GENETIC DIVERSITY OF DAHONGJUN, THE COMMERCIALLY IMPORTANT 'BIG RED MUSHROOM' FROM SOUTHERN CHINA

By

Mochan Li, B. SC.

A Thesis

Submitted to the school of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2009)
(Biology)

McMaster University
Hamilton, Ontario

TITLE: Genetic Diversity of Dahongjun The Commercially Important 'Big Red Mushroom" From Southern China

AUTHOR: Mochan Li, B. Sc. (McMaster University)

SUPERVISORS: Professor Dr. Jianping Xu

NUMBER OF PAGES: 106

Abstract

The genus *Russula* is among the most numerous and complex mushroom genera. They form symbiotic associations with a variety of plants such as pines, oaks, spruce, and birches and play vital roles in forest ecosystem. Species in this genus are widely distributed throughout the world, from the tropics to subtropics, temperate regions, and the arctic zone. Intensive studies on this genus have been carried out on European and North American *Russula.* However, there are very few studies on Asian *Russula.*

In southwestern China the favorable climate stimulates the production of a large population of *Russula.* A wild mushroom commonly called "dahongjun" or the "big red mushroom" by the locals, has been harvested, consumed and/or exported as an exotic food for many years. Although ecologically and economically important, very little is known about this mushroom, including its basic ecology and population genetic structure. In this study, we investigated the genetic diversity and geographic structure of dahongjun using sequences of the internal transcribed spacer (ITS) regions of the ribosomal RNA gene cluster. We found that this exotic mushroom contained at least three divergent lineages, with one corresponding to the recently described *Russula griseocarnosa* and two others likely representing closely related novel species within the genus *Russula.* Both phylogenetic and population genetic analyses suggested that dahongjun populations were geographically structured and gene flow among regions were limited. We also estimated the size of the genets for one *Russula* lineage by comparing haplotypes derived from four different DNA markers. Small genets suggest that sexual reproduction is likely more dominant than clonal reproduction in dahongjing. We found that the size of dahongjun genets ranged from 30cm to less than 3m, and widespread recombination within a local population. This result was consistent with the findings from other *Russula* species. Our study provided the first insights into the diversity and population biology of this endemic gourmet mushroom in southern China. This work serves as a great referencing resource on conducing conservation practices for wild *Russula.*

Acknowledgements

My graduate study in McMaster was extremely enjoyable. Six-year study (including my undergraduate study) I spent in McMaster has made up a lot of great excitement. They are life-time memories. McMaster provides great resources for learning and self-development. The knowledge I gained and the skills I developed were both closely associated with my mother school: McMaster. Many thanks should be given to her!

First of all, I would like to thank all my professors starting with my supervisor Dr. Jianping Xu for his generous help on my thesis works both experimentally and knowledgably, great leading skills, and English writing skills. Thanks for giving me the opportunity to experience the real science! I appreciate it. Thanks for sharing your scientific knowledge with your students and help them grow up to become true scientists! I would also like to thank Dr. Timothy Y. James for his constructive suggestions to my thesis. Thanks for introducing new methods to my data analysis. Your valuable suggestions were truly inspiring. The *Russula* T-shirt you brought in has been encouraging me to work hard all the way on my research. What a great support from you! It makes me never forget! The next thank you should be given to Dr. Sheng Sun for showing me new experimental techniques. The working ethics I inherited from you will certainly benefit me throughout my future career no matter what I do. I would also like to thank Dr. Junfeng Liang for his generous help on the cloning experiments and great ideas brought into the experimental design. Even though working with you was considered the hardest working time ever, it was very productive. I rated it as one of my greatest research experiences. Lastly, I would like to thank my committee member Dr. Ben Evans. He was available all the time and willing to help whenever needed. Thanks for all the valuable suggestions from you.

My research success was not possible without the help and care from everybody in Dr. Xu's lab during the year of 2007-2009. Thank you: Dr. Jianrong Wu, Dr. Ping Zhang, Dr. Anuradha Chowdhary, Irina Skosireva, Sanjay Hiremath, David Bartfai, Rahul Nayak, and Hari Maganti (in no particular order).

Also, big thanks should be given to funding agencies for the support of my research. They are: Natural Science and Engineering Research Council of Canada, Premier's Research Excellence Award, the Joint Funds of the National Natural Science Foundation of China and Yunnan Provincial Government, the National Basic Research Program of China, and the Key Laboratory of Biodiversity and Biogeography of the Chinese Academy of Sciences.

Lastly, but definitely not least, I would like to give respect and credit to my family: my parents, uncle, aunt and my boyfriend Songtao Yang. I would have made none of these accomplishments without their support, and encouragements. Many thanks to them!

Mochan Li, B. Sc.

McMaster University

September 2009

Thesis Format

This thesis is organized into three chapters. Chapter one provides a detailed introduction of ectomycorrihzal fungi (EcM). We did a general literature review on the genet sizes and ecological importance of ectomycorrhizal (EcM) fungi and the common approaches for studying their taxonomy and phylogeny. Chapters two contains experimental works in the form of manuscripts that under review (chapter two) for publication in scientific journals. Chapter three is a general discussion, providing a brief review of major results and conclusions that are established in the previous two chapters. References have been complied into one list, and are included at the end of the thesis following the last chapter.

Chapter 1: Molecular ecology of ectomycorrhizal fungi: molecular markers, genets and ecological importance

Authors: Mochan Li, Jianping Xu

Comments: This review article has been accepted by the Journal of Yunnan Botanical Research.

Chapter 2: Genetic diversity of dahongjun, the commercially important 'big red mushroom" from southern China

Authors: Mochan Li, Junfeng Liang, Yanchun Li, Bang Feng, Zhu L. Yang, Timothy Y. James, and Jianping Xu

Comments: This laboratory portion of this study was performed mainly by Mochan Li under the supervision and guidance of Dr. Jianping Xu. This paper has been submitted to PlosONE.

Chapter 3: General Discussion

Table of Contents

List of Tables

List of Figures

CHAPTER 1
Molecular Ecology of Ectomycorrhizal Fungi: Molecular Markers, Genets and Ecological Importance

Introduction of Ectomycorrhizal Fungi

Many of the known fungi are found with close associations of the roots of plant species, forming mutually beneficial symbiotic relationships. These fungi can colonize the plant roots and derive nutrients such as soluble carbohydrates, amino acids and vitamins from the plants. In return, plant hosts use the expanded surface of fungal mycelia to absorb water and minerals from soil (Read, 1991; Simard *et al.,* 1997, 2002; Table 1.1). The fungi that colonize plant roots and form mutualistic relationships with their plant hosts are called mycorrhizae, a name derived from Greek that means "fungus roots". In a mycorrhizal association, the fungus may colonize host roots either intracellularly or extracellularly, corresponding respectively to two groups of mycorrhizae: endomycorrhzae and ectomycorrhizae. In endomycorrhizae, fungal hyphae penetrate plant cell walls and form vesicles or arbuscules inside host plant cells. In contrast, in ectomycorrhizae, fungal hyphae grow extracellularly in the inter-cellular spaces and form sheaths around plant roots.

The symbiotic association between plants and their mycorrhizae can be traced back to 400-460 million years ago, when the first plant appeared on land (Remy *et al.,* 1994). Through the long history, both the plants and mycorrhizal fungi have adapted to benefit each other from the symbiosis. For example, many studies have demonstrated that mycorrhizal fungi play important roles for maintaining the normal growth of their plant partners (Jeffries *et al.,* 2003).

Without mycorrhizal colonization, plant hosts showed slower growth rate than those with EcM fungi (Smith & Read *et al.,* 1997). Recent research suggests that the reason for the poor performance of mycorrhizae-free plants could be due to their inadequate water and mineral uptake (Richard *et al.,* 2005; Selosse *et al.,* 2006). Similarly, it has been demonstrated that plants with mycorrhizae are often more resistant to draught and to infectious diseases, such as those caused by microbial soil-borne pathogens (Bledsoe *et al.,* 1982). However, for many plant-mycorrhizal associations, it still remains to be empirically determined whether this relationship is necessary and to what extent plants rely on fungi to maintain their normal growth in natural environments.

One significant issue in mycorrhizal research is the specificity of plant-mycorrhizal association. Some plants can form mycorrhizae with many different fungi while others with only a few. Similarly, some fungi can form mycorrhizae on many plant hosts while others with only one or a few. For example, trees such as oak, beech and birch form mycorrhizal relationships with only one or a few fungal partner(s). *Lactarius deliciosus* is typically associated with *Pinus pinea* while *Suillus granulatus* and *Russula emetica* with *Pinus pinaster* (Gardes & Bruns, 1996a, Taylor & Bruns, 1999). However, many photosynthetic plants seem to be able to form symbiosis with multiple unrelated EcM fungal species and many EcM fungi seem to be able to form symbiotic associations with multiple unrelated plants. Such broad associations provide significant potentials for mutual support among plants and fungi in natural ecosystems (Horton & Bruns 2001).

An estimated > 95% of all terrestrial plant species form mycorrhizal associations (Trappe, 1987). The mycorrhizal fungi are broadly distributed across different phylogenetic groups and ecological niches. Such broad distributions are indicative of the importance of mycorrhizae in plant communities. As a result, mycorrhizal research has attracted a lot of attention from scientists in many disciplines such as mycology, plant biology, population biology, and ecology (Trappe, 1987). The focus of this review will be on ectomycorrhizae.

There are several notable features about EcM fungi. First, EcM fungi are polyphyletic. They are composed of many phylogenetically diverse species. More than 5,000 fungal species have been estimated to form ectomycorrhizae worldwide (Amaranthus, 1998). They include species from many genera, family, class and order in the phyla Basidiomycetes, Ascomycetes, and Zygomycetes (LePage *et al.,* 1997). Secondly, EcM fungi have broad geographic ranges and are found in most regions of the global ecosystem. The main mycorrhizal centers of diversity and significant research are in temperate and tropical forests, such as those in southwestern China, northeastern China, northwest pacific coast of North America, and northern Europe. Thirdly, EcM fungi contain many of the high-valued gourmet mushrooms such as truffles, matsutake, and chanterelles. Therefore, understanding their basic biology, systematics, population structures, and major reproductive strategies for EcM fungi can aid us develop better strategies to maintain their biodiversity and growth in natural ecosystems.

In contrast to ectomycorrhizal fungi, endomycorrhizal fungi fall into a single phyletic group, the Glomeromycota. The contrasting pattern and the

polyphyletic nature of EcM fungi have raised many fundamental questions of EcM fungi themselves (Horton & Bruns, 2001), including (i) how diverse are EcM fungi in typical ecosystems? (ii) how many species are there and what is the most abundant species in EcM community? (iii) how specific are mycorrhizal fungi–plant symbioses? (iv) what effects do EcM fungal populations exert on their local ecosystem? And (v) what is the history of EcM fungi? How have they evolved and diversified?

To address the above questions, both intrinsic and extrinsic difficulties need to be overcome with EcM community studies. Intrinsic hurdles come from EcM fungi themselves. First, as EcM fungi do not grow normally without their hosts, it is extremely difficult to simulate their growth in laboratory settings. Secondly, large numbers of underground EcM fungal species are still not described or identified. In early mycorrhizal research, only fruit bodies were used for study. This is because the vegetative structures of these fungi (mycorrhizae and mycelia) lying underneath the ground are small and morphologically inconspicuous, making it hard to obtain and distinguish. In addition, some EcM fungi do not produce fruiting bodies and they are rarely investigated (Horton & Bruns, 2001). Indeed, modern taxonomy of fungi was constructed mostly based on the analyses of fruiting structures, without genetic markers, matching species names to underground structures is problematic, and in certain instances, close to impossible. The multiple stages of EcM development also complicate research. Often, different stages and aspects of EcM fungi are investigated by different methods, and sometimes by different scientific disciplines. Some of the methods

differentially used by different groups of scientists include morphological versus molecular methods in terms of identifying species, choices of different molecular markers, sensitivity in different molecular methods, dilemma in phylogenetic analyzing, etc. These challenges are discussed in the following text.

Characters and Techniques Used for Studying EcM Fungal Communities

Morphology – Visible Characters for Sorting EcM Fungi

Morphological features are generally the first pieces of information we use for macro-fungi identifications. These features can be very easy to apply to and require little equipment and investment. Analysis of different morphotypes in the root tips can also be used to sort different groups of EcM fungi. This is because certain mycorrhizae contain signature attributes shared by a specific group(s) of fungi. However, more often than not, morphological features of mycorrizhae are ambiguous, especially among closely related species, making it hard to resolve unknowns to different species or even genera.

One common method to identify EcM fungi is to analyze morphological features of the fruiting bodies, relying on traditional fungal taxonomy that are based on characteristics of sexual reproductive structures. Commonly seen macroscopic characteristics of mycorrhizal fruiting bodies include size, shape, color of the cap, stem, flesh, gill, and concentric rings, etc; microscopic features include spore size, shape, and hyphae type, as well as the arrangements of spores and hyphae within the fruiting bodies etc. Consistent differences in these features have been used to identify unknowns, often down to the genus level. However,

identification of individual species usually requires more effort, since many morphological features are undistinguishable between closely related species. The presence of juices upon breaking, bruising reactions, spore prints are considered as additional methods to sort species as well as identify unknowns.

When conducting population level analyses, macroscopic morphological features are extremely useful for sorting distinctively different specimens. However, sometimes microscopic morphological sorting could become impractical when analyzing a large number of samples. In addition, if morphological analysis takes too long, the DNA in the samples may become degraded, resulting in failure in the subsequent molecular analyses. There are drawbacks in using morphological features alone for EcM fungi identification. Specifically, while macroscopic morphology can be used to sort fungi into discrete groups rapidly, it is often not sufficient enough to differentiate closely related species. Often, molecular analyses reveal multiple reproductively isolated cryptic species within morphological species (Bidochka *et al.*, 2005; Dettman *et al.,* 2003; Geml *et al.,* 2003; Kauserud *et al.,* 2006, 2007; Taylor *et al.,* 2000). Furthermore, convergent evolution of certain trait features among unrelated species could lead to misidentification, when only morphological features are used.

Early mycologists developed many valuable morphological identification keys for fungal taxonomy. Some of these features were later compiled and analyzed for their usefulness. For example, in the analyses by Luoma *et al.,* 1997, 200 morphological types of ectomycorrhizal truffles and mushrooms from 189

soil cores were provided that included detailed descriptions for each morphotype (Luoma *et al.,* 1997). Unfortunately, unknown samples' classification remains unclear due to a lack of report of their specific morphologies.

Molecular Methods – Fine Scale Sorting of EcM Fungi

Due to the limitations in morphological analyses, molecular analyses have become increasingly common in fungal taxonomic and ecological research, including research into EcM fungi. In certain cases, morphotyping may be skipped entirely. The emerging field of metagenomics analyzes DNA samples directly from the environment, including both cultivable and un-cultivable ones (for a review in this area, see Xu 2006). These developments are allowing us unprecedented access to microorganisms in nature. Below we describe and discuss some of the common molecular methods in EcM research.

Restriction Fragment Length Polymorphisms (RFLP)

RFLP is one of the most popular molecular methods to discriminate species or strains of fungi. It usually involves using restriction enzymes to digest genomic DNA, and analyzing the resulting patterns, with or without specific probes. Although sequencing can characterize DNA more thoroughly, RFLP analysis has been very popular, especially during the early years of applying molecular markers to fungal studies. Depending on the specific DNA fragment and restriction enzyme combination, RFLP can cluster unknown specimens into different groups, sometimes to the species level. The early success of RFLP includes low cost, fast and efficient. In whole genome digestion, RFLP cannot be used to detect polymorphisms for low copy number genetic elements. However,

polymorphisms within high copy number genetic elements, for instance ribosomal DNA gene and mitochondrial DNA, can be detected using RFLP (Xu, 2005). However, there are disadvantages associated with RFLP. For example, when the number of DNA bands is high, it is hard to distinguish two bands with similar migration abilities on the gel. In order to identify individual bands, specific labeled probes need to be used to recognize unknown bands through DNA-DNA hybridization (Xu, 2005). Alternatively, specific DNA fragments can be amplified using highly selective primers through PCR and the PCR products can be then digested, and the resulting patterns compared side-by-side. Because of the highly conserved nature of the ribosomal RNA gene clusters, there is generally little or no difference among strains within the same species for this gene region. As a result, different restriction banding patterns for this gene region are usually indicative that the analyzed strains are of different species.

The main shortcoming of the RFLP typing is that when it is used alone, some samples cannot be successfully distinguished. In addition, there are several issues associated with RFLP pattern matching. Firstly, the databases are almost exclusively constructed using sporocarp samples but rarely nonsporocarp samples (i.e. in our case the EcM mycelia). The second problem is accuracy of RFLP pattern matching. For example, in the case where we could match a specific RFLP pattern to a species in the RFLP database, there might be some minor variations in fragment sizes that we fail to detect. Factors such as the choice of primers, the types of restriction enzymes, homogeneity of the gel, variation in electrical current can also affect the fragments' migrations through the gel matrix. Thirdly,

RFLP is sensitive to intraspecific genetic variation due to single nucleotide polymorphism (SNP). While such polymorphisms can be useful for strain typing, they can also confound species identification in the absence of a robust database. Lastly, typical RFLP databases are limited in scope, often by the investigator's personal interests. Large improvement in many aspects such as increasing the sample size, standardizing the use of restriction enzymes, primers and intrinsic conditions of gel matrix to allow cross-lab comparisons. One of the most commonly used RFLP typing genomic region is the internal transcribed spacer (ITS) region of the ribosomal RNA gene cluster. Both species and strain-specific ITS-RFLP patterns may be identified. As morphotyping and RFLP typing are becoming more inclusive, the integration of these two types of information could significantly enhance our ability for species and strain identifications.

PCR Based Molecular Method— Fast and Accurate

To complement the shortcomings of the morphology-based approach and RFLP method, rapid PCR-based molecular analyses are often adopted and implemented in many fungal genetics research. PCR allows amplification of specific DNA fragments with either specific or non-specific primers that recognizes defined regions of the genomes (Mullis & Faloona, 1987). Because it requires very little genetic material and is typically gene specific in its amplification, the PCR methodology provides tremendous advantages over traditional molecular typing techniques. The large amount of products obtained in a typical PCR reaction allows scientists to conduct further analyses, including obtaining sequence information of the target genes. With sequence information in

hand, many analyses can be performed including database searching and unknown-target mapping, sequence comparison and phylogenetic analysis. In EcM fungal community research, PCR has been used primarily for identification. However, other types of information, including the relative abundance of individual species are also possible from analyzing PCR products. For example, the ITS regions can be amplified through PCR from an ectomycorrhizal fungal community. The PCR products would contain a mixture of ITS sequences from different fungal species in the community. The PCR products can be directly sequenced using the pyrosequencing technique or cloned first into a host bacterium and then sequenced individually through conventional sequencing techniques. The obtained sequences can be analyzed that allow calculations of relative abundances of individual ITS sequence types (Xu, 2006).

DNA Sequencing — The Finest Scale Taxonomic Identification

Sequencing one or several DNA markers is the most sensitive and robust way to identify species. There are many advantages of using DNA sequence based analysis for species and strain identifications. First, DNA sequences are unambiguous and some regions are highly conserved within species and can be used as molecular markers to characterize unknowns (Xu, 2005). Second, compared to morphological features, which represent individual phenotypes, DNA sequences make up genotypes at a very fine scale. Compared to pattern-based typing (RFLP, PCR fingerprinting), DNA sequence based typing can give more accurate and robust results (Xu, 2005). Third, DNA sequences can be stored in and retrieved from public databases. Such databases can serve as valuable

resources for searching and editing (Xu, 2005). It also allows us to do fine scale genome typing, for example, intraspecific strain typing. Indeed, DNA-sequence based approaches are the future of EcM research.

ITS—An Excellent Molecular Marker to Identify Species

The ITS region is a section of DNA located within the nuclear ribosomal RNA gene cluster (Figure 1.1). Each unit of the gene cluster comprises one small subunit 18S rRNA, a large subunit 28S rRNA, and a 5.8 S rRNA flanked by two non-coding DNA sequences known as ITS1 and ITS2 on either side. The combined length of ITS regions (including the 5.8S rRNA) in fungi is typically between 650-900bp. Within an individual cell, there may be 50-200 copies of this unit cluster, linked by an intergenic sequence (IGS) that can be highly variable in length and sequence composition among species.

During the past two decades, the ITS region has become the most commonly used molecular marker in fungal ecology and systematic research (Egger *et al.,* 1995; Chambers *et al.,* 1998; McKendrick *et al.,* 2000). Its utility and popularity are due to the following reasons. First, it is present within the high copy number ribosomal RNA gene cluster as tandem repeats. As a result, a small number of cells may be sufficient for PCR and subsequent analyses. Second, it is highly conserved within species but can be highly variable between closely related species. Third, the ITS regions are relatively short and flanked by highly conserved sequences. Consequently, conserved primers applicable to broad groups of species can easily be designed and used to amplify the ITS regions by PCR. Two primers, ITS1 and ITS4, have been used widely in fungal systematics

and population genetic studies. These two primers were originally designed from plant sequences and are used as universal primers across all fungi (White *et al.,* 1990, Gardes *et al.,* 1991).

Aside from ITS1 and ITS4, other taxa - specific primers have also been developed. For example, ITS1f and ITS4 are fungal specific; ITS1f and ITS4b are basidiomycete specific (Gardes & Bruns 1993). These group specific primers are desirable in terms of enhancing specificity of one group and discriminating against other fungal groups. For instance, primer ITS4b has been designed to amplify sequences from all known basidiomycetes, while excluding fungi outside the group (Taylor *et al.,* 1999). Nevertheless, even though ITS1f/ ITS4b are specifically designed ITS primers for basidiomycetes, we do not expect that they will amplify every species within this group. This is because mutations could be present at the primer annealing sites for some potentially unknown basidiomycetes in nature. Indeed, non-amplified individuals with the above primers have been identified in previous experiments using other techniques. *Rhizoctonia* is one of the groups. Some species in the *Rhizoctonia* group cannot be amplified by the ITS1f/ ITS4b primer pair, even though it belongs to basidiomycetes (Robinson *et al.,* 2009). Indeed, due to their unknown nature in our databases, group-specific primers based on known taxa in that group will always have its limitations when working with unknown samples. One way to overcome the difficulty faced by group-specific primers is to use universal primers such as ITS1/ITS4 to reanalyze non-amplified species and to confirm the results. Bearing these caveats in mind, it's important to realize that care should be

taken when interpreting the results.

The ITS region is currently the most popular DNA marker used in RFLP, especially for studying EcM fungal diversity. Matching the polymorphisms at the ITS regions has successfully separated many EcM fungal species (Egger *et al.,* 1995, Kårén *et al.,* 1997; Pritsch *et al.,* 1997). Indeed, ITS-RFLP has turned out to be the most convenient molecular method to sort fungal species with minimal technical requirements. As mentioned above, there are several advantages associated with ITS. In combination with PCR and RFLP, ITS PCR-RFLP has allowed generating species-specific banding patterns and making it possible to classify EcM fungi at the species level. This is because the ITS region usually contains restriction sites where common restriction enzymes could target. Commonly at least 2 to 3 frequent cutting restriction enzymes are needed for species identification. Since sequence differences between species can be due to both insertions/deletions and nucleotide substitutions, different restriction enzymes may produce very different patterns. If several ITS PCR-RFLP patterns are associated with a single morphotype, additional samples with the same morphortype need to be selected and verified by RFLP. If such patterns are confirmed, taxonomic revisions may be necessary.

Nowadays, sequencing ITS regions has been routinely used to identify fungal species, and to study genetic diversity among different strains within one species. The ITS region is now perhaps the most widely sequenced DNA region in fungi (Bruns *et al.,* 1998; Pritsch *et al.*, 2000; Sha *et al.,* 2008; Xu, 2005). A large number of ITS sequences have been stored in GenBank. It has become a

valuable resource for identifying unknown fungal species and for investigating the evolutionary relationships among fungal species (Bruns *et al.,* 1998; See also below for the UNITE database). Because of these and the previously mentioned features, the ITS regions have been adopted by the international mycological community as the barcode region for fungal identifications.

Other than the ITS sequences, additional molecular markers such as LSU rRNA, mtLSU, SSU rRNA, 5.8S nuclear rRNA are also widely used to discriminate fungal groups (Table 1.2). Due to the difference in the degree of resolution, unknowns can be placed into different taxonomic levels using different rRNA gene fragments (Table 1.2). The choice of which marker to use depends on the experimental goal, feasibility of PCR amplification of each region, and the availability of the sequence database of that particular marker for comparison and analyses. In many investigations, fine scale species and strain identifications are the final objectives. However, sometimes tracing the specimens down to family levels can provide sufficient information to answer certain phylogenetic questions. Species and strain - level identifications often require large databases (large taxa sets and multiple genes) to compare. Sometimes, even with large databases, there might be no matching with those in databases. Indeed, it is not uncommon that the ITS-based molecular identification can only reach the genus level (Gardes & Bruns, 1996b; Ciardo *et al.,* 2006).

When considering markers in addressing specific research questions, we need to realize that there is always a trade–off between cost and effectiveness in the type and number of markers selected for investigating certain issues.

Relatively sensitive markers such as the ITS region works well for identification to the species level. However, for genotype-specific fine-scale identifications, other types of markers (e.g. single copy genes or inter-genic regions) are needed. With the increasing availability of candidate molecular markers, the best-fit candidate marker(s) may require prior screening for individual species before we can determine their appropriateness. Conversely, a less sensitive marker (e.g. the SSU rRNA gene) does a cruder sorting, but mapping and identification steps are easier and more accurate, since fewer samples may be needed to take into account. Another shortcoming for the ITS region is its high copy number, which makes it hard to distinguish whether heterozygosity is resulted by different copies of ITS on one chromosome or on two different chromosomes (in diploid or dikaryotic individuals). Therefore it is hard to identify genotypes for individual alleles when the unknown is heterozygous.

Single-Gene Poorly Revealing Phylogeny

The main drawback of single-gene typing is that it may not truly represent organism's phylogeny (Doyle, 1992; Maddison, 1997). For example, using a single gene as molecular marker to study organism phylogeny can be highly biased, since the gene chosen for analysis may evolve fast in some groups but relatively slow in other groups. Two species mapped together by a single marker could be actually very different from each other if analyzed with other genes. Simply speaking, each single gene has its intrinsic evolutionary bias. One way to minimize the bias is to use multiple molecular markers. Using multiple genes to

study phylogeny is always more reliable than using only one gene since it is more representative for the entire genome (Xu *et al.*, 2006).

Approaches for Examining Species Diversity in EcM Communities

If the research focus was to examine the diversity of EcM fungi within a selected geographical area, combining morphology and PCR-RFLP of the ITS region should initially provide a rough grouping among a large number of fungal isolates. For individuals with nearly identical PCR-RFLP patterns, they can be further characterized by sequencing the ITS or other particular DNA markers.

If the research aim was to study one type of EcM fungi within a selected area, morphology or PCR-RFLP of the ITS region can be used as an initial step to identify the target EcM fungi. Further identification and confirmation can be achieved by PCR amplification and sequencing of marker gene fragments at other loci. After obtaining the sequence information, blast search can be carried out in order to determine the sequence identity with known specimens in databases. If the unknown specimens have high sequence similarities (usually >97% for ITS regions) to sequences of known species deposited in the GenBank, the identity of the unknowns can be inferred. However, if a high degree of sequence identity is not found, individual sequences are then used as a query to retrieve closely related sequences with comparable length from the GenBank or other publicly available databases. Representative sequences for each closely related species from GenBank were then included to compare potential intra-specific variation within different phylogenetic groups to known species. These sequences and all the

retrieved sequences were then aligned using appropriate computer softwares. Subsequently, phylogenetic analyses applying different algorisms are used to reveal the phylogenetic relationship among the species. Usually, the most closely related known species are used as outgroups for references. Sometimes unknown fungi cannot be easily identified by examining a single gene. Sequence information of additional molecular markers could be helpful to identify unknowns. Since different genes evolve at different rates, multiple gene genealogy can limit the biases created by single genes and reveal the true phylogenetic relationships among species. In many cases, the relationships among strains and populations within species can also be revealed using sequences from multiple genes. Such sequence information can be used to determine the potential mode of reproduction and geographic patterns of molecular variations (e.g. Lan & Xu, 2006).

While GenBank, DNA Data Bank of Japan (DDBJ), and the European Molecular Biology Library (EMBL) databases contain almost all publicly deposited DNA sequences that are cross-referenced with each other through unique accession numbers, the ectomycorrhizal research community has been benefited tremendously by another database called UNITE (http://unite.ut.ee/). UNITE is an rDNA sequence database focused specifically on ectomycorrhizal fungi in two phyla, the ascomycetes and the basidiomycetes. The database was established because of widespread sequence misidentifications and inaccurate reporting in GenBank, DDBJ and EMBL. To establish an accurate and reliable database, sequences in the UNITE database are generated from fruit bodies

collected and identified by specialists and deposited in public herbaria that can be accessed by others. In addition, type specimens are used whenever possible. As of April 2009, the UNITE database contains 112363 fungal ITS sequences. Among these, 2736 were barcoding sequences from 1202 species in 128 genera. Aside from providing a robust database for identification of sequences from curate specimens, UNITE also has other features to facilitate the identification of fungal DNA directly from environmental sources (i.e. DNA sequences without specific specimens attached). This search tool has become increasingly important because direct metagenomic analysis of environmental DNA is becoming increasingly common in EcM research.

Bioinformatic analyses of EcM fungi

Based on the aligned sequences, the relationships among sequences, strains, and/or species can be revealed. The most common form of presentation for such relationships is through phylogenetic trees. Phylogenetic trees can be constructed using various algorisms implemented in different phylogenetic softwares. The most commonly used ones are maximum parsimony, neighbor-joining, maximum likelihood and Bayesian approaches. The trees generated by applying different algorisms are usually consistent with each other. However, in certain cases, they can be different from each other. The different patterns can be generated for a couple of reasons. One is the weighting scheme of polymorphic nucleotides. For example, the weights of transitional substitutions and transversional ones can have a significant effect on the final outcome of the analyses. The same can be said about insertions and deletions. Indeed, dozens of

weighting methods have been developed to try to reflect the relative importance of different types of mutations during the evolution of specific lineages.

The second reason for different tree topologies generated by different methods using the same dataset relate to the differences in algorithms among the phylogenetic tree-construction methods. Different algorisms process data in different ways. The neighbor-joining method measures the genetic distance between each pair of taxa and joins taxa with the shortest distance first, followed by progressively more distantly – related taxa (Saitou & Nei, 1987); maximum parsimony produces the most preferred phylogenetic tree invoking the fewest number of evolutionary changes (Felsenstein 1978); maximum likelihood method searches for the tree with the highest probability or likelihood that matches the data (Fisher 1978). Sometimes tree topologies are divergent at a great extent, making interpretation of phylogeny challenging. When such cases occur, the preliminary data can be used to generate specific hypotheses. Targeted additional sequence information can be then collected to test the hypotheses. In general, the greater the sample sizes and the more genes analyzed, the more robust the inferred evolutionary relationships among strains and species will be.

In conclusion, a combined morphological and molecular approach is generally used to address issues related to species diversity in EcM research. Morphological grouping is typically carried out at the initial stage of identification. Usually EcM fungi from each soil sample are classified into as many groups as possible based on their morphological characteristics. The morphological characterization can be then followed by ITS PCR-RFLP banding

pattern matching with an established reference database. However, for fungi remaining unknown after comparison to PCR-RFLP database, DNA sequencing and phylogenetic analysis of ITS and various other molecular markers are needed. Consequently, the additional data can then be used to update the ITS-RFLP database. If more than one RFLP pattern is found associated with a single morphotype, additional samples with the same morphotype are further analyzed to exclude the possibilities of contamination and potential heterozygosity within individuals for the ITS regions (Horton & Bruns, 2001).

Approaches for Examining Intraspecific Variations in a EcM Community

Studies involving the examination of genetic variation within EcM fungal populations typically rely on multilocus polymorphisms using several types of markers, including single-copy gene based RFLP, PCR fingerprinting, DNA sequencing, and microsatellite DNA. These markers have been described in detail in an earlier review (e.g. Xu, 2005). Here, we briefly mention microsatellite markers. Microsatellite markers are those with variation in the number of simple sequence repeats within a DNA fragment that can be found among individuals within species. Because of their repetitive nature, they typically mutate much faster than single nucleotide substitutions, due to high frequency slippage during DNA replication. The high variability of microsatellite DNA makes them excellent molecular makers for strain typing and population analysis, especially for recently evolved populations. Together with the increasingly popular single nucleotide polymorphism, microsatellite DNA are helping scientists working on EcM fungi to determine species boundary, population structure, and reproductive

strategy etc. Below we review recent studies on one specific issue, that of genet size of EcM fungi, in natural environments.

Genets – Size and Distribution

A genet is a group of individuals (fruiting bodies and/or underground mycelia for EcM fungi) produced from one mating event and that occupy the same geographic area (Dahlberg & Stenlid, 1994; Xu, 2005). Except for a small number of mutations, the individuals within a genet should be all genetically identical. Because they arise vegetatively from a single mating event, this shared descent among the individual fruiting bodies and mycelia implies connectivity both nutritionally and/or genetically. Nutritional connectivity refers to the nutrients absorbed by mycelia from one location of the genet can be transported to another parts of the genet. Similarly, genetic connectivity means that mutations and horizontally transferred genes obtained in one area of the genet can be spread to other parts of the genet. At present, though neither connectivity has been demonstrated conclusively in nature, the existence of genets of various sizes suggests their likely importance in nature.

Significant research activities have been devoted to characterize the size of genets in EcM fungi. A variety of molecular markers mentioned above have been used to determine whether individual fruiting bodies or mycelia of one species from a defined geographic area are genetically identical with each other but different from other individuals within the general region. Due to their high mutation rates, the polymorphisms at multiple microsatellite markers have been used extensively to robustly identify the size of individual genets. Typically, a

single genet is identified by a unique microsatellite polymorphic type for individuals sampled at a defined geographic area. Similarly, genetic identity inferred using other types of markers such as random amplified polymorphic DNA (RAPD) is also widely used, especially during the early phase of ectomycorrhizal research.

The current data suggest that the size of genets vary widely among fungi. Small genets in species of the ectomycorrhizal genus *Russula* can be less than half a meter. In contrast, large genets have been found among plant pathogenic basidiomycete genus *Armillaria.* For example, in *A. gallica*, some genets can occupy up to 2,200 acres (15ha). For many fungi, sporocarps belonging to one genet usually are spatially arranged as a ring, known as a fairy ring. The fairy ring structure is generated from a single mating event. As mycelia grow outwards, the genet expands. Furthermore, if the spatial environmental conditions are relatively homogeneous, the genet expands at similar rates in all directions across the terrain. At certain times of the year when conditions favorable for mushroom fruiting appear, fruiting bodies are produced along the edge of the mycelial growth, forming a ring structure. Because mycelial growth depletes nutrients in the center of the ring, the size of the ring also expands over time. One striking feature of EcM fairy rings is the low species richness along the rings, dominated by the ring-forming species. However, fungal diversity can be high on both sides of the rings, with similar species diversities and compositions (Lian *et al.,* 2006; Hirose *et al.,* 2004). This fact suggests that EcM communities are capable of recovering soon after genet passage. However, questions such as how recovery is

established soon after passage and how the dominance is achieved by the ring-forming species are still unclear.

Current molecular ecological investigations suggest significant variation among EcM fungi in their genet sizes. Table 1.3 summarizes the genet sizes of recently investigated EcM basidiomycete fungi. Genet size is typically measured in the largest distance between sporocarps that are produced by one mating event. As shown in Table 1.3, genet sizes vary greatly among species. *Laccaria* spp. and *Russula brevipes* have so far the smallest identified genets of less than 1m in diameter (Baar *et al.,* 1994; Bergemann & Miller, 2002; Selosse *et al.,* 1999). Genets of *Tricholoma matsutake* and *Suillus grevillei* are less than 3m in diameter (Lian *et al.,* 2006; Zhou *et al.,* 2001). Genet sizes greater than 20m are found in several species of the genus *Suillus* (Hirose *et al.,* 2004). To estimate genet expansion rate, spatial distribution of sporocarps within the same genet is typically monitored through several years. The results from previous studies show that genet expansion rates also vary among species. For example, *Tricholoma matsutake* expands relatively slowly with an average rate of 10.3cm/yr, while *Hebeloma cylindrosporum* has a much higher rate of 45-60cm/yr (Gryta *et al.,* 2000; Lian *et al.,* 2006). Some of *Laccaria* spp. genets have extremely high expansion rates, around 100cm/yr (Gryta *et al.,* 1997, 2000; Bonello *et al.,* 1998; Guidot *et al.,* 2001). Taken together, these estimates suggest that EcM genets can live up to about 30-40 years in natural ecosystems.

Early-stage versus Late-stage EcM Fungi

Typical natural ecological communities, including the EcM communities, are

established through a series of changes of biological compositions and abiotic factors. The process of ecological change involving a series of natural communities that are established and replaced over time is called a succession. There are two kinds of ecological successions, primary succession and secondary succession. Primary succession occurs in an environment in which new substrates, devoid of any living organism and usually lacking soil, is deposited (for example a lava flow) and allows the establishment of an ecological community. In primary succession, pioneer species like mosses, lichen, algae and fungi first colonize the substrate. Together with abiotic factors such as wind, water and the heat/cold cycle, these pioneer species change the substrates, making them suitable for subsequent growth of plants. The plants then dominate but often replaced successively by plants better adapted to less austere conditions. Examples of primary succession can be found on a new lava flow, an area left from retreated glacier, or abandoned strip mine.

In contrast, secondary succession is a response to a disturbance, for example, forest fire, tsunami, hurricane, flood, or an abandoned field. While primary succession takes place in an area that is originally without any living organism, secondary succession occurs in an area where life once existed but has since been destroyed or disturbed, for example by fire, tornado, harvesting or other human activities such as agriculture, that reduces an already established ecosystem (e.g. a forest or a wheat field) to a smaller population of species. As such, secondary succession occurs on preexisting soil, unlike primary succession that usually occurs in a place lacking soil.

EcM fungi can be found associated with both primary and secondary successions. EcM fungi associated with primary succession plants are called early-stage EcM fungi. Some typical early colonizing fungal species are *Hebeloma cylindrosporum* (Gryta *et al.,* 1997, 2000), *Laccaria bicolor* (Baar *et al.,* 1994) and *Pisolithus tinctorius* (Anderson *et al.,* 1998). EcM fungi found in secondary successions are called late-stage EcM fungi. Typical EcM fungi in late-stage succession are those in families Russulaceae, Cortinariaceae and Amanitaceae. Early colonizing fungal species typically colonize by spores. Because primary succession niches are typically poor in nutrients these early EcM fungi are expected to have relatively small and non-persistent genets (Deacon *et al.,* 1983; Fox 1983). Early-stage fungi are considered as R-selected species, because they produce many offspring and are capable of filling available niches in an environment very quickly (Deacon & Fleming, 1992). In contrast, late-stage EcM fungi may also colonize initially by spores or by dormant mycelia underground, but because of nutrient availability from the soil, they are able to spread by mycelial growth and expand. As a result, their genets are expected to be large and temporally persistent (Dahlberg & Stenlid, 1990). Late-stage EcM fungi may be considered K-selected species, more capable of dealing with environmental stresses (Cooke & Rayner, 1984; Grime *et al.,* 1979). While the typical late-stage EcM fungi in genera *Russula, Amanita, Lactarius*, are expected to propagate by mycelial growth, researchers have found that their genet sizes and age vary significantly, suggesting that proliferation by sexual spores in these late-stage fungi might be more important than previously expected (Bergemann &

Miller, 2002). Even in mature plant communities with relatively homogeneous ecological conditions, the genet sizes can vary significantly, suggesting reproduction by spores is a prominent feature of late-stage EcM fungi. Thus, the appearance of a species in secondary succession communities cannot be used to draw conclusions about their genet size and modes of reproduction and colonization (Bonello *et al.,* 1998).

Reproductive Modes and Persistence of Mycelia Networks

The mode of reproduction for EcM fungi is typically inferred based on the associations of alleles at the same or different loci. Populations with alleles randomly associating with each other are considered sexually reproducing populations while those with significant signatures of non-random associations are considered asexual populations (Xu, 2005). Another indicator in EcM fungi about the relative importance of sexual and asexual reproduction is the size of genets. By examining the genet sizes of EcM fungal species in nature, the reproductive strategies of individual species may be predicted. It has been suggested that species forming many small genets likely reproduce mainly by spore colonization and the species that form large genets mainly propagate by underground mycelial extension (Anderson *et al.,* 1998; Bonello *et al.,* 1998; Dahlberg & Stenlid, 1990, 1994; Dahlberg, 1997; Zhou *et al.,* 1999). Most species use a combination of both strategies. For example, *Suillus* spp. use a mixed strategy: they form many small genets by spore colonization at the early stage of lifecycle, while then produce large genets by mycelia expansion at late stages (Dahlberg & Stenlid, 1990).

Studies of *Suillus* spp. also revealed that the size of genets is correlated with age of host stands (Dahlberg & Stenlid, 1994). Small genets are mainly found in young-aged stands, while large genets are found in mature stands. However, exceptions have been found in *Lactarius xanthogalctus* and *Russula cremoricolor*, which form many small genets in mature tree forests (Redecker *et al.,* 2001). The age of stands is thus not sufficient to predict colonization strategies across all EcM species. The spatial pattern of EcM fungi also depends on many other factors, such as (a) environmental physical parameters, e.g. temperature, moisture, and light; (b) edaphic factors, e.g. soil moisture, depth of organic matter, and soil pH (Erland and Taylor, 2002); (c) biological traits of the fungi; and (d) competition with other fungal species (Anderson *et al.,* 1998; Dahlberg & Stenlid, 1990, 1994; Dhlberg, 1997). Guidot *et al.* demonstrated that *Hebeloma cylindroporum* forms large genets under conditions with low competition with their neighbor fungi (Guidot *et al.,* 2001). Conversely, small genets are observed under high intensities of competition. The capacity to survive and expand for a long period of time in the soil as generative mycelia and the ability to colonize using spores differ greatly among species. In general, in the laboratory settings, fungi tend to reproduce sexually when environmental conditions are unfavorable, and reproduce asexually when environmental conditions become favorable. The fruiting seasons for many EcM mushrooms are associated with the onset of stressful environmental conditions, consistent with laboratory observations.

Genet expansion rates, also known as mycelial expansion rate vary among species in a great range from 10cm/yr to 110cm/yr (Table 1.3). The variation could be due to a combination of genetic and environmental factors, such as mentioned above (Bonello *et al.,* 1998; Lian *et al.,* 2006). In the studies of most EcM species, genet expansion rates are relatively constant within species through several years of growth (Bonello *et al.,* 1998; Dahlberg & Stenlid 1994; Lian *et al.,* 2006; Zhou *et al.,* 1999). While EcM genet establishment and growth rates are affected by environmental factors, however, the specific mechanisms remain largely unknown. Future research should be focusing on monitoring genet progression in various environments.

The ability of EcM mycelia to expand is a unique characteristic for filamentous fungi, therefore studying genet size can understand the ecological features of these species in nature. The characteristics of EcM genet can be used as indicators of succession stage of forests, individual population boundary, and the required area of study site in EcM research (Liang *et al.,* 2004a). It can also help understand the roles of spore dispersal and mycelial growth in the life histories of EcM fungi. For example, as a late-stage EcM fungi, species in *Russula* maintain reproduction by sexual spores even in suitable environmental conditions with minimum competitors, indicating that spore colonization plays a much more important role in the life history of *Russula* even in mature forests (Redecker *et al.,* 2001).

Though still preliminary, the data in Table 1.3 also suggest that evolutionary relationships of species may play a role in genet size. Specifically, if the

phylogenetic background were important, we should expect to see similar genet sizes among closely related species. Based on the limited data available, species in certain genera, e.g. *Suillus*, all seemed to have relatively large genets. However, this may not hold in other taxonomic groups. For example, distinct genet sizes and propagation strategies are often found among closely related taxa in other genera, e.g. between two sister species of *Rhizopogon* (Kretzer *et al.,* 2003; Table 1.3). Further analyses of more taxa in representative environmental conditions are needed to critically test this hypothesis.

Ecological Importance of EcM Fungi Diversity

Aside from genet size differences, EcM fungi can vary significantly in other aspects, including the substrates they inhabit, the ability to uptake various nutrients, and the tolerance to water stress and temperature extremes. Because different groups of EcM fungi have different adaptive features, we expect that plants associated with a greater diversity of EcM fungi in their roots to be fitter and more capable of surviving and reproducing. In contrast, plants associated with a single type of EcM fungi may be vulnerable to environmental stresses. However this hypothesis has not been tested either in laboratory or field conditions, therefore whether high diversity of colonizing fungi could enhance the survival of plants remains unclear.

Previous researches on EcM fungi community structure have demonstrated that the abundance and rarity in EcM community seemed to show reverse relationships (Horton & Bruns 2001). This pattern means that the EcM

samples that we randomly select in nature are likely highly biased. The high number of rare EcM species could mask the actual abundance of common EcM species in a selected EcM community. The bigger the sample size, the more species and genotypes will likely be identified. Whether a sample size would be representative of the EcM community structure can be determined using saturation analysis. In this analysis, when the number of samples goes up, the number of species also go up, to a point when it becomes stable (McCune & Mefford, 1999). The number of samples before the curve reaches plateau is the suggested one that can robustly represent the EcM fungal diversity. Jackknife estimate can also be used to predict the number of taxa needed to resolve the community structure (Mueller-Dombois & Ellenberg, 1974; Palmer *et al.,* 1991).

Based on studies on EcM community conducted so far, the most dominant EcM fungi fruiting bodies are found belonging to Russulaceae, Thelephoraceae, and non-thelephoroid resupinates (Gardes & Bruns, 1996a; Jonsson *et al.,* 1999c; Luoma *et al.,* 1997). At present, the reasons for their numerical and potentially functional dominance are unknown. However, it has been suggested that for most EcM fungi, the ones that fruit the most abundant are not necessarily the ones most abundant as mycorrhizae (Gardes & Bruns, 1996a; Jonsson *et al.,* 1999 a,b; Mehmann 1995). One possibility is that species that are abundant as mycorrhizae on roots but fruits infrequently likely require more energy for sexual reproduction than for vegetative growth. Direct analyses of metabolisms and energy consumptions are needed to test this hypothesis.

A better understanding the community structure and the relationship between hosts and symbiotic EcM fungi will allow us better able to develop sound strategies to conserve the diversity of EcM fungi. This in turn will help construct a suitable forest management practices to protect forest ecosystems. In these developments, two specific issues warrant consideration: (1) different EcM fungi play different roles in plant growth; and (2) different EcM communities are associated with different aged forests (Molina & Trappe 1982; Sylvia *et al.,* 1997). Often, diverse EcM fungi are found associated with mature trees, and this association can benefit nearby seedlings (Sylvia *et al.,* 1997). It has been demonstrated that a greater EcM diversity exists for seedlings located closer to mature trees than those closer to immature or newly planted trees (Sylvia *et al.,* 1997). This result suggests that mature trees can be used as a reservoir to help maintain diverse EcM fungi for the broader forest community. Furthermore, Taylor *et al.* (1989) identified a high diversity of EcM fungi in the trees whose nearby trees are from other species. Therefore, planting a mixture of tree species can also be a great way to improve EcM diversity. In these mixed forest ecosystems, species with broad host ranges will likely be more competitive than those with narrow host ranges. Therefore, during or after timber harvesting, representative host plants should be retained and/or replanted.

Concluding Remarks

We outlined a set of methods for identifying EcM fungi and illustrated their utilities in EcM research on genet size, reproductive modes, and ecology. A

combination of morphological and DNA-based molecular techniques is often required to identify large numbers of unknown EcM fungi in natural ecosystems. Samples of both sporocarps and mycorrhizae are needed to help obtain a clear picture of the richness and abundance of EcM fungal species in a particular study site. At present, the full extent of EcM fungal diversity remains largely unknown, therefore directly examination of the relationships between EcM fungal diversity and function in representative ecosystems should be carried out. In addition, the patterns and mechanisms for the specificity and inter-connectedness among different host plants, different mycorrhizal fungi, and between mycorrhizal fungi and plants will be of great interest for both basic scientific investigations and applied research. The molecular ecological analyses of more EcM fungi should help us understand the broad patterns of genet sizes and their modes of reproduction. Such understandings will help develop suitable forest management strategies.

Project Objectives

The genus *Russula* is among the most numerous and complex mushroom genera. Species in this genus are widely distributed throughout the world, from the tropics to subtropics, temperate regions, and the arctic zone. They form symbiotic associations with a variety of plants such as pines, oaks, spruce, and birches and play vital roles in forest ecosystems. Many species of *Russula* are collected as exotic food for human consumption in North America, Europe, Asia and elsewhere. This is specially true in Yunnan Province in southwestern China

where favorable climate stimulates the production of many highly prized edible mushrooms including Chinese *Russula,* also called "Da Hong Jun". Despite its economical and ecological importance, relatively little is known about this species. No formal taxonomic investigation has been conducted on dahongjun and there is little molecular information for Chinese *Russula* in any public databases. My thesis project aims to understand the commercially harvested *Russula* in Yunnan in the following three aspects.

The first goal is to identify what *Russula* species dahongjun belongs to, and whether it is one species or a species complex. We also tried to determine the phylogenetic relationships between dahongjun and other known *Russula* species.

The second goal of this thesis is to investigate the genetic diversity and population differentiation among different dahongjun populations from their broad distribution ranges. We hoped to gain an understanding of the population structure and gene flow of dahongjun at both local and regional levels.

The third goal of this thesis is to estimate genetic variations within a local dahongjun population in order to infer the normal genet size for Dahongjun. The haplotypes of ITS and other self-devoloped DNA markers were used for analysis of gentic variations.

Because species in *Russula* cannot be artificially grown, in recent years, there has been a noticeable decline of production for many of the species due to over-exploitation of their natural populations. The results of this research could be valuable for helping maintain and conserve the natural populations of these species.

Table 1.1 Contrast between hosts and EcM fungi in terms of living benefits from each other

Plant hosts benefit from EcM fungi	**EcM fungi benefit from its host**
Supplying phosphorus, nitrogen, water	Supplying soluble carbohydrates
Protecting against pathogens	Providing niche space
Creating strong soil structure	Supplying amino acids
Facilitating nutrient transfer among plants	Ensuring ecological stability and evolutionary selectivity in nature
Enhancing cooperation and lowering competition among plants	Associating and interacting with other symbiotic microbes such as nitrogen fixers

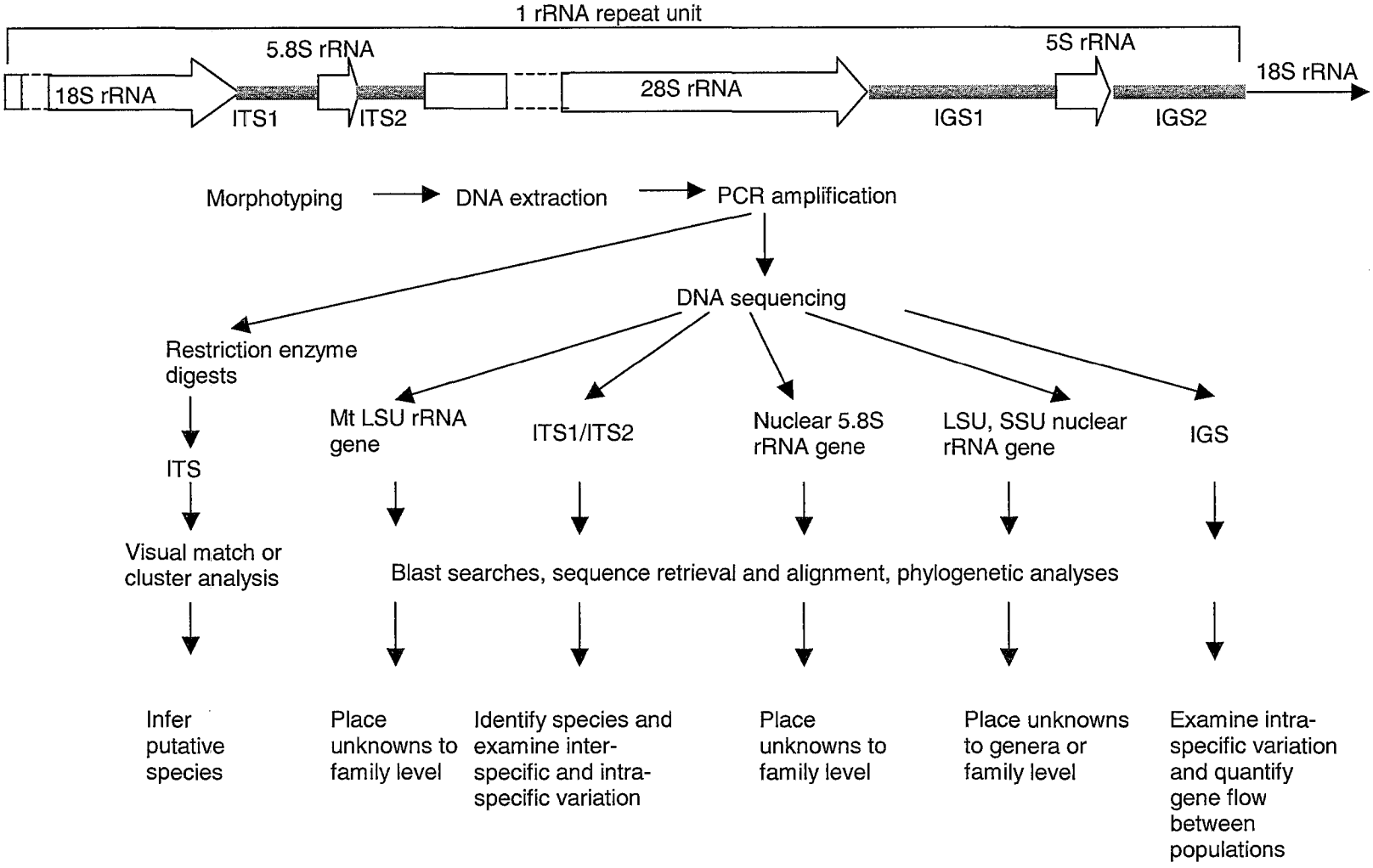
Table 1.2 Comparison of molecular markers within the ribosomal RNA gene cluster used for identification of unknown EcM fungi.

	Advantage	Disadvantage	Popularity
SSU rRNA	Most conserved, excellent for high level taxonomic investigations	Too conserved for low taxonomic level investigations	Heavily used for identifying bacterial and archaeal diversity, but rarely for studying EcM fungi
LSU rRNA	More variable than SSU rRNA, can place to generic level	Database is limited	Few EcM studies used this gene fragment
5.8S rRNA	Very conserved, but can help resolve to phylum level	Sequences too short	Usually co-analyzed with ITS sequences
mtLSU rRNA	Somewhat variable, can help place to family level	Can have introns	Some studies used this gene fragment
ITS	Highly variable, extensive database, can place sample to species level	Too variable for high level taxonomic analyses and sometimes sequence alignment is difficult	Extensively used and a large database already available

Table 1.3 Comparison of genet properties among EcM species. Genet size is measured based on the above-ground fruit body sampling. "–" indicates missing data.

EcM species	Genet size	Estimated expansion rates (cm/year)	Early or late stage?	References
Laccaria amethystina	<1m^2	60-110	Early	Gherbi *et al.*, 1999; Selosse *et al.*, 1999
Laccaria bicolor	<12.5m^2	20-100	Early	Baar *et al.*, 1994; Selosse *et al.*, 1999
Pisolithus tinctorius	<40m	-	Early	Anderson *et al.*, 1998
Hebeloma cylindrosporum	<3.5m	52.5	Early	Gryta *et al.*, 1997, 2000; Guidot *et al.*, 2001
Suillus pictus	3.4-21m	50	Both	Hirose *et al.*, 2004
Suillus variegatus	29m		Both	Dahlberg, 1997
Suillus pungens	40m(maximum 300^2)	50	Both	Bonello *et al.*, 1998
Suillus grevillei	3m	-	Both	Zhou *et al.*, 1999
Suillus bovinu	3m-30m	-	Both	Dahlberg & Stenlid, 1990
Russula brevipes	<3m	-	Late	Bergemann & Miller, 2002
Russula cremoricolor	0.38m-1.27m	-	Late	Redecker *et al.*, 2001
Russula vinosa	<1m	-	Late	Liang *et al.*, 2004b
Amanita francheti	1.5m^2	-	Late	Redecker *et al.*, 2001
Cortinarius rotundisporus	30m	-	Late	Sawyer *et al.*, 1999
Cantharellus formosus	2-13m	-	Late	Dunham *et al.*, 2003
Lactarius xanthogalactus	9.3m^2	-	Late	Redecker *et al.*, 2001
Tricholoma matsutake	2m	10.3	Late	Lian *et al.*, 2006

Figure 1.1 A combined approach to identify unknown EcM fungi using ribosomal RNA genes to different taxonomic level


1 rRNA repeat unit
5.8S rRNA
5S rRNA
18S rRNA
18S rRNA
28S rRNA
ITS1
ITS2
IGS1
IGS2
Morphotyping
DNA extraction
PCR amplification
DNA sequencing
Restriction enzyme digests
ITS
Mt LSU rRNA gene
ITS1/ITS2
Nuclear 5.8S rRNA gene
LSU, SSU nuclear rRNA gene
IGS
Visual match or cluster analysis
Blast searches, sequence retrieval and alignment, phylogenetic analyses
Infer putative species
Place unknowns to family level
Identify species and examine inter-specific and intra-specific variation
Place unknowns to family level
Place unknowns to genera or family level
Examine intra-specific variation and quantify gene flow between populations

CHAPTER 2

Genetic diversity of dahongjun, the commercially important 'big red mushroom" from southern China

M.C. Li, J.F. Liang, Y.C. Li, B. Feng, Z.L. Yang, T.Y. James, and J.P. Xu

Abstract

In southern China, a wild mushroom commonly called "dahongjun" or the "big red mushroom" by the locals, has been harvested, consumed and/or exported as an exotic food for many years. Although ecologically and economically important, very little is known about this mushroom, including its basic ecology and population genetic structure. In this study, we investigated the genetic diversity and geographic structure of dahongjun using sequences of the internal transcribed spacer (ITS) regions of the ribosomal RNA gene cluster. We found that this exotic mushroom contained at least three divergent lineages, with one corresponding to the recently described *Russula griseocarnosa* and two others likely representing closely related novel species within the genus *Russula*. Both phylogenetic and population genetic analyses suggested that dahongjun populations were geographically structured. Our study provides the first insights into the diversity and population biology of this endemic gourmet mushroom in southern China.

Introduction

Fungi are important constituents of the global biosphere. In forest ecosystems, many fungi form symbiotic associations with plant roots, establishing a structure called mycorrhizae. Over 90% of land plants form mycorrhizal associations with fungi. These fungi help plants obtain essential minerals and water from the soil and can contribute to plants' disease resistance and drought tolerance (Brundrett, 2004). While most of the fungal materials are subterranean and are not typically observed by the naked eye, some mycorrhizal fungi, especially those belonging to Basidiomycota, are easily noticeable during certain times of the year because they produce conspicuous mushrooms. Some of these mushrooms are collected as a source of exotic and highly - prized food for humans.

The mushroom genus *Russula* is a highly diverse group of ectomycorrhizal (EcM) fungi (Russulaceae; Russulales; Agaricomycetes; Basidiomycota). Globally, about 750 species have been identified and characterized so far (Kirk *et al.,* 2001). The *Russula* genus is considered important both economically and ecologically. Ecologically, as EcM fungi, they promote plant healthy growth by delivering water, minerals and phosphates. In return, they obtain carbohydrates from their host plants. This symbiosis likely plays a critical role for maintaining biodiversity and carbon cycle in forest ecosystems (Li and Xu, 2009). In typical ecosystems, *Russula* constitutes a significant amount of the EcM biomass, and its distribution ranges from the tropics to subtropics, temperate regions, and even arctic zones. It has been showed that *Russula* is one of the most dominant EcM

fungi in the forest EcM communities in terms of both frequency and abundance (Horton and Bruns, 2001). In terms of edibility, only few *Russula* species, such as *R. subnigricans* Hongo, are deadly poisonous. Many are excellent food sources for human consumption, including *R. cyanoxantha* (Schaeff.) Fr., *R. virescens* (Schaeff.) Fr., and "dahongjun" in this study (Yang and Piepenbring, 2004).

The genus *Russula* is characterized by several easily identifiable features: the brittle consistency of the fleshy basidiocarp due to the presence of sphaerocysts, the lack of latex in the basidiocarp, and basidiospores with amyloid incrustations (Romagnesi, 1967; Singer, 1986) However, distinguishing individual species within the genus *Russula* is often very difficult due to the large number of closely related species and the lack of macro-morphological features to separate them. Similar problems also exist for many other groups of fungi. Consequently, there is an increasing use of molecular information to help define and identify new species or separate groups of morphologically indistinguishable species. At present, most studies on *Russula* infrageneric classification and population biology have been carried out using European and North American samples (Bergemann and Miller, 2002; Bergenann *et al.,* 2005a; Bergemann *et al.,* 2005b; Buyck *et al.,* 2006; Miller and Buyck, 2002; Richardson, 1970). The diversity and genetic structure of *Russula* species in other parts of the world remain poorly understood.

Due to its highly variable climate and diverse topography, southwestern China is ranked as one of the world's 34 biodiversity hotspots (Myers *et al.,* 2000). For example, about 600 out of 2,000 edible fungal species worldwide

occur in Yunnan Province in southwestern China (Yang, 2002). Some of the economically important mushrooms from this region include *Boletus edulis* sensu lato, *Thelephora ganbajun* M. Zang, *Tricholoma matsutake* (S. Ito & S. Imai) Singer, and *Russula* spp. (Lian *et al.,* 2008; Sha *et al.,* 2007; Sha *et al.,* 2008; Xu *et al.,* 2008; Wang *et al.,* 2009). Indeed, large populations of *Russula* are found throughout Yunnan Province. Many *Russula* species are harvested and sold at local markets or exported for human consumption. Among these *Russula* mushrooms, one called 'dahongjun' [the Chinese word that means the "big red mushroom" by indigenous people is probably the most prominent. This mushroom has been harvested and traded in local, national and international markets for over 20 years. Like many other pricey gourmet wild mushrooms such as matsutake, ganbajun and chanterelles, *Russula* cannot be artificially cultivated. Therefore natural populations in the forests are the only source for the market.

Like many wild mushrooms in southern China, accurate official statistics of the market are not available for dahongjun. However, our personal contacts with mushroom collectors over the last few years indicated that collecting this mushroom in June and July each year often provides up to half of a household's annual income for many people in southern and central Yunnan Province. During these months, fresh dahongjun may be sold for US$8.00 per kilogram and dried products are often sold for US$50-80/kg. However, the significant profits, uncontrolled harvesting practices in recent years, and disturbances of the forest ecosystem could threaten the wild populations of dahongjun in Yunnan Province. To help sustain its natural populations and design effective conservation and

utilization strategies, it is important to understand the populations of this species in their native distribution range.

Historically, mushroom enthusiasts and mycologists have regarded dahongjun in southern China as *R. vinosa* Lindblad, which was originally described in Europe (Romagnesi, 1967; Miller and Buyck, 2002; Legon *et al.*, 2005). The reliability of this identification for Chinese *R. vinosa* is dubious as no formal taxonomic investigation has been conducted and there is little molecular information for Chinese *Russula* in any public databases. Recently, a population of dahongjun from southern Yunnan was found to have DNA sequences significantly different from those of the typical *R. vinosa* from Europe (Miller and Buyck, 2002; Tedersoo *et al.*, 2003), and this group was recently described as a new species *Russula griseocarnosa* (Wang *et al.*, 2009).

The objective of the current study is to investigate the genetic diversity and relationships among strains and populations of dahongjun from their broad distribution ranges and to investigate their phylogenetic relationships with other known *Russula* species. Through analyzing ITS sequences for individual strains and populations, as well as the patterns of distribution of specific ITS genotypes among local and regional populations, we hope to gain an understanding of the genetic structure of dahongjun at both the local and regional level.

Materials and Methods

Sampling

Fruiting body samples of dahongjun were collected from tropical and subtropical evergreen forests dominated by *Castanopsis* spp. and *Lithocarpus* spp. in Yunnan and Guangxi Provinces during the summer of 2006 and 2007. Because of their commercial value and not-cultivable nature, there has been significant competition among local residents within each area for the wild mushroom. Most local mushroom pickers guard their own secret mushroom fruiting spots in the forests and they are unwilling to share the information with others. In addition, the most desirable dahongjun fruiting bodies are the immature ones (those with their caps closed and veils intact), picked while still buried underground and not visible by the naked eyes. As a result, it was extremely difficult to find the fruiting bodies not familiar with the local terrains and fruiting spots. The mushrooms analyzed here were collected with the help of one picker at each location who, through various means, were assured the confidentiality of their fruiting spots. All the fruiting bodies collected during two-day forays at each location were included. These specimens were identified based on their morphological features as determined by the local mushroom hunters and tradesmen. A total of 122 fruiting bodies were collected from 4 areas in Yunnan Province and one area in Guangxi Province. These regions stretched about 1500 km from east to west and about 500 km from north to south. The geographical locations of the sampling sites are shown in Figure 2.1. The geographic coordinates and the sample size from each site are presented in Table 2.1.

DNA extraction and ITS sequencing

Genomic DNA was extracted from dried fruiting bodies, using a CTAB miniprep method (Xu *et al.,* 2000). Detailed experimental procedures were shown in supplementary methods. The extracted DNA was suspended in 50µl TE buffer, and then diluted 1:20 for use in PCR reactions. The primers ITS1 (5'-TCCGTAGGTGAACCTGCGG-3')and ITS4 (5'-TCCTCCGCTTATTGATAGC-3') were used to amplify the ITS regions of the ribosomal RNA gene cluster that included ITS1, 5.8S, and ITS2 (White *et al.,* 1990). Each PCR reaction contained 10µl reaction volume: 5µl GoTaq® Green Master Mix 2× (Promega), 1.5µl diluted genomic DNA, 1µM of each primer, and 1.5µl ddH2O. The PCR was conducted with the following program: 4 min at 94°C, followed by 35 cycles of 30s at 94°C, 30s at 50°C, and 90s at 72°C, and a final extension step at 72°C for 10 min. PCR products were then cleaned and sequenced using cycle sequencing with BigDye®Terminator v3.1 (Applied Biosystems) following the manufacturer's protocol (shown in supplementary methods). Sequencing was done in both forward and reverse directions for each sample using the ITS1 and ITS4 primers.

Genotype identification

DNA sequence data obtained from the forward and reverse directions were assembled for each strain using the SeqMan sequence analysis software (DNASTAR Inc.). All the heterozygous sites were coded using the following universal ambiguity codes: T/C=Y, A/G= R, A/C= M, G/T= K. Sequences were then aligned using Clustal_X version 2 (Thompson *et al.,* 1997), and the

alignment was manually adjusted as needed. The aligned sequences were then input into PAUP*4.0b10 (Swofford, 2002) to identify unique ITS genotypes. All aligned nucleotide sites, including the insertions/deletions, were used to examine ITS sequence identities. Strains with the same ITS sequences were identified as belonging to the same ITS genotype. The identified genotypes were saved and used for subsequent phylogenetic and population genetic analyses.

Phylogenetic analysis

After obtaining ITS sequences, the unique ITS genotypes were then used as queries to retrieve closely related sequences with comparable lengths from GenBank. Multiple representative ITS sequences for each of the closely related *Russula* species from GenBank were included to compare potential intra-specific variation of the dahongjun phylogenetic groups to known *Russula* species. Sequences of *Russula griseocarnosa*, which was recently identified as a new species representing dahongjun from southern Yunnan (Wang *et al.*, 2009) were also included as reference sequences in our phylogenetic analysis. Our own sequences and all the retrieved sequences were then aligned using Clustal_X version 2 (Thompson *et al.*, 1997). The aligned sequences were visually inspected, adjusted and imported into PAUP*4.0b10 (Swofford, 2002). Maximum parsimony analysis was carried out using the heuristic search option with 500 random sequence additions. Topological support was estimated by bootstrap resampling of 1000 randomized datasets.

Analysis of population structure

All ITS genotype sequences were imported into Arlequin version 1.1 (Schneider et al., 2000). The ITS genotype diversity was estimated for each local and regional population. Wright's measure of genetic differentiation F_{ST} (Wright, 1943) was used to determine the contributions of geographic separation to the total genetic variation among local populations. F_{ST} was calculated using genetic distances. An analysis of molecular variance (AMOVA) (Excoffier *et al.,* 1992) was used to measure the variation in gene sequences among populations at three different levels: among regions, among local populations within regions, and within local populations. Here, we arbitrarily define a local population to include all individuals collected within 1km of each other. Regional populations were also arbitrarily defined to include local populations separated by less than 100km of each other. In addition, the genetic distances were calculated both between pairs of divergent lineages and within each lineage as identified based on phylogenetic analysis (see below).

The mean ITS sequence diversity within each strain and between pairs of strains were calculated to examine potential evidence for mating and recombination. Briefly, if mating and recombination were rare or and the dominant modes of reproduction were primarily asexual or inbreeding, we should expect the sequence diversity (i.e. the percentage of polymorphic nucleotide sites) within individual strains to be lower than those between strains. In contrast, if the mean sequence diversity within strains were comparable to that between strains, such a result would be consistent with genetic exchange between strains, likely

due to outcrossing and sexual recombination. For this analysis, because sample sizes for four of the five local populations were relatively small, only the Ailaoshan (AL) population was examined. Student's t-test was used to determine the statistical significance in the differences of sequence diversity for within and between strains within the AL population.

Genomic DNA Library Construction (within AL dahongjun population)

We constructed a shortgun genomic library using the PCR-Script™ Amp Cloning Kit. Firstly, genomic DNA was partialy digested by restriction enzyme HaeIII to generate a length range of 500bp -1.5kb. Digested genomic DNA was then run on 1% agarose gel and the desired fragments of lengths between 500bp-1.5kb were extracted using the Qiagen MinElute Gel Extraction Kit. 2μl of the gel-purified DNA (80ng/μl) was used for ligation reaction. The ligation and transformation precedures followed the protocols provided by PCR-Script™ Amp Cloning Kit. Positive clones were identified based on the blue-white color screening, then directly used for PCR amplication. M_{13} primer pairs were used to amplify the cloned fragments for sequencing. Species-specific primers were designed by SGD Web Primer (http://seq.yeastgenome.org/cgi-bin/web-primer) to amplify the newly obtained DNA sequences, which serves as DNA markers to identify nucleotide polymorphisms between strains.

Haplotype identification (within AL dahongjun population)

All variable sites were selected and compiled into a required format for input into PHASE 2.1.1 (Stephens *et al.* 2001). PHASE 2.1.1 uses Bayesian analysis to infer possible haplotypes and estimate the frequencies for each haplotype within the samples. Mushroom fruiting bodies are diploid, and the possible haploid genotype could be inferred by performing PHASE analysis. Due to the highly diverged ITS sequences in our total samples (see below), all the sequences were first divided into different groups (three groups or lineages, see below) based on their clustering using the composite ITS sequences for each strain. Strains in individual clusters were then separately imported into PHASE 2.1.1 to deduce haplotype compositions for individual strains. The separated haplotypes were saved and used for subsequent intra-specific population studies.

RESULTS

ITS sequence variation

A total of 122 dahongjun specimens were successfully amplified using the ITS1 and ITS4 primer pair, and the resulting PCR products were sequenced in both directions. The GenBank accession numbers for these sequences are FJ613896-FJ614017. Among the 683 aligned nucleotides for the 122 specimens, 82 were variable due to either insertion/deletion (indels) or single base substitutions. Our preliminary analyses of these genotype sequences identified three phylogenetically distinct groups. Among the 82 variable nucleotide sites, 46 were found among the groups and 39 were found within individual groups. The first group contained 85 strains from the Ailaoshan (AL) region in central Yunnan and 5 strains from Cangwu (CW) in eastern Guangxi. In total, 19 variable

nucleotide sites were found within this group. Eight of them were in the ITS1 region, 10 in the ITS2 region, and 1 in the 5.8S region. These variable sites contained 16 transitional substitutions, 2 transversional substitutions, and 1 single-base indels. Among these variable sites, 15 were phylogenetically informative and 4 were phylogenetically uninformative.

The second group identified in our preliminary phylogenetic analyses contained a total of 27 isolates, including 9 from Mengla (ML), 8 from Jinuo (JN), 8 from Dadugang (DDG), and 2 from Cangwu (CW). The alignment of this group of sequences identified 12 variable sites, including 6 transitional substitutions, 2 transversional substitutions and 4 single nucleotide indels. Noticeably, a poly-T section located in the ITS2 region contained 6 variable sites, one of which contained both a transitional substitution and an indel. Among the 14 variable sites within this group of strains, 8 were phylogenetically informative and 6 were uninformative.

The third group of sequences contained 3 strains from DDG and 2 strains from CW. Five variable sites were observed among these five strains, including 3 transitions, 1 transversion and 1 single-base indels. All of these variable sites were phylogenetically informative.

The sequence comparisons showed two notable stretches of sequence differences within the ITS2 region among the three phylogenetic groups. Specifically, the first stretch includes a 3bp-5bp insertion/deletion in a poly-T region, with the insertions present among strains only in the second phylogenetic group. The second stretch of sequence is an 18bp insertion/deletion, with the

deletion present among strains only in the third phylogenetic group. Overall, the difference among the 63 ITS sequence types ranged between 1 and 90 nucleotide sites (out of 683 aligned nucleotides). The distribution of ITS genotypes among the 5 local populations is presented in Table 2.1.

Distribution of genotypes

The raw ITS sequences of the three different phylogenetic groups were imported into PAUP*4.0b10 (Swofford, 2002) for identification of unique ITS genotypes. Our analyses identified a total of 63 unique genotypes. The detailed information of polymorphic sites among genotypes and the exact number of polymorphic sites within ITS regions of each Dahongjun genotype were shown in Tables 2.2 and supplementary Table 4A.1. In the first phylogenetic group, a total of 37 unique genotypes were identified from among the 85 AL specimens and 5 unique genotypes among the 5 CW specimens respectively. The most common ITS genotype in the AL population was shared by 16 specimens while 25 genotypes were present in only one specimen each. In the second phylogenetic group, 17 unique ITS genotypes were identified among 27 strains from ML, JN, DDG, and CW. The most common genotype in this group was shared by 3 strains, while 10 genotypes were found in only one strain each. Among the 5 specimens in the third phylogenetic group, we identified 3 ITS genotypes from 3 specimens from DDG, and 1 genotype from 2 specimens from CW.

Among the 122 mushroom specimens analyzed in this study, 82 contained heterozygous nucleotide sites and 40 showed no heterozygous sites within the

sequenced ITS regions. When looking at each local population individually, 25 (out of 85) in the AL population, 5 (out of 9) in ML, 2 (out of 8) in JN, 5 (out of 8) in DDG, and 3 (out of 9) in CW were completely homozygous. The specific nucleotide sites that are variable within the total sample are presented in Supplementary Table 4A.2.

Population genetic analysis of the ITS genotypes

ITS genotype sequences for all 122 strains from the 5 local populations were analyzed by the population genetics software Arlequin version 1.1 (Schneider *et al.,* 2000). Population differentiation between pairs of samples was determined by calculating Wright's F_{ST} (Wright, 1943). The results of pairwise F_{ST} showed significant differentiation between populations from different regions, and moderate differentiation between local populations within the same region in southern Yunnan (Table 2.4). All the F_{ST} values were statistically significant ($P<0.05$), except the one between JN and DDG.

Among the three phylogenetic lineages, we calculated the sequence diversity within strains and between strains using the software Arlequin version 1.1 for the AL lineage population (Schneider *et al.,* 2000). The mean nucleotide difference between pairs of strains was 2.267 ± 1.032 (out of 683 nucleotides; N=13944 pairwise comparisons), and the mean number of heterozygous nucleotides within strains was 1.742 ± 0.340 (N=85). The t-test result supported that the mean polymorphic nucleotide sites within strains were significantly lower than those from between strains ($P<0.001$). This analysis suggested that while

mating and recombination might be occurring, there was little evidence for random mating and frequent outcrossing within the AL population. Similarly, we found lower levels of sequence diversity within strains than between strains for the other two phylogenetic groups and four local populations (detailed data not shown).

Phylogenetic analysis

After BLAST searches against GenBank, we retrieved sequences that were closely related to our dahongjun sequences with an overall sequence identity ≥90%. We would like to note that only sequences from GenBank with identification to the species level were selected for comparison. GenBank sequences from direct environmental sampling and with no attached species names were not included in the phylogenetic analysis or comparison. To compare the relative divergence of our sequences among each other with those in other closely related species, we also included multiple ITS sequences from within several other *Russula* species whenever possible. In total, 16 reference sequences were retrieved from GenBank for comparisons with our sequences. These included three ITS sequences of *R. griseocarnosa* recently identified based on three specimens of dahongjun in southern Yunnan (Wang *et al.*, 2009) and 13 ITS sequences from GenBank representing 5 closely related species in the genus *Russula* (*R. vinosa, R. clavoflova, R. occidentalis, R. decolorans, R. xerampelina*; Supplementary Figure 4A.1).

Our analyses of the above sequences confirmed the distinctiveness of the three phylogenetic groups within dahongjun from southern China. The numbers of total segregating sites, fixed and shared polymorphic sites within and between individual lineages were shown in Table 2.3. The separation of these three lineages was supported by 98%, 84% and 100% bootstrap. The distinctness of the three lineages was also supported based on the analyses of genetic distances, with the pairwise genetic distances among strains from different lineages much greater than those from within the same lineages (data not shown). The separation of dahongjun from the one of the most closely related species *R. vinosa* had 100% bootstrap support as well, including large phylogenetic distances between them (Figure 2.2 and Supplementary Figure 4A.1).

The phylogenetic analyses identified that lineage 1 contained 41 genotypes shared by 90 strains, lineage 2 contained 17 genotypes shared by 28 strains, and lineage 3 contained 4 genotypes shared by 5 strains. The strains in lineage 1 were from AL and CW. The strains in lineage 2 were from three local populations in southern Yunnan (ML, JN, and DDG) and one population from CW in eastern Guangxi. Lineage 3 comprised strains from DDG and CW. In addition, we found distinct phylogenetic subgroups within both lineages 1 and 2 (Figure 2.2). Specifically, in lineage 1, samples from AL and CW were partitioned into two separate subgroups. In lineage 2, sixteen genotypes from ML, JN, DDG, CW and one genotype of *R. griseocarnosa* (from DDG) were clustered together as subgroup 1. Subgroup 2 within lineage 2 included one genotype (#54) representing one strain from ML as well as two genotypes of *R. griseocarnosa*

from Xishuanbana in southern Yunnan (Wang *et al.,* 2009). This result showed that the recently identified *R. griseocarnosa* belonged to one lineage (i.e. lineage 2) of the southern China dahongjun. Samples from CW and DDG were found distributed across several lineages, indicating that multiple distinct lineages could coexist in one area. However, there was no sharing of ITS genotypes among strains from different regional populations.

Population genetic analyses

Three AMOVA tests were conducted for different groups of samples. In the first, geographic populations including all strains were compared irrespective of their phylogenetic lineage affiliations. This analysis identified that geographic populations were significantly different from each other (Table 2.5). However, despite the apparent contribution of geography to the distribution of ITS sequence types, Mantel test identified no significant correlation between genetic differentiation and geographical distance among the analyzed population samples ($p>0.10$; data not shown). In the second AMOVA analysis, we examined the contributions of sequence variation due to lineage divergence. Our analyses identified 60.18%, 30.25%, and 9.57% of the overall ITS sequence variation came from comparisons among the three lineages, among sublineages within the three lineages, and within sub-lineages respectively. The contributions from each of the three sources were all statistically significant ($p<0.05$ for F_{ST}, F_{SC} and F_{CT} in the AMOVA tests, Table 2.6). In the third AMOVA analysis, we analyzed populations within lineages and found that the geographic differentiation among

local populations within a lineage was still evident (Table 2.7). Due to the small and/or highly skewed sample sizes from different geographic areas for lineages 1 and 3, the AMOVA test was conducted for only lineage 2. Specifically, among the three local populations within lineage 2 from southern Yunnan, two of the three pairwise comparisons were significantly different, consistent with geographic differentiation for populations over distances ~100km apart (Table 2.7).

Genet size estimation

In order to identify the size for Dahongjun genets, we looked at four molecular markers, ITS and the other three (C205, C27, C210), which were identified from cloning.

Detailed haplotype information for these four markers is shown in Table 2.7. PHASE 2.1.1 (Stephens *et al.* 2001) was used to separate diploid sequences for each marker to haplotypes. Based on haplotype information, and spatial distribution (from collecting records) of dahonjun, the size of the genets could be estimated. Our results showed that the sizes of Dahongjun genets were extremely small ranging from 30cm to 3m.

Discussion

In this study, we used the ITS sequence information to analyze the diversity and population structure within and among local and regional samples of

dahongjun in southern China. Our analysis identified a total 63 ITS sequence types among 122 strains collected from 5 areas that spanned 2 provinces.

Phylogenetic analysis

Our phylogenetic analysis based on ITS genotypes identified 3 distinct phylogenetic lineages among the commercially collected wild dahongjun samples (Figure 2.2). The distinctiveness of these lineages was strongly supported by both phylogenetic and population genetic analyses. Indeed, the sequence divergence between lineages was significantly greater than those between strains within any of the closely related known *Russula* species included in our phylogenetic analyses (Figure 2.2 and supplementary Figure 4A.1). The significant divergence among these lineages suggests that each lineage may represent a distinct species, supporting the hypothesis that dahongjun in southern China is likely a species complex, containing potentially multiple cryptic species. Such a pattern is consistent with recent population genetic studies for many fungal species that identified multiple phylogenetically distinct species (Taylor *et al.,* 2000; Pujol *et al.,* 2005; Hedh *et al.,* 2008). For example, the endemic gourmet mushroom from southwestern China *Thelephora ganbajun*, the mycorrhizal fungus *Cenococcum geophilum* and the wood decay Boletales *Serpula himantioides* were all identified as complexes of several cryptic species using DNA sequence markers (Sha *et al.,* 2008; Douhan and Rizzo, 2005; Kauserud *et al.,* 2006). Multiple gene genealogical analyses would be needed to determine whether the distinct lineages identified here are in fact reproductively isolated.

Our analyses also identified that none of our samples had ITS sequences identical to *Russula vinosa*, a species previously thought to encompass dahongjun (Figure 2.2). The distinctiveness of dahongjun from other known *Russula* species had 100% bootstrap support and large phylogenetic distances (Figure 2.1). These results thus indicate that none of the three dahongjun lineages belonged to *Russula vinosa* or likely any other *Russula* species described from Europe and North America. In addition, the recently proposed *R. griseocarnosa* represented only one lineage (lineage 2 in Figure 2.2) within the broadly distributed dahongjun. Detailed morphological, ecological, and further molecular characterizations of lineages 1 and 3 are in progress.

Importantly, both lineages 1 and 2 likely contain additionally reproductively isolated sub-lineages. The two subgroups within lineage 2 were both from southern Yunnan. In contrast, the two subgroups in lineage 1 were geographically separated. The mechanisms for the different patterns are unknown. However, if geographic separation was the dominant factor driving sequence divergence, given their close geographic distances, these two subgroups within lineage 2 likely evolved from genetically different founders. Alternatively, limited sampling and/or incomplete lineage sorting can also result in the observed geographic patterns of sub-lineage distributions.

Our analyses suggest that the phylogenetic groups identified here seemed to have limited geographical ranges in southern China. For example, within Yunnan Province, lineage 2 was found only in the south; lineage 1 subgroup1 was restricted to central Yunnan; lineage 1 subgroup2 was restricted to eastern

Guangxi region; and lineage 3 was found in southern Yunnan as well as in eastern Guangxi. Limited geographic distribution has also been observed in *Russula brevipes*, where distinct alleles were observed in geographically separated populations (Bergemann and Miller, 2002). Such a geographic distribution was similarly observed in another gourmet mushroom from central and southern Yunnan, *Thelephora ganbajun* (Sha *et al.,* 2008). In *T. ganbajun*, many genotypes were unique to local populations and the distribution for each cryptic species was geographically restricted (Sha *et al.,* 2008). Non-overlapping of geographical habitats for distinct lineages and sub-lineages are suggestive of allopatric speciation, considered by many to be the most common mode of speciation in cellular organisms, including fungi (Lamb, 1979; Allen and Fhipps, 1984). However, we would like to note that even though the distribution of different dahongjun lineages identified here did not overlap geographically, it is possible that they might be overlapping if more extensive sampling is conducted across greater geographic areas.

Despite the restricted geographical distribution of different lineages, multiple distinct lineages were found coexisting in one geographical area. Specifically, DDG contains strains in both lineages 2 and 3 and CW contains strains in lineages 1, 2 and 3. This observation suggests that strains from different lineages could encounter each other in nature. However, our analysis showed clear distinctions among ITS sequences from different lineages, suggesting limited or no gene flow among these lineages. The existence of inter-sterility barriers and cryptic speciation is a common phenomenon in fungal species

complexes (Kauserud *et al.,* 2006). More molecular markers are needed to reveal the extent of reproductive isolation and hybridization among the three lineages. At present, the mechanism for the potential reproductive isolation among these lineages remains unknown. Factors such as preference in habitat, climate, and host tree species could all contribute to their differentiation.

We found relatively little evidence of ITS sequence type sharing and gene flow between geographic populations of dahongjun (Table 2.1). Our result is consistent with several previous studies of ectomycorrhizal mushrooms that suggested common short distance spore dispersals and relatively rare long distance spore dispersals in this group of fungi (Lamb, 1979; Liang *et al.,* 2004). However, long distance dispersal has also been found for basidiomycete fungi, by wind, animals, and often by humans (Xu, 2005). For example, empirical observations have shown that basidiospores can travel distances of up to by air 500 m in *Tricholoma matsutake* (Lian *et al.,* 2006). Similarly, spore dispersals up to 2090 m by squirrels were reported in several hypogeous and epigeous fungi (Bertolino *et al.,* 2004). If such events are frequent, significant gene flow among local and regional populations should be found. Indeed, gene flows over hundreds of kilometers has been inferred for several mushroom species from Yunnan, including the gourmet mushrooms *T. matsutake* and *T. ganbajun* (Sha *et al.,* 2007; Xu *et al.,* 2007).

Extensive genetic variation within local populations

Large numbers of unique genotypes and high genotype diversities were

found in most of the local populations (Table 2.1). By examining the genotypes for individual strains in the AL population, we observed several recombinant genotypes, suggesting evidence of sexual recombination. However, Hardy-Weinberg disequilibrium in local populations was inconsistent with random mating. For example, the AL local population showed an excess of homozygotes and greater sequence similarity between genotypes from within strains than that between strains. Both results are indicative of inbreeding. Inbreeding as a common reproductive event was observed in several other mushroom species, such as *Russula brevipes, Laccaria amethystina, Hebeloma cylindrosporum* Romagn. where positive inbreeding coefficients were detected, supporting non-random mating within local populations (Bergemann *et al.*2002; Gryta *et al.*, 2000; Wadud *et al.* 2006). Excess homozygosity and high sequence similarity within strains can be explained by high rate of somatic recombination and gene conversion that could homogenize the alleles within individual cells or allele drop-out during PCR, which result in false-positive estimates of homozygosity, contributing to skewed results that favor the hypothesis for inbreeding.

Estimation of genet size for AL Dahongjun population

A genet is a group of individuals (fruiting bodies and/or underground mycelia for EcM fungi) produced from one mating event and that occupy the same geographic area (Dahlberg & Stenlid, 1994; Xu, 2005). Bearing a small number of mutations, the individuals within a genet should be all genetically identical. Because they arise vegetatively from a single mating event, this shared

descent among the individual fruiting bodies and mycelia implies connectivity both nutritionally and/or genetically. Previous studies on *Russula* genets showed that they tended to form small genets (*R. cremoricolor,* <1.5m; *R. vinosa,* < 1m) (Liang *et al.* 2004, Redecker *et al.* 2001). Small genets suggest that *Russula* predominately reproduce sexually. Our research found that dahongjun also tends to form small genets. This finding is consistent with other *Russula* species.

In conclusion, our study suggests that the commercially collected dahongjun is a species complex, composed of at least three distinct lineages. Reproductive isolation seems to have occurred among these lineages. More extensive sampling may reveal additional distinct lineages as well as novel distribution patterns within this complex in southern China. While outcrossing might occur, our analyses suggest little evidence for random mating within local populations. Instead, we hypothesize that inbreeding might be a common mode of reproduction in nature. Overall, long-distance gene flow seemed limited, mostly only between neighboring local populations and very rarely among regional populations. The limited gene flow and the existence of distinct alleles within most sampled areas suggest that care must be taken to ensure that these lineages will not be permanently lost due to overexploitation and mis-management. Our results suggest that individual populations should be separately managed and conserved to ensure the survival of individual local populations for this ecologically and commercially important *Russula* species complex.

Acknowledgements

This study was supported by Natural Science and Engineering Research Council of Canada, Premier's Research Excellence Award, the Joint Funds of the National Natural Science Foundation of China and Yunnan Provincial Government (No. U0836604), the National Basic Research Program of China (No. 2009CB522300), and the Key Laboratory of Biodiversity and Biogeography of the Chinese Academy of Sciences (No. 0806371111).

Table 2.1 Geographic distribution and genetic diversity of dahongjun samples collected from Yunnan and Guangxi Provinces in Southern China. ITS genotypes shared between local populations are highlighted in bold.

Region/ District	County/ Community	Lat. (North)	Long. (East)	Sample size	ITS genotype (# isolates in each genotype)	Genotype diversity
Central Yunnan	Ailaoshan (AL)	24.32	101.01	85	1(2); 2(4); 3(1) 4(1); 5(1); 6(1) 7(2); 8(1); 9(1) 10(1); 11(2); 12(1) 13(3); 14(1); 15(9) 16(1); 17(1); 18(1) 19(1); 20(16); 21(2) 22(1); 23(1); 24(1); 25(5); 26(1); 27 (4); 28(5); 29(1); 30 (2); 31(1); 32(2); 33(1); 34(4); 35(1); 36(1); 37(1)	0.945
Southern Yunnan	Mengla (ML)	21.28	101.35	9	**38(1)**; **39(1)**; 41(1); 42(1); **43(1)**; 44(1); 45(1); 46(1); 54(1);	1.000
	Jinuo (JN)	22.01	101.03	8	**39(1)**; **43(1)**; **49(1)**; **50(1)**; 47(2); 48(1); 51(1)	0.964
	Dadugang (DDG)	22.20	100.58	11	**38(2)**; 40(2); **49(1)**; **50(2)**; 51(1); 55 (1); 56(1); 57(1);	0.945
Eastern Guangxi	Cangwu (CW)	23.29	111.18	9	58(2); 53(1); 52(1); 59(1); 60(1); 61(1); 62(1); 63(1)	0.972

Table 2.2 Within-strain sequence polymorphism at the ITS region for samples of dahongjun analysed in this study. AL: Ailaoshan; DG: Dadugang; ML: Mengla; JN: Jinuo; CW: Cangwu.

Isolate	Group	No. of heterozygous sites within ITS regions	Geographic Origin
DahongjunGeno01	Lineage 1	1	AL
DahongjunGeno02	Lineage 1	2	AL
DahongjunGeno03	Lineage 1	3	AL
DahongjunGeno04	Lineage 1	3	AL
DahongjunGeno05	Lineage 1	5	AL
DahongjunGeno06	Lineage 1	4	AL
DahongjunGeno07	Lineage 1	2	AL
DahongjunGeno08	Lineage 1	3	AL
DahongjunGeno09	Lineage 1	5	AL
DahongjunGeno10	Lineage 1	4	AL
DahongjunGeno11	Lineage 1	6	AL
DahongjunGeno12	Lineage 1	4	AL
DahongjunGeno13	Lineage 1	4	AL
DahongjunGeno14	Lineage 1	3	AL
DahongjunGeno15	Lineage 1	2	AL
DahongjunGeno16	Lineage 1	3	AL
DahongjunGeno17	Lineage 1	2	AL
DahongjunGeno19	Lineage 1	1	AL
DahongjunGeno21	Lineage 1	3	AL
DahongjunGeno22	Lineage 1	1	AL
DahongjunGeno23	Lineage 1	2	AL
DahongjunGeno24	Lineage 1	2	AL
DahongjunGeno27	Lineage 1	1	AL
DahongjunGeno28	Lineage 1	1	AL
DahongjunGeno29	Lineage 1	1	AL
DahongjunGeno32	Lineage 1	1	AL
DahongjunGeno34	Lineage 1	1	AL
DahongjunGeno35	Lineage 1	1	AL
DahongjunGeno36	Lineage 1	1	AL
DahongjunGeno61	Lineage 1	1	CW
DahongjunGeno62	Lineage 1	4	CW
DahongjunGeno63	Lineage 1	5	CW
DahongjunGeno39	Lineage 2	2	ML, JN
DahongjunGeno40	Lineage 2	3	DDG
DahongjunGeno41	Lineage 2	1	ML
DahongjunGeno42	Lineage 2	1	ML
DahongjunGeno47	Lineage 2	1	JN
DahongjunGeno48	Lineage 2	1	JN
DahongjunGeno49	Lineage 2	1	JN, DDG
DahongjunGeno50	Lineage 2	1	JN, DDG
DahongjunGeno52	Lineage 2	2	CW
DahongjunGeno55	Lineage 3	1	DDG
DahongjunGeno56	Lineage 3	1	DDG

Table 2.3 DNA sequence divergence between and within the phylogenetic lineages 1, 2, and 3 (as indicated in Figure 2.2). The total number of segregating sites, fixed and shared polymorphic sites within and between lineages and sublineages are shown here. The number of sites that are polymorphic found in the first group (second group) but monomorphic in the second group (first group) was also presented.

	No. of segregating sites	Fixed differences	Polym. 1 Monom. 2	Polym. 1 Monom. 2	Shared differences
Lineage 1 (sub1) – Lineage 1 (sub 2)	29	9	12	5	3
Lineage 2 (sub1) – Lineage 2 (sub 2)	15	7	7	1	0
Lineage 1- Lineage 2	63	21	25	15	2
Lineage 1- Lineage 3	77	47	21	8	1
Lineage 2- Lineage 3	54	37	11	4	2

Table 2.4 Pairwise F_{ST} values between geographic populations of dahongjun in southern China.

Geographic region	Ailaoshan (AL)	Mengla (ML)	Jinuo (JN)	Dadugang (DDG)	Cangwu (CW)
Ailaoshan (AL)	0.000				
Mengla (ML)	0.941***				
Jinuo (JN)	0.931***	0.097*			
Dadugang (DDG)	0.877***	0.147**	0.106		
Cangwu (CW)	0.711***	0.499***	0.468***	0.298***	0.000

*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$

Table 2.5 Analysis of molecular variance (AMOVA) of dahongjun samples from different local and regional populations in southern China.

Source of variation	d.f.*	Sum of squares	Variance components	Percent variation	Fixation index	P
Among regions	2	1342.399	9.69856	60.18%	F_{ST}: 0.63045	<0.0001
Among local populations within regions	2	169.196	4.06009	25.16%	F_{SC}: 0.85253	<0.0001
Within local populations	232	552.144	2.37993	14.75%	F_{CT}: 0.60095	<0.05
Total	237	2063.739	16.13858	100%		

Table 2.6 Analysis of molecular variance (AMOVA) of dahongjun samples from different lineages and sub-lineages in southern China.

Source of variation	d.f.*	Sum of squares	Variance components	Percent variation	Fixation index	P
Among lineages	2	1472.796	11.31023	60.18%	F_{ST}: 0.75986	<0.0001
Among sub-lineages within lineages	2	172.294	5.68545	30.25%	F_{SC}: 0.90439	<0.01
Within sub-lineages	233	418.650	1.79678	9.57%	F_{CT}: 0.60185	<0.05
Total	237	2063.739	18.79245	100		

*d.f., degree of freedom

Table 2.7 Analysis of molecular variance (AMOVA) of dahongjun samples from lineage 2. These samples were from two regions in southern China, one in Guangxi province (containing one local population CW) and one in southern Yunnan (containing three local populations ML, JN, and DDG).

Source of variation	d.f.*	Sum of squares	Variance components	Percent variation	Fixation index	P
Among regions	1	6.950	0.64398	28.62%	F_{ST}: 0.09557	<0.0001
Among local populations within regions	2	7.550	0.15350	6.82%	F_{SC}: 0.35442	<0.01
Within local populations	46	66.819	1.45260	64.56%	F_{CT}: 0.28620	>0.05
Total	49	81.320	2.25007	100		

*d.f., degree of freedom

Table 2.8 Detailed information of each DNA marker used for intra-specific studies.

	ITS	C27	C201	C210
Length	670bp	569bp	536bp	370bp
Variable sites	12	11+2 indels	15	14
# of individuals	84	75	80	84
Haplotypes	24	13	42	49

Figure 2.1 Sites of sample collection in southern China. Yunnan and Guangxi provinces are highlighted in red. Dadugang (DDG), Jinuo(JN) and Mengla (ML) in southern Yunnan region are marked as black triangles; Ailaoshan (AL) in central Yunnan is marked as a blue triangle, and Cangwu (CW) in eastern Guangxi region is marked as a yellow triangle.

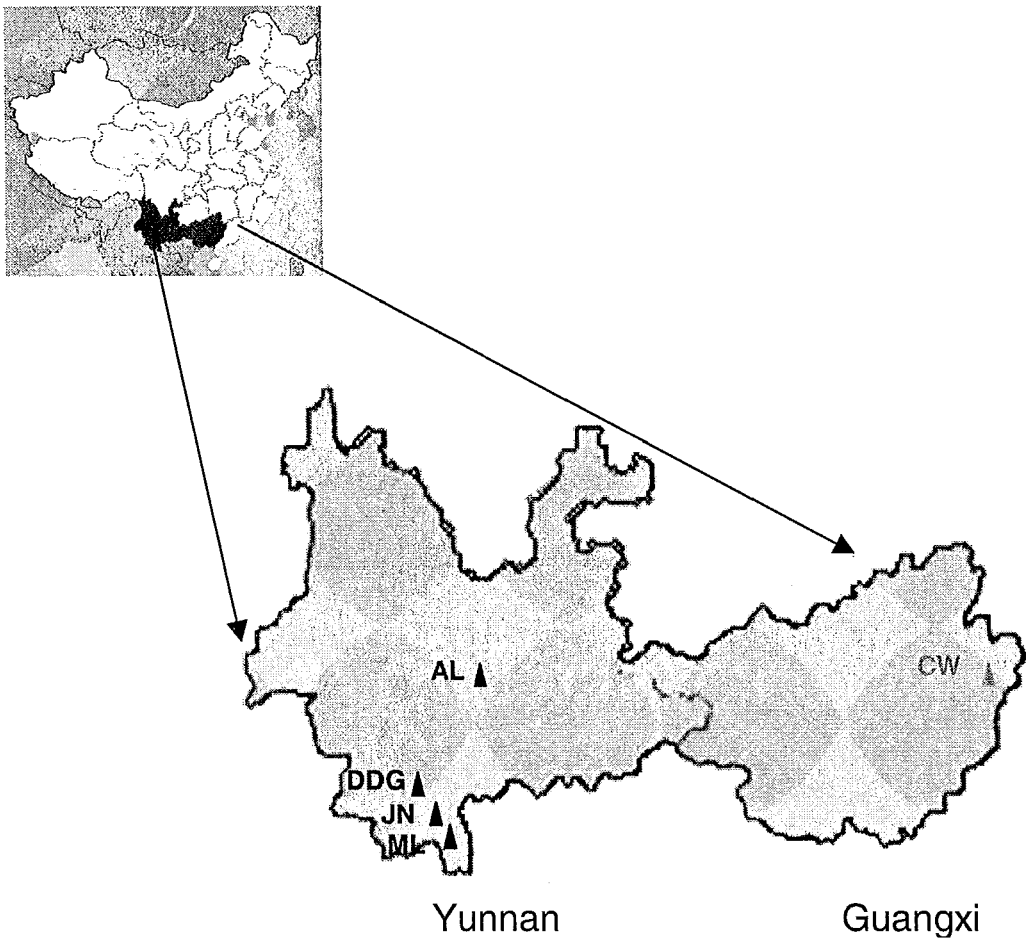
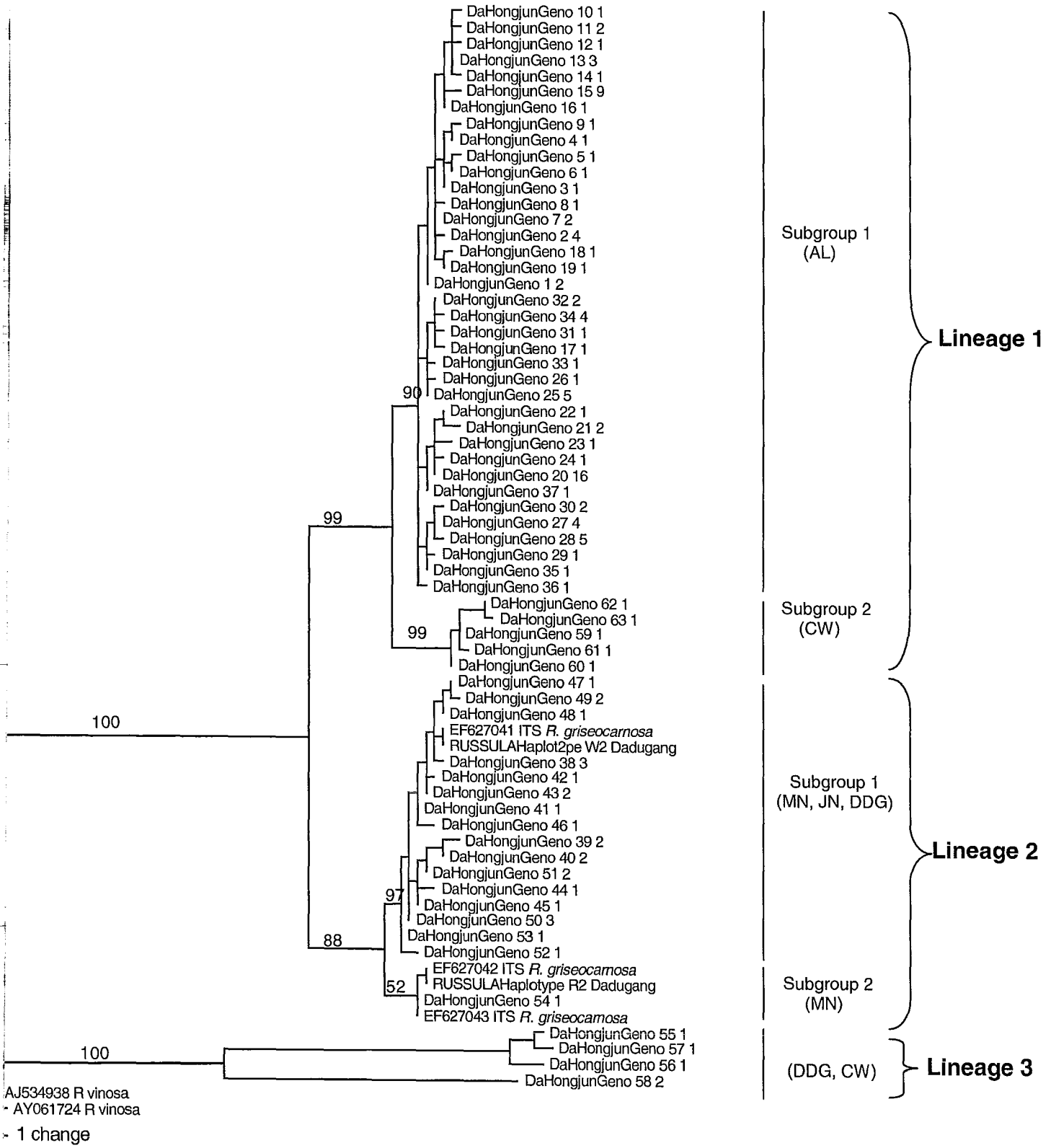

AL
CW
DDG
JN
MI
Yunnan
Guangxi

Figure 2.2 Maximum parsimony tree based on unique ITS-genotype sequences from 122 isolates of dahongjun. For each ITS genotype of dahongjun, the first number represents the genotype assignment corresponding to those in Table 2.1; the second number represents the total number of strains (out of 122) belonging to the specific genotype. Numbers along branches are bootstrap values greater than 90% obtained from 1000 replicates. The heterozygous sites, coded as T/C=Y, A/G= R, A/C= M, G/T= K were considered as new states. Sequences of *Russula vinosa* are used as outgroups. Tree length=183, Consistency index=0.716, Retention index=0.939.


DaHongjunGeno 10 1
DaHongjunGeno 11 2
DaHongjunGeno 12 1
DaHongjunGeno 13 3
DaHongjunGeno 14 1
DaHongjunGeno 15 9
DaHongjunGeno 16 1
DaHongjunGeno 9 1
DaHongjunGeno 4 1
DaHongjunGeno 5 1
DaHongjunGeno 6 1
DaHongjunGeno 3 1
DaHongjunGeno 8 1
DaHongjunGeno 7 2
DaHongjunGeno 2 4
DaHongjunGeno 18 1
DaHongjunGeno 19 1
DaHongjunGeno 1 2
DaHongjunGeno 32 2
DaHongjunGeno 34 4
DaHongjunGeno 31 1
DaHongjunGeno 17 1
DaHongjunGeno 33 1
DaHongjunGeno 26 1
90
DaHongjunGeno 25 5
DaHongjunGeno 22 1
DaHongjunGeno 21 2
DaHongjunGeno 23 1
DaHongjunGeno 24 1
DaHongjunGeno 20 16
DaHongjunGeno 37 1
99
DaHongjunGeno 30 2
DaHongjunGeno 27 4
DaHongjunGeno 28 5
DaHongjunGeno 29 1
DaHongjunGeno 35 1
DaHongjunGeno 36 1
DaHongjunGeno 62 1
DaHongjunGeno 63 1
99
DaHongjunGeno 59 1
DaHongjunGeno 61 1
DaHongjunGeno 60 1
100
DaHongjunGeno 47 1
DaHongjunGeno 49 2
DaHongjunGeno 48 1
EF627041 ITS R. griseocarnosa
RUSSULAHaplot2pe W2 Dadugang
DaHongjunGeno 38 3
DaHongjunGeno 42 1
DaHongjunGeno 43 2
DaHongjunGeno 41 1
DaHongjunGeno 46 1
DaHongjunGeno 39 2
DaHongjunGeno 40 2
DaHongjunGeno 51 2
97
DaHongjunGeno 44 1
DaHongjunGeno 45 1
DaHongjunGeno 50 3
88
DaHongjunGeno 53 1
DaHongjunGeno 52 1
EF627042 ITS R. griseocarnosa
52
RUSSULAHaplotype R2 Dadugang
DaHongjunGeno 54 1
EF627043 ITS R. griseocarnosa
DaHongjunGeno 55 1
100
DaHongjunGeno 57 1
DaHongjunGeno 56 1
DaHongjunGeno 58 2
AJ534938 R vinosa
AY061724 R vinosa
1 change
Subgroup 1 (AL)
Subgroup 2 (CW)
Lineage 1
Subgroup 1 (MN, JN, DDG)
Subgroup 2 (MN)
Lineage 2
(DDG, CW)
Lineage 3

CHAPTER 3

Concluding Remarks

Our phylogenetic analysis based on ITS genotypes identified 3 distinct phylogenetic lineages among the commercially collected wild dahongjun samples. The separation of these three lineages was supported by 98%, 84% and 100% bootstraps. Our analyses have confirmed the distinctiveness of the three phylogenetic groups within dahongjun. We also found a large phylogenetic distance between dahongjun and other *Russula* species found in U.S. and Europe. Our analyses also identified that none of our samples had ITS sequences identical to *Russula vinosa*, a species previously thought to encompass dahongjun and the separation of dahongjun from *R. vinosa* had 100% bootstrap support (Figure 2.2). All of these results clearly suggested that none of the three dahongjun lineages belonged to *Russula vinosa* or likely any other *Russula* species described from Europe and North America. In addition, the recently proposed *R. griseocarnosa* represented only one lineage (lineage 2 in Figure 2.2) within the broadly distributed dahongjun.

Dahongjun is a species complex, composed of at least three distinct lineages. Reproductive isolation seems to have occurred among these lineages, as can be seen by no genotypes sharing among lineages. Our research did not detect any hybridization happening cross lineages. However, more extensive sampling may reveal additional distinct lineages as well as their distribution patterns within this complex in southern China. Additionally, our analysis identified that the phylogenetic groups identified here seemed to have limited geographical ranges and geographic populations were significantly different from each other. We

found relatively little evidence of ITS sequence type sharing and gene flow among regional dahongjun populations, suggesting dahongjun populations are structured at regional level.

Multiple distinct lineages could coexist in one area. However, there was no sharing of ITS genotypes among strains from different regional populations. Despite the restricted geographical distribution of different lineages, multiple distinct lineages were found coexisting in one geographical area. However, our analysis showed clear distinctions among ITS sequences from different lineages, suggesting limited or no gene flow among these lineages. The existence of inter-sterility barriers and cryptic speciation is a common phenomenon in fungal species complexes (Kauserud *et al.,* 2006). More molecular markers are needed to reveal the extent of reproductive isolation and hybridization among the three lineages. At present, the mechanism for the potential reproductive isolation among these lineages remains unknown. Factors such as preference in habitat, climate, and host tree species could all contribute to their differentiation.

Very high genetic diversity was found in local dahongjun populations. At present, the mechanisms for the generation and maintenance of such high sequence diversity for the ITS regions remain unknown. Our sequence diversity analyses suggest that the polymorphisms within strains were lower than those between strains within each geographic population and each lineage, consistent with non-random mating and inbreeding in nature. Inbreeding as a common reproductive strategy has been observed in several other mushroom species.

Our analysis suggested that the size of the dahongjun genets were extremely

small. It is consistent with the sizes found in other *Russula* species. We found a large number of small genets ranging from 30cm-3m in diameter and many of them are single-sporocarp genets. It suggests dahongjun is good at growing sexually, rather than doing vegetative growth. Its fact reflects one of the typical features of late-stage EcM fungi. Sexual reproduction seems to be the dominant reproductive strategy for dahongjun. This explains the reason of existence of so many different genotypes within a local population.

In future research, a larger sample size in CW and GX would help determine a more correct geographical distribution for each lineage. Multiple gene genealogical analyses would be helpful to precisely distinguish distinct lineages. Potentially, there are many dahongjun lineages out there. A more intensive sampling could help identify new lineages as well as identify more accurate genet sizes.

Overall, long-distance gene flow seemed quite limited, mostly only between neighboring local populations and very rarely among regional populations. The limited gene flow and the existence of distinct genetic elements within most sampled areas suggest that care must be taken to ensure that these lineages will not be permanently lost due to overexploitation and mis-management. Our results suggest that individual populations should be separately managed and conserved to ensure the survival of individual local populations for this ecologically and commercially important *Russula* species complex.

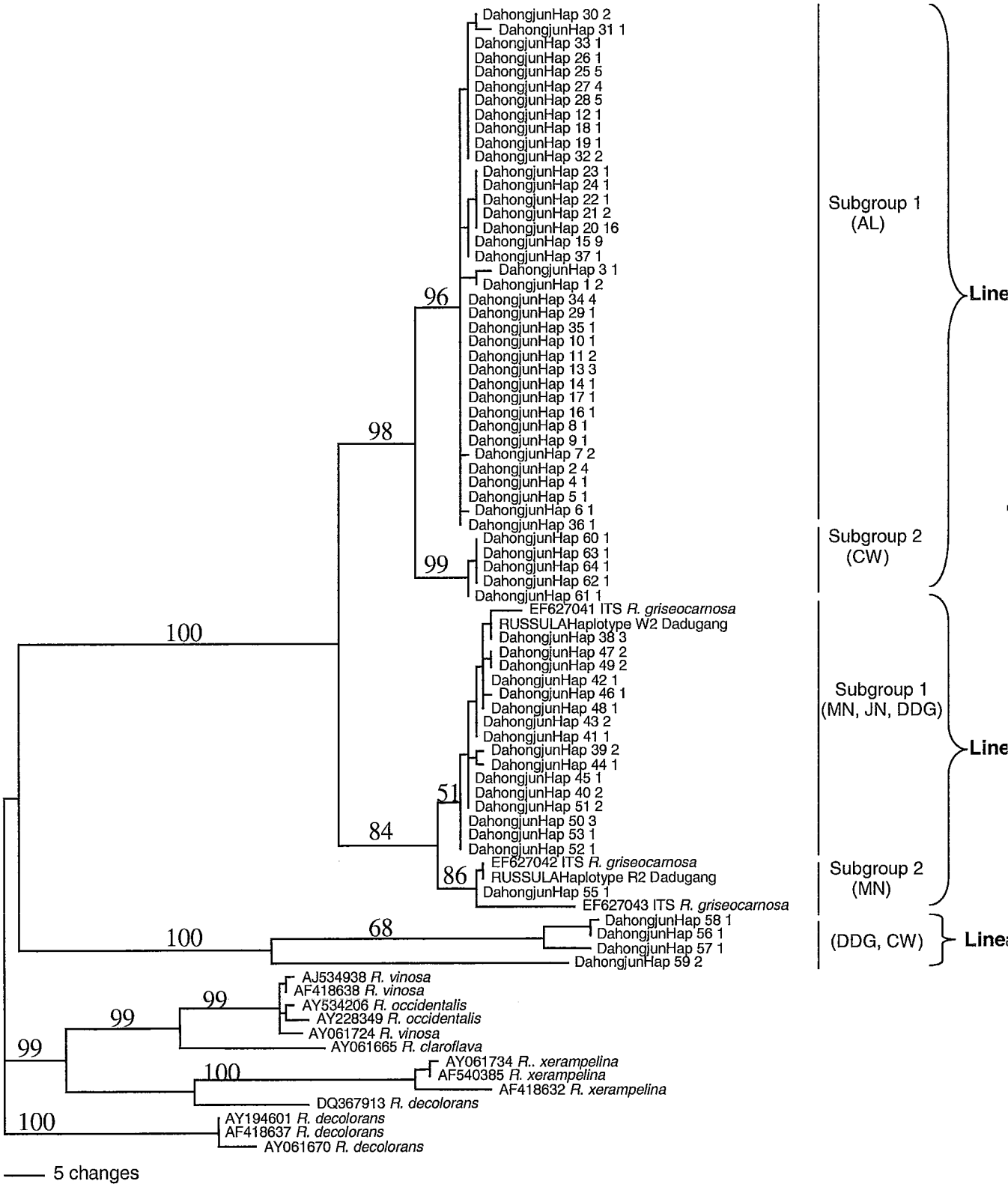
Appendix A

Supplementary Tables and Figures

Table 4A.1 Polymorphic nucleotide sites among the 63 ITS genotypes of Dahongjun. The corresponding positions for the variable nucleotides are shown. "-" represents indels.

	Nucleotide positions
	0000000000001111111111111111111133333333333334444444444444444444444444444455555
	3555566677781233333456666778888894778888999999000000011111111222344566666677934569
Genotypes	4025824617939601234552349361378930453456024567035678901234567026046902678905358594
Geno_56	ACTTGYTCTAATATTCGCATTTCGCTAT-CCTCACCC-----T----------------GGATTCCCACTTGCGTTTG-GC
Geno_57	ACTTGTTCTAATATTCGCATTTCGCTAT-CCTCACCC-----T----------------GGATTCCCACTTGCGTTTG-GC
Geno_55	ACTTGYTCTAATATTTGCATTTCGCTAT-CCTCACCC-----T----------------GGATTCCCACTTGCGTTTG-GC
Geno_58	ACT-TCTCCAATACTCGCATTTCGCTAT-CCTCACCC-----T----------------GGATTCCCACTTGCGTTTG-GC
Geno_39	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTY-Y-CCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_42	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT-YTCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_46	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT-CTCCTCCTTT-TTGGTCAGGGAAAGGACTTCACTGCGCTTG-GC
Geno_38	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTT--CTCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_44	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTTTTCCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_45	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT--CCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_43	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT--TCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_48	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT-CYCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_47	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT-CYC-TCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_49	ACTTGYTCTAATATTTGCATT-CGTTGT-ACTCCTTT-CYC-TCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_41	ACTTGYTCTAATATTTGCATT-CGTTGT-ACTCCTTT--TCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_40	ACTTGYTCTAATATTTGCATT-CGTTGT-ACTCCTTY--YCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_50	ACTTGYTCTAATATTTGCATT-CGTTGT-ACTCCTTT--CCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_51	ACTTGTTCTAATATTTGCATT-CGTTGT-ACTCCTTT--CCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_53	ACTTGCTCTAATATTTGCATT-CGTTGT-ACTCCTTT--CCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_52	ACTTGCTCTAATATTTGCATT-CGTTGT-ACYCCTTT--CCCTCCTTT-TTGGTCAGGGAARGGACT-CACTGCGCTTG-GC
Geno_54	ACTTGCTCTAATACTTGCACT-CGTTAT-ACTACATT-CCCCTCCTTT-TTGGTCAGGGAAAGGACT-CACTGCGCTTG-GC
Geno_33	GCT-CCTCCGTAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_32	GCT-CCTCCRTAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_26	GCT-CCTCCGTAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GM
Geno_25	GCT-CCTCCGTAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_34	GCT-CCTCCRTAGCTTGCATT-CACTAYTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_30	GCT-CCTCCGTAGCTTGCATC-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_31	GCT-CCTCCGTAGCTTGCATC-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_27	GCT-CCTCCGTAGCTTGCATY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_28	GCT-CCTCCGTAGCTTGCATY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_29	GCT-CCTCCATAGCTTGCATY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GM
Geno_23	GCT-CCTCCATAGCTTGCGTY-CACTATTCTTCCTCC-----YCCTTCATCGGTCAGGGAAAGGACTTCACTGCTCTCG-GC
Geno_35	GCT-CCTCCATAGCTTGCATY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_10	GCT-CCTCCRTAGCTTGCRTY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_11	GCT-CCYCCRTAGCTTGCRTY-CACTATTCTTCCTCC-----YCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_12	GCT-CCTCCRTAGCTTGCRTY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_13	GCT-CCTCCRTAGCTTGCRTY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_14	GCT-CCTCCRTAGCTTGCRTY-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_17	GCT-CCTCCRTAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_15	GCT-CCTCCRTAGCTTGCGTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_16	GCT-CCTCCRTAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GM
Geno_18	GCT-CCTCCGTAGCTTGCGTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_19	GCT-CCTCCGTAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_08	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-RC
Geno_09	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-RC
Geno_24	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCTCTCG-GC
Geno_07	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_02	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_03	GCT-CCTCCRTAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_04	GCT-CCTCCATAGCTTGCRTT-CACTAYTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCG-GC
Geno_05	GCT-CCTCCRTAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCR-RC
Geno_06	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCKCTCR-RC
Geno_01	GCT-CCTCCATAGCTTGCRTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_22	GCT-CCTCCATAGCTTGCGTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCTCTCG-RC
Geno_21	GCT-CCTCCATAGCTTGCGTT-CACYATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCTCTCR-RC
Geno_20	GCT-CCTCCATAGCTTGCGTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCTCTCG-GC
Geno_37	GCT-CCTCCATAGCTTGCGTT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-GC
Geno_36	GCT-CCTCCATAGCTTGCATT-CACTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGACTTCACTGCGCTCG-RC
Geno_61	GCT-CCTCCATAGCGCAAATT-CGTTATTCTTCCTCC-----YCCTTCATCGGTCAGGGAAAGGATTTCACCGCGCCTGTGC
Geno_59	GCT-CCTCCATAGCGCAAATT-CGTTATTCTTCCTCC-----CCCTTCATCGGTCAGGGAAAGGATTTCACCGCGCCTGTGC
Geno_60	GCT-CCTCCATAGCGCAAATT-CGTTATTCTTCCTCC-----TCCTTCATCGGTCAGGGAAAGGATTTCACCGCGCCTGTGC
Geno_62	GCT-CCTCCATAGCGCAAATT-CGTTATTCTTCCTCC-----CCCTTCATCGGTCAGGGAAAGGAYYTCACYGCGCYTGTGC
Geno_63	GCT-CCTCCATAGCGCAAATT-CGTTATTCTTCCTCC-----YCCTTCATCGGTCAGGGAAAGGAYYTCACYGCGCYTGTGC

Figure 4A.1 Maximum parsimony tree based on ITS-genotype sequences from 122 isolates of *Russula* spp. collected from 5 study sites in 3 regions in southern China. Thirteen isolates that have overall sequence identity equal or greater than 90% to our samples were retrieved from Genbank and used as outgroups. Bootstrap support values (1000 replicates) were given above branches. For each label, the first number represents the unique ITS genotype identification number and the second number in parentheses represents the total number of samples with that genotype (out of a total of 63 ITS genotypes). The distribution of these individual genotypes is indicated in Table 2.1. Tree length=268, Consistency index=0.806, Retention index=0.961.


DahongjunHap 30 2
DahongjunHap 31 1
DahongjunHap 33 1
DahongjunHap 26 1
DahongjunHap 25 5
DahongjunHap 27 4
DahongjunHap 28 5
DahongjunHap 12 1
DahongjunHap 18 1
DahongjunHap 19 1
DahongjunHap 32 2
DahongjunHap 23 1
DahongjunHap 24 1
DahongjunHap 22 1
DahongjunHap 21 2
DahongjunHap 20 16
DahongjunHap 15 9
DahongjunHap 37 1
DahongjunHap 3 1
DahongjunHap 1 2
DahongjunHap 34 4
DahongjunHap 29 1
DahongjunHap 35 1
DahongjunHap 10 1
DahongjunHap 11 2
DahongjunHap 13 3
DahongjunHap 14 1
DahongjunHap 17 1
DahongjunHap 16 1
DahongjunHap 8 1
DahongjunHap 9 1
DahongjunHap 7 2
DahongjunHap 2 4
DahongjunHap 4 1
DahongjunHap 5 1
DahongjunHap 6 1
DahongjunHap 36 1
DahongjunHap 60 1
DahongjunHap 63 1
DahongjunHap 64 1
DahongjunHap 62 1
DahongjunHap 61 1
EF627041 ITS R. griseocarnosa
RUSSULAHaplotype W2 Dadugang
DahongjunHap 38 3
DahongjunHap 47 2
DahongjunHap 49 2
DahongjunHap 42 1
DahongjunHap 46 1
DahongjunHap 48 1
DahongjunHap 43 2
DahongjunHap 41 1
DahongjunHap 39 2
DahongjunHap 44 1
DahongjunHap 45 1
DahongjunHap 40 2
DahongjunHap 51 2
DahongjunHap 50 3
DahongjunHap 53 1
DahongjunHap 52 1
EF627042 ITS R. griseocarnosa
RUSSULAHaplotype R2 Dadugang
DahongjunHap 55 1
EF627043 ITS R. griseocarnosa
DahongjunHap 58 1
DahongjunHap 56 1
DahongjunHap 57 1
DahongjunHap 59 2
AJ534938 R. vinosa
AF418638 R. vinosa
AY534206 R. occidentalis
AY228349 R. occidentalis
AY061724 R. vinosa
AY061665 R. claroflava
AY061734 R.. xerampelina
AF540385 R. xerampelina
AF418632 R. xerampelina
DQ367913 R. decolorans
AY194601 R. decolorans
AF418637 R. decolorans
AY061670 R. decolorans
96
98
99
100
51
84
86
68
100
99
99
99
100
100
Subgroup 1 (AL)
Subgroup 2 (CW)
Subgroup 1 (MN, JN, DDG)
Subgroup 2 (MN)
(DDG, CW)
Line
Line
Line
5 changes

Appendix B

Supplementary Methods

DNA Extraction From Dry Mushroom Samples

1. Grind 1.5g each sample with blue micropestle, leaving pestle in tube.
2. Apply 700μL 2X CTAB to each sample in a1.5mL tube. Place in 65°C for 1 hour.
3. Centrifuge for 7 minutes at 13K rpm.
4. Transfer upper phase liquid into a new 1.5mL tube.
5. Add equal volume (~700μL) chloroform to each tube and finger vortex. Solution should be cloudy.
6. Centrifuge for 15 minutes at 13K rpm.
7. Repeat step 5, 6 one more time.
8. Transfer upper phase into a new 1.5mL tube containing 750μL cold isopropanol.
9. Place tubes in freezer overnight.
10. Centrifuge for 30 minutes at13K rmp.
11. Decant supernatent into a beaker.
12. Add 200μL cold 70% ethanol (in freezer) to wash DNA pellet.
13. Add 200μL cold 100% ethanol (in freezer) to wash DNA pellet again.
14. Centrifuge for 5 minutes at 7K rmp.
15. Decant supernatant into a beaker and invert uncapped tubes on Kimwipes for 15-30 minutes, until dry.
16. Resuspend DNA pellet in 50μL TE buffer. Store DNA solutions at -20°C for long term usage.

Restriction Digestion

HaeIII was used for partial digestion of *Russula* genomic DNA to yield blunt cutting ends. Incubate the reaction mix at 37°C for 18min then 65°C for 15min to deactivate the enzyme. The desirable DNA concentration is at least 500μg/ml. If you know that your DNA concentration is low, increase the amount of DNA used for digestion proportionally.

Total Volume: 10μl
10× buffer: 1μl
dH_2O: 3.8μl
Enzyme: 2 unit (take 2μl of 1unit/μl enzyme)
DNA: 5μl

Operational Notes:
1. The original Enzyme concentration is 10U/μl. we should first dilute 10× buffer to 1×, then use it to dilute the enzyme from 10U/μl to 1U/μl.

2. Components should be added in the following order (make sure mixing them well) 10×Buffer-dH_2O-DNA-Enzyme

Cloning

Part 1. Ligation reaction

Genomic DNA was run on 1% agarose gel, and then extracted from the gel using PureLink™ Gel Extraction Kit from Invitrogen.

1. To prepare the ligation reaction, add the following components in order in a 0.5-ml microcentrifuge tube:

1 µl of the pPCR-Script Amp $^{SK(+)}$ cloning vector (10ng/µl)
1 µl of PCR-Script 10× reaction buffer
0.5 µl of 10 mM rATP
5.5 µl of the blunt-ended digested genomic DNA
1 µl of *Srf* I restriction enzyme (5 U/µl)
1 µl of T4 DNA ligase (4U/µl)
Final volume of 10 µl

2. Mix the ligation reaction gently and incubate this reaction for about 2 hour at room temperature.
3. Heat the ligation reaction for 10 minutes at 65°C.
4. Store it on ice until you are ready to perform transformation

Part 2. Transformation

1. Thaw the 100ul Maximum efficiency competent cells on ice.
2. Gently mix the cells by hand
3. Add 2 µl of the cloning reaction to the transformation reaction and swirl the reaction gently.
4. Incubate the tubes on ice for 30 minutes
5. Heat pulse the tubes in a 42°C water bath for 45 seconds. The duration of the heat pulse is *critical* for obtaining the highest efficiencies. Do not exceed 42°C.
6. Incubate the tubes on ice for 2 minutes.
7. Add 1.25 ml of preheated (42°C) LB broth to each tube and incubate the tubes at 37°C for 1 hour with shaking at 225–250 rpm.
8. Plate it out on LB-Amp-X-gal-IPTG plates. We plate 50ul or 100ul on each plates. Incubate them at 37°C overnight, then select white colonies.

The agar plates were made with the following compounds:

10g tryptone
5g yeast extract

5g NaCl
1ml 1M NaOH
15g agar
C (X-gal) = 64g/ml (so add 1.6ml of 40mg/ml X-gal into 1L LB media)
C (IPTG) = 0.1mM (so add 2.0ml of 0.5M IPTG to 1L LB media)

Part 3. PCR amplify our clones

Single white colonies were chosen and dipped into PCR reaction tubes.

PCR reaction (total volume 25μl)
Green Taq 2 ×: 12.5μl
Primer F (M13 F): 1μl
Primer R (M13 R): 1μl
dH2O: 10.5μl
DNA: dipped in

Alternatively,
Total 25μl
Green Taq 2 ×: 12.5μl
Primer F (M13 F): 1μl
Primer R (M13 R): 1μl
dH2O: 8.5μl
DNA: 2ul (colony dissolved in 20μl dH_2O)

PCR program
Step 1. 95°C: 10min
Step 2. 95°C: 30sec
50°C: 30 sec
72°C: 1.5 min
Repeat step 2 for 25 cycles.
Step 3. 72°C: 7min
Step 4. 4°C stand by

Big Dye Termination Sequencing Reaction Preparation

Materials:
Exonuclease I, Shrimp Alkaline Phosphatase (SAP), 5x Sequencing Buffer, BigDye® Terminator v3.1 Ready Reaction Mix, sephadex G-70.

Methods
1. Prepare sephadex 24 hours before doing the experiments. 2.5g sephadex dissolved in 40ml dH_2O, which is good for one 96-well plate. Always prepare more than you expect to use.
2. To clean off the free dNTP and phosphates, we prepare a "clean mix"
0.125μl Exo I (20mol/μl)

1μl SAP (1μ/μl)
5μl H_2O (adjust the amount of H_2O according to the amount of PCR products available)
3. Add 6.125μl clean mix to 15μl PCR products. Incubate them at 37 °C for 30min, then 80°C for 10min.
4. Dilute PCR products by 10 fold. Use the diluted one for the subsequent sequencing reaction.
5. Prepare sequencing reaction
Total volume: 12μl
DNA 2μl
5× dilution buffer: 2.5μl
big dye: 0.5μl
dH_2O: 6μl
forward or reverse primer: 1μl
6. PCR program
Step 1: 96°C 2min
Step 2: 96°C 30sec
50°C 15sec
60°C 4min (repeat step 2 for 25 cycles)
Step 3: 4°C stand by
7. Add sephadex to each well on the sephadex plates. Spin at 1500rpm for 2min to remove excess water from sephadex. Apply 12μl our samples on the center of each well. Spin it down at 1500rpm for 2min, then cover the collecting plates with aluminum foil.

References

Amaranthus, M.P., (1998) *The importance and conservation of ectomycorrhizal fungal diversity in forest ecosystems lessons from Europe and the Pacific Northwest.* Pacific Northwest Research Station General Technical Report, PNW-QTR-431.

Allen, M., FHipps, L.E. (1984) Comparative microlimates of some mycorrhizal fungi: requirements for long-range dispersal of spores. *Mycological Society of America Newsletter.* **35**: 17.

Baar, J., Ozinga, W.A., Kuyper, T.W. (1994) Spatial distribution of *Laccaria bicolor* genets reflected by sporocarps after removal of litter and humus layers in a *Pinus sylvestris* forest. *Mycological Research.* **98**: 726–728.

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

Bergemann, S.E., Miller, S.L., Garbelotto, M. (2005a) Microsatellite loci from *Russula brevipes*, a common ectomycorrhizal associate of several tree species in North America. *Molecular Ecology Notes.* **5**: 472–474.

Bergemann, S.E., Douhan, G.W., Garbelotto, M., Miller, S.L. (2005b) No evidence of population structure across three isolated subpopulations of *Russula brevipes* in oak/pine woodland. *New Phytologist.* **170**: 177-184.

Bertolino, S., Vizzini, A., Wauters, L.A., Tosi, G. (2004) Consumption of hypogeous and epigeous fungi by the red squirrel (*Sciurus vulgaris*) in subalpine conifer forests. *Forest Ecology and Management.* **202**: 227–233

Bidochka, M.J., Small, C.L.N., Spironello, M., (2005) Recombination within sympatric cryptic species of the insect pathogenic fungus *Metarhizium anisopliae. Environmental Microbiology.* **7**: 1361-1389.

Bledsoe, C.S., Tennyson, K., Lopushinsky, W. (1982) Survival and growth of mycorrhizal Douglas-fir seedlings outplanted on burned-over sites in eastern WA. *Canadian Journal of Forest Research.***12**: 720-723.

Bonello, P., Bruns, T.D., Gardes, M. (1998) Genetic structure of a natural population of the ectomycorrhizal fungus *Suillus pungens. New Phytologist.* **38**: 533-542.

Brundrett, M. (2004) Diversity and classification of mycorrhizal associations. *Biological Review - Cambridge Philosophical Society.* **79**: 473-495.

Bruns TD, Gardes M. (1993). Molecular tools for the identification of ectomycorrhizal fungi: taxon-specific oligonucleotide probes for suilloid fungi. *Molecular Ecology*, 2: 233–242.

Bruns, T.D., Szaro, T.M., Gardes, M. (1998). A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. *Molecular Ecology*. **7**: 257–272.

Buyck, B., Mitchell, D., Parrent, J. (2006) *Russula parvovirescens* sp. nov., a common but ignored species in the eastern United States. *Mycologia.* **98**: 612-615.

Ciardo, D.E., Schär, G., Böttger, E.C. (2006) Internal transcribed spacer sequencing versus biochemical profiling for identification of medically important yeasts. *Journal of Clinical Microbiology*. **44**: 77-84.

Dahlberg, A., (1997) Population ecology of *Suillus variegate*s in old Swedish Scots pine forests. *Mycological Research.* **101**: 47-54.

Dahlberg, A., Strenlid, J., (1990) Population structure and dynamics in *Suillus bovinus* as indicated by spatial disturbance of fungal clones. *New Phytologis.,* **115**: 487-493.

Dahlberg A, Strenlid J, 1994. Size, distribution and biomass of genets in population of. *Suillus bovinus* (L.Fr.) Roussel revealed by somatic incompatibility. *New Phytologist.* **12**: 225-234.

Dettman, J.R., Jacobson, D.J., Taylor, J.W. (2003) A multilocus genea-logical approach to phylogenetic species recognition in the model eukaryote *Neurospora. Evolution.* **57**: 2703–2720.

Dunham, S.M., Kretzer, A., Pfrender, M.E. (2003) Characterization of Pacific golden chanterelle (*Cantharellus formosus*) genet size using co-dominant microsatellite markers. *Molecular Ecology*. **12**: 1607–1618.

Douhan, G.W., Rizzo, D.M. (2005) Phylogenetic divergence in a local population of the ectomycorrhizal fungus *Cenococcum geophilum. New Phytologist.* **166**: 263–271.

Doyle, J.J. (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany.* **17**: 144–163.

Egger, K.N. (1995) Molecular analysis of ectomycorrhizal fungal communities. *Canadian Journal of Botany*. **73**: S1415–S1422.

Excoffier, L., Smouse, P.E., Quattro, J.M. (1992) Analysis of molecular variance

inferred from metric distances among DNA genotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**: 479-491.

Felsenstein, J. (1978) Cases in which parsimony and compatibility methods will be positively misleading. *Systematic Zoology* **27**: 401-410.

Fisher, J. (1978) *R. A. Fisher: The Life of a Scientist.* New York: Wiley, ISBN 0-471-09300-9.
Gardes, M., Bruns, T.D., (1993) ITS primers with enhanced specificity for basidiomycetes — application to the identification of mycorrhizae and rusts. *Molecular Ecology.* **2**: 113–118.

Gardes, M., Bruns, T.D. (1996a) Community structure of ectomycorrhizal fungi in a *Pinus muricata* forest: above- and below-ground views. *Canadian Journal of Botany*, **74**: 1572–1583.

Gardes, M., Bruns, T.D. (1996b) ITS-RFLP matching for identification of fungi. *Methods in Molecular Biology.* **50**: 177–186.

Gardes, M., White, T.J., Fortin, J.A. (1991) Identification of indigenous and introduced symbiotic in ectomycorrhizae by amplification of the nuclear and mitochondrial ribosomal DNA. *Canadian Journal of Botany.* **69**: 180–190.

Geml, J., Laursen, G.A., O'neill, K. (2006) Beringian origins and cryptic speciation events in the fly agaric (*Amanita muscaria*). *Molecular Ecology.* **15**: 225–239.

Gherbi, H., Delaruelle, C., Selosse, M.A., Martin, F. (1999) High genetic diversity in a population of the ectomycorrhizal basidiomycete *Laccaria amethystina* in a 150-year-old beech forest. *Molelucar Ecology.* **8**: 2003–2013.

Guidot, A., Debaud, J.C., Marmeisse, R. (2001) Correspondence between genet diversity and spatial distribution of above- and below-ground populations of the ectomycorrhizal fungus *Hebeloma cylindrosporum*: mycelial persistence and inbreedin. *Heredity.* **84**: 294-302.

Gryta, H., Debaud, J.C., Effosse, A. (1997) Fine-scale structure of populations of the ectomycorrhizal fungus *Hebeloma cylindrosporum* in coastal sand dune forest ecosystems. *Molecular Ecology.* **6**: 353–364.

Gryta, H., Debaud, J.C., Marmeisse, R. (2000) Population dynamics of the symbiotic mushroom *Hebeloma cylindrosporum*: mycelial persistence and inbreeding. *Heredity.* **84**: 294–302.

Hedh, J., Samson, P., Erland, S., Tunlid, A. (2008) Multiple gene genealogies and species recognition in the ectomycorrhizal fungus *Paxillus involutus. Mycological*

research. **112**: 965-975.

Hirose, D., Kikuchi, J., Kanzaki, N., (2004) Genet distribution of sporocarps and ectomycorrhizas of *Suillus pictus* in a Japanese white pine plantation. *New Phytologist.* **164**: 527–541.

Horton, T.R., Bruns, T.D. (2001) The molecular revolution in ectomycorrhizal ecology: peeking into the black-box . *Molecular Ecology.* **10**: 1855-1871.

Huai, W.X., Guo, L.D., He, W. (2003) Genetic diversity of an ectomycorrhizal fungus *Tricholoma terreum* in a Larixprincipis-rupprechtii stand assessed using random amplified polymorphic DNA. *Mycorrhiza.* **13**: 265–270.

Jeffries, P., Gianinazzi, S., Perotto, S. (2003) The contribution of arbuscular mycorrhizal fungi in sustainable maintenance of plant health and soil fertility. *Biology and Fertility of Soils.* **37**: 1–16.

Jonsson, L., Dahlberg, A., Nilsson, M.C. (1999a) Continuity of ectomycorrhizal fungi in self-regenerating boreal *Pinus sylvestris* forests studied by comparing mycobiont diversity on seedlings and mature trees. *New Phytologist.* **142**: 151–162.

Jonsson, L., Dahlberg, A., Nilsson, M.C., (1999b) Ectomycorrhizal fungal communities in late-successional Swedish boreal forests, and their composition following wildfire. *Molecular Ecology.* **8**: 205–215.

Jonsson, T., Kokalj, S., Finlay, R. (1999c) Ectomycorrhizal community structure in a limed spruce forest. *Mycological Research.* **103**: 501–508.

Kårén, O., Hogberg, N., Dahlberg, A. (1997) Inter- and intraspecific variation in the ITS region of rDNA of ectomycorrhizal fungi in Fennoscandia as detected by endonuclease analysis. *New Phytologist.* **136**: 313–325.

Kauserud, H., Stensrud, Ø., DeCock, C., Shalchian-Tabrizi, K., Schumacher, T. (2006) Multiple gene genealogies and AFLPs suggest cryptic speciation and long-distance dispersal in the basidiomycete *Serpula himantioides* (Boletales). *Molecular Ecology.* **15:** 421–431.

Kauserud H, Svegarden IB, Decock C, Hallenberg, N. (2007) Hybridization among cryptic species of the cellar fungus *Coniophora puteana* (Basidiomycota). *Molecular Ecology.* **16**: 389–399.

Kirk, P.M., Cannon, P.F., David, J.C., Stalpers, J.A. (2001) *Dictionary of the Fungi, 9th edition.* Oxon: CAB International.

Kretzer, A.M., Dunham, S., Molina, R. (2003). Microsatallite markers reveal the

below ground distribution of genets in two speices of *Rhizopogon* forming tuberculate ectomycorrhizas on Douglas fir. *New Phytologist.* **161**: 313-320.

Lamb, R.J. (1979) Factors responsible for the distribution of mycorrhizal fungi of Pinus in eastern Australia. *Australian Journal of Forest Research.* **9**: 25–34.

Lan, L., Xu, J.P. (2006) Multiple gene genealogical analyses suggest divergence and recent clonal dispersal in the opportunistic human pathogen *Candida guilliermondii. Microbiology.* **152**: 1539-1549.

Legon, N.W., Henrici, A., Roberts, P.J., Walting, R. (2005) *Checklist of the British and Irish Basidiomycota.* Royal Botanic Gardens, Kew 517.

Lepage, O.M. (1997) He´morragie dans les poches gutturales. Anatomie, diagnostic et e´ tiologie. *Pratique Vete´ rinaire Equine.* **26**: 225–261.

Li, M.C., Xu, J.P. (2009) Molecular ecology of ectomycorrhizal fungi: molecular markers, genets and ecological importance. *Acta Botannica Yunnanica.* In press.

Lian, B., Zang, J.P., Hou, W.G., Yuan, S., Smith, D.L. (2008) PCR-based sensitive detection of the edible fungus *Boletus edulis* from rDNA ITS sequences. *Journal of Biotechnology.* **11**:1-8.

Lian, C.L., Narimatsu, M., Nara, K., Hogetsu, T. (2006) *Tricholoma matsutake* in a natural *Pinus densiflora* forest: correspondence between above- and below-ground genets, association with multiple host trees and alteration of existing ectomycorrhizal communities. *New Phytologist.* **171**: 825–836.

Liang, Y., Guo, L.D., Ma, K.P. (2004a) Spatial pattern of the most common late-stage ectomycorrhizal fungi in a subtropical forest in Dujiangyan, southwest of China. *Acta Botanica Sinica.* **46**: 29- 34.

Liang, Y., Guo, L.D., Ma, K.P. (2004b) Genetic structure of a population of the ectomycorrhizal fungus *Russula vinosa* in subtropical woodlands in southwest China. *Mycorrhiza* **14**: 235–240.

Luoma, D.L., Eberhart, J.L., Amaranthus, M.P. (1997) *Biodiversity of ectomycorrhizal types from southwest Oregon.* In: Kaye TN, Liston A, Love RM, Luoma DL, Meinke RJ, Wilson MV(ed). Conservation and Management of Native Plants and Fungi. Oregon: Native Plant Society of Oregon, 249–253.

Maddison, W.P. (1997) Gene trees in species trees. *Systematic Biology.* **46**: 523–536.

McCune, B., Mefford, M.J. (1999) *PC-ORD: multivariate analysis of ecological data. Version 4. User's guide.* In: MjM Software Design. Oregon: Gleneden

Beach.

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

Molina, R., Trappe, J.M. (1982) Patterns of ectomycorrhizal host specificity and potential among pacific northwest conifers and fungi. *Forest Science.* **28**: 423-458.

Mueller-Dombois, D., Ellenberg, H. (1974) *Aims and Methods of Vegetation Ecology.* New York: John Wiley and Sons.

Mullis, K.B., Faloona, F.A. (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymlogy.* **155**: 335-350.

Myers, N., Mittermeier, R.A., Mittermerier, C.G. (2000) Biodiversity hotspots for conservation priorities. *Nature.* **403**: 853-858.

Palmer, M.W. (1991) Estimating species richness: the second-order jackknife reconsidered. *Ecology*, **72**: 1512-1513.

Pritsch, K., Boyle, H., Munch, J.C., (1997) Characterization and identification of black alder ectomycorrhizas by PCR-RFLP analyses of the rDNA internal transcribed spacer (ITS). *New Phytologist.* **137**: 357–369.

Pritsch, K., Munch, J.C., Buscot, F. (2000) Identification and differentiation of mycorrhizal isolates of black alder by sequence analysis of the ITS region. *Mycorrhiza.* **10**: 87–93.

Pujol, C., Dodgson, A., Soll, D.R. (2005) *Population genetics of ascomycetes pathogenic to humans and animals. In Evolutionary Genetics of Fungi*, Edited by J.P. Xu. UK: Horizon Scientific Press Norwich. pp. 149-188.

Read, D.J. (1991) Mycorrhizas in ecosystems. *Experentia.* **47**: 376–391.

Redecker, D., Szaro, T.M., Bowman, R.J., Bruns, T.D. (2001) Small genets of *Lactarius xanthogalactus, Russula cremoricolor* and *Amanita francheti* in late-stage ectomycorrhizal successions. *Molecular Ecology.* **10**: 1025–1034.

Remy, W., Taylor, T.N., Hass, H. (1994) Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proceedings of the National Academy of Sciences (USA),* **91**:11841-11843.

Richard, F., Millot, S., Gardes, M. (2005) Diversity and specificity of

ectomycorrhizal fungi retrieved from an old-growth Mediterranean forest dominated by Quercus ilex. *New Phytologist.* **166**: 1011–1023.

Richardson, M.J. (1970) Studies of *Russula emetica* and other agarics in a Scots pine plantation. *Transactions of The British Mycological Society.* **55**: 217-229.

Robinson, C.H., Szaro, T.M., Izzo, A.D. (2009) Spatial distribution of fungal communities in a coastal grassland soil. *Soil Biology and Biochemistry.* **41**: 414–416.

Romagnesi, H. (1967) *Les Russules d'Europe et d'Afrique du Nord. Parigi*: Bordas.
Singer, R. (1986) *The Agaricales in Modern Taxonomy. 4th edition.* Koenigstein: Koeltz Scientific Books.

Saitou, N., Nei, M., (1987) The neighbor- joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution.* **4**: 406–425.

Sawyer, N.A., Chambers, S.M., Cairney, J.W.G. (1999) Molecular investigation of genet distribution and genetic variation of *Cortinarius rotundisporus* in eastern Australian sclerophyll forests. *New Phytologist.* **142**: 561–568.

Schneider, S., Roessli, D., Excoffier, L. (2000) *Arlequin: A software for population genetics data analysis. Ver 2.000.* Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Selosse, M.A., Martin, F., Bouchard, D. (1999) Structure and dynamics of experimentally introduced and naturally occurring *Laccaria sp.* Discrete genotypes in a Douglas fir plantation. *Applied and Environmental Microbiology.* **65**: 2006-2014.

Selosse, M.A., Richard, F., He, X. (2006) Mycorrhizal networks: des liaisons dangereuses? *Trends in Ecology and Evolution* **21**: 621–628.

Sha, T., Xu, J.P. Palanichamy, M. Zhang, H.B., Li, T., Zhao, Z.W., Zhang, Y.P. (2008) Genetic diversity of the endemic gourmet mushroom *Thelephora ganbajun* from southwestern China. *Microbiology.* **154**: 3460-3468.

Sha, T., Zhang, H., Ding, H., Li, Z., Cheng, L. (2007) Genetic diversity of *Tricholoma matsutake* in Yunnan Province. *Chinese Science Bulletin.* **52**: 1212-1216.

Simard, S.W., Perry, D.A., Jones, M.D. (1997) Net transfer of carbon between ectomycorrhizal tree species in the field. *Nature.* **388**: 579-582.

Simard, S.W., Jones, M.D., Durall, D.M. (2002) Carbon and nutrient fluxes

within and between mycorrhizal plants. *Ecological Studies.* **157**: 33-61.

Smith, S.E., Read, D.J. (1997) *Mycorrhizal Symbiosis (2nd edition).* London: ISBN 0-12-652840-3.

Stephens, M., Smith N.J., Donnellly P. (2001) A new statistical method for genotype reconstruction form population data. *The American Journal of Human Genetics.* **68**: 978-989.

Swofford, D.L. (2002) PAUP*: *Phylogenetic analysis using parsimony (and other methods)*, version 4. Sunderland, MA: Sinauer Associates.

Sylvia, D.M., Jasrstfer, A.G. (1997) Distribution of mycorrhizae on competing pines and weeds in a southern pine plantation. *Soil Science Society of America Journal.* **61**:139-144.

Taylor, A.F.S., Alexander, I.J. (1989) Demography and population dynamics of ectomycorrhiza of Sitka spruce fertilized with N. *Agriculture Ecosystems and environment.* **28**: 493-496.

Taylor, D.L., Bruns, T.B. (1999) Community structure of ectomycorrhizal fungi in a pinus muricata forest: minimal overlap between the mature forest and resistant propagule communities. *Molecular Ecology.* **8**:1837-1850.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S., Fisher, M.C. (2000) Phylogenetic species recognition and species concepts in fungi . *Fungal Genetics and Biology.* **31**:21–32.

Tedersoo, L., Kõljalg, U., Hallenberg, N., Larsson, K.H. (2003) Fine scale distribution of ectomycorrhizal fungi and roots across substrate layers including coarse woody debris in a mixed forest. *New Phytologist.* **159**:153-165.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997) The CLUSTAL_X–Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **24**: 4876–4882.

Trappe, J.M. (1987) *Phylogenetic and ecologic aspects of mycotrophy in the angiosperms from an evolutionary standpoint.* In: Safir GR (ed). Ecophysiology of VA Mycorrhizal Plants. Florida: CRC Press.

Wadud, M.A., Lian, C.L., Nara, K., Ishida, T.A., Hogetsu, T. (2006) Development of microsatellite markers from an ectomycorrhizal fungus, *Laccaria amethystina*, by a dual-suppression-PCR technique. *Molecular Ecology Notes.* **6**: 130-132.

Wang, X.H., Yang, Z.L., Li, Y.C., Knudsen, H., Liu, P.G. (2009) *Russula griseocarnosa* sp. nov. (Russulaceae, Russulales), a commercially important

edible mushroom in tropical China: mycorrhiza, phylogeneic position, and taxonomy. *Nova Hedwigia.* **88**: 269-282.

White, T.J., Bruns, T.D., Lee, S.B. (1990) *Amplicaion and direct sequencin of fungal ribosomalRNA genes for phylogenetics.* In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds). In PCR protocols: a Guide to Methods and Applications. London: Academic Press, 315-322.

Wright, S. (1943) Isolation by distance. *Genetics.* **28**:114–138.

Xu, J.P. (2005) *Evolutionary Genetics of Fungi.* U.K.: Horizon Bioscience.

Xu, J.P. (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular Ecology.* **15**: 1713-1731.

Xu, J.P., Guo, H., Yang, Z.L. (2007) Single nucleotide polymorphisms in the ectomycorrhizal mushroom *Tricholoma matsutake. Microbiology.* **53**: 2002-2012.

Xu, J.P., Sha, T., Li, Y.C., Zhao, Z.W., Yang, Z.L. (2008) Recombination and genetic differentiation among natural populations of the ectomycorrhizal mushroom *Tricholoma matsutake* from southwestern China. *Molecular Ecology* **17**: 1238-1247.

Xu, J.P., Vilgalys, R., Mitchell, T.D. (2000) Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans. Molecular Ecology.* **9**: 1471–1481.

Yang, Z.L. (2002) On wild mushroom resources and their utilization in Yunnan Province, Southwest China. *Journal of Natural Resources.* **17**: 463-469.

Yang, Z.L., Piepenbring, M. (2004) Wild edible fungi in the Yunnan Province, southwestern China. In: Agerer R., M. Piepenbring & Blanz (eds.). *Frontiers in Basidiomycote Mycology.* IHW-Verlag. pp. 419-426.

Zhou, Z., Miwa, M., Hogetsu, T. (1999) Analysis of genetic structure of a *Suillus grevillei* population in a *Larix kaempferi* stand by polymorphism of inter-simple sequence repeat (ISSR) *New Phytologist.* **144**: 55–63.