ, Ohsawa M (1997) Zonal transition of evergreen, deciduous, and coniferous forests along the altitudinal gradient on a humid subtropical mountain, Mt. Emei, Sichuan, China. Plant Ecology, 133, 63–78.

Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006) Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. Philosophical Transactions of the Royal Society B: Biological Sciences, 361, 1947–1963.

Tedersoo L, Nara K (2010) General latitudinal gradient of biodiversity is reversed in ectomycorrhizal fungi. New Phytologist, 185, 351–354.

Tedersoo L, Põlme S (2012) Infrageneric variation in partner specificity: multiple ectomycorrhizal symbionts associate with *Gnetum gnemon* (Gnetophyta) in Papua New Guinea. Mycorrhiza, 22, 663–668.

Tedersoo L, Smith ME (2013) Lineages of ectomycorrhizal fungi revisited: foraging strategies and novel lineages revealed by sequences from belowground. Fungal Biology Reviews, 27, 83–99.

Tedersoo L, Bahram M, Toots M et al. (2012) Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. Molecular Ecology, 21, 4160–4170.

Tedersoo L, Bahram M, PõlmeS et al. (2014a) Global diversity and geography of soil fungi. Science, 346, 1–10.

Tedersoo L, Bahram M, Ryberg M et al. (2014b) Global biogeography of the ectomycorrhizal/sebacina lineage (Fungi, Sebacinales) as revealed from comparative phylogenetic analyses. Molecular Ecology, 23, 4168–4183.

Thiers B (continuously updated) Index herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden's virtual herbarium. http://sweetgum.nybg.org/ih.

Van de Putte K, Nuytinck J, Das K, Verbeken A (2012) Exposing hidden diversity by concordant genealogies and morphology—a study of the *Lactifluus volemus* (Russulales) species complex in Sikkim Himalaya (India). Fungal Diversity, 55, 171–194.

Verbeken A, Buyck B (2002) Diversity and ecology of tropical ectomycorrhizal fungi in Africa. In: Tropical Mycology. Vol. 1 Macromycetes (eds Watling R, Frankland JC, Ainsworth M et al.), pp. 11–24. CABI Publishing, Wallingford, UK.

Verbeken A, Nuytinck J, Buyck B (2011) New combinations in *Lactifluus*. 1. *L.* subgenera *Edules*, *Lactariopsis*, and *Russulopsis*. Mycotaxon, 118, 447–453.

Vilgalys R, Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several Cryptococcus species. Journal of Bacteriology, 172, 4238–4246.

Von Humboldt A (1807) Essai sur la géographie des plantes: accompagné d'un tableau physique des régions équinoxiales, fondé sur des mesures exécutées, depuis le dixième degré de latitude boréale jusqu'au dixième degré de latitude australe, pendant les années 1799, 1800, 1801, 1802 et 1803.

Wang L, Schneider H, Zhang X-C, Xiang Q-P (2012) The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. BMC Plant Biology, 12, 1–9.

Webb CO, Ackerly DD, Kembel SW (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics, 24, 2098–2100.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: A Guide to Methods and Applications (eds Innis MA, Gelfand DH, Sninsky JJ, White TJ), pp. 315–322. Academic Press, New York City, New York.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. Trends in Ecology & Evolution, 19, 639–644.

Wilson AW, Binder M, Hibbett DS (2012) Diversity and evolution of ectomycorrhizal host associations in the Sclerodermatineae (Boletales, Basidiomycota). New Phytologist, 194, 1079–1095.

Appendix

Table 4. Taxon sampling and Genbank accession numbers for constraint phylogeny.

Taxon	Herbarium	Collection voucher	ITS	LSU	*rpb1*	*rpb2*
R. aeruginea[a]	UPS	AT2003017	DQ421999	DQ421999	DQ421946	–
R. aff. azurea	TENN067625	BPL274	**KT933973**	**KT933834**	**KT957346**	**KT933905**
R. aff. chloroides	GENT FH12-273	FH12273	**KT934015**	**KT933876**	**KT957386**	**KT933947**
R. aff. delica	GENT FH12-266	FH12266	**KT934014**	**KT933875**	**KT957385**	**KT933946**
R. albonigra[a]	UPS	AT2002064	DQ422029	DQ422029	DQ421966	–
R. amara	GENT FH12-213	FH12213	**KT933998**	**KT933859**	**KT957370**	**KT933930**
R. amoenolens	TENN067119	BPL232	**KT933954**	**KT933813**	–	**KT933884**
R. aquosa[d]	TENN067620	BPL271	KF810138	**KT933831**	**KT957343**	**KT933902**
R. betularum	TENN067623	BPL269	**KT933969**	**KT933829**	**KT957341**	**KT933900**
R. cadaverolens	TENN067226	BPL239	**KT933957**	**KT933816**	**KT957327**	**KT933887**
R. camarophylla[a]	PC	PAM01081108	DQ421982	DQ421982	DQ421938	–
R. cf. appalachiensis	TENN067318	BPL250	–	–	**KT957333**	**KT933893**
R. cf. atroglauca	GENT FH12-248	FH12248	**KT934009**	**KT933870**	**KT957380**	**KT933941**
R. cf. compacta[a]	PC	AV04130	DQ422001	DQ422001	DQ421948	–
R. cf. cyanoxantha	TENN067627	BPL280	**KT933976**	**KT933837**	**KT957349**	**KT933908**
R. cf. decipiens	TENN067417	BPL266	**KT933967**	**KT933827**	**KT957339**	**KT933899**
R. cf. delica[a]	UPS	UE24.08.2004-20	DQ422005	DQ422005	DQ421950	–
R. cf. firmula	TENN067124	BPL229	–	**KT933811**	**KT957323**	**KT933882**
R. cf. foetens[a]	UPS	UE18.07.2003-7	DQ422023	DQ422023	DQ421962	–
R. cf. fragilis	TENN067621	BPL273	**KT933972**	**KT933833**	**KT957345**	**KT933904**
R. cf. integra	TENN070022	BPL288	**KT933982**	**KT933843**	**KT957354**	**KT933914**
R. cf. ochrophylla	TENN067108	BPL231	**KT933953**	**KT933812**	**KT957324**	**KT933883**
R. cf. pseudolepida	TENN067297	BPL247	**KT933962**	**KT933821**	**KT957332**	**KT933892**
R. cf. rugulosa	TENN067224	BPL237	**KT933955**	**KT933814**	**KT957325**	**KT933885**
R. cf. rugulosa	TENN067225	BPL238	**KT933956**	**KT933815**	**KT957326**	**KT933886**
R. cf. silvestris	GENT FH12-225	FH12225	**KT934004**	**KT933865**	**KT957375**	**KT933936**
R. cf. smithii	TENN070302	LUK12327	**KT934017**	**KT933878**	**KT957388**	**KT933949**
R. cf. versicolor	GENT FH12-259	FH12259	**KT934012**	**KT933873**	**KT957383**	**KT933944**
R. claroflava	GENT FH12-212	FH12212	**KT933997**	**KT933858**	**KT957369**	**KT933929**
R. columbicolor[b]	GENT	H-2010BT108A	–	JN389003	JN375606	JN389200
R. compacta	TENN067133	BPL227	**KT933952**	**KT933810**	–	**KT933881**
R. compacta	TENN067303	BPL242	**KT933960**	**KT933819**	**KT957330**	**KT933890**
R. compacta[a]		Duke s.n.	AF287888	AF287888	AY218514	–
R. cremeirosea	TENN069929	BPL289	**KT933983**	**KT933844**	**KT957355**	**KT933915**
R. crustosa	TENN070180	BPL251	**KT933963**	**KT933822**	**KT957334**	**KT933894**
R. crustosa	TENN067418	BPL265	**KT933966**	**KT933826**	**KT957338**	**KT933898**
R. cuprea	GENT FH12-250	FH12250	**KT934010**	**KT933871**	**KT957381**	**KT933942**
R. curtipes	GENT FH12-206	FH12206	**KT933995**	**KT933856**	**KT957367**	**KT933927**
R. cyanoxantha[a]	UPS	UE29.09.2002-2	DQ422033	DQ422033	DQ421970	–
R. decolorans	GENT FH12-196	FH12196	**KT933992**	**KT933853**	**KT957364**	**KT933924**
R. dissimulans	TENN070021	BPL285	**KT933979**	**KT933840**	–	**KT933911**
R. earlei	TENN067260	BPL245	**KT933961**	**KT933820**	**KT957331**	**KT933891**
R. earlei[a]	PC	WCRW00-412	DQ422025	DQ422025	DQ421963	–
R. emetica[a]	UPS	UE05.10.2003-11	DQ421997	DQ421997	DQ421943	–
R. emeticolor	GENT FH12-253	FH12253	**KT934011**	**KT933872**	**KT957382**	**KT933943**
R. farinipes[a]	UPS	UE28.09.2002-4	DQ421983	DQ421983	DQ421939	–
R. fellea	GENT FH12-185	FH12185	**KT933989**	**KT933850**	**KT957361**	**KT933921**
R. firmula[a]	UPS	AT2004142	DQ422017	DQ422017	DQ421958	–
R. foetens	GENT FH12-277	FH12277	**KT934016**	**KT933877**	**KT957387**	**KT933948**
R. fontqueri	GENT FH12-223	FH12223	**KT934003**	**KT933864**	**KT957374**	**KT933935**
R. fragilis	GENT FH12-197	FH12197	**KT933993**	**KT933854**	**KT957365**	**KT933925**
R. gracillima[a]	UPS	UE23.08.16-01	DQ422004	DQ422005	DQ421949	–
R. granulata	TENN067622	BPL272	**KT933971**	**KT933832**	**KT957344**	**KT933903**
R. grisea	GENT FH12-234	FH12234	**KT934006**	**KT933867**	**KT957377**	**KT933938**
R. grisea[a]	UPS	UE2005.08.16-01	DQ422030	DQ422030	DQ421968	–
R. heterophylla[a]	UPS	UE20.08.2004-2	DQ422006	DQ422006	DQ421951	–

Table 4 Continued

Taxon	Herbarium	Collection voucher	ITS	LSU	*rpb1*	*rpb2*
R. illota[a]	UPS	UE26.07.2002-3	DQ422024	DQ422024	DQ421967	–
R. integra	GENT FH12-172	FH12172	**KT933984**	**KT933845**	**KT957356**	**KT933916**
R. khanchanjungae[b]	GENT	AV-KDKVP09-106	–	JN389004	JN375607	JN389201
R. krombholzii	GENT FH12-186	FH12186	**KT933990**	**KT933851**	**KT957362**	**KT933922**
R. laurocerasi	GENT FH12-178	FH12178	**KT933988**	**KT933849**	**KT957360**	**KT933920**
R. lepida[a]	UPS	HJB9990	DQ422013	DQ422013	DQ421954	–
R. luteotacta	GENT FH12-187	FH12187	**KT933991**	**KT933852**	**KT957363**	**KT933923**
R. maculata[a]	UPS	HJB10019	DQ422015	DQ422015	DQ421956	–
R. mairei	GENT FH12-262	FH12262	**KT934013**	**KT933874**	**KT957384**	**KT933945**
R. mustelina	GENT FH12-226	FH12226	**KT934005**	**KT933866**	**KT957376**	**KT933937**
R. nauseosa	GENT FH12-173	FH12173	**KT933985**	**KT933846**	**KT957357**	**KT933917**
R. nigricans[a]	UPS	UE20.09.2004-07	DQ422010	DQ422010	DQ421952	–
R. nitida	GENT FH12-218	FH12218	**KT934001**	**KT933862**	–	**KT933933**
R. ochroleuca	GENT FH12-211	FH12211	**KT933996**	**KT933857**	**KT957368**	**KT933928**
R. ochrospora[a]	UPS	GD20.07.2004	DQ422012	DQ422012	DQ421953	–
R. pallescens[a]	TUR	PL146/2002	DQ421987	DQ421987	DQ421941	–
R. paludosa	GENT FH12-216	FH12216	**KT934000**	**KT933861**	**KT957372**	**KT933932**
R. parazurea[a]	UPS	MF01.10.2003	DQ422007	DQ422007	DQ421945	–
R. peckii	TENN067447	BPL270	**KT933970**	**KT933830**	**KT957342**	**KT933901**
R. pectinatoides	TENN067626	BPL276	**KT933975**	**KT933836**	**KT957348**	**KT933907**
R. pectinatoides[a]	UPS	AT2001049	DQ422026	DQ422026	DQ421964	–
R. persicina[a]	UPS	UE21.09.2003-01	DQ422019	DQ422019	DQ421960	–
R. pulchra	TENN067117	BPL226	**KT933951**	**KT933809**	**KT957322**	**KT933880**
R. pusilla	TENN067416	BPL267	**KT933968**	**KT933828**	**KT957340**	–
R. queletii	GENT FH12-237	FH12237	**KT934007**	**KT933868**	**KT957378**	**KT933939**
R. raoultii	GENT FH12-222	FH12222	**KT934002**	**KT933863**	**KT957373**	**KT933934**
R. redolens	TENN069923	BPL141	**KT933950**	**KT933808**	**KT957321**	**KT933879**
R. redolens	TENN067593	BPL260	**KT933965**	**KT933825**	**KT957337**	**KT933897**
R. risigallina[a]	UPS	UE03.07.2003-08	DQ422022	DQ422022	DQ421961	–
R. romellii	GENT FH12-177	FH12177	**KT933987**	**KT933848**	**KT957359**	**KT933919**
R. rubellipes	TENN067227	BPL240	**KT933958**	**KT933817**	**KT957328**	**KT933888**
R. sanguinea	GENT FH12-240	FH12240	**KT934008**	**KT933869**	**KT957379**	**KT933940**
R. sardonia	GENT FH12-215	FH12215	**KT933999**	**KT933860**	**KT957371**	**KT933931**
R. sp. 1	TENN067379	BPL255	**KT933964**	**KT933823**	**KT957335**	**KT933895**
R. sp. 2	TENN069926	BPL283	**KT933977**	**KT933838**	**KT957350**	**KT933909**
R. sp. 3	TENN069927	BPL286	**KT933980**	**KT933841**	**KT957352**	**KT933912**
R. sp. 4	TENN069928	BPL287	**KT933981**	**KT933842**	**KT957353**	**KT933913**
R. sp.[a]	PC	BB99.250	DQ422028	DQ422028	DQ421965	–
R. subtilis	TENN067624	BPL275	**KT933974**	**KT933835**	**KT957347**	**KT933906**
R. tsokae[b]	GENT	KD-KVP1283	–	JN389006	JN375608	JN389203
R. variata	TENN067302	BPL241	**KT933959**	**KT933818**	**KT957329**	**KT933889**
R. velutipes	GENT FH12-203	FH12203	**KT933994**	**KT933855**	**KT957366**	**KT933926**
R. vesca	TENN070020	BPL284	**KT933978**	**KT933839**	**KT957351**	**KT933910**
R. vesca[a]	UPS	AT2002091	DQ422018	DQ422018	DQ421959	–
R. vinacea[d]	TENN067365	BPL257	KF810139	**KT933824**	**KT957336**	**KT933896**
R. virescens[a]	UPS	HJB9989	DQ422014	DQ422014	DQ421955	–
R. zvarae	GENT FH12-175	FH12175	**KT933986**	**KT933847**	**KT957358**	**KT933918**
Multifurca zonaria[a]	PC	DED7442	DQ421990	DQ421990	DQ421942	–
M. ochricompacta*[a]	PC	BB02107	DQ421984	DQ421984	DQ421940	–
Lactarius lignyotus[c]	CUW	PBM2424	DQ221107	AY631898	DQ408128	–
Lactifluus deceptivus[c†]	CUW	PBM2462	AY854089	AY631899	AY803749	AY864884

[a]Sequences generated by Buyck et al. (2008).
[b]Sequences generated by Van de Putte et al. (2012).
[c]Sequences generated by Matheny et al. (2007).
[d]Sequences generated by Looney (2015).

* Annotated as *Russula ochricompacta* in Buyck et al. (2008).
† Annotated as *Lactarius deceptivus* in Matheny et al. (2007).
° Sequences generated for this study in bold.

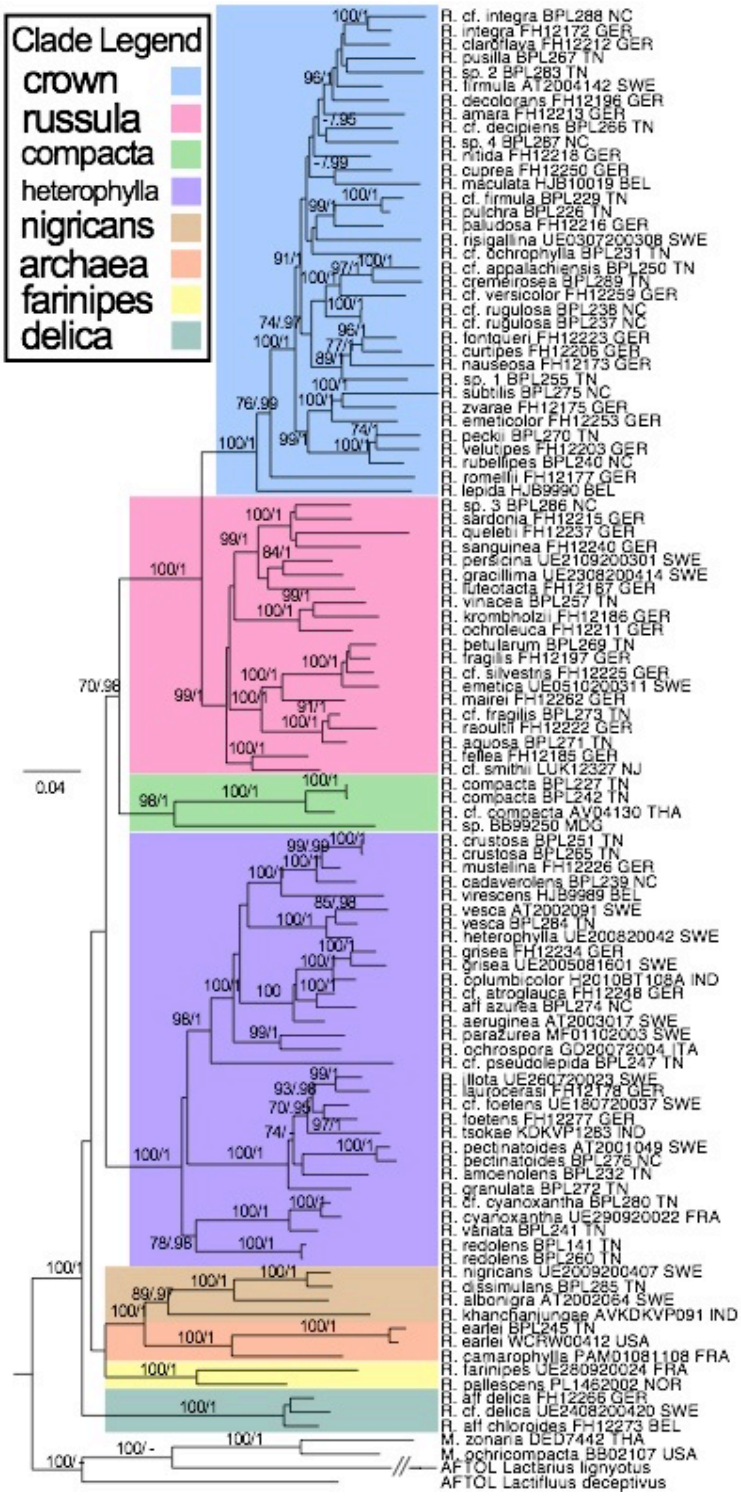


Figure 5. Phylogenetic relationships of *Russula* inferred from nuclear ribosomal and single-copy (ITS, nrLSU, *rpb1* and *rpb2*) sequences derived from a maximum-likelihood analysis.

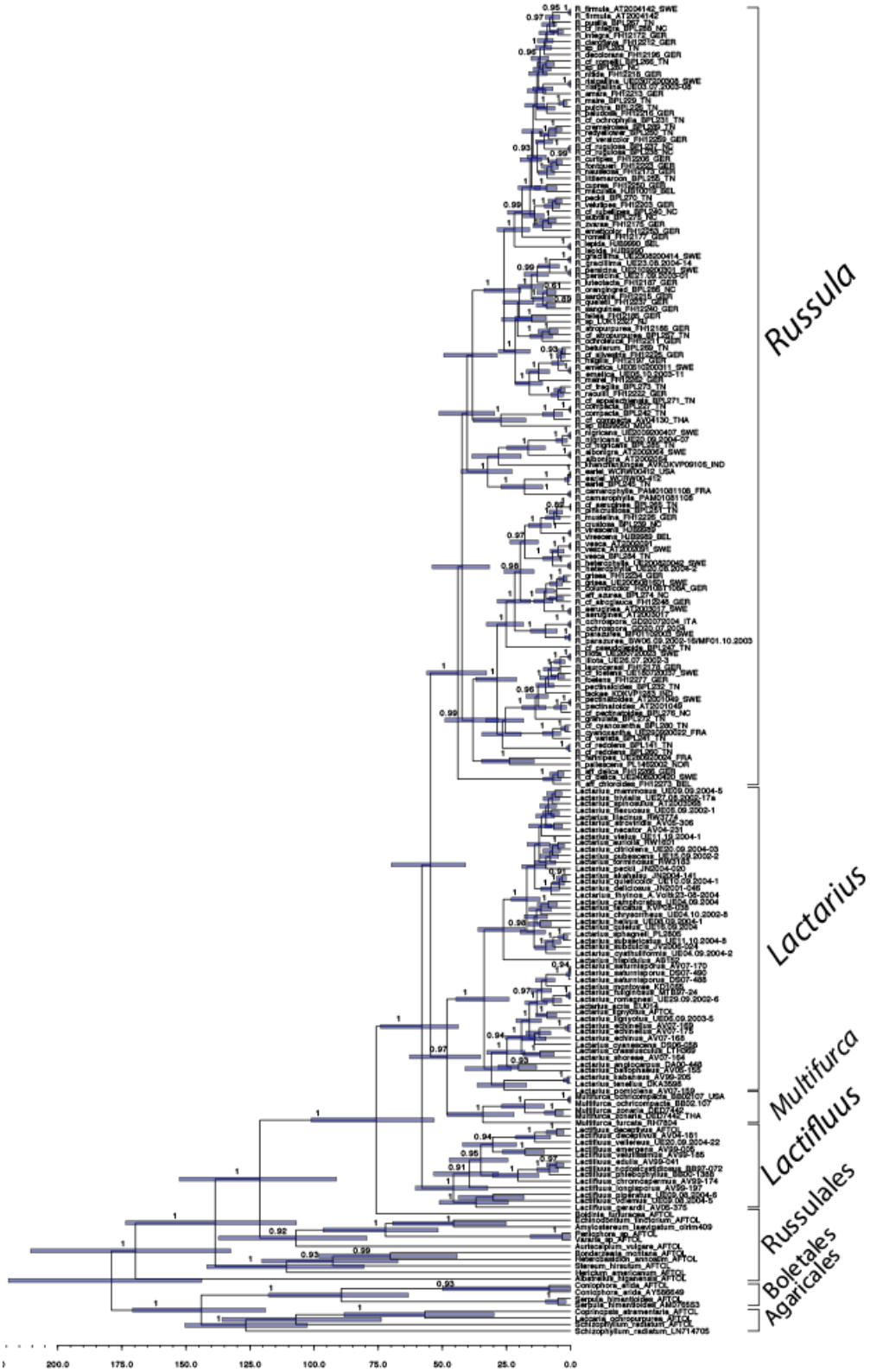


Figure 6. Chronogram of Russulaceae inferred in BEAST 2. The 95% HPD posterior probabilities >0.9 are reported.

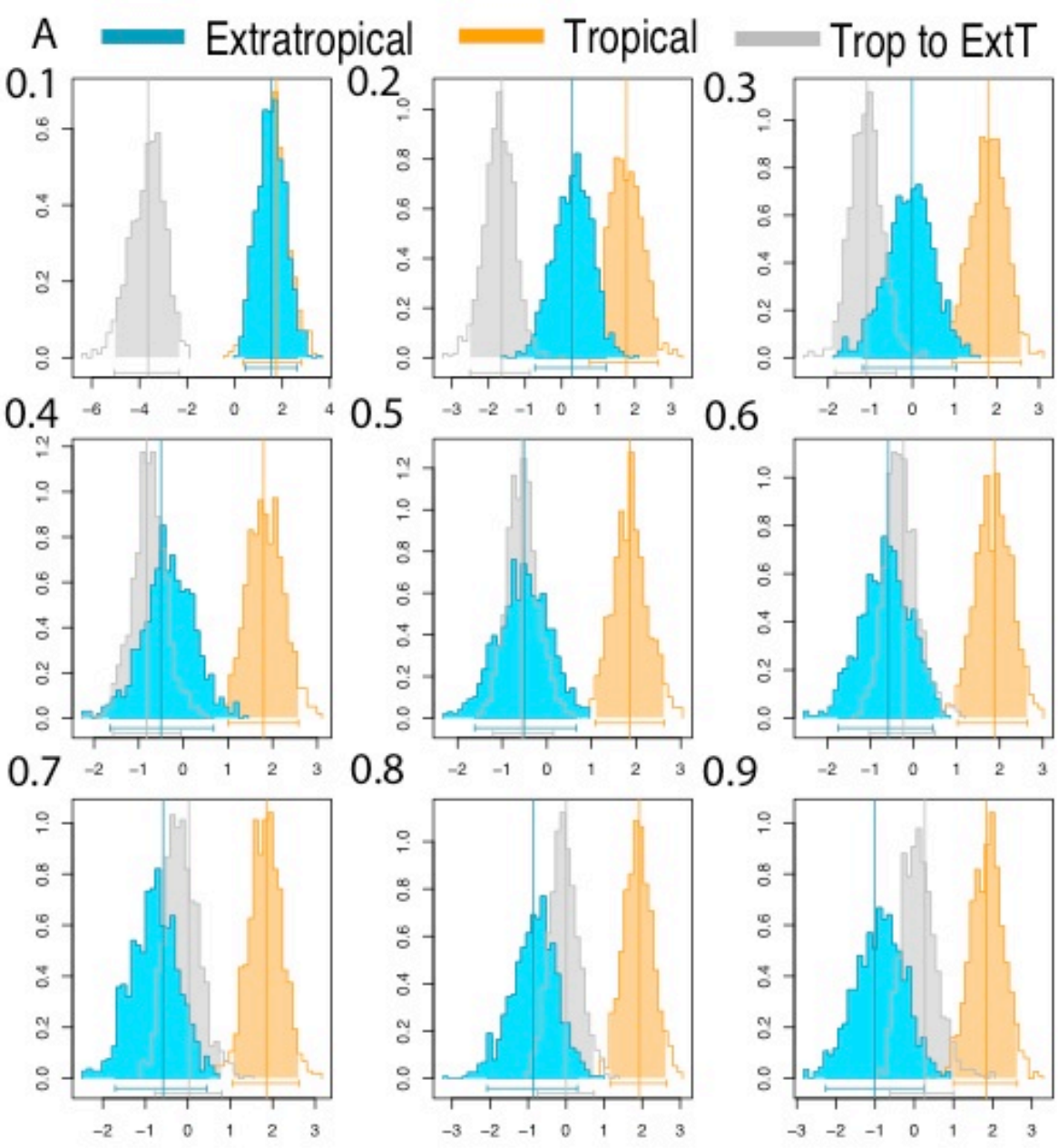


Figure 7. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for GeoSSE tropical vs. extratropical distribution.

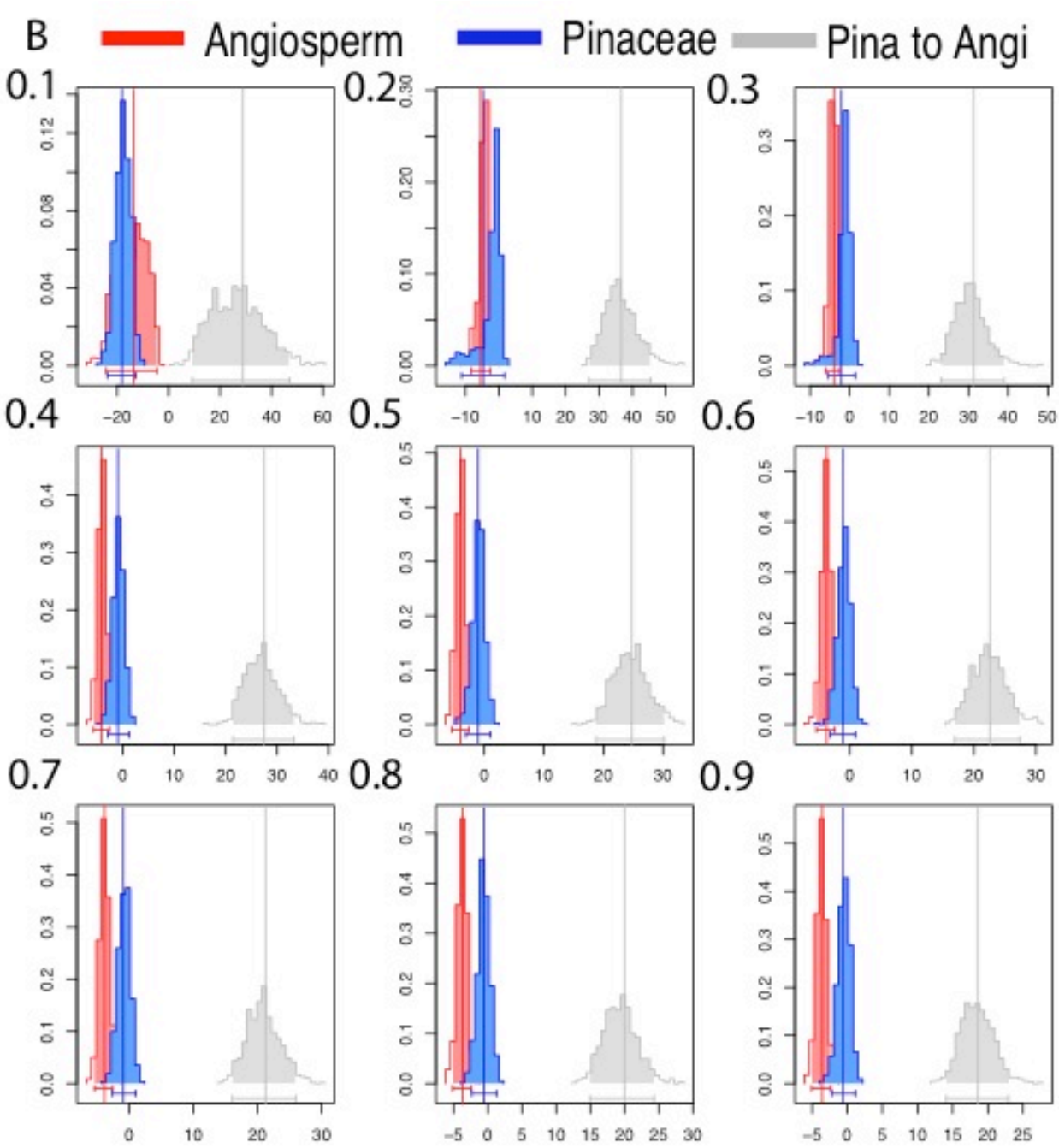


Figure 8. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for GeoSSE angiosperm vs. Pinaceae association.

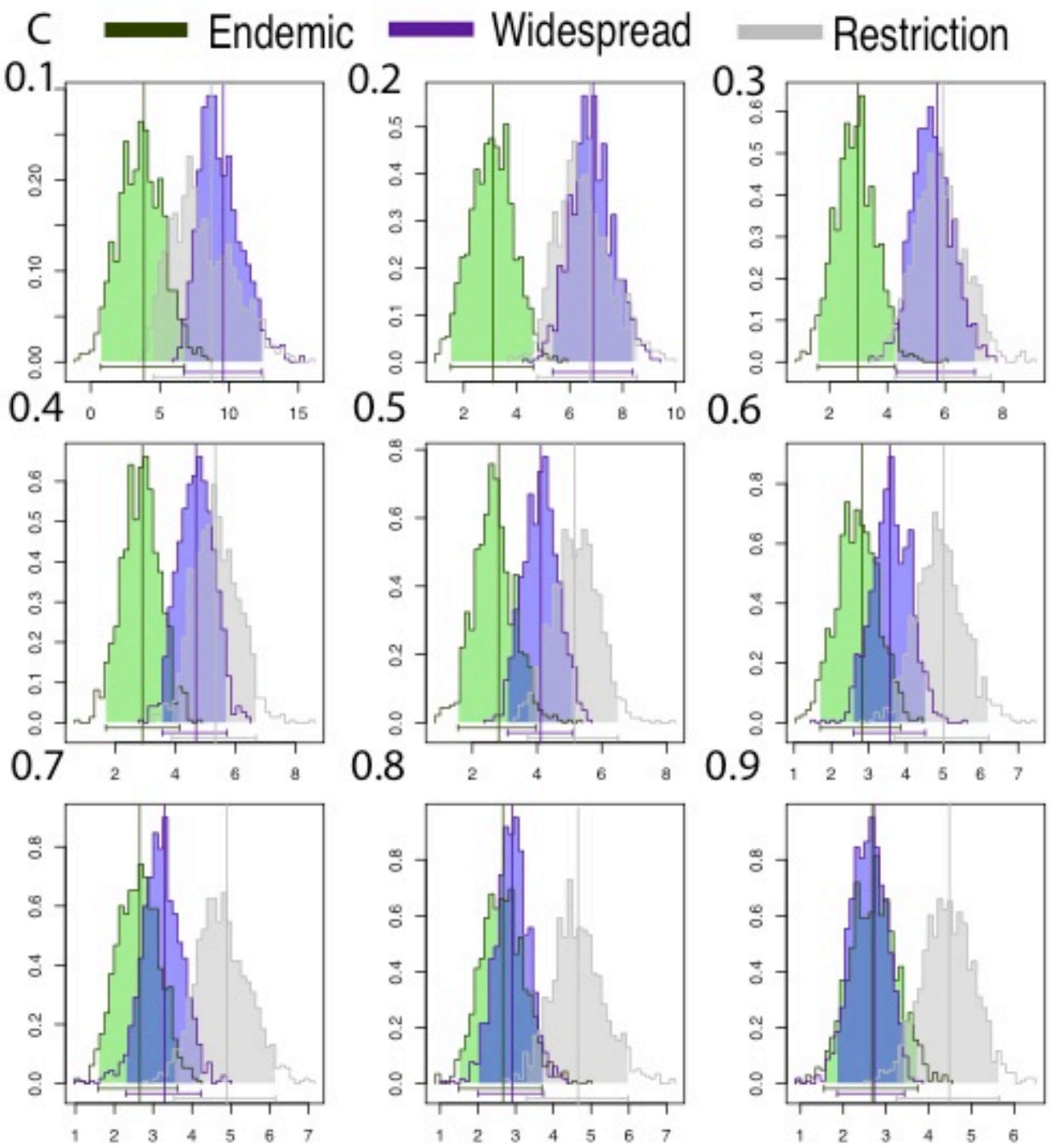


Figure 9. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for BiSSE binary tropical/temperate endemicity vs. widespread.

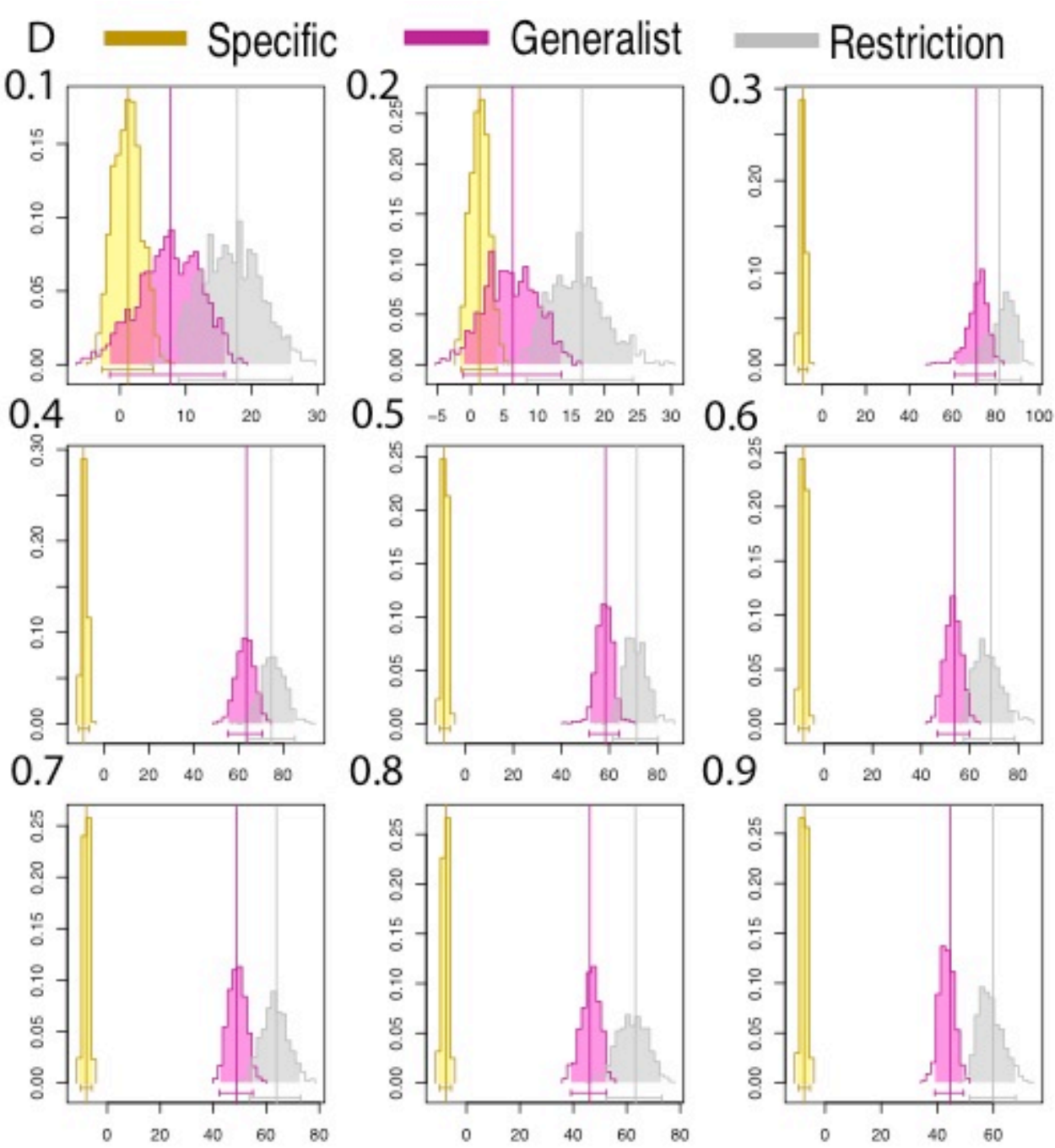


Figure 10. Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations for BiSSE binary angiosperm/Pinaceae specificity vs. generalist association.

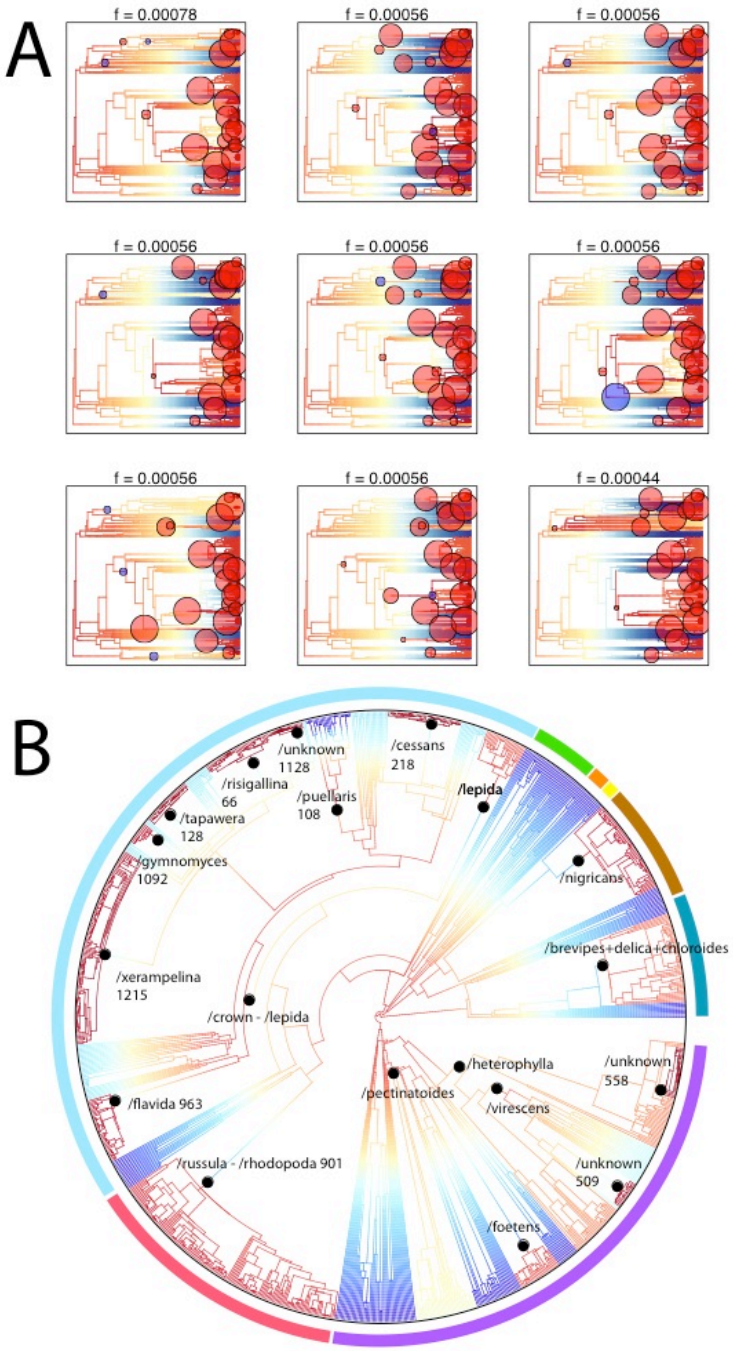


Figure 11. BAMM analysis showing (A) the top nine shift configurations of the most credible shift configuration set. Red circles represent increases in diversification rate, while blue circles represent slowdowns in diversification. The size of the circle indicates how significant the shift is. The f values indicate what proportion of the confidence can be assigned to that particular scenario; and (B) a circle phylogeny of the best shift configuration.

Table 5. Final MOTU clusters with associated metadata for mycoheterotrophic parasitism, gasteroid morphology, plant association and geographic distribution.

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
1	GU234024	0	0	Various	Ang	1	NA/EU	0
2	GU997922	0	0	Betula nana	Bet	1	NA	0
3	FJ845428	0	0	Mix	All	2	NA/EU	0
4	GU997841	0	0	Angiosperm	Ang	1	NA/AS	0
5	GU998048	0	0	Betula nana	Bet	1	NA	0
6	GU997885	0	0	Betula nana	Bet	1	NA	0
7	GU997928	0	0	Betula nana	Bet	1	NA	0
8	EU711838	0	0			3	NA	0
9	GU997971	0	0	Betula nana	Bet	1	NA	0
10	GU997943	0	0	Betula nana	Bet	1	NA	0
11	GU998198	0	0	Betula nana	Bet	1	NA	0
12	EU711809	0	0			3	NA	0
13	GU966632	0	1			3	AS	0
14	EU711822	0	0			3	NA	0
15	EU711808	0	0			3	NA	0
16	EU711761	0	1			3	NA	0
17	EU711793	0	0			3	NA	0
18	EF218813	0	0	Pseudotsuga menziezii	Pin	0	NA	0
19	EU711804	0	0			3	NA	0
20	EU711817	0	0			3	NA	0
21	EU711824	0	0			3	NA	0
22	JF834365	0	0			3	NA	0
23	AY822746	0	1	Pinus contorta	Pin	0	NA	0
24	AY239310	0	1	Various	Pin	0	NA	0
25	HQ204709	0	0	Quercus ilex	Fag	1	EU	0
26	HQ204712	0	0	Quercus ilex	Fag	1	EU	0
27	EU403089	0	0	Quercus rotundifolia	Fag	1	EU	0
28	DQ061902	1	0	Mix	All	2	EU	0
29	EU569270	0	0	Quercus	Fag	1	NA	2
30	AY061668	0	0	Pinaceae	Pin	0	NA/EU	0
31	EU569269	0	0	Quercus	Fag	1	NA	1
32	EU711768	0	0			3	NA	0
33	EU711866	0	0			3	NA	0
34	GU998043	0	1	Betula nana	Bet	1	NA	0
35	EU711908	0	0			3	NA	0
36	GU998008	0	0	Angiosperm	Ang	1	NA/EU	0
37	GU234047	0	0	Various	Ang	1	NA/EU	0
38	EU711888	0	0	Various	Ang	1	NA/EU	0
39	GU998132	0	0	Betula nana	Bet	1	NA	0
40	GU998177	0	0	Betula nana	Bet	1	NA	0
41	GU998010	0	0	Betula nana	Bet	1	NA	0
42	GU998398	0	0	Betula nana	Bet	1	NA	0
43	AB597704	1	0			3	AS	0
44	FJ789625	0	0	Pinaceae	Pin	0	NA	0
45	FJ789624	0	0	Pinaceae	Pin	0	NA	0
46	HQ604850	0	0	Pinaceae	Pin	0	NA	0
47	HQ022268	0	0			3	NA	0
48	AY061716	0	0	Pinaceae	Pin	0	NA/EU	0
49	FJ660474	0	0	Pinus albicaulis	Pin	0	NA	0
50	AM231797	1	0	Mix	All	2	NA	0
51	AY061720	1	0	Pinaceae	Pin	0	NA/EU	0
52	JF834354	0	0			3	NA	0
53	EF372407	1	0	Pinus sylvestris	Pin	0	EU	0
54	JF834351	0	0			3	NA	0
55	AB597714	1	0			3	AS	0
56	EU597057	0	0	Various	Pin	0	NA	0
57	DQ273397	0	0	Notholithocarpus densiflorus	Fag	1	NA	0
58	AB218086	0	0	Fagus crenata	Fag	1	AS	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
59	FJ946975	0	0	Quercus	Fag	1	EU	0
60	FN669242	0	0	Populus temula	Sal	1	EU	0
61	AY061713	0	0			3	EU	0
62	JF908685	0	0			3	EU	0
63	JF908702	0	0	Fagus sylvatica	Fag	1	EU	0
64	FJ196945	0	0	Quercus	Fag	1	NA	1
65	HM105560	0	0	Quercus liaotungensis	Fag	1	AS	0
66	HQ604848	0	0	Betula papyrifera	Bet	1	NA	0
67	AF230898	1	0	Mix	All	2	EU	0
68	FJ196944	0	0	Quercus	Fag	1	NA	1
69	EU569271	0	0	Quercus	Fag	1	NA	1
70	DQ493564	0	0			3	NA	0
71	EU711770	0	0			3	NA	0
72	EU711769	0	0			3	NA	0
73	DQ273396	0	0	Notholithocarpus densiflorus	Fag	1	NA	0
74	FJ897211	0	0	Quercus rotundifolia	Fag	1	EU	0
75	JF908663	0	0			3	EU	0
76	GQ166871	0	0	Quercus	Fag	1	NA	0
77	DQ493586	0	0			3	NA	0
78	FJ454968	1	0			3	AS	1
79	JF908662	0	0			3	EU	0
80	GU327497	1	0			3	EU	0
81	EU569263	0	0	Quercus	Fag	1	NA	1
82	EU598178	0	0			3	NA	0
83	HM146862	0	0	Pinus sylvestris	Pin	0	EU	0
84	FN565339	1	0	Pinus sylvestris	Pin	0	EU	0
85	AF335442	0	0	Mix	All	2	NA	0
86	EU375714	0	0	Quercus	Fag	1	NA	0
87	GU256186	0	0	Quercus	Fag	1	EU	0
88	AF349708	1	0			3	AS	0
89	GU907803	0	0	Quercus rubra	Fag	1	NA	0
90	AF418634	0	0	Mix	All	2	NA/EU /ME	0
91	AY061699	0	0			3	EU	0
92	EU284013	0	0			3	NA	0
93	HQ204707	0	0	Quercus ilex	Fag	1	EU	0
94	AF418635	0	0			3	EU	0
95	FM999647	0	0	Fagus grandifolia	Fag	1	NA	0
96	FM999715	0	0	Various	Ang	1	NA	0
97	FM999630	0	0	Fagus grandifolia	Fag	1	NA	0
98	GQ359818	1	0	Various	Ang	1	AS	0
99	AM113956	0	0	Fagus sylvatica	Fag	1	EU	0
100	GQ221640	0	0	Notholithocarpus densiflorus	Fag	1	NA	0
101	EU819511	1	0	Castanea dentata	Fag	1	NA	0
102	EU569268	0	0	Quercus	Fag	1	NA	1
103	HQ703024	0	0	Mix	All	2	EU	0
104	AY061709	0	0	Mix	All	2	EU	0
105	EU819424	0	0	Quercus rubra	Fag	1	NA	0
106	AJ633571	0	0	Pinus taeda	Pin	0	NA	0
107	GQ268640	0	0	Dipterocarpaceae	Dip	1	AS	1
108	GQ268644	0	0	Dipterocarpaceae	Dip	1	AS	1
109	JF960819	0	0	Eucalyptus delegatensis	Myr	1	AU	0
110	DQ328136	0	0			3	AU	0
111	AY702073	1	0			3	AU	0
112	JF960820	0	0	Eucalyptus delegatensis	Myr	1	AU	0
113	FJ656018	0	0	Angiosperm	Ang	1	AS	1
114	DQ388847	0	0	Eucalyptis	Myr	1	AU	0
115	JF960815	0	0	Eucalyptus delegatensis	Myr	1	AU	0
116	JF960825	0	0	Eucalyptus delegatensis	Myr	1	AU	0
117	EU019917	0	0			3	AU	0
118	EU019914	0	0			3	AU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
119	EU019927	0	0			3	AU	0
120	EU019940	0	0			3	AU	0
121	EU019948	0	1			3	AU	0
122	EU019942	0	0	Various	Ang	1	AU	0
123	JF960816	0	0	Eucalyptus delegatensis	Myr	1	AU	0
124	JF960821	0	1	Eucalyptus delegatensis	Myr	1	AU	0
125	EU019947	0	1			3	AU	0
126	EU019946	0	1			3	AU	0
127	JF960809	0	0	Eucalyptus delegatensis	Myr	1	AU	0
128	GU222285	0	0			3	AU	0
129	GU222324	0	0			3	AU	0
130	JF960817	0	0	Eucalyptus delegatensis	Myr	1	AU	0
131	GU222261	0	0			3	AU	0
132	AY702070	1	0			3	AU	0
133	GU222292	0	0			3	AU	0
134	GU222263	0	0			3	AU	0
135	AB594967	1	0			3	AS	0
136	GQ268641	0	0	Dipterocarpaceae	Dip	1	AS	1
137	EU711835	0	0			3	NA	0
138	AY061687	0	0			3	EU	0
139	FJ454956	1	0			3	AS	1
140	JF960822	0	0	Eucalyptus delegatensis	Myr	1	AU	0
141	JN182871	0	0			3	AS	0
142	AB218199	0	0	Betula maximowicziana	Bet	1	AS	0
143	AB458684	0	0	Intsia bijuga	Fab	1	AF	1
144	GU981744	0	0			3	NA	0
145	FR731481	0	0	Brachystegia spiciformis	Fab	1	AF	1
146	FR731620	0	0	Julbernardia paniculata	Fab	1	AF	1
147	AB629022	1	0			3	AS	0
148	AY061737	0	0			3	AF	1
149	FR731760	0	0	Caesalpinioideae	Fab	1	AF	1
150	FR731768	0	0	Caesalpinioideae	Fab	1	AF	1
151	FJ455032	1	0			3	AS	1
152	FR731655	0	0	Caesalpinioideae	Fab	1	AF	1
153	FR731769	0	0	Caesalpinioideae	Fab	1	AF	1
154	AB451976	0	0	Dipterocarpus	Dip	1	AS	1
155	GQ268653	0	0	Dipterocarpaceae	Dip	1	AS	1
156	AB597651	1	0			3	AS	0
157	FR731893	0	0	Uapaca heudelotti	Phy	1	AF	1
158	GU391444	0	0	Pinus densiflora	Pin	0	AS	0
159	HM044516	0	0	Larix decidua	Pin	0	EU	0
160	JN887982	0	0	Pinus montezumae	Pin	0	NA	1
161	FJ946947	0	0	Quercus	Fag	1	EU	0
162	FJ454982	1	0			3	AS	1
163	GU256185	0	0	Quercus suber	Fag	1	EU	0
164	GU391432	0	0	Pinus densiflora	Pin	0	AS	0
165	JN168751	0	0	Dicymbe corymbosa	Fab	1	SA	1
166	FR731506	0	0	Brachystegia spiciformis	Fab	1	AF	1
167	GQ268633	0	0	Dipterocarpaceae	Dip	1	AS	1
168	FR731380	0	0	Asteropeia micraster	Fab	1	AF	1
169	HM069496	0	0	Pinus	Pin	0	EU	0
170	AB253519	0	0	Pinus thunbergii	Pin	0	AS	0
171	AB354282	0	0	Pinus thunbergii	Pin	0	AS	0
172	AB587768	0	0	Pinus thunbergii	Pin	0	AS	0
173	GU371293	0	0	Pinus	Pin	0	AS	2
174	HM044536	0	0	Larix decidua	Pin	0	EU	0
175	HM044475	0	0	Larix decidua	Pin	0	EU	0
176	FJ845437	0	0	Various	Pin	0	NA	0
177	AY061685	0	0			3	EU	0
178	JF834326	0	0			3	NA	0
179	JF834352	0	0			3	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
180	JF834374	0	0			3	NA	0
181	HQ650754	0	0	Pseudotsuga menziesii	Pin	0	NA	0
182	HM146861	0	0	Pinus sylvestris	Pin	0	EU	0
183	HM044548	0	0	Larix decidua	Pin	0	EU	0
184	EF619749	0	0	Pinus taeda	Pin	0	NA	0
185	AY061730	0	0			3	NA	0
186	DQ061928	1	1	Pinus	Pin	0	EU	0
187	HQ667811	1	0			3	NA	0
188	AF350065	0	0			3	AS	1
189	AB453021	0	0	Dipterocarpus	Dip	1	AS	1
190	AF345250	0	0	Dipterocarpus	Dip	1	AS	1
191	FJ454929	1	0			3	AS	1
192	EU598153	0	0	Various	Ang	1	NA/AS	0
193	EU598193	0	0	Various	Ang	1	NA	0
194	AB629011	1	0			3	AS	0
195	AM087264	0	0	Fagus sylvatica	Fag	1	EU	0
196	GQ268649	0	0	Dipterocarpaceae	Dip	1	AS	1
197	AB458895	0	0	Dipterocarpaceae	Dip	1	AS	1
198	GQ268645	0	0	Various	Ang	1	AS	1
199	GU134509	0	0	Pinus densiflora	Pin	0	AS	0
200	DQ974758	0	0	Quercus douglasii	Fag	1	NA	0
201	FJ803981	0	0	Pinus banksiana	Pin	0	NA	0
202	DQ777996	1	0	Various	Ang	1	NA	0
203	AM087258	0	0	Fagus sylvatica	Fag	1	EU	0
204	GU134499	0	0	Pinus densiflora	Pin	0	AS	0
205	DQ054553	0	0	Various	Ang	1	EU	0
206	DQ422033	0	0			3	EU	0
207	GQ219877	0	0	Fagus	Fag	1	EU	0
208	EU819436	0	0	Quercus	Fag	1	NA	2
209	AB597648	1	0			3	AS	0
210	AJ937993	0	0	Fagus sylvatica	Fag	1	EU	0
211	AB568440	1	0			3	AS	0
212	AB597713	1	0			3	AS	0
213	GQ268652	0	0	Dipterocarpaceae	Dip	1	AS	1
214	JF273557	0	0	Various	Ang	1	AS	0
215	AF350061	0	0			3	AS	1
216	AB458685	0	0	Dipterocarpaceae	Dip	1	AS	1
217	AF345251	0	0			3	AS	1
218	AF345252	0	0			3	AS	1
219	FJ454924	1	0			3	AS	1
220	FJ454927	1	0			3	AS	1
221	FJ454926	1	0			3	AS	1
222	AB571510	0	0	Pinus densiflora	Pin	0	AS	0
223	EU598194	0	0			3	NA	0
224	DQ777993	1	0	Fagus grandifolia	Fag	1	NA	0
225	AY061693	0	0			3	NA/EU	0
226	FJ196950	0	0	Quercus	Fag	1	NA	1
227	JF834343	0	0			3	NA	0
228	AY969522	0	0	Pinaceae	Pin	0	NA	0
229	AY970053	0	0	Pinus taeda	Pin	0	NA	0
230	HQ541830	0	0	Angiosperm	Ang	1	NA	0
231	AB597662	1	0			3	AS	0
232	AB218174	0	0	Quercus crispula	Fag	1	AS	0
233	FJ865563	0	0	Quercus rubra	Fag	1	EU	0
234	AB594937	1	0			3	AS	0
235	HQ022078	0	0			3	NA	0
236	GU222299	0	0	Various	Ang	1	AU	0
237	DQ422018	1	0	Mix	All	2	NA/EU /AS	0
238	AY061723	0	0	Quercus	Fag	1	EU	0
239	AY534202	1	0	Pinus mericata	Pin	0	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
240	AY061681	0	0			3	EU	0
241	GQ268658	0	0	Dipterocarpaceae	Dip	1	AS	1
242	FR731899	0	0	Uapaca guineensis	Phy	1	AF	1
243	FR731869	0	0	Uapaca guineensis	Phy	1	AF	1
244	AB571504	0	0	Pinus densiflora	Pin	0	AS	0
245	AY061652	0	0	Pinaceae	Pin	0	NA/EU	0
246	HM069440	0	0	Pinus	Pin	0	EU	0
247	FJ845430	0	0	Pinaceae	Pin	0	NA/EU	0
248	DQ990850	0	0			3	NA/EU	0
249	DQ273398	0	0	Mix	All	2	NA	0
250	HM488502	0	0	Pseudotsuga menziesii	Pin	0	NA	0
251	AB597710	1	0			3	AS	0
252	GU143030	0	0	Pinus densiflora	Pin	0	AS	0
253	GU981745	0	0			3	AS	0
254	EF218808	0	0	Mix	All	2	NA	2
255	EU819428	0	0	Quercus liaotungensis	Fag	1	NA/AS	0
256	HQ439176	0	0	Mix	All	2	EU/ME/AS	0
257	EF101773	1	0			3	NA	0
258	EU526012	0	0	Pinaceae	Pin	0	NA	0
259	JF834370	0	0			3	NA	0
260	EU526009	0	0	Pinaceae	Pin	0	NA	0
261	AF461607	0	0	Myrtaceae	Myr	1	AU	0
262	EU019918	0	0			3	EU	0
263	EU598197	0	0			3	NA	0
264	AB291753	0	0	Pinus densiflora	Pin	0	AS	0
265	AB291749	0	0			3	AS	0
266	AB291743	0	0			3	AS	0
267	EF218806	0	0	Pseudotsuga menziezii	Pin	0	NA	0
268	EU645647	0	0	Pseudotsuga menziesii	Pin	0	NA	0
269	JF834356	1	0	Pseudotsuga menziesii	Pin	0	NA	0
270	HM488591	0	0	Pinaceae	Pin	0	NA	0
271	DQ061901	0	0	Quercus robur	Fag	1	EU	0
272	EU816643	0	0	Quercus rotundifolia	Fag	1	EU	0
273	DQ061882	1	0	Mix	All	2	EU	0
274	DQ061903	1	0	Mix	All	2	EU	0
275	FN669240	0	0	Populus tremula	Sal	1	EU	0
276	FJ946961	0	0	Mix	Ang	2	NA/EU	0
277	EU232105	0	0	Pinus taeda	Pin	0	NA	2
278	EF611147	0	0	Tilia vulgaris	Mal	1	EU	0
279	DQ367912	0	0			3	NA	0
280	EF611150	0	0	Tilia vulgaris	Mal	1	EU	0
281	AY061663	0	0			3	EU	0
282	DQ422016	0	0			3	EU	0
283	DQ658888	0	0			3	EU	0
284	HE601890	0	0	Quercus ilex	Fag	1	EU	0
285	HQ667803	1	0			3	NA	0
286	FJ845429	0	0	Various	Pin	0	NA	0
287	GU234042	0	0	Dryas octopetala	Ros	1	NA/EU	0
288	EU668265	0	0	Pinus nigra var.maritima	Pin	0	NA/EU	0
289	EF411120	0	0	Quercus wislizeni	Fag	1	NA	0
290	GQ166868	0	0	Quercus	Fag	1	NA	0
291	AY061671	0	0			3	EU	0
292	EU819422	0	0	Various	Ang	1	NA	0
293	DQ061886	1	0	Castanea sativa	Fag	1	EU	0
294	DQ777985	1	0			3	NA	0
295	HQ667805	1	0	Quercus	Fag	1	NA	1
296	EF411133	1	0	Mix	All	2	NA	0
297	FR852101	0	0	Various	Ang	1	ME	0
298	FR731504	0	0	Isoberlinia angolensis	Fab	1	AF	1
299	JF960818	1	0	Eucalyptus delegatensis	Myr	1	AS/AU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
300	FJ803978	0	0	Pinus	Pin	0	NA	0
301	HM141050	1	0			3	AS	0
302	FJ378823	0	0	Mix	Ang	1	AS	0
303	FJ378827	0	0	Kobresia	Cyp	1	EU	0
304	FJ378832	0	0	Kobresia	Cyp	1	AS	0
305	FJ378831	0	0	Kobresia	Cyp	1	AS	0
306	HM100658	0	0	Pinus densiflora	Pin	0	AS	0
307	FJ378822	0	0	Kobresia	Cyp	1	AS	0
308	EF218798	0	0	Pseudotsuga menziesii	Pin	0	NA	0
309	DQ061893	1	0	Mix	All	0	EU	0
310	FJ196949	0	0	Quercus	Fag	1	NA	1
311	DQ061884	1	0	Mix	All	2	EU/ME	0
312	DQ061883	0	0	Quercus	Fag	1	EU	0
313	AF096987	1	0	Mix	All	2	EU/ME	0
314	GQ268636	0	0	Dipterocarpaceae	Dip	1	AS	1
315	FJ688108	1	0			3	EU	0
316	JF834332	1	0	Fagus grandifolia	Fag	1	NA/EU	0
317	DQ061905	1	0	Quercus ilex	Fag	1	EU	0
318	DQ061921	0	0	Quercus ilex	Fag	1	EU	0
319	FR731266	0	0			3	NA	0
320	FR731507	0	0	Uapaca kirkiana	Phy	1	AF	1
321	FR731623	0	0	Brachystegia longifolia	Fab	1	AF	1
322	FR731744	0	0	Caesalpinioideae	Fab	1	AF	1
323	FR731389	0	0	Angiosperm	Ang	1	AF	1
324	FR731755	0	0	Caesalpinioideae	Fab	1	AF	1
325	FR731758	0	0	Caesalpinioideae	Fab	1	AF	1
326	FR731897	0	0	Uapaca heudelotii	Phy	1	AF	1
327	FR731756	0	0	Angiosperm	Ang	1	AF	1
328	FR731757	0	0	Caesalpinioideae	Fab	1	AF	1
329	DQ421985	0	0			3	EU	0
330	FJ454990	1	0			3	AS	1
331	FJ454996	1	0			3	AS	1
332	EF534352	1	0			3	AS	0
333	AB291728	0	0			3	AS	0
334	JF273537	0	0	Various	Ang	1	AS	0
335	AB597630	1	0			3	AS	0
336	GU289649	0	0	Pinus taeda	Pin	0	NA	0
337	EU597075	0	0	Mix	All	2	NA/EU/AS	0
338	DQ422010	0	0	Mix	All	2	NA/EU	0
339	AY061695	0	0	Quercus suber	Fag	1	EU	0
340	FM995572	0	0	Fagus sylvatica	Fag	1	EU	0
341	EF126734	0	0			3	AS	0
342	JF519171	0	0	Fagus sylvatica	Fag	1	EU	0
343	JF834364	0	0	Pinaceae	Pin	0	NA/EU	0
344	AB291767	0	0			3	AS	0
345	JF908707	0	0			3	EU	0
346	GQ268661	0	0	Dipterocarpaceae	Dip	1	AF	1
347	AB291763	0	0	Pinus massoniana	Pin	0	AS	0
348	GQ268650	0	0	Dipterocarpaceae	Dip	1	AS	1
349	AB291762	0	0			3	AS	0
350	AF350067	0	0			3	AS	1
351	AB458686	0	0	Dipterocarpaceae	Dip	1	AS	1
352	FJ454969	1	0			3	AS	1
353	FM999505	0	0	Fagus grandiflolia	Fag	1	NA	0
354	EU019919	0	0			3	AU	0
355	AB218194	0	0	Betula maximowicziana	Bet	1	AS	0
356	AB291760	1	0	Fagus crenata	Fag	1	AS	2
357	HQ021867	0	0			3	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
358	AY606961	0	0	Mix	All	2	EU	0
359	FJ845426	0	0	Pinus densiflora	Pin	0	NA/AS/AF	2
360	JN168745	0	0	Dicymbe corymbosa	Fab	1	SA	1
361	DQ422032	0	0			3	EU	0
362	FR852099	1	0	Various	Ang	1	EU/ME	0
363	EU598152	0	0			3	NA	0
364	FR731896	0	0	Uapaca guineensis	Phy	1	AF	1
365	FR731505	0	0	Cryptosepalum exfoliatum	Fab	1	AF	1
366	FR731889	0	0	Marquesia excelsa	Dip	1	AF	1
367	FR731887	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
368	EU019934	0	0			3	AU	0
369	EU019938	0	0			3	AU	0
370	JF960823	0	0	Eucalyptus delegatensis	Myr	1	AU	0
371	AB459514	0	0	Dipterocarpus alatus	Dip	1	AS	1
372	GU371290	0	0	Pinus	Pin	0	AS	2
373	AY061726	0	0			3	EU	0
374	AY061655	0	0	Fagus sylvatica	Fag	1	EU	0
375	AJ633577	0	0	Pinus taeda	Pin	0	NA	0
376	HQ667813	1	0			3	NA	1
377	GQ359821	1	0	Various	Ang	1	AS	0
378	JF273558	0	0	Various	Ang	1	AS	0
379	EU819426	0	0	Fagaceae	Fag	1	NA	0
380	FJ455025	1	0			3	AS	1
381	FJ455026	1	0			3	AS	1
382	FR731771	0	0	Caesalpinioideae	Fab	1	AF	1
383	AB459511	0	0	Dipterocarpaceae	Dip	1	AS	1
384	AB459518	0	0	Dipterocarpaceae	Dip	1	AS	1
385	AB206535	0	0	Dipterocarpaceae	Dip	1	AS	1
386	GQ268635	0	0	Dipterocarpaceae	Dip	1	EU	0
387	AF345247	0	0			3	AS	1
388	AF345248	0	0			3	AS	1
389	FR731316	0	0	Uapaca densifolia	Phy	1	AF	1
390	FR731727	0	0	Caesalpinioideae	Fab	1	AF	1
391	FR731423	0	0	Angiosperm	Ang	1	AF	1
392	FR731281	0	0	Angiosperm	Ang	1	AF	1
393	JF273535	1	0	Castanopsis fargesii	Fag	1	AS	0
394	DQ398092	0	0	Myrtaceae	Myr	1	AU	0
395	AF350057	0	0			3	AS	1
396	GQ268639	0	0	Dipterocarpaceae	Dip	1	AF	1
397	GQ268632	0	0	Dipterocarpaceae	Dip	1	AS	1
398	FR731823	0	0	Anthonotha macrophylla	Fab	1	AF	1
399	EF218802	0	0	Pseudotsuga mensiezii	Pin	0	NA	0
400	EF218803	0	0	Pseudotsuga mensiezii	Pin	0	NA	0
401	AY702750	0	0	Pinaceae	Pin	0	NA	0
402	AB291748	0	0			3	AS	0
403	AB597671	1	0			3	AS	0
404	AB594932	1	0			3	AS	0
405	EU598202	0	0			3	NA	0
406	GQ268654	0	0	Dipterocarpaceae	Dip	1	AS	1
407	EU019920	0	0			3	AU	0
408	JF960810	0	0	Eucalyptus delegatensis	Myr	1	AU	0
409	GU222265	0	0			3	AU	0
410	EF634147	0	0	Nothofagus menziesii	Fag	1	AU	0
411	AB597705	1	0			3	AS	0
412	AF350063	0	0			3	AS	1
413	DQ422001	0	0			3	EU	0
414	DQ778002	1	0	Pinus	Pin	0	NA	0
415	FJ196294	1	0			3	AS	0
416	GU229820	0	0			3	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
417	DQ778001	1	0			3	NA	0
418	FR731885	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
419	FR731882	0	0	Uapaca guineensis	Phy	1	AF	1
420	FR731908	0	0	Uapaca guineensis	Phy	1	AF	1
421	FR731729	0	0	Caesalpinaceae	Fab	1	AF	1
422	FR731277	0	0	Angiosperms	Ang	1	AF	1
423	FR731487	0	0	Brachystegia spiciformis	Fab	1	AF	1
424	FR731251	0	0	Uapaca densifolia	Phy	1	AF	1
425	FR731900	0	0	Marquesia excelsa	Fab	1	AF	1
426	AM113427	0	0	Various	Ang	1	AF	1
427	FR731906	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
428	FR731490	0	0	Caesalpinioideae	Fab	1	AF	1
429	DQ422028	0	0			3	EU	0
430	FR731178	0	0	Uapaca bojeri	Phy	1	AF	1
431	FR731905	0	0	Caesalpinioideae	Ang	1	AF	1
432	FR731472	0	1	Uapaca densifolia	Phy	1	AF	1
433	FR731909	0	0	Uapaca guineensis	Phy	1	AF	1
434	FR731247	0	1	Uapaca densifolia	Phy	1	AF	1
435	FR731245	0	0	Angiosperm	Ang	1	AF	1
436	FR731310	0	1	Uapaca densifolia	Phy	1	AF	1
437	FR731692	0	0	Caesalpinioideae	Fab	1	AF	1
438	FR731753	0	0	Uapaca staudtii	Phy	1	AF	1
439	FR731268	0	0	Intsia bijuga	Fab	1	AF	1
440	FR731818	0	0	Anthonotha macrophylla	Fab	1	AF	1
441	FR731881	0	0	Aphanocalyx sp.	Fab	1	AF	1
442	AM113433	0	0	Caesalpinioideae	Fab	1	AF	1
443	FR731912	0	0	Uapaca guineensis	Phy	1	AF	1
444	FR731740	0	0	Caesalpinioideae	Fab	1	AF	1
445	FR731654	0	0	Angiosperms	Ang	1	AF	1
446	FR731872	0	0	Uapaca guineensis	Phy	1	AF	1
447	FR731879	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
448	JN168739	0	0	Angiosperm	Ang	1	SA	1
449	FR731336	0	0	Uapaca	Phy	1	AF	1
450	FR731246	0	0	Uapaca densifolia	Phy	1	AF	1
451	FR731883	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
452	FR731730	0	0	Microberlinia bisulcata	Fab	1	AF	1
453	FR731880	0	0	Gilbertiodendron ogoouense	Fab	1	AF	1
454	JN168741	0	0	Dicymbe corymbosa	Fab	1	SA	1
455	FR731734	0	0	Uapaca staudtii	Phy	1	AF	1
456	FR731820	0	0	Anthonotha macrophylla	Fab	1	AF	1
457	FR731262	0	0	Angiosperm	Ang	1	AF	1
458	FR731225	0	0	Angiosperm	Ang	1	AF	1
459	EU712085	0	0			3	NA	0
460	HQ667808	1	0			3	NA	0
461	EU712034	0	0			3	NA	0
462	EU712058	0	0			3	NA	0
463	EU712024	0	0			3	NA	0
464	EU712060	0	0			3	NA	0
465	EU712059	0	0			3	NA	0
466	GU083052	0	0			3	NA	0
467	EU712020	0	0			3	NA	0
468	EU712016	0	0			3	NA	0
469	EU712077	0	0			3	NA	0
470	EU712039	0	0			3	NA	0
471	EU712032	0	0			3	NA	0
472	EU712027	0	0			3	NA	0
473	EU712062	0	0			3	NA	0
474	EU712073	0	0			3	NA	0
475	EU712056	0	0			3	NA	0
476	GU083123	0	0	Betula papyrifera	Bet	1	NA	0
477	EU712017	0	0			3	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
478	GU083106	0	0			3	NA	0
479	EU712051	0	0			3	NA	0
480	EU712065	0	0			3	NA	0
481	EU712048	0	0			3	NA	0
482	EU712079	0	0			3	NA	0
483	EU712052	0	0			3	NA	0
484	EU712053	0	0			3	NA	0
485	GU083162	0	0			3	NA	0
486	EU712076	0	0			3	NA	0
487	EU712094	0	0			3	NA	0
488	EU712081	0	0			3	NA	0
489	EU712043	0	0			3	NA	0
490	GU981742	0	0			3	NA	0
491	FJ455004	1	0			3	AS	1
492	EU019922	0	0			3	AU	0
493	FN610945	0	0	Fagus sylvatica	Fag	1	EU	0
494	AY061679	0	0	Mix	All	2	EU	0
495	EU598196	0	0	Various	Ang	1	NA	0
496	DQ422030	0	0			3	EU	0
497	DQ974761	0	0	Mix	All	2	NA	0
498	JF834366	0	0			3	NA	0
499	EU712012	0	0	Peudotsuga menziesii	Pin	0	NA	0
500	EU712071	0	0			3	NA	0
501	JF300767	0	0	Pinus sylvestris	Pin	0	NA/EU	0
502	EU712070	0	0			3	NA	0
503	EU712011	0	0			3	NA	0
504	EU712033	0	0			3	NA	0
505	FM992958	0	0	Picea abies	Pin	0	EU	0
506	DQ421999	0	0	Picea mariana	Pin	0	NA/EU	0
507	HM146850	0	0	Pinus sylvestris	Pin	0	EU	0
508	EF619750	0	0	Pinus taeda	Pin	0	NA	0
509	EU569273	0	0	Quercus	Fag	1	NA	2
510	JF908705	0	0			3	EU	0
511	FJ946960	0	0	Quercus	Ang	1	EU	0
512	DQ061911	0	0	Quercus ilex	Fag	1	EU/ME	0
513	FR852113	0	0	Various	Ang	1	ME	0
514	DQ061912	0	0	Pinus	Pin	0	EU	0
515	HQ604836	0	0			3	NA	0
516	HQ650736	0	0			3	NA	0
517	FM999569	0	0	Fagus grandifolia	Fag	1	NA	0
518	FR731187	0	0	Angiosperm	Ang	1	AF	1
519	FR731877	0	0	Uapaca heudelotii	Phy	1	AF	1
520	FR731951	0	0	Uapaca guineensis	Phy	1	AF	1
521	FR731646	0	0	Aphanocalyx sp.	Fab	1	AF	1
522	FR731694	0	0	Caesalpinioideae	Fab	1	AF	1
523	FR731878	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
524	GQ219838	0	0	Fagus	Fag	1	EU	0
525	AY061675	0	0			3	EU	0
526	DQ421983	0	0			3	EU	0
527	EU057119	0	0	Tsuga heterophylla	Pin	0	NA	0
528	JF834327	0	0			3	NA	1
529	EU597082	0	0	Various	Pin	0	NA	0
530	DQ421987	0	0	Mix	All	2	NA/EU/AS	0
531	AB597642	1	0			3	AS	0
532	AB218100	0	0	Quercus crispula	Fag	1	AS	0
533	AY061715	0	0	Various	Ang	1	EU	0
534	EU819429	0	0	Fagus grandifolius	Fag	1	NA/EU	0
535	FM999567	0	0	Fagus grandifolius	Fag	1	NA/EU	0
536	EU598173	0	0			3	NA	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
537	EU598175	0	0			3	NA	0
538	AY061688	1	0			3	EU	0
539	DQ422015	0	0			3	EU	0
540	AY878657	1	0	Pseudotsuga menziesii	Pin	0	NA	0
541	FR852097	0	0	Various	Ang	1	ME	0
542	EU019939	0	0			3	AU	0
543	EU019933	0	0			3	AU	0
544	DQ422004	0	0			3	EU	0
545	DQ422019	0	0			3	EU/AS	0
546	EU598156	0	0			3	NA	0
547	EU598158	0	0			3	NA	0
548	EU598159	0	0			3	NA	0
549	DQ422031	0	0			3	EU	0
550	AY061662	0	0			3	EU	0
551	EU598165	0	0			3	NA	0
552	DQ422025	0	0			3	EU/AS	0
553	FN557555	0	0	Coccoloba	Pol	1	SA	1
554	EU019931	0	0			3	AU	0
555	EU019930	0	0			3	AU	0
556	EU019936	0	0			3	AU	0
557	EU019915	0	0			3	AU	0
558	EU019932	0	0			3	AU	0
559	FN557558	0	0	Neea comun	Nyc	1	SA	1
560	EU598167	0	0			3	NA	0
561	DQ422006	0	0			3	EU	0
562	EU598164	0	0			3	NA	0
563	DQ422021	0	0			3	EU	0
564	EU019928	0	0			3	AU	0
565	EU019929	0	0			3	AU	0
566	DQ422012	0	0			3	EU	0
567	DQ422007	0	0	Fagus sylvatica	Fag	0	EU	0
568	EU598163	0	0			3	NA/AS	0
569	DQ422027	0	0			3	EU	0
570	DQ422029	0	0			3	EU/AS	0
571	DQ421998	1	0			3	NA/EU /AS	0
572	FN557552	0	0	Guapira clasica	Nyc	1	SA	1
573	JN168743	0	0	Aldina insignis	Fab	1	SA	1
574	JN168746	0	0	Dicymbe altsonii	Fab	1	SA	1
575	EU019916	0	0			3	AU	0
576	FR731291	0	1	Angiosperm	Ang	1	AF	1
577	AB600187	1	0			3	AS	0
578	AB594977	1	0			3	AS	0
579	AB600188	1	0	Pinus massoniana	Pin	0	AS	0
580	AB594933	1	0			3	AS	0
581	HM240161	0	0	Mix	All	2	NA	0
582	HQ667810	0	0			3	NA	0
583	DQ377406	0	0	Pinus taeda	Pin	0	NA	0
584	AB211275	1	0	Angiosperm	Ang	1	NA/AS	2
585	FJ348386	0	0	Quercus garryana	Fag	1	NA	0
586	JN681168	0	0	Pinus muricata	Pin	0	NA	0
587	DQ822824	1	0	Pinus	Pin	0	NA/AS	0
588	AY880930	0	0	Mix	All	2	NA	0
589	EU819432	0	0	Quercus	Fag	1	NA	0
590	AY061732	0	0			3	NA	0
591	AY061706	0	0			3	EU	0
592	HM146852	0	0	Mix	All	2	EU/AU	0
593	AJ438036	0	1	Mix	All	2	EU	0
594	AF230891	0	1			3	EU	0
595	AF230890	0	1			3	EU	0
596	JN887986	0	0	Pinus montezumae	Pin	0	NA	1

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
597	AB636426	0	0	Pinus massoniana	Pin	0	AS	0
598	EU819534	0	0	Various	Ang	1	NA	0
599	EU819425			Various	Ang	1	NA	0
600	EU819493	0	0	Mix	All	2	NA	0
601	EU816667	0	0	Various	Ang	1	EU	0
602	DQ061930	1	0	Mix	All	2	EU	0
603	DQ422026	0	0	Various	Ang	1	EU/ME	0
604	JF273538	0	0	Quercus liaotungensis	Fag	1	AS	0
605	AY239338	0	1	Pinaceae	Pin	0	NA	0
606	FR852096	1	0	Various	Ang	1	EU/ME	0
607	AY239337	0	1			3	NA	0
608	AY970122	0	0	Pinus taeda	Pin	0	NA	0
609	AB211276	0	0	Salix rainii	Sal	1	AS	0
610	AY061700	0	0	Quercus ilex	Fag	1	EU	0
611	GU371297	0	0	Mix	All	2	AS	0
612	AY061736	0	0			3	NA	0
613	EU019941	0	0			3	AU	0
614	DQ388874	0	0	Eucalyptis	Myr	1	AU	0
615	EU019945	0	0	Eucalyptis	Myr	1	AU	0
616	JF960808	0	0	Eucalyptus delegatensis	Myr	1	AU	0
617	DQ178934	1	0			3	AU	0
618	DQ178932	1	1			3	AU	0
619	EF090512	1	1			3	AU	0
620	JF960812	0	0	Eucalyptus delegatensis	Myr	1	AU	0
621	DQ388809	0	0	Eucalyptus	Myr	1	AU	0
622	DQ388830	0	1	Eucalyptis pilularis	Myr	1	AU	0
623	DQ093423	0	1	Dipterocarpaceae	Dip	1	AS	1
624	AB453035	0	0	Dipterocarpus	Dip	1	AS	1
625	EU712095	0	1			3	NA	0
626	AY239335	0	1			3	NA	0
627	AY969536	0	1	Various	Ang	1	NA	0
628	AY239319	0	1			3	NA	0
629	HQ204656	0	0	Quercus ilex	Fag	1	EU	0
630	AY239349	1	1	Pinaceae	Pin	0	NA	0
631	HM021172	0	1	Pinus murcata	Pin	0	NA	0
632	AF230894	0	1			3	EU	0
633	GQ219836	0	1	Fagus	Fag	1	EU	0
634	EU598184	0	0	Pinus	Pin	0	NA	0
635	EU284011	0	0	Mix	All	2	NA	0
636	HQ021868	0	0			3	NA	0
637	GQ219888	0	0	Fagus	Fag	1	EU	0
638	FJ454957	1	0			3	AS	1
639	FJ454960	1	0			3	AS	1
640	AY061735	0	0			3	NA	0
641	DQ422024	0	0			3	EU	0
642	FJ623066	1	0			3	AS	1
643	JF908666	0	0			3	EU	0
644	HQ677769	0	0			3	EU	0
645	EU598188	0	0			3	NA	0
646	EU598187	0	1	Castanea dentata	Fag	1	NA	0
647	HE647707	0	1	Abies pindro	Pin	0	AS	0
648	AJ937998	0	0	Fagus sylvatica	Fag	1	EU	0
649	GQ219932	0	0	Fagus sylvatica	Fag	1	EU	0
650	FJ403509	0	0	Fagus sylvatica	Fag	1	EU	0
651	GQ219911	0	0	Fagus	Fag	1	EU	0
652	GQ219922	0	1	Fagus	Fag	1	EU	0
653	AY351624	1	0	Fagus sylvatica	Fag	1	EU	0
654	EU284010	0	0	Pinaceae	Pin	0	NA	0
655	JF960824	0	1	Eucalyptus delegatensis	Myr	1	AU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
656	AB259126	1	1	Mix	All	2	AS	0
657	FJ454978	1	1	Various	Ang	1	AS	1
658	EU880224	0	1	Various	Ang	1	NA	0
659	EF611133	0	0	Tilia cordata	Mal	1	EU	0
660	EF611135	0	0	Tilia cordata	Mal	1	EU	0
661	AY061677	1	0			3	EU/AS	0
662	DQ422023	0	0			3	EU	0
663	FJ845427	0	1	Mix	All	2	NA/EU	0
664	DQ146386	0	0	Cotylelobium lanceolata	Dip	1	AS	1
665	DQ388829	0	0	Eucalyptis	Myr	1	AU	0
666	GU222258	0	0			3	AU	0
667	DQ388850	0	1	Eucalyptus pilularis	Myr	1	AU	0
668	GQ268646	0	0	Dipterocarpaceae	Dip	1	AS	1
669	AF418609	0	0			3	NA	0
670	AY061682	1	0			3	EU	0
671	AB594963	1	0			3	AS	0
672	AB594965	1	0	Fagus crenata	Fag	1	AS	0
673	AF349709	1	0	Quercus	Fag	1	NA	2
674	AB568431	1	0	Pinus massoniana	Pin	0	AS	0
675	HQ022145	0	0			3	NA	0
676	JF273536	0	0	Various	Ang	1	AS	0
677	AY061692	0	0			3	EU	0
678	AM113429	0	0	Various	Ang	1	AF	1
679	DQ178935	1	0			3	AU	0
680	FR731494	0	0	Cryptosepalum exfoliatum	Fab	1	AF	1
681	FR731754	0	0	Caesalpinioideae	Fab	1	AF	1
682	FR731819	0	0	Anthonotha macrophylla	Fab	1	AF	1
683	FR731895	0	0	Uapaca	Phy	1	AF	1
684	FR731359	0	0	Uapaca	Phy	1	AF	1
685	FR731411	0	0	Asteropeia micraster	Ast	1	AF	1
686	FR731815	0	0	Anthonotha macrophylla	Fab	1	AF	1
687	FR731821	0	0	Anthonotha macrophylla	Fab	1	AF	1
688	FR731615	0	0	Julbernardia paniculata	Fab	1	AF	1
689	FR731741	0	0	Caesalpinioideae	Fab	1	AF	1
690	FR731888	0	0	Tetraberlinia bifoliolata	Fab	1	AF	1
691	FR731656	0	0	Fabaceae	Fab	1	AF	1
692	AB594957	1	0			3	AS	0
693	AB594961	1	0			3	AS	0
694	AB597698	1	0			3	AS	0
695	EF661990	0	0	Quercus douglasii	Fag	1	NA	0
696	JN172982	0	0	Larix decidua	Pin	0	EU	0
697	GU220371	0	0	Angiosperm	Ang	1	NA/EU	0
698	AY061729	0	0			3	NA/EU	0
699	AJ534937	0	0			3	EU	0
700	HM196031	0	0			3	EU	0
701	EU563497	0	0	Quercus	Fag	1	NA	2
702	DQ273395	0	0	Notholithocarpus densiflorus	Fag	1	NA	0
703	GU234011	0	0	Mix	All	2	NA/EU	0
704	JF834375	0	0			3	NA	0
705	AM930237	0	0	Quercus	Fag	1	EU	0
706	AF230897	0	0			3	EU	0
707	EF040856	0	0	Fagaceae	Fag	1	EU	0
708	DQ990846	0	0	Quercus	Fag	1	EU	0
709	DQ778000	1	0			3	NA	0
710	HQ604846	0	0	Mix	All	2	NA	0
711	DQ777971	1	0	Various	Ang	1	NA	0
712	GU391436	0	0	Pinus densiflora	Pin	0	AU	0
713	HQ604845	0	0	Mix	All	2	NA/EU	0
714	JF899570	0	0	Tsuga heterophylla	Pin	0	NA	0
715	DQ421997	0	0			3	EU	0
716	GU234120	0	0	Mix	All	2	NA/EU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
717	FJ158067	0	0	Pinus sylvestris	Pin	0	EU	0
718	EU597051	0	0	Various	Pin	0	NA	0
719	JF908658	0	0			3	EU	0
720	HQ650737	0	0	Mix	All	2	NA/EU	0
721	EF634131	0	0	Angiosperm	Ang	1	AU	0
722	DQ777999	1	0			3	NA	0
723	HQ204702	0	0	Quercus ilex	Fag	1	EU	0
724	HQ021792	0	0			3	NA	0
725	HQ703018	0	0			3	EU	0
726	AY061676	0	0	Various	Ang	1	EU	0
727	JF908675	0	0			3	EU	0
728	DQ054560	0	0	Fagaceae	Fag	1	EU	0
729	AY281091	0	0			3	NA	0
730	AY061666	0	0			3	EU	0
731	DQ974760	0	0	Quercus douglasii	Fag	1	NA	0
732	GQ221653	0	0			3	NA	0
733	EF411090	0	0	Quercus wislizeni	Fag	1	NA	0
734	JF908644	0	0	Quercus	Fag	1	EU	0
735	FJ196956	0	0	Quercus	Fag	1	NA	1
736	DQ377401	0	0	Pinus taeda	Pin	0	NA	0
737	JN084192	0	0			3	AS	0
738	JF748076	0	0	Quercus liaotungensis	Fag	1	AS	0
739	AB597695	1	0			3	AS	0
740	GU134512	0	0	Pinus densiflora	Pin	0	AS	0
741	EU569272	0	0	Quercus	Fag	1	NA	1
742	GU981743	0	0			3	NA	0
743	AY061674	1	0	Mix	All	2	NA/EU	0
744	EU711957	0	0			3	NA	0
745	EU711964	0	0	Betula papyrifera	Bet	1	NA	0
746	AY061678	0	0	Betula papyrifera	Bet	1	NA/EU	0
747	EU057100	0	0	Pinaceae	Pin	0	NA	0
748	FJ378820	0	0	Kobresia	Cyp	1	AS	0
749	EF434062	0	0	Picea mariana	Pin	0	NA	0
750	AY061707	0	0	Mix	All	2	NA/EU	0
751	AY239346	0	1			3	NA	0
752	JF908641	0	0			3	EU	0
753	JF908668	1	0	Various	Pin	0	EU	0
754	AF418625	1	0			3	EU	0
755	AY822742	0	0	Mix	All	2	NA	0
756	GU180318	0	0	Pinus mericata	Pin	0	NA	0
757	FJ845434	0	0	Pseudotsuga menziezii	Pin	0	NA	0
758	HQ604841	0	0	Pseudotsuga menziezii	Pin	0	NA	0
759	AY061711	0	0			3	EU	0
760	DQ367914	1	0	Pseudotsuga menziesii	Pin	0	NA/EU	0
761	EF619751	0	0	Pinus taeda	Pin	0	NA	0
762	HQ604839	0	0			3	NA	0
763	HQ604843	0	0			3	NA	0
764	JF908649	0	0			3	EU	0
765	EF619753	0	0	Mix	All	2	NA	0
766	FJ803976	0	0	Pinus banksiana	Pin	0	NA	0
767	FJ816746	0	0	Pinus pinaster	Pin	0	EU	0
768	AY061718	0	0			3	EU	0
769	AF418626	0	0	Pinus	Pin	0	EU/SA	0
770	JF908689	0	0			3	EU	0
771	AF418623	0	0			3	EU	0
772	DQ974757	1	0	Quercus douglasii	Fag	1	NA/EU	0
773	AY061680	0	0			3	EU	0
774	JF908667	0	0			3	EU	0
775	AB218078	0	0	Fagus crenata	Fag	1	AS	0
776	AF418620	0	0	Fagus sylvatica	Fag	1	EU	0
777	AB597703	1	0			3	AS	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
778	AY061712	0	0	Mix	All	2	NA	0
779	EU569277	0	0	Mix	All	2	NA	2
780	AB594944	1	0			3	AS	0
781	AB594953	1	0			3	AS	0
782	AB594954	1	0			3	AS	0
783	AB594948	1	0			3	AS	0
784	AB594970	1	0			3	AS	0
785	AB597697	1	0			3	AS	0
786	AB594946	1	0			3	AS	0
787	GQ219923	0	0	Fagus	Fag	0	EU	0
788	GQ219949	0	0	Fagus	Fag	1	EU	0
789	FJ188358	0	0	Pinus pinaster	Pin	0	EU	0
790	GU391440	0	0	Pinus densiflora	Pin	0	AS	0
791	FJ816748	0	0	Mix	All	2	EU	0
792	FJ158068	0	0	Pinus sylvestris	Pin	0	EU	0
793	EU700257	0	0	Quercus	Fag	1	EU	0
794	GQ219862	0	0	Mix	All	2	EU	0
795	AY254872	0	0	Pinus sylvestris	Pin	0	EU	0
796	JF519103	0	0	Fagus sylvatica	Fag	1	EU	0
797	HM015478	0	0	Pinus nigra	Pin	0	EU	0
798	AY061728	0	0			3	EU	0
799	FM993279	0	0	Alnus	Bet	1	EU	0
800	FJ627039	0	0			0	NA	0
801	FR852109	0	0	Various	Ang	1	ME	0
802	DQ777992	1	0	Various	Ang	1	NA/EU	0
803	GU391443	0	0	Pinus densiflora	Pin	0	AS	0
804	AY061654	0	0	Mix	All	2	NA	0
805	DQ493554	0	0			3	NA	0
806	DQ493555	1	0			3	NA	0
807	DQ493552	0	0			3	NA	0
808	DQ493558	0	0			3	NA	0
809	AB218203	0	0	Betula grossa	Bet	1	AS	0
810	EF619756	0	0	Pinus taeda	Pin	0	NA	0
811	AF418618	0	0	Various	Ang	1	EU	0
812	HQ604847	0	0			3	NA	0
813	AY656977	0	0	Quercus	Fag	1	NA	0
814	AJ937992	0	0	Abies alba	Pin	0	EU	0
815	EF040867	0	0	Castanea sativa	Fag	1	EU	0
816	DQ990849	0	0	Various	Ang	1	EU	0
817	AF418621	0	0			3	EU	0
818	JF519017	0	0	Fagus sylvatica	Fag	1	EU	0
819	FJ196951	0	0	Quercus	Fag	1	NA	1
820	EU569262	1	0	Quercus	Fag	1	NA	1
821	HQ022216	0	0			3	NA	0
822	DQ481997	0	0	Pseudotsuga menziesii	Pin	0	NA	0
823	AY061657	0	0	Picea abies	Pin	0	EU	0
824	GU966633	0	0			3	AS	0
825	FJ845435	0	0	Pinaceae	Pin	0	NA/EU	0
826	HQ022215	1	0			3	NA	0
827	FJ152483	0	0	Tsuga heterophylla	Pin	0	NA	0
828	JF908661	0	0			3	EU	0
829	AB218161	0	0	Fagus japonica	Fag	1	AS	0
830	EU569264	0	0	Quercus	Fag	1	NA	1
831	EU598170	0	0			3	NA	0
832	EU598171	0	0			3	NA	0
833	HQ667807	1	0			3	NA	0
834	DQ778005	1	0			3	NA	0
835	EU569278	0	0	Quercus	Fag	1	NA	1
836	AY061659	0	0			3	EU	0
837	AF345249	0	0			3	AS	1
838	AM087280	0	0	Fagus sylvatica	Fag	1	EU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
839	AY061714	0	0			3	EU	0
840	FJ196946	0	0	Mix	All	2	NA	2
841	JF834334	0	0			3	NA	0
842	AB568435	1	0			3	AS	0
843	AB597661	1	0			3	AS	0
844	FJ196947	0	0	Quercus	Fag	1	NA	1
845	AY969918	0	0	Pinus taeda	Pin	0	NA	0
846	DQ403804	0	1			3	NA	0
847	EU019921	0	0			3	AU	0
848	GU222260	0	0			3	AU	0
849	GQ268648	0	0	Dipterocarpaceae	Dip	1	AS	1
850	JF960811	0	0	Eucalyptus delegatensis	Myr	1	AU	0
851	GQ268642	0	0	Dipterocarpaceae	Dip	1	AS	1
852	FR731739	0	0	Caesalpinioideae	Fab	1	AF	1
853	FR731746	0	0	Caesalpinioideae	Fab	1	AF	1
854	FR731822	0	0	Anthonotha macrophylla	Fab	1	AF	1
855	FR731833	0	0	Microberlinia bisulcata	Fab	1	AF	1
856	AM113428	0	0	Fabaceae	Fab	1	AF	1
857	FR731618	0	0	Isoberlinia angolensis	Fab	1	AF	1
858	FR731750	0	0	Brachystegia longifolia	Fab	1	AF	1
859	FR731485	0	0	Monotes glaber	Dip	1	AF	1
860	DQ990845	0	0			3	EU	0
861	DQ422013	0	0			3	EU	0
862	AF418641	0	0			3	EU	0
863	HQ022146	0	0			3	NA	0
864	AY061708	0	0			3	EU	0
865	FR852107	0	0	Various	Ang	1	ME	0
866	GQ240916	0	0	Pinus massoniana	Pin	0	AS	0
867	GQ268662	0	0	Dipterocarpaceae	Dip	1	AF	1
868	JF273539	0	0	Castanopsis fargesii	Fag	1	AS	0
869	AJ937985	0	0	Fagus sylvatica	Fag	1	EU	0
870	JF960807	0	0	Eucalyptus delegatensis	Myr	1	AU	0
871	AF096978	0	0			3	EU	0
872	JF834367	0	0			3	NA	0
873	EU598162	1	0	Quercus	Fag	1	NA	0
874	EU019944	0	0			3	AU	0
875	GQ268651	0	0	Dipterocarpaceae	Dip	1	AF	1
876	FR731643	0	0	Caesalpinioideae	Fab	1	AF	1
877	AY239303	0	1			3	NA	0
878	FJ789601	0	1	Pinaceae	Pin	0	NA	0
879	AY239306	0	1	Pseudotsuga menziesii	Pin	0	NA	0
880	DQ028476	0	1	Quercus	Fag	1	NA	0
881	AY878656	0	1	Pseudotsuga menziesii	Pin	0	NA	0
882	DQ403803	0	1			3	NA	0
883	GU222323	0	1			3	AU	0
884	HQ285393	0	0	Pinus banksiana	Pin	0	NA	0
885	AY061686	0	0			3	EU	0
886	DQ990848	0	0	Castanea sativa	Fag	1	EU	0
887	GU222325	0	0			3	AU	0
888	AB597682	1	0			3	AS	0
889	EU266067	0	0			3	AS	0
890	AY061690	0	0			3	NA	0
891	AB218190	0	0	Betula maximowicziana	Bet	1	AS	0
892	AY061665	0	0	Betula	Bet	1	NA/EU	0
893	AY061724	0	0	Mix	All	2	NA/EU	0
894	EU711941	0	0			3	NA	0
895	FJ152485	0	0	Tsuga heterophylla	Pin	0	NA	0
896	FJ152484	0	0	Tsuga heterophylla	Pin	0	NA	0
897	HM044604	0	0	Pinus cembra	Pin	0	EU	0
898	HM044550	0	0	Larix decidua	Pin	0	EU	0
899	AY839221	0	0			3	EU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
900	AY194601	0	0	Pinus sylvestris	Pin	0	EU	0
901	AY061670	0	0	Pinus cembra	Pin	0	EU	0
902	JF300824	0	0	Pinus sylvestris	Pin	0	EU	0
903	HM044598	0	0	Pinus cembra	Pin	0	EU	0
904	EF521212	0	0	Pinus	Pin	0	EU	0
905	DQ777986	1	0			3	NA	0
906	AY061717	0	0			3	EU	0
907	AB597665	1	0	Carpinus japonica	Bet	1	AS	0
908	JF834349	0	0	Quercus suber	Fag	1	NA/EU	0
909	DQ422017	0	0	Mix	All	2	NA/EU	0
910	AF418630	0	0	Various	Ang	1	EU/ME	0
911	AY061661	0	0			3	EU	0
912	JF908678	0	0			3	EU	0
913	AF418633	0	0			3	EU	0
914	FJ196954	0	0	Quercus	Fag	1	NA	1
915	EF641837	0	0	Mix	All	2	NA	0
916	JF908700	0	0	Quercus ilex	Fag	1	EU	0
917	AY245542	0	0	Pinus mericata	Pin	0	NA	0
918	FM999541	0	0	Fagaceae	Fag	1	NA	0
919	FM999531	0	0	Fagus grandifolia	Fag	1	NA	0
920	FM999515	0	0	Fagus grandifolia	Fag	1	NA	0
921	EF218811	0	0	Betula papyrifera	Bet	1	EU	0
922	AJ534905	1	0	Mix	All	2	EU	0
923	JF834344	0	0			3	NA	0
924	AF349711	1	0	Fagaceae	Fag	1	NA	0
925	EF627042	0	0	Fagaceae	Fag	1	AS	1
926	FJ613980	0	0	Fagaceae	Fag	1	AS	2
927	FJ613988	0	0	Fagaceae	Fag	1	AS	1
928	FJ613927	0	0	Fagaceae	Fag	1	AS	1
929	DQ777988	1	0	Various	Ang	1	NA/EU	0
930	AY656975	0	0	Quercus	Fag	1	NA	0
931	JF834371	0	0	Quercus wislizeni	Fag	1	NA	0
932	JF834368	0	0			3	NA	0
933	AY061683	0	0	Pinus	Pin	0	EU	0
934	AY061684	0	0	Mix	All	2	NA/EU/ME	0
935	EU563492	0	0	Quercus	Fag	1	NA	1
936	FJ196953	0	0	Quercus	Fag	1	NA	1
937	GU997804	0	0	Betula nana	Bet	1	NA	0
938	GU997738	0	0	Betula nana	Bet	1	NA	0
939	GU997933	0	0	Betula nana	Bet	1	NA	0
940	GU997734	0	0	Betula nana	Bet	1	NA	0
941	GU997963	0	0	Betula nana	Bet	1	NA	0
942	AY656944	0	0	Quercus	Fag	1	NA	0
943	GU998321	0	0	Betula nana	Bet	1	NA	0
944	HQ260227	1	0	Various	Ang	1	NA	0
945	AF349713	1	0			3	NA	0
946	AJ971402	1	0	Pinaceae	Pin	0	EU	0
947	DQ367916	0	0	Tsuga heterophylla	Pin	0	NA	0
948	DQ777982	1	0			3	NA	0
949	AY239343	0	1			3	NA	0
950	AY239342	0	1			3	NA	0
951	FJ789600	0	1	Pinaceae	Pin	0	NA	0
952	HQ667804	1	0			3	NA	0
953	GQ221638	0	0			3	NA	0
954	EF411125	0	0	Quercus wislizeni	Fag	1	NA	0
955	GQ219844	0	0			3	EU	0
956	AF418639	0	0			3	EU	0
957	GU134506	0	0	Pinus densifolia	Pin	0	AS	0
958	DQ061918	0	0	Castanea sativa	Fag	1	EU	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
959	HQ204705	1	0	Mix	All	2	EU	0
960	JF908710	0	0	Dryas octopetala	Ros	1	NA/EU	0
961	EU569276	0	0	Quercus	Fag	1	NA	2
962	AY918955	0	0	Quercus garyanna	Fag	1	NA	0
963	GU234031	0	0	Salix reticulata	Sal	1	NA/EU	0
964	JF304348	0	0	Salix reticulata	Sal	1	NA/EU	0
965	FM999623	0	0	Fagus grandifolia	Fag	1	NA	0
966	AB597680	1	0			3	AS	0
967	AB597681	1	0			3	AS	0
968	AY061725	0	0			3	EU	0
969	FN669244	0	0	Populus tremula	Sal	1	EU	0
970	HE601886	0	0	Quercus ilex	Fag	1	EU	0
971	AY061664	0	0			3	EU	0
972	HE601889	0	0	Quercus ilex	Fag	1	EU	0
973	FJ897214	0	0	Quercus rotundifolia	Fag	1	EU	0
974	FM999500	0	0	Fagus grandiflolia	Fag	1	NA	0
975	AY061691	0	0			3	EU	0
976	DQ493553	0	0	Quercus rubra	Fag	1	NA	0
977	HM057197	0	0	Quercus rubra	Fag	1	NA	0
978	DQ777974	1	0			3	NA	0
979	GQ268643	0	0	Dipterocarpaceae	Dip	1	AS	1
980	GQ240917	0	0	Castanopsis fargesii	Fag	1	AS	0
981	AB568439	1	0			3	AS	0
982	AB629052	1	0			3	AS	0
983	AB628993	1	0	Various	Ang	1	AS	0
984	AB629037	1	0			3	AS	0
985	AB629054	1	0			3	AS	0
986	AB629038	1	0			3	AS	0
987	AB629032	1	0			3	AS	0
988	AB629046	1	0			3	AS	0
989	AB629033	1	0			3	AS	0
990	AB629018	1	0			3	AS	0
991	AB629050	1	0			3	AS	0
992	AB629040	1	0			3	AS	0
993	AB629039	1	0			3	AS	0
994	AB629030	1	0			3	AS	0
995	AB629056	1	0			3	AS	0
996	AB629042	1	0			3	AS	0
997	AB629057	1	0			3	AS	0
998	AB629000	1	0	Various	Ang	1	AS	0
999	AB629044	1	0			3	AS	0
1000	AB629045	1	0			3	AS	0
1001	AB629010	1	0			3	AS	0
1002	AB629048	1	0			3	AS	0
1003	GQ359819	1	0	Various	Ang	1	AS	0
1004	AY061689	0	0	Populus tremula	Sal	1	EU	0
1005	FJ454917	1	0			3	AS	1
1006	GU328588	0	0	Various	Ang	1	NA	0
1007	AY310852	0	0	Pinaceae	Pin	0	NA	0
1008	GQ221634	0	0			3	NA	0
1009	DQ384581	0	0			0	NA	0
1010	AY239348	0	1	Pinaceae	Pin	0	NA	0
1011	AY061651	0	0			3	EU	0
1012	AY061667	0	0			3	EU	0
1013	FR852115	0	0	Various	Ang	1	ME	0
1014	JF495176	0	0	Fagus sylvatica	Fag	1	EU	0
1015	EU569275	0	0	Quercus	Fag	1	NA	1
1016	HQ604851	0	0	Kobresia	Cyp	1	NA/AS	0
1017	FN610934	0	0	Fagus sylvatica	Fag	1	EU	0
1018	DQ054555	0	0	Fagus sylvatica	Fag	1	EU	0
1019	FR852116	0	0	Various	Ang	1	ME	0

Table 5 Continued

Clusters	Genbank accessions	Orchid	Truffle	Host	Host Family	Cluster Host	Geog	Cluster Geog
1020	AF418627	0	0	Various	Ang	1	EU/ME	0
1021	EU375711	0	0	Querucs	Fag	1	NA	0
1022	FJ389445	0	0	Quercus rubra	Fag	1	NA	0
1023	EU819515	0	0	Quercus alba	Fag	1	NA	0
1024	FJ389447	0	0	Quercus rubra	Fag	1	NA	0
1025	AF418629	0	0			3	EU	0
1026	AY061698	0	0	Quercus	Fag	1	EU	0
1027	AY061722	0	0	Mix	All	2	NA/EU/AS/ME	0
1028	FR852100	0	0	Various	Ang	1	ME	0
1029	HQ604852	0	0	Pinaceae	Pin	0	NA/EU	0
1030	DQ822825	0	0	Pinus	Pin	0	NA	2
1031	DQ777980	1	0	Mix	All	2	NA	0
1032	DQ777978	1	0			3	NA	0
1033	DQ195594	0	0	Alnus acuminata	Bet	1	SA	0
1034	AY061719	0	0	Betula	Bet	1	NA/EU	0
1035	GQ900534	0	0	Castanopsis fargesii	Fag	1	AS	0
1036	AB597667	1	0	Mix	All	2	AS	0
1037	EU569279	0	0	Quercus	Fag	1	NA	1
1038	JF834373	0	0			3	NA	0
1039	AY061705	0	0	Various	Ang	1	NA/EU	0
1040	EU711826	0	0			3	NA	0
1041	EU711843	0	0			3	NA	0
1042	EU711818	0	0			3	NA	0
1043	EU711823	0	0			3	NA	0
1044	FJ803947	1	0	Larix gmelinii	Pin	0	NA/EU/AS	0
1045	FJ845433	0	0	Mix	All	2	NA/EU	0
1046	JF273534	0	0	Castanopsis fargesii	Fag	1	AS	0
1047	JF834333	0	0			3	NA	0
1048	AY534210	0	0	Pinus sylvestris	Pin	0	NA/EU	0
1049	EU645649	0	0	Pseudotsuga menziesii	Pin	0	NA	0
1050	AY061734	0	0	Mix	All	2	NA/EU	2
1051	EU645601	0	0	Pseudotsuga menziesii	Pin	0	NA	0
1052	JF908704	0	0	Pinus sylvestris	Pin	0	EU	0
1053	HM240542	0	0	Larix decidua	Pin	0	NA/EU	0
1054	JF908684	0	0			3	EU	0
1055	EF530944	0	0			3	NA	0
1056	DQ777977	1	0			3	NA	0
1057	FJ389453	0	0	Quercus rubra	Fag	1	NA	0
1058	HQ667806	1	0			3	NA	1
1059	AY061656	0	0			3	EU	0
1060	AY656978	0	0	Quercus	Fag	1	NA	0
1061	EU819421	0	0	Various	Ang	1	NA	0
1062	DQ777976	1	0			3	NA	0
1063	EU711841	0	0			3	NA	0
1064	EU597074	0	0	Pseudotsuga menziesii	Pin	0	NA	0

CHAPTER II
SYSTEMATICS, BIOGEOGRAPHY, AND CHARACTER EVOLUTION OF *RUSSULA* SUBSECTION *ROSEINAE*

The doctoral student collected specimens, sequenced all specimens, performed all the analyses, and wrote the manuscript. S. Adamčík assisted in fieldwork, provided additional samples, performed morphometric analysis for ancestral state characters, produced drawings of microscopic features, and offered comments on the manuscript. P. Matheny provided lab facilities and facilitated field excursions for collecting samples.

Abstract

Numerous clades of mushroom-forming fungi have been subject to hyperdiversification events throughout their evolutionary history. Inferring phylogenetic relationships and recognizing species with confidence within these clades can be difficult and requires the proper selection of informative loci to resolve relationships at different scales. Here I combine morphological, phylogenetic, biogeographic, and evolutionary evidence as a model for good practices in species delimitation and description of novel taxa. This study uses a multi-locus approach for species delimitation in the hyper-diverse genus Russula in order to compare the efficacy of proposed genetic markers for resolving interspecific relationships. The targeted group is composed of species morphologically placed in *Russula* subsection *Roseinae*, comprised of seven morphological species described from North America and two from Europe. Species hypotheses based on morphological differentiation and phylogenetic analyses are used to evaluate different approaches to applying the multi-species coalescent model. Biegeographic and host association history are reconstructed to determine events that may have driven diversification in this clade, including presence in refugia during the glacial cycles of the Pleistocene. Subsection *Roseinae* is found to have a Laurasian distribution with an evolutionary origin in the Appalachian Mountains of eastern North America. Coalescent approaches failed to recapitulate a reasonable species delimitation scenario given other lines of evidence. Given morphological, biogeographic, and evolutionary evidence, we delimit a total of fourteen species with molecular sampling and recommend that the subsection *Roseinae* to correspond to the core Roseinae clade and the Roseinae clade be considered a section of *Russula*.

Introduction

The genus *Russula* is one of the most speciose genera of mushroom-forming fungi with around 750-900 species accepted worldwide (Buyck and Atri, 2011; Kirk et al., 2008). Members of this group are an important part of microbial biota in forested ecosystems since they form ectomycorrhizal symbioses with select lineages of trees and shrubs, and they occur in a wide range of environments from arctic tundra to tropical rainforests. Ecological studies attempting to better understand community or ecosystem processes are increasingly acknowledging microbial communities as playing key roles in these processes (Van Der Heijden et al., 2008). As a result, characterizing these communities using next generation sequencing approaches is becoming a new

standard for modeling ecosystem functioning (Graham et al., 2016). These studies rely on accurate characterization of species as taxonomic units that are estimated based on thorough systematic assessments by multiple sources of evidence using morphological and ecological traits as phenotypic indicators for genotypic divergence or reproductive isolation (Jayasiri et al., 2015). The current state of systematics of the genus *Russula* does not satisfy these criteria and is widely based on morphology-based classification using a combination of field and microscopic characters and macro-chemicals reactions (Romagnesi, 1967).

Early systematic studies on North American species of *Russula* by C.H. Peck, C.H. Kauffman, H.C. Beardslee, G.S. Burlingham, and W.A. Murrill briefly defined more than 230 species (Buyck, 2007) and were soon followed by more elaborate morphological analyses (often based on type studies) trying to classify existing species and understand their concept (e.g. (Shaffer 1962, Shaffer 1964, Shaffer 1972, Singer 1986, Adamčík et al., 2013). These studies, however, have not been evaluated through modern molecular phylogenetics. Recent studies of *Russula* are beginning to use the internal transcribed spacer (ITS) of the ribosomal region as a universal marker for fungi (Schoch et al., 2012) to resolve species complexes and place these groups into a larger systematic context (Adamčík et al. 2016a, Adamčík et al. 2016b). This is a step forward, however, the ITS region alone is often inadequate for inferring relationships with strong support and only constitutes a single sampling of the evolutionary history (Vellinga et al., 2015). To date there have been no molecular systematic revisions of a clade of *Russula* using multiple molecular markers for the purpose of species delimitation. For resolution of relationships between larger clades it will be important to apply multiple phylogenetically informative markers, which has been lacking in previous studies (Miller and Buyck 2002, Larsson and Larsson 2003).

An advanced and potentially fruitful method for species delimitation in fungi is coalescent-based species delimitation, which allows evolutionary models that incorporate estimations of genetic drift and past population size to determine evolutionary independence (Fujita et al., 2012). This is especially true for fungi that are not able to be readily cultured, precluding mating studies to test for the biological species concept. A number of methods for applying coalescent theory to species delimitation have been developed (Kubatko et al., 2009; Liu et al., 2009; O'Meara, 2010; Yang and Rannala, 2010), and these approaches have revealed cryptic diversity in a number of lineages (Carstens and Dewey, 2010; Ruane et al., 2014; Satler et al., 2013; Singh et al., 2015). Only a few studies in mushroom-forming fungi have utilized coalescent approaches (Sanchez-Ramirez et al., 2015) and few have compared these to traditional character-based approaches (Aldrovandi et al. 2015).

The Appalachian Mountains of eastern North America (NA) are one of the oldest mountain ranges on the planet and are known for hosting a very high species diversity and endemism of different taxonomic groups (Stein et al., 2000). In fact, the mountain range is considered by some experts to be the diversity center for taxonomic groups such as plethodontid salamanders (Kozak

and Wiens, 2010), *Trillium* (Griffin and Barrett, 2004), hickory trees (Latham and Ricklefs, 1993), crayfish (Crandall and Buhay, 2008), darters (Lundberg et al., 2000), and freshwater mussels (Parmalee and Bogan, 1998). This pattern of high richness and endemicity has also been observed in mushroom-forming fungi with over 3,000 species of Basidiomycota and Ascomycota reported from the Great Smoky Mountains National Park alone (Lickey et al., 2007; Walker et al., 2005)(TENN Herbarium Database on mycoportal.org unpublished). Many of these groups that are diverse in the Appalachian mountains, including macrofungi, often exhibit a close evolutionary relationship to the biota of eastern Asia (Mueller et al., 2001; Qian and Ricklefs, 2000; Wen, 1999). Unlike plants, which share sister genera between the two regions, macrofungi are generally related as sister species (Mueller et al., 2001). It has also been demonstrated that plant genera exhibit an Arcto-Tertiary disjunction after the closing off of migration routes via the Beringia land bridge (Tiffney, 1985). In contrast, it has been hypothesized that sister species of macrofungi are a result of multiple recent migration events facilitated by similarities in climate and habitat between the two regions (Mueller et al., 2001).

Glaciation has also been proposed as having importance for the distribution and diversification of taxa in the Appalachian Mountains, as the southern part of the range has been hypothesized as a refugium during the Last Glacial Maximum (LGM) (Church et al., 2003; Hughes and Petersen, 2004; Lickey et al., 2002; Soltis et al., 2006). Also, changes in past climate may have allowed the mountains to act as a species pump, in accordance with the "montane species-pump" hypothesis, which points to heterogeneity and complexity in both topography and climatic zonation as driving speciation by way of allopatry and parapatry (Kozak and Wiens, 2010). This may also be driven by reintroduction of populations from different refugia after becoming reproductively isolated (Petit et al., 2003). This may explain apparent sympatry of sister species without invoking sympatric speciation, which is considered extremely rare in fungi (Giraud et al., 2008). In this study we seek to test whether glaciation or the Arcto-Tertiary disjunction may have affected the diversification of a clade of macrofungi that appear to be have their highest species diversity in the Appalachian Mountains.

For our study we focus on a putatively small clade of *Russula* that contains about eight species morphologically placed in *Russula* subsection *Roseinae* Singer ex Sarnari (Figure 12), defined by red or pink pilei, white spore print, mild taste, context turning bright red in the macrochemical sulfovanillin, and a pseudoparenchymatic subpellis of the pileipellis that is composed of inflated elements (Adamčík and Buyck, 2012). This group is ideal for studies in evolution and phylogeography in *Russula* because it has a well-defined delimitation. Traditionally, this group is restricted to two species in Europe (*R. velutipes* Velen. and *R. minutula* Velen.) and two in North America (*R. albida* Peck and *R. peckii* Singer). However, type studies of historic North American species have placed additional taxa (*R. rimosa* Murrill and *R. nigrescentipes* Peck) into this group (Adamčík and Buyck, 2012) and additional species (*R. rubellipes* Fatto and *R.*

pseudopeckii Fatto) have been described from the Appalachian Mountains (Fatto, 1998). Recent surveys in the Great Smoky Mountains National Park have led to the discovery of even more potential species in this group, indicating that this group may have its highest diversity in this area. Our objectives with this group are the following: 1) Determine the number of species and distribution of subsection *Roseinae* in eastern NA; 2) test molecular markers to determine the most effective ones for resolving relationships within *Russula* at the species and subsection level; 3) reconstruct morphological and ecological characters to determine which character are conserved and may be important for species delimitation; and 4) test whether glaciation in the Pleistocene or an Arcto-Tertiary disjunction has driven this group's diversification.

Materials and Methods

Taxon Sampling

Specimens that morphologically match subsection *Roseinae* and its sister group *Russula* subsect. *Lilaceinae* (Melzer & Zvára) Jul. Schäff. (Looney et al., 2016) were collected throughout five field seasons in the United States and Europe. Members of subsection *Lilaceinae* were included for testing systematic placement of putative new members of subsection *Roseinae* and are traditionally recognized by lacking a red color change of their stipe tissue in sulfovanillin and having a subpellis of filamentous, narrow hyphae. Efforts were made to sample species placed in subsection *Roseinae* based on monographical works (Sarnari, 1998; Singer, 1986) and type studies (Adamčík and Buyck, 2012).

Sporocarps were collected from forested sites in the eastern U.S. centered on New York, the region of type localities for species described by C.H. Peck, G.S. Burlingham, and R.M. Fatto, Mississippi and Florida, for species described by W.A. Murrill, and Tennessee and North Carolina, where we might expect northern and southern species to overlap. The two known European species of subsection *Roseinae* and members of subsection *Lilaceinae* were collected from central Europe (Slovakia). A number of species have been described from temperate Asia as putative members of subsection *Roseinae*, including *R. dhakuriana* K. Das, J.R. Sharma & S.L. Mill., *R. sharmae* K. Das, Atri & Buyck, *R. minutula* var. *robusta* Saini, Atri & Singer, and *R. rosea* Pers. sensu Romagnesi , which has been reported from Japan (Das et al., 2013, 2006; Hongo, 1960; Saini et al., 1982). To our knowledge, no species are known or described in subsection *Roseinae* from Africa, South America, southeast Asia or Australasia. To test relationships of extraterritorial species, all sequences from GenBank that "blasted" within 95% identity of sampled members for both datasets were included in alignments to increase global taxon sampling. All field collections were described and photographed in the fresh condition with color designations given by Kornerup and Wanscher (1967).

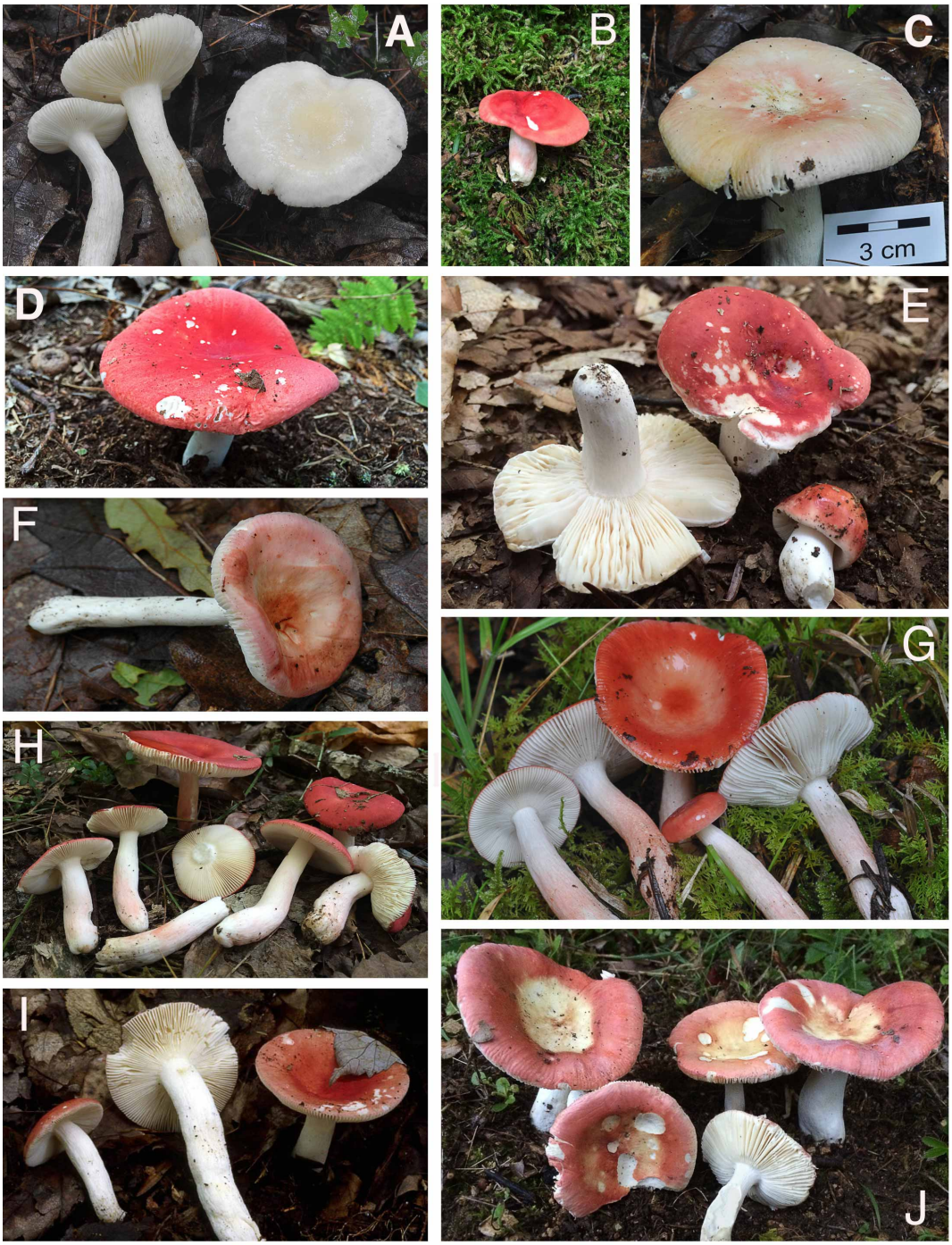


Figure 12. North American and European members of *Russula* subsect. *Roseinae*: A) *Russula albida* (photo Per Marstad); B) *Russula leporina* nom. prov.; C) *Russula magnorosea* nom. prov. (photo Per Marstad); D) *Russula* cf. *magnorosea*; E) *Russula peckii*; F) *Russula minutula* (photo Per Marstad); G) *Russula peckii* (photo Per Marstad); H) *Russula niveopersonata* nom. prov.; I) *Russula austrorubellipes* nom. prov. (photo Steve Trudell); J) *Russula velutipes*

All dried collections are deposited at in the herbarium of the University of Tennessee (TENN) and the Slovak Academy of Sciences (SAV) (herbarium abbreviations per Thiers [continuously updated]). Additional historical collections including available type material were received from herbaria and examined on loan from the Florida Museum of Natural History (FLAS), New York Botanical Gardens (NY), New York State Museum (NYS), and University of Michigan (MICH).

Molecular Sampling

Genomic DNA was extracted using an E.Z.N.A. High Performance Fungal DNA Kit for historical collections and an Extraction Solution-based method for fresh collections. For historical collections, a pie wedge of the pileus weighing approximately 20 mg was ground into a fine powder using a mortar and pestle in liquid nitrogen and a pinch of sterile sand. Buffer was added and samples were further ground using centrifugation at 17,000 xG followed by grinding with a micropestal. Two microliters of beta-mercaptoethanol was added to samples and left to incubate at 65°C for 24 hours. Other modifications to the prescribed protocol follow Looney (2015). The Extraction Solution-based protocol for fresh or recently dried collections started placing one partial piece of lamellar tissue into 100 microliters of filter-sterilized Extraction Buffer (10 mL of 1M Tris stock, 1.86 g KCl, 0.37 g EDTA, and 80 mL DI H_2O) and macerated using a fresh toothpick. Samples were then incubated at room temperature for at least 24 hours and then incubated at 90°C for 10 minutes. Finally, 100 microliters of a shaken and filter-sterilized Dilution Solution (3 g BSA, DI H_2O added until 100 mL solution) was added. DNA solutions were then diluted to a 1:10 ration with double distilled water and 2 microliters were added to the amplification master mix. Polymerase chain reaction (PCR), gel electrophoresis, PCR clean-up, and sequencing reaction protocols follow that of Birkebak et al. (2013). Sequencing was performed on an ABI 3730 capillary electrophoresis instrument at the UT Genomics Core. Five nuclear loci were sequenced: ITS using ITS1F – ITS4 (White et al. 1990), *rpb1* using gAf – fCr (Matheny et al. 2002), *rpb2* using b6F – b7.1R (Matheny 2005), *tef1* using EF1-983F – EF1-2218R (Rehner and Buckley 2005), and *mcm7* using mcm7-709for – mcm7-1348rev (Schmitt et al. 2009). Sequencing products were assembled and edited using Sequencher 5.1 (Gene Codes, Ann Arbor, MI, USA). Outgroup sequences were retrieved from Mycocosm from the genome of *Russula rugulosa* BPL 654 v1.0 sequenced by the Joint Genome Institute (Walnut Creek, CA). All sequences have been deposited in GenBank (accession Nos. KY509431-KY509517 [ITS]; KY701434-KY701467 [*rpb1*]; KY701345-KY701392 [*rpb2*]; KY701393-KY701433 [*tef1*]; KY701468-KY701513 [*mcm7*]).

Phylogenetic Inference and Time Calibration

Two datasets were constructed for different analyses. The Incrustatula dataset, which includes all samples determined as either subsection *Lilaceinae* or closely related, was constructed for inferring broad phylogenetic relationships and assessing phylogenetic markers. The Roseinae dataset include those clades that were inferred as part of a Roseinae clade including residual clades inferred as monophyletic with the core Roseinae clade of known members of subsection *Roseinae*. This Roseinae dataset is used for coalescent species delimitation approaches, phylogeographic reconstruction, and ancestral character reconstruction. Single gene alignments were constructed using MAFFT ver. 7 (Katoh and Toh 2008) and then manually aligned in AliView ver. 1.18 (Larsson 2014). Hyper-variable or conserved regions of the different loci were excluded from phylogenetic analysis, including the 5.8S gene of the ITS region and a coding region of *rpb2* composed of mostly repeating codons of variable length. Individual gene trees were inferred using raxmlGUI (Stamatakis et al. 2008, Silvestro and Michalak 2012).

Individual gene trees and an ultrametric chronogram based on the concatenated Incrustatula dataset were inferred in BEAST 2 ver. 2.4.2 (Bouckaert et al., 2014). For the concatenated dataset, genes were first partitioned by introns and codons and analyzed by PartitionFinder v. 1.1.1 (Lanfear et al., 2012) to detect the best partitioning scheme and evolutionary models implemented in BEAST. The suggested partitioned matrices were imported into BEAUTi 2 ver. 2.4.2 with site models unlinked and clock models and tree linked for the concatenated dataset. Suggested models were set for partitioned matrices with substitution rates estimated with a fixed mean. A relaxed molecular clock with a log normal distribution was selected and the tree was modeled under the birth-death prior with tertiary calibrations. Calibrations were taken from Looney et al. (2016) for the Crown clade node at 15.15 [95% posterior density (HPD) 11.3-19.6] million years (MY), Incrustatula clade node at 14.04 [HPD 10.3-18.4] MY, and Lilaceinae clade node at 6.91 [HPD 4.24-10.2] MY. For GTR models, transition and transversion rates were modeled under a Poisson distribution. Three independent Markov chain Monte Carlo (MCMC) were run for 50 million generations, sampling states/trees every 1 000 generations. Log files for all three chains were jointly inspected in Tracer v1.6 (Rambaut et al., 2015) to ensure estimated sample size (ESS) values reached above 200 and that all three runs had converged. The three runs were then combined in LogCombiner 2.4.2 using a burnin of 10% for each chain to drop pre-convergent values for a final total of 135 000 trees. A consensus tree was constructed in TreeAnnotator v2.4.2 as a maximum clade credibility (MCC) tree with ages given as mean node heights. Trees were inspected in FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Gene markers used in the phylogenetic analyses of the Incrustatula clade dataset and resulting tree were used to generate phylogenetic informativeness profiles using PhyDesign (López-Giráldez and Townsend, 2011).

Coalescent species delimitation approaches

To detect species, two approaches were used to apply the multispecies coalescent model. The first approach was implemented in the program BP&P v3.1 (Yang, 2015), to compare different models of species delimitation and species trees in a Bayesian framework that accounts for incomplete lineage sorting due to ancestral polymorphism (Rannala and Yang, 2013, 2003, Yang and Rannala, 2010). We used the approach of (Yang and Rannala, 2014) for unguided species delimitation using the reversible-jump Markov chain Monte Carlo (MCMC) algorithm (Yang and Rannala, 2010: algorithm 1) and assigned equal probabilities to the rooted species trees as a species model prior. For population size parameters (*qs*) we assigned the gamma prior G(2, 1000), with a mean of 2/2000 = 0.001 as these have worked well for similar groups of diverse ECM species clades. The divergence time at the root of the species tree (τ) was assigned the gamma prior G(2, 1000) and all other divergence time parameters were assigned the Dirichlet prior (Yand and Rannala, 2010: equation 2). The analyses were run twice to confirm the consistency between runs.

A likelihood approach to the multispecies coalescent model was implemented in the program SpedeSTEM v. 2.0 (Ence and Carstens, 2011). SpedeSTEM takes a group partition scheme and a number of gene trees and uses an information-theoretic approach to calculate all hierarchical permutations within species groupings. To do this it uses the species tree estimation using maximum likelihood (STEM) package (Kubatko et al., 2009) to infer species trees and calculates the likelihood of the species tree given the provided gene trees and these models of lineage composition are compared using the Akaike Information Criteria (AIC). A discovery run was performed with a theta value of 0.123 and a beta value of 0.005. All gene trees were set with a scaling of 1.0. A validation analysis was run for 200 permutations.

Species Tree Estimation

For all ancestral state reconstructions (ASR) and diversification analyses a species tree for the Roseinae dataset was inferred using *BEAST in BEAST 2 (Bouckaert et al., 2014). Species tree clades were guided by results from BP&P based on the highest supported species model. Population size for species trees was modeled as changing linearly over branches with the sum of the population size of the two species as equal to the population size of the ancestral species at the time of the split, with the root constrained as a constant population size (Heled and Drummond, 2010). Models were again estimated in PartitionFinder for gene loci, and the TrNef +G model was selected for ITS, *rpb1*, *rpb2*, and *tef1* and the K2P +I model for *mcm7*. The multi-species coalescent model analysis was time-calibrated using quaternary calibrations based on mean estimate and confidence interval of the root node for the Incrustatula clade and the root of the Roseinae clade from the Incrustatula dataset time reconstruction. Three

independent Markov Chain Monte-Carlo (MCMC) analyses were run for one hundred million generations, storing and logging every five thousand trees.

Phylogeographic and host analyses

Ancestral geographic states for the Roseinae clade were inferred using the R package (R Core Team, 2015) 'BioGeoBEARS' (Matzke, 2013). The package allows for model testing between popular biogeographic models including DEC, DIVA, and BAYAREA with the inclusion of an additional parameter called the jump (*j*) parameter, which simulates founder-event speciation events. The package also uses probabilistic inference of historical biogeography using a ML estimation of parameters with the quasi-Newton method with box constraints, and then it calculates the ancestral states under the globally optimum model (Matzke, 2014). Geographic states for North American species were coded based on their recovered ranges and whether these ranges overlap with areas that were glaciated during the Last Glacial Maximum (LGM), restricted to the southern Appalachian mountains, or occurring in coastal plains (Pielou, 1991).

Ancestral State Reconstruction

A number of macro- and microscopic features were observed for each species and coded as tip character states. Macroscopic characters were coded as consensus measurements taken from fresh collections as well as photographs of the collections *in situ* for verification. Microscopic characters were measured from representative collections of each species, using the type species where possible. Microscopic structures were examined from desiccated herbarium specimens in Congo red solution with ammonia after a short treatment in aqueous 10% KOH. They were viewed under an Olympus CX-41 light microscope with an oil immersion lens at a magnification of 1000×. Drawings were made with a camera lucida using an Olympus U-DA drawing attachment at a projection scale of 2000×. Basidiospores micrographs were produced by an Artray Artcam 300MI camera and measured by Quick Micro Photo (version 2.1) software. Spores were observed in Melzer's reagent. Enlarged scanned pictures of spores were used for measuring with an accuracy of 0.1 µm and for making line drawings. The statistical values in the analysis are based on 20 or 30 measurements.

Ancestral state reconstruction performed on a set of 1 000 trees randomly sampled from the posterior distribution of species trees in Mesquite v. 2.74 (Maddison and Maddison, 2001). Character history was traced across the tree topologies using the Trace Character Over Trees function using a ML approach with stored probability models. Significant support for character states was assessed using a 2.0 cut-off difference in log-likelihoods between states. The character state values were transformed to only display differences in states and were mapped onto the majority-rule consensus tree.

Results

Phylogeographic reconstruction, dating, and phylogenetic informativeness of genes

A total of 252 sequences were generated from 86 samples across the Incrustatula clade. No well-supported incongruencies between gene trees were detected, so a concatenated gene matrix was used for all phylogenies. Four major clades of the Incrustatula clade dataset were recovered with good support from at least one inference method (Figure 13). Overall support for the 5-gene phylogeny was high as 45 of 51 major nodes received good support of either 70 bootstrap support or a posterior probability of 0.95. The crown age for the Incrustatula clade has a mean age of 13.9 [HPD 10.6-17.3] MY with the Lilaceinae clade splitting off from the Roseinae clade (including the Albida clade, the Residual Roseinae clade, and the Core Roseinae clade) 12.86 [HPD 9.9-15.9] MY ago. The crown age of the Roseinae clade is 12.2 [HPD 9.2-15.2] MY, while the crown age for the Lilaceinae clade is younger at 8.49 [HPD 6.6-11.1] MY old. The crown age for the Core Roseinae clade is slightly younger at 6.96 [HPD 4.9-9.2] MY old. According to the phylogenetic infomativeness profile, *tef1* was the best gene marker for resolving clades at least 4.3 MY old, which was then replaced by *mcm7* as the best marker for younger clades. An uptick of informativeness at around 0.3 MY likely indicates the initial threshold for interspecific divergence, which can be detected in both *tef1* and *mcm7*. The *rpb1* locus was the second most informative gene for clades at least 10.5 MY old, whereas the ITS barcode marker's performance was only average compared to the other markers and *rpb2* was the least phylogenetically informative locus.

Species delimitation in the Roseinae clade

A total of 176 sequences were generated from 55 samples across the Roseinae clade, including the type collections of *R. rubellipes* Fatto, *R. pseudopeckii* Fatto, and *R. purpureomaculata* Shaffer. The type collections for *R. peckii* Singer, *R. nigrescentipes* Peck, and *R. rimosa* Murrill were extracted but failed to amplify with PCR. No well-supported incongruencies between gene trees were detected, so a concatenated gene matrix was used for all phylogenies. The ML phylogeny recovered 28 well-supported clades with at least 70 bootstrap support (Figure 14).

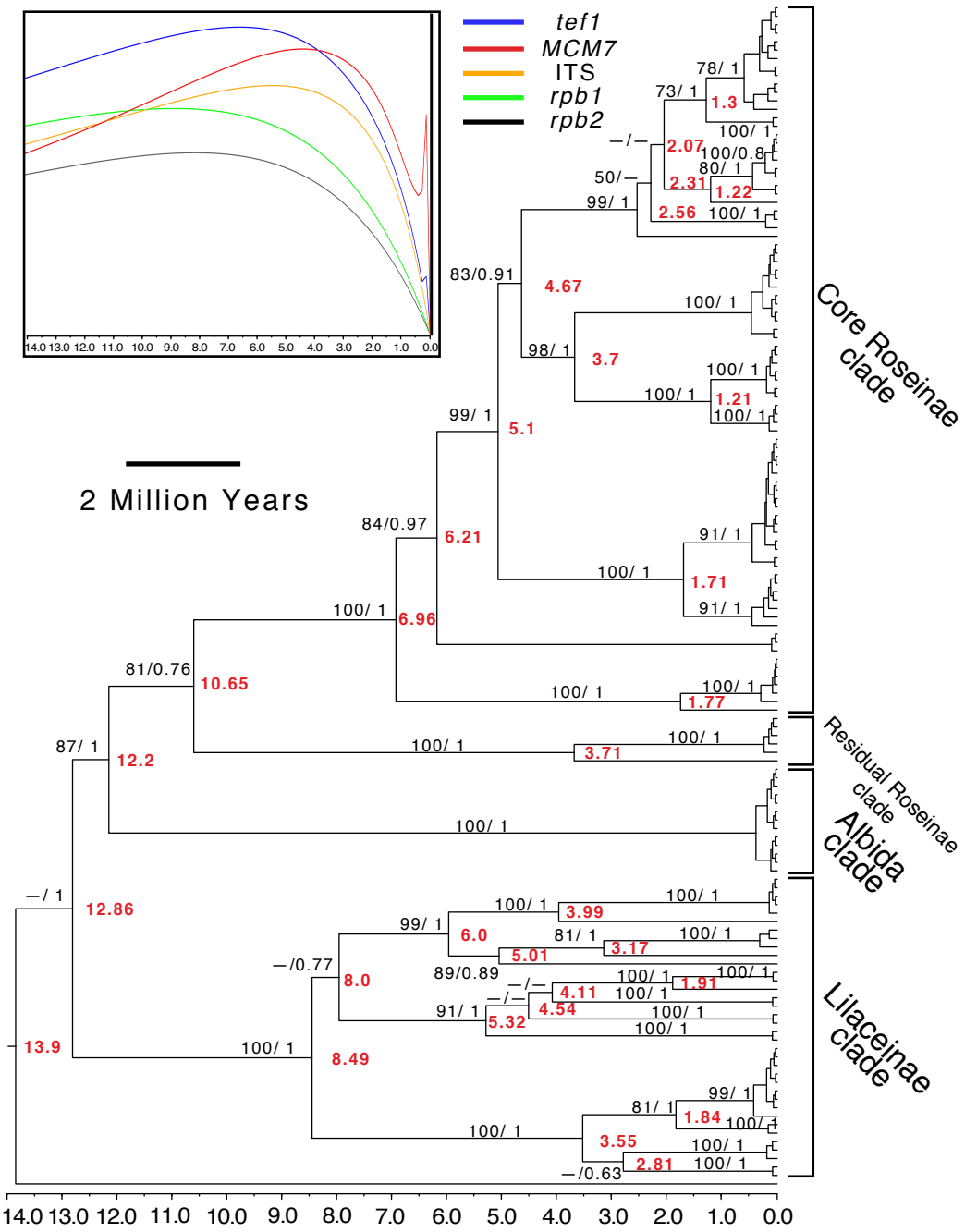


Figure 13. Ultrametric chronogram of the Incrustatula clade inferred in BEAST. Bootstrap support is reported along branches (black) from ML reconstruction in RAxML and is followed by posterior probabilities. Mean estimations of ages in million years are reported at nodes (red). Hyphens are used when bootstrap support is below 50%, posterior probability is below 0.5, or if the node was not recovered by either method. Inset shows phylogenetic informativeness inferred in PhyDesign of the five nuclear markers used for inferring the phylogeny.

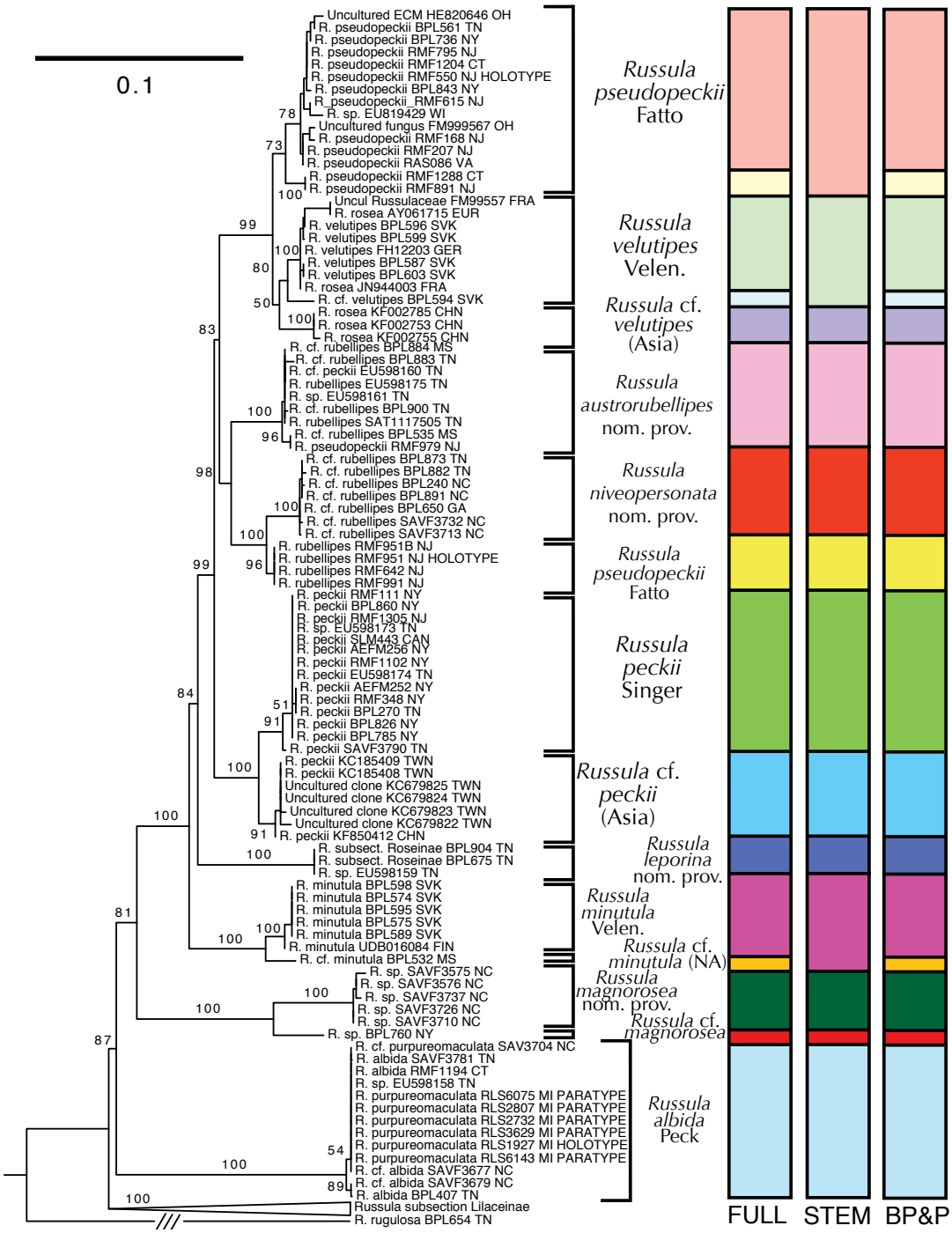


Figure 14. Maximum Likelihood phylogeny of the Roseinae clade inferred in RaXML with bootstrap support reported along branches. Collections sampled are listed with collector number and locality. Best supported species models for SpedeSTEM and BP&P are given in colored bars with the full species model (FULL) given first. Species epithets are given for well-resolved clades recovered from both coalescent models, with questionable taxa given *confer* status.

To test the phylogenetic recovery of putative species clades, two different coalescent approaches to species delimitation were applied to the Roseinae clade dataset. A total of 16 potential species units were evaluated in the BP&P analysis as the full model. The full set of proposed species were recovered as the highest supported species model in the BP&P analysis (0.92 pp, 16-species model). A more conservative estimate of species was recovered in SpedeSTEM, which found the highest model likelihood for a 13-species model that excluded *R. pseudopeckii* NA clade B, *R. velutipes* Europe clade B, and the *R.* cf. *minutula* NA clade. For clades that are well supported and recovered in both coalescent approaches, provisional names were assigned.

Sequences of *R. purpureomaculata* were recovered in the same clade as specimens identified as *R. albida* with high support. After morphological comparisons with type collections of North American members of subsection *Roseinae* it was determined that *R. praeumbonata* Burl. and *R. rimosa* Murrill were not recovered in our sampling and remain unsampled species. Also based on morphological comparisons and sequences from type material, six described species from subsection *Roseinae* were recovered, including *R. peckii*, *R. albida*, *R. pseudopeckii*, *R. rubellipes*, and the two European representatives, *R. minutula* and *R. velutipes*. The type collections for *R. peckii* and *R. nigrescentipes* were both determined to be mixed collections, so given the popular concept of *R. peckii* and the obscurity and confusion associated with *R. nigrescentipes*, we adopt one and exclude the other species. A total of nine terminal clades with good support were recovered to which no published species name could be attributed, including two clades from Asia.

Ancestral range and host reconstruction of the Roseinae clade

The best model based on model likelihood under the 5-state geographical analysis was DEC+J, which estimates a jumping parameter simulating founder events along with parameters for dispersal, extinction, sympatry as a subset or in the narrow sense, and vicariance in the narrow sense (Figure 15). The ancestral area of subsection *Roseinae* is inferred as most likely eastern NA with jump migration to the southern Appalachian Mountains. Multiple sympatric diversification events are inferred in the southern Appalachian Mountains with one notable jump dispersal event to Europe around 6.7 [HPD 5.1-8.5] MY ago and two jumps to Asia at around 4.6 [HPD 3.4-5.8] and 1.5 [HPD 0.4-2.7] MY ago. Part of the early Asian lineage spread across Eurasia around 2.2 [HPD 1.5-3.0] MY ago. Two jump dispersal events from Europe/Eurasia to NA occurred around the same time at 2.0 [HPD 1.3-2.6] and 1.7 [HPD 1.0-2.4] MY ago.

The best model for the 2-state host analysis was the DIVALIKE model, which estimates dispersal, extinction rate, sympatry in the narrow sense, and vicariance in a narrow and widespread sense (Figure 15). The ancestral host of subsection *Roseinae* is inferred as most likely an angiosperm. Four host expansion events were inferred, with a host specialization event on Pinaceae occurring at 5.0 [HPD 3.7-6.3] MY ago.

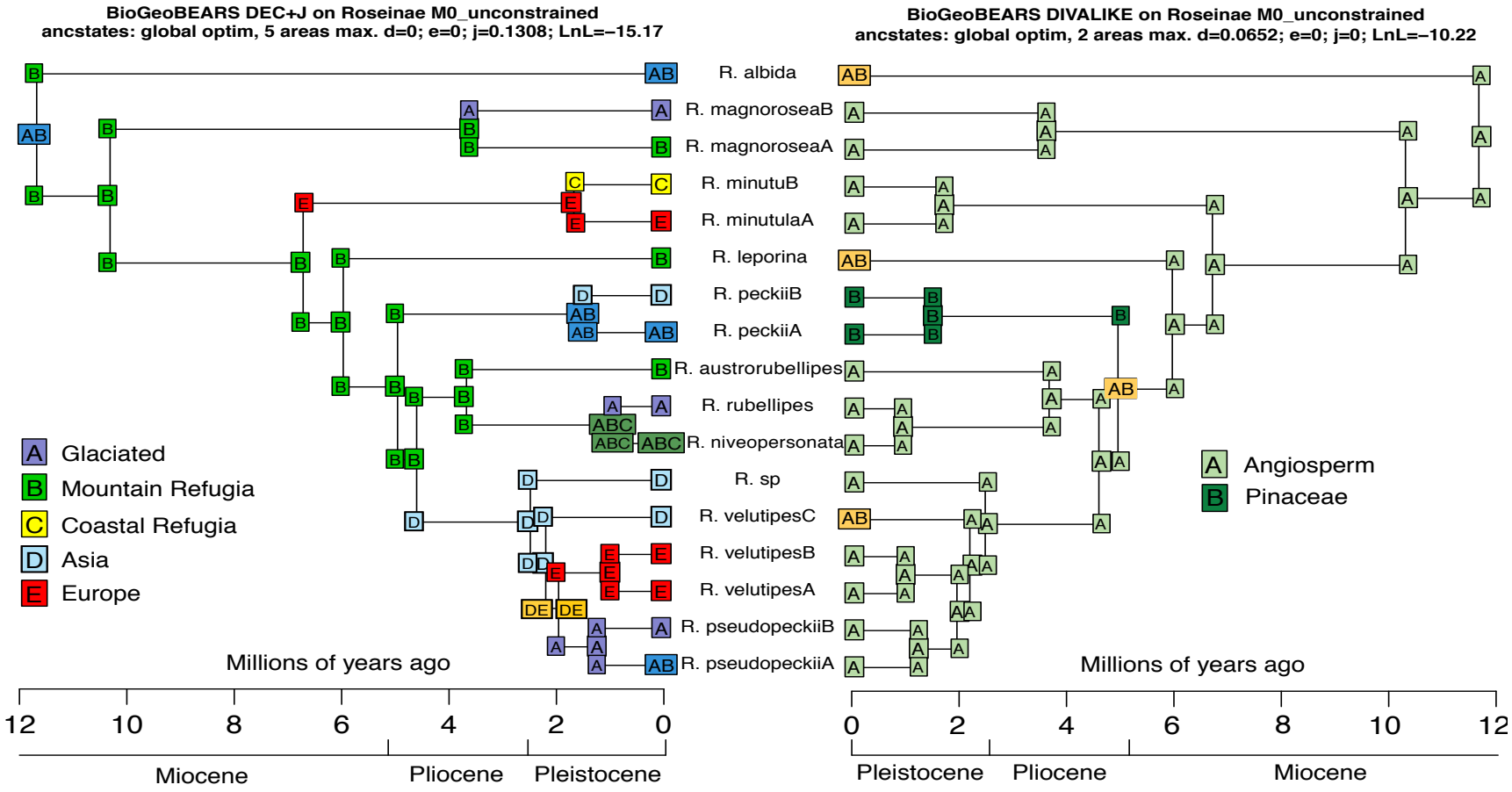


Figure 15. Ancestral area and host reconstruction of the best-supported model estimated in the R package 'BioGeoBEARS'

Ancestral reconstruction of traits in the Roseinae Clade

Twelve microscopic and macroscopic traits were deemed potentially important for species delimitation and were characterized using available individual collections (Figure 16). The ancestor of the Roseinae clade as well as the clade including the Core Roseinae clade and Residual Roseinae clade were recovered as ambiguous for all character states (Figure 17). The ancestor of the Core Roseinae clade was resolved with significant support as having a pink flush on the stipe, intermediate pileus width, red pileus color, a fruity or nutty odor, frequent branching of terminations in pileipellis, and cylindrical terminal cells of hyphae in the pileipellis near the pileus margin. The ancestor of the "*R. magnorosea*" clade encountered an expansion of pileus size and transition to attenuated terminal cells in the pileipellis. In contrast, the ancestor of the *R. minutula* clade saw a shift to a small, pink pileus, frequent branching of hyphal terminations in the pileipelllis, narrow, cylindrical terminal cells of hyphae in the pileipellis near the pileus margin, and narrow primordial hyphae, which are terminal cells with acid-resistent crystals that stain in the reagent carbolfuschin. The ancestor of the *R. rubellipes* clade, including "*R. nivopersonata*", saw a shift to a mostly pink/red stipe, a loss of odor, loss of branching in the pileipellis, an intermediate spore ornamentation and width, a positive reaction to guaiac in the gills, pleurocystidia with heteromorphous to dispersed contents, and wide terminal cells. A shift was detected in the pileipellis type, stipe color, odor, terminal cell branching, spore ornamentation height, pleurocystidia contents, spore, and wide terminal cells in the pileipellis forming an epithelium at least near the pileus center.

Discussion

Systematics of the Incrustatula clade

The future challenge of systematics in diverse groups such as Russulaceae is to resolve clades that correspond to infrageneric ranks. This is an essential organizational tool for binning species into evolutionary groups and facilitating biodiversity studies. Though barcoding species with the ITS region is useful for stabilizing species, no single locus is sufficient to resolve these higher-level relationships. This is demonstrated in our study and our attempt to resolve a *Russula* clade at the subsection level. No single locus resolved the final topology of the Incrustatula clade with high bootstrap support (Figure 18; Appendix). Only *mcm7* resolved the Roseinae clade with high support and only *rpb2* resolved the final topology without high support. Only when genes were concatenated did a well-supported relationship between the four major clades become resolved, indicating the power and importance of a multi-locus approach. With the advent of next generation sequencing technologies, obtaining multi-gene datasets is becoming much easier and less costly.

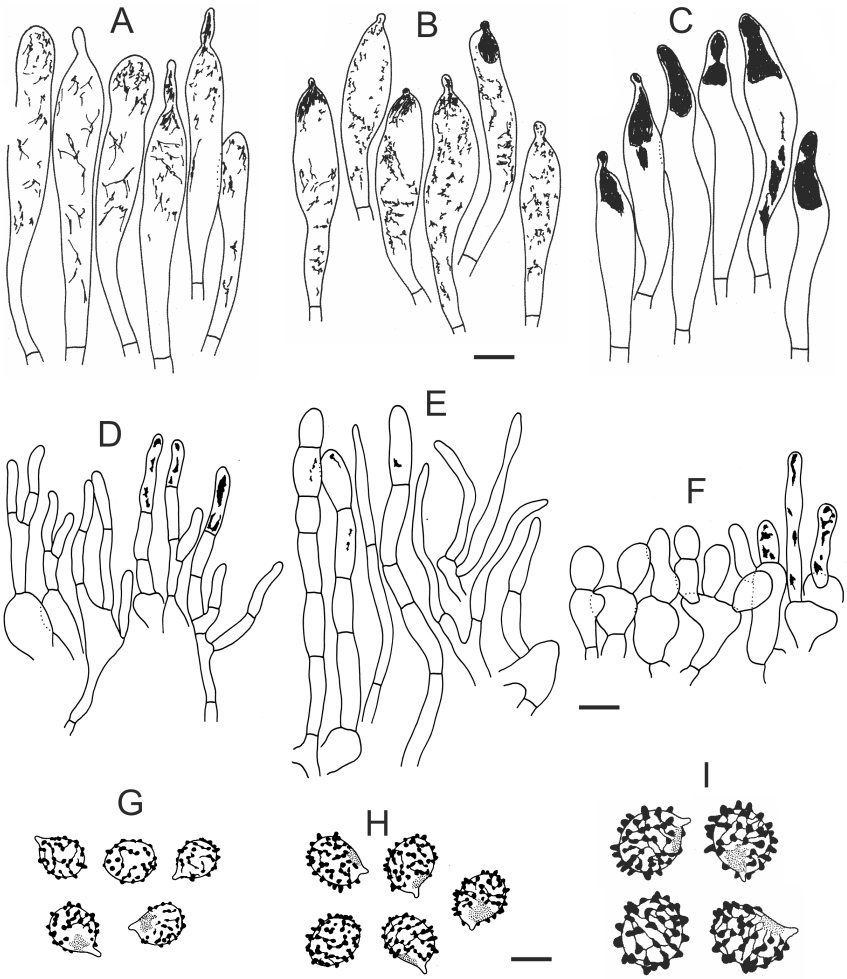


Figure 16. Microscopic characters used in ancestral trait reconstruction of the Roseinae clade: A) Pleurocystidia with dispersed elements (*R. rubellipes*, NY00253510, holotype); B) Pleurocystidia with partly dispersed elements and a refringent body (*R. pseudopeckii*, NY00253511, holotype); C) Pleurocystidia with large refringent bodies only in apex (*R. peckii*, NYS f3630.7, holotype); D) Dense trichoderm pileipellis composed of narrow, cylindrical, frequently branched hyphal terminations and narrow primordial hyphae (*R. pseudopeckii*, NY00253511, holotype); E) Loose trichoderm pileipellis composed of moderately wide, attenuated, occasionally branched hyphal terminations and wide, short-celled primordial hyphae (*R.* sp., SAV F-3576); F) Epithelium pileipellis of inflated, ellipsoid, rarely branched elements and moderately wide primordial haphae (*R. rubellipes*, NY00253510, holotype) ; G) Spores exhibiting short width and ornamentation height (*R. pseudopeckii*, NY00253511, holotype); H) Spores exhibiting medium width and ornamentation height (*R. rubellipes*, NY00253510, holotype); and I) Spores exhibiting tall width and ornamentation height (*R. praeumbonata*, NY00760516, syntype). Scale bar equals 10 µm but only 5 µm for spores.

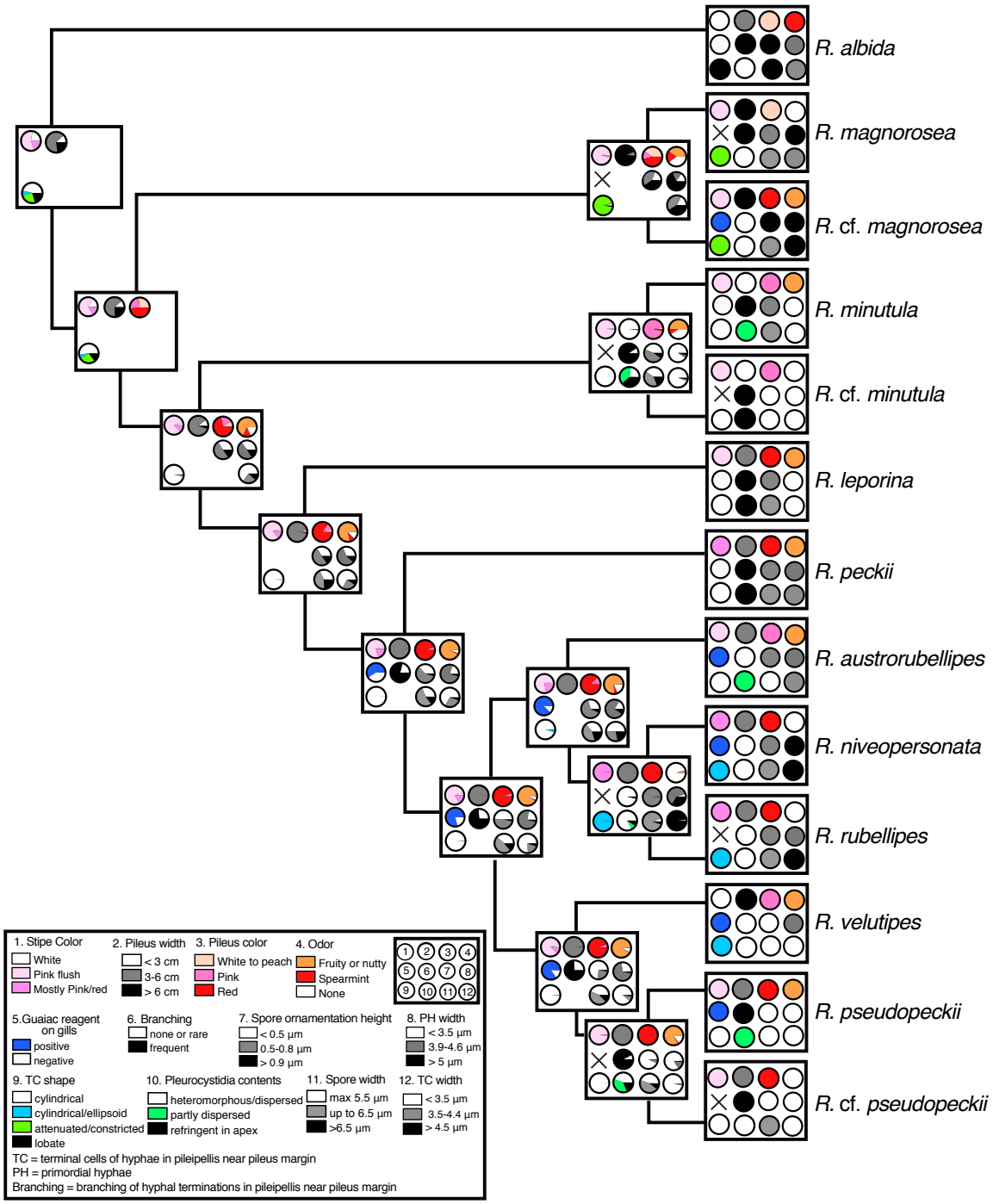


Figure 17. Ancestral character reconstruction using ML in Mesquite for both macro- and microcharacters. Characters that were not measured for given species are marked as an X. Reconstructions where no character was resolved as less than 25% significant were excluded.

However, identification of suitable gene markers is important for studies in countries without research support where researchers must balance gene sampling with taxon sampling. It is therefore important that the best genes are identified for multi-gene analyses to resolve different levels of diversity. Based on the results of our study, for the genus *Russula* we recommend the use of *tef1* for resolving clades at the subsection level or above, whereas *mcm7* appears to be ideal for resolving species. Both of these genes should be used in concordance with the barcode ITS region that does an intermediate job at resolving relationships at both scales.

Since *R. albida* has been considered a member of subsection *Roseinae*, we propose the Roseinae clade to correspond to this subsection, which will need to be emended with a new morphological diagnosis. The recovery of the Residual Roseinae clade as nested in the Roseinae clade makes this necessary, as it is made up of collections that do not match the morphological diagnosis of subsection *Roseinae* and potentially constitute a novel clade of under-explored species. Another well-resolved clade is the Lilaceinae clade, which includes species such as *R. lilacea* Quél., *R. subtilis* Burl., *R. corallina* Burl., *R. emeticicolor* (Jul. Schäff.) Singer, and *R. zvarae* Velen., all species traditionally placed in subsection *Lilaceinae*. Both of these groups are included in *Russula* subgenus *Incrustatula* Romagn., which is typified by *R. lilacea*. Increased sampling and studies of type specimens in the Lilaceinae clade should lead to at least one other well-resolved subclade. In our opinion, since both Lilaceinae and Roseinae clades are nested in the crown clade of *Russula* (Looney et al. 2016), these should be classified in two ranks that best fit the classification model proposed by Sarnari (1998). Following this classification, both groups together should be placed in subgenus *Incrustatula* with the type species *R. lilacea*. The concept of the subgenus presented by Sarnari covers all species with incrusted primordial hyphae, but this and previous phylogenetic studies (Looney et al. 2016) demonstrated that dark yellow-spored species, e.g. *Russula* subsect. *Amethystinae* (Romagn.) Bon, likely do not belong to this lineage. This study suggests that the subgenus *Incrustatula* includes only species with pale white or cream spore prints, mild taste, and primordial hyphae.

We think that the appropriate rank for the Roseinae clade is the section level because of the recovery of a morphologically divergent Residual Roseinae clade and Albida clade. Each subclade of the Roseinae clade should constitute a subsection. The proposed section should be recognized by having a pseudoparenchymatic subpellis composed of inflated elements and either small or obtuse and short-celled primordial hyphae. *Russula albida* representing the whole Albida clade differs from all studied members of Roseinae core clade by larger spores with more prominent ornamentation and a corraloid trichoderm pileipellis of short, frequently lobate or branched elements. The Residual Roseinae clade represented by two species clades potentially constitute a novel clade of under-explored species defined by the absence of red staining of the context in sulfovanillin, long attenuated terminal cells of hyphae in the pileipellis and broad, obtuse primordial hyphae often composed of chains of ellipsoid cells.

Species delimitation in the Roseinae clade

For the four criteria on which we based our species delimitation on (phylogeny, coalescence, geography, and morphology) no single delimitation scheme was supported by all methods. The BP&P method for multispecies coalescence was the most sensitive method, resolving all proposed species as distinct. This scenario is appealing as it was able to distinguish all species that have support from the three other species criteria; however, we detect some potential over-splitting in the *R. pseudopeckii* and *R. velutipes* clades. These segregate clades are very similar morphologically to their sister clades, there is not much phylogenetic divergence between them, and they are found in the same general location/region. In contrast, the SpedeSTEM approach was more conservative and did not resolve these questionable splits as species; however, it did lump a collection from Mississippi with the European species *R. minutula*. This would be the only transcontinental species recovered in this study and there is morphological evidence that this should represent its own species. Based on coalescent analysis in conjunction with geography and morphology (when available), we recognize here fourteen species in this clade (Fig. 3).

Phylogeography and host association of the Roseinae clade

Phylogeographical analysis of the Roseinae clade has recovered the Appalachian Mountains as the ancestral origin of this group, which, as far as we know, is the first evidence for ancient endemism of a group of fungi from this area. Only a few studies have looked at Plethodontid salamanders were thought to have an Appalachian origin, however, recent phylogeographical assessments have refuted this claim (Vieites et al., 2007). One area where sampling is lacking is central Mexico, which is known as another hotspot for macrofungal diversity and we lack sampling from reported species from eastern Asia (Das et al., 2013, 2006; Hongo, 1960; Saini et al., 1982; Sanchez-Ramirez et al., 2015). We also find evidence that the mountain refugia of the southern Appalachian Mountains have been a place of many apparent sympatric speciation events, with at least four diversification events. It is likely that these species were spatially isolated at either different elevations or different refugia within the mountains, though actual sympatric speciation cannot be rejected. Rare jump dispersal events have been important in this lineage's diversification, which has allowed it to spread both east and west to Europe and Asia respectively. Though rare, these dispersal events are important to the spread and diversification of ectomycorrhizal fungi, like in the *Cortinarius violaceus* group, which saw a jump dispersal from Australasia to South America (Harrower et al., 2015).

A number of phylogeographic patterns have been detected in plants and animals in the southern Appalachian Mountains, which have been attributed to Pleistocene glaciation events as well as earlier events of the Pliocene (Soltis et al., 2006). The Appalachian Mountain discontinuity refers to sister species divergent across the ridgeline of the mountains, which does not match any of the

patterns we have detected for the Roseinae clade. Differences in watersheds also do not seem to coincide with patterns in this clade. Soltis et al. (2006) proposed pseudocongruence as a possible reason for many of these conflicting patterns in distributions between different taxa, and perhaps it is a combination of many overlapping factors occurring at different times that results in these complex distributions.

Geographic isolation of the populations due to the appearance and disappearance of the Bering land bridge may explain an apparent Arcto-Tertiary disjunction of both European and Asian taxa with eastern NA taxa that occurred 6 MY and 4 MY ago respectively (Hopkins et al., 1967). However, the more recent shift to Asia and multiple shifts back from Europe cannot be explained by spatial interaction of continents, giving further evidence for the necessity of long-distance dispersal. Glaciation of the northern Appalachian Mountains began towards the beginning of the Pleistocene about 2.6 MY ago, which according to our dated phylogeny, post-dates all diversification events occurring in North America except for the split between *R. rubellipes* and *R. niveopersonata*.

If glaciation has not driven diversification in the Roseinae clade, then the apparent sympatry of species like *R. niveopersonata* with *R. austrorubellipes* and *R. rubellipes* is difficult to explain. The admixture of endemic species in the southern Appalachian Mountains with a more widespread sister species has been documented in a number of fungal genera, like *Hygrocybe*, *Armillaria*, *Amanita*, and *Sparassis* (Hughes et al., 2014, 2013). In the genus *Auricularia* we see a primarily southeastern U.S. species as closely related to a northeastern species that are differentiated by their substrate, one on hardwood and the other on the wood of conifers (Looney et al., 2013). An endemic genus of mushroom-forming members of Tricholomataceae, called *Albomagister*, has been described from the southern Appalachian Mountains with at least three species that are seemingly sympatric (Sánchez-García et al., 2014). The southern Appalachian Mountains has been highlighted as an area of hybridization for agaric fungi (Hughes et al., 2013). It is perhaps possible that ancient hybridization events have resulted in speciation of agarics in the southern Appalachian Mountains, though we would want to look at heterozygosity across entire genomes to determine this.

Also difficult to explain is the lack of species recovered from the west coast of the U.S. Mycologists, including the authors here, have done much sampling in the Pacific Northwest region however, no species of subsection *Roseinae* have been recovered from this area. It is possible that a species may be undiscovered in the rich mycota of California that may extend its range down into Mexico, but it is highly unlikely that there is an unrecovered species from the northwest even through Alaska, given a high sampling effort of that region. Only 49 of 332 native North American *Russula* species have been described based on the material from the Pacific Coast, and most of these species are centered in California (Buyck et al., 2015). This distribution does indeed imitate the same Arcto-Tertiary disjunction we see in many plant groups.

Life history is important to consider when examining phylogeographic patterns, and here we are looking at a group of ectomycorrhizal fungi that require a plant host. Given that the ancestral host of subsection *Roseinae* was likely deciduous, the distribution of species currently found in the southern Appalachian Mountains must have had their ranges shifted southward into the Gulf Coast region, Florida, and maybe all the way to Mexico during the maximum extent of Pleistocene glaciation. This suggests that, given extinction was not to play a large role in species distributions, *Russula* species are able to track the range shifts of their hosts during phases of climate change. Four host expansion events occurred in the group with one complete switch to Pinaceae association. In an analysis of host effects on diversification of *Russula*, Looney et al. (2016) found that diversification rates were higher with Pinaceae associates and host generalists. If this is the case, perhaps this shift in diversification seen in younger clades like *R. peckii* that may be shifting due to the cooling climates of the last 5 MY.

Morphological trait evolution in the Roseinae clade

Russula has traditionally been a focus for anatomical and chemical studies for species delimitation in macrofungi (Singer, 1986). This is primarily a result of its promiscuity of species richness and distribution as well as the unreliability of the group's most obvious characters to separate species due to their intraspecific variability, which includes pileus color, size of fruitbodies, and general gestalt. This has led to the characterization of over one hundred traits associated just with sporocarps to delimit species morphologically (Romagnesi, 1967). It is the challenge of modern taxonomists to identify what morphologic characters are most phylogenetically significant. To this end, detailed descriptions of sporocarps in the field as well as close examination and measurements of micro-features have been documented for members of subsection *Roseinae*. It is important that detailed descriptions be performed on fresh fruitbodies, including photos taken *in situ*, standardized color descriptions, a spore print taken and scraped *en masse* to be compared to standardized spore color charts (ideally Romagnesi or Crawshay's system), notes on taste and distinct odors, measurements of stipe and pileus, macrochemical tests, surface irregularities, and any change in appearance such as bruising over time.

A number of characters have been used to traditionally unite members of subsection *Roseinae* (Romagnesi, 1967; Sarnari, 1998; Singer, 1986). These species are supposed to have either a white, pink, or red pileus, typically white to cream spore print *en masse*, a lack of pileocystidia, the presence of primordial hyphae with acid-resistant incrustations, a taste that is either bitter or mild, and a strong positive reaction to sulfovanillin that turns "Eosine red" on the stipe of dried material. For the Residual Roseinae clade and Albida clade, species were recovered that contradict this traditional concept of characters. Members of the Residual Roseinae clade possess a spore print with much darker color yellow color (IIIb) than those previously placed in this group. Also, despite being placed

in this subsection by Singer, *Russula albida* does not exhibit the characteristic "Eosine red" reaction to sulfovanillin though still shows purple-red staining contrary to nearly neutral reactions of most *Russula* members. This expanded concept of the group opens the possibility that species that belong to this group may have been described by taxonomists like C.H. Peck, G.S. Burlingham, and W.A. Murrill, who all described many species which are poorly understood today (Buyck, 2007). Our study did not reveal contradictory evidence for pileus color, the lack of pileocystidia, presence of primordial hyphae with acid-resistant incrustations, and taste that is either mild or bitter. In addition, a new character was discovered with the positive reaction of Ehrlich's Reagent or p-(Dimethylamino) benzaldehyde (PDAB) to the stipe surface, which produces a magenta to lilac color. This positive reaction has been demonstrated in other isolated groups as well as a separate positive reaction that turns blue, however, most species of *Russula* do not produce any reaction.

Identifying the traits that the group shares is important, but it is also important to identify which traits are useful in species delimitation. As might be expected, members of the Albida clade and Residual Roseinae clade appear to be the most morphologically divergent of the group. *Russula albida* is the only member of the Roseinae clade to possess a completely white or yellow pileus, spearmint taste, pileipellis that is a corraloid trichoderm, and spores wider than 6.5 µm. This species and "*Russula* cf. *magnorosea*" are the only species to possess spores with ornamentation higher than 0.9 µm. The members of the Residual Roseinae clade are united by their pileipellis that is a loose trichoderm of long-celled and apically attenuated hyphal terminations and large pileus diameter that is only matched by *R. velutipes*. The two species of the Residual Roseinae clade can be separated by a combination of their difference in pileus color, odor, and branching and width of terminal cells in the pileipellis. The Core Roseinae clade is united by a number of characters shared by their common ancestor. This common ancestor likely had a pink flush on the stipe, pileus of intermediate width, a red cap, a fruity or nutty odor, and a palisade trichoderm. The *R. minutula* clade saw a likely transition from red to pink colored pileus and a reduction of pileus diameter and primordial hyphal width. The two species in this group can be separated by differences in their odor, spore width and ornamentation height, and pleurocystidia contents. Though not forming a clade, "*R. leporina*" and *R. peckii* share many features, however they can be differentiated by differences in stipe color, primordial hyphae width, and terminal cell width. In addition, *R. peckii* possesses lamellae whose margins are finely serrated. "*Russula austrorubellipes*" is fairly distinct with features most resembling *R. velutipes*, but differing by possessing a mostly pink stipe, smaller pileus diameter, partly dispersed pleurocystidia contents, more prominent spore ornamentation, and lacking any yellow color in the disc of the pileus. The ancestor of the *R. rubellipes* and "*R. niveopersonata*" clade transitioned to ellipsoid terminal cells of hyphae in the pileipellis forming, at least near the pileus center, an epithelium pileipellis. Both of these species are extremely similar and are the best candidates for cryptic speciation that we have detected, with the only

detected difference being in the width and cell number of primordial hyphae. Another potential case of cryptic speciation may lie within the *R. pseudopeckii* clade, which includes a segregate clade of two collections that was not supported as different by the coalescent delimitation methods. These two clades, however, possess character divergence in their pleurocystidia contents, odor, cell number of primordial hyphae, and spore width. While cryptic speciation may seem common in phylogenetic and coalescent species delimitation studies (Sanchez-Ramirez et al., 2015; Singh et al., 2015), a rigorous and attentive look at morphological features can often yield identifiable diagnostic characters (Adamčík et al., 2016b).

Conclusion

Here we have explored species delimitation in a clade of red *Russula* species using a multi-faceted approach incorporating phylogenetics, geography, ecology, morphology, and the evolution through the multispecies coalescent model. These analyses resulted in the identification of fourteen species in the Roseinae clade, including eight species that have not been formally described yet. Through model testing of different phylogeographic approaches and reconstruction through the DEC +J model, we have inferred an eastern North American origin of subsection *Roseinae*, indicating that the Appalachian Mountains may have acted as a biological hotspot in recent geologic time. Also using a phylogeographic approach, we reconstructed the ancestral host association in this clade to be with deciduous trees, indicating both the possibility of refugia in the southeast U.S. for hardwood tree species as well as the ability of *Russula* species to track range shifts of their host due to climate change. The species recovered here and supported by species delimitation analyses will be formally described along with re-description of type specimens of described species in a future publication.

Acknowledgments

We would like to acknowledge the participants and organizers of the Russulales Workshop 2014 in Slovakia, including Annemieke Verbeken. We would also like to acknowledge the Gulf States Mycological Society for hosting B.P. Looney at their yearly summer foray to collect specimens. The studies of S. Adamčík were funded by the Slovak National Project APVV-15-0210.

References

Adamčík, S., Buyck, B., 2012. Type-studies in American *Russula* (Russulales, Basidiomycota): in and out subsection *Roseinae*. Nov. Hedwigia 94, 413–428. doi:10.1127/0029-5035/2012/0003

Adamčík, S., Carteret, X., Buyck, B., 2013. Type studies on some *Russula* species described by C.H. Peck. Cryptogam. Mycol. 34, 367–391. doi:10.7872/crym.v34.iss2.2013.367

Adamčík, S., Slovák, M., Eberhardt, U., Ronikier, A., Jairus, T., Hampe, F., Verbeken, A., 2016a. Molecular inference, multivariate morphometrics and ecological assessment are applied in concert to delimit species in the *Russula clavipes* complex. Mycologia 108, 716–730. doi:10.3852/15-194

Adamčík, S., Cabon, M., Eberhardt, U, Saba, M., Hampe, F., Slovák, M., Kleine, J., Marxmuller, H., Jančovičova, S.S., and Pfister, D.H., 2016b. A molecular analysis reveals hidden species diversity within the current concept of *Russula maculata* (Russulaceae, Basidiomycota). Phytotaxa 270, 71-88.

Aldrovandi, M.S.P., Johnson, J.E., O'meara, B., Petersen, R.H., Hughes, K.H., 2015. The *Xeromphalina campanells/ kaufmanii* complex: species delineation and biogeographical patterns of speciation. Mycologia 107, 1270-1284. doi:10.3852/15-087

Birkebak, J.M., Mayor, J.R., Ryberg, K.M., Matheny, P.B., 2013. A systematic, morphological and ecological overview of the Clavariaceae (Agaricales). Mycologia 105, 896-911.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10, 1–6. doi:10.1371/journal.pcbi.1003537

Buyck, B., 2007. A new initiative towards the study of *Russula* in the eastern USA. Pagine Di Micol. 27, 81–86.

Buyck, B., Atri, N.S., 2011. A *Russula* (Basidiomycota, Russulales) with an Unprecedented Hymenophore Configuration from Northwest Himalaya (India). Cryptogam. Mycol. 32, 185–190. doi:10.7872/crym.v32.iss2.2011.185

Buyck, B., Jančovičová, S., Adamčík, S., 2015. The Study of *Russula* in the Western United States. Cryptogam. Mycol. 36, 193–211. doi:10.7872/crym/v36.iss2.2015.193

Carstens, B.C., Dewey, T.A., 2010. Species delimitation using a combined coalescent and information-theoretic approach: An example from North American myotis bats. Syst. Biol. 59, 400–414. doi:10.1093/sysbio/syq024

Church, S.A., Kraus, J.M., Mitchell, J.C., Church, D.R., Taylor, D.R., 2003. Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. Evolution (N. Y). 57, 372–383. doi:10.1554/0014-3820(2003)057[0372:efmpri]2.0.co;2

Crandall, K.A., Buhay, J.E., 2008. Global diversity of crayfish (Astacidae, Cambaridae, and Parastacidae - Decapoda) in freshwater. Hydrobiologia 595, 295–301. doi:10.1007/s10750-007-9120-3

Das, K., Atri, N., Buyck, B., 2013. Three new species of *Russula* (Russulales) from Sikkim (India). Mycosphere 4, 722–732. doi:10.5943/mycosphere/4/4/9

Das, K., Miller, S.L., Sharma, J.R., 2006. *Russula* in Himalaya 2: Four new taxa. Mycotaxon 95, 205–215.

Ence, D.D., Carstens, B.C., 2011. SpedeSTEM: A rapid and accurate method for species delimitation. Mol. Ecol. Resour. 11, 473–480. doi:10.1111/j.1755-0998.2010.02947.x

Fatto, R.M., 1998. Notes on four little red Russulas. Mycotaxon (USA).

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27, 480–488. doi:10.1016/j.tree.2012.04.012

Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D.M., Hood, M.E., 2008. Speciation in fungi. Fungal Genetics and Biology 45, 791-802. doi:10.1016/j.fgb.2008.02.001

Graham, E.B., Knelman, J.E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., Beman, J.M., Abell, G., Philippot, L., Prosser, J., et al., 2016. Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? Front. Microbiol. 7, 1–10. doi:10.3389/fmicb.2016.00214

Griffin, S.R., Barrett, S.C.H., 2004. Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in eastern North America: Inferences from phylogeography. Am. J. Bot. 91, 465–473.

Harrower, E., Bougher, N.L., Henkel, T.W., Horak, E., Matheny, P.B., 2015. Long-distance dispersal and speciation of Australasian and American species of *Cortinarius* sect. *Cortinarius*. Mycologia 107, 697–709. doi:10.3852/14-182

Heled, J., Drummond, A.J., 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol. Biol. Evol. 27, 570–580. doi:10.1093/molbev/msp274

Hongo, T., 1960. The Agaricales of Japan 1-3 Russulaceae. Acta Phytotaxon. Geobot. 18, 129–146. doi:10.2307/4117675

Hongo, T., 1970. The Agaricales of Japan 1-3: Russulaceae. Acta Phytotax. Geobot. 18, 129-146.

Hopkins, D.M., others, 1967. The Bering land bridge. Stanford University Press.

Hughes, K.W., Petersen, R.H., 2004. Origins of Fungal Biodiversity in the Southern Appalachians. Inoculum 55, 18.

Hughes, K.W., Petersen, R.H., Lodge, D.J., Bergemann, S.E., Baumgartner, K., Tulloss, R.E., Lickey, E., Cifuentes, J., 2013. Evolutionary consequences of putative intra-and interspecific hybridization in agaric fungi. Mycologia 105, 1577–94. doi:10.3852/13-041

Hughes, K.W., Segovia, A.R., Petersen, R.H., 2014. Transatlantic disjunction in fleshy fungi. I. The *Sparassis crispa* complex. Mycol. Prog. 13, 407–427. doi:10.1007/s11557-013-0927-1

Jayasiri, S.C., Hyde, K.D., Ariyawansa, H.A., Bhat, J., Buyck, B., Cai, L., Dai, Y.-C., Abd-Elsalam, K.A., Ertz, D., Hidayat, I., et al., 2015. The Faces of Fungi database: fungal names linked with morphology, phylogeny and human impacts. Fungal Divers. 74, 3–18.

Katoh, K., H. Toh, 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings in Bioinformatics 9, 286-298.

Kirk, P.M., Cannon, P.F., Stalpers, J.A., 2008. The Dictionary of the Fungi, 10th ed. CABI Publishing, Wallingfork, UK.

Kornerup, A. Wanscher, J.H., 1967. Methuen handbook of colour.

Kozak, K.H., Wiens, J.J., 2010. Niche conservatism drives elevational diversity patterns in Appalachian salamanders. Am. Nat. 176, 40–54. doi:10.1086/653031

Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM: Species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25, 971–973. doi:10.1093/bioinformatics/btp079

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701. doi:10.1093/molbev/mss020

Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276-3278.

Larsson, E., Larsson K.-H., 2003. Phylogenetic relationships of russuloid basidiomycetes with emphasis on aphyllophoralean taxa. Mycologia 95, 1037-1065.

Latham, R.E., Ricklefs, R.E., 1993. Continental Comparisons of Temperate-Zone Tree Species Diversity. Species Divers. Ecol. Communities Hist. Geogr. Perspect. 294–314.

Lickey, E.B., Hughes, K.W., Petersen R.H., 2002. Biogeographical patterns in *Artomyces pyxidatus*. Mycologia 94, 461-471.

Lickey, E.B., Tieken, S.M., Hughes, K.W., Petersen, R.H., 2007. The mushroom TWiG: A marvelous mycological menagerie in the mountains. Southeast. Nat. 6, 73–82. doi:10.1656/1528-7092(2007)6[73:TMTAMM]2.0.CO;2

Liu, L., Yu, L., Pearl, D.K., Edwards, S. V., 2009. Estimating species phylogenies using coalescence times among sequences. Syst. Biol. 58, 468–477. doi:10.1093/sysbio/syp031

Looney B.P., 2015. Molecular annotation of type specimens of *Russula* species described by W.A. Murrill from the southeast United States. Mycotaxon 129, 255–268.

Looney, B.P., Birkebak, J.M., Matheny, P.B., 2013. Systematics of the genus *Auricularia* with an emphasis on species from the southeastern United States. North Am. Fungi 8, 1–25. doi:10.2509/naf2013.008.006

Looney, B.P., Ryberg, M., Hampe, F., Sánchez-García, M., Matheny, P.B., 2016. Into and out of the tropics: Global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi. Mol. Ecol. 25, 630–647. doi:10.1111/mec.13506

López-Giráldez, F., Townsend, J.P., 2011. PhyDesign: an online application for profiling phylogenetic informativeness. BMC Evol. Biol. 11, 152. doi:10.1186/1471-2148-11-152

Lundberg, J.G., Kottelat, M., Smith, G.R., Stiassny, M.L.J., Gill, A.C., Fishes, S.O.M., Little, S.O., Lundberg, J.G., Kottelat, M., An, T., Of, O., Smith, G.R., Melanie, L.J., Stiassny, R.I., Gill, A.C., 2000. So Many Fishes, So Little Time : An Overview of Recent Ichthyological Discovery in Continental Waters. Annals of the Missouri Botanical Garden 87, 26–62.

Maddison, W.P., Maddison, D.R., 2001. Mesquite: a modular system for evolutionary analysis.

Matheny, P.B., 2005. Improving phylogenetic inference of mushrooms with RPB1 and RPB2 nucleotide sequences (*Inocybe*; Agaricales). Molecular Phylogenetics and Evolution 35, 1-20.

Matheny, P.B., Liu, Y.J., Ammirati, J.F., Hall, B.D., 2002. Using RPB1 sequences to improve phylogenetic inference among mushrooms (*Inocybe*, Agaricales). American Journal of Botany 89, 688-698.

Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63, 951–970. doi:10.1093/sysbio/syu056

Matzke, N.J., 2013. BioGeoBEARS: biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts [WWW Document]. R Packag. version 0.2.1. URL http://cran.r-project.org/package=BioGeoBEARS (accessed 1.16.17).

Miller, S.L., Buyck, B., 2002. Molecular phylogeny of the genus *Russula* in Europe with a comparison of modern infrageneric classifications. Mycological Research 106, 259-276.

Mueller, G.M., Wu, Q.-X., Huang, Y.-Q., Gou, S.-Y., Aldana-Gomez, R., Vilgalys, R., 2001. Assesing biogeographic relationchips between North American and Chinese macrofungi. J. Biogeo. 28, 271–281. doi:10.1046/j.1365-2699.2001.00540.x

O'Meara, B.C., 2010. New heuristic methods for joint species delimitation and species tree inference. Syst. Biol. 59, 59–73. doi:10.1093/sysbio/syp077

Parmalee, P.W., Bogan, A.E., 1998. Freshwater mussels of Tennessee. University of Tennessee Press.

Petit, R.J., Aguinagalde, I., de Beaulieu, J.-L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B., Palmé, A., Martín, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300, 1563–1565. doi:10.1126/science.1083264

Pielou, E.C., 1991. After the ice age: the return of life to glaciated North America. University of Chicago Press, Chicago.

Qian, H., Ricklefs, R.E., 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. Nature 407, 180–182. doi:10.1038/35025052

R Core Team, 2015. R: A Language and Environment for Statistical Computing.

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2015. Tracer v1. 6. 2014.

Rannala, B., Yang, Z., 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics 194, 245–253. doi:10.1534/genetics.112.149039

Rannala, B., Yang, Z., 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164, 1645–1656.

Rehner, S.A., Buckley, E., 2005. A *Beauveria* phylogeny inferred from nuclear ITS and EF1-α sequences: evidence for cryptic diversification and links to *Cordyceps* teleomorphs. Mycologia 97, 84-98.

Romagnesi, H., 1967. Les russules d'Europe et d'Afrique du Nord. Bordas Paris.

Ruane, S., Bryson, R.W., Pyron, R.A., Burbrink, F.T., 2014. Coalescent species delimitation in Milksnakes (Genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. Syst. Biol. 63, 231–250. doi:10.1093/sysbio/syt099

Saini, S.S., Atri, N.S., Singer, R., 1982. North Indian Agaricales. II. Sydowia 35, 236–241.

Sánchez-Garcíia, M., Matheny, P.B., Palfner, G., Lodge, D.J., 2014. Deconstructing the Tricholomataceae (Agaricales) and introduction of the new genera *Albomagister*, *Corneriella*, *Pogonoloma* and *Pseudotricholoma*. Taxon 63, 993–1007.

Sánchez-Ramirez, S., Tulloss, R.E., Guzman-Davalos, L., Cifuentes-Blanco, J., Valenzuela, R., Estrada-Torres, A., Ruan-Soto, F., Diaz-Moreno, R., Hernandez-Rico, N., Torres-Gomez, M., Leon, H., Moncalvo, J.M., 2015. In and out of refugia: Historical patterns of diversity and demography in the North American Caesar's mushroom species complex. Mol. Ecol. 24, 5938–5956. doi:10.1111/mec.13413

Sarnari, M., 1998. Monografia illustrata del genere *Russula* in Europa.

Satler, J.D., Carstens, B.C., Hedin, M., 2013. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). Syst. Biol. 62, 805–823. doi:10.1093/sysbio/syt041

Schmitt, I., Crespo, A., Divakar, P.K., Fankhauser, J.D., Herman-Sackett, E., Kalb, K., Nelsen, M.P., Nelson, N.A., Rivas-Plata, E., Shimp, A.D., Widhelm, T., 2009. New primers for promising single-copy genes in fungal phylogenetics and systematics. Persoonia 23, 35-40.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Consortium, F.B., Bolchacova, E., Voigt, K., et al.,

2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc. Natl. Acad. Sci. U. S. A. 109, 1–6. doi:10.1073/pnas.1117018109

Shaffer, R. L., 1962. The subsection compactae of *Russula*. Brittonia 14, 254-284.

Shaffer, R. L., 1964. The subsection Lactarioideae of *Russula*. Mycologia, 202-231.

Shaffer, R. L., 1972. North American Russulas of the subsection *Foetentinae*. Mycologia, 1008-1053.

Silvestro, D., Michalak, I., 2012. raxmlGUI: a graphical front-end for RAxML. Organisms Diversity & Evolution 12, 335-337.

Singer, R., 1986. The Agaricales in Modern Taxonomy, 4th ed. Koeltz Scientific Books, Koenigstein, Germany.

Singh, G., Dal Grande, F., Divakar, P.K., Otte, J., Leavitt, S.D., Szczepanska, K., Crespo, A., Rico, V.J., Aptroot, A., Da Silva Cáceres, M.E., Lumbsch, H.T., Schmitt, I., 2015. Coalescent-based species delimitation approach uncovers high cryptic diversity in the cosmopolitan lichen-forming fungal genus *Protoparmelia* (Lecanorales, Ascomycota). PLoS One 10, 1–20. doi:10.1371/journal.pone.0124625

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of unglaciated eastern North America. Mol. Ecol. 15, 4261–4293. doi:10.1111/j.1365-294X.2006.03061.x

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML Web servers. Systematic biology 57, 758-771.

Stein, B.A., Kutner, L.S., Adams, J.S., 2000. Precious heritage: the status of biodiversity in the United States. Oxford University Press on Demand.

Thiers, B. from http://sweetgum.nybg.org/ih/

Tiffney, B.H., 1985. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. J. Arnold Arbor. 66, 73–94.

Van Der Heijden, M.G.A., Bardgett, R.D., Van Straalen, N.M., 2008. The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol. Lett. 11, 296–310. doi:10.1111/j.1461-0248.2007.01139.x

Vellinga, E., Kuyper, T.W., Vellinga, E.C., Kuyper, T.W., Ammirati, J., Desjardin, D.E., Halling, R.E., Justo, A., Læssøe, T., Lebel, T., 2015. Six simple guidelines for introducing new genera of fungi. IMA Fungus 6, 65–68.

Vieites, D.R., Min, M.-S., Wake, D.B., 2007. Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. Proc. Natl. Acad. Sci. U. S. A. 104, 19903–19907. doi:10.1073/pnas.0705056104

Walker, J.F., Miller, O.K., Horton, J.L., 2005. Hyperdiversity of ectomycorrhizal fungus assemblages on oak seedlings in mixed forests in the southern Appalachian Mountains. Mol. Ecol. 14, 829–838. doi:10.1111/j.1365-294X.2005.02455.x

Wen, J., 1999. Evolution of Eastern Asian and Eastern North American Disjunct Distributions in Flowering Plants. Annu. Rev. Ecol. Syst. 30, 421–455.

White, T.J., Bruns, T., Lee, S., Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18, 315-322.

Yang, Z., 2015. A tutorial of BPP for species tree estimation and species delimitation. Curr. Zool. 61, 854–865. doi:10.1093/czoolo/61.5.854

Yang, Z., Rannala, B., 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 31, 3125–3135. doi:10.1093/molbev/msu279

Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. … Natl. Acad. Sci. 107, 1–6. doi:10.1073/pnas.0913022107

Appendix



Figure 18. Gene tree reconstructions of Incrustatula clade for five loci (ITS, *rpb1*, *rpb2*, *tef1*, *mcm7*) used in multi-gene analyses showing major clade relationships (Red-Core Roseinae; Pink-Residual Roseinae; Orange-Albida; and Purple-Lilaceinae). Asterisks and green stars mark well-supported clade.

CHAPTER III
RUSSULACEAE AS A MODEL SYSTEM TO RECOGNIZE FUNCTIONAL DIVERSITY IN AN ECTOMYCORRHIZAL LINEAGE

The doctoral student executed the literature search, reconstructed phylogenetic trees, supervised and assisted a student intern in designing and executing growth studies, and was the primary author of the manuscript. P. Meidl worked with doctoral student on growth studies. O. Miettinen provided culture material for the Gloeopeniophorella convolvens genome. P. Matheny offered comments on manuscript. J. Labbé also offered comments on manuscript.

Abstract

Ectomycorrhizal fungi (EmF) are ubiquitous members of the microbiota of most temperate and some tropical forested ecosystems. Their important ecological as role mutualists of many tree and shrub species has been explored in the context of ecosystem functioning and forest health, however, recent studies interested in carbon cycling and its effect on global climate change have begun to look at EmF and their contribution to decomposition of organic carbon sources. Oxidative enzymatic activity in soil has been attributed to mostly ectomycorrhizal rather than saprotrophic fungi, the guild previously assumed to be the key players in soil lignin decomposition. EmF have evolved independently at least some 78 times, and niche specialization has likely occurred within and between these lineages, potentially driving their diversification. Here we outline an ongoing sequencing project through the Joint Genome Institute that seeks to densely sample a single, diverse lineage of EmF to explore the evolution of plant-carbon degradative enzymes characterized as a potential niche specialization. The targeted group for this endeavor is the family Russulaceae, which includes the genera *Lactarius*, *Lactifluus*, *Multifurca*, and *Russula*. This group has been highlighted as a principal producer of laccases in temperate soil, a class of copper containing enzymes known to degrade lignin through oxidative reaction. Class II peroxidases, the enzymes primarily used in lignin degradation, have been recovered from forest soils that originate from members of Russulaceae. Here we discuss expectations from this multi-genome study as well as preliminary data on growth characteristics of the group, with discussion on the inherent ecological features of Russulaceae that we seek to understand through their functional genomic composition.

Introduction

A.B. Frank was one of the first scientists to unambiguously identify the symbiotic nature of mycorrhizal fungi in the late 19th century (Trappe, 2005). This shifted the paradigm of functional roles for fungi from the view of all fungi as either antagonistic parasites or nutrient recycling decomposers to one where fungi can be essential mutualistic partners for other lifeforms. We know today that fungi form intimate mutualistic associations with a plethora of other lifeforms including mycorrhizal plants (Brundrett, 2004), orchids (Dearnaley et al., 2012), animals (Currie et al., 1999; Aanen et al., 2002), grasses (Clay, 1988; Busby et al., 2016),

green algae and cyanobacteria (Hawksworth, 1988), bryophytes (Pressel et al., 2010), and bacteria (Partida-Martinez et al., 2007). Though these associations share similarities, most have been subject to convergent evolution and resulted in novel interactions through coevolutionary changes in morphology, behavior, chemical signaling, or gene expression. For example, while the arbuscular mycorrhizal habit has evolved only once in the ancestor of Glomeromycota, ectomycorrhizal fungi have evolved some 78 times independently in members of Ascomycota, Basidiomycota, and Zygomycota (Tedersoo & Smith, 2013). Recognition of the diversity of specialized functions within these specific groups or guilds is an imperative next step towards modeling their associated effects on community dynamics, and integral ecosystem level processes.

Interest in ectomycorrhizal fungi (EmF), one of at least seven subdivisions of mycorrhizal fungi (Smith & Read, 2010), has recently increased as ecological studies attempt to model the flux of carbon through ecosystems in the face of a global accumulation of atmospheric carbon (Treseder & Allen, 2000; Lindahl et al., 2007). Ectomycorrhizal fungi are essential mutualistic partners to many trees and shrubs in forested ecosystems, where fungi act as a substantial extension of the root system of plants, providing scavenged water and nutrients in exchange for carbon allocation from the plant. Only recently have EmF been recognized as significant contributors to the decomposition processes of soil organic matter, leading to the novel term “mycorrhizal decomposition” (Talbot et al., 2008). A broad range of EmF have been found to contain class-II-peroxidase genes responsible for lignin degradation in humus (Bödeker et al., 2009). Furthermore, peroxidase activity in soil has been statistically correllated with EmF species richness (Talbot et al., 2013) and their relative abundance to saprotrophic fungi (Phillips et al., 2014) rather than saprotrophic species richness as one might expect. Recent evidence suggests that the capability of EmF to decompose has been retained from their saprotrophic ancestry (Rineau et al., 2012; Kohler et al., 2015). This has led to some debate over whether EmF are capable of carbon uptake from organic matter in times of stress (i.e. facultative saprotrophy) or, more likely, that EmF use primarily oxidative reactive enzymes to scavenge nitrogen or other limiting nutrients from recalcitrant biopolymers in soil (Lindahl & Tunlid, 2014). The ability to scavenge nutrients from a diverse array of biopolymers found in soil expands the functional diversity potential of EmF, however, knowledge of this potential functional diversity is limited by our current genome sampling.

The age of genome-enabled mycology, looking at broad trends between the genomes of multiple lineages of fungi, was initiated at a large scale by the Fungal Genomics Program (FGP) through the Joint Genome Institute (Walnut Creek, CA) with a focus on energy and environmental science applications (Martin et al., 2011; Hibbett et al., 2013). Through support from the FGP, the Saprotrophic Agaricomycotina Project (SAP) has elucidated the evolution of decay mechanisms in over 30 different species of saprotrophic fungi representing 12 orders. This project has contributed to a much more complex understanding of decomposition that now transcends the brown rot/white rot paradigm (Floudas

et al., 2012; Hibbett et al., 2013; Riley et al., 2014). Also emerging from the FGP is the Mycorrhizal Genomics Initiative (MGI), which has targeted 25 lineages of mycorrhizal fungi to explore the underlying genetic mechanisms for host communication and association (Plett & Martin, 2011). While the MGI seeks to elucidate the genetic machinery behind the mycorrhizal synthesis among its independent lineages, here we outline a genomic study, the Russulaceae Genome Initiative (RGI) that takes an evolutionary perspective to examine trends of functional variation within a single densely sampled and diverse lineage of EmF, the family Russulaceae.

Russulaceae have been nominated as a genomic model EmF group for several attractive reasons. First, like many EmF, Russulaceae are slow growing and difficult to isolate in pure culture. However, members from diverse clades within Russulaceae have frequently been isolated and exploited in manipulative experiments. In addition, species of Russulaceae are common in forested ecosystems and often produce large-bodied mushrooms that are ideal for extracting mostly axenic tissue from their inner context. Second, given the evolutionary age of the group, only three other EmF lineages *Cortinarius*, Inocybaceae, and Amanitaceae can be said to have gone through as rapid diversification as Russulaceae making it ideal for studying speciation processes and potential adaptive radiation (Ryberg and Matheny, 2011). Unlike these other lineages, Russulaceae is a rather isolated EmF lineage with only one other major EmF lineage in its order. Finally, we know that this group is ecologically important for EmF communities, evidenced by its pervasiveness in both species diversity and transcript abundance in soils (Liao et al., 2014). The genus *Russula*, in particular, is known as a late-stage colonizer of forests, indicating *Russula* is likely important for stabilizing nutrient networks in mature and old-growth forests (Twieg et al., 2007). In this review, we provide an overview of Russulaceae in context of its global diversity, evolution, ecology, growth characteristics, and known functional diversity. We highlight several hypotheses ideal to test with the proposed dataset, in the hopes of spurring investigation and collaboration focused on unlocking the functional diversity within this fascinating group.

Systematics and Sampling Strategy

Russulaceae are a species rich lineage of EmF that have traditionally been composed of the genera *Russula* Pers. and *Lactarius* Pers. (Singer, 1986). Mushroom-forming *Russula* species are commonly referred to as brittlegill mushrooms for their easily broken lamellae and flesh due to an abundance of specialized cells called sphaerocytes, a synapomorphy of the family (Miller *et al.*, 2006). Species of *Russula* are typically characterized as having a white to orange spore deposit, a mild to very acrid taste, ornamented and amyloid spores, and usually a brightly colored pileus. *Lactarius*, known as the milkcap mushrooms for the latex exuded from the flesh and lamellae where damaged, share many features with *Russula* but can have darker orange colored spore deposits, higher spore ornamentation, fewer sphaerocytes in the context,

concentric zonations on the pileus, and strobicules, or pits, on the stipe. These two genera were recently split into four genera: the elevated genus *Lactifluus* (Pers.) Roussel made up of mostly former tropical members of *Lactarius*; and the new genus *Multifurca* Buyck & V. Hofstetter, made up of former members of both *Lactarius* and *Russula* (Buyck *et al.*, 2008; Verbeken *et al.*, 2011). No stable synapmorphies have been identified to separate these segregate genera morphologically, but both *Multifurca* and *Lactifluus* can be distinguished by a combination of traits: *Multifurca* by the presence of zonations on the pileus that continue into the context and lamellae with many furcations, or forks; and *Lactifluus* by the presence of thick-walled elements, lamprocystidia, and a hymenophoral trama composed of sphaerocytes (Buyck *et al.*, 2008; De Crop *et al.*, 2016). Phylogenetic relationships between the four core genera of Russulaceae have not been well-resolved using multi-gene phylogenies (Buyck *et al.*, 2008; Verbeken *et al.*, 2014)(Figure 19). Included in Russulaceae are a number of polyphyletic genera, including *Arcangelliella* Cavara, *Cystangium* Singer & A.H. Sm., *Elasmomyces* Cavara, *Gastrolactarius* R. Heim ex J.M. Vidal, *Gymnomyces* Massee & Rodway, *Macowanites* Kalchbr., *Martellia* Mattir. and *Zelleromyces* Singer & A.H. Sm., which all comprise only species with sequestrate basidiocarp morphologies (Miller *et al.*, 2001). Also, a cluster of six species belonging to the genera *Boidinia* Stalpers, *Gloeocystidiellum* Donk, and *Gloeopeniophorella* Rick have been recovered as part of Russulaceae in a study of corticioid fungi hypothesized to be in the order Russulales (Larsson & Larsson, 2003). Although such corticioid species have a resupinate habit on wood and are putatively assigned as white-rot saprotrophs, this has yet to be confirmed (Miller *et al.*, 2006).

Ectomycorrhizal members of Russulaceae comprise one of the most species diverse lineages of EmF and are frequently dominant members of EmF communities. Currently we recognize an estimated 900 species of *Russula* (Buyck & Atri, 2011), 300 species of *Lactarius,* 150 species of *Lactifluus* (De Crop *et al.*, 2016), and 6 species of *Multifurca* (Buyck *et al.*, 2008; Lebel *et al.*, 2013) are accepted worldwide. Many species have yet to be described (Buyck & Thoen, 1996; Buyck, 2007). A global metanalysis of *Russula* sequences in GenBank recovered almost 1200 MOTUs of *Russula* alone (Looney *et al.*, 2016). Russulaceae are a dominant ectomycorrhizal lineage in a multitude of different habitats including arctic shrublands (Geml *et al.*, 2012), boreal forests (Geml *et al.*, 2010), beech hardwood forests (Burke *et al.*, 2009), *Notholithocarpus* forests in the western United States (Bergemann & Garbelotto, 2006), neotropical caesalipinoid forests (Henkel *et al.*, 2012), and dipterocarp tropical rainforests (Peay *et al.*, 2010b). Russulaceae are also present in depauperate EmF communities like those associated with *Alnus* (Põlme *et al.*, 2013), Nyctaginaceae (Haug *et al.*, 2005), and *Gnetum* (Tedersoo & Põlme, 2012).

The Russulaceae Genome Initiative seeks to target the phylogenetic breadth of the lineage to capture as much evolutionary divergence as possible.

Figure 19. Maximum Likelihood phylogenetic reconstruction of Russulaceae with sampling from Buyck et al. (2008) and Looney et al. (2016). Clades are collapsed based on clades from Looney et al. (2016), the molecular subgeneric classification of *Lactarius* from (Verbeken & Nuytinck, 2013), and the subgeneric classification of *Lactifluus* from De Crop et al. (2017). Taxon sampling for the Mycorrhizal Genome Initiative is highlighted in blue and RGI samples are highlighted in red. Samples marked with * are still being prepared for submission to sequence.

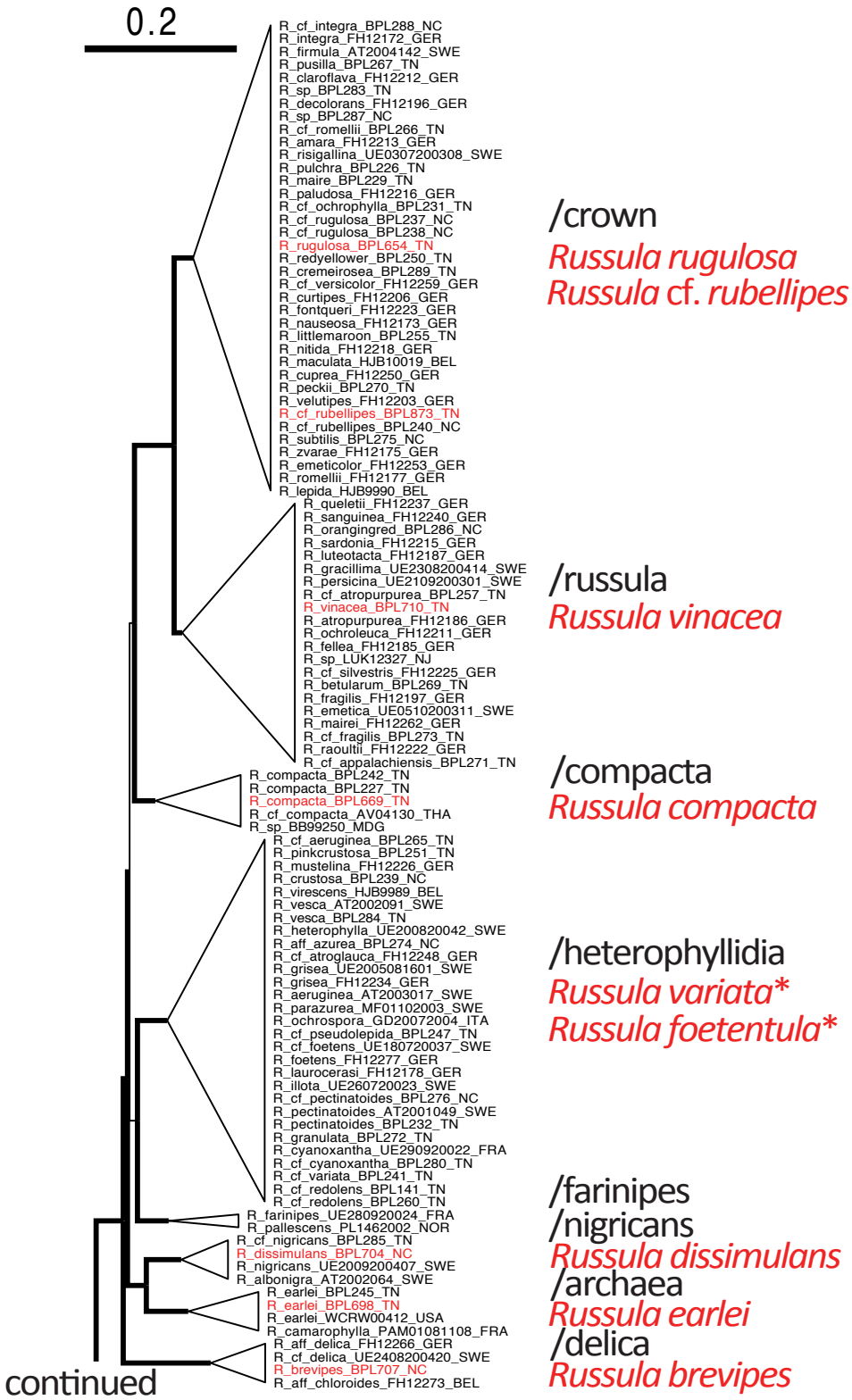


Figure 19 continued

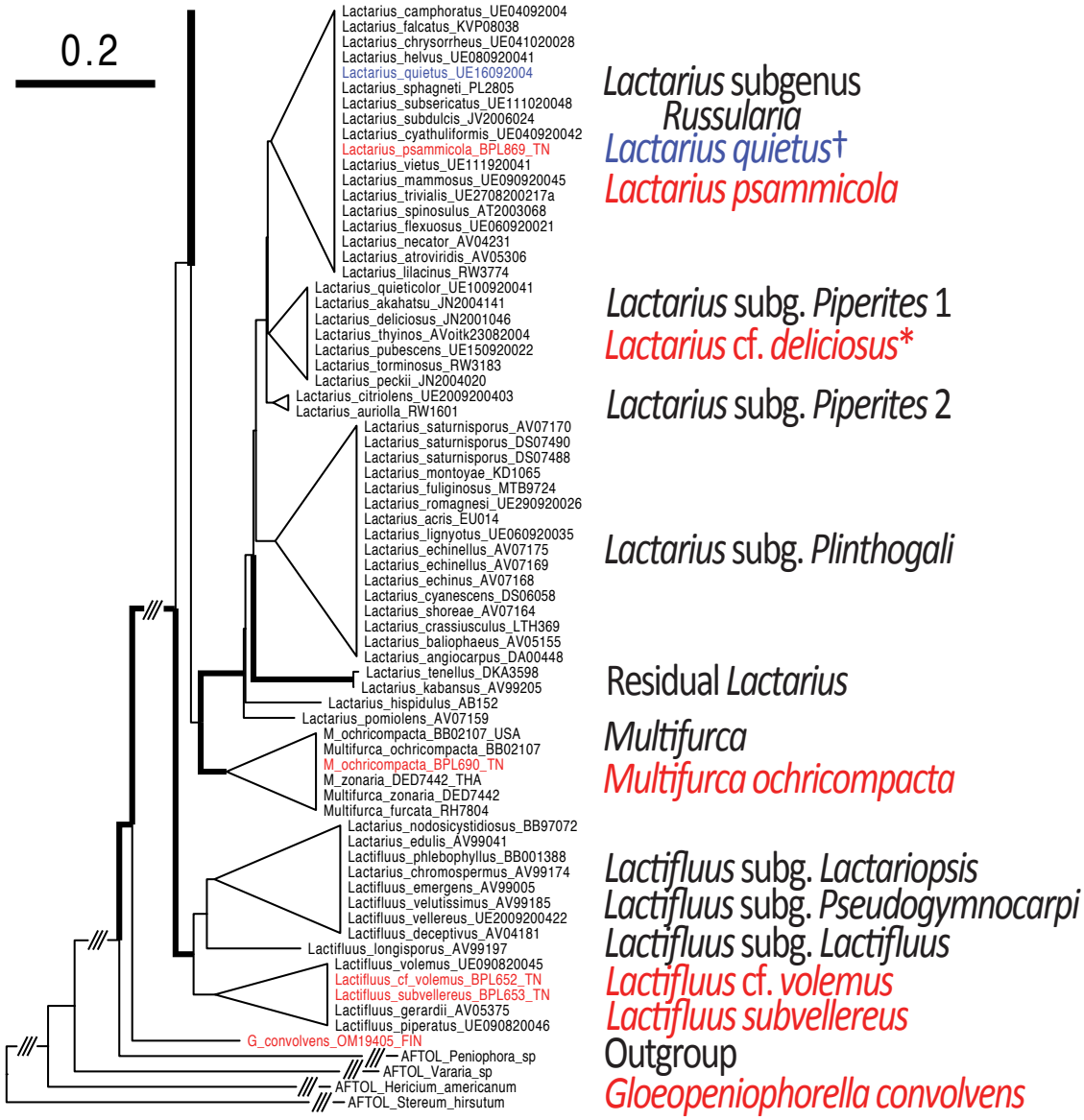


Figure 19 continued

All EmF species sampled for the sequencing project are from the same geographic region of eastern North America from temperate hardwood or mixed forests. Therefore they share the same or similar ecological communities, thereby controlling for geographic variation and hopefully capturing diverse functional roles. Phylogenetic sampling is based on published phylogenies for *Russula* (Looney *et al.*, 2016) and Russulaceae (Buyck *et al.*, 2008), balancing sampling based on the species richness of the different genera. For the initial dataset, 14 species are targeted for genome and transcriptome sequencing, including seven species of *Russula*, two species of *Lactifluus*, one species of *Multifurca*, two species of *Lactarius* in addition to the released genome of *Lactarius quietus* (Fr.) Fr., and a saprotrophic outgroup, *Gloeopeniophorella convolvens* (Table 6).

Table 6. Russulaceae genome and transcriptome sampling.

Taxon	Isolate	Source	Country	Date collected	GenBank
Gloeopeniophorella convolvens	OM19405	culture	Finland	9/16/15	KY848506
Lactarius quietus	S23C	culture	France	unknown	NA
Lactarius cf. *deliciosus*	BPL912	culture	U.S.A.	10/28/15	NA
Lactarius psammicola	BPL869	culture	U.S.A.	9/6/14	KY848507
Lactifluus subvellereus	BPL653	sporocarp	U.S.A.	6/27/15	KY848508
Lactifluus cf. *volemus*	BPL652	sporocarp	U.S.A.	6/27/15	KY848509
Multifurca ochricompacta	BPL690	sporocarp	U.S.A.	7/13/15	KY848510
Russula brevipes	BPL707	sporocarp	U.S.A.	7/16/15	KY848511
Russula compacta	BPL669	sporocarp	U.S.A.	7/8/15	KY848512
Russula dissimulans	BPL704	sporocarp	U.S.A.	7/16/15	KY848513
Russula earlei	BPL698	sporocarp	U.S.A.	7/15/15	KY848514
Russula cf. *rubellipes*	BPL873	sporocarp	U.S.A.	10/16/15	KY848515
Russula rugulosa	BPL654	sporocarp	U.S.A.	6/27/15	KY848516
Russula vinacea	BPL710	sporocarp	U.S.A.	7/17/15	KY848517

Genomes are being sequenced at the Joint Genome Institute with a PacBio RS (Pacific Biosystems, California) or the Illumina Hi-Seq 2500 platform (Illumina, California). Two species, *Lactarius psammicola* A.H. Sm. and *Lactarius* cf. *deliciosus*, were successfully cultured and grown on Modified Melin-Norkrans media to derive the axenic cultured material for sequencing. Another species, *Multifurca ochricompacta* (Bills & O.K. Mill.) Buyck & V. Hofstetter, was successfully cultured but could not be grown in enough quantity for genome sequencing, so material from the basidiocarp from which the culture was derived was used instead.

The family Russulaceae is a member of the order Russulales (Basidiomycota), a group diverse in nutritional strategies, basidiocarp morphologies, and ecological habits (Miller *et al.*, 2006). To date, 10 genomes have been sequenced from this order including the saprotrophic genera *Artomyces* P. Karst, *Auriscalpium* Gray, *Dentipellis* Donk, *Hericium* Pers.,

Heterobasidion Bref., *Lentinellus* P. Karst, *Peniophora* Cooke, *Stereum* Hill ex Pers., and *Vararia* P. Karst. One genome from a representative ectomycorrhizal species of *Lactarius* has been produced through the MGI (Figure 20).

The most focused work on genomes from Russulales has been on *Stereum hirsutum* (Willd.) Pers., a white-rot saprotroph on usually attached but dead branches of living trees, and *Heterobasidion irregulare* Garbel. & Otrosina (annotated as *H. annosum*), which is considered both a saprotroph and a pathogen on mostly coniferous trees (Olson *et al.*, 2012). *Heterobasidion irregulare* has been shown to switch metabolic strategies given different environmental conditions. In the presence of cellulose and lignin the transcriptome of *H. irregulare* showed up-regulation of carbohydrate-active enzymes (CAzymes) typical for saprotrophs, whereas a higher expression of metabolite genes involved in toxin production and protection against plant defense in parasitic conditions were detected in the presence of the plant host, showing a trade-off between nutritional modes (Olson *et al.*, 2012).

As representatives of Russulales, both species have shown to contain specialized gene content indicative of their evolutionary isolation. *Stereum hirsutum* utilizes a powerful arsenal of enzymes for lignin degradation, and of all wood-rot genomes analyzed so far, *S. hirsutum* has the highest number of glucose-methanol-choline (GMC) oxidoreductase genes, a family of enzymes with diverse catalytic activities (Riley *et al.*, 2014). An analysis of the mating type loci found that the hormone receptor gene family involved in fungal mating, *STE3*, was found to have arisen three times in the ancestor of *Stereum* and *Heterobasidion*, which may be correlated with a switch from a bipolar to tetrapolar mating system in the ancestor of Russulales (James *et al.*, 2013). In an analysis of small secreted proteins called hydrophobins implicated in a number of functions including pathogen virulence factors, mycorrhiza formation, and cell wall assembly, seventeen hydrophobin genes were found in the genome of *Heterobasidion*, whereas *Stereum* was found to be the only genome analyzed that lacked these genes entirely (Mgbeahuruike *et al.*, 2013).

Phylogenetic relationships between families and other major groups in Russulales have not been well-resolved using single or multi-gene approaches, and there is still extensive work to do to understand the functional diversity of this order (Miller *et al.*, 2006; Zhou & Dai, 2013). Other than Russulaceae, important lineages with an array of nutritional modes to be targeted for future genome sequencing include the genus *Albatrellus* Gray, the other major ectomycorrhizal lineage in the Russulales, *Bondarzewia* Singer, a small genus of root parasites, *Entomocorticium* H.S. Whitney, Bandoni & Oberw., a monotypic genus symbiotic with insects, and *Echinodontium* Ellis & Everh., a genus of root parasites.

Diversification, Biogeographic, and Host Association Patterns

What factors are driving the high species diversity we see in EmF lineages like Russulaceae is a major question being addressed by evolutionary biologists.

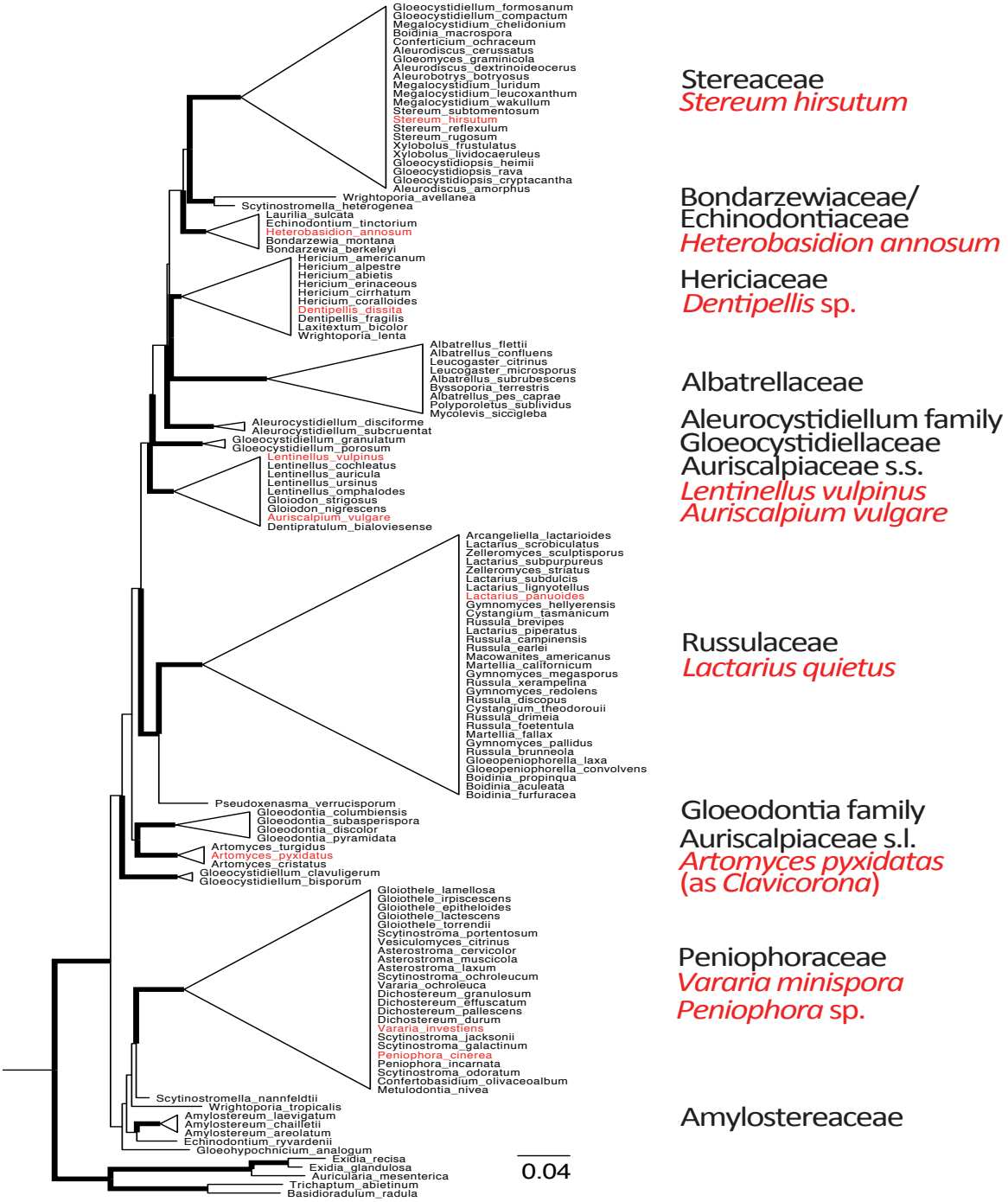


Figure 20. Maximum Likelihood phylogenetic reconstruction of Russulales dataset from Miller et al. (2006) with a backbone constraint topology of well-supported relationships from the Mycocosm Portal phylogeny (Grigoriev *et al.*, 2014). Clades are collapsed based on clades recovered from Miller et al. (2006) and bootstrap values ≥ 50 are reported. Representative samples of species and/or genera for which genomes are currently available are highlighted in red.

At a local scale, dozens of EmF lineages and species of the same lineage can be found coexisting in the same environment and even with a single plant host individual (Bahram *et al.*, 2011). Allopatric speciation due to dispersal limitation has been favored as the primary process for fungal speciation as we see few co-occurring sister species as well as conserved enzymatic regimes in geographically segregated fungal communities (Geml *et al.*, 2008; Peay *et al.*, 2010a; Talbot *et al.*, 2014), though some examples of sympatric sister species are documented (Sánchez-Garcíia et al. 2016; Van Dorp *et al.*, 2016). At the global scale EmF exhibit a diversity pattern counter to the latitudinal biodiversity gradient, where EmF diversity peaks near the temperate/boreal interface and declines towards the tropics (Tedersoo *et al.*, 2012). This pattern is likely governed by innate factors causing a higher net diversification rate, rather than dispersal from tropical regions (Sánchez-Ramírez *et al.*, 2015; Looney *et al.*, 2016). Several factors were found significant in predicting this large-scale pattern, including the mean annual temperature; mean annual precipitation; edaphic factors; and plant host diversity. If functional diversity exists within Russulaceae, we should expect these to be main factors driving diversification however the exact mechanisms or how these factors have driven adaptation, specialization, and evolution of ecosystem function still needs further exploration.

Russulaceae began diversification around 55-61 Mya during the early Paleogene period, before the early Eocene climatic optimum when global climates began the gradual cooling trend that continued through the ice ages of the late Pleistocene (Zachos et al. 2001; Looney *et al.*, 2016; Wisitrassameewong *et al.*, 2016). It is possible that the group originated in the paleotropics as *Lactifluus*, the clade that appears to be sister to the rest of Russulaceae, is a largely tropical group found mostly in Africa (De Crop *et al.*, 2016), however, this still needs to be tested. *Russula*, the most diverse genus in Russulaceae, likely originated in the north temperate region, with at least one major shift to the tropics and multiple shifts back, though a tropical ancestry is still possible (Buyck *et al.*, 1996; Looney *et al.*, 2016). Recent molecular systematic treatments have shown that many north temperate species are widespread, segregated by host or soil type, and may be extremely genetically similar in the ITS marker region ((Adamčík *et al.*, 2016a; Adamčík *et al.*, 2016b).

Plant host diversification has likely been an important driver of diversification and evolution of functional diversity in Russulaceae. Russulaceae exhibit a pattern of host generalism despite some species having narrow or specific host preferences (De Crop *et al.*, 2016; Looney *et al.*, 2016). Host association was investigated in a global survey of *Russula*, which found that association with conifers and frequent host switching from conifers to hardwoods with subsequent host expansion has driven *Russula* diversification (Looney *et al.*, 2016). This could indicate that Russulaceae can take advantage of similar niches in different geographical regions to fulfill important roles for a phylogenetically wide range of hosts. A close look at *Lactarius* species in Alaska showed that species were highly partitioned by habitat type yet not necessarily by host (Geml

et al., 2009). Plant hosts may therefore act as bridges for Russulaceae to expand and diversify globally without being niches for functional specialization.

Ecology, Life History, and Ecosystem Function

Russulaceae play an important role in ecosystems with ectomycorrhizal hosts as symbionts, but what their specific role in nutrient cycling and plant health will be a major focus of the RGI project. One major trend is that with the addition of nitrogen, EmF communities can come to be dominated by members of Russulaceae (Lilleskov *et al.*, 2002; Avis *et al.*, 2003; Allison *et al.*, 2008). To explain this, it has been suggested that Russulaceae have a competitive advantage when nitrogen is not limiting because they are adapted to acquire phosphorus, perhaps through oxalates produced from their mycorrhizal cystidia (Avis *et al.*, 2003). Russulaceae, along with Amanitaceae, were shown to grow slowly on nitrate and genes encoding for nitrate reductase (*nar* genes) could not be amplified using PCR primers possibly due to a loss of selective constraints (Nygren *et al.*, 2008). Nygren et al. (2008) suggest that Russulaceae are specialized in the uptake of ammonium instead of nitrite as a nitrogen source, giving a competitive advantage as ammonium is a less energy-intensive source of nitrogen that does not require active transport across the cell membrane. It was also shown by Lilleskov et al. (2002) that Russulaceae do not grow well on nitrogen from protein sources except for glutamine, an amino acid that is incorporated as ammonium is being taken up. To explain the shift in EmF community composition with nitrogen deposition, Nygren et al. (2008) point out that the hydrophilic character of the mycorrhizal mantle in Russulaceae allow nitrates to be passively absorbed directly into the host, thereby avoiding nitrate toxicity for the fungus (Nygren *et al.*, 2008). It will be important to verify the presence of *nar* genes in the genomes of Russulaceae and to determine if the genes involved in nitrogen and phosphorus acquisition have seen expansions or contractions indicating functional specialization.

Niche differentiation has been proposed to explain the apparent functional redundancy in EmF fungi at both spatial and temporal scales (Koide *et al.*, 2007; Courty *et al.*, 2008). In soil, Russulaceae have been shown to be partitioned by soil horizon, indicating spatial differentiation (Geml *et al.*, 2010). Most members of Russulaceae exhibit the contact exploration type of ectomycorrhizae with little to no emanating hyphae, though some *Lactarius* have a medium-distance smooth type (Agerer, 2001). This distribution of hyphae around the roots has been correlated with the production of phenoloxidases (Agerer *et al.*, 2000), nitrogen isotope content (Hobbie & Agerer, 2010), and preference for root density (Peay *et al.*, 2011), which suggests some functional conservation and potential competition for dense root colonization. Temporal partitioning has been found between *Lactarius* and *Russula*, with *Lactarius* species abundant as mycelium during fall and *Russula* being variably present throughout the year (Koide *et al.*, 2007). The opposite was found for Courty et al. (2008) where the *Lactarius* species was found year-round and *Russula* peaking in the spring. This gives

some evidence that temporal preferences are not conserved at the genus level and may have contributed to species diverisification within Russulaceae rather than functional differentiation. However, Russulaceae have a high species diversity in tropical ecosystems where soil profiles are not pronounced, which suggests that functional and not spatial niche differentiation may be more important.

Russulaceae sporocarps are occasionally observed fruiting on rotting wood (Kropp, 1982; Roberts *et al.*, 2004), however, there is little evidence that Russulaceae can be considered as capable of saprotrophy. Traditional assays for enzymatic activity have determined that species in Russulaceae are unable to degrade biopolymers like cellulose to the extent that obligate saprotrophs are (Lamb 1974, Oort 1981, Hutchison 1990a), supporting their status as obligate EmF. It has also been demonstrated that Russulaceae have the typical nitrogen:carbon isotopic signature of mycorrhizal fungi, indicating they are receiving their carbon directly from the plant host and accessing older nitrogen from well decayed soil organic matter (Hobbie *et al.*, 2001). It is also likely that Russulaceae have lost the genes capable of degrading and accessing carbon from cellulose, as has been shown for the ectomycorrhizal genus *Amanita* (Wolfe *et al.*, 2012). It has, however, been demonstrated that Russulaceae have retained the ability to produce polyphenolic compounds including laccase and tyrosinase that are implicated in plant matter degradation (Lindeberg 1948, Giltrap 1982, Hutchison 1990b).

Ontogeny of sporocarp development has been highlighted as a future prospect for genomic and transcriptomic studies (Hibbett *et al.*, 2013; Nowrousian, 2014). Russulaceae are typically gymnocarpic, but certain groups, especially in the tropics have been shown to be pilangiocarpic, mixangiocarpic, pileostipitocarpic, and even monovelangiocarpic (Singer, 1986). Species with a velum have only been found in the genus *Lactifluus* (De Crop *et al.*, 2016). Russulaceae contain members that produce alternate sporocarp morphologies, including pleurotoid, secotioid, and gasteroid habits (Miller *et al.*, 2001). All known secotioid ‘milk cap’ taxa currently belong to *Lactarius* sensu stricto, whereas all pleurotoid ‘mild cap’ taxa belong to *Lactifluus* (De Crop *et al.*, 2016). A few south temperate and neo-tropical species of pleurotoid *Russula* have been reported and (Buyck & Horak, 1999; Henkel *et al.*, 2000). Spore ontogeny has been investigated in both agaricoid and gasteroid *Russula* species, which has found that the spores have four walls and an identical early development, indicating that ballistosporic modifications to the hilar appendage come later in development which are missing in gasteroid species (Miller, 1988a,b). Genomes and transcriptomes sequenced for this initiative have primarily been sequenced from sporocarps, so in cases where cultures have been obtained, there is potential for comparative transcriptomic studies with gasteroid representatives. *Lactarius deliciosus* has been shown to be capable of sporocarp fruiting *in planta* in a greenhouse (Guerin-Laguette *et al.*, 2000) showing that Russulaceae have potential for evo-devo studies of EmF.

Russulaceae comprises a lineage known to be an important host for mycoheterotrophic plants in Orchidaceae and Ericaceae, whose members rely on the mycorrhizal fungal network to move fixed carbon from the ectomycorrhizal plant host to the usually achlorophyllous, hemiparasitic plant (Bidartondo & Bruns, 2001, 2005; Girlanda *et al.*, 2006). As these plants rely on mycorrhization for their nutrition, they also require these fungi to be present for seed germination, which can only be induced by the same fungus that associates with the mature plants. This extreme specificity has been hypothesized to be the result of a coevolutionary interaction where specific volatile chemical cues of closely related fungal species sharing particular genotypes promote germination and growth of the plant but constrain its ability to switch fungal or germinate seeds successfully (Bidartondo & Bruns, 2005). With reference genomes produced by this project, potential genetic controls for seed germination and mycorrhizal association with heterotrophic plants can be identified and explored in controlled environments using RNAseq approaches. This could potentially lead to a greater understanding of fungal symbiosis using a tripartite system.

In addition to their diverse, and complex relationship with plants, members of Russulaceae have been shown to exhibit beneficial bacterial-fungal relationships. The interactions are increasingly being investigated with mycorrhizal fungi, which are known to frequently harbor both mycorrhization helper bacteria (MHB) and endosymbiotic bacteria that can facilitate host colonization, nutrient acquisition and reproduction (Frey-Klett *et al.*, 2007; Kobayashi & Crouch, 2009). Co-inoculation of symbiotic bacteria and EmF with plant hosts have shown synergistic benefits to both fungal and plant growth (Wu *et al.*, 2012). It is likely that the plant host controls for associated bacteria in these tripartite mutualisms, as shown with *Lactarius deliciosus* where mycosphere-associated bacteria promote nutrient mobilization for *Pinus pinaster* and root growth rate in *Pinus pinea* (Barriuso *et al.*, 2005). Another study examined bacterial associates with *Lactarius rufus* and found variation in growth-promoting properties with some isolates better able to spread to root tips and others able to increase root colonization up to twice as much as a control group (Poole *et al.*, 2001). Metabolites from a specific MHB species, *Paenibacillus* sp., increased *Lactarius deliciosus* hyphal branching yet decreased hyphal radial growth, suggesting a particular mechanism for increasing plant mycorrhization (Aspray *et al.*, 2013). To our knowledge, no study has yet looked for endosymbiotic bacteria in Russulaceae.

Growth Characteristics *in vitro*

Though often considered difficult to culture, studies have documented growth characteristics for species of both *Lactarius* and *Russula* (Hutchison, 1990). Spores of EmF in Russulaceae do not germinate on media, so all cultural studies of Russulaceae are sourced from tissue of either a sporocarp or EmF root tip. *Lactarius deliciosus*, a species complex which we have successfully cultured,

has been shown to easily form mycorrhizae with *Pinus sylvestris,* even producing fruitbodies (Guerin-Laguette *et al.*, 2000). Within Russulaceae, *Lactarius* has been scrutinized in more detail than other genera, due to the ease of its culturability *in vitro* (Hutchison 1999). Melin (1924) first confirmed the ectomycorrhizal status of *Lactarius* when he successfully synthesized *L. deliciosus* with *Pinus montana*.

In anticipation of the RGI first genome release, a growth study was initiated utilizing three species, *Multifurca ochricompacta, Lactarius* cf. *deliciosus*, and *Lactarius psammicola* to test several variables on growth characteristics of Russulaceae sampled through the JGI. The cultures used in these experiments were sourced from sporocarps and therefore represent dikaryotic tissue. For six weeks, five treatment groups were applied to the species including temperature, pH, light exposure, organic nutrient sources, and introduction of a known Mycorrhizal Helper Bacteria (MHB), which are bacteria species that have been shown to increase mycelial growth rate. Different temperatures and pH had drastic effects on the growth of the three Russulaceae species (Figure 21). *Lactarius psammicola* has a more cycrophilic growth pattern, with a preference for 18–21 C, in comparison to *L.* cf. *deliciosus* and *M. ochricompacta* with preferences ranging from 21–25 C (Figure 22). For pH we see the two *Lactarius* species sharing a preference for a high pH of 7, whereas *M. ochricompacta* prefers a lower pH of 5 (Figure 23 & 24). This agrees with previous studies that have examined the closely related *L. deliciosus*, which had a pH preference of 7 (Sanchez et al. 2001). Exposure to a 12-hour light cycle resulted in the death of all cultures examined. Light exposure has been shown negatively affect fungal growth by causing damage to DNA (Rodriguez-Romero et al. 2010). The addition of soil organic matter and macerates from Populus roots had no impact on the growth for any of the Russulaceae species. The introduction of the MHB, Pseudomonas sp. GM41, produced varying responses from the three Russulaceae species. GM41 negatively affected the growth of *Multifurca ochricompacta*, *Lactarius* cf. *deliciosus's* growth was not significantly affected by GM41, and *Lactarius psammicola* failed to grow in the presence of GM41. This result is not surprising, as MHB have been shown to have variable effects on different species of EmF (Bowen & Theodorou, 1979; Garbaye & Bowen, 1989). It will be useful to apply more putative MHB species to determine if Russulaceae exhibit a higher growth rate with other soil microbes.

Gene Content and Molecular Composition

Numerous novel compounds have been identified and isolated from members of Russulaceae. Sesquiterpenes are one of the most well-studied group of compounds isolated from Russulaceae, which give the sporocarps their characteristic sharp or acrid taste (Sterner *et al.*, 1985a,b; Bergendorff & Sterner, 1988; Anke *et al.*, 1989; Clericuzio *et al.*, 2012).

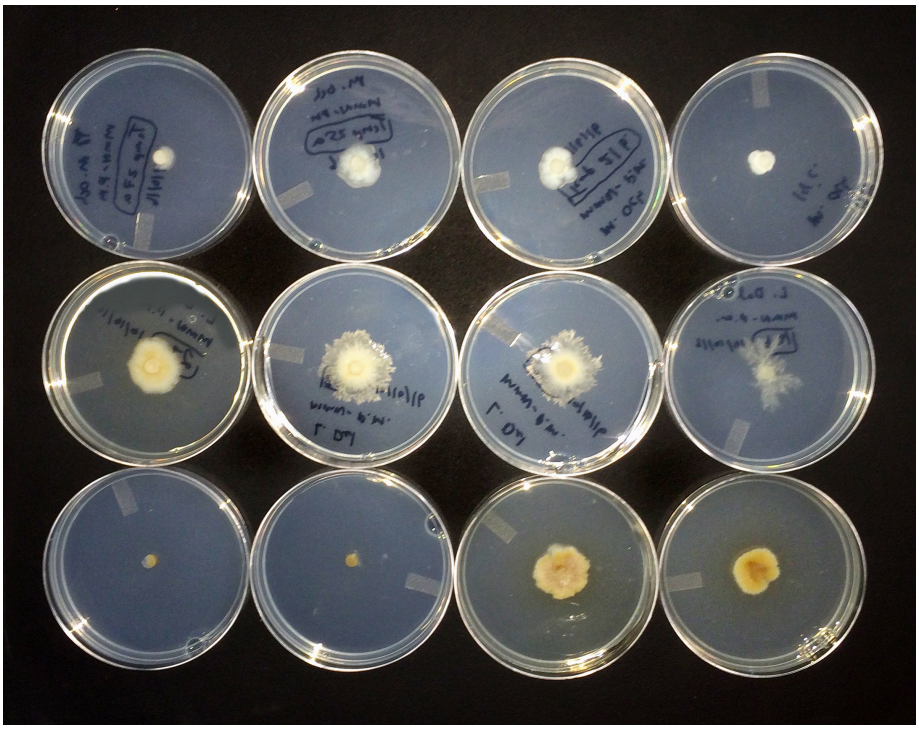


Figure 21.Growth and colony morphology of *M. och* (*Multifurca ochricompacta*), *L.del* (*Lactarius* cf. *deliciosus*), and *L.psa* (*Lactarius psammicola* with response to a temperature gradient in MMN.

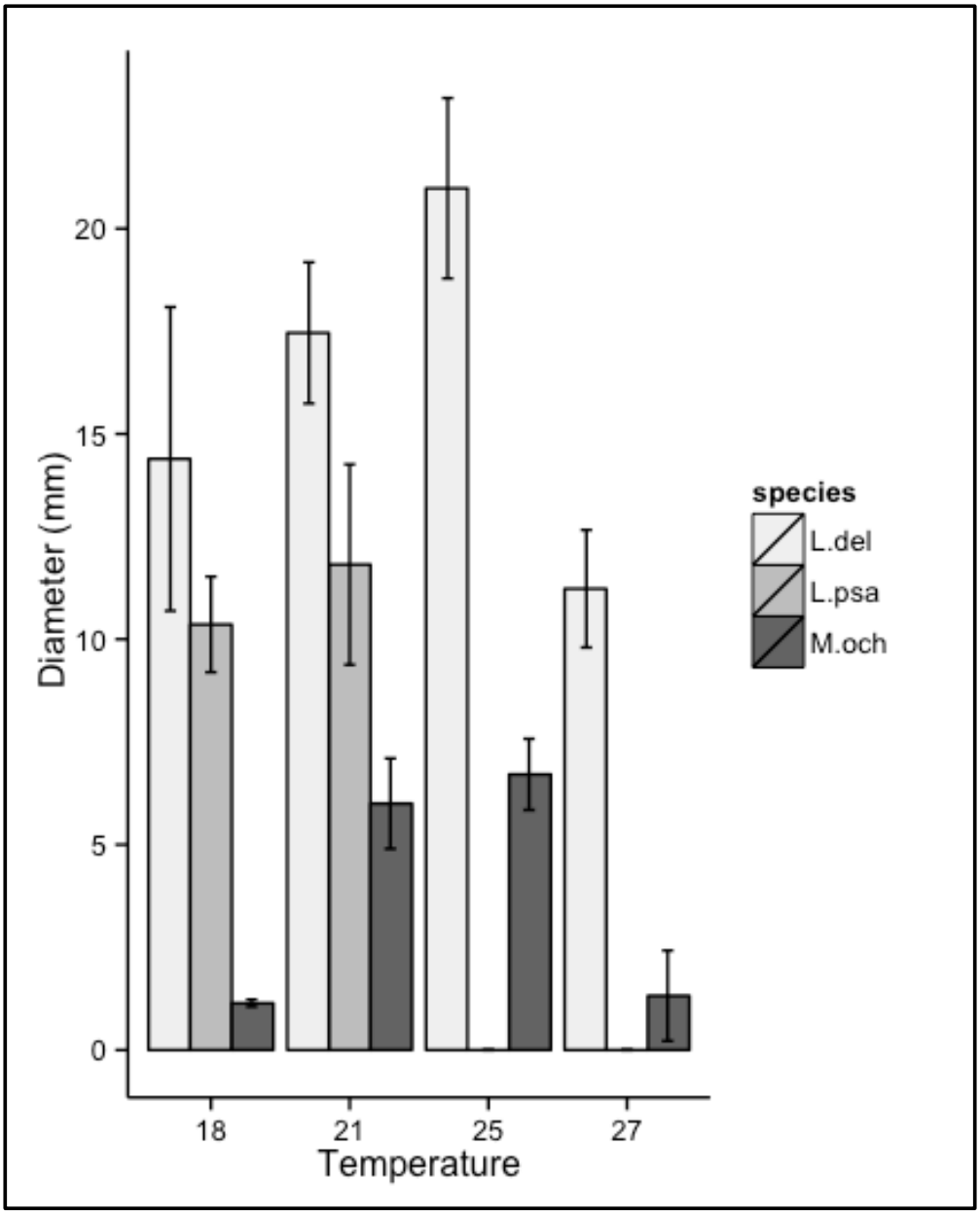


Figure 22.Growth of *L.del* (*Lactarius* cf. *deliciosus, M.och* (*Multifurca ochricompacta*), and *L.psa* (*Lactarius psammicola*) in terms of colony diameter (mm) in MMN in response to different temperature treatments. Each data point is a mean of three replicates with a standard deviation of the mean.

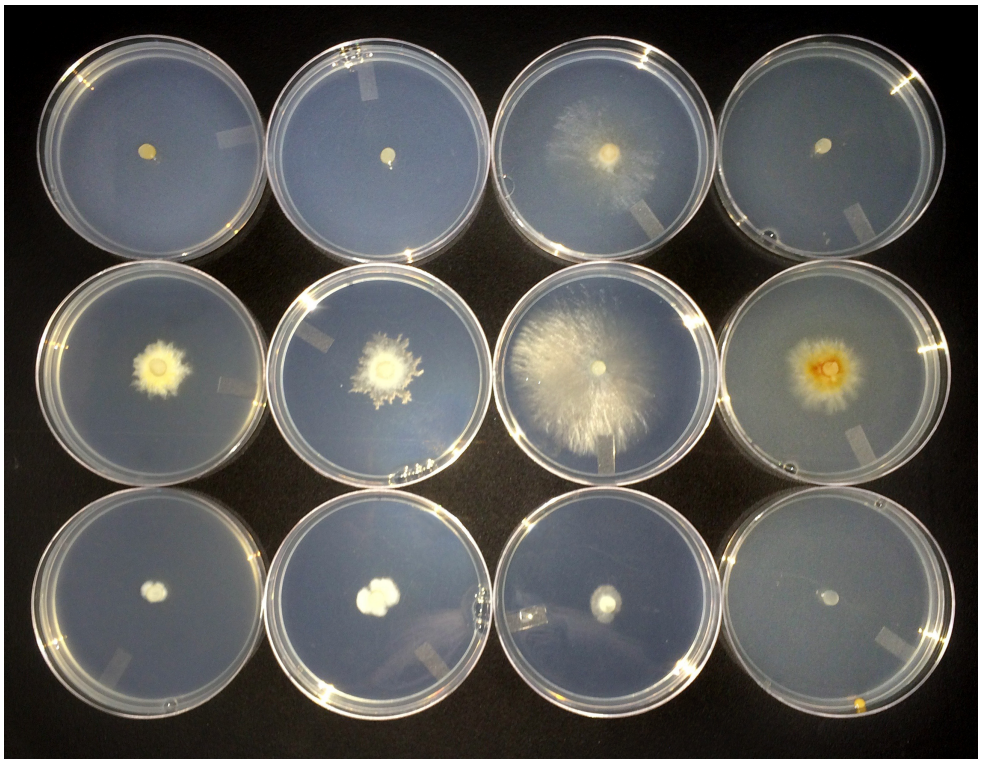


Figure 23. Growth and colony morphology of *M. och* (*Multifurca ochricompacta*), *L.del* (*Lactarius* cf. *deliciosus*), and *L.psa* (*Lactarius psammicola* with response to a temperature gradient in MMN.

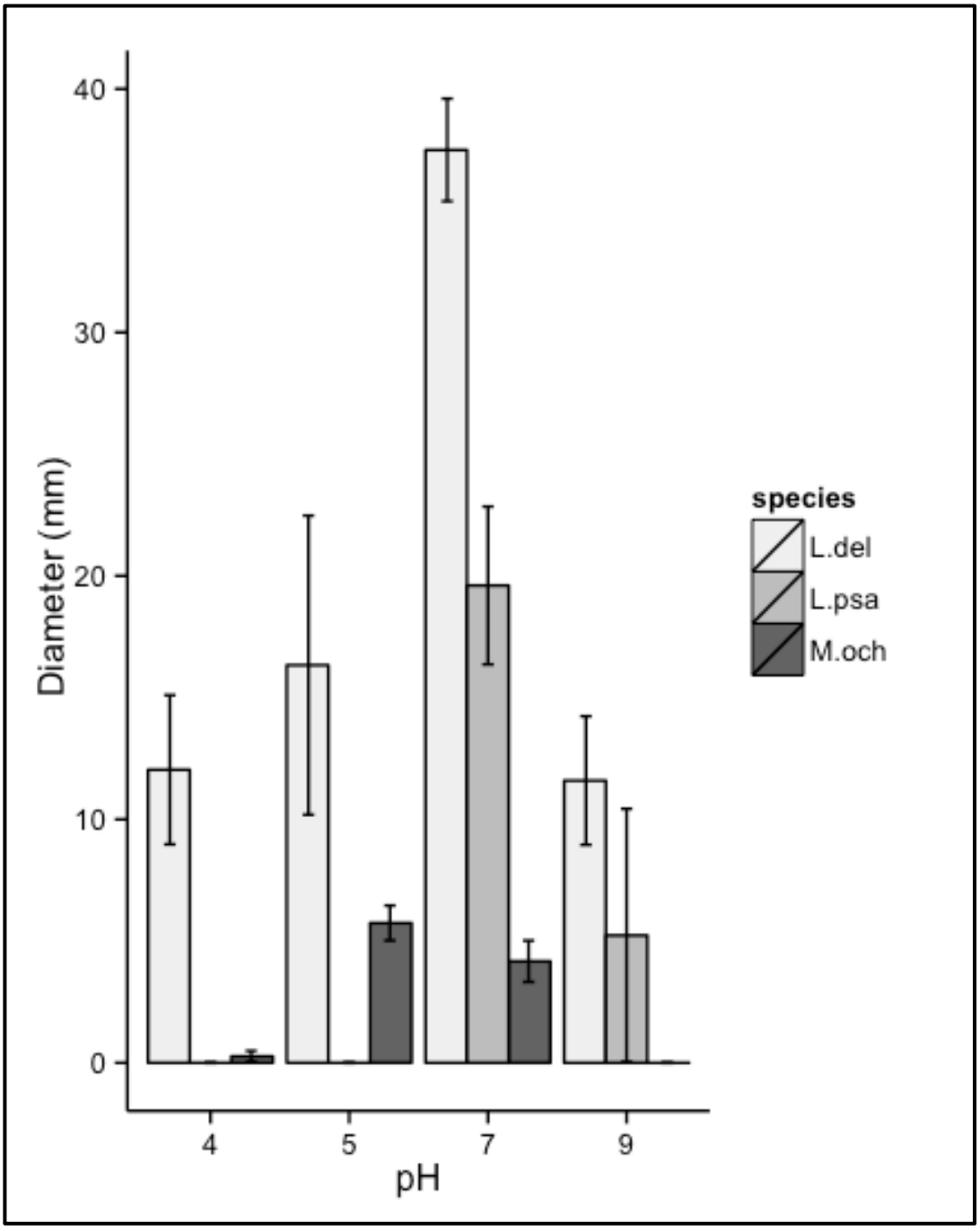


Figure 24. Growth of *L.del* (*Lactarius* cf. *deliciosus, M.och* (*Multifurca ochricompacta*), and *L.psa* (*Lactarius psammicola*) in terms of colony diameter (mm) in MMN in response to different pH treatments. Each data point is a mean of three replicates with a standard deviation of the mean.

Sesquiterpinoid compounds in Russulaceae are considered part of a chemical defense system that is enzymatically activated by physical trauma to the fungus, which then converts dialdehydes into iso-vellerol or vellerol for deterring mycophagy by both insects and small mammals (Sterner *et al.*, 1985b; Bergman *et al.*, 1990; Daniewski *et al.*, 1993; Hansson *et al.*, 1993). The presence of such a defense system seems to be in contrast to the lifestyle of sequestrate species of Russulaceae that rely on animal vectors for dispersal, however, sequestrate species of *Lactarius* have been observed as having fewer lactifers as their mushroom-forming relatives, suggesting evolutionary loss of this defense system (Eberhardt & Verbeken, 2004). Some sesquiterpines are found across many species of Russulaceae, while many are species specific. Wide phylogenetic sampling of Russulaceae genomes can help determine evolutionary trends in the development of fungal chemical defense systems and may potentially be linked to animal-fungal interactions that may have important implications for spore dispersal.

The molecular basis for the establishment and maintenance of the mycorrhizal symbiosis is only recently being addressed using an –omics approach, which has highlighted some key genetic components involved in this complex interaction. Transcriptomics of the EmF model *Laccaria bicolor* has led to the discovery of mycorrhizal small secreted proteins (MiSSPs), which are effectors responsible for shutting down plant defenses and facilitating mycorrhizal formation (Martin *et al.*, 2008). Another set of effectors implicated in mycorrhizal establishment are auxin signaling molecules that alter host auxin metabolism to promote lateral root proliferation and arrest root meristem growth to produce short roots (Vayssieres *et al.*, 2015). Sesquiterpenes have also been implicated in lateral root proliferation before colonization, an effect hypothesized to improve plant nutrient uptake and plant exudate allocation to EmF (Ditengou *et al.*, 2015). The presence of MiSSPs and other effectors evolved in mycorrhizal symbiosis have not been investigated in Russulaceae. It will be important to determine whether these model system findings are recapitulated in Russulaceae or whether novel mechanisms for symbiosis establishment and maintenance have arisen in this lineage.

Many oxidative enzymes, including lignin peroxidases (Chen *et al.*, 2001; Bödeker *et al.*, 2009) and laccases (Gregg & Miller, 1940; Chen *et al.*, 2003) have been identified from members of Russulaceae, indicating a probable white-rot ancestry and potential to mobilize nutrients from lignin. Lignin peroxidase primers developed from *Phanerochaete chrysosporium* were used in a screening of 44 EmF species (Chen *et al.*, 2001). These authors demonstrated that three of four Russulaceae species contained lignin peroxidases (Chen *et al.*, 2001). In another study using degenerate primers targeting class II peroxidases, Russulaceae made up 1/3 of the species containing these genes with *Cortinarius*, another very diverse EmF lineage, actually showing evidence of gene duplication (Bödeker *et al.*, 2009). In the case of laccases, other than one amplicon from *Rhizopogon,* only ectomycorrhizal members of Russulaceae and Atheliaceae were found to contain either *lac2* or *lac3* gene, indicating that

Russulaceae might be specialized in laccase production for oxidation and extraction of nutrients (Chen *et al.*, 2003). Coincidentally, both of these groups were found to be the transcriptionally dominant groups in the soil of a natural loblolly pine system (Liao *et al.*, 2014). In fact, Russulaceae were found to be the dominant producers of laccases in a temperate forest system, showing more vertical stratification of expression and a having a higher gene diversity than even saprotrophic fungi (Luis *et al.*, 2005). In a specific study on *Lactarius quietus,* one of the sequenced Russulaceae genomes, secretion of extracellular enzymes, including laccase, correlated with the bud break of oak trees, where the fungus may be supplying carbon to the sapling and switching to oxidative reactions as the sapling leafs out (Courty *et al.*, 2007). According to Kohler et al. (2015) the gene composition of EmF lineages was found to be variable but held to a pattern of convergent losses of peroxidase genes. Within the Russulaceae we expect to see less variation in gene copy number within families of oxidoreductases and carbohydrate active enzymes (CAZys; www.cazy.org) then found between different lineages of EmF (Kohler *et al.*, 2015) and saprotrophic fungi (Eastwood *et al.*, 2011). We hypothesize that gene duplication of oxidative enzymes like peroxidases and laccases will be present in the genomes of Russulaceae and will be evidence of specialization of function showing variation in conservation of these genes at the species, genus, and family level.

Russulaceae have been an experimental focus for macrochemical tests used in systematics and taxonomy, reinforcing the idea that this group is biochemically diverse. A shared trait of Russulaceae is the amyloid reaction of Melzer's reagent to a layer of the spore wall, indicating the deposition of amylose missing in many other groups of fungi (Miller, 1988b). An early chemical test developed by Bourquelot (1896) to test for the presence and activity of oxidases gives a positive blue reaction in all *Russula* fruitbodies at variable time intervals for different species. The unique conducting system of Russulaceae often reacts to sulfobenzaldehyde or sulfovanilline, turning dark blue to black (Singer 1986). The application of iron salts to *Russula* turns the flesh typically orange, but in some groups, a bright blue-green reaction can be seen. Benzaldahydes, such as para-dimethylamino-benzaldehyde (PDAB) and sulfobenzaldehyde, have been developed as stains and macrochemical tests in Russulaceae, turning different kinds of cystidia black and causing a metachromatic reaction with the flesh of different species (Singer, 1986). These two compounds are thought to detect the presence of different classes of indole alkaloids. While all of these chemical reactions have been developed as diagnostic characters, they are evidence of a high level of bioactivity in Russulaceae sporocarps that may hold the key to understanding their potential functional complexity. With the sequenced genomes we may be able to infer which particular gene families are responsible for coding for the different enzymes, which would allow the chemical tests to be diagnostic for differential expression of these genes.

The recent release of the first set of genome sequences and transcriptomes from diverse Russulaceae will not only expand ongoing research on the functional diversity of those species but will enhance the value of other

fungal sequences through comparative studies of genome evolution, structure, and metabolic pathways and achieving a better insight into the symbiotic lifestyle. Six of the fourteen dikaryotic genomes are completed to date and range from ~40 to 116 Mbps respectively for *Russula dissimulans* to *Lactarius quietus*, revealing a very diverse genome size and suite of genes. A search for peroxidases in *L. quietus* yields 88 predicted genes, which include lignin peroxidases, haem peroxidases, peroxiredoxins, and thioredoxins. There are also 38 laccase gene models predicted in this genome as well. A diverse array of 16 sesquiterpene gene models is also recovered. We predict that Russulaceae genomes with variable genome sizes will harbor variation in gene copy number, indicating a lineage of diverse functions and gene history.

Conclusions and Future Prospective

We expect Russulaceae fulfill important roles in large-scale ecosystem processes as plant associates and nutrient cyclers, but before we can begin to understand their cumulative affects, we must first characterize representative species in isolation. The focus of the RGI is to take an evolutionary approach to better understand the functional diversity of this single lineage of EmF. Are specific functional traits conserved between clades of the family or is functional differentiation a driver of their diversification? Do Russulaceae have a particular functional niche in forested ecosystems and can this be characterized in the pan-genome? How have functional roles in Russulaceae evolved and in what geographic and environmental context did this occur? Given their functional evolution can we predict how changing climates may affect their impact on ecosystem functioning? This dataset will provide a rich context for exploring a number of large biological questions spanning the gamut of these fungi's varied lifestyles and their shared evolutionary history. It is hoped that the mycological community as a whole will consider utilizing this data to develop Russulaceae as a useful model group.

Acknowledgments

We thank Dr. Jorinde Nutinck for consultation on milk cap phylogenetics and species identities. We would also like to acknowledge NSF DEB-1501293 for research support for BPL. This work was supported by the Plant-Microbe Interfaces Scientific Focus Area in the Genomic Science Program, the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the United States Department of Energy under contract DE-AC0500OR22725. The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the US Department of Energy under contract number DE-AC02-05CH11231. This material is also based upon work supported by the U.S. Department of Energy, Office of Science, Office of

Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE-SC0014664.

References

Aanen DK, Eggleton P, Rouland-Lefevre C, Guldberg-Froslev T, Rosendahl S, Boomsma JJ (2002) The evolution of fungus-growing termites and their mutualistic fungal symbionts. Proceedings of the National Academy of Sciences of the United States of America, 99, 14887–14892.

Adamčík S, Caboň M, Eberhardt U, Saba M, Hampe F, Slovák M, Kleine J, Marxmüller H, Jančovičová S, Pfister DH, et al. (2016a) A molecular analysis reveals hidden species diversity within the current concept of *Russula maculata* (Russulaceae, Basidiomycota). Phytotaxa, 270, 71.

Adamčík S, Slovák M, Eberhardt U, Ronikier A, Jairus T, Hampe F, Verbeken A (2016b) Molecular inference, multivariate morphometrics and ecological assessment are applied in concert to delimit species in the *Russula clavipes* complex. Mycologia, 108, 716–730.

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.

Agerer R, Schloter M, Hahn C, et al. (2000) Fungal enzymatic activity in fruitbodies. Nova Hedwigia, 71, 315–336.

Allison SD, Czimczik CI, Treseder KK (2008) Microbial activity and soil respiration under nitrogen addition in Alaskan boreal forest. Global Change Biology 14: 1156–1168.

Anke H, Bergendorff O, Sterner O (1989) Assays of the Biological Activities of Guaiane Sesquiterpenoids Isolated from the Fruit Bodies of Edible *Lactarius* Species. Food and Chemical Toxicity, 27, 393–397.

Aspray TJ, Jones EE, Davies MW, Shipman M, Bending GD (2013) Increased hyphal branching and growth of ectomycorrhizal fungus Lactarius rufus by the helper bacterium *Paenibacillus* sp. Mycorrhiza, 23, 403–410.

Avis PG, McLaughlin DJ, Dentinger BC, Reich PB (2003) Long-term increase in nitrogen supply alters above- and below-ground ectomycorrhizal communities and increases the dominance of *Russula* spp. in a temperate oak savanna. New Phytologist, 160, 239–253.

Bahram M, Põlme S, Kõljalg U, Tedersoo L (2011) A single European aspen (Populus tremula) tree individual may potentially harbour dozens of Cenococcum geophilum ITS genotypes and hundreds of species of ectomycorrhizal fungi. FEMS Microbiology Ecology, 75, 313–320.

Barriuso J, Pereyra MT, García JAL, Megías M, Mañero FJG, Ramos B (2005) Screening for putative PGPR to improve establishment of the symbiosis *Lactarius deliciosus*-*Pinus* sp. Microbial Ecology, 50, 82–89.

Bergemann SE, Garbelotto M (2006) High diversity of fungi recovered from the roots of mature tanoak (*Lithocarpus densiflorus*) in northern California. Canadian Journal of Botany, 84, 1380–1394.

Bergendorff O, Sterner O (1988) The sesquiterpenes of *Lactarius deliciosus* and *Lactarius deterrimus*. Phytochemistry, 27, 97–100.

Bergman R, Hansson T, Sterner O, Wickberg B (1990) A Total Synthesis of (+)-Isovelleral. The Absolute Configuration of the Russulaceae Sesquiterpenes. Journal of the Chemical Society, Chemical Communications, 865–867.

Bidartondo MI, Bruns TD (2001) Extreme specificity in epiparasitic Monotropoideae (Ericaceae): Widespread phylogenetic and geographical structure. Molecular Ecology, 10, 2285–2295.

Bidartondo MI, Bruns TD (2005) On the origins of extreme mycorrhizal specificity in the Monotropoideae (Ericaceae): Performance trade-offs during seed germination and seedling development. Molecular Ecology, 14, 1549–1560.

Bödeker ITM, Nygren CMR, Taylor AFS, Olson A, Lindahl BD. 2009. Class II peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. The ISME journal 3: 1387–95.

Bourquelot É (1896) Actions successives d'un ferment soluble hydratant et d'un ferment soludble oxydant. Comptes Rendus de la Société de Biologie, 48, 314.

Bowen GD, Theodorou C (1979) Interactions between bacteria and ectomycorrhizal fungi. Soil Biology and Biochemistry, 11, 119-126.

Brundrett M (2004) Diversity and classification of mycorrhizal associations. Biological Reviews, 79, 473–495.

Burke DJ, López-Gutiérrez JC, Smemo KA, Chan CR (2009) Vegetation and soil environment influence the spatial distribution of root-associated fungi in a mature beech-maple forest. Applied and Environmental Microbiology, 75, 7639–7648.

Busby PE, Ridout M, Newcombe G (2016) Fungal endophytes: modifiers of plant disease. Plant Molecular Biology, 90, 645–655.

Buyck B (2007) A new initiative towards the study of *Russula* in the eastern USA. Pagine Di Micologia, 27, 81–86.

Buyck B, Atri NS (2011) A *Russula* (Basidiomycota, Russulales) with an Unprecedented Hymenophore Configuration from Northwest Himalaya (India). Cryptogamie, Mycologie, 32, 185–190.

Buyck B, Hofstetter V, Eberhardt U, Verbeken A, Kauff F (2008) Walking the thin line between *Russula* and *Lactarius*: The dilemma of *Russula* subsect. *Ochricompactae*. Fungal Diversity, 28, 15–40.

Buyck B, Horak E (1999) New Taxa of Pleurotoid Russulaceae. Mycologia, 91, 532–537.

Buyck B, Thoen D, Watling R (1996) Ectomycorrhizal fungi of the Guinea--Congo region. Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences, 104, 313–333.

Chen DM, Bastias BA, Taylor AFS, Cairney JWG (2003) Identification of laccase-like genes in ectomycorrhizal basidiomycetes and transcriptional regulation by nitrogen in *Piloderma byssinum*. New Phytologist, 157, 547–554.

Chen DM, Taylor AFS, Burke RM, Cairney JWG (2001) Identification of genes for lignin peroxidases and manganese peroxidases in ectomycorrhizal fungi. New Phytologist, 152, 151–158.

Clay K (1988) Fungal Endophytes of Grasses : A Defensive Mutualism between Plants and Fungi. Ecology, 69, 10-16.

Clericuzio M, Cassino C, Corana F, Vidari G (2012) Terpenoids from *Russula lepida* and *R. amarissima* (Basidiomycota, Russulaceae). Phytochemistry, 84, 154–159.

Courty PE, Breda N, Garbaye J. 2007. Relation between oak tree phenology and the secretion of organic matter degrading enzymes by *Lactarius quietus* ectomycorrhizas before and during bud break. Soil Biology and Biochemistry, 39, 1655–1663.

Courty PE, Franc A, Pierrat JC, Garbaye J (2008) Temporal changes in the ectomycorrhizal community in two soil horizons of a temperate oak forest. Applied and Environmental Microbiology, 74, 5792–5801.

De Crop E, Nuytinck J, Van de Putte K, Wisitrassameewong K, Häckel J, Stubbe D, Hyde KD, Roy M, Halling RE, Wang X-H, et al. (2016) A multi-gene phylogeny of *Lactifluus* (Basidiomycota, Russulales) translated into a new infrageneric classification of the genus. Persoonia, 58–80.

Currie CR, Scott J a., Summerbell RC, Malloch D (1999) Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. Nature, 398, 701–704.

Daniewski WM, Gumulka M, Ptaszynska K, Skibicki P, Bloszyk E, Drozdz B, Stromberg S, Norin T, Holub M (1993) Antifeedant activity of some sesquiterpenoids of the genus *Lactarius* (Agaricales: Russulaceae). European Journal of Entomology, 90, 65–70.

Dearnaley JDW, Martos F, Selosse M (2012) Orchid Mycorrhizas: Molecular Ecology, Physiology, Evolution and Conservation Aspects. Fungal Associations, 216.

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.

Van Dorp CH, Beiler KJ, Durall DM (2016) Dominance of a *Rhizopogon* sister species corresponds to forest age structure. Mycorrhiza, 26, 169–175.

Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, Bendiksby M, et al. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science, 333, 762–765.

Eberhardt U, Verbeken A (2004) Sequestrate *Lactarius* species from tropical Africa: *L. angiocarpus* sp. nov. and *L. dolichocaulis* comb. nov. Mycological Research, 108, 1042–1052.

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 from 31 Fungal Genomes. Science, 336, 1715–1719.

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

Garbaye J, Bowen GD (1989) Stimulation of ectomycorrhizal infection of *Pinus radiata* by some microorganisms associated with the mantle of ectomycorrhizas. New Phytologist, 112, 383-388.

Geml J, Laursen GA, Herriott IC, McFarland JM, Booth MG, Lennon N, Chad Nusbaum H, Lee Taylor D (2010) Phylogenetic and ecological analyses of soil and sporocarp DNA sequences reveal high diversity and strong habitat partitioning in the boreal ectomycorrhizal genus *Russula* (Russulales; Basidiomycota). New Phytologist, 187, 494–507.

Geml J, Laursen GA, Timling I, McFarland JM, Booth MG, Lennon N, Nusbaum C, Taylor DL (2009) Molecular phylogenetic biodiversity assessment of arctic and boreal ectomycorrhizal *Lactarius* Pers. (Russulales; Basidiomycota) in Alaska, based on soil and sporocarp DNA. Molecular Ecology, 18, 2213–2227.

Geml J, Timling I, Robinson CH, Lennon N, Nusbaum HC, Brochmann C, Noordeloos ME, Taylor DL (2012) An arctic community of symbiotic fungi assembled by long-distance dispersers: Phylogenetic diversity of ectomycorrhizal basidiomycetes in Svalbard based on soil and sporocarp DNA. Journal of Biogeography, 39, 74–88.

Geml J, Tulloss RE, Laursen GA, Sazanova NA, Taylor DL (2008) Evidence for strong inter- and intracontinental phylogeographic structure in *Amanita muscaria*, a wind-dispersed ectomycorrhizal basidiomycete. Molecular Phylogenetics and Evolution, 48, 694–701.

Giltrap NJ (1982) Production of polyphenol oxidases by ectomycorrhizal fungi with special reference to *Lactarius* spp. Transactions of the British Mycological Society, 78, 75-81.

Girlanda M, Selosse MA, Cafasso D, Brilli F, Delfine S, Fabbian R, Ghignone S, Pinelli P, Segreto R, Loreto F, et al. (2006) Inefficient photosynthesis in the Mediterranean orchid *Limodorum abortivum* is mirrored by specific association to ectomycorrhizal Russulaceae. Molecular Ecology, 15, 491–504.

Gregg C, Miller H (1940) A Laccase from the Wild Mushroom, *Russula foetens*. Journal of the American Chemical Society, 62, 1374–1379.

Grigoriev I V., 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, 1–6.

Guerin-Laguette A, Plassard C, Mousain D (2000) Effects of experimental conditions on mycorrhizal relationships between *Pinus sylvestris* and *Lactarius deliciosus* and unprecedented fruit-body formation of the Saffron milk cap under controlled soilless conditions. Canadian Journal of Microbiology, 46, 790–799.

Hansson T, Pang Z, Sterner O (1993) The Conversion of [12-2H3]-Labelled Velutinal in Injured Fruit Bodies of *Lactarius vellereus*. Further Insight into the Biosynthesis of the Russulaceae Sesquiterpenes. Acta Chemica Scandinavica, 47, 403–405.

Haug I, Weiß M, Homeier J, Oberwinkler F, Kottke I (2005) Russulaceae and Thelephoraceae form ectomycorrhizas with members of the Nyctaginaceae

(Caryophyllales) in the tropical mountain rain forest of southern Ecuador. New Phytologist ,165, 923–936.

Hawksworth DL (1988) The variety of fungal-algal symbioses, their evolutionary significance, and the nature of lichens. Botanical Journal of the Linnean Society, 96, 3–20.

Henkel TW, Aime MC, Chin MML, Miller SL, Vilgalys R, Smith ME (2012) Ectomycorrhizal fungal sporocarp diversity and discovery of new taxa in *Dicymbe* monodominant forests of the Guiana Shield. Biodiversity and Conservation, 21, 2195–2220.

Henkel TW, Aime MC, Miller SL, Henkel TW, Aime MC (2000) Systematics of Pleurotoid Russulaceae from Guyana and Japan, with Notes on Their Ectomycorrhizal Status. Mycologia, 92, 1119–1132.

Hibbett DS, Stajich JE, Spatafora JW (2013) Toward genome-enabled mycology. Mycologia,105, 1339–1349.

Hobbie EA, Agerer R(2010) Nitrogen isotopes in ectomycorrhizal sporocarps correspond to belowground exploration types. Plant and Soil, 327, 71–83.

Hobbie EA, Weber NS, Trappe JM (2001) Mycorrhizal vs saprotrophic status of fungi: The isotopic evidence. New Phytologist, 150, 601–610.

Hutchison LJ (1990a) Studies on the systematics of ectomycorrhizal fungi in axenic culture: V. Linear growth response to standard extreme temperatures used as a taxonomic character. Canadian Journal Of Botany, 68, 2179–2184.

Hutchison LJ (1990b) Studies on the systematics of ectomycorrhizal fungi in axenic culture. II. The enzymatic degradation of selected carbon and nitrogen compounds. Canadian Journal of Botany, 68, 1522-1530.

Hutchison LJ (1999) *Lactarius*. In Ectomycorrhizal Fungi Key Genera in Profile (pp. 269-285). Springer Berlin Heidelberg.

James TY, Sun S, Li W, Heitman J, Kuo H-C, Lee Y-H, Asiegbu FO, Olson A (2013) Polyporales genomes reveal the genetic architecture underlying tetrapolar and bipolar mating systems. Mycologia, 105, 1374–90.

Kobayashi DY, Crouch JA (2009) Bacterial/Fungal interactions: from pathogens to mutualistic endosymbionts. Annual review of phytopathology, 47, 63–82.

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, 410–5.

Koide RT, Shumway DL, Xu B, Sharda JN (2007) On temporal partitioning of a community of ectomycorrhizal fungi. New Phytologist, 174, 420–429.

Kropp BR (1982) Fungi from decayed wood as ectomycorrhizal symbionts of western hemlock. Canadian Journal of Forest Research, 21, 36–39.

Lamb, RJ (1974) Effect of D-glucose on utilization of single carbon sources by ectomycorrhizal fungi. Transactions of the British Mycological Society, 63, 295-306.

Larsson E, Larsson K-H (2003) Phylogenetic relationships of russuloid basidiomycetes with emphasis on aphyllophoralean taxa. Mycologia, 95, 1037–1065.

Lebel T, Dunk CW, May TW (2013) Rediscovery of *Multifurca stenophylla* (Berk.) T.Lebel, C.W.Dunk & T.W.May comb. nov. (Russulaceae) from Australia. Mycological Progress, 12, 497–504.

Liao HL, 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.

Lilleskov EA, Hobbie EA, Fahey TJ (2002) Ectomycorrhizal fungal taxa differing in response to nitrogen deposition also differ in pure culture organic nitrogen use and natural abundance of nitrogen isotopes. New Phytologist, 154, 219–231.

Lindahl BD, Ihrmark K, Boberg J, Trumbore SE, Högberg P, Stenlid J, Finlay RD (2007) Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. New Phytologist, 173, 611–620.

Lindahl BD, Tunlid A (2014) Ectomycorrhizal fungi – potential organic matter decomposers, yet not saprotrophs. New Phytologist, 205, 1443–1447.

Lindeberg G (1948) On the occurrence of polyphenol oxidases in soil-inhabiting basidiomycetes. Physiologia Plantarum, 1, 196-205.

Looney BP, Ryberg M, Hampe F, Sánchez-García M, Matheny PB (2016) Into and out of the tropics: Global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi. Molecular Ecology, 25, 630–647.

Luis P, Kellner H, Zimdars B, Langer U, Martin F, Buscot F (2005) Patchiness and spatial distribution of laccase genes of ectomycorrhizal, saprotrophic, and unknown basidiomycetes in the upper horizons of a mixed forest cambisol. Microbial Ecology, 50, 570–579.

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, Cullen D, Hibbett DS, Pisabarro A, Spatafora JW, Baker SE, Grigoriev IV (2011) Sequencing the fungal tree of life. New Phytologist, 190, 818–821.

Melin E (1924) Uber den Einfluss der Wasserstoffionenkonzentration auf die Virulenz der Wurzelpilze von Kiefer und Fichte. Bot. Not., 3.

Mgbeahuruike AC, Kovalchuk A, Asiegbu FO (2013) Comparative genomics and evolutionary analysis of hydrophobins from three species of wood-degrading fungi. Mycologia, 105, 1471–8.

Miller SL (1988a) Early Basidiospore Formation in *Lactarius lignyotellus*. Mycologia, 80, 99–107.

Miller S. 1988b. A systematic evaluation of basidiospore symmetry and tegumentation in hypogeous and gasteroid Russulales. Canadian journal of botany, 66, 2561–2573.

Miller SL, Larsson E, Larsson K-H, Verbeken A, Nuytinck J (2006) Perspectives in the new Russulales. Mycologia, 98, 960–970.

Miller SL, Mcclean TM, Walker JF, Buyck B, Mcclean TM, Walker JF, Buyck B (2001) A molecular phylogeny of the Russulales including agaricoid , gasteroid and pleurotoid taxa. Mycologia, 93, 344–354.

Nowrousian M (2014) Genomics and Transcriptomics to Analyze Fruiting Body Development. In: Esser K, ed. Fungal Genomics. Berlin: Springer, 149–172.

Nygren CMR, Eberhardt U, Karlsson M, Parrent JL, Lindahl BD, Taylor AFS (2008) Growth on nitrate and occurrence of nitrate reductase-encoding genes in a phylogenetically diverse range of ectomycorrhizal fungi. New Phytologist, 180, 875–889.

Olson Å, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canbäck B, Coutinho PM, Cullen D, Dalman K, et al. (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytologist, 194, 1001–1013.

Oort AJP (1981) Nutritional requirements of *Lactarius* species, and cultural characters in relation to taxonomy. North-Holland Publishing Co., Amsterdam.

Partida-Martinez LP, Monajembashi S, Greulich KO, Hertweck C (2007) Endosymbiont-Dependent Host Reproduction Maintains Bacterial-Fungal Mutualism. Current Biology, 17, 773–777.

Peay KG, Garbelotto M, Bruns TD (2010a) Evidence of dispersal limitation in soil microorganisms: Isolation reduces species richness on mycorrhizal tree islands. Ecology, 91, 3631–3640.

Peay KG, Kennedy PG, Bruns TD (2011) Rethinking ectomycorrhizal succession: are root density and hyphal exploration types drivers of spatial and temporal zonation? Fungal Ecology, 4, 233–240.

Peay KG, Kennedy PG, Davies SJ, Tan S, Bruns TD (2010b) Potential link between plant and fungal distributions in a dipterocarp rainforest: Community and phylogenetic structure of tropical ectomycorrhizal fungi across a plant and soil ecotone. New Phytologist, 185, 529–542.

Phillips LA, Ward V, Jones MD (2014) Ectomycorrhizal fungi contribute to soil organic matter cycling in sub-boreal forests. The ISME Journal, 8, 699–713.

Plett JM, Martin F (2011) Blurred boundaries: Lifestyle lessons from ectomycorrhizal fungal genomes. Trends in Genetics, 27, 14–22.

Põlme S, Bahram M, Yamanaka T, Nara K, Dai YC, Grebenc T, Kraigher H, Toivonen M, Wang PH, Matsuda Y, et al. (2013) Biogeography of ectomycorrhizal fungi associated with alders (*Alnus* spp.) in relation to biotic and abiotic variables at the global scale. New Phytologist, 198, 1239–1249.

Poole EJ, Bending GD, Whipps JM, Read DJ (2001) Bacteria associated with *Pinus sylvestris-Lactarius rufus* ectomycorrhizas and their effects on mycorrhiza formation in vitro. New Phytologist, 151, 743–751.

Pressel S, Bidartondo MI, Ligrone R, Duckett JG (2010) Fungal symbioses in bryophytes : New insights in the Twenty First Century. Phytotaxa, 9, 238–253.

Riley R, Salamov AA, Brown W, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Lindquist EA, et al. (2014) Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot / brown-rot paradigm for wood decay fungi. Proceedings of the National Academy of Sciences of the United States of America, 111, 9923–9928.

Rineau F, Roth D, Shah F, Smits M, Johansson T, Canbäck B, Olsen PB, Persson P, Grell MN, Lindquist E, et al. (2012) The ectomycorrhizal fungus Paxillus involutus converts organic matter in plant litter using a trimmed brown-rot mechanism involving Fenton chemistry. Environmental Microbiology, 14, 1477–1487.

Roberts C, Ceska O, Kroeger P, Kendrick B (2004) Macrofungi from six habitats over five years in Clayoquot Sound, Vancouver Island. Canadian Journal of Botany, 82, 1518–1538.

Rodriguez-Romero J, Hedtke M, Kastner C, Müller S, & Fischer R (2010) Fungi, hidden in soil or up in the air: light makes a difference. Annual Review of Microbiology, 64, 585-610.

Ryberg M, Matheny PB (2012) Asynchronous origins of ectomycorrhizal clades of Agaricales. Proceedings of the Royal Society B, 279, 2003-2011.

Sanchez F, Honrubia M, Torres P (2001) Effects of pH, water stress and temperature on in vitro cultures of ectomycorrhizal fungi from Mediterranean forests. Cryptogamie Mycologie, 22, 243-258.

Sánchez-García M, Henkel TW, Aime MC, Smith ME, Matheny PB (2016) *Guyanagarika*, a new ectomycorrhizal genus of Agaricales from the Neotropics. Fungal Biology, 120, 1540-1553.

Sánchez-Ramírez S, Etienne RS, Moncalvo JM (2015) High speciation rate at temperate latitudes explains unusual diversity gradients in a clade of ectomycorrhizal fungi. Evolution, 69, 2196–2209.

Singer R (1986) The Agaricales in Modern Taxonomy. Koenigstein, Germany: Koeltz Scientific Books.

Smith SE, Read DJ (2010) Mycorrhizal symbiosis. Academic press.

Sterner O, Bergman R, Fanzen C, Wickberg B (1985a) New Sesquiterpenes in a Proposed Russulaceae Chemical Defense System. Tetrahedron Letters, 26, 3163–3166.

Sterner O, Bergman R, Kihlberg J, Wickberg B (1985b) The sequiterpenes of *Lactarius vellereus* and their role in a proposed chemical defense system. Journal of Natural Products, 48, 279–288.

Talbot JM, Allison SD, Treseder KK (2008) Decomposers in disguise: Mycorrhizal fungi as regulators of soil C dynamics in ecosystems under global change. Functional Ecology, 22, 955–963.

Talbot JM, Bruns TD, Smith DP, Branco S, Glassman SI, Erlandson S, Vilgalys R, Peay KG (2013) Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. Soil Biology and Biochemistry, 57, 282–291.

Talbot JM, Bruns TD, Taylor JW, Smith DP, Branco S, Glassman SI, Erlandson S, Vilgalys R, Liao H, Smith ME, et al. (2014) Endemism and

functional convergence across the North American soil mycobiome. Proceedings of the National Academy of Sciences of the United States of America, 111, 6341–6.

Tedersoo L, Bahram M, Toots M, Diédhiou AG, Henkel TW, Kjoller R, Morris MH, Nara K, Nouhra E, Peay KG, et al. (2012) Towards global patterns in the diversity and community structure of ectomycorrhizal fungi. Molecular Ecology, 21, 4160–4170.

Tedersoo L, Põlme S (2012) Infrageneric variation in partner specificity: Multiple ectomycorrhizal symbionts associate with *Gnetum gnemon* (Gnetophyta) in Papua New Guinea. Mycorrhiza, 22, 663–668.

Tedersoo L, Smith ME (2013) Lineages of ectomycorrhizal fungi revisited: Foraging strategies and novel lineages revealed by sequences from belowground. Fungal Biology Reviews, 27, 83–99.

Trappe JM (2005) A.B. Frank and mycorrhizae: The challenge to evolutionary and ecologic theory. Mycorrhiza, 15, 277–281.

Treseder KK, Allen MF (2000) Mycorrhizal fungi have a potential role in soil carbon storage under elevated CO and nitrogen deposition. New Phytologist, 147, 189–200.

Twieg BD, Durall DM, Simard SW (2007) Ectomycorrhizal fungal succession in mixed temperate forests. New Phytologist, 176, 437–447.

Vayssieres A, Pencik A, Felten J, Kohler A, Ljung K, Martin F, Legue V (2015) Development of the Poplar-*Laccaria bicolor* Ectomycorrhiza Modifies Root Auxin Metabolism, Signaling, and Response. Plant Physiology, 169, 890–902.

Verbeken A, Nuytinck J (2013) Not every milkcap is a *Lactarius*. Scripta Botanica Belgica, 51, 162–168.

Verbeken A, Nuytinck J, Buyck B (2011) New combinations in *Lactifluus*. 1. *L*. subgenera *Edules*, *Lactariopsis*, and *Russulopsis*. Mycotaxon, 118, 447–453.

Verbeken A, Stubbe D, van de Putte K, Eberhardt U, Nuytinck J (2014) Tales of the unexpected: Angiocarpous representatives of the Russulaceae in tropical South East Asia. Persoonia: Molecular Phylogeny and Evolution of Fungi, 32, 13–24.

Wisitrassameewong K, Looney BP, Le HT, De Crop E, Das K, Van de Putte K, Eberhardt U, Jiayu G, Stubbe D, Hyde KD, et al. (2016) *Lactarius* subgenus *Russularia* (Basidiomycota, Russulales): novel Asian species, worldwide phylogeny and evolutionary relationships. Fungal Biology, 120(12), 1554–1581.

Wolfe BE, Tulloss RE, Pringle A (2012) The irreversible loss of a decomposition pathway marks the single origin of an ectomycorrhizal symbiosis. PLoS ONE, 7(7), e39597.

Wu XQ, Hou LL, Sheng JM, Ren JH, Zheng L, Chen D, Ye JR (2012) Effects of ectomycorrhizal fungus *Boletus edulis* and mycorrhiza helper *Bacillus cereus* on the growth and nutrient uptake by *Pinus thunbergii*. Biology and Fertility of Soils, 48, 385–391.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. Science, 292, 686-693.

Zhou L-W, Dai Y-C (2013) Taxonomy and phylogeny of wood-inhabiting hydnoid species in Russulales: two new genera, three new species and two new combinations. Mycologia, 105, 636–649.

CHAPTER IV
PHYLOGENOMICS AND COMPARATIVE GENOMICS OF RUSSULACEAE

The doctoral student submitted the initial proposal to JGI, collected specimens, extracted DNA and RNA for sequencing, performed phylogenomic analyses, performed gene evolution analyses and wrote the manuscript. P. Meidl assisted with SSR analysis. M.J. Piatek performed pan-genome analysis and assisted doctoral student with data manipulation, D. Weighill performed the Proportional Similarity analysis. JGI sequenced, assembled and annotated all genomes. J. Labbé provided lab space and materials for labwork, assisted in lab work, acted as liason with JGI, and performed TE analysis. P. Matheny provided lab facilities and facilitated field excursions for collecting samples.

Abstract

Russulaceae is a diverse fungal family including the genera *Russula*, *Lactarius*, *Lactifluus*, and *Multifurca*, and is one of the most widespread and species rich ECM lineages. In a recent collaborative effort, the Joint Genome Institute has sequenced genomes and transcriptomes of representative groups across Russulaceae, including a saprotrophic outgroup. Presented here is an overview of the first insight into the dense genome sampling within the family to capture specific genomic features and investigate i) compare genome size and structure within the order Russulales, ii) to examine functional diversity within this ecologically important clade, iii) reconstruct the pan-genome of Russulales to look for genetic patterns associated with trophic mode, and iv) to what extent genes involved in plant biopolymer degradation have been maintained within a single, diverse, ECM lineage. Indeed, preliminary evidence suggests that members of this family, though being mutualists, have retained a restricted set of genes coding for lignin peroxidases and copper oxidoreductases which may be responsible for the degradation of lignin derivatives accumulating in soil organic matter.

Introduction

The age of genome-enabled mycology has arrived with the US Department of Energy (DOE) Joint Genome Institute (JGI) and their Fungal Genomics Program (http://jgi.doe.gov/fungi), which seeks to support community projects for sequencing fungal genomes (Grigoriev et al., 2011; Hibbett et al., 2013). The largest sequencing project to date, the 1000 Fungal Genome Project (1KFGP), set out with the goal to genomically sample the breadth of diversity across Fungi to further resolve the fungal tree of life (Grigoriev et al., 2014). Having made significant progress to this end, this landmark project has now been split into two projects, the Zygomycetes Geneology of Life (ZyGoLife) project (zygolife.org), seeking to sample and reconstruct the early branches of life in Fungi, and the Deep Sequencing of Dikarya project (jgi.doe.gov/deep-sequencing-of-dikarya), which seeks denser genome sampling in groups of ecologically relevant groups. The Agaricomycotina is a diverse subphylum in Dikarya, which includes wood

and litter decomposers (Floudas et al., 2012; Riley et al., 2014), ectomycorrhizal fungi (Kohler et al., 2015; Martin et al., 2010, 2008), plant pathogens (Collins et al., 2013; Olson et al., 2012), mushroom-forming fungi (Morin et al., 2013; Wawrzyn et al., 2012), and many other biologically diverse groups (Hibbett, 2006). This has been a targeted group for dense genome sequencing because of its multitudes of potential ecological, medicinal, and industrial applications (Martin et al., 2011). A majority of sampling in this group has been focused on three orders; Agaricales, Boletales, and Polyporales. Here we present comparative genomics of the diverse order of Russulales, a thus far relatively underexplored clade of Dikarya with one of the most diverse lineages of ectomycorrhizal fungi, Russulaceae.

Russulales is an order characterized by diverse nutritional strategies, sporocarp morphologies, hymenophore conformation, and biochemical activity (Miller et al., 2006). Current taxonomy has Russulales divided into twelve families and eighty genera, for which JGI currently has sampling for 10 taxa representing the families Auriscalpiaceae, Bondarzewiaceae, Hericiaceae, Lachnocladiaceae, Peniophoraceae, Russulaceae, and Stereaceae. This sampling includes three of the four primary nutritional strategies from the order; saprotrophs, ectomycorrhizal, and plant pathogens. Unsampled is the only known animal pathogenic lineage in the order, the genus *Entomocorticium* in Peniophoraceae. Current sampling also covers many sporocarp morphologies, including agaricoid (*Auriscalpium vulgare* and *Lactarius quietus*), bracket-like (*Heterobadisium irregulare* and *Stereum hirsutum*), coralloid (*Artomyces pyxidata*), corticioid (*Dentipellis* sp., *Peniophora* sp., *Peniophora cinereus*, and *Vararia minispora*), and pleurotoid (*Lentinellus vulpinus*). While diverse in these aspects, species richness in Russulales is dominated by a single group, the family Russulaceae. Out of 1750 species of Russulales, about 1250 are species from four genera of Russulaceae; *Russula*, *Lactarius*, *Lactifluus*, and *Multifurca* (Kirk et al., 2008).

Ectomycorrhizal (ECM) fungi have been sequenced through JGI and the Mycorrhizal Genomics Initiative (mycor.nancy.inra.fr/IMGC/MycoGenomes), which seeks to sample broadly among the 80 or so independent lineages of ECM fungi as well as other forms of mycorrhizal associations (ericoid and orchid endomycorrhizas)(Kohler et al., 2015; Tedersoo and Smith, 2013). ECM fungi are biotrophic mutualists of certain trees and shrubs, providing essential trace elements, nutrients, water, and protection against pathogens in exchange for photosynthates from the host. ECM fungi have codiversified with their plant hosts and a single plant host can harbor over 100 different species from a multitude of lineages at any one time (Bahram et al., 2011). While ECM fungi have been thought to have mostly lost the capability to degrade recalcitrant soil organic matter like chitin and lignin, recent studies have been recovered oxidative enzymes capable of decomposition in high abundance from ECM fungi in natural systems (Bodeker et al., 2014; Luis et al., 2005; Talbot et al., 2013). In fact, it has been demonstrated that ECM fungi often retain the decomposition system of their ancestors, whether lignin degrading or not (Rineau et al., 2012; Shah et al., 2016). It is hypothesized that this capability to decompose is directly related to

their ability to scavenge nutrients from the environment (Bodeker et al., 2014). The genes responsible for coding these enzymes, however, are reduced in genomes of ECM fungi across Dikarya (Kohler et al., 2015). While this seems to be fairly conserved across ECM fungi, these are independent lineages and ECM lineages have not been sampled densely to see whether there are conserved patterns within ECM lineages or if functional diversity may be driving ECM diversification.

The first release of the Russulaceae Genome Initiative introduces nine newly sequenced genomes from Russulaceae, including the first samples from the largest genus *Russula* as well as *Lactifluus*, *Multifurca*, and one of the few extant saprotrophic members of Russulaceae, *Gleoeopeniophorella convolvens*. All eighteen representative species from Russulales are combined for comparative analyses across the order to detect trends in genomic architecture and gene content. With sampling across the order and a closely related saprotrophic outgroup to ECM Russulaceae, we seek to test the following hypotheses: 1) Russulaceae have retained the genes necessary to degrade lignin from its saprotrophic ancestor; 2) Russulaceae are united by a conserved genetic toolset derived from its common ancestor; 3) Russulaceae are functionally divergent at the genus level.

Materials and Methods

Taxon sampling and nucleic acid extraction

Newly sequenced genomes and transcriptomes were derived from phylogenetically distinct lineages within the family Russulaceae according to (Looney et al., 2016). Representative species were sampled as mushroom sporocarps from forested habitat in the Great Smoky Mountains National Park and surrounding areas. To retrieve high molecular weight DNA and undegraded RNA, the inner flesh of the sporocarps was extracted in the field using a sterilized scalpel and placed in a 50mg Falcon tube. Material was then flash-frozen in the field using liquid nitrogen. Tissue samples were also attempted on Melin-Norkrans Modified media with collections for experimental applications. A member of the closest related extant outgroup, *Gloeopeniophorella convolvens*, was also sampled for comparative analyses of different trophic modes.

Extraction of high molecular weight DNA was done using a cetyl trimethylammonium bromide (CTAB) based protocol. Frozen sporocarp material was first ground into a fine powder in liquid nitrogen using a mortar and pestle. The powder was added to 1.5 mL Eppendorf tubes weighed at 90 mg increments. A pre-warmed (~55° C) lysis buffer was added to each sample at a volume of 700 µL. The lysis buffer consisted of a mixture of 260 µL of buffer A (0.35 M sorbitol, 0.1 M Tris HCl pH 9, and 5 mM EDTA pH 8), 260 µL of buffer B (0.2 M Tris HCl pH 9, 50 mM EDTA pH 8, 2M NaCl, and 2% CTAB), 104 µL buffer C (5% Sarkosyl [N-lauroylsarcosine sodium salt], and 70 µL of a 0.1% solution of Polyvinylpyrrolidone (PVP). Samples were then centrifuged at 14 000

rpm for 3 minutes to compact rehydrated biomass. A micropestle was used for additional grinding and this process was repeated at least one more time. Protein digestion was performed by adding 5 µL of Proteinase K (10mg/mL), vortexing, and incubation of samples for 30 min. at 65° C. For sodium dodecyl sulfate precipitation, 230 µL of 5 M KAc was added to samples, inverted to mix, and incubated for at least 30 minutes in ice or for 16 hours in a 4° C refrigerator. Following incubation, samples were centrifuged at 14 000 rpm for 10 minutes and 1 mL of supernatant was transferred to 2 mL Eppendorf tubes. An equal volume of Chloroforme:Isoamylalcohol (24:1) was added and the tubes were centrifuged for 10 minutes at 14 000 rpm. A conservative amount of supernatant (~850 µL) was drawn avoiding the top and bottom layers and added to additional 2 mL Eppendorf tubes. Again, an equal amount of Chloroforme:Isoamylalcohol (24:1) was added and centrifuged for 10 minutes. A final volume of 675 µL was added to new 1.5 mL Eppendorf tubes and treated with an RNase digestion with 10 µL of RNAseA (100mg/mL) and incubated at 37° C for 10 minutes. DNA precipitation was done by adding 67.5 µL of 3 M NaAc pH 8 and 675 µL of absolute isopropanol and incubated for 5 minutes at room temperature. The samples were then centrifuged for 10 minutes at 4° C and the surnageant was eliminated by gently pouring it off. Ethanol washing was done with 200 µL 70% ethanol followed by centrifugation at 14 000 rpm at 4° C. Ethanol was then carefully drawn out using a double pipet tip method making sure not to disturb the pellet. Samples were then dried for 5 minutes in a vacuum pump to completely dry the pellet. The pellets were then resuspended in 10 µL of TE buffer and stored at 4° C for quality assessment.

Extraction of RNA was performed using a Sigma™ Plant Total RNA Kit. Surfaces were first sterilized with 70% ETOH and D/RNAse Free™ decontaminant to prevent enzyme contamination. Frozen sporocarp material was again ground into a fine powder in liquid nitrogen using a decontaminated mortar and pestle. The powder was added to enzyme-free 1.5 mL Eppendorf tubes weighed at 100 mg increments. The provided lysis buffer was added to each sample at a volume of 500 µL. Samples were then centrifuged at 14 000 rpm for 3 minutes to compact rehydrated biomass. A micropestle was used for additional grinding and this process was repeated at least one more time. Once samples were sufficiently ground, 5 µL of 2-mercaptoethanol was added to each sample and incubated at 55° C. The rest of the protocol followed the provided protocol of the kit, using Protocol A for the binding step and following the optional On-Column DNase Digestion procedure. Once product was eluted, 1 µL of Roche Protector RNase Inhibitor was added to stabilize the product. An aliquot of 9 µL was stored at 4° C° for quality control and the remaining sample was stored at -80° C.

Quality assessment followed the recommendations of JGI for DNA and RNA. First, nucleic acids were visualized using gel electrophoresis on a 1% agarose gel with Roche DNA Molecular Weight Marker II as ladder. Bands were evaluated based on brightness, amount of smearing, and presence or absence of contamination (i.e. RNA or DNA). Concentrated and undegraded DNA samples

were pooled after centrifugation at low speed for one minute to homogenize and without pumping the pipet. Assessment of concentration and total amount for genomic DNA was assessed using the Qubit® DNA BR Assay Kit on a Qubit® 2.0 fluorometer. Assessment for RNA concentration and quality was done using an Experion™ RNA Analysis kit analyzed using the Experion™ Automated Electrophoresis System. RNA with clear bands that achieved an RQI score of at least 6.5 was deemed adequate for JGI submission.

Genome Sequencing and Annotation

Genomes were sequenced using the PacificBiosciences (PacBio) platform at the Joint Genome Institute (JGI) in Walnut Creek, CA. PacBio >10kb with AMPure Bead Size Selection with 1x240 bp kb was used for representatives from Russulaceae as this method has been shown to result in fewer contigs that are also longer than HiSeq Illumina sequencing. Filtered subread data were assembled using the Falcon ver. 0.4.2 assembler (https://github.com/PacificBiosciences/FALCON) to generate an initial assembly. Mitochondria were assembled separately from the Falcon pre-assembled reads (preads) using an in-house tool (assemblemito.sh), used to filter the preads, and polished with Quiver version smrtanalysis_2.3.0.140936.p5 (https://github.com/PacificBiosciences/GenomicConsensus). A secondary Falcon assembly was generated using the mitochondria-filtered preads with Falcon version 0.4.2, and polished with Quiver version smrtanalysis_2.3.0.140936.p5. Statistics based on 1 N to denote a gap. Contigs less than 1000 bp were excluded. Completeness of the euchromatic portion of the genome assembly was assessed by aligning assembled consensus RNA sequence data with ESTmapper at 90% identity and 85% coverage. This is a routine test to determine whether we are missing significant portions of the genome. Contaminant contigs were identified via BLAST/tetramer analysis/GC/coverage and removed from the assembly prior to release. Contaminant contigs were indicated as ribosomal in origin, suggesting insect contamination might be at low levels.

Transcriptomes were sequenced using the Illumina HiSeq-2500 sequencing platform at the Joint Genome Institute (JGI) in Walnut Creek, CA. Stranded RNASeq library(s) were created and quantified by qPCR. Raw fastq file reads were filtered and trimmed using the JGI QC pipeline resulting in the filtered fastq file (*.filter-RNA.fastq.gz). Using BBDuk (https://sourceforge.net/projects/bbmap/), raw reads were evaluated for artifact sequence by kmer matching (kmer=25), allowing 1 mismatch and detected artifact was trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads and reads containing any Ns were removed. Quality trimming was performed using the phred trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length 25 bases or 1/3 of the original read length - whichever is longer). Assembly for transcriptomes were done de novo. Filtered fastq files were used as input for de

novo assembly of RNA contigs. Reads were assembled into consensus sequences using Trinity (ver. 2.1.1) (Grabherr et al., 2011). Trinity partitions the sequence data into many individual de Bruijn graphs, each representing the transcriptional complexity at at a given gene or locus, and then processes each graph independently to extract full-length splicing isoforms and to tease apart transcripts derived from paralogous genes. Trinity combines three independent software modules: Inchworm, Chrysalis, and Butterfly, applied sequentially to process large volumes of RNA-seq reads. Trinity was run with the --normalize_reads (In-silico normalization routine) and --jaccard_clip (Minimizing fusion transcripts derived from gene dense genomes) options.

Annotation for genomes followed the JGI Annotation Pipeline. This procedure follows the step of gene prediction, functional annotation, and then a comparative analysis. For gene prediction, assembly scaffolds are masked using RepeatMasker (Smit et al., 1996) with the standard RepBase library (Jurka et al., 2005), frequent repeats recognized by Repeat Scout (Price et al., 2005), and using manually curated libraries of transposons when available. Expressed sequence tags (ESTs) generated from transcriptome are mapped to the assembly using the BLAST-Like Alignment Tool (BLAT) and filtered by identity and coverage. Genes were then predicted from the repeat-masked assembly using *ab initio*, homology-based, and EST-based methods like FGENESH (Salamov et al., 2000), GeneMark (Ter-hovhannisyan et al., 2008), FGENESH+ (Salamov et al., 2000), Genewise (Birney et al., 2004), and EST_map (http://www.softberry.com/). To detect or estimate coding or untranslated regions, estExt (I. Gregoriev, unpublished) was used. Predicted proteins are functionally annotated using SignalP (Nielsen and Engelbrecht, 1997) for signal sequences, TMHMM (Melén et al., 2003) for transmembrane domains, InterProtScan (Quevillon et al., 2005) for integrated collection of functional and structure protein domains, NCBI nr, SwissProt (http://www.expasy.org/sprot/), KEGG (Kanehisa et al., 2006), and KOG (Koonin et al., 2004) for eukaryotic clusters of orthologs. Definition lines for each protein were inferred from the top BLASTp protein hit when meeting coverage and e-value thresholds or else it is replaced with 'hypothetical protein'. A number of comparative tools were made available through the MycoCosm workbench (www.genome.jgi.doe.gov/programs/fungi/index.jsf) including genome browsers, interactive dot-plots, synteny analysis tools, classification schemas from annotation databases (e.g. KOG, KEGG, and GO), and a number of other tools for managing, curating, and downloading genomic data.

Phylogenomics and gene family reconstructions

Protein families that were present in all 18 genomes with only a single gene copy were extracted as amino acid fasta files for phylogenetic reconstruction. Each gene was automatically aligned using the E-INS-i alignment strategy in MAFFT version 7 (Katoh and Standley, 2013). All genes were then concatenated using SequenceMatrix ver. 1.8 (Vaidya et al., 2011), leaving gaps coded as gaps to

represent AA gain/loss and not missing information. Phylogenetic reconstruction was performed in raxmlGUI ver. 1.8 (Silvestro and Michalak, 2011) using the PROTGAMMI substitution model and WAG model of protein evolution with 1,000 bootstrap iterations. The resulting tree was visualized in FigTree v. 1.4.0 (Rambaut, 2012) and mid-point rooted along the longest internode.

For gene tree reconstruction, an *in silico* search for representative genes of targeted gene families seeded with annotated gene models from the *Stereum hirsutum* and *Heterobasidion irregulare* genomes was conducted using pBLAST. Amino acid sequences were loaded into AliView v. 1.17.1 (Larsson, 2014) and sequences were automatically aligned using MUSCLE (Edgar, 2004) and manually adjusted. Sequences that were clearly divergent were excluded and ends were trimmed. Gene trees were inferred in raxmlGUI using the PROTGAMMI substitution model and WAG model of protein evolution with 1,000 bootstrap iterations. The *Phanerochaete chrysosporium* genome was selected as an outgroup for Russulales.

Pan-genome assembly and network analyses

All-against-all BLASTP comparisons were performed amongst proteomes of all 18 genomes of Russulales. Based on best reciprocal sequence similarity, putative orthologous relationships between proteome pairs were identified with OrthoMCL (Li et al., 2003). Paralogs were defined as sequences that are reciprocally more similar to each other within the same proteome than to any other proteome. Based on empirical evidence, a p-value cut-off of 1e-5 was applied for putative orthologs or paralogs. Protein families were then counted and the count of the number of proteins in each homologous protein family for each isolate were subsequently binarized and the results graphed as a heat map. Protein families that were only found in one representative species were further binned to infer number of unique protein families for each species.

Gene family profiles from the pan-genome assembly were used in a similarity analysis where similarity was calculated between the profiles for all potential species pairings. Proportional similarity was used as the similarity metric. The resulting similarity relationships were represented as a network in which nodes represent species and edges represent similarity between species. Incremental stringency thresholds were applied, causing the network to fragment. The order of fragmentation was used as indication of the strongest and weakest relationships among species.

To determine specific functional similarities between different trophic modes (ECM, saprotrophic, and switching between saprotrophic and parasitic), a gene network analysis of enrichment for Pfam domains was performed using an upper-tail Fisher's exact test for each species. A Benjamin-Hochberg correction fo multiple hypothesis bias was applied using a false discovery rate (FDR) less than 0.01 as the statistical significance threshold. Networks were visualized in Cytoscape ver. 3.4.0 (Shannon et al., 2003).

Transposable elements and microsatellite analysis

Transposable elements (TE) were annotated using the CENSOR tool (http://www.girinst.org/censor/) in Repbase RMBLR procedure from the TE annotation pipeline described by Quesneville et al. (Quesneville et al., 2005). Consecutive fragments on both the genome and the reference TE were automatically joined if they were separated by a sequence of which more than 80% consisted of other TE insertions. Simple repeats were found using the Tandem Repeat Finder program (Benson, 1999) and used to filter out spurious hits. All TE annotations that were less than 20 bp, after removing any regions that overlapped simple repeat regions, were eliminated. Finally, consensus sequences (complete and incomplete) belonging to various classes or types of TEs were obtained. LTR retrotransposons sequences were confirmed using a second identification procedure based on the program LTR_STRUC (McCarthy and McDonald, 2003). To facilitate manual curation we promoted the consensus sequences identified via RMBLR to a candidate annotation set defined as a set of one or more joined fragments and by using sequence alignments with CLUSTALW (http://www.ebi.ac.uk/clustalw) that were then validated or modified by the curator in Artemis v11. The copy number for each TE type was calculated on the basis of the elements obtained from the TE annotation pipeline, TBLASTN searches and RepeatMasker analysis.

All genomes were scanned for simple sequence repeats (SSR) utilizing the Microsatellite Identification Tool (MISA) (Thiel, 2003). MISA was then used to scan and identify motifs (1–6 bp in length) with the minimum number of repeats for the six classes being set to 10, 6, 5, 5, 5, and 5 for each bp length respectively and the maximal number bases between adjacent microsatellites was set to 100 bps.

Genome size and TE content were compared between trophic modes using the ‘OUwie’ package in R (Beaulieu and O’Meara, 2016). Genome size was first log transformed and trophic mode was scored as 0 and 1. Brownian motion and Ornstein-Uhlenbeck models were compared to test whether rates differed between trophic modes. Differences in AICc were used to determine the best model. Robustness of the best model were analyzed using diagnostic eigenvalues and standard errors of rate estimates.

Results

Phylogenomic reconstruction shows trends in genome size

To date, nine genomes and transcriptomes from Russulaceae have now been sequenced through JGI, with eight being produced specifically for this study. Combined with previous sampling, this makes eighteen currently available genomes for the order Russulales (Table 1). A phylogeny inferred from a 1.67 million amino acid alignment of 2,518 single-copy genes present in all of the Russulales genomes illustrates a topology that is fully supported by 1,000

bootstrap iterations (Figure 25). A number of different nutritional modes are represented in the Russulales sampling, including saprotrophic species, a species that can switch from being a root parasite to saprotrophic (*Heterobasidion irregulare*), and now eight ectomycorrhizal species from Russulaceae. Given the reconstruction of the genomes, Russulaceae are resolved as monophyletic and so are the ECM members of Russulaceae. Under mid-point rooting the sister clade of Russulaceae is resolved as Auriscalpiaceae with *A. vulgare, A. pyxidata,* and *L. vulpinus* as representatives. The Stereaceae is resolved as sister to the Bondarzewiaceae, which is sister to Hericiaceae.

The largest genome size of all Russulales genomes is *Lactarius quietus* with 115.9 MB, which is almost twice as large as any other Russulales genome and the fifth largest sequenced ECM genome after *Tuber melanosporum*, *Tuber borchii*, *Tricholoma matsutake*, and *Cantharellus anzutake* (JGI Mycocosm). Both *Peniophora* species, however, possess the highest number of genes with over 18,000, but *Lactarius quietus* possesses the highest number of exons. *Vararia minispora* possesses the lowest number of genes with 9,397 genes, however, *Multifurca ochricompacta*, despite having a relatively large genome of 54.7 MB, has the second fewest genes with 9,990 genes. Comparison of genome size between ECM Russulaceae and saprotrophic members of Russulales supported ECM genomes as significantly larger ($p=0.04$). No other comparisons of trophic modes were significant. However, the best supporting model for rate differences taking into account phylogeny was an Ornstein-Uhlenbeck process with rates constrained as the same.

Transposable elements and simple sequence repeats analyses

Transposable elements and variable number tandem repeats (VNTRs) were estimated as a proportion of the total genome (Figure 26). Repeated elements make up a significant proportion of Russulales genomes, especially for the ECM Russulaceae. ECM Russulaceae members possess a larger proportion of their genome composed of repeated elements than saprotrophic Russulales ($p<0.001$). *Multifurca ochricompacta* possesses the largest percent of its genome as repeated elements (71%), with *L. quietus* second the second most (61%). VNTR content is highest in *R. compacta* (11%) and *R. brevipes* (11%). The plant pathogen, *Heterobasidion irregulare*, has a higher repeated element content than purely saprotrophic members of Russulales. A majority of the genome made up of repeated elements were TEs, which are those repeated elements with sequences longer than 100 bases. Long terminal repeats (LTRs) made up a large proportion of the TE content for *R. brevipes* (62%), *R. dissimulans* (50%), *R. compacta* (49%), and especially *Auriscalpium vulgare* (87%) but not for other species

Table 7. Genome assembly and annotation statistics. Green indicate the smallest values and red the largest.

Species	ID	Scaffolds	Genes	CDS	Exons	Mean Exon [bp]	Proteins	Genome
Russula compacta	Ruscom1	984	10,288	10,834	62,141	242	10,834	45,678,814
Russula brevipes	Rusbre1	1,320	13,133	14,000	74,096	243	14,000	48,529,157
Russula dissimulans	Rusdis1	489	11,268	11,865	65,780	242	11,865	44,689,478
Russula rugulosa	Rusrug1	1,389	12,453	13,281	72,889	228	13,281	43,452,339
Lactifluus subvellereus	Lacsub1	1,468	12,090	12,802	69,367	243	12,802	44,824,221
Lactifluus cf. *volemus*	Lacvol1	712	14,078	14,970	81,221	229	14,970	62,975,844
Multifurca ochricompacta	Muloch1	772	9,990	10,514	62,625	236	10,514	54,697,930
Lactarius quietus	Lacqui1	2,812	17,677	18,943	104,511	223	18,943	115,901,997
Gloeopeniopherella convolvens	Glocon1	654	12,364	12,904	73,681	247	12,904	37,988,413
Auriscalium vulgare	Aurvu1	1,349	15,739	16,945	88,492	251	16,945	39,647,923
Lentinellus vulpinus	Lenvul1	578	12,877	13,477	75,630	253	13,477	34,711,142
Artomyces pyxidata	Clapy1	477	14,269	15,130	85,828	244	15,130	37,987,417
Stereum hirsutum	Stehi1	159	13,340	14,072	91,696	239	14,072	46,511,623
Heterobasidion irregulare	Hetan2	15	13,383	13,405	72,148	231	13,405	33,649,967
Dentipellis sp.	Densp1	425	13,695	14,320	78,260	262	14,320	36,706,823
Peniophora sp.	Lopni1	217	18,385	18,999	98,980	287	18,999	48,435,708
Peniophora cinereus	Ricme1	1,092	18,084	18,952	94,316	287	18,952	46,030,792
Vararia minispora	Varmin1	1,435	9,397	10,962	63,850	219	10,962	36,812,209

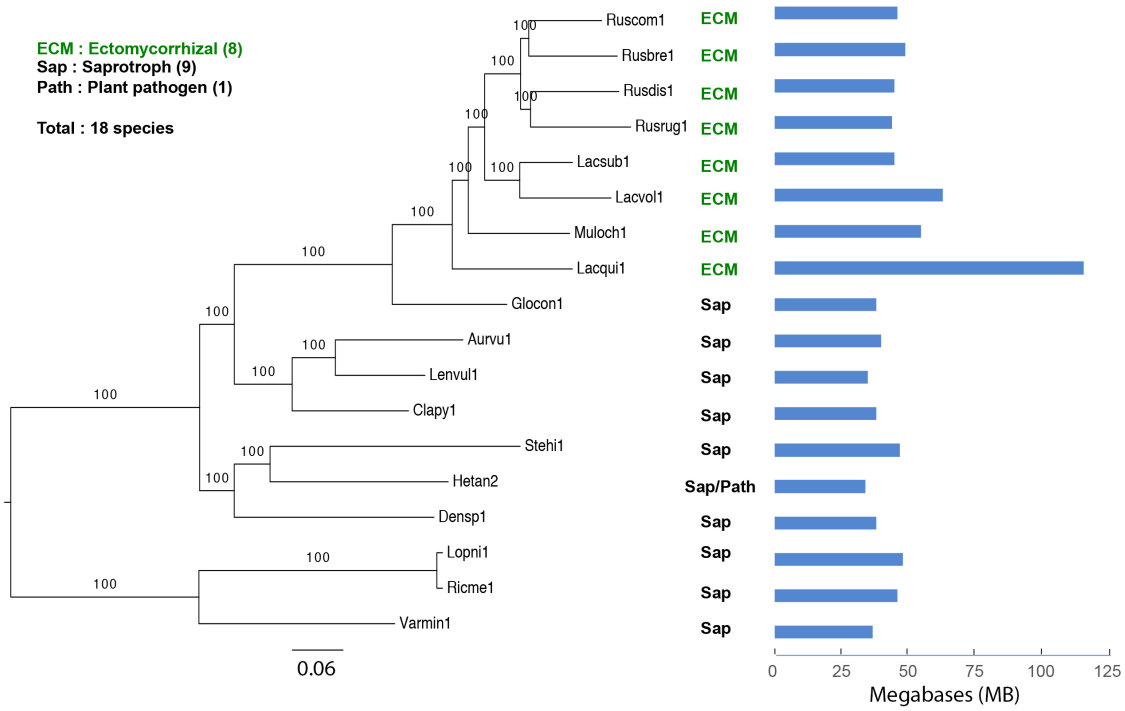


Figure 25. Phylogenetic reconstruction of Russulales genomes using 2,518 single-copy genes in RaXML with 1,000 bootstrap iterations. Taxon labels correspond to JGI identifiers (Ruscom1–*Russula compacta*; Rusbre1–*Russula brevipes*; Rusdis1–*Russula dissimulans*; Rusrug1–*Russula rugulosa*; Lacsub1–*Lactifluus subvellereus*; Lacvol1–*Lactifluus* cf. *volemus*; Muloch1–*Multifurca ochricompacta*; Lacqui1–*Lactarius quietus*; Glocon1–*Gloeopeniophorella convolvens*; Aurvu1–*Auriscalpium vulgare*; Lenvul1–*Lentinellus vulpinus*; Clapy1–*Artomyces pyxidata*; Stehi1–*Stereum hirsutum*; Hetan2–*Heterobasidion irregulare*; Densp1–*Dentipellis sp.*; Lopni1–*Peniophora sp.*; Ricme1–*Peniophora cinereus*; Varmin1–*Vararia minispora*). Genomes are coded for nutritional mode as Ectomycorrhizal (ECM), Saprotrophic (Sap), and Pathogen (Path) with genome size in Megabases (Mb) graphed as bars.

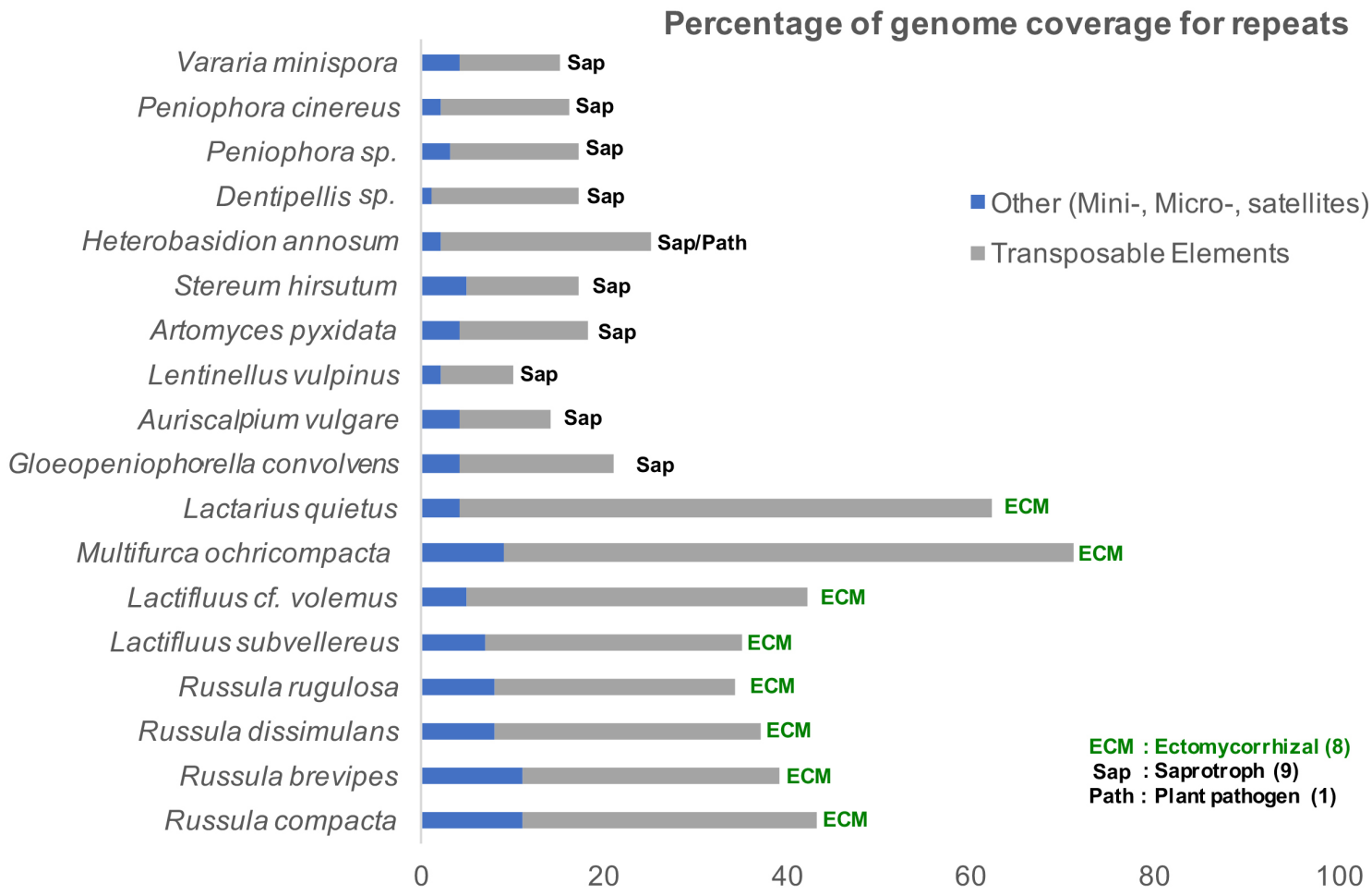


Figure 26. Comparison of the proportion of Russulales genomes that are comprised of transposable elements and other repeated elements.

Simple sequence repeats (SSRs) were compiled for each Russulales genome and classified by base pair length (Figure 27). Again, *Multifurca ochricompacta* contained the highest number of SSRs with a total of 15,167 SSRs of varying lengths. A majority of these were mononucleotide SSRs. *Lentinellus vulpinus* had the fewest number of SSRs with only 1,041 SSRs of varying lengths. Comparison of total SSEs between ECM Russulaceae and saprotrophic members of Russulales supported ECM genomes as having significantly more ($p<0.001$). However, when comparing rate differences between trophic modes, a Brownian motion model with rates constrained as equal received the highest support. ECM members also had more SSEs in the mononucleotide ($p<0.001$), dinucleotide ($p=0.03$), and trinucleotide ($p=0.005$) classes. Saprotrophic members of Russulales were deplete in these three classes of SSRs but still contained an equivalent number of larger SSRs.

Pangenome and network analyses

A presence/absence binarized analysis of Pfam gene families based on sequence homology resulted in a pan-genome reconstruction of Russulales (Figure 28). The entire pan-genome (core, variable, and unique) composes 24,629 gene families with 4,216 or 17.12% being shared among all members. *Lactarius quietus* had the most gene copies in a single gene family with 320 total copies. The number of gene families found to be unique among members of Russulales varied (Figure 29). *Lactarius quietus* was found with the most unique gene families and *Lactifluus* cf. *volemus* with the second most. The two species of *Peniophora* had the fewest unique gene families, with *Russula compacta* and *Multifurca ochricompacta* containing the third and fourth least number of unique gene families respectively. There was no statistically significant difference in unique gene family numbers between ECM Russulaceae and the rest of Russulales ($p=0.35$).

A similarity network was constructed based on similarity, measured as Proportional Similarity (PS), between the gene family content profiles of all species pairs (Figure 30). Under a PS threshold of 0.5, all species pairs are indistinguishable in the network. At a PS threshold of 0.55, *Lactarius quietus* and both *Peniophora* genomes can be distinguished from the network. A PS threshold of 0.6 separates *L. quietus* from the network entirely and polarizes the *Peniophora* genomes with *Peniophora* sp. being the most different. The 0.65 threshold has all ECM Russulaceae except *L. quietus* as a network with *G. convolvens* and most of the other saprotrophic members of Russulales in a network. The two *Peniophora* genomes together, the *Stereum hirsutum* genome, and *L. quietus* are separated from the network entirely. At a PS threshold of 0.68, *Lf.* cf. *volemus*, *H. irregulare*, and *Dentipellis* sp. are removed from the network. The remaining ECM Russulaceae are now a distinct unit, connected to saprotrophic members of Russulales through the *G. convolvens* genome.

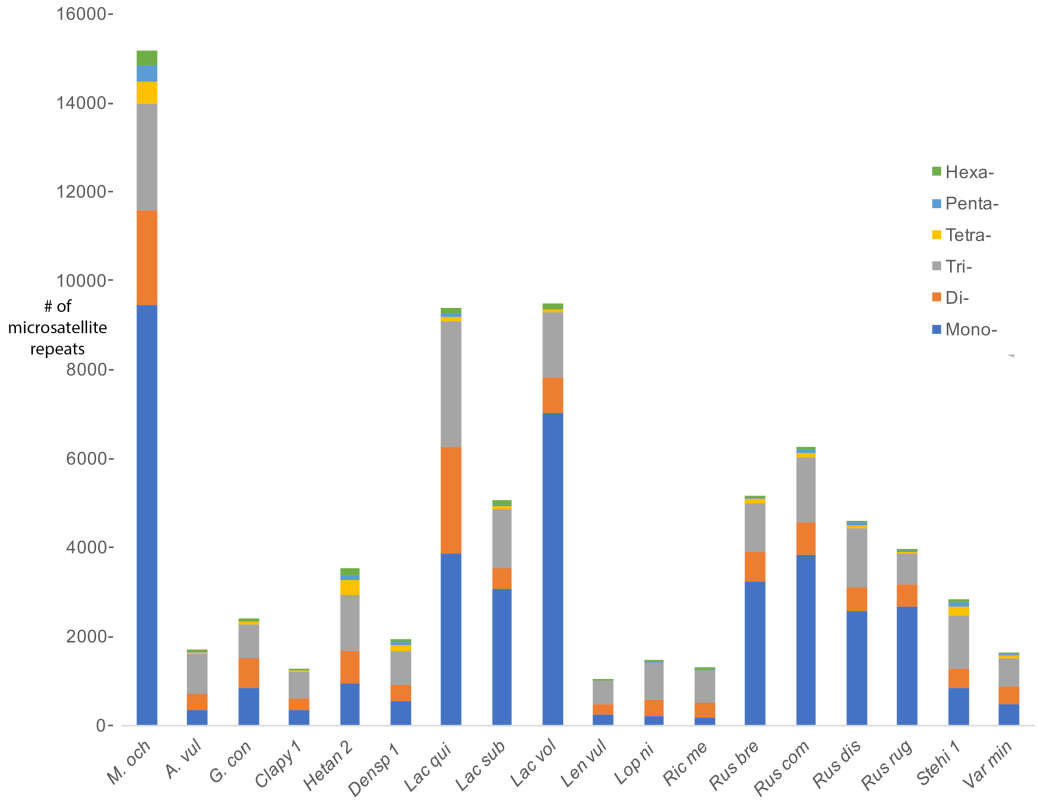


Figure 27. Comparison of the number and different classes of microsatellite repeats found in Russulales genomes. Colors correspond to different classes (dark blue–mononucleotide; orange–dinucleotide; gray–trinucleotide; yellow–tetranucleotide; light blue–pentanucleotide; green–hexanucleotide microsatellite).

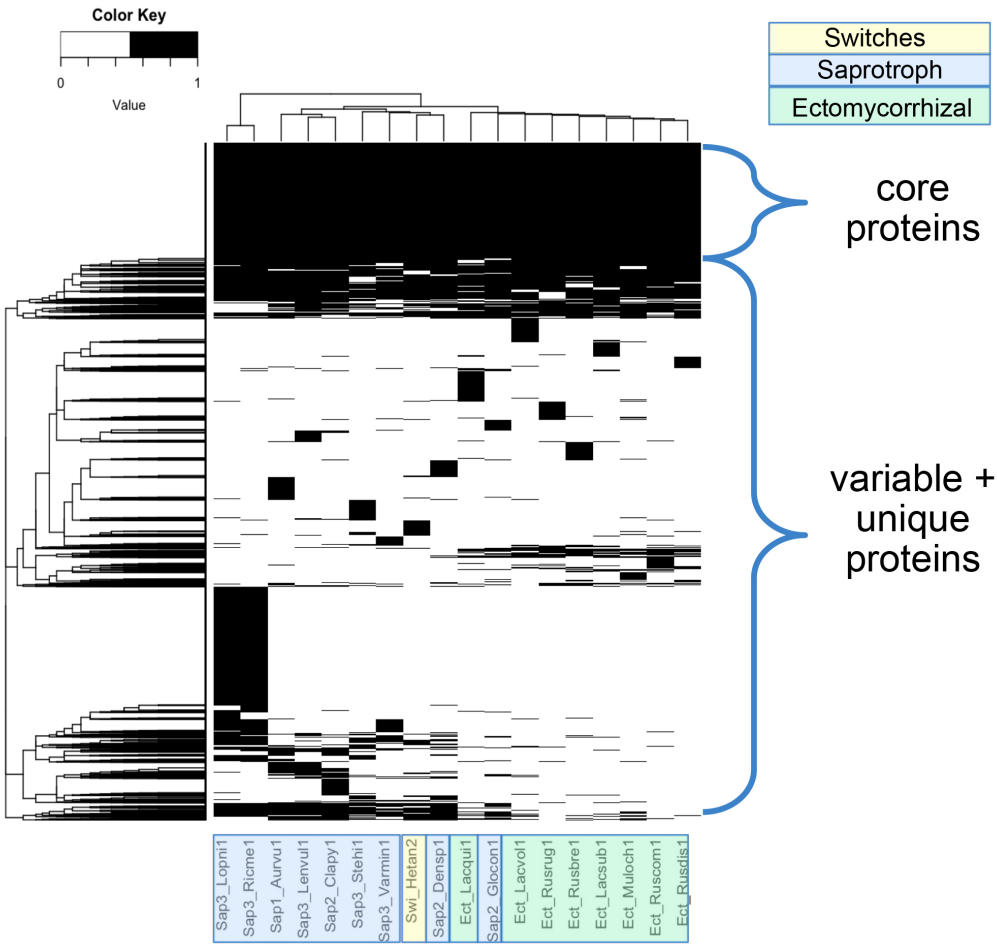
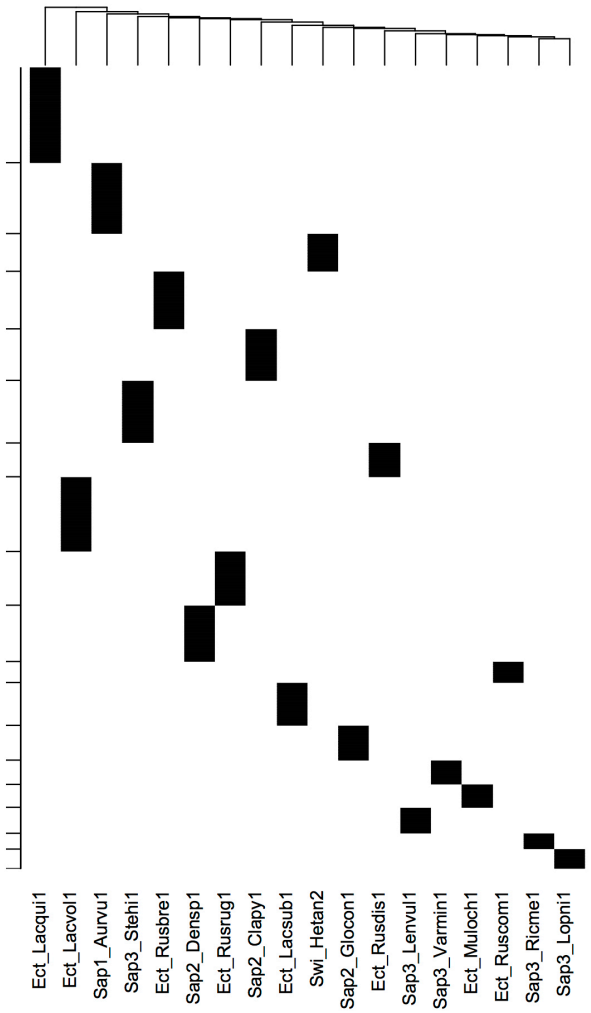


Figure 28. Heat-map reconstruction of the pan-genome of Russulales from a sequence homology-based protein families approach highlighting presence and absence of Pfam gene families based on Markov Clustering Algorithm. Core proteins shared by all members are clustered at the top with variability in presence and absence found throughout the rest. Abbreviations for genomes are JGI identifiers (see Fig. 25).



species	# protein families
Ect_Lacqui1	967
Ect_Lacsub1	434
Ect_Lacvol1	757
Ect_Muloch1	232
Ect_Rusbre1	583
Ect_Ruscom1	213
Ect_Rusdis1	342
Ect_Rusrug1	544
Sap1_Aurvu1	718
Sap2_Clapy1	522
Sap2_Densp1	570
Sap2_Glocon1	352
Sap3_Lenvul1	263
Sap3_Lopni1	198
Sap3_Ricme1	160
Sap3_Stehi1	634
Sap3_Varmin1	246
Swi_Hetan2	382

Figure 29. Heat-map reconstruction of unique Pfam gene families recovered in pan-genome analysis. Abbreviations for genomes are JGI identifiers (see Fig. 25).

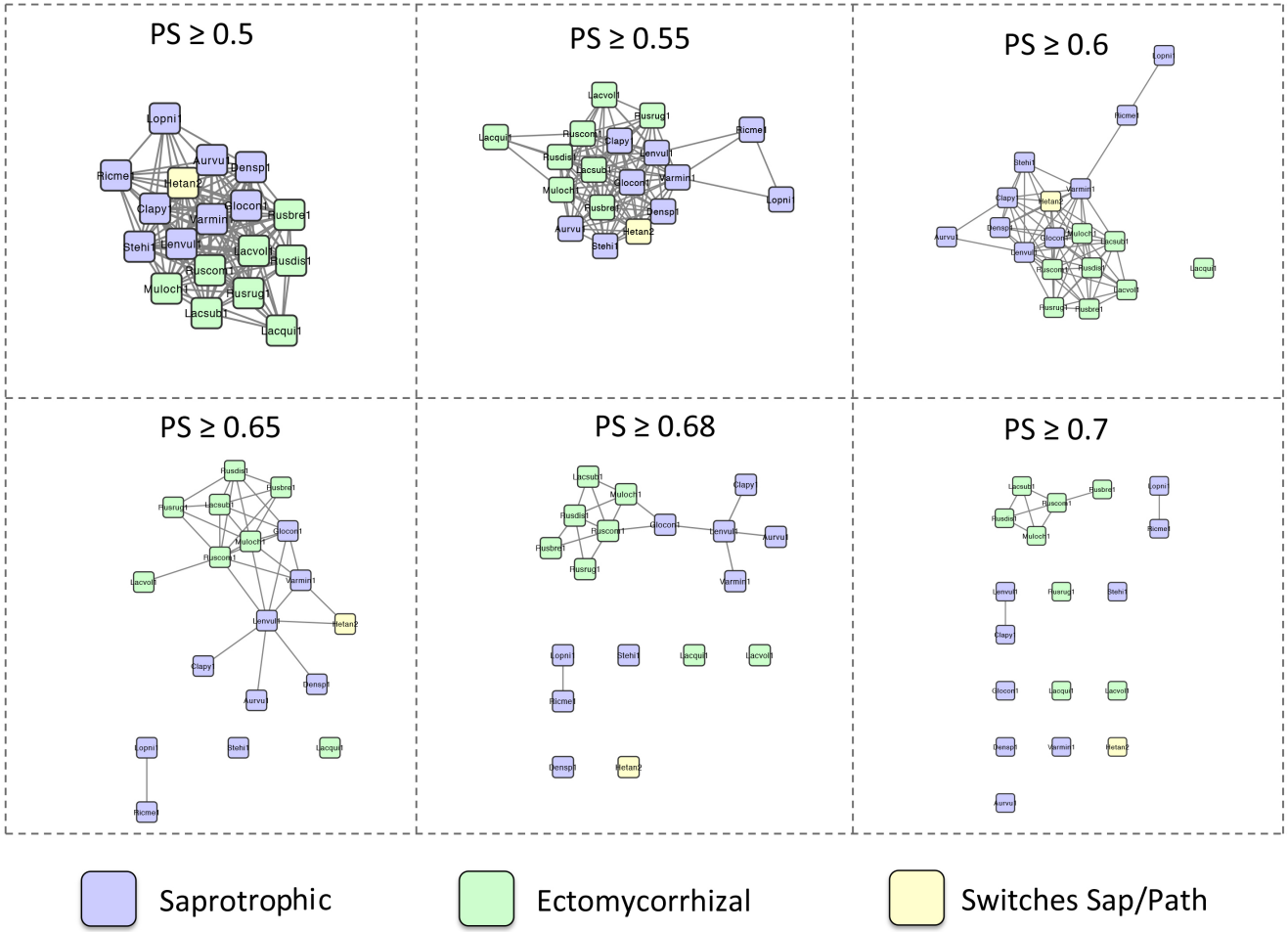


Figure 30. Proportional Similarity networks of the gene family content profiles of 18 Russulales species at different Proportional Similarity thresholds. Each node represents a species, and edges represent the similarity between the gene family profiles of species.

At the final PS threshold of 0.7, the network mostly collapses, with the ECM Russulaceae minus *R. rugulosa* still forming a network and paired connections between *L. vulpinus* and *A. pyxidata* as one set and the *Peniophora* genomes as the other.

A total of 1,281 Pfam domains were found to be positively enriched within Russulales (Figure 31). A majority of the enriched domains (721) were shared among members of Russulales. The second largest number of domains was enrichment for ECM Russulaceae (243). ECM Russulaceae enrichment includes methyl-associated, transferase, and transportase domains. A large number of enriched domains were shared between ECM and *H. irregulare* (96) and saprotrophs and *H. irregulare* (54). Only 13 domains were found enriched in only *H. irregulare*. Saprotrophic genome domains were enriched for polyketide activity, catabolic process, necrotic cell death, and others.

Evolution of lignin degrading enzymes

Gene histories for two lignin degrading gene families were reconstructed, including the lignin peroxidases (class-II peroxidases) and laccases (multi-copper oxidases). A phylogenetic reconstruction of class-II peroxidases from the Russulales resolved seven clades of saprotrophic Russulaceae (*G. convolvens*) and two clades of ECM Russulaceae with good support (bootstrap ≥70)(Figure 32). All ECM members of Russulaceae are represented with at least one lignin peroxidase gene, with *Russula brevipes* and *Multifurca ochricompacta* having representative genes in each clade. It appears that *Multifurca ochricompacta* has recently undergone a gene duplication event.

A phylogenetic reconstruction of the multi-copper oxidase family (MCO) for Russulaceae has resolved many gene copies with a complex history (Figure 33). All ECM Russulaceae members except *R. compacta* have more gene copies of laccases than their nearest extant saprotrophic ancestor, *G. convolvens*. Five ECM members, including *R. brevipes*, *Lf.* cf. *volemus*, *L. quietus*, *Lf. subvellereus*, and *R. rugulosa*, have ten copies of MCOs, which is the highest copy number in the family. The lowest copy number among ECM members is *R. compacta* with 7 copies. Major clades within the family received low support (bootstrap 50-69) except for the two largest clades, which received good support (bootstrap ≥70), and a clade containing Glocon MCO1, which received no support. Multiple gene copies for all members are recovered in clades, indicating the potential for having undergone gene duplication. *Gloeopeniophorella convolvens* genes occupy a sister position within most major clades that are present, either with or without support.

Discussion

As a first in-depth look at genomes of Russulales, we find a biochemically diverse order with differences in genomic architecture between nutritional modes.

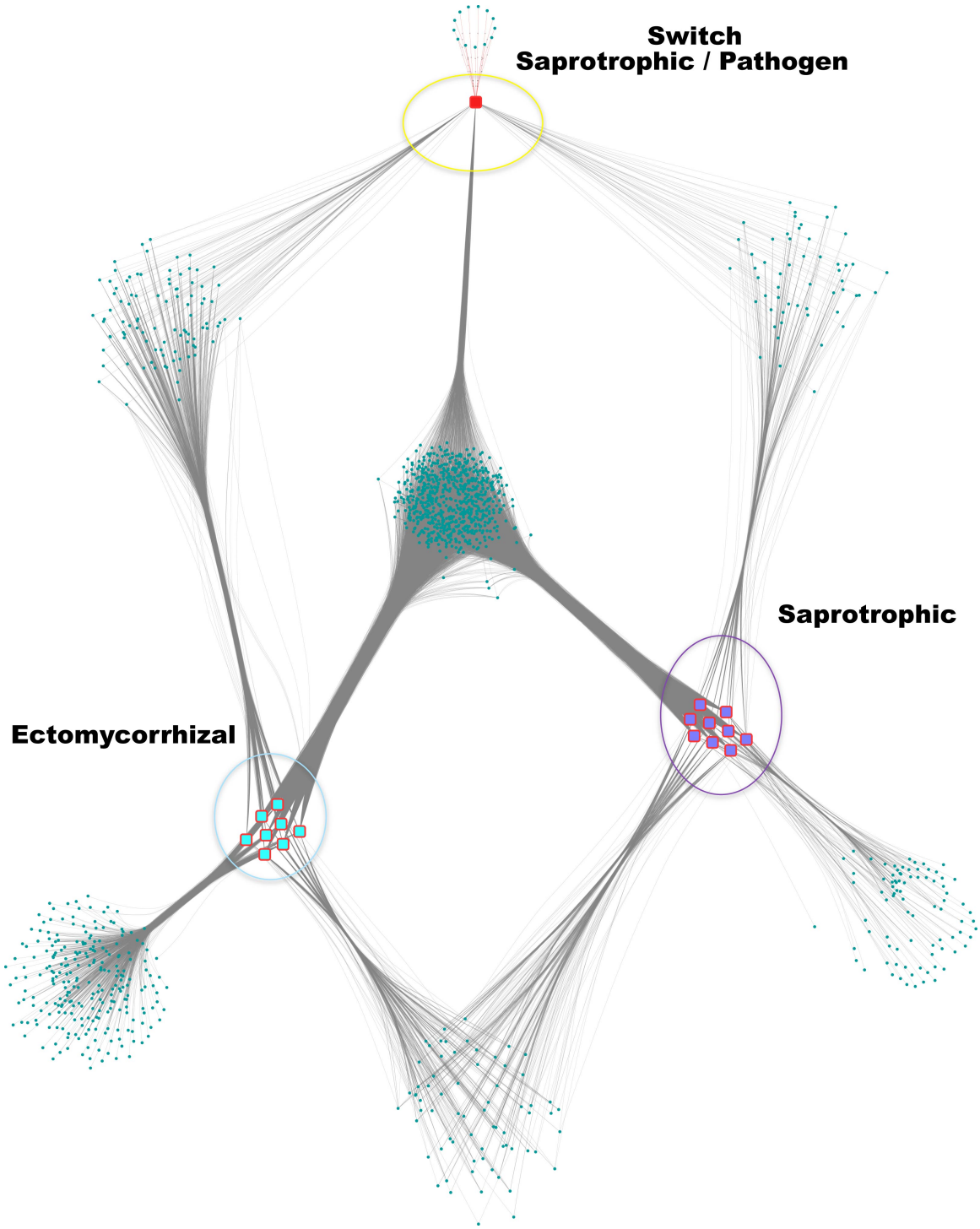


Figure 31. Network of Pfam domains that are positively enriched for nutritional mode within Russulales. Squares represent different species, while dots indicated enriched domains.

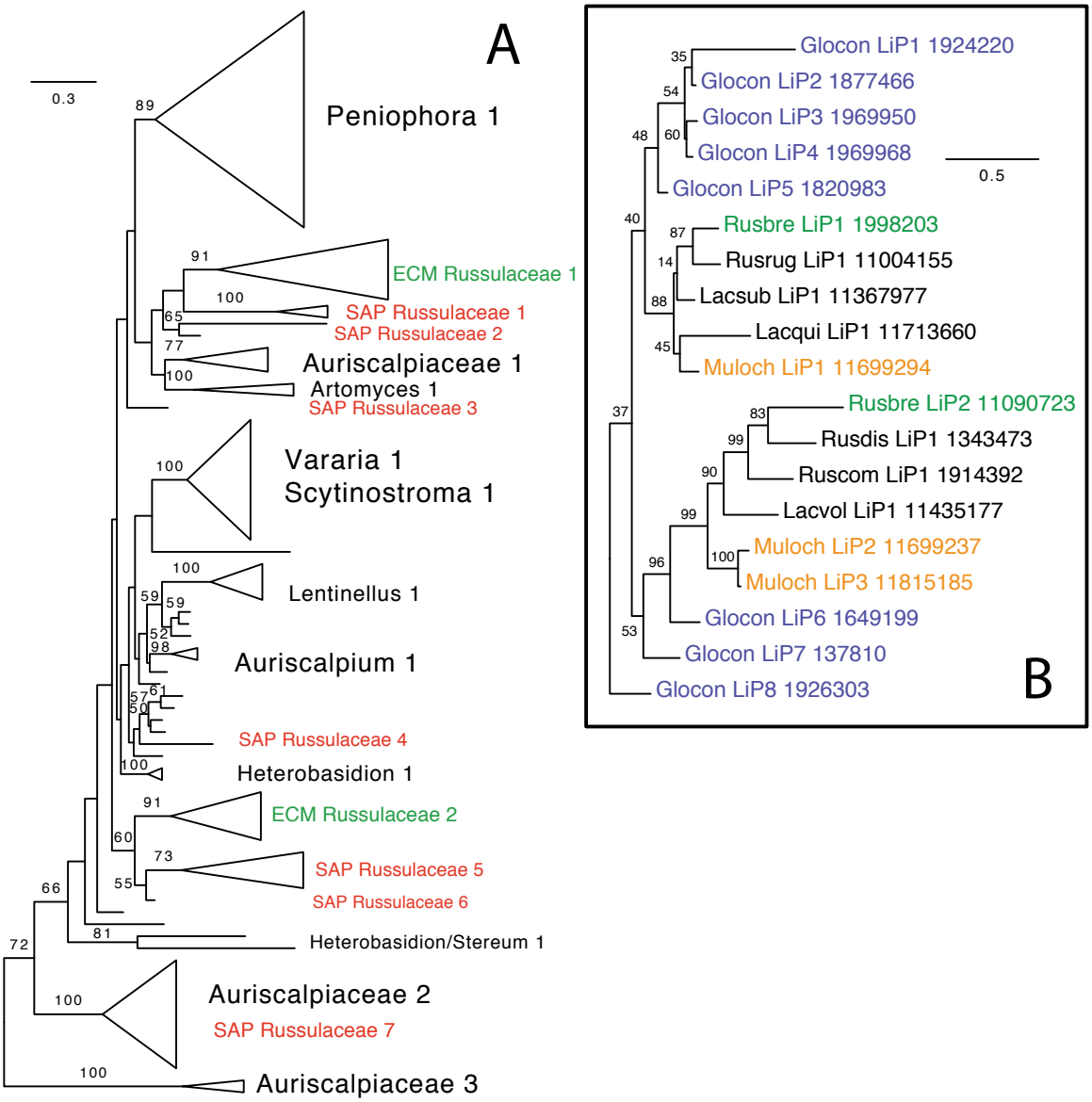


Figure 32. A) Phylogenetic reconstruction of the lignin peroxidase gene family from Russulales that codes for lignin peroxidase. Well-supported clades (bootstrap ≥70) are named based on included taxa. Clades including or made up of members of Russulaceae are colored based on nutritional mode (green–ectomycorrhizal; red–saprotrophic). B) Phylogenetic reconstruction of the lignin peroxidase gene family only from Russulaceae. Species with multiple gene copies are highlighted in color, including two ectomycorrhizal species (*M. ochricompacta* & *R. brevipes*) and the saprotrophic representative (*G. convolvens*). Abbreviations for genomes are JGI identifiers (see Fig. 25).

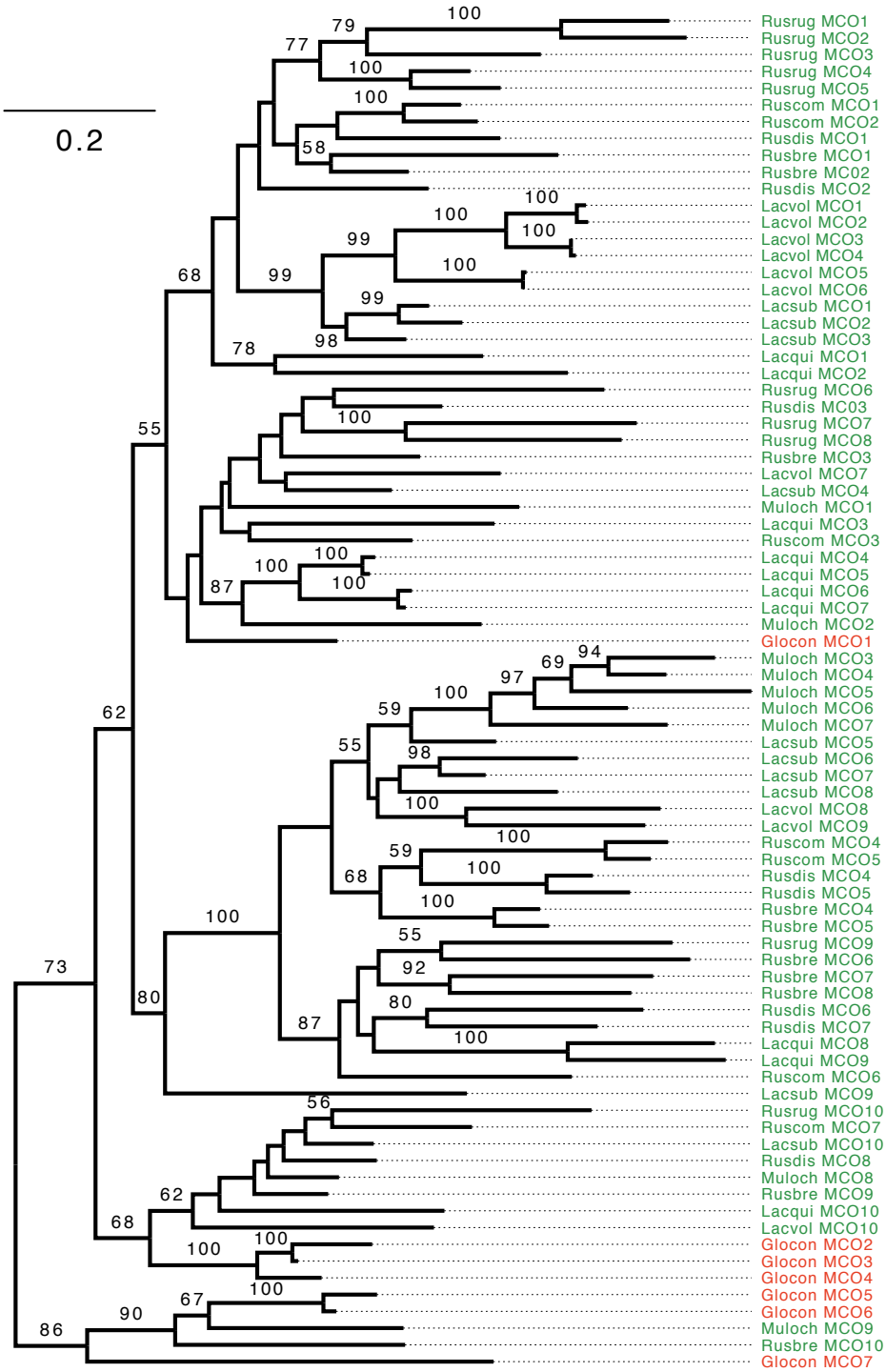


Figure 33. Phylogenetic reconstruction of the multi-copper oxidase (MCO) gene family from Russulaceae that codes for laccases. Gene copy numbers are numbered with JGI identifiers as abbreviations for genomes (see Fig. 1). Members of Russulaceae are colored based on nutritional mode (green–ectomycorrhizal; red–saprotrophic).

Phylogenetic support for evolutionary relationships based on single genes between major clades within Russulales has been lacking in past studies (Miller et al., 2006). Gene sampling is important for phylogenetic inference, and genomes provide the gene sampling power to fully resolve phylogenies (Binder et al., 2013; Rokas and Carroll, 2005).It is likely that a subset of genes may be all that is required to resolve this topology with maximum support, which would provide candidate markers for further phylogenetic reconstruction of the order. Russulaceae remains monophyletic, likely with a single origin of ECM for the ECM Russulaceae. Phylogenomic reconstruction also recovered the close relationships between *Auriscalpium*, *Lentinellus*, and *Artomyces* detected in Miller et al. (2006). The close relationship of *Stereaceae* with *Hericiaceae* and *Bondarzewiaceae* is a novel discovery and will need to be further investigated to determine if morphological or biochemical synapomorphies can be found that support this relationship. It will also be important to incorporate an outgroup from another order of Agaricomycotina, such as Agaricales or Polyporales, in order to infer ancestral evolution of genes across the phylogeny, though this will likely greatly reduce the number of candidate gene markers. Phylogenetic analyses of Russulaceae has also recovered poorly supported relationships (Buyck et al., 2008; Looney et al., 2016; Verbeken et al., 2014), which, at the genus level, are now fully resolved. Buyck et al. (2008) resolved *Lactifluus* as sister to the rest of the ECM Russulaceae clade, with *Multifurca* and *Lactarius* being closely related as sister clades. We see that *Lactarius* is resolved as sister to the rest of Russulaceae, with *Russula* and *Lactifluus* as sister clades. With *Lactarius* being a mostly north temperate group that is sister to the rest of Russulaceae, it is likely that Russulaceae have a north temperate origin like Sebacinales, though this will need to be explicitly tested (Looney et al., 2016; Tedersoo et al., 2014). Finally, relationships within four major groups of *Russula* have been resolved, repudiating previous phylogenetic reconstructions (Looney et al., 2016; Miller and Buyck, 2002). Perhaps with more taxon sampling, highly supported relationships can be used for examining trait evolution and biogeography at large scales within and between these clades.

In our examination of genomic architecture, a pattern emerged where ECM members of Russulaceae tend to have larger genomes than their most closely related saprotrophic relatives. Some of the largest genomes sequenced so far have been biotrophic, with the largest fungal genome currently belonging to the plant pathogen *Melampsora allii-populina* at 335.7 MB. When the *Laccaria bicolor* genome was released, it was the largest published genome at 65 MB, and this was attributed to the presence of TEs making up about 21% of the genome (Martin et al., 2008). When modeling the evolution of genome size within the Russulaceae, a model with no rate differences was supported. It is likely that only having sampling from one ECM lineage biased model selection, and a larger sampling from other orders may support this trend. We also see an expansion of TEs in the ECM Russulaceae. This trend of expanded TEs has been identified in biotrophic fungi in general, especially in pathogenic fungi (Duplessis, 2011) and also in ECM fungi (Hess et al., 2014). Again, model testing supported a model

with no rate difference, and the same consideration applies here as it does for genome size. Alterations in TE-related processes can lead to changes in virulency or host-specifity (Kang et al., 2001) and have been associated with rapid diversification events, like in the most species rich genus of mammals, *Myotis* (Ray et al., 2008). For ECM fungi, host interactions are largely mediated by small secreted effectors (SSEs) (Plett et al., 2011), which have been associated with repeat-rich regions of the genome (Zuccaro et al., 2011). An expansion in TEs may be a necessary adaptation for navigating complex interactions with plant hosts and could be a good evolutionary mechanism for host-mediated diversification. Unlike plant pathogens, ECM fungi like Russulaceae are not specific to a single plant host species but are usually able to colonize a large phylogenetic breadth of host lineages. The presence and arrangement of TEs in relation to SSEs should be an important aspect for examining host-specificity in ECM fungi.

The *Multifurca ochricompacta* genome is noteworthy for a number of reasons. This species is a member of an unusually species poor major clade in Russulaceae, with only six species known worldwide (Mycobank.org). The *M. ochricompacta* genome has an average size for ECM Russulaceae (54 MB), yet it has the largest proportion of its genome composed of repeatable elements and the lowest gene content of any ECM Russulaceae. Instead of spurring diversification, the ‘stress-induced’ hypothesis purports environmental disruption as a cause for direct activation of TEs, whereby the genome will restructure itself in order to overcome a threat to its survival (McClintock, 1984). Increased activity of TEs in this case can lead to extinction due to fitness loss and purifying selection against deleterious mutations (Belyayev, 2014). A past disruption caused by viral attack may have resulted in an accumulation of TEs coinciding with gene loss through purifying selection, leading to extinction in the ancestry of *Multifurca*.

Unlike the genome of *M. ochricompacta*, the *L. quietus* genome contains both a high TE count and a high gene count. Likely because of its size and unique gene content, *L. quietus* was the first species of Russulales to break from the gene family similarity analysis. The functional profile of unique gene families is of great interest Future analyses of the *Lactarius quietus* genome will seek to test whether a whole genome duplication event took place in its recent evolutionary history or if there has been a high rate of gene diversification and expansion in the *Lactarius* subgenus *Russularia* lineage. Except for in the unusual cases of *Multifurca* and *L. quietus*, we do not detect functional conserved divergence between genera of ECM Russulaceae at the genome scale. More sampling of both *Lactarius* and *Multifurca* will help determine if genome or TE expansion has occurred in the distant past, defining these groups, or if there is variation at the species or population level.

The pan-genome of Russulales has revealed the potential for significant functional differentiation between trophic modes that is mediated by evolutionary distance. We see a clade of conserved gene families uniting the saprotrophic members of Russulales (Figure 4). A similar region of conserved gene families

unites the ECM Russulaceae but also includes the saprotrophic Russulaceae outgroup. We also detected this same pattern in the functional similarity analysis, with *G. convolvens* acting as an anchor uniting a majority of ECM Russulaceae with a cluster of saprotrophic Russulales. This supports the idea of a gradual transition, rather than an abrupt switch in gene family composition between saprotrophy and biotrophy. Genome modification within major clades also surpasses those between nutritional modes, as unique protein families made up a larger proportion of genomes than those conserved for nutritional modes. We see this in functional similarity network analyses, where *L. quietus* is the first to break from the pan-genome network, even though the greatest amount of sampling comes from that family.

Genomes of ECM fungi have been characterized by a reduction or loss of enzymatic genes required for decomposition of plant cell walls, including lignin (Kohler et al., 2015). Whether or not ECM fungi retain the capability to degrade lignin is dependent on their saprotrophic ancestry (Kohler et al., 2015). This pattern has now been detected in Russulaceae, where ECM Russulaceae have retained lignin peroxidases from their white-rot (lignin decomposition) ancestry but have far fewer copies compared to their closest saprotrophic relative. Dense sampling within the ECM clade has allowed us to detect that these species have not retained the same lignin peroxidase gene, however, with two clades of lignin peroxidase genes having been recovered. These genes are split between both members of *Russula* and *Lactifluus*, indicating that there were multiple independent losses of these genes in both clades. In both cases, the same gene has likely been retained in the closest extant saprotrophic ancestor, *G. convolvens* and makes it less likely that they are functionally redundant. We have also detected a recent gene duplication event in one the clades for *M. ochricompacta*, further complicating this species' trophic status. Both species that possess copies of both lignin peroxidase genes are able to grow in culture, but the *L. quietus* genome was derived from culture and possesses only one lignin peroxidase. A transcriptomic analysis of *M. ochricompacta* can help detect under what condition each gene is expressed and whether they are co-expressed.

In an environmental analysis of fungal derived laccases in litter, Russulaceae were highlighted as being the most dominant in laccase genes (Luis et al., 2005). ECM Russulaceae genomes were found to contain multiple copies of MCOs with many potential instances of gene duplication (Figure 8). While this seems to suggest specialization in laccase production for this clade, gene copy numbers for other ECM species, like *Laccaria bicolor*, possess a similar number of genes, and saprotrophic members of Russulales possess even more (Floudas et al., 2012). Luis et al. (2005) points to the patchiness of genes recovered from Russulaceae, which is likely due to their contact mycorrhizal exploration type which concentrates their mycelial growth to only near to root tips (Agerer, 2001). If this is the case, Russulaceae may have been over-sampled by chance. Further exploration of gene families may yield functional specialization for this group.

In conclusion, we found support for a white-rot ancestry for Russulaceae with each member possessing at least one representative of two clades of lignin peroxidase genes. There is even an example of gene duplication in *M. ochricompacta*, which may indicate neofunctionalization. Though we did find evidence for expansion in the MCO family, this does not seem to indicate a unique specialization of Russulaceae. The ECM Russulaceae do not necessarily possess functional conservation at the genome scale that unites the group but rather supports a gradual functional gradation through ancestry towards functionalization. Finally, we find little evidence for functional specialization between major clades of ECM Russulaceae as sampling is lacking within *Multifurca* and *Lactarius* for infrageneric comparisons.

Acknowledgments

We would like to thank Jorinde Nuytinck at the Naturalis Biodiversity Center in Leiden and Otto Mietenin at the University of Helsinki for taxonomic expertise and providing sample support. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

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. doi:10.1007/s005720100108

Bahram, M., Põlme, S., Kõljalg, U., Tedersoo, L., 2011. A single European aspen (Populus tremula) tree individual may potentially harbour dozens of Cenococcum geophilum ITS genotypes and hundreds of species of ectomycorrhizal fungi. FEMS Microbiol. Ecol. 75, 313–320. doi:10.1111/j.1574-6941.2010.01000.x

Beaulieu, J.M., O'Meara, B.C., 2016. OUwie: analysis of evolutionary rates in an OU framework, v. 1.5.

Belyayev, A., 2014. Bursts of transposable elements as an evolutionary driving force. J. Evol. Biol. 27, 2573–2584. doi:10.1111/jeb.12513

Benson, G., 1999. Tandem Repeats Finder: a program to analyse DNA sequences. Nucleic Acids Res. 27, 573–578.

Binder, M., Justo, A., Riley, R., Salamov, a, Lopez-Giraldez, F., Sjokvist, E., Copeland, a, Foster, B., Sun, H., Larsson, E., Larsson, K.-H., Townsend, J., Grigoriev, I. V, Hibbett, D.S., 2013. Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105, 1350–1373. doi:10.3852/13-003

Birney, E., Clamp, M., Durbin, R.E., 2004. GeneWise and Genomewise. Genome Res. 14, 988–995. doi:10.1101/gr.1865504.quickly

Bodeker, I.T.M., Clemmensen, K.E., de Boer, W., Martin, F., Olson, A., Lindahl, B.D., 2014. Ectomycorrhizal Cortinarius species participate in enzymatic oxidation of humus in northern forest ecosystems. New Phytol. 203, 245–256. doi:10.1111/nph.12791

Buyck, B., Hofstetter, V., Eberhardt, U., Verbeken, A., Kauff, F., 2008. Walking the thin line between *Russula* and *Lactarius*: The dilemma of *Russula* subsect. *Ochricompactae*. Fungal Divers. 28, 15–40.

Collins, C., Keane, T.T.M., Turner, D.J., O'Keeffe, G., Fitzpatrick, D. a, Doyle, S., 2013. Genomic and proteomic dissection of the ubiquitous plant pathogen, *Armillaria mellea*: toward a new infection model system. J. Proteome Res. 12, 2552–70. doi:10.1021/pr301131t

Duplessis, S., 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc. Natl. Acad. Sci. U. S. A. 1–23. doi:10.1073/pnas.1019315108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1019315108

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. doi:10.1093/nar/gkh340

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R.A., Henrissat, B., Martínez, A.T., Otillar, R., Spatafora, J.W., Yadav, J.S., Coutinho, P.M., Vries, R.P. De, Ferreira, P., Findley, K., Foster, B., 2012. The Paleozoic Origin of

Enzymatic from 31 Fungal Genomes. Science 336, 1715–1719. doi:10.1126/science.1221748

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–52. doi:10.1038/nbt.1883

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. Nucleic Acids Res. 42, 1–6. doi:10.1093/nar/gkt1183

Grigoriev, I. V, Cullen, D., Goodwin, S.B., Hibbett, D., Jeffries, T.W., Kubicek, C.P., Kuske, C., Magnuson, J.K., Martin, F., Spatafora, J.W., Tsang, A., Baker, S.E., Grigoriev, I. V, Cullen, D., Goodwin, S.B., Hibbett, D., Jeffries, T.W., Kubicek, C.P., Kuske, C., Magnuson, J.K., Martin, F., Spatafora, J.W., Tsang, A., Baker, S.E., 2011. Fueling the future with fungal genomics. Mycology 2, 192–209. doi:10.1080/21501203.2011.584577

Hess, J., Skrede, I., Wolfe, B.E., Butti, K. La, Ohm, R.A., Grigoriev, I. V., Pringle, A., 2014. Transposable element dynamics among asymbiotic and ectomycorrhizal Amanita fungi. Genome Biol. Evol. 6, 1564–1578. doi:10.1093/gbe/evu121

Hibbett, D.S., 2006. A phylogenetic overview of the Agaricomycotina. Mycologia 98, 917–925. doi:10.3852/mycologia.98.6.917

Hibbett, D.S., Stajich, J.E., Spatafora, J.W., 2013. Toward genome-enabled mycology 105, 1339–1349. doi:10.3852/13-196

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., Walichiewicz, J., 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–467. doi:10.1159/000084979

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K., Itoh, M., Kawashima, S., Katayama, T., Araki, M., Hirakawa, M., 2006. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 34, D354-7. doi:10.1093/nar/gkj102

Kang, S., Lebrun, M.H., Farrall, L., Valent, B., 2001. Gain of virulence caused by insertion of a Pot3 transposon in a Magnaporthe grisea avirulence gene. Mol. Plant. Microbe. Interact. 14, 671–4. doi:10.1094/MPMI.2001.14.5.671

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. doi:10.1093/molbev/mst010

Kirk, P.M., Cannon, P.F., Stalpers, J.A., 2008. The Dictionary of the Fungi, 10th ed. CABI Publishing, Wallingfork, UK.

Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., Canbäck, B., Choi, C., Cichocki, N., Clum, A., Colpaert, J., Copeland, A., Costa, M.D., Doré, J., Floudas, D., Gay, G., Girlanda, M., Henrissat, B., Herrmann, S., Hess, J., Högberg, N., Johansson, T., Khouja, H.-R., LaButti, K., Lahrmann, U.,

Levasseur, A., Lindquist, E.A., Lipzen, A., Marmeisse, R., Martino, E., Murat, C., Ngan, C.Y., Nehls, U., Plett, J.M., Pringle, A., Ohm, R.A., Perotto, S., Peter, M., Riley, R., Rineau, F., Ruytinx, J., Salamov, A., Shah, F., Sun, H., Tarkka, M., Tritt, A., Veneault-Fourrey, C., Zuccaro, A., Tunlid, A., Grigoriev, I. V, Hibbett, D.S., Martin, F., 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat. Genet. 47, 410–5. doi:10.1038/ng.3223

Koonin, E.V, Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Krylov, D.M., Makarova, K.S., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Rogozin, I.B., Smirnov, S., Sorokin, A. V, Sverdlov, A. V, Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D.A., 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biol. 5, R7. doi:10.1186/gb-2004-5-2-r7

Larsson, A., 2014. AliView: A fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30, 3276–3278. doi:10.1093/bioinformatics/btu531

Li, L., Stoeckert, C.J.J., Roos, D.S., 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes -- Li et al. 13 (9): 2178 -- Genome Research. Genome Res. 13, 2178–2189. doi:10.1101/gr.1224503.candidates

Looney, B.P., Ryberg, M., Hampe, F., Sánchez-García, M., Matheny, P.B., 2016. Into and out of the tropics: Global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi. Mol. Ecol. 25, 630–647. doi:10.1111/mec.13506

Luis, P., Kellner, H., Zimdars, B., Langer, U., Martin, F., Buscot, F., 2005. Patchiness and spatial distribution of laccase genes of ectomycorrhizal, saprotrophic, and unknown basidiomycetes in the upper horizons of a mixed forest cambisol. Microb. Ecol. 50, 570–579. doi:10.1007/s00248-005-5047-2

Martin, F., Aerts, a, Ahrén, D., Brun, a, Danchin, E.G.J., Duchaussoy, F., Gibon, J., Kohler, a, Lindquist, E., Pereda, V., Salamov, a, Shapiro, H.J., Wuyts, J., Blaudez, D., Buée, M., Brokstein, P., Canbäck, B., Cohen, D., Courty, P.E., Coutinho, P.M., Delaruelle, C., Detter, J.C., Deveau, a, DiFazio, S., Duplessis, S., Fraissinet-Tachet, L., Lucic, E., Frey-Klett, P., Fourrey, C., Feussner, I., Gay, G., Grimwood, J., Hoegger, P.J., Jain, P., Kilaru, S., Labbé, J., Lin, Y.C., Legué, V., Le Tacon, F., Marmeisse, R., Melayah, D., Montanini, B., Muratet, M., Nehls, U., Niculita-Hirzel, H., Oudot-Le Secq, M.P., Peter, M., Quesneville, H., Rajashekar, B., Reich, M., Rouhier, N., Schmutz, J., Yin, T., Chalot, M., Henrissat, B., Kües, U., Lucas, S., Van de Peer, Y., Podila, G.K., Polle, a, Pukkila, P.J., Richardson, P.M., Rouzé, P., Sanders, I.R., Stajich, J.E., Tunlid, a, Tuskan, G., Grigoriev, I.V., 2008. The genome of Laccaria bicolor provides insights into mycorrhizal symbiosis. Nature 452, 88–92. doi:10.1038/nature06556

Martin, F., Cullen, D., Hibbett, D.S., Pisabarro, A., Spatafora, J.W., Baker, S.E., Grigoriev, I. V, 2011. Sequencing the fungal tree of life. New Phytol. 190, 818–821.

Martin, F., Kohler, A., Murat, C., Balestrini, R., Coutinho, P.M., Jaillon, O., Montanini, B., Morin, E., Noel, B., Percudani, R., Porcel, B., Rubini, A., Amicucci, A., Amselem, J., Anthouard, V., Arcioni, S., Artiguenave, F., Aury, J.-M., Ballario,

P., Bolchi, A., Brenna, A., Brun, A., Buée, M., Cantarel, B., Chevalier, G., Couloux, A., Da Silva, C., Denoeud, F., Duplessis, S., Ghignone, S., Hilselberger, B., Iotti, M., Marçais, B., Mello, A., Miranda, M., Pacioni, G., Quesneville, H., Riccioni, C., Ruotolo, R., Splivallo, R., Stocchi, V., Tisserant, E., Viscomi, A.R., Zambonelli, A., Zampieri, E., Henrissat, B., Lebrun, M.-H., Paolocci, F., Bonfante, P., Ottonello, S., Wincker, P., 2010. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature 464, 1033–8. doi:10.1038/nature08867

McCarthy, E.M., McDonald, J.F., 2003. LTR STRUC: A novel search and identification program for LTR retrotransposons. Bioinformatics 19, 362–367. doi:10.1093/bioinformatics/btf878

McClintock, B., 1984. The Significance of Responses of the Genome to Challenge. Science (80-.). 226, 792–801.

Melén, K., Krogh, A., Von Heijne, G., 2003. Reliability measures for membrane protein topology prediction algorithms. J. Mol. Biol. 327, 735–744. doi:10.1016/S0022-2836(03)00182-7

Miller, S.L., Buyck, B., 2002. Molecular phylogeny of the genus *Russula* in Europe with a comparison of modern infrageneric classifications. Mycol. Res. 106, 259–276. doi:10.1017/S0953756202005610

Miller, S.L., Larsson, E., Larsson, K.-H., Verbeken, A., Nuytinck, J., 2006. Perspectives in the new Russulales. Mycologia 98, 960–970. doi:10.3852/mycologia.98.6.960

Morin, E., Kohler, A., Baker, A.R., et al., 2013. Genome sequence of the button mushroom Agaricus bisporus reveals mechanisms governing adaptation to a humic-rich ecological niche. Proc. Natl. Acad. Sci. U. S. A. 110, 4146. doi:10.1073/pnas.1300201110

Nielsen, H., Engelbrecht, J., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites Artificial neural networks have been used for many biological. Protein Eng. 10, 1–6. doi:10.1142/S0129065797000537

Olson, Å., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., Canbäck, B., Coutinho, P.M., Cullen, D., Dalman, K., Deflorio, G., van Diepen, L.T.A., Dunand, C., Duplessis, S., Durling, M., Gonthier, P., Grimwood, J., Fossdal, C.G., Hansson, D., Henrissat, B., Hietala, A., Himmelstrand, K., Hoffmeister, D., Högberg, N., James, T.Y., Karlsson, M., Kohler, A., Kües, U., Lee, Y.H., Lin, Y.C., Lind, M., Lindquist, E., Lombard, V., Lucas, S., Lundén, K., Morin, E., Murat, C., Park, J., Raffaello, T., Rouzé, P., Salamov, A., Schmutz, J., Solheim, H., Ståhlberg, J., Vélëz, H., de Vries, R.P., Wiebenga, A., Woodward, S., Yakovlev, I., Garbelotto, M., Martin, F., Grigoriev, I. V., Stenlid, J., 2012. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol. 194, 1001–1013. doi:10.1111/j.1469-8137.2012.04128.x

Plett, J.M., Kemppainen, M., Kale, S.D., Kohler, A., Legué, V., Brun, A., Tyler, B.M., Pardo, A.G., Martin, F., 2011. A secreted effector protein of laccaria

bicolor is required for symbiosis development. Curr. Biol. 21, 1197–1203. doi:10.1016/j.cub.2011.05.033

Price, A.L., Jones, N.C., Pevzner, P.A., 2005. De novo identification of repeat families in large genomes. Bioinformatics 21, 351–358. doi:10.1093/bioinformatics/bti1018

Quesneville, H., Bergman, C.M., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M., Anxolabehere, D., 2005. Combined evidence annotation of transposable elements in genome sequences. PLoS Comput. Biol. 1, 0166–0175. doi:10.1371/journal.pcbi.0010022

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: Protein domains identifier. Nucleic Acids Res. 33, 116–120. doi:10.1093/nar/gki442

Rambaut, A., 2012. FigTree v1. 4. Mol. Evol. phylogenetics Epidemiol. Edinburgh, UK Univ. Edinburgh, Inst. Evol. Biol.

Ray, D.A., Feschotte, C., Pagan, H.J.T., Smith, J.D., Pritham, E.J., Arensburger, P., Atkinson, P.W., Craig, N.L., 2008. Multiple waves of recent DNA transposon activity in the bat, Myotis lucifugus. Genome Res. 18, 717–728. doi:10.1101/gr.071886.107

Riley, R., Salamov, A.A., Brown, W., Nagy, L.G., Floudas, D., Held, B.W., Levasseur, A., Lombard, V., Morin, E., Lindquist, E.A., Sun, H., Labutti, K.M., Jabbour, D., Luo, H., Baker, S.E., Antonio, G., Walton, J.D., Blanchette, R.A., Martin, F., Cullen, D., Hibbett, D.S., Riley, R., Salamov, A.A., Brown, D.W., Nagy, L.G., Floudas, D., Held, B.W., 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot / brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. U. S. A. 111, 9923–9928. doi:10.1073/pnas.1418116111

Rineau, F., Roth, D., Shah, F., Smits, M., Johansson, T., Canbäck, B., Olsen, P.B., Persson, P., Grell, M.N., Lindquist, E., Grigoriev, I. V., Lange, L., Tunlid, A., 2012. The ectomycorrhizal fungus Paxillus involutus converts organic matter in plant litter using a trimmed brown-rot mechanism involving Fenton chemistry. Environ. Microbiol. 14, 1477–1487. doi:10.1111/j.1462-2920.2012.02736.x

Rokas, A., Carroll, S.B., 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22, 1337–1344. doi:10.1093/molbev/msi121

Salamov, A.A., Solovyev, V.V., 2000. Ab initio Gene Finding in *Drosophila* Genomic DNA, Genome research, 10, 516–522. doi:10.1101/gr.10.4.516

Shah, F., Nicolás, C., Bentzer, J., Ellström, M., Smits, M., Rineau, F., Canbäck, B., Floudas, D., Carleer, R., Lackner, G. and Braesel, J., 2016. Ectomycorrhizal fungi decompose soil organic matter using oxidative mechanisms adapted from saprotrophic ancestors. New Phytol. 209, 1705–1719. doi:10.1111/nph.13722

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., 2003. Cytoscape: a software environment for integrated models of biomolecular

interaction networks. Genome Res. 13, 2498–2504. doi:10.1101/gr.1239303.metabolite

Silvestro, D., Michalak, I., 2011. raxmlGUI: a graphical front-end for RAxML. Org. Divers. Evol. doi:10.1007/s13127-011-0056-0

Smit, A.F.A., Hubley, R., Green, P., 1996. RepeatMasker Open-3.0.

Talbot, J.M., Bruns, T.D., Smith, D.P., Branco, S., Glassman, S.I., Erlandson, S., Vilgalys, R., Peay, K.G., 2013. Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. Soil Biol. Biochem. 57, 282–291. doi:10.1016/j.soilbio.2012.10.004

Tedersoo, L., Bahram, M., Ryberg, M., Otsing, E., Kõljalg, U., Abarenkov, K., 2014. Global biogeography of the ectomycorrhizal /sebacina lineage (Fungi, Sebacinales) as revealed from comparative phylogenetic analyses. Mol. Ecol. 23, 4168–4183. doi:10.1111/mec.12849

Tedersoo, L., Smith, M.E., 2013. Lineages of ectomycorrhizal fungi revisited: Foraging strategies and novel lineages revealed by sequences from belowground. Fungal Biol. Rev. 27, 83–99. doi:10.1016/j.fbr.2013.09.001

Ter-hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., Borodovsky, M., 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training 1979–1990. doi:10.1101/gr.081612.108

Thiel, T., 2003. MISA—Microsatellite identification tool. Website http//pgrc.ipk-gatersleben.de/misa/[accessed 7 March 2017].

Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180. doi:10.1111/j.1096-0031.2010.00329.x

Verbeken, A., Stubbe, D., van de Putte, K., Eberhardt, U., Nuytinck, J., 2014. Tales of the unexpected: Angiocarpous representatives of the Russulaceae in tropical South East Asia. Persoonia Mol. Phylogeny Evol. Fungi 32, 13–24. doi:10.3767/003158514X679119

Wawrzyn, G.T., Quin, M.B., Choudhary, S., Lopez-Gallego, F., Schmidt-Dannert, C., 2012. Draft genome of Omphalotus olearius provides a predictive framework for sesquiterpenoid natural product biosynthesis in Basidiomycota. Chem. Biol. 19, 772–783. doi:10.1038/jid.2014.371

Zuccaro, A., Lahrmann, U., Güldener, U., Langen, G., Pfiffi, S., Biedenkopf, D., Wong, P., Samans, B., Grimm, C., Basiewicz, M., Murat, C., Martin, F., Kogel, K.H., 2011. Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont Piriformospora indica. PLoS Pathog. 7. doi:10.1371/journal.ppat.1002290

CONCLUSIONS

It has been the primary objective of my dissertation research to elucidate the evolutionary relationships of *Russula* and use the power of molecular phylogenetics to develop a better understanding of the biogeography, evolution, and ecology of Russulaceae. This approach has resulted in a phylogenetic study of diversity at multiple scales that includes species, infra-generic groupings, genera, and even families. Each level presents its own challenges and rewards. A genus-wide phylogeny of *Russula* was an essential first step for these studies in order to identify where a targeted study on species relationships might be fruitful and also what groups need to be sampled for studies at a larger scale to more accurately represent the diversity of the genus. This approach has also provided a powerful tool for examining global diversity patterns, whereby a multi-gene phylogeny was used as a constraint topology for a megaphylogeny approach to inferring a global phylogeny. Multiple gene markers at the near-species level not only give one access to the phylogenetic species concept by helping to infer a phylogeny with high support but also the evolutionary species concept through coalescence theory that predicts past population sizes and estimates whether gene flow has ceased between divergent populations. Phylogenetics as a field has also provided effective tools for comparative genomics for infering the evolutionary history of genes and predicting their function. Species within ECM Russulaceae have retained a different suite of oxidative enzymes, with some gene expansions and some gene loss events that are not conserved within lineages. This indicates functionally diverse roles in nutritional scavenging that may be driving diversification within Russulaceae. Also, when taking into account phylogeny, we see that certain trends in genome size and TE content cannot be attributed to nutritional mode given current sampling. By combining all of these approaches, we now have a systematic framework for the genus *Russula* based on evolutionary relationships, a global perspective of the distribution and diversification patterns of *Russula* in relation to the reversal of the latitudinal diversity gradient in ectomycorrhizal fungi, the first multigene systematic revision of a clade of *Russula* species, a fully-resolved phylogeny of Russulales, and the evolution of functional traits within and between saprotrophic and ectomycorrhizal members of Russulales.

VITA

Brian Patrick Looney was born on an Air Force base in Peru, Indiana on March 6th, 1984. He spent a good portion of his childhood in Virginia Beach, VA playing in the marshes of the Chesapeake Bay and catching fiddler crabs. From an early age he developed a fascination with the sense of adventure one gets when interacting with nature. He continued this trend into his teenage years with the Boy Scouts of America, where his spent summers hiking around the Cumberland Plateau and visiting different areas of America on high adventure trips. His education at Earlham College, a liberal arts college in Richmond, IN, was diverse and his studies covered literature, philosophy, psychology, geology, and history. He received a Bachelors of Arts degree in English in 2006. It was not until he was teaching English in a small, coastal city in Japan that his fascination with biology and mushroom-forming fungi in particular was piqued. He began studying biology at the University of Tennessee in 2010 and joined the Matheny Lab as a post-graduate researcher in 2011. His first research project was a molecular systematic study of the genus *Auricularia* in the southeast United States. In 2012, he joined the graduate program in the Department of Ecology and Evolutionary Biology to pursue a PhD with a focus on mycology and molecular systematics.