Naturalis Repository

DNA barcoding of tuberous Orchidoideae: a resource for identification of orchids used in Salep

Abdolbaset Ghorbani, Barbara Gravendeel, Sugirthini Selliah, Shahin Zarré, Hugo de Boer

Downloaded from:
https://doi.org/10.1111/1755-0998.12615

Article 25fa Dutch Copyright Act (DCA) - End User Rights

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with consent from the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available following a reasonable period after the work was first published, provided that reference is made to the source of the first publication of the work.

This publication is distributed under the Naturalis Biodiversity Center 'Taverne implementation' programme. In this programme, research output of Naturalis researchers and collection managers that complies with the legal requirements of Article 25fa of the Dutch Copyright Act is distributed online and free of barriers in the Naturalis institutional repository. Research output is distributed six months after its first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and copyrights owner(s) of this work. Any use of the publication other than authorized under this license or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the department of Collection Information know, stating your reasons. In case of a legitimate complaint, Collection Information will make the material inaccessible. Please contact us through email: collectie.informatie@naturalis.nl. We will contact you as soon as possible.

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2017) **17**, 342–352 doi: 10.1111/1755-0998.12615

DNA barcoding of tuberous Orchidoideae: a resource for identification of orchids used in Salep

ABDOLBASET GHORBANI,*† BARBARA GRAVENDEEL,‡§ SUGIRTHINI SELLIAH,¶
SHAHIN ZARRÉ** and HUGO DE BOER*‡¶
*Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236 Uppsala, Sweden, †Traditional Medicine and Materia Medica Research Center, Shahid Beheshti University of Medical Sciences, No 19, Tavanir Street, Hemmat Highway, P.O. Box 14155-6153, Tehran, Iran, ‡Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands, §University of Applied Sciences Leiden, Zernikedreef 11, 2333 CK Leiden, The Netherlands, ¶The Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, 0318 Oslo, Norway, **Department of Plant Sciences, School of Biology, College of Science, University of Tehran, 14155-6455 Tehran, Iran

Abstract

Tubers of terrestrial orchids are harvested and traded from the eastern Mediterranean to the Caspian Sea for the traditional product Salep. Overexploitation of wild populations and increased middle-class prosperity have escalated prices for Salep, causing overharvesting, depletion of native populations and providing an incentive to expand harvesting to untapped areas in Iran. Limited morphological distinctiveness among traded Salep tubers renders species identification impossible, making it difficult to establish which species are targeted and affected the most. In this study, a reference database of 490 nrITS, *trn*L-F spacer and *mat*K sequences of 133 taxa was used to identify 150 individual tubers from 31 batches purchased in 12 cities in Iran to assess species diversity in commerce. The sequence reference database consisted of 211 nrITS, 158 *trn*L-F and 121 *mat*K sequences, including 238 new sequences from collections made for this study. The markers enabled unambiguous species identification with tree-based methods for nrITS in 67% of the tested tubers, 58% for *trn*L-F and 59% for *mat*K. Species in the genera *Orchis* (34%), *Anacamptis* (27%) and *Dactylorhiza* (19%) were the most common in Salep. Our study shows that all tuberous orchid species in this area are threatened by this trade, and further stresses the urgency of controlling illegal harvesting and cross-border trade of Salep tubers.

Keywords: CITES, molecular identification, orchid conservation, overharvesting, plant DNA barcoding, wildlife trade

Received 5 June 2016; revision received 26 August 2016; accepted 6 September 2016

Introduction

Tuberous terrestrial orchids have long been used as medicine and dietary supplements in different parts of the world (Bulpitt 2005; Bulpitt *et al.* 2007; Chinsamy *et al.* 2011; Hossain 2011). In the eastern Mediterranean, Asia Minor and the Middle East, tubers of different orchid species are collected indiscriminately from the wild and are traded for production of Salep tuber powder (Kasparek & Grimm 1999; Ece Tamer *et al.* 2006; Sandal Erzurumlu & Doran 2011; Ghorbani *et al.* 2014a; Kreziou *et al.* 2015). Harvested tubers are washed in water, boiled in either water or milk, sun-dried and traded as dried tubers (Kasparek & Grimm 1999). The tubers are ground into a powder and used in preparing a hot beverage known as Salep or Salepi and also in ice cream production (Sezik 2002a; Ece Tamer *et al.* 2006; Starin 2012). Salep drink was once common in Europe (Landerer 1850) but is now consumed mainly in Turkey and Greece (Bulpitt 2005; Ece Tamer *et al.* 2006; Starin 2012). It is estimated that as much as 30 tons of orchid tubers are harvested annually in Turkey, which requires the destruction of 30–120 million orchid plants (Kasparek & Grimm 1999; Sezik 2006). Increasing popularity of Salep has increased the demand for Salep tubers, which in turn has led to further overharvesting of wild orchid populations (Sezik 2002b; Kreziou *et al.* 2015). Scarcity of wild orchids in Turkey has forced traders to tap into new sources in adjacent countries (Ghorbani *et al.* 2014b). In Iran, where orchid tubers are traditionally hardly consumed, an orchid boom is underway in which an estimated 5.5–6.1 million orchids are harvested annually for export to Turkey (Ghorbani *et al.* 2014a). Conservation concerns have made orchid tuber collection illegal in

Correspondence: Hugo de Boer, Fax: +47 228 51875;
E-mail: hugo.deboer@nhm.uio.no

Greece, Turkey and Iran, but collection bans are poorly enforced (Ghorbani *et al.* 2014b; Kreziou *et al.* 2015). All orchid species are included by the Convention on International Trade of Endangered Species of Fauna and Flora (CITES) on Appendices I or II (CITES 2014), which means that international trade of these species and derived products is regulated. Most of the Salep tuber trade from Iran to Turkey takes place without CITES permits, and tubers are often mislabelled as low-value nuts or other products to circumvent taxes and permit requirements (Kasparek & Grimm 1999; Ghorbani *et al.* 2014b; Kreziou *et al.* 2015). This large-scale, yet poorly visible trade makes it difficult to ascertain which species are targeted and in what quantities. Morphology-based approaches for identification are insufficient and cannot even accurately distinguish dried tubers from different genera. Other methods for salep identification, such as GCMS, HPLC, gravimetric, absorbance and rheological analyses, all indicate that identification to species level is not possible using only chemical analyses (Dogan *et al.* 2007; Tekinşen & Güner 2010; Babbar & Singh 2016). Adequate monitoring would enable identification of priority species for conservation measures such as curbing overexploitation, and targeting high-value species for cultivation.

DNA barcoding provides an accurate and reliable alternative to morphology-based identification of biological material (Hebert *et al.* 2003). As a method, it can be used to identify and discern species at any developmental or processing stage from which DNA can be extracted (Hebert *et al.* 2003; Hajibabaei *et al.* 2007), and even from the minute amounts such as those found in dung (Hibert *et al.* 2013), pollen (Richardson *et al.* 2015), degraded herbarium vouchers (Särkinen *et al.* 2012), permafrost preserved subfossils (Van Geel *et al.* 2008) and ancient sediment cores (Williams *et al.* 2000; Posadzki *et al.* 2012). Plant DNA barcoding has been applied in many fields, for example molecular systematics (Liu *et al.* 2011; Van Velzen *et al.* 2012), biodiversity inventories (Aubriot *et al.* 2013; Thompson & Newmaster 2014), wildlife forensics (Deguilloux *et al.* 2002; Ogden *et al.* 2009), biopiracy control (Parveen *et al.* 2012) and authentication of herbal products (Coghlan *et al.* 2012; Kool *et al.* 2012; Newmaster *et al.* 2013; De Boer *et al.* 2014; Vassou *et al.* 2015).

Several genetic regions have been proposed as standard barcodes for land plants, the ideal barcode being both easily amplifiable and efficiently retrievable from any of the 300 000+ species of plants (Kress *et al.* 2005; Fazekas *et al.* 2008). Most studies now employ a tiered multilocus approach, which is based on the use of a common, easily amplified and aligned region such as *rbc*L, *rpo*C1, *trn*L or *trn*L-F spacer that can act as a scaffold on which to place data from a more variable noncoding region such as *mat*K, *trn*H-*psb*A, nrITS or nrITS2. Most species (approximately 75%–85%) can be identified using such an approach, and the subsequent addition of surrogate regions can increase barcoding success to over 90% in some floras (Ebihara *et al.* 2010; Burgess *et al.* 2011; De Vere *et al.* 2012; Kuzmina *et al.* 2012; Liu *et al.* 2015). In Orchidaceae, several plastid and nuclear molecular markers including *rbc*L, *psa*B, *psb*C-*trn*S, *rpl*16, *mat*K, *ycf*1, *trn*H-*psb*A, *trn*H-*trn*K, *trn*L-F and nrITS have been applied for phylogenetic analysis (Cameron 2004; Xiang *et al.* 2011; Inda *et al.* 2012; Parveen *et al.* 2012; Kim *et al.* 2014). These studies suggest that a multilocus combination of coding and noncoding regions with different evolutionary rates is necessary for effective identification of species in Orchidaceae.

This study tests the hypothesis that molecular identification using DNA barcoding can be used for identification of orchid species comprising boiled and dried tuber samples traded in the main export market hubs in Iran. We address the following research questions: (i) Can DNA be extracted, amplified and sequenced from boiled and dried Salep tubers? (ii) What marker or markers are optimal for the identification of Salep tubers traded in the markets of Iran? (iii) What genera and species are most common among the tubers included in our sampling? (iv) Can the most common traded species be used to predict the main source areas of orchid tubers exported to Turkey? The aim was to test and establish a DNA barcoding protocol to identify dried orchid tubers from markets and to show the potential of this technique to curb illegal trade of CITES listed orchid tubers.

Methods

Collection of reference and market material

Flora Iranica vol. 126 (Renz 1978), Flora of Iran vol. 57 (Shahsavari 2008) and Orchids of Europe, North Africa and the Middle East (Delforge 2006) were used to estimate that a total of 47 orchid species occur in Iran, including 32 species with tuberous roots that could potentially be targeted for Salep collection. During fieldwork in 2013–2014, a total of 127 herbarium vouchers representing 30 species and subspecies of orchids were collected from natural populations in different parts of Iran (Appendix S1, Supporting information). Vouchers were identified (Renz 1978; Delforge 2006; Shahsavari 2008) and deposited at the herbarium of Tehran University (TUH). Sequences generated from these vouchers (Appendix S1, Supporting information) as well as selected vouchered sequences from NCBI GenBank were used to construct a DNA barcode reference library (Appendix S2, Supporting information).

Markets in 12 cities and towns in Iran (Tehran, Kermanshah, Sanandaj, Tabriz, Urmia, Mahabad, Shahindezh, Kashan, Ardabil, Aq-Emam, Marave-Tappe and Kalaleh) were visited, and 31 batch samples of unidentified Salep tubers containing 15–50 tubers each were purchased. Figure 1 shows the distribution of orchids in Iran at genus level based on indexed vouchers from TUH and W, plus the location of the 12 main Salep markets. Per sample, tubers were subsequently categorized based on shape and size, and a total of 150 random tubers were selected as query tubers for DNA barcoding. Salep tubers in trade are hard to identify, although palmate *Dactylorhiza* tubers differ from those of other tuberous genera (Fig. 2).

DNA extraction, amplification and sequencing

For reference samples, total genomic DNA was extracted from silica gel-dried leaf material using a modified CTAB protocol (Doyle & Doyle 1987). The query tubers were ground into powder using liquid nitrogen, and subsequently DNA was extracted using a STE-CTAB protocol (Shepherd & McLay 2011). The STE-CTAB protocol was necessary to reduce gel formation due to the high glucomannan content of tubers. A gelatinous layer, which was formed after adding CTAB buffer, caused difficulties in extraction procedures and low DNA yields. Extracted DNA was purified using a GE Illustra GFX™ PCR DNA and Gel Band Purification kit following the manufacturer's protocol (GE Healthcare, Buckinghamshire, UK).

Three barcode regions, nrITS (ITS1-5.8S-ITS2), *trn*L-F spacer and *mat*K were amplified by a standard polymerase chain reaction (PCR). The nrITS (ITS1-5.8S-ITS2) region was amplified using the following primers: 17SE_F (5′-ATGGTCCGGTGAAGTGTTC-3′), 26SE_R (5′-CCCGGTTCGCTCGCCGTTAC-3′) and 5.8I-1_R (5′-GTTGCCGAGAGTCGT-3′) and 5.8I-2_F (5′-GCCTGGGCGTCACGC-3′) (Sun *et al*. 1994). The *trn*L-F spacer was amplified using the following primers: C_F (5′-CGAAATCGGTAGACGCTACG-3′), C2_F (5′-GGATAGGTGCAGAGACTCAAT-3′) and F_R (5′-ATTTGAACTGGTGACACGAG-3′) (Taberlet *et al*. 1991; Bellstedt *et al*. 2001). *Mat*K was amplified using the following four primers: 19_F (5′-CGTTCTGACCATATTGCACTATG-3′) and 881R (5-TMTTCATCAGAATAAGAGT-3) (Gravendeel *et al.* 2001);

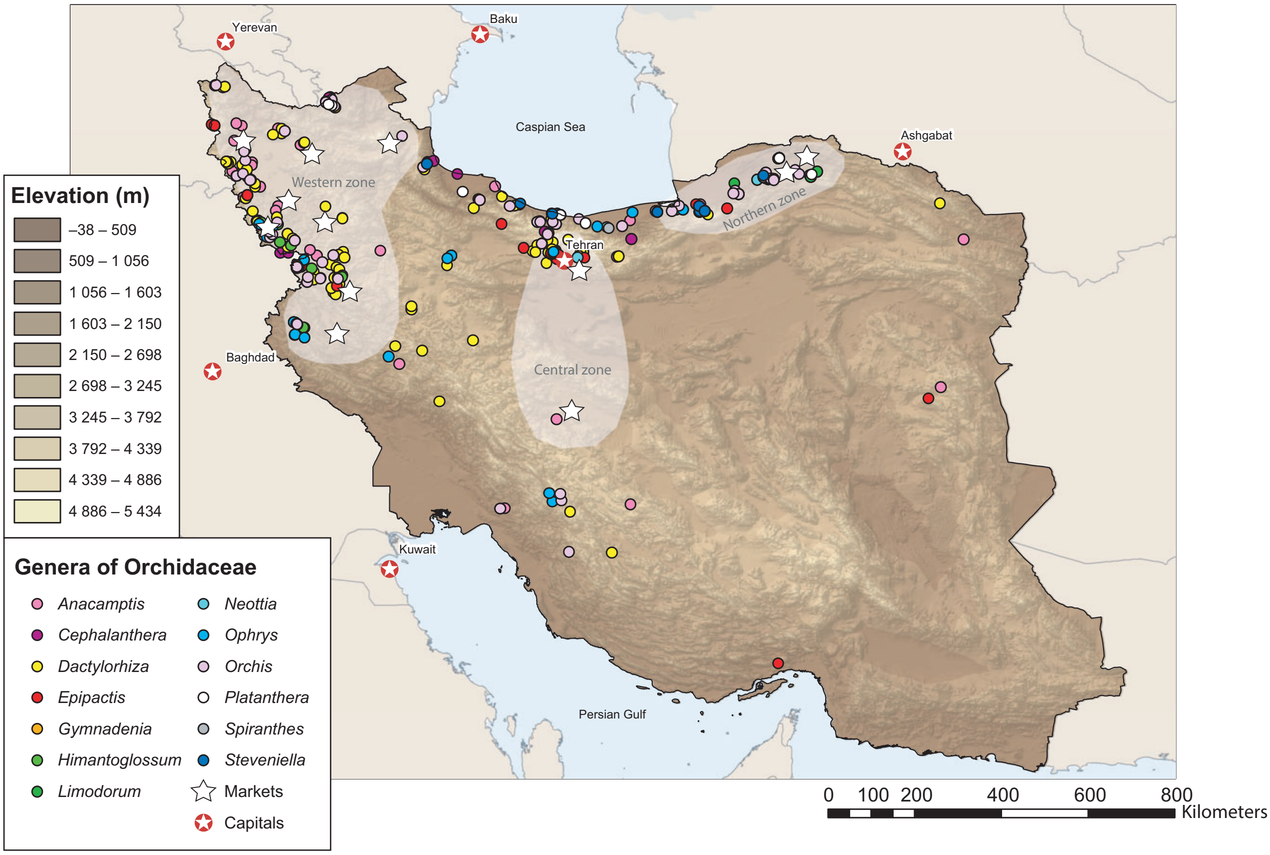

Fig. 1 Distribution of Salep genera in Iran and location of main markets. Shaded areas show the three zones of origin: western, central and northern zones. [Colour figure can be viewed at wileyonlinelibrary.com].

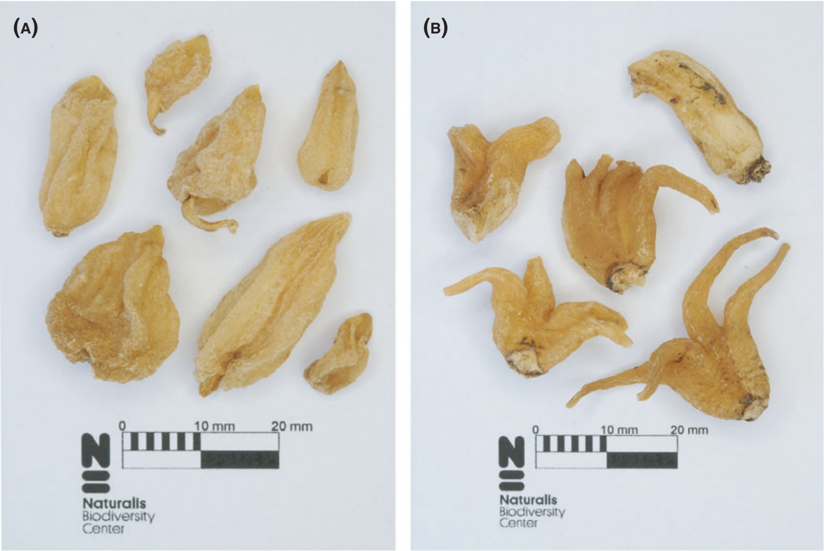

Fig. 2 Tuber samples of different morphology purchased from the markets. (A) Samples of *Orchis*/*Anacamptis* type tubers. (B) Samples of *Dactylorhiza* type tubers. [Colour figure can be viewed at wileyonlinelibrary.com].

F2_F (5′-CTAATACCCCATCCCATCCAT-3′) (Steele & Vilgalys 1994) and R1_R (5′-CATTTTTCATTGCACACGRC-3′) (Kocyan *et al.* 2004). PCR amplification was performed in a 50-μL reaction volume containing 5 μL reaction buffer IV (10×), 5 μL $MgCl_2$ (25 mM), 1 μL dNTP (10 μM), 0.25 μL Taq-polymerase (5 U/μL), 05 μL BSA, 1 μL of each primer (10 mM) and 1 to 4 μL of template DNA. The PCR protocols of 95 °C 3 min, (95 °C 20 s, 55 °C 1 min, 72 °C 2 min) × 35, 72 °C 10 min, 8 °C ∞ for nrITS, 95 °C 3 min, (95 °C 15 s, 55 °C 50 s, 72 °C 4 min) × 35, 72°C 8 min, 8 °C ∞ for *trn*L-F spacer and 95 °C 3 min, (95 °C 34 s, 59 °C 45 s, 72 °C 1 min) × 35, 72 °C 7 min, 8 °C ∞ for *mat*K were applied. Sanger sequencing was performed by Macrogen Europe Inc. (Amsterdam, the Netherlands) on an ABI3730XL automated sequencer (Applied Biosystems). Primers used for PCR amplification were also used for sequencing reactions.

Reference database preparation

The reference database was compiled from a total of 490 source sequences of 133 taxa, including both voucher specimens collected from the field including 85 nrITS sequences (19 species), 90 *trn*L-F (26 species), 63 *mat*K (20 species) and publicly available DNA sequences from NCBI GenBank including 126 nrITS sequences (102 species), 68 *trn*L-F (56 species) and 58 *mat*K (55 species) (Table 1). All sequences were downloaded from the listed tuberous genera in the tribe Orchideae (Orchidaceae), including synonymous genera and/or species: *Anacamptis* Rich., *Cephalanthera* Rich., *Chamorchis* Rich., *Dactylorhiza* Neck. ex Nevski (including *Coeloglossum* Hartm.), *Gennaria* Parl., *Gymnadenia* R.Br., *Himantoglossum* W.D.J.Koch (incl. *Barlia* Parl. and *Comperia* K.Koch), *Neotinea* Rchb.f., *Neottia* Guett. (incl. *Listera* R.Br.), *Neottianthe* Schltr., *Ophrys* L., *Orchis* L. (incl. *Aceras* R.Br.), *Serapias* L. *Limodorum* Boehm., *Platanthera* Rich., and *Steveniella* Schltr. Representative accessions were included for nontuberous genera and tuberous species occurring close to the study area: *Corallorhiza trifida* Châtel, *Epipactis helleborine* (L.) Crantz, *Goodyera repens* (L.) R.Br., *Habenaria macroceratitis* Willd., *Herminium monorchis* (L.) R.Br., *Pecteilis gigantea* (Sm.) Raf., *Peristylus densus* (Lindl.) Santapau & Kapadia, *Pseudorchis albida* (L.) Á.Löve & D.Löve, *Satyrium bicorne* (L.) Thunb., *Spiranthes aestivalis* (Poir.) Rich., *Spiranthes spiralis* (L.) Chevall. and *Zeuxine strateumatica* (L.) Schltr (Appendix S2, Supporting information). Where there were more than two accessions per marker per species, only two accessions were selected, giving priority to those accessions with associated vouchers plus optimal read length and quality. Representative accessions of *Brownleea parvflora* Harv. ex Lindl., *Disa uniflora* P.J.Bergius and *Disperis lindleyana* Rchb.f. were selected as outgroups based on Inda *et al.* (2012).

Data analysis

Contigs were assembled and edited in SEQTRACE (Stucky 2012). All sequences including reference sequences and query tuber sequences were aligned using MUSCLE (Edgar 2004) as implemented in ALIVIEW v. 1.15 aligner (Larsson 2014). Final manual inspections were performed and adjustments were carried out if necessary. Sequences generated for this study were submitted to NCBI GenBank (Appendices S1, S3, Supporting information).

Bayesian inference (BI) and maximum-likelihood (ML) analysis were performed for each marker separately and on concatenated datasets, using RAXML-HPC v.8

Table 1 Species and samples per genus in sequence reference library

	Identification reference resource			
	GenBank		Field collections	
Genus	# Samples	# Species	# Samples	# Species
Anacamptis	6	5	17	3
Brownleea	1	1	NA	NA
Cephalantera	3	3	10	5
Chamorchis	1	1	NA	NA
Corallorhiza	1	1	NA	NA
Dactylorhiza	20	18	17	3
Disa	1	1	NA	NA
Disperis	1	1	NA	NA
Epipactis	1	1	7	2
Gennaria	1	1	NA	NA
Goodyera	1	1	NA	NA
Gymnadenia	8	7	NA	NA
Habenaria	1	1	NA	NA
Herminium	1	1	NA	NA
Himantoglossum	10	10	7	2
Limodorum	NA	NA	2	1
Neotinea	5	4	NA	NA
Neottia	3	3	3	1
Neottianthe	3	2	NA	NA
Ophrys	41	23	35	4
Orchis	8	7	21	4
Pecteilis	1	1	NA	NA
Peristylus	1	1	NA	NA
Platanthera	2	2	5	2
Pseudorchis	1	1	NA	NA
Satyrium	1	1	NA	NA
Serapias	8	7	NA	NA
Spiranthes	2	2	NA	NA
Steveniella	1	1	3	1
Zeuxine	1	1	NA	NA

Table 2 Sequence matrix and Bayesian analysis data

	No. of sequences			
Markers	Reference (R)	Leaf (L)	Tuber (T)	Total
nrITS	126	85	104	315
*trn*L-F	68	90	94	252
*mat*K	58	63	28	149
Concatenated	138	135	150	423

	Alignment matrix		Bayesian analysis
Markers	Seq length incl. gaps (bp)	Min/max length without gaps (bp)	No. trees retained
nrITS	822	209/722	30 002
*trn*L-F	1663	287/1032	30 002
*mat*K	1173	365/1105	15 002
Concatenated	3658	209/2677	30 002

(Stamatakis 2014) and MRBAYES v.3.2.2 (Ronquist *et al.* 2012) on CIPRES SCIENCE GATEWAY v.3.3 (Miller *et al.* 2010) and the high-performance computing facility available at University of Oslo, Lifeportal (https://lifeportal.uio.no/root). Gaps were treated as missing data.

For Bayesian analyses, the model GTR + G was selected for all datasets. Two independent runs with sixteen MCMC chains were simultaneously performed for 20 million rearrangements initiated with a random starting tree, and sampling one tree every 1000 generations, except for *mat*K. For *mat*K, we performed eight MCMC chains and a total of 10 million generations using the default heating temperature. Convergence of runs with default parameters was assessed on preliminary analyses. Where convergence did not occur, the heating parameter was adjusted to reach a convergence. Convergence of runs was assessed using TRACER v. 1.6 (Rambaut *et al.* 2014). Twenty-five percent of trees were discarded as burn-in, and the remaining trees were used to generate a consensus tree with Bayesian posterior probabilities (PP) values. Only PP values over 0.95 were considered and included for each marker and concatenated topologies. The number of trees retained for each analysis is presented in Table 2.

For maximum-likelihood analyses with RAXML, the model GTR + G was selected for all datasets, and a rapid bootstrap analysis with 1000 trees was conducted. Single marker trees were compared for incongruence prior to concatenation. Datasets were concatenated using GENEIOUS v. 6.1.8 (Kearse *et al.* 2012). Multiple GenBank reference sequences for a single species were merged to obtain one consensus species sequence (cf. Appendix S2, Supporting information). The unlink option was used to estimate the parameters for each partition.

The BI and ML phylogenetic trees were used to identify the query tubers (Appendices S4–S11, Supporting information). The tubers were considered successfully identified to species level when they were monophyletically clustered with related individuals of the same species. When tubers were clustered with individuals of different species of the same genus, only a genus level identification was assigned (Appendix S12, Supporting information).

Sequence similarity search using Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990) is often used in DNA barcoding (Little & Stevenson 2007; Sass *et al.* 2007; Kool *et al.* 2012; De Boer *et al.* 2014). BLAST+ (Camacho *et al.* 2009) features implemented in NCBI BLAST were used to query unknown tuber sequences against the compiled reference database. All top hits <15 points lower than the max score were considered for identification: if the retained top hits (max score −15 points) included only a single species then a species level identification was estimated; if the retained top hits (max score ×15 points) included multiple species in the same genus then a genus level identification was estimated; if the retained top hits (max score −15 points) included multiple species in different genera then a family level identification was estimated (Appendix S12, Supporting information).

Final consensus identifications were made based on the results from all markers and methods, BLAST, ML and BI (Appendix S12, Supporting information). Species level identification was assigned if all markers with species level identifications yielded the same species identification. Genus level identification was assigned if identifications resulted in multiple species of the same genus.

Results and discussion

Amplification and sequencing success

Sequencing success rates were different for reference samples and market tuber samples. For the reference leaf samples (L), sequencing success was 67% (85 samples) for nrITS, 71% (90) for the *trn*L-F spacer and 47% (63) for *mat*K (Table 2). Of the 127 samples, all three markers could be sequenced for 34 samples, solely nrITS for 29, solely *trn*L-F spacer for 14 and solely *mat*K for 7. For tuber samples (T), sequencing success was 69% for nrITS (104 samples), 63% for the *trn*L-F spacer (94) and 19% for *mat*K (28) (Table 2). Of 150 tuber samples, all three markers could be sequenced for eight samples, solely nrITS for 53, solely *trn*L-F spacer for 29 and none for *mat*K only. In general, low sequencing success might be due to degraded DNA as a result of boiling and drying the tubers during processing. Sequencing success for nrITS might be affected by fungal contamination during the drying process and orchid mycorrhizal associations producing a mix of plant and fungal nrITS sequences. *Mat*K had the lowest amplification success, and it has been shown that this locus cannot be amplified with 'universal' orchid primers due to the presence of alternative translation initiation codons in orchids (Barthet *et al.* 2015), and therefore requires 'case by case' optimization for each genus.

Species identifications

The similarity-based approach using BLAST using nrITS marker data identified 59 of 104 tuber samples (57%) to genus level and 45 (43%) to species level. Using *trn*L-F spacer, 61 of 94 tuber samples (65%) were identified to genus level and 33 (35%) to species level. Using *mat*K, 11 of 28 tuber samples (39%) were identified to genus and 17 (61%) to species level. The consensus of the BLAST identification of the three markers resulted in genus level identification in 93 samples (62%) and species level in 57 samples (38%) (Table 3; Appendix S12, Supporting information).

The tree-based approach using RAXML maximum-likelihood using nrITS marker data identified 34 of 104 tubers (33%) to genus level and 70 (67%) to species level (Appendices S4, S12, Supporting information). Using *trn*L-F spacer, 39 of 94 samples (42%) were identified to genus level and 55 (58%) to species level (Appendices S5, S12, Supporting information). Using *matK*, 12 of 28 tuber samples (43%) were identified to genus and 16 (57%) to species level (Appendices S6, S12, Supporting information). Concatenated data identified 87 samples (58%) to genus level and 63 (42%) to species level (Appendices S7, S12, Supporting information). The ML consensus identification of the three markers identified 60 samples (40%) to genus level and 90 samples (60%) to species level (Table 3; Appendix S12, Supporting information).

Table 3 Molecular identification of Salep tuber to species and genus level

	Samples for which sequences were obtained									
	ITS		*trn*L-F		*mat*K		Concatenated		Consensus	
Sequenced samples	104		94		28		150		150	
	Similarity (BLAST) identification									
Species	45	43%	33	35%	17	61%	–	–	57	38%
Genus	59	57%	61	65%	11	39%	–	–	93	62%
	Maximum-likelihood (RAXML) identification									
Species	70	67%	55	59%	16	57%	63	42%	90	60%
Genus	34	33%	39	41%	12	43%	87	58%	60	40%
	Bayesian inference (MRBAYES) identification									
Species	71	68%	55	59%	19	68%	102	68%	97	65%
Genus	33	32%	39	41%	9	32%	48	32%	53	35%

The tree-based approach using MRBAYES Bayesian inference using nrITS marker data identified 33 of 104 tubers (32%) to genus level and 71 (68%) to species level (Appendices S8, S12, Supporting information). Using *trn*L-F spacer, 39 of 94 samples (42%) were identified to genus level and 55 (58%) to species level (Appendices S9, S12, Supporting information). Using *mat*K, nine of 28 tuber samples (32%) were identified to genus and 19 (68%) to species level (Appendices S10, S12, Supporting information). Concatenated data identified 48 samples (32%) to genus level and 102 (68%) to species level (Appendices S11, S12, Supporting information). The BI consensus identification of the three markers identified 53 samples (35%) to genus level 97 samples (65%) to species level (Table 3; Appendix S12, Supporting information).

The final identification that combines consensus identification results of ML, BI and BLAST approaches produced an identification of 49 tubers (32.7%) to the genus level and 101 (67.3%) to the species level (Appendix S12, Supporting information).

Species composition of Salep

Similarity-based identifications using BLAST showed that *Orchis* (51 samples), *Anacamptis* (40 samples), *Dactylorhiza* (29 samples), *Ophrys* (18 samples), *Himantoglossum* (11 samples) and *Steveniella* (1 sample) were constituents of the studied Salep samples from Iran (Appendix S12, Supporting information). *Orchis simia* Lam. and *O. mascula* (L.) L. were the main *Orchis* species in Salep. *Anacamptis pyramidalis* (L.) Rich., *A. coriophora* (L.) R.M.Bateman, Pridgeon & M.W.Chase and *A. palustris* (Jacq.) R.M.Bateman, Pridgeon & M.W.Chase were the main *Anacamptis* species. *Dactylorhiza umbrosa* (Kar. & Kir.) Nevski was the only identified *Dactylorhiza* species. However, 24 of 29 *Dactylorhiza* samples were identified only to genus level. It is known that *Dactylorhiza* has a dynamic system of hybridization and allopolyploidy formation (Hedrén *et al.* 2001, 2008). These allopolyploids show no clear genetic differentiations despite phenotypic differences (Balao *et al.* 2015), and it is therefore difficult to identify these samples to species level using the applied markers. Similarly, *Ophrys* was found to be one of the constituents of Salep but discerning the species used as Salep with the BLAST similarity search was not possible. Species delimitation in closely related taxa of the genus *Ophrys* has been challenging because of continuous introgression and absence of complete lineage sorting (Devey *et al.* 2008).

Tree-based identifications using ML and BI showed similar results: *Orchis* (51 samples for ML and 52 for BI), *Anacamptis* (40 ML; 39 BI), *Dactylorhiza* (29 ML; 29 BI), *Ophrys* (18 ML; 18 BI), *Himantoglossum* (11 ML; 11 BI) and *Steveniella* (1 ML; 1 BI) were the constituents of Salep (Appendix S12, Supporting information). *Anacamptis* species in Salep samples are *A. palustris*, *A. morio* (L.) R.M.Bateman, Pridgeon & M.W.Chase, *A. pyramidalis* and *A. coriophora*. *Orchis* species contributing to Salep are *O. mascula*, *O. militaris* L. and *O. simia*. It was not possible to identify *Ophrys* and *Dactylorhiza* samples to species level using the applied markers.

Figure 3 shows the species composition of studied Salep tubers based on final consensus identifications including all markers and methods (Appendix S12, Supporting information). The phylogenetic relationships among genera is based on Inda *et al.* (2012). Based on final identification results the genera *Orchis* (51 samples), *Anacamptis* (40 samples), *Dactylorhiza* (29 samples), *Ophrys* (18 samples), *Himantoglossum* (11 samples) and *Steveniella* (1 sample) are the main the constituents of studied Salep samples. All tuberous orchid species are used for Salep with a preference for species in the genera *Orchis*, *Anacamptis* and *Dactylorhiza*.

Generic composition of tubers per geographic origin

The analyzed tubers can be geographically categorized into three zones of origin: a western zone (Ardabil, Eastern and Western Azarbaijan, Kurdistan and Kermanshah provinces), a northern zone (Golestan) and a central zone (Tehran and Esfahan). Sixty-five tubers originate from the western zone, and these include 26 tubers (38%) of *Anacamptis*, 22 tubers (32%) of *Dactylorhiza* and 11 tubers (16%) of *Himantoglossum*. The generic composition of the 66 tubers from the northern zone is different, and these include 42 tubers (64%) of *Orchis* and 15 tubers (23%) of *Ophrys*. The 15 tubers from the central zone are mainly *Anacamptis* (8 samples, 53%) and *Dactylorhiza* (5 samples, 33%). Although distribution and abundance of orchids in Iran is poorly documented, the results show that *Dactylorhiza* tubers, that trade at a lower value in the market, are harvested in the western and central zones, whereas high-value *Orchis* tubers are most commonly collected in the northern zone. Kasparek & Grimm (1999) report the presence of Iranian Salep in eastern Turkey in the 1990s, and Ghorbani *et al.* (2014a,b) writes that orchid tuber collection in western Iran has a longer history than in the north and east of Iran, where a recent boom is escalating harvesting and trade. The results could indicate that the resources for superior quality Salep tubers from *Orchis* species have been depleted in the western zone, and that Salep collection is now targeting the more inferior quality *Dactylorhiza* tubers. In the northern zone, *Orchis* tubers are still readily available, but as natural populations dwindle collectors will target other genera.

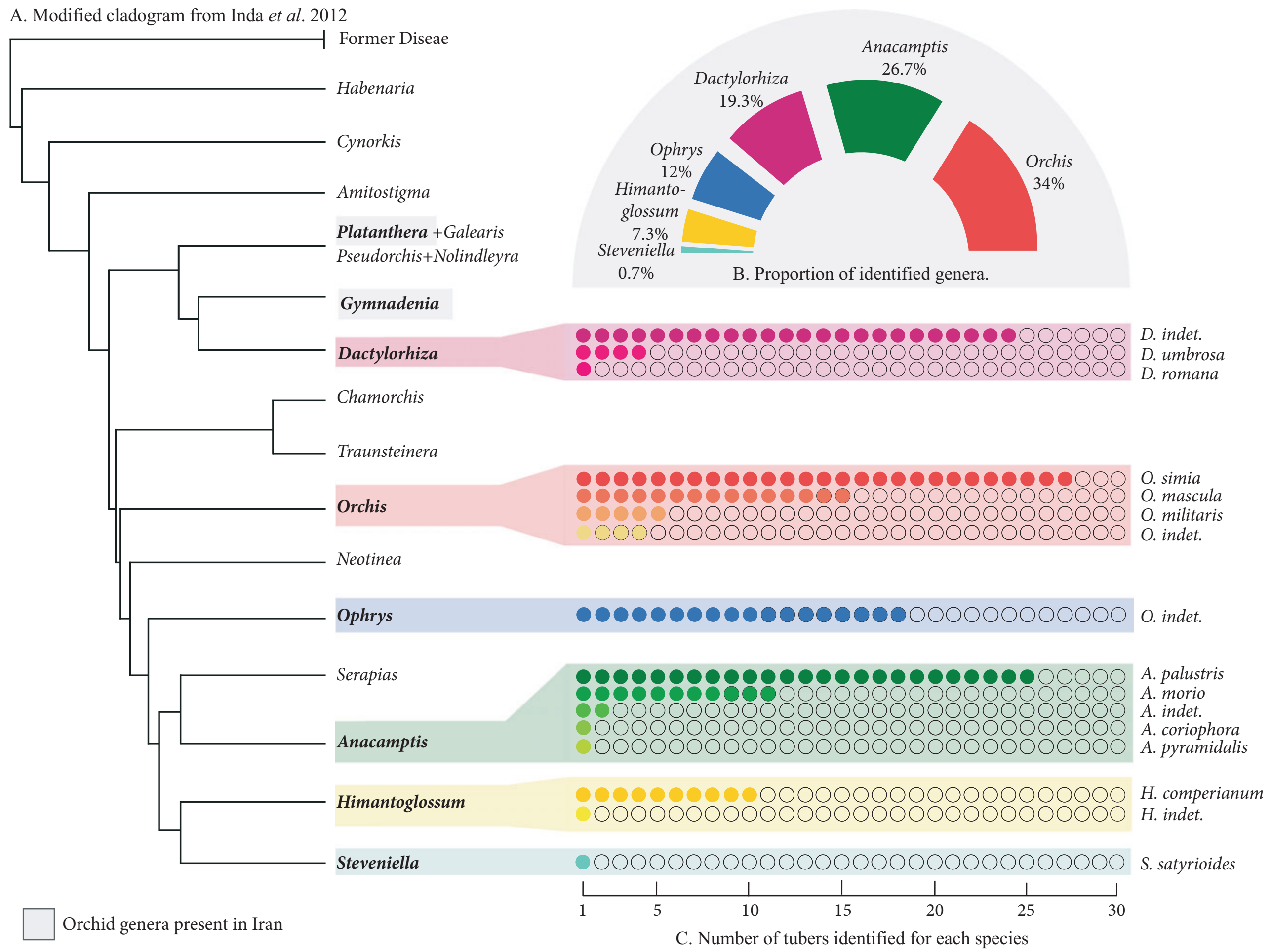

Fig. 3 Identifications of screened Iranian Salep tubers. (A) Shaded genera occur in Iran. Phylogeny adapted from Inda et al. 2012. (B) Proportion of identified genera. (C) Filled circles represent the number of tubers identified to a particular species. [Colour figure can be viewed at wileyonlinelibrary.com].

Conclusions

This study has produced a resource of 238 reference sequences and 226 tuber sequences that can be used for identification of Orchidaceae species in the poorly documented Salep trade in Turkey, Greece and Albania. It also shows that genomic DNA of sufficient quality can be extracted and sequenced from highly processed Salep tubers. However, extraction of DNA is accompanied with some difficulties as a result of gel formation due to the high glucomannan content in the tubers. Postharvest storage time of the tubers and boiling time during processing may also affect the quality of extracted DNA. Among the applied markers, nrITS and *trn*L-F spacer were easier to amplify and sequence than *mat*K, and these markers also show a higher discriminatory power for most of the genera. However, *Dactylorhiza* and *Ophrys*, that are known for allopolyploidy and hybridization, are challenging for barcoding using the applied markers, and a high-throughput sequencing gene capture approach would probably yield the right read depth for phasing of alleles and accurate species identification (Weitemier *et al*. 2014; Schmickl *et al*. 2015). The results also show that the genera most affected by Salep harvesting are *Orchis*, *Anacamptis*, *Dactylorhiza* and *Ophrys*. Geographic clustering of Salep tubers show clear differences in generic composition per zone with significant implications for harvesting pressure and resource depletion. *Dactylorhiza* and *Anacamptis* are more abundant as Salep tubers from the western zone, whereas *Orchis* and *Ophrys* are more abundant as Salep tubers from the northern zone. *Himantoglossum* was only present in Salep from the western zone. The results expose the overharvested species in each region that should be targeted for tailored conservation activities, and confirm the finding by Ghorbani *et al.* (2014a) that overharvesting of superior value *Orchis* tubers in western parts has led Salep middlemen and traders to tap into new areas in northern parts of the country. Conservation measures should be implemented in western, central and northern Iran to protect wild orchid populations from immediate threats due to unsustainable over-exploitation and to prevent their disappearance before many of them have even been studied properly.

Acknowledgements

Vincent Manzanilla is acknowledged for the graphic design and creation of Fig. 3. Caroline Nilsson is acknowledged for the graphic design and creation of Fig. 1. Financial support from the Swedish Research Link program of the Swedish Science Council is gratefully acknowledged. The Carl Tryggers Foundation for Scientific Research provided a postdoctoral research stipend to AG (through HdB). Further support for fieldwork was gratefully received from the Anne S. Chatham fellowship of the Garden Club of America (AG), Sven och Dagmar Saléns stiftelse (HdB, AG) and Helge Ax:son Johnsons Stiftelse (HdB).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Aubriot X, Lowry PP, Cruaud C, Couloux A, Haevermans T (2013) DNA barcoding in a biodiversity hot spot: potential value for the identification of Malagasy *Euphorbia* L. listed in CITES Appendices I and II. *Molecular Ecology Resources*, **13**, 57–65.

Babbar SB, Singh DK (2016) Protocols for In Vitro Mass Multiplication and Analysis of Medicinally Important Phenolics of a Salep Orchid, Satyrium nepalense D. Don ("Salam Mishri"). In: *Protocols for In Vitro Cultures and Secondary Metabolite Analysis of Aromatic and Medicinal Plants*, 2nd edn (ed Mohan Jain S.), pp. 1–11. Springer Protocols, Springer, US.

Balao F, Tannhäuser M, Lorenzo MT, Hedrén M, Paun O (2015) Genetic differentiation and admixture between sibling allopolyploids in the *Dactylorhiza majalis* complex. *Heredity*, **116**, 351–361.

Barthet MM, Moukarzel K, Smith KN, Patel J, Hilu KW (2015) Alternative translation initiation codons for the plastid maturase MatK: unraveling the pseudogene misconception in the Orchidaceae. *BMC Evolutionary Biology*, **15**, 210.

Bellstedt DU, Linder HP, Harley EH (2001) Phylogenetic relationships in *Disa* based on non-coding *trnL-trnF* chloroplast sequences: evidence of numerous repeat regions. *American Journal of Botany*, **88**, 2088–2100.

Bulpitt CJ (2005) The uses and misuses of orchids in medicine. *QJM*, **98**, 625–631.

Bulpitt CJ, Li Y, Bulpitt PF, Wang J (2007) The use of orchids in Chinese medicine. *Journal of the Royal Society of Medicine*, **100**, 558–563.

Burgess KS, Fazekas AJ, Kesanakurti PR *et al.* (2011) Discriminating plant species in a local temperate flora using the *rbcL+matK* DNA barcode. *Methods in Ecology and Evolution*, **2**, 333–340.

Camacho C, Coulouris G, Avagyan V *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.

Cameron KM (2004) Utility of plastid *psaB* gene sequences for investigating intrafamilial relationships within Orchidaceae. *Molecular Phylogenetics and Evolution*, **31**, 1157–1180.

Chinsamy M, Finnie JF, Van Staden J (2011) The ethnobotany of South African medicinal orchids. *South African Journal of Botany*, **77**, 2–9.

CITES (2014) The Convention on International Trade in Endangered Species of Wild Fauna and Flora Appendices.

Coghlan ML, Haile J, Houston J *et al.* (2012) Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. *PLoS Genetics*, **8**, e1002657.

De Boer HJ, Ouarghidi A, Martin G, Abbad A, Kool A (2014) DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. *PLoS ONE*, **9**, e84291.

De Vere N, Rich TC, Ford CR *et al.* (2012) DNA barcoding the native flowering plants and conifers of Wales. *PLoS ONE*, **7**, e37945.

Deguilloux M-F, Pemonge M-H, Petit RJ (2002) Novel perspectives in wood certification and forensics: dry wood as a source of DNA. *Proceedings of the Royal Society B*, **269**, 1039–1046.

Delforge P (2006) *Orchids of Europe, North Africa and the Middle East*. A&C Black, London, UK.

Devey DS, Bateman RM, Fay MF, Hawkins JA (2008) Friends or relatives? Phylogenetics and species delimitation in the controversial European orchid genus *Ophrys*. *Annals of Botany*, **101**, 385–402.

Dogan M, Kayacier A, Ic E (2007) Rheological characteristics of some food hydrocolloids processed with gamma irradiation. *Food Hydrocolloids*, **21**, 392–396.

Doyle J, Doyle J (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

Ebihara A, Nitta JH, Ito M (2010) Molecular species identification with rich floristic sampling: DNA barcoding the pteridophyte flora of Japan. *PLoS ONE*, **5**, e15136.

Ece Tamer C, Karaman B, Utku Copur O (2006) A traditional Turkish beverage: Salep. *Food Reviews International*, **22**, 43–50.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Fazekas AJ, Burgess KS, Kesanakurti PR *et al.* (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE*, **3**, e2802.

Ghorbani A, Gravendeel B, Naghibi F, de Boer HJ (2014a) Wild orchid tuber collection in Iran: a wake-up call for conservation. *Biodiversity and Conservation*, **23**, 2749–2760.

Ghorbani A, Gravendeel B, Zarre S, de Boer HJ (2014b) Illegal wild collection and international trade of CITES-listed terrestrial orchid tubers in Iran. *TRAFFIC Bulletin*, **26**, 52–58.

Gravendeel B, Chase MW, de Vogel EF *et al.* (2001) Molecular phylogeny of *Coelogyne* (Epidendroideae; Orchidaceae) based on plastid RFLPS, matK, and nuclear ribosomal ITS sequences: evidence for polyphyly. *American Journal of Botany*, **88**, 1915–1927.

Hajibabaei M, Singer GA, Hebert PD, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**, 167–172.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings. Biological Sciences/The Royal Society*, **270**, 313–321.

Hedrén M, Fay MF, Chase MW (2001) Amplified fragment length polymorphisms (AFLP) reveal details of polyploid evolution in *Dactylorhiza* (Orchidaceae). *American Journal of Botany*, **88**, 1868–1880.

Hedrén M, Nordström S, Ståhlberg D (2008) Polyploid evolution and plastid DNA variation in the *Dactylorhiza incarnata/maculata* complex (Orchidaceae) in Scandinavia. *Molecular Ecology*, **17**, 5075–5091.

Hibert F, Taberlet P, Chave J *et al.* (2013) Unveiling the diet of elusive rainforest herbivores in next generation sequencing era? The tapir as a case study. *PLoS ONE*, **8**, e60799.

Hossain MM (2011) Therapeutic orchids: traditional uses and recent advances—an overview. *Fitoterapia*, **82**, 102–140.

Inda LA, Pimentel M, Chase MW (2012) Phylogenetics of tribe Orchideae (Orchidaceae: Orchidoideae) based on combined DNA matrices: inferences regarding timing of diversification and evolution of pollination syndromes. *Annals of Botany*, **110**, 71–90.

Kasparek M, Grimm U (1999) European trade in Turkish Salep with special reference to Germany. *Economic Botany*, **53**, 396–406.

Kearse M, Moir R, Wilson A *et al.* (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)*, **28**, 1647–1649.

Kim HM, Oh S-H, Bhandari GS, Kim C-S, Park C-W (2014) DNA barcoding of Orchidaceae in Korea. *Molecular Ecology Resources*, **14**, 499–507.

Kocyan A, Qiu Y-L, Endress PK, Conti E (2004) A phylogenetic analysis of Apostasioideae (Orchidaceae) based on ITS, trnL-F and matK sequences. *Plant Systematics and Evolution*, **247**, 203–213.

Kool A, de Boer HJ, Krüger A *et al.* (2012) Molecular identification of commercialized medicinal plants in southern Morocco. *PLoS ONE*, **7**, e39459.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences*, **102**, 8369–8374.

Kreziou A, de Boer H, Gravendeel B (2015) Harvesting of salep orchids in north-western Greece continues to threaten natural populations. *Oryx*, **50**, 393–396.

Kuzmina ML, Johnson KL, Barron HR, Hebert PD (2012) Identification of the vascular plants of Churchill, Manitoba, using a DNA barcode library. *BMC Ecology*, **12**, 25.

Landerer X (1850) Beiträge zur Pharmakognosie. *Archiv der Pharmazie*, **112**, 177–180.

Larsson A (2014) ALIVIEW: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, **30**, 3276–3278.

Little DP, Stevenson DW (2007) A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. *Cladistics*, **23**, 1–21.

Liu JIE, Moeller M, Gao LM, Zhang D, Li DZ (2011) DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. *Molecular Ecology Resources*, **11**, 89–100.

Liu J, Yan H-F, Newmaster SG *et al.* (2015) The use of DNA barcoding as a tool for the conservation biogeography of subtropical forests in China. *Diversity and Distributions*, **21**, 188–199.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Gateway Computing Environments Workshop (GCE), 2010*.

Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. *BMC Medicine*, **11**, 222.

Ogden R, Dawnay N, McEwing R (2009) Wildlife DNA forensics—bridging the gap between conservation genetics and law enforcement. *Endangered Species Research*, **9**, 179–195.

Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB (2012) DNA barcoding of endangered Indian Paphiopedilum species. *Molecular Ecology Resources*, **12**, 82–90.

Posadzki P, Watson L, Ernst E (2012) Contamination and adulteration of herbal medicinal products (HMPs): an overview of systematic reviews. *European Journal of Clinical Pharmacology*, **69**, 295–307.

Rambaut A, Suchard M, Xie D, Drummond J (2014) TRACER v1.6.

Renz Z (1978) *Flora Iranica. Part 126: Orchidaceae*. Naturhistorisches Museum, Vienna, Austria.

Richardson RT, Lin C-H, Sponsler DB *et al.* (2015) Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Applications in Plant Sciences*, **3**, 1400066.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MRBAYES 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Sandal Erzurumlu G, Doran I (2011) Türkiye'de salep orkideleri ve salep kültürü. *Journal of the Agricultural Faculty of the University of Harran*, **15**, 29–34.

Särkinen T, Staats M, Richardson JE, Cowan RS, Bakker FT (2012) How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE*, **7**, e43808.

Sass C, Little DP, Stevenson DW, Specht CD (2007) DNA barcoding in the cycadales: testing the potential of proposed barcoding markers for species identification of cycads. *PLoS ONE*, **2**, e1154.

Schmickl R, Liston A, Zeisek V *et al.* (2015) Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Molecular Ecology Resources*, **16**, 1124–1135.

Sezik E (2002a) Turkish orchids and Salep. *Acta Pharmaceutica Turcica*, **44**, 151–157.

Sezik E (2002b) Destruction and Conservation of Turkish Orchids. In: *Biodiversity* (ed Şener B.), pp. 391–400. Kluwer Academic/Plenum Publishers, New York.

Sezik E (2006) Destroying of Ophrys species to obtain Salep in turkey. *Journal Europäischer Orchideen*, **38**, 290.

Shahsavari A (2008) *Flora of Iran. Part 57: Orchidaceae*. Research Institute of Forests and Rangelands, Tehran, IR Iran.

Shepherd LD, McLay TGB (2011) Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research*, **124**, 311–314.

Stamatakis A (2014) RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

Starin D (2012) Salepi extinction, Salepi survival: how a change in ingredients could help safeguard orchids. *Orchids: the Bulletin of the American Orchid Society*, **81**, 490–494.

Steele KP, Vilgalys R (1994) Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene matk. *Systematic Botany*, **19**, 126–142.

Stucky BJ (2012) SEQTRACE: a graphical tool for rapidly processing DNA sequencing chromatograms. *Journal of Biomolecular Techniques*, **23**, 90–93.

Sun Y, Skinner DZ, Liang GH, Hulbert SH (1994) Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics*, **89**, 26–32.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Tekinşen KK, Güner A (2010) Chemical composition and physicochemical properties of tubera salep produced from some Orchidaceae species. *Food Chemistry*, **121**, 468–471.

Thompson KA, Newmaster SG (2014) Molecular taxonomic tools provide more accurate estimates of species richness at less cost than traditional morphology-based taxonomic practices in a vegetation survey. *Biodiversity and Conservation*, **23**, 1411–1424.

Van Geel B, Aptroot A, Baittinger C *et al.* (2008) The ecological implications of a Yakutian mammoth's last meal. *Quaternary Research*, **69**, 361–376.

Van Velzen R, Weitschek E, Felici G, Bakker FT (2012) DNA barcoding of recently diverged species: relative performance of matching methods. *PLoS ONE*, **7**, e30490.

Vassou SL, Kusuma G, Parani M (2015) DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*. *Gene*, **559**, 86–93.

Weitemier K, Straub SC, Cronn RC *et al.* (2014) HYB-SEQ: combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, **2**, 1400042.

Williams VL, Balkwill K, Witkowski ETF (2000) Unraveling the commercial market for medicinal plants and plant parts on the Witwatersrand, South Africa. *Economic Botany*, **54**, 310–327.

Xiang X-G, Hu H, Wang W, Jin X-H (2011) DNA barcoding of the recently evolved genus Holcoglossum (Orchidaceae: Aeridinae): a test of DNA barcode candidates. *Molecular Ecology Resources*, **11**, 1012–1021.

A.G., Hd.B. and B.G. devised the project. A.G. carried out the vast majority of the fieldwork, assisted by S.Z., H.d.B. and B.G. A.G., S.S. and H.d.B. analyzed the data. A.G., S.S. and H.d.B. wrote the first draft of the manuscript. All authors have read and approve the final manuscript.

Data accessibility

The concatenated sequence matrix of all reference and tubers species, as well as resulting BI and ML phylogenetic trees (Appendices S4–S11, Supporting information) are deposited in Dryad, http://dx.doi.org/10.5061/dryad.qb36g.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Reference sequences derived from vouchers collected for this study.

Appendix S2 Reference sequences derived from external NCBI GenBank accessions.

Appendix S3 GenBank accession numbers of the Salep tubers.

Appendix S4 RAXML maximum-likelihood phylogenetic tree for nrITS.

Appendix S5 RAXML maximum-likelihood phylogenetic tree for *trn*L-F spacer.

Appendix S6 RAXML maximum-likelihood phylogenetic tree for *mat*K.

Appendix S7 RAXML maximum-likelihood phylogenetic tree for the concatenated matrix.

Appendix S8 MRBAYES bayesian phylogenetic tree for nrITS.

Appendix S9 MRBAYES bayesian phylogenetic tree for *trn*L-F spacer.

Appendix S10 MRBAYES bayesian phylogenetic tree for *mat*K.

Appendix S11 MRBAYES bayesian phylogenetic tree for the concatenated matrix.

Appendix S12 Molecular identifications of tubers based on similarity- and tree-based approaches.