Fungal Genetics and Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Regular Articles

Phylogeography and population genetic analyses reveal the speciation of the *Tuber indicum* complex

Peng Qiao[a,b,1], Wei Tian[c,1], Peigui Liu[b,*], Fuqiang Yu[d], Juan Chen[e], Xiaojuan Deng[f], Shanping Wan[g], Ran Wang[d], Yun Wang[h], Hongen Guo[a,*]

[a] *Shandong Institute of Sericulture, Shandong Academy of Agricultural Sciences, Yantai 264002, Shandong, PR China*
[b] *Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, Yunnan, PR China*
[c] *Agriculture College, Ludong University, Yantai 264025, Shandong, PR China*
[d] *Key Laboratory of Economic Plants and Biotechnology, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, PR China*
[e] *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, PR China*
[f] *College of Biological Science and Engineering, Beifang University of Nationalities, Yinchuan 750021, PR China*
[g] *College of Resource and Environment, Yunnan Agricultural University, Kunming 650100, PR China*
[h] *The New Zealand Institute for Plant and Food Research Limited, Christchurch 8140, New Zealand*

ARTICLE INFO

Keywords:
Phylogeography
Population genetics
ITS
SSR
Tuber indicum
Truffle

ABSTRACT

Tuber indicum is an ectomycorrhizal ascomycete that produces edible ascocarps. Based on a number of specimens with known exact origin, we investigate the speciation of the *Tuber indicum* complex in southwest China. Internal transcribed spacer (ITS) and simple sequence repeat (SSR) markers were used in the study. Phylogeography and population genetics analyses were combined to detect 31 wild populations of the *T. indicum* complex. Two distinct lineages, *Tuber* cf. *indicum* and *Tuber* cf. *himalayense*, were identified in the *T. indicum* complex that exhibited significant phylogeographic structures and genetic differentiation. The characteristics of haplotypes distributing along the river demonstrate that the diffusion and modern distribution pattern of species was influenced by river expansion. These findings are critical for the protection of the diversity of truffles in this region.

1. Introduction

The first Asian truffle, *Tuber indicum*, was described from a dried sample harvested in January 1892 near Mussooree (now Mussoorie) in India, in north-western Himalaya, at about 2000 m (Cooke and Massee, 1892). Later, the new species of Asian black truffles were discovered and reported, including *T. himalayense*, *T. sinense*, *T. pseudohimalayense*, *T. pseudoexcavatum*, *T. formosanum* and *T. sinoaestivum* (Zhang and Minter, 1988; Tao et al., 1989; Moreno et al., 1997; Wang et al., 1998; Hu, 1992; Zhang et al., 2012; Qiao et al., 2013). Although there have some studies that have discussed the phylogenetic relationships of Asian black truffles (*T. indicum*, *T. sinense*, *T. himalayense*, *T. pseudohimalayense* and *T. pseudoexcavatum*), classification and identification of these species has remained problematic (Chen et al., 2011; Chen and Liu, 2011).

Paolocci et al. (1997) analyzed the relationships between Chinese black truffles and Périgord truffles from the European market, and the results showed that these species were clearly split into different branches; the truffles from China were clearly divided into two branches that did not correspond to the various known species. Similar results and conclusions were obtained in subsequent studies, but the taxonomic status of these fungi is still controversial. Roux et al. (1999) suggested that the two branches corresponded to *T. indicum* and *T. himalayense*. This view was accepted by Zhang et al. (2005). However, Mabru et al. (2001) suggested that the black truffle from China might be a different manifestation of *T. indicum*. Based on the analysis of internal transcribed spacer (ITS) and β-tubulin sequences, Wang et al. (2006) accepted the point view of Mabru et al. (2001) and suggested that *T. sinense*, *T. himalayense*, and *T. pseudohimalayense* were synonyms of *T. indicum*. Based on morphological characteristics and the examination of nrDNA-LSU, ITS, and β-tubulin sequences in separate and joint analyses, Chen and Liu (2011) confirmed that *T. pseudoexcavatum* was a synonym of *T. pseudohimalayense*, while *T. sinense* was a synonym of *T. indicum*. Hence, the *T. indicum* complex was divided into two branches in a phylogenetic tree, which were considered to represent two independent phylogenetic species. These discrepancies might originate from the high morphological and molecular variability within the *T. indicum* complex and the insufficient number of specimens

* Corresponding authors.
E-mail addresses: pgliu@mail.kib.ac.cn (P. Liu), guohongen163@163.com (H. Guo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.fgb.2018.02.001
Received 13 September 2017; Received in revised form 5 February 2018; Accepted 11 February 2018

studied (Chen et al., 2011).

Because it is too old, it is not possible to extract DNA from *T. indicum* type specimens. Furthermore, no additional specimens were subsequently used for the comparative analysis of molecular phylogeny. So the taxonomic and phylogenetic status and the biogeography of the *T. indicum* complex from China have not been studied carefully.

All known *Tuber* species are ectomycorrhizal fungi that exhibit symbiotic relationships with many plants in northern temperate forests, such as Fagaceae, Corylaceae, Betulaceae, Salicaceae, and Pinaceae species. There is gene transfer and coevolution between fungi and host plants (Morran et al., 2011.). So the diversity of host plants played a key role in the diversity of *Tuber* species (Liu et al., 2011). Previous studies have identified the dramatic topographic changes triggered by the rapid uplifting of the Qinghai-Tibetan Plateau (QTP) and the Quaternary climate oscillations as two key evolutionary factors that drive the patterns of genetic diversity and genetic structures of plants in the QTP and its adjacent regions (Qiu et al., 2011). As a widely distributed species in Southwest China, the speciation of the *T. indicum* complex may be influenced by the events. Recently, Feng et al. (2016) reconstruct the demographic history of this species complex by using species distribution modeling and proposed the role of Quaternary climate oscillations in *T. indicum* genetic structure.

In this study, we sampled extensively from southwest China, ITS and simple sequence repeat (SSR) markers were used in this study. The objectives were: to quantify the genetic differences between the two lineages of the *T. indicum* complex; to unravel the pattern of distribution of genetic diversity within and among natural populations; and to discuss the demographic history of the complex and relate it to geological and climatic events.

2. Materials and methods

2.1. Sample collection and DNA isolation

A total of 476 ascocarps were sampled from 31 populations of the *T. indicum* complex distributed in different counties of Yunnan and Sichuan Provinces during 2010–2012 (Table S1). The collecting sites with the elevations ranging from 1258 to 2469 m covered the known areas of this species complex in southwest China. All the ascocarps were first macroscopically and microscopically checked and identified, as described by Chen et al. (2011), and were then cut into slices, dried in silica gel and stored at −70 °C before DNA extraction. Voucher specimens of each sampled individual were deposited at the Herbarium of the Kunming Institute of Botany, Chinese Academy of Sciences (KUNHKAS). Total genomic DNA was isolated from approx. 0.3 g of each ascocarp using the CTAB method (Doyle and Doyle, 1987).

2.2. PCR amplification and sequence analysis of the ITS

PCR amplification of the ITS region was performed using the fungal-specific ITS1-F (Gardes and Bruns, 1993) – ITS4 (White, 1990) primer pair. PCR amplification was performed in 25 μl reaction mixtures containing 2.5 μl of 10× PCR buffer [100 mM Tris-HCl (pH 8.3); 500 mM KCl; 15 mM $MgCl_2$], 2.0 μl of a dNTP mixture [each 2.5 mM], 0.5 μl of each primer [10 μM], 0.2 μl of Taq polymerase [5 U μl^{-1}] (Takara, Dalian, China), 1.0 μl of template DNA [30–50 ng genomic DNA], and finally, 18.3 μl of distilled deionized water. PCR was performed in a GeneAmp PCR System 9700 thermal cycler (Perkin Elmer, Foster City, CA, USA) with the following profile: initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 1 min, 58 °C for 1 min and 40 s, 72 °C for 2 min, and a final step at 72 °C for 10 min. The ITS amplicons were either directly sequenced or purified using the JetQuick PCR purification kit (Sangon, Shanghai, China) and then sequenced on an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA). All newly acquired sequences have been submitted to GenBank with accession numbers from MG881220-MG881695.

The chromatograms of sequences were carefully checked with Chromas v2.6.5 (http://www.technelysium.com.au/wp/). ITS sequences were aligned in CLUSTAL X v2.0 (Larkin et al., 2007) and subsequently manually adjusted on Bioedit v7.2.6.1 (Hall, 1999) while necessary. All indels were treated as single mutation events. Relationships among haplotypes were also visualized through NJ clustering using MEGA v6.0 (Tamura et al., 2011), with *T. melanosporum* (GenBank: KM659874.1) chosen as outgroup. Molecular diversity indices, including the number of haplotypes, haplotype diversity (H_d) and nucleotide diversity (π), were estimated using DNASP v5.10 (Librado and Rozas, 2009). A median-joining haplotype network was constructed using the program NETWORK v4 (Polzin and Daneshmand, 2003). The spatial genetic structuring of the haplotypes was analyzed using a spatial analysis of molecular variance analysis (SAMOVA v1.0; Dupanloup et al., 2002). Analysis of molecular variance (AMOVA) was performed to partition the variation within and among defined lineages and populations using Arlequin v3.5 (Excoffier and Lischer, 2010). The existence of phylogeographic structure was tested following Pons and Petit (Pons and Petit, 1996). Two estimates of population diversity (H_S, H_T) and two of differentiation (G_{ST}, N_{ST}) were obtained using PermutCpSSR v2.0 (http://www.pierroton.inra.fr/genetics/labo/Software/) with 1000 permutations.

2.3. Isolation and analysis of SSR loci

T. indicum SSR loci were isolated from the whole-genome sequence of *T. melanosporum* (Martin et al., 2010) using SSRIT software (Temnykh et al., 2001). *T. indicum* SSR primers were designed using Primer Premier V5.0 (http://www.premierbiosoft.com/).

All SSR fragments were amplified with the primer pairs (Table S3) and preliminarily tested under the PCR conditions and cycling parameters reported in Rubini et al. (2004). The following PCR cycling conditions were used: a denaturation step at 94 °C for 2 min and 30 s, followed by 35 cycles of 30 s at 94 °C, 30 s at the appropriate annealing temperature (Table 2) and 30 s at 72 °C, with a final extension at 72 °C for 20 min (Riccioni et al., 2008). The PCR products were separated via capillary electrophoresis in an ABI 3730 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). Sizing of fragments was performed using GeneMapper v3.2 (Applied Biosystems).

For all ten loci, MicroChecker v2.2.3 (Oosterhout et al., 2004) was used to detect the presence of null alleles and genotyping errors such as large allele dropout or stuttering using 1000 randomizations. Tests for departures from Hardy-Weinberg equilibrium (HWE) were performed using Genepop V4.2 (http://www.genepop.curtin.edu.au/). A two-locus linkage disequilibrium (LD) analysis of SSR loci was conducted using FSTAT v2.9.3 (Goudet, 2001). Significance levels were adjusted using the sequential Bonferroni correction for multiple comparisons (Rice, 1989).

Genetic diversity indices (number of total alleles, N_A; observed heterozygosity, H_O; expected heterozygosity, H_E) were calculated with GenALEx v6.5 (Peakall and Smouse, 2012) for each population and locus.

We analyzed total genetic variation within and among populations by AMOVA with 1000 permutations (Excoffier et al., 1992) and performed total differentiation F_{ST} (Weir and Cockerham, 1984) using Arlequin v3.5 (Excoffier and Lischer, 2010). In addition, population differentiation F_{ST} was estimated for each locus across all populations using GenALEx v6.5. Standardized genetic differentiation (G'_{ST}) (Hedrick, 2005) was calculated with GenALEx v6.5. G'_{ST} is a more suitable measure of differentiation than traditional measures (e.g., F_{ST} and G_{ST}) for highly polymorphic markers, such as microsatellites (Hedrick, 2005).

The existence of genetic structure was also tested with a Bayesian approach using the program STRUCTURE v2.3 (Pritchard et al., 2000). The program was employed to estimate the likelihood of a given number of clusters (K). The simulation was run with the number of