江西农业大学学报 2010,32(5):1010 – 1017 http://xuebao.jxau.edu.cn
Acta Agriculturae Universitatis Jiangxiensis E – mail: ndxb7775@sina.com

DNA Barcoding, Fungal Diversity, and Authentication of Wild Gourmet Mushrooms

XU Jianping

(Department of Biology, McMaster University, 1280 Main St W, Hamilton, Ontario, L8S 4K1, Canada)

Abstract: The application of molecular tools for the identification of strains, populations and species has become a common practice in a variety of applied and basic investigations in many parts of the world. However, standardization of such applications varies widely among organisms and scientific fields. In this mini – review, the author provides a brief introduction to one of the most prominent effort for species identification – the international barcode of life (iBOL) project, discusses the features of fungal diversity including the proposed fungal barcode DNA fragment-the intergenic spacer regions (ITS) of the nuclear ribosomal RNA gene clusters, and illustrates the potential promises and problems of using ITS for barcoding and for analyzing the phylogeographic pattern of the wild gourmet mushroom *Tricholoma matsutake* species complex. The analyses show that identification to species level is often insufficient for practical applications and that sequences from multiple genes in combination with critical morphological and physiological evaluations are needed to identify strains, populations and species.

Key words: barcoding; ITS; gourmet mushrooms; population genetics; counterfeit

中图分类号: Q756; Q789 **文献标志码:** A **文章编号:** 1000 – 2286(2010)05 – 1010 – 08

1 DNA barcoding

Accurate species identification is a prerequisite for almost all biological investigations. The description and classification of organisms have traditionally been the work of specialist taxonomists, who also generally provide nomenclatural backbones for non-specialists for identification purposes. In today's society, while the number of traditional taxonomists has decreased for most groups of organisms in most fields, the need for accurate and efficient species identification has increased significantly. This has been driven by multiple factors such as for critically assessing and maintaining biodiversity, ensuring bio-security in the age of terrorism and anti-terrorism, protecting engendered species, preventing emerging infections, and avoiding pandemics.

Date of receiving: 2010 – 06 – 02

About the author: 徐建平(1965—),男,江西樟树人,博士,加拿大麦克马斯特大学教授。1981 年考入江西农业大学农学系,1985 年本科毕业后考入南京农业大学硕士研究生;1988 年 7 月—1989 年 5 月在浙江省农业科学研究院农业微生物研究所工作;1989 年 6 月—1991 年 2 月在江西省科学研究院微生物研究所工作;1991 年 3 月—8 月在加拿大农业研究站任研究助理;1991 年 9 月—1996 年 12 月先后在加拿大多伦多大学植物学系读理学硕士及理学博士学位;1997 年 1 月—2000 年 7 月在美国杜克大学微生物系任博士后研究助理;2000 年 8 月—2005 年 6 月在加拿大麦克马斯特大学生物系任助理教授;2005 年 7 月升任终身教授;2010 年 7 月始任副系主任。十多年来,在加拿大麦克马斯特大学独立主持一个微生物生态和遗传进化实验室,参与 20 部本领域专著撰写,主编两部专著,发表 SCI 研究论文 70 多篇,包括本领域高影响因子的学术期刊《Current Biology》,《Trends in Genetics》,《PLoS One》及《PNAS》等,受邀为《Emerging Infectious Diseases》,《Current Issues in Molecular Biology》,《Molecular Ecology》等期刊撰写综述文章 7 篇。在国际学术会议上做报告 60 余次,组织过 2 次国家及国际学术会议。2001 年荣获加拿大安大略"Premier's Research Excellence Award",2005 年荣获"Young Scientist Award of the Genetics Society of Canada",主持中国国家基金委海外杰出青年基金(2007—2009)。E-mail: jpxu@mcmaster.ca.

2010 is the International Year of Biodiversity. People from all walks of life are encouraged to participate in a variety of efforts to help understand, appreciate, and conserve our rapidly declining biological resources on Earth. To care about biodiversity is not only an academic exercise or for aesthetic reasons, but also for practical issues such as dealing with invasive species, developing appropriate diagnostic tools for the treatments of infectious diseases, and monitoring global trades in foods and other biological materials. For example, invasive plants and pathogens cause tens of billions of dollars of losses to crops and forest products each year. The lack of monitoring of the Zebra mussel and its introduction from Europe to North America has already caused tens of millions of dollars of damages to ships and water drainage pipes and an estimated over $5 billion will be needed to fully repair the damage. Furthermore, each year, millions of lives and billions of hours of productivity are lost due to infection by emerging human pathogens dispersed from one region to another. The increased global trade in food also demands fast and accurate species identifications.

To deal with these issues, the method of 'DNA Barcode of Life' was developed as a standardized, rapid and inexpensive species identification system accessible to both specialists and non-specialists. DNA barcode refers to a short standardized sequence that enables species discrimination among a large group of closely related organisms [1]. The scientific basis of barcoding is simple and has been noted for a long time: (i) every organism on Earth requires a set of genes to survive and reproduce; (ii) some of these genes are shared among a large number of organisms; and (iii) within those genes are short sequences unique to the species that allow specific identification of the organisms to the species level. Though using DNA sequence for species identification had existed for several decades, the concept of barcoding all organisms was pioneered and developed by researchers at the University of Guelph in Canada [1]. The initial focus of barcoding was on animals but the effort has subsequently extended to plants, fungi and other groups of eukaryotic organisms.

The DNA barcoding effort has now evolved into the overarching International Barcode of Life project (iBOL; www. ibolproject. org). iBOL is among the largest biodiversity genomics initiative in the world and its objective was to create a digital identification system for all living organisms. The iBOL global partnership currently involves 25 nations arranged in three collaborative configurations. Specifically, based on funding and infrastructure commitments from national governments, iBOL is organized to four Central Nodes, ten Regional Nodes, and six National Nodes (Figure 1). The four Central Nodes are Canada, China, European Union (including France, Germany, Italy, Netherlands, Portugal, Spain, United Kingdom) and the United States. They will have primary responsibility in funding and coordinating iBOL and in supporting its core analytical facilities and data archives. The ten Regional Nodes are Argentina, Australia, Brazil, India, Korea, Mexico, New Zealand, Norway, Russia, and South Africa. These countries will oversee regional barcode efforts. In addition, New Zealand and Norway will lead barcoding programs in the Antarctic and Arctic respectively. The six National Nodes including Argentina, Colombia, Costa

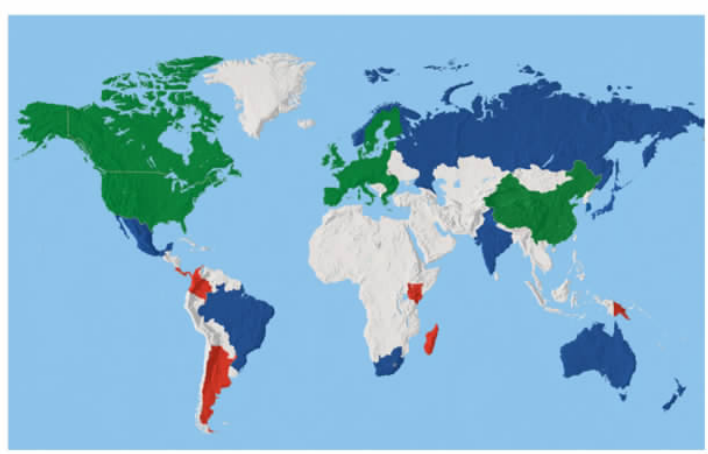

Fig. 1　Participating nations for the international barcode of life (iBOL) project. Colored countries represent those whose governments officially contribute to the iBOL project. Differentially colored countries contribute different amount of resources to the barcoding effort. Green: Central Nodes ($25million); Blue: Regional Nodes ($5million); and Red: National Nodes ($1million). Please refer to text for details of the specific countries.

Rica, Kenya, Madagascar and Panama will augment national biodiversity surveys via barcoding. Researchers from National Node countries will assemble specimens and interpret results but will often use sequencing facilities in the Central Nodes. Aside from official government participations, hundreds of individual labs and institutions from over 50 countries are actively involved in barcoding, including natural history museums, zoos, herbaria, botanical gardens, university as well as private companies and governmental organizations (www.barcodeoflife.org).

The standard approach for generating the barcode for individual species is relatively simple and straightforward (Figure 2). However, it does demand large accessory information related to individual specimens and to the quality of the DNA sequence. All these data are stored and managed in the Barcode of Life Data Systems (BOLD, http://www.boldsystems.org), an informatics workbench for the acquisition, storage, analysis and publication of DNA barcode records. Different from other DNA databases such as GenBank, BOLD contains not only molecular data but also detailed morphological and distributional data. As a result, it has the great potential to bridge a long-standing bioinformatics chasm typical of most databases for species identification. At present, BOLD is the go-to site for DNA-based identification for several large groups of taxa and is freely available to anyone interested in DNA barcoding. In the brief seven-year period, a large number of articles have been published on barcoding various groups of organisms. As of July 2010, its data repository contains 844,163 specimens from 70,291 species (http://www.boldsystems.org). However, the current data is highly biased towards animals and relatively few eukaryotic microbe data are present in the database.

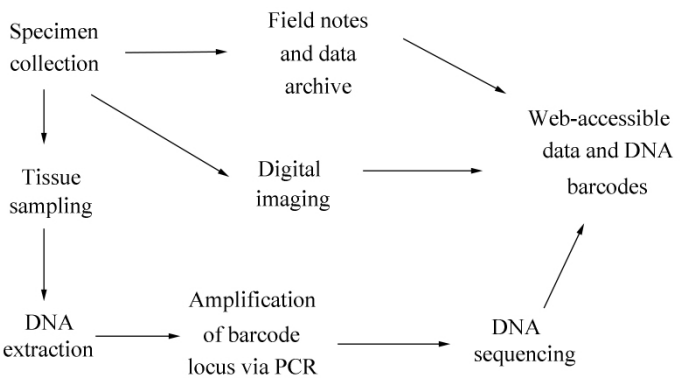

Fig. 2 The flow chart for assembling typical barcode data

Among the barcode loci, the mitochondrial cytochrome oxidase subunit 1 (*COX*1) has been the marker of choice for animal barcoding [10]. This gene is universally distributed in animals, is often highly polymorphic among species but relatively homogeneous within species, and is easy to extract and amplify due to its high copy number in most cells. In addition, specific primers are available to amplify large groups of organisms. In contrast, this marker is not effective for most plants. Instead, a combination of two genes in the plastid, *rbcL* and *matK*, have emerged as the preferred marker choice for plant barcoding, at least during initial screening [5]. In the case of a lack of discriminatory polymorphism at these two loci among closely related plant species, additional markers appropriate for the specific group of plants will need to be developed to help generate diagnostic features. As shown below, a locus different from the three mentioned above (*COX*1, *rbcL* and *matK*) is proposed for fungal barcoding.

2 Fungal diversity and fungal barcoding

There are about 80 000 named fungal species. These fungi have been traditionally divided into six main Divisions (the equivalent of phyla in plants and animals), differentiated by their sexual reproductive structures as well as other morphological, structural and biochemical properties. However, based on multiple gene sequence analyses, two of the six main Divisions, Myxomycetes and Oomycetes are now considered not true fungi and are grouped into protists and Straminipila respectively. The true fungi are now classified into five main Divisions Chytrids, Zygomycetes, Glomeromycetes, Ascomycetes, and Basidiomycetes [14-15] (Figure 3). In some textbooks, another Division, the Deuteromycetes or Fungi Imperfecti, is still presented. This group refers to moulds and yeasts which when they were first described, had only the asexual form (also called the anamorphic

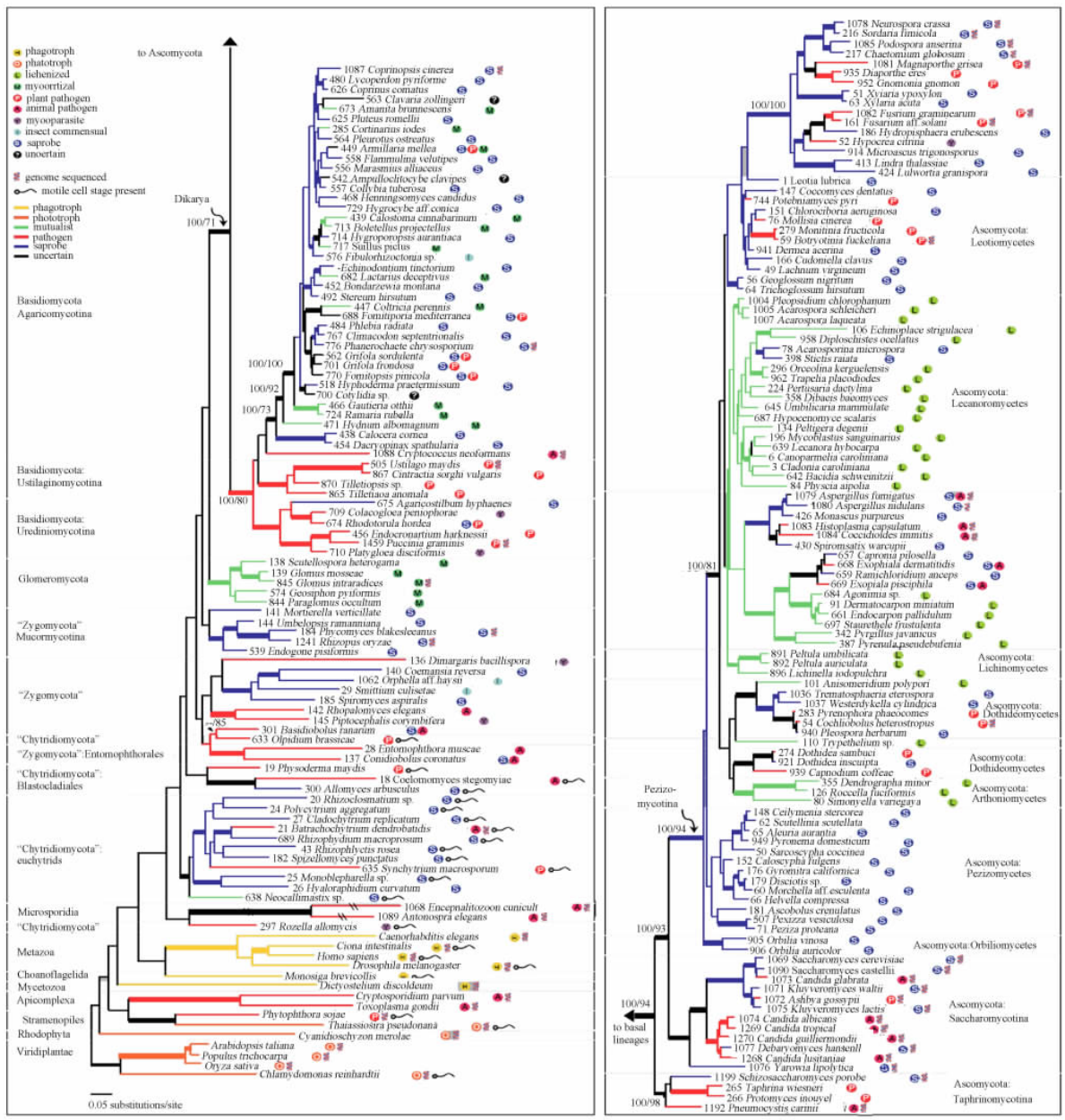

Fig. 3　An overview of the fungal tree of life based on the analyses of six gene fragments (from James et al, 2006)

fungi). However, based on molecular phylogenetic studies, most fungi in this Division are phylogenetically related to sexual Ascomycetes. A minority of these asexual fungi belonged to the Basidiomycetes [15].

Based on estimates from ecological and environmental DNA sampling, a large number of fungi, estimated up to 90%-99% of all fungi are not yet described. The main difficulties for describing and for studying these fungi have been our inability to observe and culture most of these fungi and the lack of distinguishing morphological and physiological features to separate closely related fungal species. The advent of DNA-based identification techniques, including the concerted fungal barcoding initiative, have accelerated fungal identification and will continue to impact this field in the near future.

Among the many loci used for fungal species delineation, the internal transcribed spacer (ITS) regions of the nuclear ribosomal RNA gene cluster have emerged as the top choice as the fungal barcode. In some groups of fungi, the barcode used in animals (i.e. *COX*1) was found effective for species identification. However, for many groups of fungi, the presence of introns within *COX*1, sometime up to 20 kb in length, makes this locus difficult to amplify and to obtain DNA sequence. However, the ITS barcode also has its own problems. For

example, there was limited variation among closely related species at the ITS regions within several groups of fungi. In addition, the multi-copy nature of ITS often results in significant variation among the copies within an individual strain, making DNA sequencing chromatograms unreadable. A recent review discussed these issues and summarized recent progresses toward fungal barcoding [17].

3 ITS sequence variation in a gourmet mushroom

The problems and promises of using ITS sequences as barcode for fungi can be illustrated using population samples of the gourmet mushroom *Tricholoma matsutake* species complex. Mushrooms have been widely used since ancient times not only as foods or food flavoring materials but also for spiritual and medicinal purposes [19]. The ectomycorrhizal matsutake (or Pine-mushroom, *Phylum Basidiomycota*) is among the most revered and valuable mushrooms in the world. Its accurate identification is thus of significant commercial importance.

The commonly used term matsutake refers to a loosely defined species complex in the genus *Tricholoma*. Like most mushrooms, the major biomass of the matsutake mushroom is underground in the soil in which their mycelia form an extensive network called Shiro. For species in *Tricholoma* and many other ectomycorrhizal fungi, their mycelia establish unique symbiotic relationships with the roots of conifer and broadleaf trees [3]. Due to the lack of clearly defined morphological or physiological characters to distinguish populations and species of the commercially available samples, it has been difficult to establish natural groups within the *T. matsutake* species complex. As a result, the informal name matsutakes commonly refer to those fungi recognized as valuable by the Japanese market-the world's major consumer market for matsutake mushrooms. Bergius and Danell [4] concluded that the European specimens of *T. nauseosum* and the Asian *T. matsutake* are considered conspecific and belonged to the 'true matsutake'. Though also consumed in Japan and North America, mushrooms of the species *Tricholoma magnivelare* are considered a 'matsutake-ally', similar to those of *Tricholoma caligatum* and *Tricholoma fulvocastaneum*. On the other hand, though morphologically very similar to *T. matsutake*, a more distantly related species *Tricholoma bakamatsutake* is not consumed and is called 'false matsutake' or sometimes 'fool's matsutake' in the popular literature.

Geographically, the loosely defined matsutake is broadly distributed from Europe to Northern Africa, from the sub-Himalayan region to the easternmost countries of Asia, from the Pacific Rim in North America to the Rocky Mountains, the Great Lakes, the east coast of the United States, and to Mexico. Matsutake has yet to be recorded from Africa outside the Atlas Range, from Australasia, or from South America [7]. In general, the range of matsutake roughly matches the distribution of coniferous genera such as *Pinus*, *Pseudotsuga*, *Tsuga*, *Picea*, *Cedrus and Abies* [21]. Nonetheless, several forms of matsutake frequently are found in mixed forests where associations with broadleaf trees cannot be ruled out. In southwestern China, matsutake is often found associated with oak (*Quercus* spp). *T. magnivelare* in the Pacific Northwest coast of North America often is found in pure stands of *Lithocarpus densiflora* (tanoak) [5]. *T. bakamatsutake*, the false matsutake sympatric with *T. matsutake*, is believed to be associated with Castanopsis, Fagus, Pasania and *Quercus* spp., despite its occurrence in forests where conifers also are present [13]. A similar case in North America applies to the matsutake-ally *T. caligatum* from the US and Mexico, suspected to associate with angiosperm hosts.

Despite our lack of understanding for the basic biology of these organisms, matsutakes are of significant importance, not only because of their market value as gourmet mushrooms but also their epitomizing non-timber forest products in coniferous forests in many parts of the world [1]. For example, the estimated value of non-timber forest products in mixed forests in the Pacific Northwest of the United States and British Columbia in western Canada is greater than the economic value of the timber in the same forest and the matsutake-ally *T. magnivelare* represents the most significant component of the non-timber forest products in this region [1-2].

In Japan-the world's preeminent consumer market, matsutake collected and imported from different parts of the world are priced very differently, from less than US $100 to over US $4 000/kg of fresh fruiting

bodies [16,25]. As mentioned briefly before, matsutake in Asia is produced in two major regions: eastern Himalaya which includes southwest (SW) China and Bhutan, and the Far East which includes Japan, the Korean Peninsula, and northeast (NE) China. The NE Chinese matsutake is distributed in the Changbai Mountain range bordering North Korea and covers parts of two provinces, Jilin and Heilongjiang [23]. The SW Chinese matsutake is distributed in eastern Tibet and western parts of Yunnan and Sichuan provinces [27]. In Japan, the NE Chinese matsutake is sold for an average of about US $200/kg, similar to those from North Korea, slightly lower than South Korean matsutake, and about half of those from within Japan. However, the matsutake from SW China is traded at about US $100/kg, significantly lower than their NE counterpart and other Far East matsutake [16]. Such a price variation is also seen within the Chinese domestic consumer market. Despite its high price tag, matsutake's unique aroma [20] and its demonstrated health benefits to humans [8,12] are attracting a growing number of people within China. In Japan, the Asian matsutake are generally more expensive and their flavors more attractive to consumers than those from North America and Europe.

The significant price differences for matsutakes from different parts of the world create conditions for counterfeiting both among countries as well as among areas within countries. However, while useful for discriminating the "true matsutake" from "matsutake-ally" and "false matsutake", the ITS region has very limited discriminating power for different geographic populations of the true matsutakes from different regions (Figure 4). In fact, as noted by Chapella and Garberlatto [6], some of the presumed *T. magnivelare* from North America were clustered with the Asian and European matsutake samples (Figure 4), suggesting a complex relationship between geography and genetic relationship among the populations.

To investigate the potential of matsutake counterfeiting, we recently analyzed matsutakes from two major production and trading regions in China and found evidence for significant counterfeiting [26]. In this analysis, we obtained DNA profiles of *T. matsutake* fruiting bodies from matsutake trading companies and compared them with known authentic specimens from NE China and SW China. The commercial matsutake included 107 fruiting bodies from four companies in NE China and 45 fruiting bodies from three companies in SW China. The geographically authentic matsutake samples included 38 mushrooms from four local populations in Jilin and Heilongjiang in NE China, and 183 samples from 18 local populations in Yunnan, Tibet, and Sichuan provinces in SW China [27]. Their DNA was extracted [28] and the genotypes of all 373 isolates were determined using a PCR genotyping system based on specific primers targeting the gypsy-type retroelements [16,26]. Our analyses showed that 67% commercial matsutake claimed from the northeast were in fact genetically identical at this marker to those from southwest China but different from authentic northeast Chinese samples. Such analyses highlight the importance of accurate identification of matsutake mushrooms, not only at the species level but also at population and strain level. Our analysis is similar to those found for counterfeited fish in New York City, USA [22]. Our finding suggests that caution should be applied to authenticate commercial matsutake from northeast China. Similar concerns about the authenticity of other gourmet mushrooms such as truffles, morels, and tubers have also been raised but the identification system is yet to be finalized.

Despite recent progresses, much remains unknown about the relationships among populations of *T. matsutake* from different geographic regions such as those from eastern North America, Europe, North Africa, Southwestern China and the Far East. Such understandings will help the conservation and sustainable utilization efforts. In addition, the precise relationships among closely related species in this group of mushrooms remain to be clarified. A powerful approach to address both within-species and between-species relationships among strains and populations is the gene genealogical analyses based on DNA sequences from multiple loci [24]. Such an approach has been applied to analyze many microbial, including fungal, populations [24-25] and data from those analyses will continue to impact the identification of strains, populations, and species of microbes, plants, and animals, beyond what DNA barcoding is offering.

4 Acknowledgements

Research in my lab on microbial population genetics has been supported by grants from McMaster University,

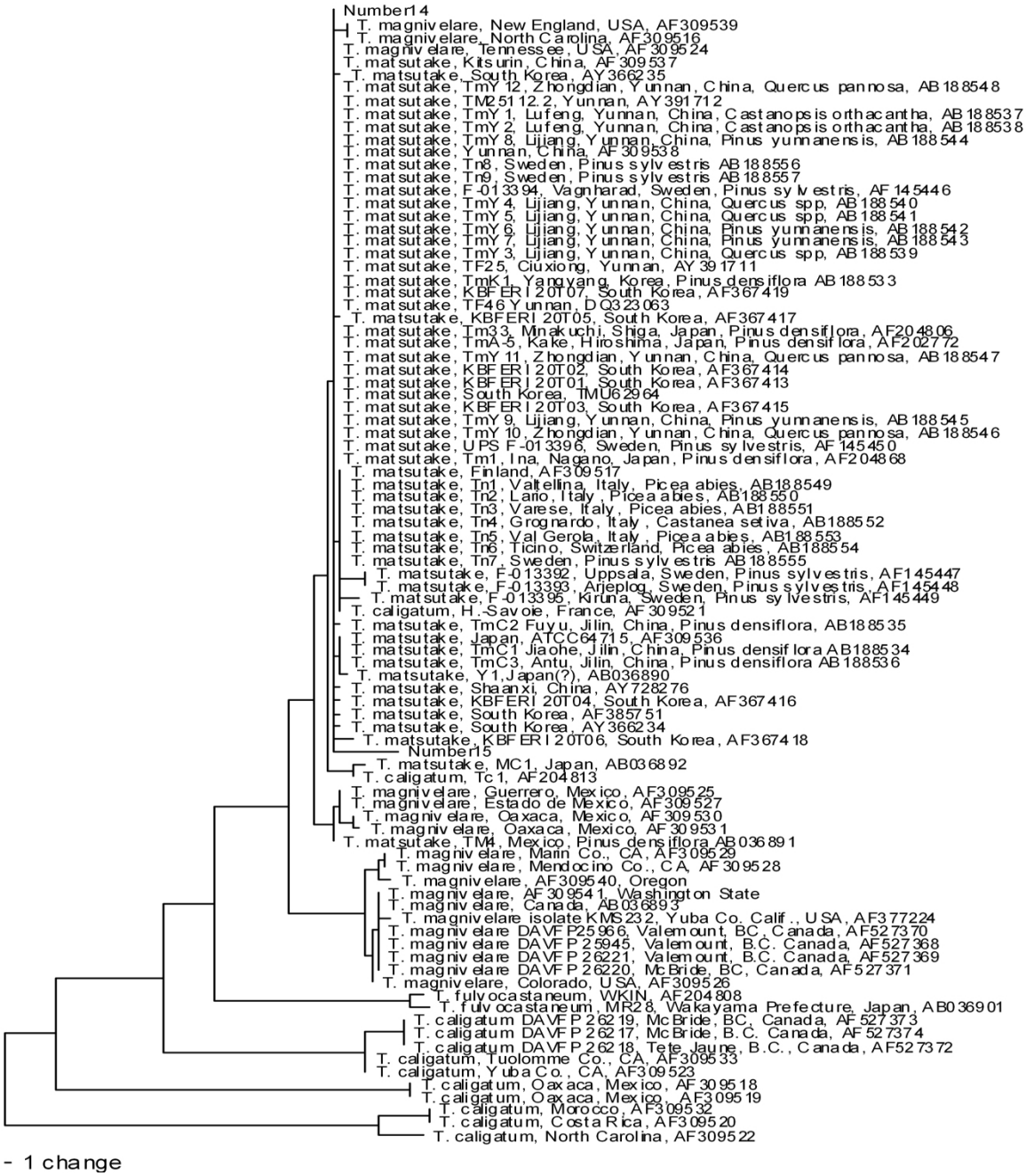

Fig. 4 Phylogenetic relationships among strains of *Tricholoma matsutake* and its close allies based on ITS sequences. Each entry contains the species name, followed by strain name, geographic location, likely associated plant host, and the GenBank accession number. Strains labeled "Number 14" and "Number 15" are two strains from Yunnan associated with *Pinus yunnensis* and *Quercus* spp respectively. Note the limited or no ITS sequence variation pattern among populations of *T. matsutake* from different geographic areas.

Genome Canada, Natural Science and Engineering Research Council (NSERC) of Canada, Premier's Research Excellence Award of Ontario, Canadian Foundation for Innovation, and Ontario Innovation Trust.

References:

[1] Amaranthus M P, Weigand J F, Abbott R. Managing high-elevation forests to produce American matsutake (*Tricholoma magnivelare*), high – quality timber, and nontimber forest products [J]. West J Appl, 1998, 13: 120 – 128.

[2] Arnolds E. Decline of ectomycorrhizal fungi in Europe [J]. Agriculture, Ecosystems and Environment, 1991, 35: 209 – 244.

[3] Arora D. Mushrooms demystified [M]. Berkeley, California: Ten Speed Press, 1986.

[4] Bergius N, Danell E. The Swedish matsutake (*Tricholoma nauseosum* syn. *T. matsutake*): Distribution, abundance and ecology [J]. Scan J Forest Res, 2000, 15: 318 – 325.

[5] CBOL Plant Barcode Working Group. A DNA barcode for land plants [J]. PNAS, 2009, 106: 12794 – 12797.

[6] Chapela I H, Garbelotto M. Phylogeography and evolution in matsutake and close allies inferred by analyses of ITS sequences and AFLPs [J]. Mycologia, 2004, 96: 730 – 741.

[7] Dunstan W A, Dell B, Malajczuk N, et al. Detection of the ectomycorrhizal fungus *Tricholoma matsutake* and some related species with specific ITS primers [J]. Mycoscience, 2000, 41: 33 – 37.

[8] Ebina T, Kubota T, Ogamo N, et al. Antitumor effect of a peptide – glucan preparation extracted from a mycelium of *Tricholoma matsutake* (S. Ito and Imai) Sing. Biotherapy, 2002, 16: 255 – 259.

[9] Gill W M, Lapeyrie F, Gomi T, et al. Tricholoma matsutake: An assessment of in situ and in vitro infection by observing cleared and stained whole roots [J]. Mycorrhiza, 1999, 9: 227 – 231.

[10] Hajibabaei M, Singer G A C, Hebert P D N, et al. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics [J]. Trends Genet, 2007, 23: 167 – 172.

[11] Hebert P D N, Ratnasingham S, deWaard J R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species [J]. Proc R Soc Lond B, 2003, 270: 96 – 99.

[12] Hoshi H, Yagi Y, Iijima H, et al. Isolation and characterization of a novel immunomodulatory R-glucan-protein complex from the mycelium of *Tricholoma matsutake* in basidiomycetes [J]. Journal of Agricultural and Food Chemistry, 2005, 53: 8948 – 8956.

[13] Imazeki R, Otani Y, Hongo T. Fungi of Japan [M]. Tokyo: Yama-Kei Publishers, 1988.

[14] James T Y, Kauff F, Schoch C, et al. Reconstructing the early evolution of the fungi using a six-gene phylogeny [J]. Nature, 2006, 443: 818 – 822.

[15] Moncalvo J M. Molecular systematics: Major fungal phylogenetic groups and fungal species concepts [M]. Xu J. Evolutionary Genetics of Fungi, Norfolk: Horizon Bioscience. U. K, 2005: 1 – 34.

[16] Murata H, Babasaki K, Saegusa T, et al. Traceability of asian matsutake, specialty mushrooms produced by the ectomycorrhizal basidiomycete *Tricholoma matsutake*, on the basis of retroelement-based DNA markers [J]. Applied and Environmental Microbiology, 2008, 74: 2023 – 2031.

[17] Seifert K A. Progress towards DNA barcoding of fungi [J]. Molecular Ecology Resources, 2009, 9(1): 83 – 89.

[18] Singer R. The Agaricales in Modern Taxonomy [M]. 2nd ed. Koeningstein: Koeltz Scientific Books, 1986.

[19] Stamets P. Growing gourmet and medicinal mushrooms [J]. Ten Speed Press, Berkeley, USA, 1993.

[20] Takama F, Ishii H, Muraki S. Flavor components in Japanese and Korean matsutake *Tricholoma matsutake* (S. Ito et Imai) Sing. and their change during storage [J]. Japanese Society for Food Science and Technology, 1984, 31: 14 – 18.

[21] Wang Y, Hall I R. Edible ectomycorrhizal mushrooms: challenges and achievements [J]. Canadian Journal of Botany, 2004, 82: 1063 – 1073.

[22] Wong E H K, Hanner R H. DNA barcoding detects market substitution in North American seafood [J]. Food Research International, 2008, 41: 828 – 837.

[23] Xu G B, Fu W J, Wei T Z, et al. Distribution and ecological characterization of *Tricholoma matsutake* in Changbai Mountain range in northeast China: a review [J]. Territory and Natural Resources Research, 2001, 2: 73 – 74.

[24] Xu J. Evolutionary Genetics of Fungi [R]. Horizon Bioscience, Norfolk, UK, 2005.

[25] Xu J. Microbial Population Genetics [R]. Horizon Bioscience, Norfolk, UK, 2010.

[26] Xu J, Cadorin M, Liang Y J, et al. DNA-based geographic typing of the gourmet mushroom *Tricholoma matsutake* traded in China [J]. Mycoscience, 2010, 51: 248 – 251.

[27] Xu J, Sha T, Li Y C, et al. Recombination and genetic differentiation among natural populations of the ectomycorrhizal mushroom *Tricholoma matsutake* from southwestern China [J]. Molecular Ecology, 2008, 17: 1238 – 1247.

[28] Xu J, Yoell H J, Anderson J B. An efficient protocol for isolating DNA from higher fungi [J]. Trends in Genetics, 1994, 10: 26 – 27.