Congruent species delineation of *Tulasnella* using multiple loci and methods

**Introduction**

Understanding evolutionary and ecological processes requires accurate delimitation of species. Species are most commonly defined under the general lineage concept (GLC), where they are considered to be segments of diverging population-level lineages (de Queiroz, 2007). Within the umbrella of the GLC, various species concepts have been developed, including the morphological species concept (MSC), the biological species concept (BSC) and the phylogenetic species concept (PSC). The BSC is characterized by species representing populations that potentially can interbreed. The PSC includes the genealogical concordance phylogenetic species recognition concept, which uses phylogenetic concordance of multiple unlinked genes to identify evolutionary independence of lineages (Taylor *et al.*, 2000). PSC is applied frequently in taxonomic groups where it is difficult to quantify morphological variation or perform mating studies. One such group is *Tulasnella*, a group of fungi where morphological identification of species is problematic.

*Tulasnella* includes putatively saprotrophic species on decayed wood (Roberts, 1999; Cruz *et al.*, 2011). Some species are also encountered as ectomycorrhizas (Tedersoo *et al.*, 2010) or orchid mycorrhizal symbionts (Dearnaley *et al.*, 2012). However, many questions remain about species delimitation in this group. We therefore need a robust multilocus method for species delimitation to establish a framework for studying the evolution, ecology and physiology of orchid–fungus relationships.

It is now recognized that the most effective approach to species delineation is the integration of multiple datasets and analytical methods (Sites & Marshall, 2004; Leaché & Fujita, 2010; Yang & Rannala, 2010; Barrett & Freudenstein, 2011). Here we use evidence from six nuclear loci, two mitochondrial loci, orchid-host association and geographical location of samples in a multifaceted approach to delineate species of *Tulasnella* associated with the Australian orchid genera *Chiloglottis*, *Drakaea*, *Paracaleana* and *Arthrochilus*. Specifically, we employ gene tree construction methods to resolve fungal species boundaries; use coalescent species tree construction methods (using the programs *BEAST* and BPP) to test for host and geographic association; and apply population genetic assignment methods to test for admixture between populations. Finally, germination data are used to explore the correlation between physiological traits and phylogenetic boundaries among *Tulasnella*. In light of the outcomes we evaluate the implications of multigene approaches for fungal species delimitation in this *Tulasnella* group.

**Materials and Methods**

**Sampling of *Tulasnella* isolates**

In total, 28 *Tulasnella* isolates from eight species of *Chiloglottis*, 17 isolates from seven *Drakaea* species, nine isolates from five species of *Paracaleana* and nine isolates from *Arthrochilus oreophilus* D.L. Jones were analysed (Supporting Information Table S1, Fig. S1). Samples from *Chiloglottis*, *Paracaleana minor* (R.Br) Blaxell and *A. oreophilus* were all collected from eastern Australia, while all other samples were collected from south-western Australia.

**Loci**

Six sequence loci for *Tulasnella*, which includes a mitochondrial (C14436; ATP) and five nuclear loci (C3304; ATP helicase, C10499; 26 proteasome regulatory complex, C12424; isocitrate dehydrogenase, C4722; CAS1, C4102; glutamate synthase), were amplified and sequenced as described previously (Ruibal *et al.*, 2013). These loci were developed using sequences from a 3 kb pair-end sequence library (Roche) on a GS FLX 454 platform using GS XL70 sequencing chemistry (Ruibal *et al.*, 2013). Additionally, the internal transcribed spacer (ITS) region, using primers ITS1 and ITS4Tul4, and the mtLSU were sequenced and edited following Roche *et al.* (2010).

**Species delimitation**

**Phylogenetic analyses** A multiple sequence alignment was constructed using the alignment tool in Geneious pro v5.6.3 (Drummond *et al.*, 2011) before performing manual checks and minor adjustments. Intronic regions that were difficult to align for loci C14436 (base pair positions 183–250), C4722 (13–86), C1242 (221–281) and C4102 (164–216) were deleted. Phylogenies of individual and concatenated loci were estimated with a maximum likelihood (ML) analysis using RAxML 7.0.3 (Stamatakis *et al.*, 2008) and using Bayesian inference with MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003). Support for nodes was assessed for ML trees using 1000 pseudoreplicates of nonparametric bootstrapping in RAxML and with Bayesian Posterior Probabilities (BPP) in MrBayes. A GTR+G substitution model was used for all analyses as all other models are nested inside this model. Trees were visualized using FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/) and mid-point rooted. Topological similarity between trees of each locus was tested with a congruency test (de Vienne *et al.*, 2007).
Sequence diversities and genetic divergence were calculated in Mega 5.05 (Tamura et al., 2011) (Table S2). For genetic divergence estimation we employed the p-distance as well as the Kimura 2-parameter (K2P) distances (Kimura, 1980) (pairwise deletion), which, despite some limitations (see Srivathsan & Meier, 2012), is recommended by the consortium for the Barcoding of Life (http://www.barcoding.si.edu/protocols.html) in order to standardize comparisons among studies.

**Population structure analysis** While assignment tests are better known for their ability to cluster individuals into K populations (Manel et al., 2005), they are also useful for inferring species boundaries (Noble et al., 2010; Barrett & Freudenstein, 2011). Simulation and empirical studies indicate that STRUCTURE is generally robust even when the assumption of Hardy–Weinberg-Equilibrium (Pritchard et al., 2000) is not met (Martien et al., 2007; Behere et al., 2013). We tested for the presence of population and species level structure within fungi across the eight-locus sequence dataset using Bayesian genetic assignment as incorporated in STRUCTURE v2.3.3 (Pritchard et al., 2000; Falush et al., 2003) (Notes S1). Separate analyses were run for three sets of fungal isolates, representing the main clades revealed in the eight locus RAxML phylogeny (Fig. 1): (1) *Chiloglottis*; (2) *Drakaea + Paracaleana*; and (3) *Arthrochilus*.

Bayesian coalescent-based species delimitation: host and/or geographic association

Reciprocal monophyly or diagnostic states (e.g. fixed differences between putative species) are often used as criteria for species delimitation using molecular data (Sites & Marshall, 2004). By contrast, coalescent-based species delimitation methods do not require reciprocal monophyly (Fujita et al., 2012). To apply this approach, we first generated coalescent species trees (guide trees) using *BEAST 1.7.4* (Heled & Drummond, 2010). The species trees were inferred alternatively by incorporating the host species from which the *Tulasnella* was isolated or using the geographic origin of isolates. We then used the guide trees in a Bayesian species delimitation approach using the program BPP 2.0 (Yang & Rannala, 2010) (see Notes S1 for details).

Germination trials

The ability of a random subset of fungal isolates to germinate orchid seed was tested with germination trials of *Drakaea, Chiloglottis* and *Paracaleana* following Roche et al. (2010) (Notes S1).

**Results and Discussion**

**Phylogenetic analyses**

Phylogenetic analyses of the eight individual loci showed similar topologies (Figs 2, S2–S8), with all loci passing the congruence test. Gene trees for individual loci (Figs 2, S2–S8) and for the combined eight locus matrix (Fig. 1) always gave strong bootstrap and Bayesian probability support for at least four *Tulasnella* clades; a clade each for *Tulasnella* from *Chiloglottis* and *Drakaea + Paracaleana*, plus three clades from *Arthrochilus*. There was also strong support for *Tulasnella* from *P. minor* as sister to *Tulasnella* from *Drakaea*, however, in loci C12424, C14436 and C3304, all grouped into one clade. *Tulasnella* isolated from *Paracaleana* in western Australia all fell within the *Drakaea* clade except the isolate from *P. lyonsii*. This isolate was sister to *Tulasnella* from *Drakaea* and all other *Paracaleana* in the combined sequence matrix (Fig. 1) and in five out of eight individual loci (Figs 2, S2–S8); however, strong support was not always achieved for this relationship. Comparing the same isolates, both the ITS and eight-locus concatenated data set resulted in strong support for three *Tulasnella* clades from *Arthrochilus* (Figs 1–2).

Most mycorrhizal studies that attempt to delineate species via ITS sequence data rely on the rule of thumb that clades showing > 3% sequence divergence belong to different species (Nilsson et al., 2008; Peay et al., 2008; Hughes et al., 2009). However, levels of ITS variation can vary within fungal species, meaning the use of a threshold may not always accurately reflect species boundaries (Nilsson et al., 2008). Further, genetic thresholds do not explicitly take into account how the timing of speciation influences patterns of genetic differentiation (e.g. Matz & Nielsen, 2005).

Across the eight loci analysed (5015 bp), mean within host group K2P sequence divergences (converted to percentage) were 0.69% (*Tulasnella* from *Drakaea + Paracaleana*), 1.74% (*Tulasnella* from *Chiloglottis*) and 6.69% (*Tulasnella* from *Arthrochilus*). The mean within host group sequence divergence for the ITS ranged from 0.43% to 7.52% (Table 1). K2P sequence divergence (for ITS) between isolates from *P. minor* and *Drakaea* was 0.60% and was 1.3% between isolates from *P. minor* and *P. lyonsii*. *Paracaleana lyonsii* showed 1.4% sequence divergence to the *Tulasnella* clade consisting of isolates from *Drakaea* and *Paracaleana* (excluding *P. minor*). The individual divergences among *Tulasnella* isolates from *Paracaleana* and *Drakaea*, fell well within the range of ITS variation deemed to be typical of ‘within species variation’ (Hebert et al., 2003; Jacquemyn et al., 2010), where clades with < 3% sequence divergence among individuals are referred to as Operational Taxonomic Units (OTUs) and are considered likely to represent species. The percentage p-distances were comparable or even < K2P distances (Table 1), further supporting low genetic divergence within host groups. Although we do not suggest a universal sequence divergence threshold for *Tulasnella*, the phylogenetic analyses and a sequence divergence of < 1.8% within host groups suggest the presence of one *Tulasnella* species (OTU) each from *Chiloglottis* and *Drakaea + Paracaleana*, and three *Tulasnella* species (OTUs) from *Arthrochilus* (Fig. 1).

**Population structure**

Assignment tests for the two mitochondrial and six nuclear loci indicated that the number of genetic clusters (K) among fungi from *Chiloglottis* was most likely K = 1 as indicated by the loge P(D) values (Fig. S10). For fungi from *Drakaea + Paracaleana* loge P(D) values indicated K = 1 or 2, with isolates from *P. minor*. *P. lyonsii
Species tree inference and posterior probabilities: host and/or geographic association

The species trees and posterior probabilities inferred with *BEAST and BPP are shown in Fig. S9. Strong posterior probability support is inferred when values are ≥ 0.99. The *BEAST analysis with the 21 host-associated populations (Fig. S9a) provides strong posterior probability support (≥ 0.99) for four populations, one each representing Tulasnella from Chiloglottis, P. minor, Drakaea + Paracaleana from western Australia, and Arthrochilus. The one isolate from P. lyonsii formed...
an unsupported sister to the *Tulasnella* populations from *Drakaea* and *Paracaleana* from western Australia. With Bayesian species delimitation using BPP, a 10 species 95% credible model (Table S3) was identified with all prior combinations (Fig. S9a, Table S3). The remaining putative species were not strongly supported and were influenced by the choice of priors. Changing the guide tree topology did not alter overall posterior probability support for the hypothesized species obtained by BPP.

The *BEAST* analysis using the geography-associated populations as priors provide strong (posterior probability = 1) support for four *Tulasnella* geographical populations including two from eastern Australia (those representing populations from *Chiloglottis* and a population from *P. minor*) and two from south-western Australia. The *P. minor* population (eastern Australia) formed a weakly supported sister to the *Tulasnella* populations from western Australia. Species delimitation analyses with BPP found further support (posterior probabilities ≥ 0.99) for geographic groups (Fig. S9b).

Bayesian species delimitation using the *Chiloglottis* guide tree supports all nodes, with posterior probabilities of one, except the *Tulasnella* from *C. aff. jeanesii* + *C. valida* node (Fig. S9c). The host-associated guide tree (Fig. S9a) also did not support the *Tulasnella* nodes from *C. aff. jeanesii* + *C. valida*. Thus, for *Tulasnella* from *Chiloglottis*, phylogenetic, coalescent (*BEAST*) and *STRUCTURE* analyses provide strong support that all *tulasnellas* analysed from *Chiloglottis* belong to a single species. This contrasts with BPP, which recognized up to four lineages depending on which guide tree (Fig. S9a–c) was used. In light of the broad congruence across the phylogenetic, *BEAST* and *STRUCTURE* analyses, we urge caution in relying on BPP alone for species delimitation.

The three step process for Bayesian species delimitation that we employed: (1) population structure analysis (*STRUCTURE*); (2) species tree reconstruction (*BEAST*); and (3) posterior probabilities of species cluster inference (*BEAST and BPP*); collectively indicated that one *Tulasnella* species (OTU) each is associated with *Chiloglottis* and *Drakaea + Paracaleana* (Fig. 1).
Germination

Although there was some variation in germination success among seed lots, all tested *Tulasnella* isolates from *Chiloglottis* germinated *Chiloglottis*, but not *Drakaea* or *Paracaleana nigrita* seed. Interestingly, all fungi from *Drakaea* and *Paracaleana*, including divergent fungi from both sides of the country, supported germination of both *Drakaea* and *Paracaleana* seed in vitro (Table S4). The results are definite, despite the well-known variability of seed germination trials between replicate isolates and seed batches (Phillips et al., 2011). We interpret the germination results as further evidence that *Tulasnella* from *Drakaea* and *Paracaleana* belong to the same species and as consistent with the phylogenetic evidence that all *Tulasnella* isolates from *Chiloglottis* belong to a single species.

Conclusions and directions

This study has combined six newly developed DNA sequence loci in combination with the ITS and mtLSU loci to delimit *Tulasnella* species associated with several genera of Australian orchids. Given the pitfalls of single-criterion species delimitation, we applied a multi-method approach over multiple loci. Further, we used specified sequence divergence estimation methods to allow comparison with future studies. Under the general lineage species concept based on phylogenetics, coalescent species trees and germination data, the multilocus analyses provide further support to our previous ITS phylogenetic studies (Roche et al., 2010; Phillips et al., 2011). All *Tulasnella* isolates analysed from *Chiloglottis* belong to one species, whereas those from *Drakaea* and *Paracaleana* belong to a sister species (Fig. 1).

What are the implications of our findings, given that an eight-locus analysis is broadly congruent with the ITS locus in this group of fungi? Although PCR amplification of the ITS is sometimes difficult, given its widespread use across fungi spanning many phyla (Schoch et al., 2012), we suggest that there may be a case in *Tulasnella*, at least, where the ITS provides good phylogenetic support for species delimitation. However, more generally the use of multiple sequence loci remains optimal for species delimitation (e.g. Knowles & Carstens, 2007). We predict that as we move towards routine multilocus analyses, which are becoming increasingly more cost efficient with high throughput sequencing methods, that these new studies will largely consolidate, rather than re-write our current understanding of phylogenetic relationships in fungi.

Acknowledgements

C.C.L. and R.P. were funded by the Australian Research Council (LP098338), and R.D.P. by the Australian Orchid Foundation and the Holsworth Wildlife Research Endowment. The Department of Environment and Conservation issued permits to collect Declared Rare species of *Drakaea*. The authors thank Don Gomez for collection of *Arthrochilus oreophilus*, the Editor Marc-André Selsosse, L. Tedersoo and two anonymous reviewers for helpful comments that improved the manuscript.

Celeste C. Linde1,*, Ryan D. Phillips1,2,3, Michael D. Crisp1 and Rod Peakall1

1Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT, 0200, Australia;

Table 1 Within host group and between group K2P- and ρ-genetic (italicised) distances for *Tulasnella* species as calculated from the eight loci concatenated data set and internal transcribed spacer (ITS)

<table>
<thead>
<tr>
<th>Species</th>
<th>K2P genotype</th>
<th>ρ genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tulasnella from Drakaea + Paracaleana</em></td>
<td>(0.69 ± 0.06%)</td>
<td>(0.69 ± 0.06%)</td>
</tr>
<tr>
<td><em>Tulasnella from Drakaea + Paracaleana</em></td>
<td>17.99 ± 0.72%</td>
<td>14.81 ± 0.45%</td>
</tr>
<tr>
<td><em>Tulasnella from Arthrochilus</em></td>
<td>(6.69 ± 0.27%)</td>
<td>(5.81 ± 0.24%)</td>
</tr>
<tr>
<td><em>Tulasnella from Arthrochilus</em></td>
<td>19.37 ± 0.75%</td>
<td>15.69 ± 0.51%</td>
</tr>
<tr>
<td><em>Tulasnella from Arthrochilus</em></td>
<td>15.68 ± 0.63%</td>
<td>12.75 ± 0.41%</td>
</tr>
<tr>
<td><em>Tulasnella from Chiloglottis</em></td>
<td>(1.23 ± 0.20%)</td>
<td>(1.17 ± 0.18%)</td>
</tr>
<tr>
<td><em>Tulasnella from Drakaea + Paracaleana</em></td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><em>Tulasnella from Drakaea + Paracaleana</em></td>
<td>22.35 ± 0.21%</td>
<td>16.78 ± 1.16%</td>
</tr>
<tr>
<td><em>Tulasnella from Drakaea + Paracaleana</em></td>
<td>18.64 ± 1.85%</td>
<td>14.46 ± 1.09%</td>
</tr>
</tbody>
</table>

Numbers in bold and in brackets represent within host group genetic distances whereas the remaining numbers represent between host group genetic distances. K2P-genetic distance between these clades and CLM031 is 9.84 ± 0.16%.

*Numbers in bold and in brackets represent within host group genetic distances whereas the remaining numbers represent between host group genetic distances.
References


Supporting Information

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

Fig. S1 Distribution of Tulanella samples obtained from orchids in Australia.

Fig. S2 Midpoint root Maximum Likelihood tree for Tulanella obtained for the mtLSU.

Fig. S3 Midpoint root Maximum Likelihood tree for Tulanella obtained for locus C4102.
Fig. S4 Midpoint rooted Maximum Likelihood tree for *Tulasnella* obtained for locus C12424.

Fig. S5 Midpoint rooted Maximum Likelihood tree for *Tulasnella* obtained for locus C14436.

Fig. S6 Midpoint rooted Maximum Likelihood tree for *Tulasnella* obtained for locus C3304.

Fig. S7 Midpoint rooted Maximum Likelihood tree for *Tulasnella* obtained for locus C4722.

Fig. S8 Midpoint rooted Maximum Likelihood tree for *Tulasnella* obtained for locus C10499.

Fig. S9 The coalescent-based species trees and Bayesian species delimitation results for *Tulasnella* Bayesian species trees inferred with *BEAST*.

Fig. S10 Bayesian model-based clustering likelihoods and $\Delta K$ model selection for 28 fungal isolates from *Chiloglottis*.

Fig. S11 Bayesian model-based clustering likelihoods and $\Delta K$ model selection for 26 fungal isolates from *Drakaea + Paracaleana*.

Fig. S12 Bayesian model-based clustering likelihoods and $\Delta K$ model selection for nine fungal isolates from *Arthrochilus*.

Table S1 Fungal symbionts from *Chiloglottis, Drakaea, Paracaleana and Arthrochilus* used in phylogenetic analyses

Table S2 Characteristics of phylogenetic markers for *Tulasnella* fungi from *Chiloglottis, Drakaea, Paracaleana and Arthrochilus* orchids

Table S3 Bayesian posterior probabilities for the speciation models sampled by BPP under different combinations of $\theta$ and $\tau_0$ priors

Table S4 Results of germination trials with *Tulasnella* isolates and orchid seed combinations used

Notes S1 Additional information on methods and results.

Key words: fungi, internal transcribed spacer (ITS), multilocus analyses, operational taxonomic units (OTUs), orchids, species delimitation, *Tulasnella*.