

Into and out of the tropics: global diversification patterns in a hyperdiverse clade of ectomycorrhizal fungi

BRIAN P. LOONEY,* MARTIN RYBERG,† FELIX HAMPE,‡ MARISOL SÁNCHEZ-GARCÍA* and P. BRANDON MATHENY*

*Department of Ecology and Evolutionary Biology, University of Tennessee, 332 Hesler Biology Building, Knoxville, TN 37996-1610, USA, †Department of Organismal Biology, Uppsala University, Evolutionsbiologiskt Centrum, Norbyväg. 18D, 75236 Uppsala, Sweden, ‡Department of Biology, Gent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

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 publically 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.

Keywords: fungi, latitudinal diversity gradient, phylogeography, phyloinformatics, Russulaceae

Received 28 July 2015; revision received 10 November 2015; accepted 22 November 2015

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).

Correspondence: Brian P. Looney, Fax: 865-974-3067; E-mail: blooney@vols.utk.edu

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). The Russulaceae has 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 myco-heterotrophic members of Ericaceae and Orchidaceae (Kennedy *et al.* 2011a), and occasionally have a

gasteroid fruit body morphology which makes up a significant 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 S1, Supporting information). 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 colour 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). PARTITION-FINDER 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://datadryad.org>).

A multigene phylogeny was inferred using RAXMLGUI 1.2 (Stamatakis *et al.* 2008; Silvestro & Michalak 2012) executing 1000 rapid ML bootstraps replicates (Table S2, Supporting information). 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 acquisition

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 data 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 (Table S3, Supporting information). 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 analysed 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 (Table S4, Supporting information). 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 (Table S5, Supporting information).

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 (Fig. 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 (Fig. 2). Most tropical MOTUs (105 or 51%) were sampled from Africa. Over 93% of MOTUs

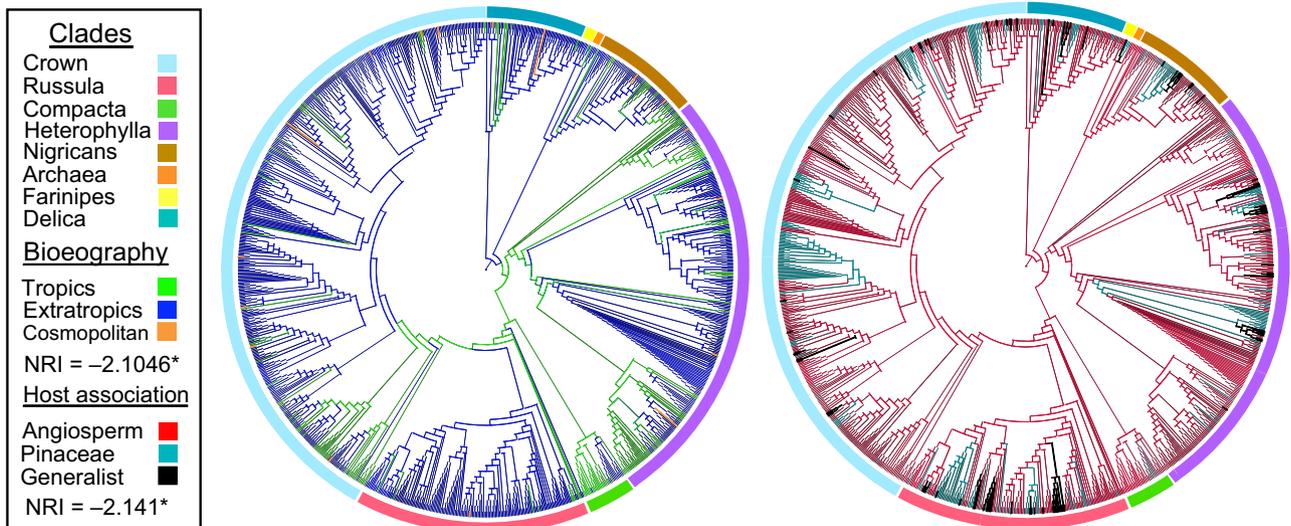


Fig. 1 Maximum parsimony ancestral reconstructions of geography (left) and host association (right) along an ultrametric mega-phylogeny of environmental *Russula* MOTUs inferred using r8s. Major clades are designated by colour. Geographical and host tree metadata associated with MOTU clusters are designated 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.

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 179 MOTUs (17%) were recovered as hosts for orchids or achlorophyllous members of the Ericaceae (Table S6, Supporting information). Fifty-three MOTUs (5%) were recovered with a gasteroid morphology, and not all of the members in these clusters shared this morphology.

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

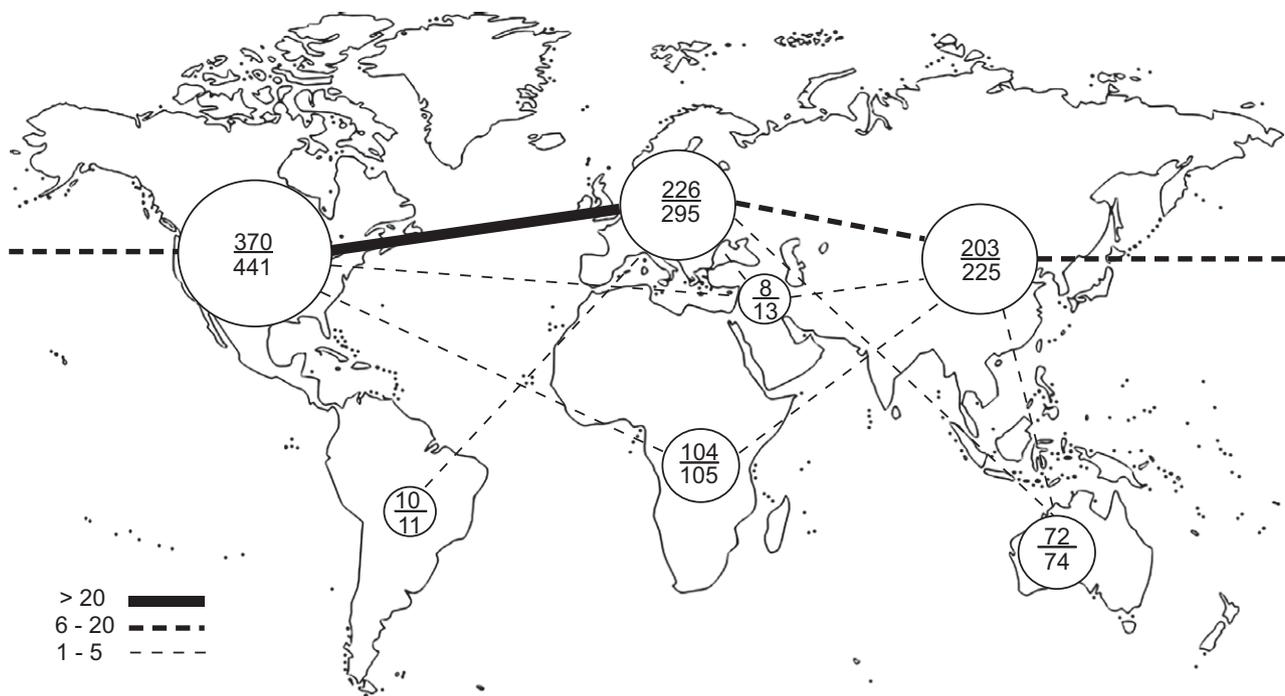


Fig. 2 Map projection of the global distribution of *Russula* MOTUs. The areas of circles are scaled by the number of MOTUs relative to total MOTUs recovered. Top numbers represent number of endemic MOTUs. Bottom numbers indicate total number of MOTUs. Lines represent number of overlapping distributions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

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 (Table S3, Supporting information). 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 (Fig. 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).

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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 (Table S4, Supporting information).

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

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

Clades	Geography binary						Host binary			Geography 4-state			Host family 6-state					
	Age		Geog		PP		Geog		Host		ML state		MP		ML state		MP	
	MY	PP	ML	MP	state	ML	MP	state	ML	MP	ML	MP	ML	MP	ML	MP	ML	MP
Root	43.96	0.88*	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.78/0.21*	Angi	0.55	Equi	0.55	Pin,Fab
Except del	43.52	0.79	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.78/0.22*	Angi	0.54	Equi	0.54	Pin,Fab
Het/rus/com/cro	43.18	0.64	Trop	Trop	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.48/0.52*	Angi	0.53	Equi	0.53	Fab
Nig	42.78	0.88*	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.88/0.12*	Angi	0.49	Equi	0.49	Pin,Fab
arc/far																		
Het	42.17	0.50	Trop	Temp	Temp	Temp	Temp	Trop	0.50	Equi	0.50	Equi	0.37/0.64*	Angi	0.5	Equi	0.5	Fab
Rus/com/cro	42.14	0.65	Trop	Temp	Temp	Temp	Temp	Trop	0.50	Equi	0.51	Angi	0.45/0.55*	Angi	0.49	Equi	0.49	Fab
Nig/arc	37.88	0.87	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.95*	Angi	0.5	Equi	0.5	Pin,Fab
Com	37.03	0.82	Trop	Temp	Temp	Temp	Temp	Trop	0.50	Equi	0.50	Equi	0.12/0.87*	Angi	0.48	Equi	0.48	Fab
Nig	36.65	0.80	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.51	Pina	0.98*	Angi	0.5	Equi	0.5	Pin,Fab
Arc	33.73	0.69	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.97*	Angi	0.5	Equi	0.5	Pin,Fab
Rus/cro	33.25	0.72	Temp	Temp	Temp	Temp	Temp	Trop	0.50	Equi	0.50	Equi	0.50/0.50*	Angi	0.5	Equi	0.5	Fab
Del	31.44	0.80	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.50	Equi	0.95*	Angi	0.5	Equi	0.5	Pin,Fab
Far	30.57	0.83	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.52	Angi	0.97*	Angi	0.32	Equi	0.32	Pin,Fab
Rus	30.29	0.85	Temp	Temp	Temp	Temp	Temp	Temp	0.50	Equi	0.51	Angi	0.91*	Angi	0.39	Equi	0.39	Fab,Fag
Cro	29.94	0.64	Trop	Temp	Temp	Temp	Trop	Trop	0.50	Equi	0.54	Angi	0.44/0.55*	Angi	0.57	Equi	0.57	Fab

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

Bold indicates states that have a higher likelihood when multiple states are found significant.

*Significance based on the difference of $-LnLik > 2$.

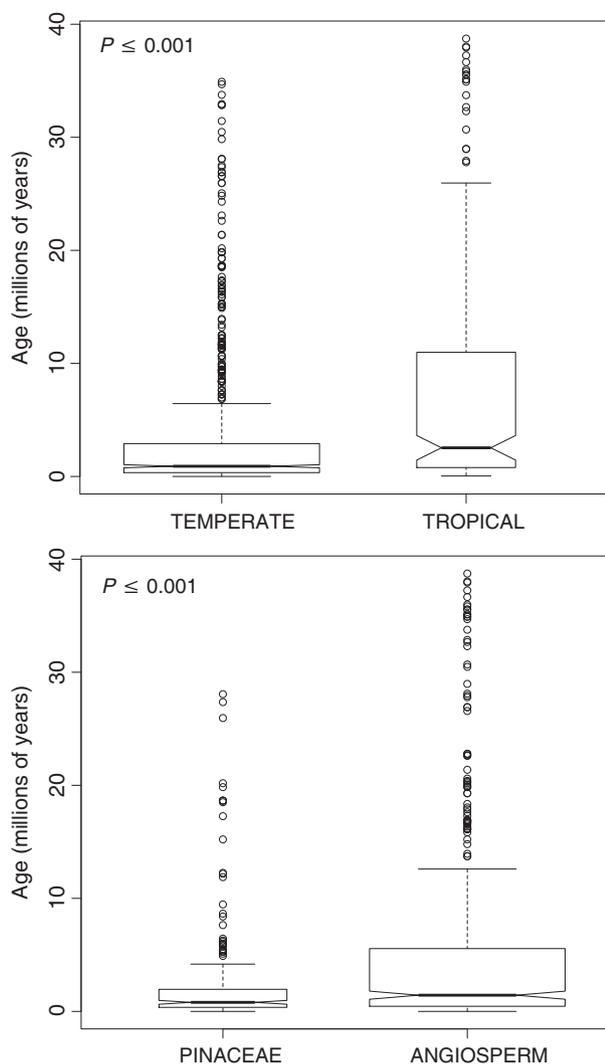


Fig. 3 Boxplot comparing average taxon age based on terminal branch lengths of taxa from a secondarily time-calibrated mega-phylogeny with ranges in the tropics or extratropics (Top) and host association with Pinaceae or angiosperms (Bottom). P -values resulted from nonparametric Mann–Whitney U -tests.

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 CO_2 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 (Brunns *et al.* 1998; Ryberg &

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***

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

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

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;

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

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

Wilson *et al.* 2012; Sánchez-Ramírez *et al.* 2015b). The palaeotropics, in particular, have been recognized as the ancestral origin of Inocybaceae, *Amanita* sect. *Caesareae*, and most ECM clades of Sclerodermatineae (Matheny *et al.* 2009; Wilson *et al.* 2012; Sánchez-Ramírez *et 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

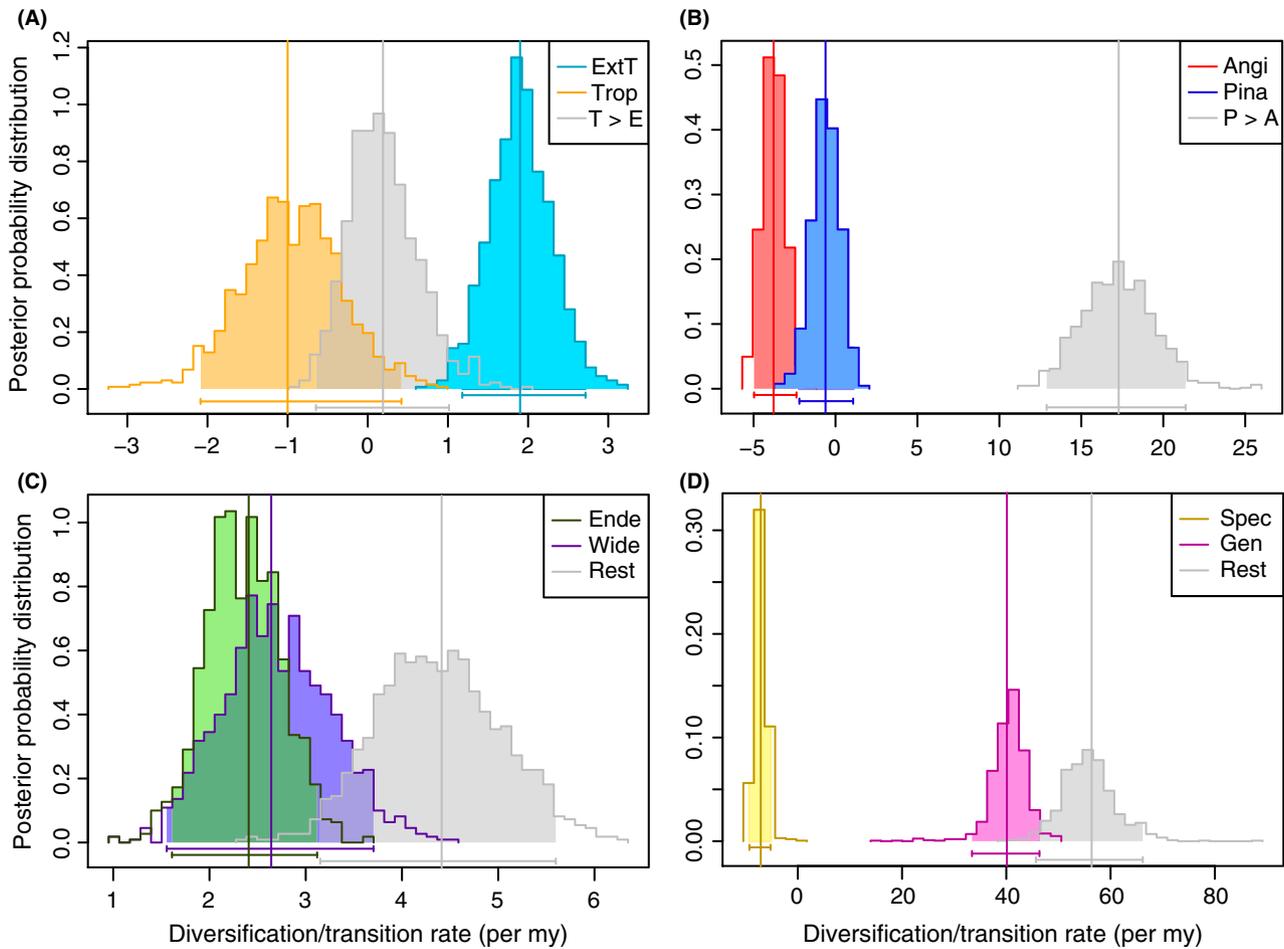


Fig. 4 Posterior probability density means and standard error representing relative diversification ($\lambda-\mu$) and dispersal rates for geographic and host state-specific models for *Russula*. (A) Estimates for extratropical MOTUs (light blue) and tropical MOTUs (orange) with differential transition 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.

'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 (Fig. 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; Pölme *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 *Cortinari*, *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; Pölme *et 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). Stud-

ies 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 (Kallioikoski *et al.* 2010), mycorrhizal root signalling (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 phenological 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 (Table S4, Supporting information), 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 S6, Supporting information).

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 & Pölme 2012). 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 under-sampled, 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.

Acknowledgements

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 Milieu-kunde*. WP Van Stockum & Zoon, The Hague, The Netherlands.
- Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-visited. *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 Tuberales 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 ectomycorrhizal 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, Fernández-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.
- Kerckhoff 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.
- Pölmö 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, Pölme 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, Pölme S *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, Sebaciales) 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.

B.L., B.M. and M.R. jointly conceived of the study; B.L. generated all sequence data; F.H. provided European specimens and assisted B.L. with species identification; B.L. analysed the data with input from B.M. and M.R.; M.S. wrote Perl scripts for data manipulation; and B.L. wrote the manuscript with input from all authors.

Data accessibility

All annotated sequences generated for this study are deposited in GenBank with accession numbers given in bold in Table S1 (Supporting information). Molecular alignments, phylogenies and the CLUSTERTREE file package are deposited in Dryad: doi:10.5061/dryad.gn4p4. All other relevant data are available through Supplementary files.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 GenBank accession numbers of DNA sequences generated for this study.

Table S2 Phylogenetic relationships of *Russula* inferred from nuclear ribosomal and single-copy (ITS, nrLSU, *rpb1* and *rpb2*) sequences derived from a maximum-likelihood analysis.

Table S3 Chronogram of Russulaceae inferred in BEAST 2. The 95% HPD Posterior probabilities >0.9 are reported.

Table S4 Taxon sampling bias runs at increments of 10% sampling bias for 1000 MCMC generations each for the four data sets depicted in Fig. 4: GeoSSE tropical vs. extratropical geography (A), GeoSSE angiosperm vs. Pinaceae association (B), BiSSE binary tropical/temperate endemism vs. widespread (C), and BiSSE binary angiosperm/Pinaceae specificity vs. generalist association.

Table S5 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 S6 Final MOTU clusters with associated metadata for mycoheterotrophic parasitism, gastroid morphology, plant association and geographic distribution.