A global meta-analysis of *Tuber* ITS rDNA sequences: species diversity, host associations and long-distance dispersal

GREGORY M. BONITO,* ANDRII P. GRYGANSKYI,* JAMES M. TRAPPE† and RYTAS VILGALYS*

*Department of Biology, Duke University, Durham, NC 27708-0338, USA, †Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331-5752, USA

Abstract

Truffles (*Tuber*) are ectomycorrhizal fungi characterized by hypogeous fruitbodies. Their biodiversity, host associations and geographical distributions are not well documented. ITS rDNA sequences of *Tuber* are commonly recovered from molecular surveys of fungal communities, but most remain insufficiently identified making it difficult to determine whether these sequences represent conspecific or novel taxa. In this meta-analysis, over 2000 insufficiently identified *Tuber* sequences from 76 independent studies were analysed within a phylogenetic framework. Species ranges, host associates, geographical distributions and intra- and interspecific ITS variability were assessed. Over 99% of the insufficiently identified *Tuber* sequences grouped within clades composed of species with little culinary value (Maculatum, Puberulum and Rufum). Sixty-four novel phylotypes were distinguished including 36 known only from ectomycorrhizae or soil. Most species of *Tuber* showed 1-3% intraspecific ITS variability and >4% interspecific ITS sequence variation. We found 123 distinct phylotypes based on 96% ITS sequence similarity and estimated that *Tuber* contains a minimum of 180 species. Based on this meta-analysis, species in Excavatum, Maculatum and Rufum clades exhibit preference for angiosperm hosts, whereas those in the Gibbosum clade are preferential towards gymnosperms. Sixteen *Tuber* species (>13% of the known diversity) have putatively been introduced to continents or islands outside their native range.

Keywords: biodiversity, biogeography, hypogeous fungi, invasive biology, ITS rDNA, phylogeny, *Tuber*

Received 1 April 2010; revision received 18 June 2010; accepted 2 July 2010

Introduction

Truffles are icons of the fungal world because of the aromatic hypogeous fruitbodies of some species in the genus *Tuber*. The ecology and host associations of a few commercialized European *Tuber* spp. have been intensively studied (Murat et al. 2004; Wedén et al. 2005; Paolocci et al. 2006; Riccioni et al. 2008); however, at the global scale, little is known about their overall species diversity or ecology. Albert B. Frank first recognized the mycorrhizal symbiosis in 1885 when he showed that fungi in the truffle-forming genus *Tuber* were found growing on the roots of living plants (Frank 2005). Since then, it has been established that *Tuber* is an obligate mycorrhizal lineage, unable to complete its life cycle apart from a host and that *Tuber* spp. are important to the nutrition and drought tolerance of host plants (Nardini et al. 2000; Bradshaw 2005; Nuñez et al. 2009a). *Tuber* spp. form ectomycorrhizae with a broad diversity of gymnosperm and angiosperm hosts in a variety of habitats including subtropical cloud forests, temperate forests, boreal forests, floodplains, tree nurseries, restoration sites and Mediterranean woodlands (Ceruti et al. 2003; Bidartondo et al. 2004; Izzo et al. 2005; Menkis et al. 2005; Bergemann & Garbelotto 2006; Frank et al. 2006a; Ishida et al. 2007; Hrynkiewicz et al. 2008; Krapa et al. 2008; Leski et al. 2008; Morris et al. 2008; Taylor et al. 2008; Southworth et al. 2009; Bulman et al. 2010).
Since the early 1990s, ITS rDNA sequence data have been used to analyse the composition and dynamics of ectomycorrhizal communities (Gardea-Torresdey et al. 1991; Horton & Bruns 2001). The rate of 'species' discovery resulting from molecular community ecology studies now outpaces that of modern taxonomy by orders of magnitude, and this trend is expected to increase exponentially as next-generation sequencing technology becomes widely available (Hibbett et al. 2009). Although ITS is a robust marker for identifying fungi and for discriminating most fungal species, species-level determinations are questionable unless the reference sequence can be verified from well-preserved publicly available specimen. In the case of Tuber, it is currently difficult to identify ITS sequences using BLAST because of species misidentifications, mislabelling of specimen packets, nomenclatural errors in the public database and because of the quantity of insufficiently identified submissions (Trappe and Halász et al. 2005; Lotti et al. 2007). It is common for sequences to be submitted to GenBank as 'unidentified' when reference taxa are lacking. These have been termed insufficiently identified sequences to distinguish them from fully identified sequences (Nilsson et al. 2006). The number of insufficiently identified sequences of Tuber in GenBank (1924 on April 1, 2010) places the genus within the top 10 insufficiently identified mycorrhizal genera in GenBank (Ryberg et al. 2008). Greater insight into Tuber's species diversity, phylogeny, ecology and distribution are expected to emerge from the synthesis of these data. We hypothesized that many insufficiently identified Tuber sequences are conspecific but that many also represent novel taxa.

The invasive biology of Tuber is of interest, because some of the economically important species are being intentionally introduced into ecosystems around the world (Hall et al. 2007; Bonito 2009a), yet the phenomena of human-mediated long-distance dispersal of mycorrhizal fungi is not well understood. Vellinga et al. (2009) have proposed a conceptual model of ectomycorrhizal species invasion that involves four stages: (i) transport of ectomycorrhizal propagules to a novel location; (ii) establishment of the ectomycorrhizal species within the landscape; (iii) spread beyond the point of introduction; and, (iv) ecological impacts, such as the displacement of native species or alterations to biogeochemical processes. However, discerning native ranges from areas of introduction is challenging, given the cryptic nature and lack of available biogeographical and natural history data for most ectomycorrhizal fungi (Vellinga et al. 2009). Molecular approaches have largely superseded morphological approaches owing to their increased sensitivity for assessing evolutionary units, fungal biogeography and introduction events (Stukenbrock et al. 2006; Hosaka et al. 2008; Matheny et al. 2009; Pringle et al. 2009). In the event of an introduction, founder populations are expected to exhibit less genetic variability than source populations. Ectomycorrhizal fungi that are host-generalists are hypothesized to have higher rates of establishment and spread because they have a greater likelihood of finding a suitable host (Vellinga et al. 2009). If hypogeous fungi are constrained by mycophagy for spore dispersal, their biogeography and invasive ecology may differ from that of epigeous fungi adapted for wind dispersal (Trappe and Claridge 2005; Hosaka et al. 2008; Nuñez et al. 2009).

Current research aimed at resolving phylogenetic relationships within Tuber has resulted in DNA sequencing from a high proportion (~70%) of known species in the genus and to the discovery of many new Tuber species and lineages (Guevara et al. 2008; Bonito 2009b; Bonito et al. 2010). These data were used together with >2000 ITS sequences from insufficiently identified Tuber spp. generated from herbarium collections, fresh fruitbodies, mycorrhizae and soil clones in order to provide the most comprehensive molecular assessment of global Tuber biodiversity to date.

Materials and methods

Material studied

ITS sequence data were generated from ~270 collections of identified and unidentified truffles, including 14 type specimens (i.e. holotypes, isotypes and paratypes), made available from the Oregon State University Mycological Collections (OSC), the National Fungus Collections (BPI), Harvard University's Farlow Herbarium (FH), and the Herbarium of Università di Bologna, Italy (BOLO). Fresh specimens collected 2006-2010 in Europe, Asia and North America were also sequenced and are available from the Duke University Herbarium (DUK). Additional ITS sequences from verifiably identified Tuber species were downloaded from GenBank and combined with our sequence database to create a Tuber ITS phylogeny to illustrate the phylogenetic diversity and major clades in Tuber.

The genus search tool in emerencia (http://www.emerencia.org/) was used to retrieve insufficiently identified ITS sequences from GenBank, whose pairwise similarity was most similar to identified Tuber species (Nilsson et al. 2005; Ryberg et al. 2009). GenBank numbers for sequences and citations for the studies in which they were generated are compiled and available as supporting information (Table S1). Metadata accompanying individual sequences (including geographical origin and host) were compiled when available.
**Molecular methods**

DNA from truffle ascomata was extracted by the CTAB miniprep (Gardes & Bruns 1993). For DNA extraction, global tissue was ground (dried or in CTAB) in sterile sand and large cubic zirconium beads in a Mini Beadbeater for 1-2 min (Biospec Products, Bartlesville, OK). The internal transcribed spacer region (ITS) was amplified with the primer set ITS5-ITS4 (White et al. 1990). Amplified fragments were viewed through agarose gel electrophoresis, cleaned with Qiagen Quick-Clean columns and sequenced with Big Dye chemistry v.3.1 (Applied Biosystems). DNA sequences were determined on an ABI3700 DNA sequencer (Applied Biosystems).

**Sequence analyses**

Generated DNA sequences were viewed and manually edited in Sequencher 4.0 (Gene Codes, Ann Arbor, MI). Sequence alignments were performed in MUSCLE (Edgar 2004). Ambiguous regions were excluded in Mesquite 2.5 (Maddison & Maddison 2009). Outgroup selection was based on previous phylogenetic studies of Tuberaceae (O’Donnell et al. 1997; Bonito 2009b). Insufficiently identified sequences were assigned to a specific Tuber clade based on initial parsimony analyses conducted in PAUP* 4d106 (Swofford 2002).

The data set of compiled ITS sequences was filtered by removing sequences of poor quality or short length as well as sequences not belonging to Tuber based on BLAST (Altschul et al. 1997). In a few cases, we excluded sequences that were determined to be chimeras based on independent BLASTing of the ITS1 and ITS2 regions. Redundant sequences and those with minor variation were removed a priori by assembling total ITS sequences into 98% similarity clusters with the dirty data assembly algorithm in Sequencher 4.0 (Gene Codes). Sequences comprising each 98% similarity cluster were recorded (Table S2, supporting information), and only one representative sequence from each was included in the final analyses.

The ITS1 region is too diverse to align unambiguously across the complete Tuber genus. To improve phylogenetic resolution on the placement of sequences from insufficiently identified Tuber, separate individual alignments and analyses were performed for each of the three Tuber clades that contained the majority of unidentified sequences. These clade designations are supported by multigene phylogenetic reconstructions of the Tuberaceae based on rDNA, elongation factor 1 alpha, and RNA polymerase 2 (Bonito 2009b). Appropriate models of nucleotide substitution were selected in PAUP* (Swofford 2002) by Akaike information criterion, penalizing more complex models by one likelihood unit per additional free parameter. Phylogenetic analyses were conducted with maximum likelihood (ML) in PAUP* (Swofford 2002) and Bayesian inference (BI) with MrBayes (Huelsenbeck & Ronquist 2001). Maximum likelihood bootstrap support based on 1000 bootstrap replicates was assessed with RAxML (Stamatakis et al. 2008) through the CIPRES web portal (http://www.phyl.org).

We also analysed 1622 putative Tuber pyrosequences accessioned as non-redundant (99% sequence identity) examples from Quercus ectomycorrhizae (Jumpponen et al. 2010). Because these were short sequences (<250) from a single study, they were analysed separately from the rest of our data set. We followed the same methods described earlier except that singletons remaining after clustering at 98% similarity were discarded to avoid possible sequencing artefacts. We analysed the remaining sequences by performing a BLAST search against our local Tuber database and discarding sequences with a bit score <150. The phylogenetic placement of the remaining 1548 sequences was assessed as described earlier.

To determine an appropriate phylotype definition for Tuber, levels of intraspecific and interspecific ITS variation were assessed for 20 Tuber species including all commercialized species and representatives from the nine major clades. Values of intraspecific and interspecific ITS variation were assessed by aligning pairs of species (multiple sequences from the species of interest and its closest sister taxon) in MUSCLE (Edgar 2004). Alignments were then manually edited, but no regions were excluded and uncorrected P values resulting from these ITS alignments were calculated in PAUP* (Swofford 2002). Based on these analyses, we have defined Tuber phylotypes a posteriori as those having at least 96% ITS sequence similarity. We consider this threshold as a species approximation. In some cases, this phylotype definition may lump morphological or ecological species. Geographical origin of sequences was documented to assess the distribution of Tuber species. Putative introduction events were diagnosed either as phylotypes occurring outside of Tuber’s native range, such as the Southern hemisphere or as phylotypes showing large continental disjuncts and having <1% ITS variation, a level often exceeded at the local scale in native populations (Smith et al. 2007b; Bonito 2009b; Southworth et al. 2009; Jumpponen et al. 2010). Estimates of global species diversity for Tuber was calculated with incidence-based estimators (i.e. ICE, Chao2, Jackknife2) in EstimateS (Colwell 2005). Sequences generated in this study have been annotated and deposited in GenBank (HM485330-HM485429), and collection numbers and herbaria where these sporocarps are accessioned are provided for future taxonomic work (Table S5, supporting information).
Results

Mining ITS sequence metadata from GenBank with enversico yielded 230 sequences (>300 bp) from 75 studies (Table S1, supporting information) and another 1622 short sequences (<250 bp) from a single pyrosequencing study (Jumpponen et al. 2010). ITS sequences from reference taxa (n = 74) and unidentified fruitbodies (n = 196) were combined with GenBank sequences for a total of 2122 sequences. By removing all sequences of poor quality, this data set was reduced to 1950 sequences. Assembling these into clusters of 98% sequence similarity resulted in 185 unique ITS types.

Phylogenetic analyses of ITS rDNA from Tuber reference taxa distinguished nine major clades (Fig. 1). The majority (98%) of unique unidentified Tuber ITS sequences grouped within three of the nine Tuber clades: Puberulum (44%), Maculatum (34%) and Rufum (20%). Of the nine sequences outside these three clades, two from Epipactis microphylla mycorrhizae were identified as T. excavatum group A (Excavatum clade); two from Epipactis microphylla and Cephalanthera damasonium mycorrhizae were identified as T. aestivum (Aestivum clade); two from angiosperm mycorrhizae were identified as T. magnatum (Aestivum clade); one from a Pinus sabiiana ectomycorrhiza and was identified as T. gibbosum (Gibbosum clade) (Fig. 1). Two sequences from an artificially established truffle orchard in North America grouped with T. melanosporum.

Intraspecific ITS variation differed among the species examined and was below 3% for all species except T. aestivum (<3.7%) (Table 1). Tuber oreonense, T. puberulum and T. castellanioides had the lowest levels of intraspecific ITS variation (0.2%). At least 4% interspecific ITS variation occurred between pairs of sister taxa examined, except for the morphological species T. excavatum, T. gemmatai, and T. candelum, which appear to be species complexes. Tuber species also differed significantly in ITS length (Table 1). Tuber brumale had the longest ITS (859 bp), while T. bellisorum had the shortest (465 bp).

Based on our phylotype definition, 25 putatively undescribed species can be added to the Puberulum clade (Fig. 2), including 15 represented only by mycorrhiza or soil clone sequences (Table S3, supporting information). Plant hosts varied considerably within the Puberulum clade. Some phylotypes are only recorded from angiosperms and others only with gymnosperms (Table 2). Four phylotypes (T. borchi, T. menseri nom. prov., Tuber sp.19, and Tuber sp. 24) in this clade were recorded as ectomycorrhiza on both angiosperm and gymnosperm hosts. Asia, Europe and North America each contain endemic Tuber species belonging to this clade. Tuber menseri nom. prov., Tuber sp.19, and T. borchi (cultivated) have been documented in New Zealand (Table 3).

Twelve putatively undescribed species can be ascribed to the Maculatum clade (Fig. 3). Five of these are represented only by mycorrhiza or soil clone sequences (Table S3, supporting information). Only one sequence (AM418469) in the Maculatum clade came from a gymnosperm host (Pinus nigra var. australis) (D. Redeker personal communication). One species (T. rupiformis) appears to occur in Europe, North America and New Zealand and was recovered in 14 studies (Table 3).

Twenty-five putatively undescribed species belong to the Rufum clade (Fig. 4), including 13 represented only by mycorrhiza or soil clone sequences (Table S3, supporting information). There was only one sequence (FJ789634) in the Rufum clade that came from a gymnosperm host (Pinus jeffreyi) (D. Southworth, personal communication). Species in the Rufum clade are distributed across Asia, Europe and North America, and one species, Tuber sp.57, has been documented in New Zealand (Table 3).

Analysis of pyrosequencing data from Quercus ectomycorrhizae (Jumpponen et al. 2010) indicated that 1517 of the 1548 Tuber sequences already had counterparts in our data set. These sequences were comprised of 11 phylotypes within the Rufum and Maculatum clades, three of which are novel to this study (Table S6, supporting information). The three most abundant sequences belonged to Tuber lyonii (n = 520), Tuber sp. 40 (n = 494) and Tuber sp. 36 (n = 342).

In total, 123 distinct phylotypes (based on 96% ITS sequence similarity) are presented in this study, including 59 singletons and 24 doubletons. Rarefaction curves did not plateau and incidence-based coverage estimators predict a minimum of 180 (Jackknife 2) to 230 (ICE) Tuber species worldwide (Fig. 5). We provide molecular evidence for the introduction of 10 Tuber species outside of their native ranges (Table 3). These consist of commercial and non-commercial species belonging to five major Tuber clades.

Discussion

Phylotype definition

The ITS rDNA region is considered a DNA barcode for fungi (Nilsson et al. 2008). Though not appropriate for all species, our data demonstrate it is reliable for discerning species in Tuber. For the species examined, intraspecific ITS variation was typically <3% and interspecific variation was 4% or greater. Therefore, a phylotype definition of 96% ITS sequence similarity appears to be a valid species approximation for Tuber. Based on this criterion, which is less stringent than the 97% phylotype threshold commonly used in fungal community studies (Smith et al. 2007b; Peay et al. 2008;
Fig. 1 Tuber ITS phylogeny showing the nine major clades in the genus. Taxa were chosen to demonstrate the phylogenetic breadth of Tuber and to represent as many described species as possible. Insufficiently identified sequences belonging in the Aestivum, Excavatum and Gibbosum clades have been included, while those belonging to the Maculatum, Puberulum and Rufum clades are shown in Figs 2–4. These analyses are based on ITS nuclear rRNA (323 included characters) and a GTR+G+I submodel of nucleotide substitution with three substitution rate classes. Maximum likelihood (ML) bootstrap values are shown above branches and posterior probabilities based on Bayesian inference (BI) below. Values of >70 for ML and ≥99 for BI are considered significant. Asterisks (*) denote support values of 100, when space was limited. Sequences are labelled with Latin binomials, GenBank accession or collection number and geographical origin. Other Tuberaceae genera were chosen as outgroups.
Hughes et al. 2009), the morphological species *T. excavatum*, *T. gemmellii* and *T. candidum* are actually species complexes. *Tuber gemmellii* is a rare species that has been described as a separate genus (Loculotuber) because of its morphology of chambers lined by asci (Alvarez et al. 1992). ITS places *T. gemmellii* basal in the *Tuber* lineage and distinguishes two distinct phylotypes (Fig. 1). *Tuber excavatum* is often considered a single species, yet sequences cluster into three distinct phylotypes (Table S4, supporting information). *Tuber candidum* contains up to three phylogenetic species, which cannot be resolved by ITS alone (Fig. 4). The highest level of intraspecific ITS variation (c.3.7%) was found for *Tuber aestivum* (=*T. uncinatum*). Mello et al. (2002) and Wedén et al. (2005) also report intraspecific variation of over 3% in *T. aestivum*. Morphological studies indicate *T. aestivum* occurs throughout Europe and Asia and has been recorded from Morocco in North Africa (Bucholtz 1901; Ceruti et al. 2003; Song et al. 2005). Although this economically important taxon is among the most studied *Tuber* species, there is still debate over whether this is a species complex (Mello et al. 2002; Paolocci et al. 2004; Wedén et al. 2004, 2005). Additional molecular studies will be needed to resolve the issue of cryptic species in these *Tuber* species complexes.

### Distribution of *Tuber* species and introductions into non-native habitats

Relying on morphological and phylogenetic species concepts Vellinga et al. (2009) conclude that at least 200 species of ectomycorrhizal fungi (including eight *Tuber* species) have been introduced into novel habitats. In most cases, introduced fungi were associated with non-native host plant species and were likely introduced in conjunction with a plant host (either on roots or in the accompanying soil). With the inclusion of our molecular data, the number of *Tuber* species introduced to novel habitats is 16 (Table 3).

### Table 1 Intraspecific and interspecific ITS variation of reference *Tuber* species

<table>
<thead>
<tr>
<th>Species</th>
<th>Number (n)</th>
<th>Maximum ITS length (bp)</th>
<th>Intraspecific variation (%)</th>
<th>Interspecific variation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tuber aestivum</em></td>
<td>104</td>
<td>653</td>
<td>&lt;3.7</td>
<td>&gt;20.9</td>
</tr>
<tr>
<td><em>Tuber borchii</em></td>
<td>72</td>
<td>501</td>
<td>&lt;2.2</td>
<td>&gt;5.1</td>
</tr>
<tr>
<td><em>Tuber brunneol</em></td>
<td>9</td>
<td>859</td>
<td>&lt;0.5</td>
<td>&gt;25.4</td>
</tr>
<tr>
<td><em>Tuber castellannai</em></td>
<td>3</td>
<td>474</td>
<td>&lt;0.2</td>
<td>&gt;6.2</td>
</tr>
<tr>
<td><em>Tuber excavatum</em></td>
<td>9</td>
<td>622</td>
<td>&lt;14.7</td>
<td>&gt;9.2</td>
</tr>
<tr>
<td><em>Tuber gemellii</em></td>
<td>3</td>
<td>666</td>
<td>&lt;16.7</td>
<td>&gt;27.9</td>
</tr>
<tr>
<td><em>Tuber gibboum</em></td>
<td>15</td>
<td>472</td>
<td>&lt;0.6</td>
<td>&gt;3.9</td>
</tr>
<tr>
<td><em>Tuber lycii</em></td>
<td>26</td>
<td>537</td>
<td>&lt;2.9</td>
<td>&gt;5.3</td>
</tr>
<tr>
<td><em>Tuber indicum</em> A</td>
<td>57</td>
<td>543</td>
<td>&lt;2.5</td>
<td>&gt;7.0</td>
</tr>
<tr>
<td><em>Tuber indicum</em> B</td>
<td>30</td>
<td>541</td>
<td>&lt;2.3</td>
<td>&gt;7.0</td>
</tr>
<tr>
<td><em>Tuber macroporum</em></td>
<td>4</td>
<td>587</td>
<td>&lt;1.7</td>
<td>&gt;9.1</td>
</tr>
<tr>
<td><em>Tuber nigromatum</em></td>
<td>64</td>
<td>538</td>
<td>&lt;0.4</td>
<td>&gt;30.0</td>
</tr>
<tr>
<td><em>Tuber mesieri</em> nom prov.</td>
<td>25</td>
<td>593</td>
<td>&lt;2.3</td>
<td>&gt;3.9</td>
</tr>
<tr>
<td><em>Tuber mesieri</em> nom prov.</td>
<td>171</td>
<td>622</td>
<td>&lt;1.0</td>
<td>&gt;20.9</td>
</tr>
<tr>
<td><em>Tuber oerognos</em></td>
<td>34</td>
<td>473</td>
<td>&lt;0.2</td>
<td>&gt;8.9</td>
</tr>
<tr>
<td><em>Tuber palustrum</em></td>
<td>4</td>
<td>488</td>
<td>&lt;0.2</td>
<td>&gt;5.1</td>
</tr>
<tr>
<td><em>Tuber raepedorum</em></td>
<td>42</td>
<td>643</td>
<td>&lt;1.6</td>
<td>&gt;4.6</td>
</tr>
<tr>
<td><em>Tuber wheelei</em></td>
<td>10</td>
<td>472</td>
<td>&lt;2.5</td>
<td>&gt;4.8</td>
</tr>
</tbody>
</table>

*Based on the number of basepairs between the end of the SSU-CATTA-motif and the beginning 28s LSU-TAGGGT motif (if present).

1Represent species complexes (see Table S4, supporting information and Fig. 1).

1Sequences accessioned as *T. formosanum* appear to be nested within *T. indicum* B, so are considered as *T. indicum* B in these calculations.

© 2010 Blackwell Publishing Ltd
Table 2 The 10 most abundant insufficiently identified *Tuber* taxa, their geographical range and host plant mycorrhiza from which they have been sequenced. Species calls are based upon a 98% sequence similarity threshold across the complete ITS rDNA region and thus are likely to underestimate host and geographical ranges. Pyrosequence data have not been included in abundance scores. See (Table S2, supporting information) for a complete list of species and accompanying metadata on host, source, and accession number for each sequence.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Abundance</th>
<th>Species</th>
<th>Geographical source</th>
<th>Host associates</th>
<th>Clade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td><em>Tuber rypacolorum</em></td>
<td>Finland, Netherlands, Germany, Hungary, Poland, Estonia, Hungary, US, Oregon, US, California, US, Texas, US, New York, Canada, Quebec, New Zealand</td>
<td><em>Epipactis helleverine</em>, <em>Salix caprea</em>, <em>Populus alba</em>, <em>Tilia cordata</em>, <em>Alnus</em>, <em>Cephalanthera damasonium</em></td>
<td>Maculatum</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td><em>Tuber menzeri</em> nom. prov.</td>
<td>US, Oregon, US, Washington, Canada, Quebec, Lithuania, Finland, Poland, Netherlands, New Zealand</td>
<td><em>Quercus garryana</em>, <em>Salix caprea</em>, <em>Betula pendula</em>, <em>Salix alba</em>, <em>Populus sp.</em>, <em>Tilia cordata</em>, <em>Pseudotsuga menziesii</em>, <em>Picea abies</em></td>
<td>Puberulum</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td><em>Tuber sp.19</em></td>
<td>Austria, Germany, Lithuania, Sweden, Quebec, US, Nebraska, US, California, New Zealand</td>
<td><em>Epipactis dunnensis</em>, <em>Populus tremula</em>, <em>Betula pendula</em>, <em>Pinus sylvestris</em>, <em>Picea abies</em>, <em>Pseudotsuga menziesii</em></td>
<td>Puberulum</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td><em>Tuber laurii</em> nom.prov.</td>
<td>US, California, US, Oregon, US, California, US, Oregon</td>
<td><em>Epipactis helleverine</em>, <em>Quercus visiizensi</em>, <em>Notholithocarpus densiiflorus</em>, <em>Quercus garryana</em>, <em>Quercus visiizensi</em>, <em>Quercus douglasii</em>, <em>Quercus sp.</em></td>
<td>Maculatum</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td><em>Tuber whetstonei</em></td>
<td>US, California, US, Oregon, US, California, US, Georgia, US, California</td>
<td><em>Quercus visiizensi</em>, <em>Quercus douglasii</em>, <em>Quercus sp.</em>, <em>Carya illinoinensis</em>, <em>Epipactis helleverine</em>, <em>Pinus pinaster</em></td>
<td>Maculatum</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td><em>Tuber sp.36</em></td>
<td>US, Missouri, US, Georgia, US, California</td>
<td><em>Epipactis helleverine</em>, <em>Pinus pinaster</em></td>
<td>Puberulum</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td><em>Tuber borchii</em> (anamorphic)</td>
<td>Austria, Spain, Italy</td>
<td><em>Castanea dentata</em>, <em>Quercus sp.</em>, <em>Quercus palustris</em></td>
<td>Puberulum</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td><em>Tuber sp.16</em></td>
<td>US, New York, US, Michigan, US, California, Canada, Quebec</td>
<td><em>Populus sp.</em>, <em>Epipactis helleverine</em></td>
<td>Puberulum</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td><em>Tuber dryophilum</em> (anamorphic)</td>
<td>Italy, Austria, Germany, Hungary</td>
<td><em>Epipactis helleverine</em>, <em>Pinus pinaster</em></td>
<td>Puberulum</td>
</tr>
</tbody>
</table>
GLOBAL META-ANALYSIS OF TUBER ITS RDNA

Fig. 2 Placement of insufficiently identified Tuber collections and sequences belonging in the Puberulum clade: most likely tree for Puberulum clade. The analysis is based on ITS nuclear rDNA (502 included characters) and a GTR+G+I submodel of nucleotide substitution with four substitution rate classes. Maximum likelihood (ML) bootstrap values are shown above branches and posterior probabilities based on Bayesian inference (BI) below. Values of >70 for ML and >99 for BI are considered significant. Asterisks (*) denote support values of 100, when space was limited. Sequences are labelled with Latin binomials, GenBank accession or collection number and geographical origin. Species are marked by alternating white and grey bands. Tuber melanosporum was chosen as an outgroup.

T. sepatus, Tuber sp.19 and Tuber sp.57) and Argentina (T. borchii, T. californicum and T. maculatum) (Barroeta et al. 2005, 2006; Bulman et al. 2010). Tuber is a Northern Hemisphere lineage and has only been found in New Zealand associated with introduced host plants and never in native forests (Chu-Chou & Grace 1983; Bulman et al. 2010).

Dispersal considerations

Epigean fungi are dispersed by wind, and therefore less constrained than hypogeous forms dependent on small mammals and insects for spore dispersal (Trappe & Claridge 2005; Frank et al. 2006a). In cases where hypogeous fungi successfully establish after introduction, they may fail to spread across the landscape because of a lack of suitable dispersal agents (Nuñez et al. 2009b). There may be exceptions to this epigeous/hypogeous dichotomy however. Anamorphic states have been documented in the hypogeous ectomycorrhizal Pezizales genera Tuber, Pachypleurotus and Rhizomarasmius and may be more common than is realized (Urban et al. 2004; Perry et al. 2007). The effect of anamorphic states on species distributions or invasion biology is unknown. The two known anamorphic Tuber species, T. borchii and T. dryophillum, do not appear to have exceptionally large ranges nor is there evidence for unintentional introductions of these species.
Host preferences for *Tuber* species

Host specificity at the genus and family level occurs in a minority of ectomycorrhizal fungi (Molina & Trappe 1982; Molina et al. 1992). However, some ectomycorrhizal clades may be restricted to particular plant phyla, such as the association between Suillineae (Boletales) and Pinaceae hosts (Taylor & Bruns 1997). In other cases, it appears that individual ectomycorrhizal taxa or clades have preferential associations with certain host groups (Murat et al. 2004; Ishida et al. 2007; Tedersoo et al. 2008; Smith et al. 2009). From this meta-analysis, which includes sequence data from over 70 molecular-based studies of ectomycorrhizal communities, we are able to document 24 *Tuber* species associated with multiple host species and genera. Clearly more data is needed to ascertain the range of host preference and specificity in *Tuber*, but from the data currently available, some speculation can be made. The Gibbosum clade appears to be the only *Tuber* lineage with strong preference for gymnosperms (Bonito et al. 2010). In contrast, species in the Rufum, Excavatum and Maculatum clades may be capable of forming ectomycorrhizal associations with gymnosperms, but show a strong preference towards angiosperm hosts in nature (Montecchi & Sarasini 2000; Halász et al. 2005; Frank et al. 2006b). Species in the Puberulum clade appear to associate with either gymnosperm or angiosperm hosts, and in some cases with both. Species in the Aestivum and Melanosporum clades are typically associated with gymnosperms, whereas *T. indicum* (A & B) commonly fruits under Pinaceae (Trappe 1971; Ceruti et al. 2003; Hall et al. 2007). Mycorrhizal associations with terrestrial orchid species were documented for thirteen *Tuber* species that belong to the Excavatum, Aestivum, Rufum, Maculatum and Puberulum clades (Bidartondo et al. 2004; Selosse et al. 2010).
GLOBAL META-ANALYSIS OF TUBER ITS rDNA

Fig. 3 Placement of insufficiently identified Tuber collections and sequences belonging in the Maculatum clade: most likely tree for Maculatum clade. The analysis is based on ITS nuclear rDNA (574 included characters) and a GTR+G substitution with four substitution rate classes. Maximum likelihood (ML) bootstrap values are shown above branches and posterior probabilities based on Bayesian inference (BI) below. Values of >70 for ML and ≥99 for BI are considered significant. Asterisks (*) denote support values of 100, when space was limited. Sequences are labelled with Latin binomials, GenBank accession or collection number and geographical origin. Species are marked by alternating white and grey bands. Tuber multifasciulatum was chosen as an outgroup.

Phylogenetic placement of insufficiently identified Tuber taxa

In this meta-analysis of over 2000 Tuber ITS sequences representing 76 studies from around the world, a total of 123 phylotypes were distinguished, 39 of which are reported from multiple studies. In some cases, Tuber species were dominant members of the ectomycorrhizal community (Walker et al. 2005; Smith et al. 2007a; Morris et al. 2009; Junpponen et al. 2010), whereas in others they were less frequently detected (Hrynkievicz et al. 2008; Krpata et al. 2008; Leski et al. 2008). Well over 99% of the insufficiently identified Tuber taxa (including sporocarp and pyrosequence data) grouped within three less studied Tuber clades: Puberulum, Maculatum and Rufum. To the best of our knowledge, 64 undescribed species are represented by these data and 36 are known only from DNA sequences (Table S3, supporting information). This raises questions concerning their taxonomy, as species are traditionally described by fruitbody characters (Hibbett et al. 2009). In a step short of sequence-based species descriptions to address taxonomic issues, we have assigned temporary numbers to undescribed species (up to Tuber sp.73) following the convention of Jeandroz et al. (2008).

© 2010 Blackwell Publishing Ltd
Herbaria contain a large number of unsampled taxa and their role in increasing the fungal biodiversity represented in GenBank has been addressed (Brock et al. 2009). However, there are known issues with the nomenclature of identified GenBank accessions (Vigayls 2003; Trappe 2004) and it is estimated that ~20% of the entries have been incorrectly identified (Nilsson et al. 2006). Therefore, a phylogenetic framework is preferable for identifying unknowns, and discretion should be used when taxonomic calls are based solely on BLAST results. Our analyses indicate that GenBank accessions for T. rufum, T. borchii, T. californicum, T. scripsorum, T. maculatum and T. excavatum are composed of multiple phylogenetic species. It is still unclear at this time which (if any) of these accessions are representative of type collections of these species.

Global estimations of Tuber biodiversity

Estimating worldwide species diversity for Tuber has been difficult because of numerous synonyms and misidentifications. Although Index Fungorum lists 256 described species of Tuber (including synonyms and varieties) and the Dictionary of the Fungi states there are 86 species (Kirk et al. 2008), only 70–75 species are believed to be valid (Ceruti et al. 2003; Jeandroz et al. 2003; Nilsson et al. 2006).
Our meta-analysis of global ITS rDNA diversity for Tuber phylotypes (defined as those sharing 96% ITS rDNA sequence similarity) distinguishes 123 phylotypes, although only ~70% of the accepted species were represented in the analyses. Projections of global Tuber species richness from these data predict a minimum of 180–230 species worldwide, depending on the estimator used (Fig. 5). To assess the accuracy of these estimators in predicting species richness when a 'true' value is known, Petersen et al. (2003) compared richness estimators using herbarium collections for Asilidae, a group of conspicuous and well-sampled beetles. They found that the estimators were internally consistent but consistently underestimated the 'true' diversity. This did not seem to be particularly sensitive to sample size or subsampling strategies. While the true number of Tuber species may never be known, it is clear that considerable Tuber diversity awaits discovery. Novel taxa continue to be encountered, even in relatively well-sampled regions (e.g. western Europe, Pacific Northwestern USA) and even more so in less studied regions of high plant endemism including central Asia, Japan and Mexico.

In summary, working within a phylogenetic framework we demonstrate that Tuber is more diverse than previously realized and that most of the diversity resides within non-economical and less studied clades (i.e. Rufum, Puberulum, and Maculatum). Most Tuber clades show strong phylum-level preference to either angiosperm or gymnosperm hosts. We also infer from these data that Tuber species in Europe, Asia and North America are endemic to their respective continents and are generally not shared between continents except in cases of human-mediated long-distance dispersal.

Acknowledgements

This research would not have been possible without the contributions from members of the North American Truffling Society and other collectors who provided specimens and sequences used in this research. For this, we thank Rosanne Healy, Efren Cazares, Gonzalo Guevara, Matthew Smith, Thomas Horton, Alessandra Zambonelli, Donald Pfister, Baldomero Moreno Arroyo, Michael Castellano, Amer Montecchi, Jeri Parment, Peter Kennedy, Kathie Hodge, Matthew Trappe, Bernardini Danilo and Tim Baroni. Thanks to Joey Spatafora and Richard Halse of the Oregon State University Department of Botany Herbarium for extensive herbarium services, including accessioning of the collections cited here plus many additional collections. The National Fungus Collections, the Farlow Herbarium and the Herbarium of the Universita di Bologna generously lent additional collections for study. Matthew Smith, Terri Porter, Jason Jackson, Hannah Reynolds and Michelle Heras provided valuable discussion and feedback. This research was supported through NSF awards DEB-0614297 and DEB-0710213 with additional funding by Friends of the Farlow and the Mycological Society of America.

References

Bonito G (2009a) Fast DNA-based identification of the black truffle Tuber melanosporum with direct PCR and species-specific primers. FEMS Microbiology Letters, 301, 171-175.
Bradshaw BP (2005) Physiological Aspects of Corylus avellana Associated with the French Black Truffle Fungus Tuber melanosporum and the Consequence for Commercial Production of Black Truffles in Western Australia, Murdoch University, Perth.

© 2010 Blackwell Publishing Ltd


© 2010 Blackwell Publishing Ltd


Supporting information

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

Table S1 Citations for reference sequences used and insufficiently identified sequences of Tuber retrieved by enclavem. GenBank number of Tuber sequences are presented at the end of each citation

Table S2 Tuber groupings based on 98% ITS rDNA sequence similarity

Table S3 Thirty-six Tuber species known only from sequence data

Table S4 Most likely tree of the Excavaturn clade based on ITS nuclear rDNA and a GTR+G+I model of nucleotide substitution with four substitution rate classes. Three phylogenetic species of 'T. excavatum' are resolved. Maximum likelihood (ML) bootstrap values are shown above branches and posterior probabilities based on Bayesian inference (BI) are below branches. Values of >70 for ML and >99 for BI are considered significant

Table S5 Collection information and GenBank accession numbers for sequences generated during this study

Table S6 Phylogenetic placement of 1548 Tuber sequence data from Quercus sp. ectomycorrhizae in Kansas, USA produced through pyrosequencing (Jumpponen et al. 2010). This most likely tree includes taxa from the Muculatum and Rufum clades and is based on ITS nuclear rDNA and a GTR+G+I model of nucleotide substitution. The eleven phylogenotypes represented by these sequences are shown in a larger font and their abundance based on binning at 96% is presented on the taxon labels. Three phylogenotypes belonging to the Rufum clade were unique to this study and are labelled with an asterisk (*). Maximum likelihood (ML) bootstrap values are shown on top of the branches. Values of >70 for ML are considered significant

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

G.B.’s research involves truffle systematics, fungal ecology, and biogeography. A.G.’s research is focused on biodiversity and fungal mating systems. J.T. has spent his career researching truffle biodiversity, taxonomy systematics. Much of R.V.’s research has focused on molecular systematics of fungi.