Plant Syst Evol (2011) 293:213–226
DOI 10.1007/s00606-011-0439-x

ORIGINAL ARTICLE

Strong genetic differentiation between *Gymnadenia conopsea* and *G. densiflora* despite morphological similarity

Christiane Stark · Stefan G. Michalski · Wiesław Babik · Grit Winterfeld · Walter Durka

Received: 28 September 2010 / Accepted: 20 February 2011 / Published online: 27 March 2011

Abstract The fragrant orchid *Gymnadenia conopsea* s.l. is a controversial taxon with two commonly distinguished species, *G. conopsea* s.str. and *G. densiflora*. Despite morphological similarity, differentiation between the taxa has been reported for several characters; however, character variation within taxa has obviated a clear consensus. We assessed ITS sequences, microsatellite variation and chromosome numbers on the European scale (1,420 samples) and conducted morphological analyses for 626 samples from Germany. ITS analysis revealed a 2% nucleotide divergence between the taxa, similar to the divergence between other *Gymnadenia* species. The ITS sequences of *G. densiflora* form a well-supported monophyletic group sharing a most recent common ancestor with *G. nigra* and *G. austriaca*. Thus, *G. conopsea* and *G. densiflora* are not sister species, and a species rank is supported for *G. densiflora* (Wahlenb.) Dietrich and *G. conopsea* (L.) R.Br. s.str. This was confirmed by the microsatellite analysis, which revealed a strong genetic differentiation between the taxa because of largely non-overlapping sets of alleles. Chromosome numbers showed that *G. conopsea* was either diploid or tetraploid, whereas *G. densiflora* was diploid throughout. Morphologically, the taxa differed significantly in the mean value of a number of diagnostic characters. However, a discriminant analysis showed that the morphological variability is substantial, and on the individual level an unequivocal assignment is not possible as 96% of *G. conopsea*, but only 77% of *G. densiflora* could be assigned correctly. Further studies are needed on character variation within and among species and ploidy levels to allow for a better identification of the genetically differentiated but morphologically similar taxa.

Keywords *Gymnadenia conopsea* · Polyploidy · ITS sequences · Microsatellites · Genetic differentiation · Sibling species

Electronic supplementary material The online version of this article (doi:10.1007/s00606-011-0439-x) contains supplementary material, which is available to authorized users.

C. Stark (✉) · S. G. Michalski · W. Durka
Department of Community Ecology,
Helmholtz Centre for Environmental Research-UFZ,
Theodor-Lieser Str. 4, 06120 Halle, Germany
e-mail: christiane.stark@ufz.de

W. Babik
Institute of Systematics and Evolution of Animals,
Polish Academy of Sciences, Slawkowska 17,
31-016 Krakow, Poland

G. Winterfeld
Department of Systematic Botany,
Martin-Luther University of Halle-Wittenberg,
Neuwerk 21, 06108 Halle, Germany

Introduction

Orchids (Orchidaceae) form one of the largest families of flowering plants, harboring approximately 25,000 species (Dressler 1987). As such, they are a prime example of diversification, and with one of the highest speciation rates of all flowering plants (Gill 1989), they offer an extraordinary opportunity to study speciation (Peakall 2007). Major drivers for orchid diversification are thought to be their symbiotic interaction with fungi for germination (Otero and Flanagan 2006; Swarts and Dixon 2009) and their highly specialized pollinator mechanisms (Cozzolino and Widmer 2005a; Micheneau et al. 2009). Plant-pollinator interactions play a decisive role in orchid speciation, and therefore major distinctions are obvious between

specialized and generalized pollination systems and between rewarding and non-rewarding taxa (Salzmann et al. 2007). In highly specialized pollination systems like sexually deceptive orchids have, pollinator-driven selection is likely to be strongly directional, resulting in floral diversity (e.g., Juillet et al. 2010; Schiestl and Schlüter 2009). In contrast, in generalized pollination systems the sharing of different pollinators may exert stabilizing selection that constrains variation of floral morphology and retards speciation (Cresswell 1998; Vereecken et al. 2010).

Orchids are known for the ease with which inter-specific and inter-generic hybrids can occur as they do not have strong prezygotic reproductive barriers or incompatibilities (Cozzolino and Widmer 2005b; Gill 1989; Mallet 2005). However, due to strong postzygotic barriers, e.g., chromosomal incompatibilities (Scopece et al. 2008), hybridization does not necessarily lead to speciation. Still, in several orchid genera and most prevalent in the subtribe Orchidinae (Gill 1989) hybridization and polyploidy are common features. Thus, for example, the high diversification of the Iberian orchid flora is partly attributed to polyploidy, with species-rich groups such as *Dactylorhiza* showing particularly high proportions of intra-generic polyploidy (83.3%) (Amich et al. 2007). While autopolyploidy does occur, e.g., in *Dactylorhiza* (Devos et al. 2005; Ståhlberg 2009) or *Vanilla* (Bory et al. 2008), allopolyploidy seems to be more common and has been reported in several genera including, e.g., *Dactylorhiza* (Hedrén et al. 2001) and *Nigritella* (now included in *Gymnadenia*) (Hedrén et al. 2000).

Gymnadenia conopsea (L.) R.Br. s.l., the fragrant orchid, is a controversial taxon. *Gymnadenia* is an Eurasian genus, covering most of Europe and parts of Asia (Tutin et al. 1980), and *G. conopsea* is distributed from Western Europe to China. In Europe it is a still relatively common orchid with a wide ecological amplitude including forest habitats and wet to dry grasslands. *G. conopsea* s.l. is a rewarding taxon providing nectar to various diurnal and nocturnal lepidopteran pollinators. Morphologically, *G. conopsea* s.l. is highly variable, which gave rise to various taxonomic treatments, ranging from one species with three varieties (Delforge 2006) to five species plus two subspecies (Dworschak 2002). The two most commonly distinguished taxa, particularly in local Flora, are *G. conopsea* (L.) R.Br. ssp. *conopsea* and *G. conopsea* ssp. *densiflora* (Wahlenb.) K. Richt. Today the two taxa are treated as distinct species *G. conopsea* (L.) R.Br. s.str. and *G. densiflora* (Wahlenb.) Dietrich (cf. Marhold et al. 2005). They are described to differ in morphology, flowering phenology, scent emission and habitat preferences (Gustafsson 2000; Gustafsson and Lönn 2003; Jersáková et al. 2010; Marhold et al. 2005; Scacchi and de Angelis 1989; Soliva and Widmer 1999). However, due to a considerable morphological variability, a clear assignment to either taxon is often difficult in the field. Additionally, flowering time has been shown to depend on the ploidy level and does not allow a clear distinction either (Jersáková et al. 2010). Previous investigations revealed genetic differentiation at allozyme loci (Scacchi and de Angelis 1989), at the DNA sequence level (Gustafsson and Lönn 2003) and between flowering-time variants in Sweden (Gustafsson and Lönn 2003) and Switzerland (Soliva and Widmer 1999). Reports on the ploidy status are complex, with authors describing *G. conopsea* as tetraploid or diploid (Marhold et al. 2005; Vöth and Sontag 2006) and *G. densiflora* as diploid (Marhold et al. 2005) or tetraploid (Hagerup in Bisse 1963; Mrkvicka 1993). Referring to a basic chromosome number of $x = 10$ rather than $x = 20$, higher ploidy levels were deduced recently (Jersáková et al. 2010; Trávníček et al. 2011). Autopolyploidy has been postulated for *G. conopsea* (Jersáková et al. 2010); however, no clear evidence has been presented yet.

So far, previous studies inherently assumed that the two taxa are sister groups and compared, among others, morphology, chromosome counts, genetic variation and phylogenetic analyses of samples from various geographic regions. However, a comprehensive analysis of these parameters is missing on one sample. We analyzed 32 sites of *G. conopsea* and *G. densiflora* in East and North Germany based on morphological and genetic parameters, including chromosome counts, ITS sequences and microsatellite analyses. We found a strong genetic, but only moderate morphological differentiation between the taxa. In order to further analyze the relationships on a larger geographic scale, we included additional samples from other European regions. In particular, we tested if the differentiation between the taxa is consistent across regions and parameters analyzed and whether they in fact share a most recent common ancestor.

Materials and methods

Sampling

Phylogenetic and microsatellite analyses were conducted on a large geographic scale, including samples from East Germany (EG, 10 sites), North Germany (NG, 6 sites), Eifel (WG, 5 sites), Saarland (SWG, 4 sites), the Bavarian Alps (SEG, 8 morphological varieties sensu Dworschak 2002) and other European regions like Sweden (SE; Gotland and Öland, 5 sites), France (FR; the French Alps, 6 sites) and Austria (AT; Austrian Alps, 5 sites) (Table 1). Individuals were assigned to either *G. conopsea* or *G. densiflora* based on the ITS sequence and microsatellite data, which allowed an unequivocal distinction of the taxa (see below).

Table 1 Study sites, number of samples analyzed for ITS (N_{ITS}), microsatellites (N_{micro}) and morphology (N_{morph}), haplotypes and total number of alleles (A) detected at loci Gc42 and Gc51, ploidy level inferred from microsatellites and the taxon to which the populations were assigned

Study site	Code	E	N	N_{ITS}	Haplotypes	N_{micro}	A	Ploidy	N_{morph}	Taxon
East Germany (EG)										
Jaucha	EG01	12°11′	51°14′	2	Gd1	28	3	D*	20	Gd
Theißen	EG02	12°07′	51°08′	3	Gd1	30	5	D*	20	Gd
Predel	EG03	12°19′	51°12′	2	Gd1, Gd3	19	6	D*	20	Gd
Domsen	EG04	12°17′	51°18′	3	Gd1, Gd2	31	3	D	5	Gd
Espenhain	EG05	12°44′	51°25′	–		7	5	D	8	Gd
Tote Täler	EG06	11°74′	51°19′	3	Gd1, Gd2	27	8	D*	20	Gd
Rothenstein	EG07	11°57′	50°86′	–		27	9	D	20	Gd
Würze	EG08	11°69′	50°84′	7	Gd1, Gd2, Gd4, Gd6, Gd7, Gd8	29	8	D*	20	Gd
Klingelsteine	EG09	11°64′	50°98′	–		27	8	D*	20	Gd
Jägertalwiese	EG10	11°73′	50°97′	–		30	6	D*	20	Gd
Domsen	EG11	12°14′	51°19′	1	Gc9	26	9	P	20	Gc
Rothenstein	EG12	11°58′	50°86′	–		26	37	P	20	Gc
Zietschkuppe	EG13	11°70′	50°97′	–		23	30	P	20	Gc
Alter Gleisberg	EG14	11°70′	50°95′	–		25	39	P	20	Gc
Rabis	EG15	11°66′	50°89′	1	Gc20	25	33	P*	20	Gc
Krawinkel	EG16[a]	11°64′	51°21′	–		29	34	P*	20	Gc
Steigra	EG17[a]	11°65′	51°30′	–		26	35	P*	20	Gc
Grockstädt	EG18	11.59′	51°33′	2	Gc12, Gc26, Gc27	19	24	P	20	Gc
Langer Berg	EG19	11.71′	51°24′	2	Gc13, Gc18, Gc19	27	22	P	20	Gc
Tote Täler	EG20[a]	11°73′	51°19′	1	Gc28, Gc29	29	33	P	19	Gc
North Germany (NG)										
Stb. Polle	NG01[a]	9°40′	51°89′	1	Gc21, Gc22	25	35	P	20	Gc
Stb. “Im Schießstand”	NG02	9°36′	51°78′	–		27	38	P	20	Gc
Stb. “Alter Steinbruch”	NG03	9°36′	51°78′	–		25	35	P	20	Gc
Stb. Hehlen	NG04	9°45′	51°98′	2	Gc11, Gc21	25	25	P	20	Gc
Stb. Bärenbrink	NG05	9°88′	51°94′	1	Gc9, Gc10	26	33	P	14	Gc
Stb. Delligsen	NG06	9°81′	51°95′	–		27	33	P	20	Gc
Burgberg	NG07[a]	9°51′	51°87′	–		25	31	P	20	Gc
Rühle	NG08[a]	9°53′	51°92′	1	Gc12	24	37	P	20	Gc
Räuschenberg	NG09	9°37′	51°81′	–		30	18	P	20	Gc
Poppenburg	NG10	9°44′	51°88′	1	Gc11	25	28	P	20	Gc
Bielenberg	NG11	9°35′	51°78′	1	Gc9	27	37	P	20	Gc
Bocksberg	NG12	9°65′	51°88′	–		25	35	P	20	Gc
Holberg	NG13	9°57′	51°91′	–		22	32	P	20	Gc
Eifel (WG)										
Alenberg	WG1	6°38′	50°22′	3	Gd1, Gd2	35	3	D	–	Gd
Ripsdorfer Moor	WG2	6°39′	50°23′	1	Gd1	40	5	D	–	Gd
Hillesheim	WG3	6°40′	50°17′	3	Gd1, Gc9, Gc10	39	21	D + P	–	Mixed
Eierberg	WG4	6°38′	50°22′	1	Gc9	20	21	D	–	Gc
Höneberg	WG5	6°40′	50°23′	1	Gc9, Gc10	20	18	D	–	Gc
Saarland (SWG)										
Niedergailbach	SWG1	7°12′	49°08′	3	Gc9	40	24	P	–	Gc
Bliesransbach	SWG2	7°05′	49°09′	3	Gc13, Gc15, Gc19, Gc22, Gc23	40	26	P	–	Gc
Nachtweide	SWG3	7°04′	49°12′	3	Gd1, Gd2	40	10	D	–	Gd
Ensheim	SWG4	7°07′	49°12′	2	Gd1, Gd2	39	8	D	–	Gd

Table 1 continued

Study site	Code	E	N	N_{ITS}	Haplotypes	N_{micro}	A	Ploidy	N_{morph}	Taxon
Bavarian Alps (SEG)										
Bavaria-South	SEG1[c]	12°00′	47°40′[b]	2	Gd1, Gd3	9	7	D	–	Gd
Bavaria-South	SEG2[d]	12°00′	47°40′[b]	1	Gc9, Gc10	10	17	D + P	–	Gc
Bavaria-South	SEG3[e]	12°00′	47°40′[b]	2	Gc15, Gc17	12	24	D + P	–	Gc
Bavaria-South	SEG4[f]	12°00′	47°40′[b]	1	Gc9, Gc10	13	28	D + P	–	Gc
Bavaria-South	SEG5 [g]	12°00′	47°40′[b]	2	Gc9, Gc10, Gc14	10	24	D	–	Gc
Bavaria-South	SEG6 [h]	12°00′	47°40′[b]	2	Gc9, Gc14, Gc9, Gc10,	5	18	D + P	–	Gc
Bavaria-South	SEG7[i]	12°00′	47°40′[b]	2	Gc16	12	25	D + P	–	Gc
Bavaria-South	SEG8[j]	12°00′	47°40′[b]	2	Gc9, Gc24, Gc25	–	–	–	–	Gc
Sweden (SE)										
Öland	SE1	16°38′	56°40′	1	Gc9, Gc10	5	9	D	–	Gc
Gotland Lojsta	SE2	18°22′	57°18′	1	Gc9, Gc10	40	41	D	–	Gc
Gotland, Lickershamn	SE3	18°30′	57°49′	3	Gd1, Gd2, Gd3	6	6	D	–	Gd
Gotland, Häftings	SE4	18°39′	57°53′	2	Gd1, Gd2	3	6	D	–	Gd
Gotland, Kallgatburg	SE5	18°40′	57°42′	4	Gd1, Gd2	11	10	D	–	Gd
Austria (AT)										
Kärnten	AT1	13°25′	46°42′	2	Gd1, Gd2	20	27	D + P	–	Mixed
Tweng	AT2	13°36′	47°11′	2	Gd3, Gd5	2	2	D	–	Gd
Gstatterboden	AT3	14°37′	47°35′	1	Gc33, Gc34	9	21	D + P	–	Gc
Obertauern	AT4	13°33′	47°15′	3	Gd1, Gd2	4	4	D	–	Gd
Steinplatte	AT5	12°05′	51°23′	–		31	11	D	–	Gd
France (FR)										
Col du Lautaret	FR1	6°24′	45°02′	2	Gc9, Gc14	15	20	D + P	–	Gc
Col du Lautaret	FR2	6°24′	45°02′	3	Gd1, Gd3	15	7	D	–	Gd
Clairvaux les lacs	FR3	5°46′	46°34′	2	Gc9, Gc13	–			–	Gc
Clairvaux les lacs	FR4	5°46′	46°34′	2	Gd2, Gc30, Gc31	–			–	Mixed
Sailles-le-Haut	FR5	6°01′	45°21′	2	Gc9, Gc10, Gc32	11	17	D	–	Gc
Sailles-le-Haut	FR6	6°01′	45°21′	2	Gd1	15	4	D	–	Gd

D diploid, *P* polyploid, *indicates confirmation by chromosome counts (Table 3)

[a] Site codes sensu Stark et al. (2009): EG16 = E1, EG17 = E2, EG20 = E3, NG07 = N1, NG01 = N2, NG08 = N3

[b] Taxa sensu Dworschak (2002), exact sampling sites unknown: [c] *G. conopsea* × *serotina;*[d] *G. vernalis,* [e] *G. splendida*, [f] *G. splendida* × *odoratissima*, [g] *G. graminea*, [h] *G. conopsea*, [i] *G. alpina*, [j] *G. alpina* × *conopsea*

Detailed population genetic and morphometric analyses were conducted in the two study regions in East and North Germany (Table 1). Here, individuals from a total of 32 sites were characterized for microsatellites (843 samples), ploidy, which was assessed by direct counting (31 samples, 10 sites) and based on microsatellite phenotypes, and morphology (626 samples).

Molecular phylogenetic analysis

DNA sequences were obtained for the nuclear ITS region, including ITS1, the 5.8 s rDNA gene and ITS2 using the primers ITS1 and ITS4 (White et al. 1990). A 50-µl PCR reaction contained 5 µl 10× PCR buffer (Fermentas, St. Leon-Rot, Germany), 2 mM $MgCl_2$, 0.16 mM of each dNTP, 1 µM of each primer, 1 U of *Taq* (Fermentas) and 1–10 ng DNA. The cycling scheme was 95°C for 5 min, followed by 40 cycles at 95°C for 40 s, 57°C for 30 s, 72°C for 40 s and the final extension step at 72°C for 10 min. PCR products were purified with MinElute (Qiagen), and 40 µl was concentrated into 10 µl. PCR products were sequenced using the BigDye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems), using the primer ITS4 (White et al. 1990). A 10-µl reaction contained 1.75 µl SB buffer, 3.2 pmol Primer, 0.5 µl BDT and 1 µl template. The cycling scheme was 96°C for 1 min, followed by 40 cycles at 96°C for 10 s, 50°C for 5 s, and 60°C for 2 min. Samples were run on an ABI 3100 genetic analyzer. Sequences were aligned using the software SeqScape v2.6 (Applied Biosystems). In case of sequences containing ambiguous sites, haplotypes were reconstructed using the PHASE algorithm (Stephens et al. 2001) as implemented in

Table 2 Unique haplotypes, overall frequency and variable nucleotide positions found for *Gymnadenia conopsea* and *G. densiflora*. Altogether 35 haplotypes were obtained, 8 haplotypes in *G. densiflora* (Gd1 to Gd8) and 27 haplotypes in *G. conopsea* (Gc9 to Gc35)

		ITS1													ITS2													
Haplotype	*n*	42	53	56	58	75	86	94	106	119	189	198	213	240	241	426	438	448	458	466	478	487	489	552	568	592	598
Gd1	35	C	C	C	T	C	C	C	C	G	G	C	C	C	T	A	G	G	C	T	C	T	T	T	A	A	T
Gd2	21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
Gd3	5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
Gd4	5	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.
Gd5	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
Gd6[a]	1	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Gd7[a]	1	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
Gd8[a]	1	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	A	.	.	.	.	.	.	.	.	.
Gc9	22	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc10	10	.	.	.	.	T	T	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc11	3	.	.	T	.	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc12	3	.	.	T	.	T	.	.	.	A	A	A	G	.	.	.	.	.	T	C	.	C	G	A	.	T	A
Gc13	3	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	A	C	.	C	G	A	.	T	A
Gc14[a]	3	T	.	.	.	T	T	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc15	2	.	.	T	.	T	.	.	.	A	A	A	G	.	.	T	.	.	T	C	.	C	G	A	.	T	A
Gc16	1	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	T	T	A
Gc17	1	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	.	.	G	A	.	T	A
Gc18[a]	2	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	A	C	T	C	G	A	.	T	A
Gc19[a]	2	.	.	.	.	T	T	.	.	A	A	A	G	.	.	.	.	.	A	C	.	C	G	A	.	T	A
Gc20	1	.	.	.	C	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc21[a]	2	.	.	T	.	T	T	.	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc22[a]	2	.	.	T	.	T	T	.	.	A	A	A	G	.	.	T	.	.	T	C	.	C	G	A	.	T	A
Gc23	1	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	T	C	.	C	G	A	.	T	A
Gc24[a]	1	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	A	.	.	C	.	C	G	A	.	T	A
Gc25[a]	1	T	.	.	.	T	.	.	.	A	A	A	G	.	.	.	A	.	.	C	.	C	G	A	.	T	A
Gc26[a]	1	.	.	T	.	T	T	.	.	A	A	A	G	.	.	.	.	.	T	C	.	C	G	A	.	T	A
Gc27[a]	1	.	.	T	.	T	.	.	G	A	A	A	G	.	.	.	.	.	T	C	.	C	G	A	.	T	A
Gc28[a]	1	.	.	.	.	T	.	T	.	A	A	A	G	.	.	.	.	.	A	C	.	C	G	A	.	T	A
Gc29[a]	1	.	.	.	.	T	.	T	.	A	A	A	G	.	.	.	.	.	A	C	T	C	G	A	.	T	A
Gc30[a]	1	.	.	.	.	T	.	T	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc31[a]	1	.	.	.	.	T	T	T	.	A	A	A	G	.	.	.	.	.	.	C	.	C	G	A	.	T	A
Gc32[a]	1	.	.	.	.	T	T	.	.	A	A	A	G	.	G	.	.	.	.	C	.	C	G	A	.	T	A
Gc33[a]	1	.	.	T	.	T	.	.	.	A	A	A	G	.	.	.	.	.	A	C	.	C	G	A	.	T	A
Gc34[a]	1	.	.	T	.	T	.	.	.	A	A	A	G	.	.	.	.	.	A	C	T	C	G	A	.	T	A
Gc35[a]	1	.	.	.	.	T	.	.	.	A	A	A	G	.	.	.	.	.	.	C	T	C	G	A	.	T	A

[a] Solely as reconstructed haplotype (for explanatory details see "Materials and methods")

DnaSP (Librado and Rozas 2009). The PHASE algorithm resolves haplotypes based on the given genotypic information using a coalescent-based Bayesian method. A list of unique haplotypes was created with the program TCS (Clement et al. 2000). Sequences are deposited at GenBank (accession no. JF414017–JF414052).

Genetic divergence between taxa was based on *p*-distance, which is the proportion of total base pair differences between two sequences (Nei and Kumar 2000), with the computer software MEGA4 2003 (Tamura et al. 2007). For tree inference the alignment included all unique haplotypes (Table 2), published sequences of *G. conopsea* and *G. densiflora* as well as published sequences of other congeneric species and a sequence of *G. odoratissima* obtained from a sample from the Swiss Alps. A *Dactylorhiza incarnata* sequence was added as outgroup. Phylogenetic trees were obtained applying two criteria, maximum parsimony and maximum likelihood. Most parsimonious trees were inferred using MEGA4 with a heuristic search and initial trees produced by the random

addition option. Alignment gaps were removed from the data set before analysis. Preceding the maximum likelihood estimation, we selected the best-fit model of nucleotide substitution by running jModelTest (Posada 2008) with default options and using the Akaike Information Criterion. Using the alignment and the estimated model of nucleotide substitution (a simple Jukes and Cantor model) and a BIONJ starting tree, we inferred a phylogeny using PhyML version 3 (Guindon and Gascuel 2003). For both methods, reliability of the branching pattern was tested by bootstrapping the dataset 500 times.

Microsatellite analysis

Microsatellite data were obtained for a total of 1,414 individuals from 62 sites (main study regions EG and NG: 843 samples, 32 sites; additional regions: 571 samples, 30 sites, for details see Table 1). Total genomic DNA was extracted with the DNeasy 96 Plant Kit (Qiagen, Hilden, Germany). Ten microsatellite loci have been described for *G. conopsea* (Campbell et al. 2002; Gustafsson and Thorén 2001). However, only five loci gave repeatable and interpretable PCR products and were used for the further analysis: Loci Gc17 (fluorescent label PET), Gc42 (FAM) and Gc77 (VIC) were amplified in a multiplex reaction, whereas Gc49 (PET) and Gc51 (PET) were amplified separately. A 10-µl PCR reaction contained 5 µl Multiplex PCR Kit (Qiagen, Hilden, Germany), 1 pmol of each primer and 1–10 ng DNA. The cycling scheme was 3 min at 94°C for initial denaturation, followed by 35 cycles of 30 s at 94°, 30 s at respective annealing temperature and 45 s at 72°C, and ended by a final elongation time of 10 min. Samples were run on an ABI 3100 genetic analyser (Applied Biosystems, Darmstadt, Germany) and genotyped manually using GeneMapper 3.7 (Applied Biosystems, Darmstadt, Germany). We used principal component analysis (PCA) to display the relationship among populations based on allele frequencies, using loci Gc42 and Gc51, which were consistently amplified in all individuals. Two analyses were conducted, first with all samples and second for the 32 sites from East and North Germany. The PCA was performed with R (R Development Core Team 2009).

Determination of ploidy level

Chromosomes were counted for 31 representative individuals from 10 sites in East Germany. Young growing buds were placed into 8-hydroxyquinoline for 2 h to synchronize cell division. Fixation was for 3 h at room temperature with a 3:1 mixture of ethanol (99%) and pure acetic acid. Chromosome preparations of mitotic and meiotic states were made from enzyme-treated buds, as described by Schwarzacher et al. (1980). Tissue was squashed on a slide in a drop of 45% propionic acid with 2% carmine according to the protocol described by Winterfeld and Röser (2007). Chromosome numbers in orchids are notoriously difficult to determine because of the occurrence of aneuploidy (e.g., Bernardos et al. 2003; Bianco et al. 1991; Greilhuber and Ehrendorfer 1975), which results in variable numbers that often depart from multiples of the basic chromosome number of $x = 20$. Thus, ploidy levels were deduced to be diploid if $n = 18$–20 or $2n = 31$–41, triploid if $2n$ was around 60 and tetraploid if $2n = 60$–80. Additionally, we used microsatellite data (see above) to infer chromosome numbers. Based on the number of alleles detected per individual, they were characterized as either polyploid or diploid with a diploid individual showing a maximum of two peaks per locus and a polyploid showing up to four alleles. These polyploids are most probably tetraploids; however, the presence of triploids cannot be excluded, as three alleles can occur in both of them. Populations were classified as either diploid if all plants had fewer than three alleles at both loci, polyploid, if at least 50% of plants had more than two alleles and of mixed ploidy when 0–50% of plants had more than two alleles.

Morphometric analysis

For morphometric analyses, 32 locations in the main study regions in East and North Germany were sampled. If possible, on 20 individuals per location (in total 626) the following parameters were measured, which are traditionally used to identify the species: plant height (ph), inflorescence length (il), number of flowers (nf), flower density (fd, number of flowers/inflorescence length), number of leaves (nl), length of the second lowermost leaf (ll), maximum width of the second lowermost leaf (lw), spur length (sl) and ovary length (ol) of a central flower (flower traits only measured for 30 locations). We tested for significant difference between the species as well as the two study regions using a mixed effect model with species/region as fixed factors and populations and individuals as random factors. If not normally distributed, data were either log- (hp, li, wl, nf, fd) or sqrt-transformed (ll) to achieve normally distributed errors.

In order to assess the level of morphological differentiation between the species, we first performed principal component analysis (PCA) on individuals and on populations, the latter using the mean values of nine morphological variables (30 populations, 408 individuals). Thereafter, a linear discriminant analysis was performed with individuals grouped to species using the data from 30 populations for which all nine above-mentioned parameters were available. Normal distribution as tested with the Lilliefors (Kolmogorov-Smirnov) test was not fully achieved after transformation for some variables because of outliers (wl, ls, lo) or kurtosis (nl). However, although discriminant analysis assumes normal distribution within

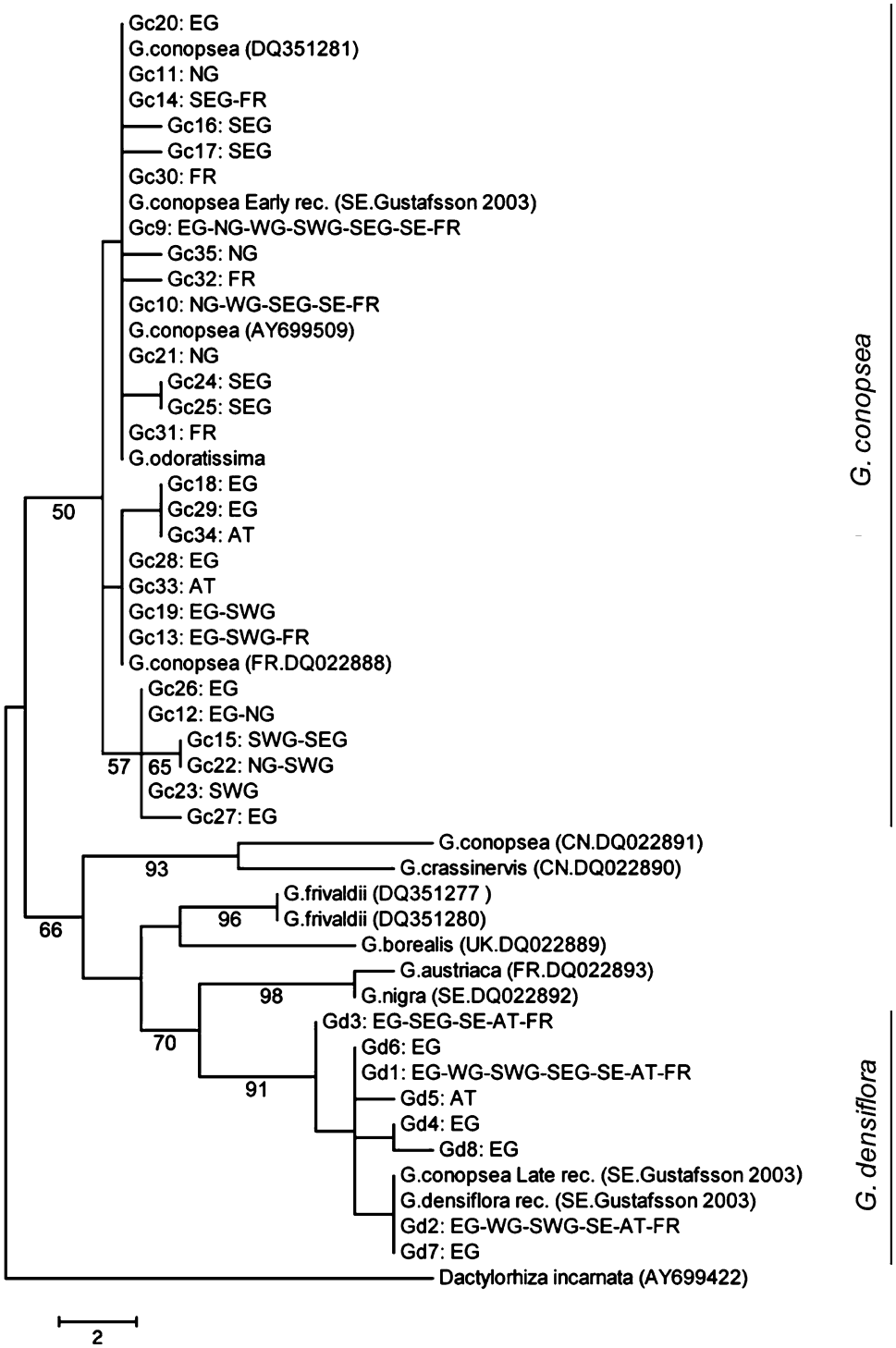

Fig. 1 One tree out of 441 most parsimonious trees (length = 62) for unique *Gymnadenia conopsea* (Gc9–Gc35) and *G. densiflora* (Gd1–Gd8) sequences of the ITS region (ITS1, 5.8 s, ITS2) from this study and published *Gymnadenia* data. The consistency index of the tree is 0.725, the retention index is 0.931, and the composite index is 0.766 for parsimony-informative sites. *Dactylorhiza incarnata* has been set as outgroup. Bootstrap support based on 500 replicates and >50% is given. Localities where the sequence types occur are presented as *EG* East Germany, *NG* North Germany, *WG* West Germany, *SWG* Southwest Germany, *SEG* Southeast Germany, *SE* Sweden, *FR* France and *AT* Austria, or by international abbreviation codes for GenBank samples

groups, violations of this assumption are not fatal and results are still trustworthy (Hill and Lewicki 2007). Subsequently, the posterior probabilities were correlated with the morphological variables, indicating the contribution of each variable to the discrimination between species. Morphometric data analysis was performed with R 2.10.1 (R Development Core Team 2009) using the MASS and nortest library.

Results

Phylogenetic analysis

ITS sequences were obtained from 100 samples from 49 sites. These resulted in a total of 140 haplotypes as 39 sequences had ambiguous nucleotide sites that were resolved applying the PHASE algorithm. We detected 35

unique haplotypes (Table 2), 27 of which were *G. conopsea* (Gc9–Gc35) and 8 were *G. densiflora* (Gd1–Gd8) haplotypes. In Germany, 29 of the 35 haplotypes were found, whereas 3 were restricted to France and Austria, respectively. In most sampling locations only one species was detected, but one site in the Eifel (WG3) and France (FR4) harbored both species (Table 1). Sequences containing ambiguous sites never included haplotypes of both species, but only those of either species.

The alignment with sequences of both species only had a length of 599 base pairs with 26 variable sites, which were located in the ITS1 (14 sites) and ITS2 (12 sites) regions (Table 2). The total alignment used for phylogenetic inference, including other congeneric species, had a length of 589 base pairs with 65 variable nucleotide sites out of which 36 were parsimony informative.

Both the maximum parsimony (Fig. 1) and the maximum likelihood method (Fig. S1 Supplement) yielded very similar results. In both trees the samples separated into two major clades, but with only low bootstrap support (50/66 and 72/77% for parsimony and maximum likelihood, respectively). The first one comprised most *G. conopsea* sequences from GenBank, our *G. conopsea* haplotypes Gc9–Gc35 as well as *G. odoratissima*, but revealed no further structuring. The second clade consisted of six congeners, but only few groupings within the clade received support >80%: *G. crassinervis* clustering with a sample of "*G. conopsea*", both from China; two *G. frivaldii* sequences, *G. austriaca* together with *G. nigra*, both formerly *Nigritella*; and lastly the clade including published *densiflora* sequences, as well as a late-flowering Swedish "*G. conopsea*" and our *G. densiflora* sequences Gd1–Gd8, originating from nearly all sampled regions. Both phylogenies strongly suggest that *G. conopsea* and *G. densiflora* are not sister species. Rather, the former *Nigritella* is the sister group of *G. densiflora*. However, the phylogenetic status of *G. conopsea* remains unclear.

Disregarding the published sequence from China and the late-flowering Swedish "*G. conopsea*", which seem not to be *G. conopsea*, the mean genetic divergence between the *Gymnadenia* species was 2% ($p = 0.02$, SD 0.007), ranging from $p_{min} = 0.002$ between *G. odoratissima* and *G. conopsea* to $p_{max} = 0.032$ between *G. crassinervis* and *G. austriaca*. The mean divergence between sequences of *G. conopsea* and *G. densiflora* from our study was considerable with 2% ($p = 0.02$, SD 0.002), whereas within-species divergence was low (*G. conopsea*: $p = 0.003$, SD 0.002; *G. densiflora*: $p = 0.003$, SD 0.002). The divergence between the two species was caused by 11 nucleotide sites, of which 8 differences were fixed and 3 sites showed variability within 1 of the clades (Table 2).

Microsatellite analysis

Three of the microsatellite markers (Gc17, Gc49, Gc77) produced hardly any bands in samples assigned to *G. densiflora*, but two loci (Gc42, Gc51) worked well in all samples and were used for the subsequent analyses. The total number of alleles detected in the two loci was three times larger in *G. conopsea* than in *G. densiflora* (data not shown). The PCA analysis of the European data set resulted in a differentiated pattern with two major clusters and a few intermediate populations (Fig. 2a). All

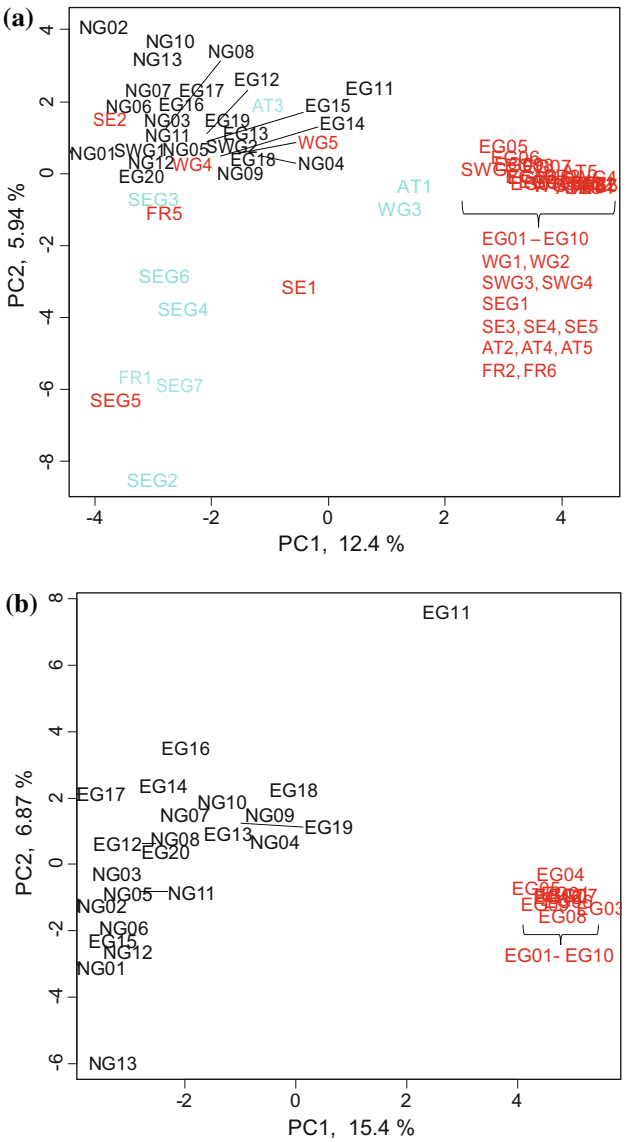

Fig. 2 Principal component analysis of *Gymnadenia conopsea* and *G. densiflora* populations based on allele frequencies of microsatellite loci Gc42 and Gc51. Colors indicate ploidy as determined from microsatellite genotypes and direct counting (*red* = 2*x*, *black* = 4*x*, *blue* = mixed ploidy); for site codes see Table 1. **a** European data set, **b** 32 sites from East and North Germany

Table 3 Chromosome counts and ploidy level of *Gymnadenia conopsea* ($N = 16$ individuals) and *G. densiflora* ($N = 15$ individuals) from 10 sites in East Germany (*D* diploid, *P* polyploid)

Study site	Code	*N*	Individual chromosome counts	Ploidy level (2*n*)	Ploidy type
Gymnadenia densiflora					
Tote Täler	EG06	2	$2n = 39$; $2n = 40$	2*x*	D
Theißen	EG02	1	$n = 19$–20	2*x*	D
Predel	EG03	1	$2n = 35$–38	2*x*	D
Jaucha	EG01	5	$2n = 36$–40; $2n = 32$–40; $n = 19$–$20/2n = 36$–40; $2n = 35$–36; $n = 18/2n = 33$–38	2*x*	D
Würze	EG08	3	$2n = 37$–41; $2n = 31$–37; $2n = 40$	2*x*	D
Jägertalwiese	EG10	3	$n = 20$; $2n = 38$–40; $2n = 36$–38	2*x*	D
Gymnadenia conopsea					
Steigra	EG17	1	$2n = 63$	3*x*	P
		5	$2n = 71$–79; $2n = 72$–78; $2n = 75$; $2n = 60$–74; $2n = 78$	4*x*	P
Krawinkel	EG16	1	$2n = 40$–60	3*x*	P
		4	$2n = 63$–80; $2n = 80$; $2n = 75$–80; $2n = 80$	4*x*	P
Rabis	EG15	1	$2n = 65$–66	3*x*?	P
		1	$2n = 69$–77	4*x*	P
Alter Gleisberg	EG14	1	$2n = 62$–64	3*x*	P
		2	$2n = 73$–77; $2n = 78$–80	4*x*	P

G. densiflora populations formed a dense cluster separated along Axis 1. This cluster also included alpine *G. c.* ssp. *serotina* sensu Dworschak (2002), which thus can be subsumed as *G. densiflora.* All *G. densiflora* were diploid. Most other populations clustered to *G. conopsea*. These populations were more strongly separated along Axis 2 with the alpine populations somewhat separated. The *G. conopsea* cluster comprised diploid, tetraploid and mixed ploidy populations, which, however, were not differentiated. Three populations took intermediate positions (WG3, SE1, AT1) in the PCA and contained individuals from both species when individual genotypes were inspected (data not shown).

The PCA analysis of the populations from East and North Germany separated the populations into two major clusters, representing the taxonomic affiliation to either *G. conopsea* or *G. densiflora* (Fig. 2b). This clear grouping is caused by strongly deviating allele frequencies at both loci. *G. conopsea* had a wider distribution along both axes, indicating more heterogeneous genotypes than *G. densiflora*. One population morphologically assigned as *G. conopsea* showed an intermediate position (EG11), which was due to low allelic diversity compared to other *G. conopsea* populations, possibly as a result of bottleneck effects. Mapping of ploidy onto the PCA showed that all *G. conopsea* were tetraploid and all *G. densiflora* were diploid (Fig. 2b).

Ploidy level

Chromosome numbers were directly counted in four populations of *G. conopsea* and six populations of *G. densiflora* (Table 3). All individuals of *G. conopsea* were either tetraploid (12 individuals) or triploid (4 individuals). In contrast, all individuals of *G. densiflora* proved to be diploid (15 individuals).

Microsatellite markers showed a maximum of two alleles per individual for *G. densiflora*, indicating diploidy in accordance with the chromosome counts. All *G. conopsea* had a re-occurring polyploid pattern, with more than two alleles for at least one microsatellite locus, indicative of polyploidy. In the whole European data set, populations of *G. densiflora* were diploid throughout (Table 1). In contrast, *G. conopsea* populations were either diploid (Eifel, Bavaria, Sweden, France), tetraploid (Saarland) or of mixed ploidy (Eifel, Alps).

Morphology

Altogether 626 samples were investigated, with 372 individuals in East Germany (*G. densiflora*: 173 samples, *G. conopsea*: 199 samples) and 254 in North Germany (all *G. conopsea*). Significant differences between the species were found for seven out of nine morphological parameters. *G. densiflora* was in general larger and showed higher

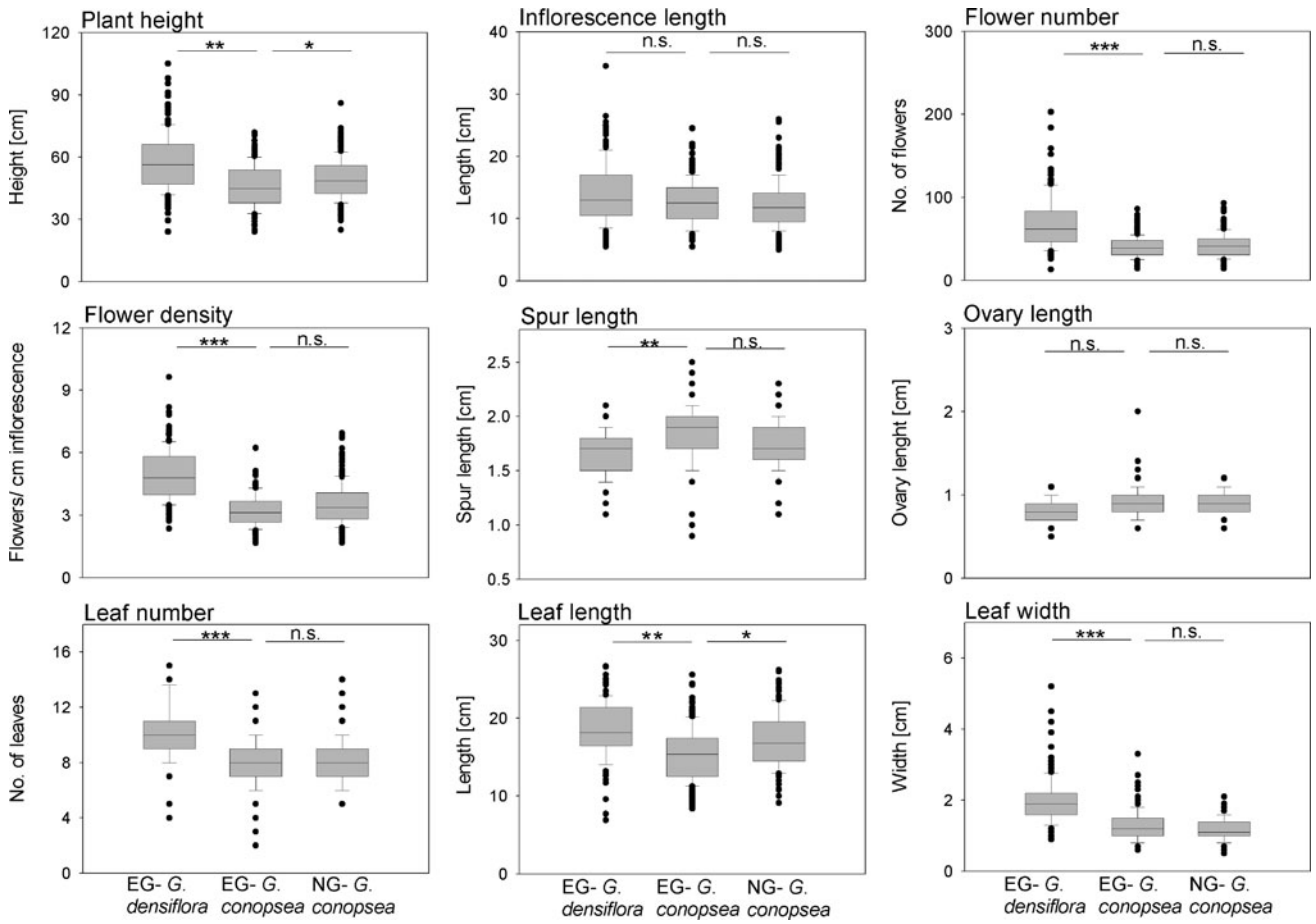

Fig. 3 Morphometric analysis of 32 sites in the study regions in East Germany (*EG*) and North Germany (*NG*), using the morphological characters that are traditionally used to identify the two species *Gymnadenia conopsea* and *G. densiflora*

levels than *G. conopsea* for plant height ($p < 0.01$), leaf number ($p < 0.001$), leaf length ($p < 0.01$), leaf width ($p < 0.001$), flower number ($p < 0.001$) and flower density ($p < 0.001$), but had a shorter spur ($p < 0.01$) (Fig. 3 and Table S1 Supplement). The largest relative differences were found for the number and density of flowers and for the number and width of the leaves. For *G. conopsea* only small differences were found between the regions for plant height ($p = 0.05$) and leaf length ($p = 0.02$).

In the PCA, 65.6 and 76.0% of the morphological variations were described by the first two axes for individuals and populations, respectively (Fig. 4). When species were mapped on the PCA scores, they were separated mainly along PC 1, however, with substantial overlap. Visual inspection of the scores of the morphological variables indicates that flower density, flower number and number of leaves were most distinct between species. *G. densiflora* showed a higher morphological variability than *G. conopsea*. In accordance with PCA the discriminant analysis was not fully successful to separate the species as 96% of *G. conopsea* individuals, but only 77% of *G. densiflora* were assigned correctly. Variables most highly correlated with the posterior probabilities were flower density ($r = 0.75$), number of flowers ($r = 0.71$), leaf width ($r = 0.71$) and number of leaves ($r = 0.68$). Note that in this data set, *G. densiflora* was found to be diploid and *G. conopsea* tetraploid.

Discussion

Our data provide unequivocal evidence for strong phylogenetic and genetic differentiation, but incomplete morphological differentiation between the two taxa of *Gymnadenia conopsea* s.l. Thus, a species rank is supported for *Gymnadenia densiflora* (Wahlenb.) Dietrich and *Gymnadenia conopsea* (L.) R.Br. s.str., as has been suggested previously (Bateman et al. 1997; Campbell et al. 2007; Marhold et al. 2005). Concerning ploidy, our findings are concordant with Marhold et al. (2005) as *G. conopsea* was found to be either diploid or tetraploid, whereas *G. densiflora* was found to be diploid throughout.

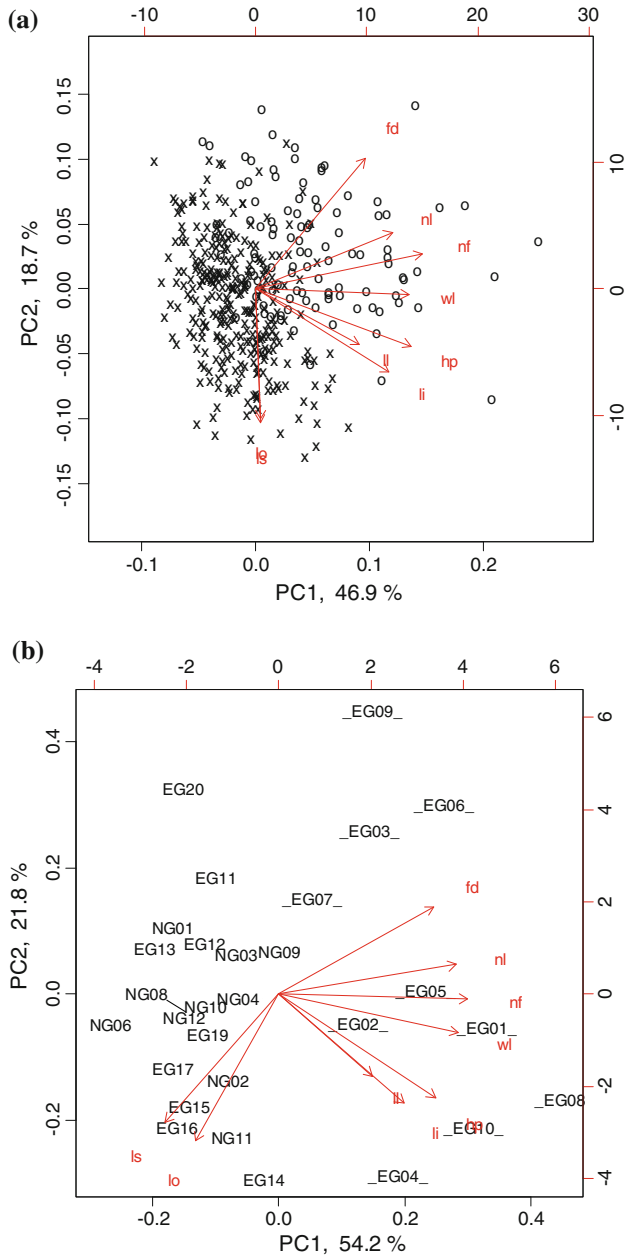

Fig. 4 Biplot of principal component scores PC1 vs. PC2 of nine morphological characters for **a** individuals and **b** populations of *Gymnadenia conopsea* (*x*) and *G. densiflora* (*o*, underlined site codes) and of underlying morphological characters. For site codes and characters, see text

Genetic differentiation

ITS sequence divergence has been suggested to be indicative of differentiation on the species level in orchids (Bateman et al. 2003). Based on substantial ITS divergence, Bateman et al. (1997) recognized *G. c.* ssp. *borealis* as a full species and suggested the same status for *G. c.* ssp. *densiflora,* which was further substantiated later (Bateman et al. 2006). Our results for ITS on a large sample of both *G. conopsea* and *G. densiflora* support the view that they in fact deserve species status, as the genetic divergence of 2% between the two clades was similar to the average genetic divergence between other species of the genus *Gymnadenia*. Also, sequences of *G. densiflora* form a well-supported monophyletic group sharing a most recent common ancestor with *G. nigra* and *G. austriaca*. Hence, *G. densiflora* and *G. conopsea* are not even sister species. Based on ITS data, *G. conopsea* cannot be separated from *G. odoratissima*, a morphologically well-distinguished species, which is in line with Gustafsson and Lönn (2003). Hence, the evolutionary history of the complex needs to be assessed with additional genetic markers. A combination of nuclear and plastid marker genes can be particularly useful in this respect, as this may allow the inference of lineages with hybrid origin (van den Berg et al. 2009).

The phylogenetic divergence between *G. conopsea* and *G. densiflora* is fully supported by microsatellite analyses presented here and previously (Campbell et al. 2007), and by allozyme data (Scacchi and de Angelis 1989). Using a different set of microsatellite loci Campbell et al. (2007) similarly found strong differentiation of allele frequencies on the British Isles and suggested to consider them as distinct species. In our study, in addition to differences in alleles, some microsatellite markers did not produce any bands in *G. densiflora*, while all markers were amplified in *G. conopsea*, which underlines the genetic differentiation between the taxa. The strong differentiation between the taxa also contrasts with the low to moderate differentiation among populations within the two taxa shown here and elsewhere for *Gymnadenia* (Chung 2009; Soliva and Widmer 1999).

Ploidy

The formerly unclear ploidy relationships of *G. conopsea* and *G. densiflora* were resolved by Marhold et al. (2005), who concluded that *G. densiflora* is diploid and that other reports are probably based on taxonomic misinterpretations, whereas *G. conopsea* can be both diploid and tetraploid. Here, using direct counting and microsatellite genotypes, we also found a clear pattern of diploidy for *G. densiflora* across all samples from the French Alps to Sweden. Recently, based on a basic chromosome number of $x = 10$ rather than $x = 20$, higher ploidy levels were deduced (Jersáková et al. 2010; Trávníček et al. 2011). However, the fact that we detected a maximum of two microsatellite alleles in plants with $2n = 40$ and a maximum of four alleles for $2n = 80$ suggests $x = 20$ as the basic chromosome number and diploidy for *G. densiflora*. For *G. conopsea* we found both diploids and polyploids with different distribution patterns. Only diploids were detected in Sweden, consistent with microsatellite analyses

(Gustafsson 2000). In contrast, only tetraploids were encountered in East Germany, North Germany and Saarland, and both ploidy levels occurred either within or among populations in the Alps and in the Eifel. The occurrence of two ploidy levels within or among populations of *G. conopsea* has also been found previously (Jersáková et al. 2010; Marhold et al. 2005). Surprisingly, so far none of the studies that used codominant molecular markers in *G. conopsea* have reported on problems with more than two alleles per locus, which would have indicated polyploidy (Campbell et al. 2007; Gustafsson 2000; Gustafsson and Lönn 2003; Gustafsson and Sjögren-Gulve 2002; Scacchi and de Angelis 1989; Soliva and Widmer 1999). Thus, in the respective regions, *G. conopsea* seems to be at least predominantly diploid, i.e., Sweden, the British Isles, Italy and Switzerland. This complex geographic distribution of ploidy levels suggests that a phylogeographic perspective is needed to achieve a more comprehensive picture (Nordström and Hedrén 2008). The two ploidy levels of *G. conopsea* were hardly differentiated in the microsatellite analysis, and the most frequent ITS haplotypes occurred in both of them. This lack of genetic differentiation between ploidy levels may indicate an autopolyploid origin of tetraploid from diploid *G. conopsea* as suggested by Jersáková et al. (2010). The presence of triploid individuals in some populations found here and elsewhere (Marhold et al. 2005; Trávníček et al. 2011) may indicate that gene flow between ploidy levels can occur via a triploid bridge (Ramsey and Schemske 1998). However, considerable phenological separation has been reported between ploidy levels, which reduces the opportunity for gene flow and may select for alternative pollinators (Jersáková et al. 2010). We cannot assess the degree of morphological differentiation of the two ploidy levels in *G. conopsea* as in the main study regions only tetraploids were found. However, spur length varies significantly between ploidy levels within *G. conopsea* (Jersáková et al. 2010), and variation in other traits is probable. Interestingly, the spur length of *G. densiflora* is intermediate between the two ploidy levels of *G. conopsea* (Jersáková et al. 2010). This shows that polyploidy has further complicated the distinction of the species, both morphologically and phenologically. Therefore, determination of ploidy levels is essential, and it will be important to investigate which diagnostic characters are affected by ploidy and which ones are not.

Morphology

While a number of morphological characters, e.g., flower number and flower density, allow a fairly good distinction between the two species, they show considerable morphological variation, which sometimes does not allow unequivocal determination. Apart from polyploidy (see above), two other aspects of orchid biology may contribute to the similarity between and variation within species, mycorrhization and pollination biology.

Orchids are obligatorily associated with mycorrhizal fungi for germination, and the availability of fungal partners might determine habitat suitability and orchid distribution. Furthermore, mycorrhizal fungi can have an effect on plant growth, and different types of the same fungal taxon have been shown to differ in their effect, e.g., for plant size (Lee 2002). *G. conopsea* has been found to associate with a large number of fungal taxa, and the spatial structure of the fungal community suggests a non-random distribution (Stark et al. 2009). Such fine-scale distribution patterns are thought to contribute to orchid diversification, for example, the phylogenetic divergence of floral variants within the *Hexalectris spicata* complex is partly attributed to differences in fungal associates (Taylor et al. 2003), or a correlation between *Corallorhiza maculata* genotypes and certain fungal associates has been found (Taylor et al. 2004). Therefore, both the overall similarity between and the morphological variation within the two species may be related to an association with mycorrhizal fungi that may either differ within or may be shared among taxa.

Orchids are prime examples of selection on flower morphology by insect pollinators (Thompson 1994). Both *G. densiflora* and *G. conopsea* have fairly specialized flowers, which provide nectar and are pollinated by the same taxa of moths. Thus, it may be hypothesized whether convergent selection of pollinators has led to similar flower morphology, although it has been shown that the two species differ in their scent bouquet (Jersáková et al. 2010). Furthermore, as *G. densiflora* is described to be strongly scented and *G. conopsea* appears to be less scented (Schmeil 1996), the latter may be under selection to morphologically resemble *G. densiflora* to attract the same pollinators.

Taking the findings together, an evolutionary scenario emerges in which *G. conopsea* and *G. densiflora* have phylogenetically separated from each other prior to the split of other groups, e.g., the former *Nigritella*. The two taxa have achieved a large distribution range and diverged ecologically, phenologically and partly also morphologically, however retaining considerable variability. The phylogenetic split also seems to have occurred before polyploidy arose in *G. conopsea*, probably by autopolyploidy, whereas *G. densiflora* stayed at the diploid level. Polyploidy in turn may have led to increased morphological and phenological variability within *G. conopsea.*

Acknowledgments We are grateful to local AHOs (Arbeitskreis Heimische Orchideen) and their members for their support during this

work. Special thanks to K. Heyde, W. Mohrman, P. Rode, J. Passin, P. Steinfeld and F. Opitz for locating plant populations; and to W. Dworschak, O. Gerbaud, A. Hoch, H. Kerschbaumsteiner, S. Klemich, P. Kuss and M. L. Perko for providing samples. Thanks to Jan Hanspach for statistical advice and two anonymous reviewers for valuable comments. This work was supported by the Federal Ministry of Education and Research (BMBF-FKz 01LC0018A) within the framework of BIOLOG-SUBICON (Successional Change and Biodiversity Conservation).

References

Amich F, García-Barriuso M, Bernardos S (2007) Polyploidy and speciation in the orchid flora of the Iberian Peninsula. Bot Helv 117:143–157

Bateman RM, Pridgeon AM, Chase MW (1997) Phylogenetics of subtribe Orchidinae (Orchidoideae, Orchidaceae) based on nuclear ITS sequences. 2. Intrageneric relationships and reclassification to achieve monophyly of *Orchis* sensu stricto. Lindleyana 12:113–141

Bateman RM, Hollingsworth PM, Preston J, Yi-Bo L, Pridgeon AM, Chase MW (2003) Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). Bot J Linn Soc 142:1–40

Bateman RM, Rudall PJ, James KE (2006) Phylogenetic context, generic affinities and evolutionary origin of the enigmatic Balkan orchid *Gymnadenia frivaldii* Hampe ex Griseb. Taxon 55:107–118

Bernardos S, Amich F, Gallego F (2003) Karyological and taxonomic notes on *Ophrys* (Orchidoideae, Orchidaceae) from the Iberian Peninsula. Bot J Linn Soc 142:395–406

Bianco P, D'Emerico S, Medagli P, Ruggiero L (1991) Polyploidy and aneuploidy in *Ophrys*, *Orchis*, and *Anacamptis* (Orchidaceae). Plant Syst Evol 178:235–245

Bisse J (1963) Ein Beitrag zur Kenntnis der deutschen Orchideenflora. Feddes Repert 67:181–189

Bory S, Catrice O, Brown S, Leitch IJ, Gigant R, Chiroleu F, Grisoni M, Duval MF, Besse P (2008) Natural polyploidy in *Vanilla planifolia* (Orchidaceae). Genome 51:816–826

Campbell VV, Rowe G, Beebee TJC, Hutchings MJ (2002) Isolation and characterization of microsatellite primers for the fragrant orchid *Gymnadenia conopsea* (L.) R. Brown (Orchidaceae). Conserv Genet 3:209–210

Campbell VV, Rowe G, Beebee TJC, Hutchings MJ (2007) Genetic differentiation amongst fragrant orchids (*Gymnadenia conopsea s.l.*) in the British Isles. Bot J Linn Soc 155:349–360

Chung MY (2009) Low levels of genetic variation within populations of the four rare orchids *Gymnadenia cucullata*, *Gymnadenia camtschatica*, *Amitostigma gracile*, and *Pogonia minor* in South Korea: indication of genetic drift and implications for conservation. Plant Syst Evol 281:65–76

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659

Cozzolino S, Widmer A (2005a) Orchid diversity: an evolutionary consequence of deception? Trends Ecol Evol 20:487–494

Cozzolino S, Widmer A (2005b) The evolutionary basis of reproductive isolation in Mediterranean orchids. Taxon 54:977–985

Cresswell JE (1998) Stabilizing selection and the structural variability of flowers within species. Ann Bot 81:463–473

Delforge P (2006) Orchids of Europe, North Africa and the Middle East, 3rd edn. A&C Black Publishers Ltd., London

Development Core Team R (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Devos N, Oh S-H, Raspé O, Jacquemart AL, Manos PS (2005) Nuclear ribosomal DNA sequence variation and evolution of spotted marsh-orchids (*Dactylorhiza maculata* group). Mol Phylogenet Evol 36:568–580

Dressler RL (1987) Die Orchideen: Biologie und Systematik der Orchidaceae. Eugen Ulmer GmbH & Co., Stuttgart

Dworschak W (2002) Gliederung der verschiedenen Erscheinungsformen der Mücken-Händelwurz in Südbayern. In: Naturwissenschaftlicher Verein Wuppertal e.V. (Ed.) Jahresberichte des Naturwissenschaftlichen Vereins Wuppertal e.V., Heft 55, Wuppertal, pp 27–45

Gill DE (1989) Fruiting failure, pollinator inefficiency and speciation in orchids. In: Otte D, Endler JA (eds) Speciation and its consequences. Sinauer Associates Inc., pp 458–481

Greilhuber J, Ehrendorfer F (1975) Chromosome numbers and evolution in *Ophrys* (Orchidaceae). Plant Syst Evol 124:125–138

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Gustafsson S (2000) Patterns of genetic variation in *Gymnadenia conopsea*, the fragrant orchid. Mol Ecol 9:1863–1872

Gustafsson S, Lönn M (2003) Genetic differentiation and habitat preference of flowering-time variants within *Gymnadenia conopsea*. Heredity 91:284–292

Gustafsson S, Sjögren-Gulve P (2002) Genetic diversity in the rare orchid, *Gymnadenia odoratissima* and a comparison with the more common congener, *G. conopsea*. Conserv Genet 3:225–234

Gustafsson S, Thorén PA (2001) Microsatellite loci in *Gymnadenia conopsea*, the fragrant orchid. Mol Ecol Notes 1:81–82

Hedrén M, Klein E, Teppner H (2000) Evolution of polyploids in the European orchid genus *Nigritella*: evidence from allozyme data. Phyton-Ann Rei Bot A 40:239–275

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

Hill T, Lewicki P (2007) STATISTICS methods and applications. StatSoft, Tulsa

Jersáková J, Castro S, Sonk N, Milchreit K, Schödelbauerová I, Tolasch T, Dötterl S (2010) Absence of pollinator-mediated premating barriers in mixed-ploidy populations of *Gymnadenia conopsea s.l.* (Orchidaceae). Evol Ecol 24:1199–1218

Juillet N, Delle-Vedove R, Dormont L, Schatz B, Pailler T (2010) Differentiation in a tropical deceptive orchid: colour polymorphism and beyond. Plant Syst Evol 289:213–221

Lee SS (2002) A review of orchid mycorrhizae in Korea. The Plant Pathol J 18:169–178

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20:229–237

Marhold K, Jongepierová I, Krahulcová A, Kučera J (2005) Morphological and karyological differentiation of *Gymnadenia densiflora* and *G. conopsea* in the Czech Republic and Slovakia. Preslia 77:159–176

Micheneau C, Johnson SD, Fay MF (2009) Orchid pollination: from Darwin to the present day. Bot J Linn Soc 161:1–19

Mrkvicka AC (1993) Statistische Untersuchungen an *Gymnadenia conopsea* (L.) R. Br. s.l. Mitt Bl Arbeitskr Heim Orch Baden-Württ 25:361–367

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York

Nordström S, Hedrén M (2008) Genetic differentiation and postglacial migration of the *Dactylorhiza majalis* ssp. *traunsteineri/lapponica* complex into Fennoscandia. Plant Syst Evol 276:73–87

Otero JT, Flanagan NS (2006) Orchid diversity—beyond deception. Trends Ecol Evol 21:64–65

Peakall R (2007) Speciation in the Orchidaceae: confronting the challenges. Mol Ecol 16:2834–2837

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu Rev Ecol Syst 29:467–501

Salzmann CC, Nardella AM, Cozzolino S, Schiestl FP (2007) Variability in floral scent in rewarding and deceptive orchids: The signature of pollinator-imposed selection? Ann Bot 100:757–765

Scacchi R, de Angelis G (1989) Isoenzyme polymorphisms in *Gymnaedenia conopsea* and its inferences for systematics within this species. Biochem Syst Ecol 17:25–33

Schiestl FP, Schlüter PM (2009) Floral isolation, specialized pollination, and pollinator behavior in orchids. Annu Rev Entomol 54:425–446

Schmeil O (1996) Flora von Deutschland, 90. Auflage. Quelle & Meyer Bestimmungsbücher, Wiesbaden

Schwarzacher T, Ambros P, Schweizer D (1980) Application of giemsa banding to orchid karyotype analysis. Plant Syst Evol 134:293–297

Scopece G, Widmer A, Cozzolino S (2008) Evolution of postzygotic reproductive isolation in a guild of deceptive orchids. Am Nat 171:315–326

Soliva M, Widmer A (1999) Genetic and floral divergence among sympatric populations of *Gymnadenia conopsea* s.l. (Orchideaceae) with different flowering phenology. Int J Plant Sci 160:897–905

Ståhlberg D (2009) Habitat differentiation, hybridization and gene flow patterns in mixed populations of diploid and autotetraploid *Dactylorhiza maculata* s.l. (Orchidaceae). Evol Ecol 23:295–328

Stark C, Babik W, Durka W (2009) Fungi from the roots of the common terrestrial orchid *Gymnadenia conopsea*. Mycol Res 113:952–959

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Swarts ND, Dixon KW (2009) Terrestrial orchid conservation in the age of extinction. Ann Bot 104:543–556

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Taylor DL, Bruns TD, Szaro TM, Hodges SA (2003) Divergence in mycorrhizal specialization within *Hexalectris spicata* (Orchidaceae), a nonphotosynthetic desert orchid. Am J Bot 90: 1168–1179

Taylor DL, Bruns TD, Hodges SA (2004) Evidence for mycorrhizal races in a cheating orchid. Proc R Soc Lond B Biol Sci 271: 35–43

Thompson JN (1994) The coevolutionary process. The University of Chicago Press, Chicago

Trávníček P, Kubátová B, Čurn V, Rauchová J, Krajníková E, Jersáková J, Suda J (2011) Remarkable coexistence of multiple cytotypes of the *Gymnadenia conopsea* aggregate (the fragrant orchid): evidence from flow cytometry. Ann Bot 107:77–87

Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA (1980) Flora Europaea Cambridge University Press, Cambridge

van den Berg C, Higgins WE, Dressler RL, Whitten WM, Soto-Arenas MA, Chase MW (2009) A phylogenetic study of Laeliinae (Orchidaceae) based on combined nuclear and plastid DNA sequences. Ann Bot 104:417–430

Vereecken NJ, Dafni A, Cozzolino S (2010) Pollination syndromes in Mediterranean orchids-implications for speciation, taxonomy and conservation. Bot Rev 76:220–240

Vöth W, Sontag S (2006) Die intraspezifischen Varietäten der *Gymnadenia conopsea* (L.). R Br J Eur Orch 38:581–624

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, Inc., San Diego, pp 315–322

Winterfeld G, Röser M (2007) Chromosomal localization and evolution of satellite DNAs and heterochromatin in grasses (Poaceae), especially tribe Aveneae. Plant Syst Evol 264:75–100

Supplementary Material for:

Plant Systematics and Evolution

Strong genetic differentiation between *Gymnadenia conopsea* and *G. densiflora* despite morphological similarity

Christiane Stark, Stefan G. Michalski, Wiesław Babik, Grit Winterfeld and Walter Durka

Corresponding author:
Christiane Stark
Helmholtz Centre for Environmental Research - UFZ, Department of Community Ecology,
Theodor-Lieser Str. 4, D-06120 Halle, Germany
E-Mail: christiane.stark@ufz.de

Table S1 Morphological parameters (mean ± SE) and sample size for the 32 sites of *G. conopsea* and *G. densiflora* investigated in the study regions in East Germany (EG01 - EG20) and North Germany (NG01 - NG13).

Study site	Code	N	Plant height (ph, cm)	Inflorescence length (il, cm)	Flower number (nf)	Flower density (fd, cm^{-1})	Leaf number(nl)	Leaf length (ll, cm)	Leaf width (lw, cm)	Spur length (sl, cm)	Ovary length (ol, cm)
Gymnadenia densiflora											
Jaucha	EG01	20	61.20 ± 1.94	16.90 ± 0.85	92.55 ± 6.53	5.49 ± 0.31	11.35 ± 0.48	17.37 ± 0.48	1.93 ± 0.08	1.80 ± 0.03	0.81 ± 0.02
Theißen	EG02	20	63.93 ± 2.10	12.78 ± 0.85	50.80 ± 3.38	4.01 ± 0.12	8.80 ± 0-31	20.31 ± 0.64	1.83 ± 0.09	1.67 ± 0.03	0.82 ± 0.01
Predel	EG03	20	48.78 ± 1.95	11.63 ± 0.71	49.90 ± 2.77	4.35 ± 0.15	10.45 ± 0.38	19.32 ± 1.37	1.85 ± 0.09	1.54 ± 0.04	0.72 ± 0.03
Domsen	EG04	5	65.20 ± 3.03	15.00 ± 1.35	60.80 ± 9.04	4.00 ± 0.37	8.20 ± 1.20	26.60	2.00 ± 0.11	1.74 ± 0.08	0.88 ± 0.06
Espenhain	EG05	8	57.31 ± 9.11	16.31 ± 3.17	78.88 ± 21.64	4.31 ± 0.45	10.13 ± 0.83	13.46 ± 3.10	2.30 ± 0.40	1.64 ± 0.06	0.84 ± 0.03
Tote Täler	EG06	20	49.70 ± 2.10	14.45 ± 0.98	79.45 ± 5.41	5.56 ± 0.19	10.20 ± 0.36	17.13 ± 1.17	1.76 ± 0.15	1.49 ± 0.04	0.71 ± 0.03
Rothenstein	EG07	20	47.60 ± 1.84	10.18 ± 0.47	51.25 ± 3.38	5.04 ± 0.24	9.80 ± 0.31	18.73 ± 0.71	1.63 ± 0.06	1.81 ± 0.03	0.80 ± 0.02
Würze	EG08	20	70.58 ± 2.37	17.33 ± 0.93	95.05 ± 7.39	5.44 ± 0.24	11.90 ± 0.53	18.61 ± 0.80	2.59 ± 0.11	1.61 ± 0.03	0.94 ± 0.02
Klingelsteine	EG09	20	49.35 ± 1.93	10.63 ± 0.58	61.70 ± 2.87	5.89 ± 0.17	9.25 ± 0.32	16.07 ± 0.72	1.48 ± 0.10	1.44 ± 0.03	0.71 ± 0.02
Jägertalwiese	EG10	20	67.90 ± 3.40	15.85 ± 1.21	70.60 ± 10.29	4.22 ± 0.36	10.15 ± 0.50	21.02 ± 0.86	2.66 ± 0.27	1.71 ± 0.04	0.86 ± 0.02
Gymnadenia conopsea											
Domsen	EG11	20	43.85 ± 2.04	13.65 ± 0.96	37.65 ± 2.99	2.77 ± 0.10	7.90 ± 0.29	12.93 ± 0.54	1.33 ± 0.06	1.72 ± 0.03	0.70 ± 0.01
Rothenstein	EG12	20	45.45 ± 1.48	12.63 ± 0.73	41.45 ± 2.20	3.35 ± 0.16	7.33 ± 0.50	15.17 ± 0.72	1.37 ± 0.13	1.73 ± 0.05	0.84 ± 0.03
Zietschkuppe	EG13	20	42.50 ± 1.92	11.73 ± 0.57	34.90 ± 2.35	3.00 ± 0.17	6.85 ± 0.33	13.61 ± 0.77	1.16 ± 0.08	1.74 ± 0.03	0.90 ± 0.02
Alter Gleisberg	EG14	20	54.25 ± 2.04	15.58 ± 0.48	50.10 ± 3.02	3.18 ± 0.12	8.00 ± 0.34	17.86 ± 0.86	1.67 ± 0.09	1.90 ± 0.04	0.98 ± 0.03
Rabis	EG15	20	46.53 ± 1.76	13.05 ± 0.79	40.65 ± 2.66	3.13 ± 0.13	7.25 ± 0.34	17.69 ± 0.76	1.32 ± 0.09	1.80 ± 0.13	1.04 ± 0.11
Krawinkel	EG16[a]	20	49.63 ± 2.35	13.78 ± 0.77	37.10 ± 1.99	2.76 ± 0.13	7.30 ± 0.34	17.46 ± 0.85	1.17 ± 0.06	2.00 ± 0.05	0.91 ± 0.03
Steigra	EG17[a]	20	49.43 ± 2.47	12.73 ± 0.85	35.65 ± 2.50	2.84 ± 0.14	8.20 ± 0.39	13.96 ± 0.81	1.21 ± 0.08	1.98 ± 0.06	0.91 ± 0.02
Grockstädt	EG18	20	46.48 ± 2.06	12.00 ± 0.60	42.45 ± 2.62	3.54 ± 0.15	7.65 ± 0.31	16.39 ± 0.84	1.20 ± 0.08	-	-
Langer Berg	EG19	20	44.20 ± 2.16	11.58 ± 0.98	43.75 ± 3.64	3.84 ± 0.18	8.40 ± 0.28	15.37 ± 0.68	1.51 ± 0.11	1.88 ± 0.05	0.99 ± 0.03
Tote Täler	EG20[a]	19	34.95 ± 1.60	9.82 ± 0.50	37.89 ± 2.17	3.89 ± 0.15	8.63 ± 0.40	13.45 ± 0.55	1.02 ± 0.07	1.68 ± 0.04	0.80 ± 0.04
Stb. Polle	NG01[a]	20	43.98 ± 1.73	9.88 ± 0.63	39.05 ± 2.61	4.04 ± 0.21	7.20 ± 0.29	14.71 ± 0.88	1.12 ± 0.10	1.79 ± 0.04	0.91 ± 0.03
Stb. “Im Schießstand”	NG02	20	52.48 ± 2.66	13.95 ± 0.89	42.80 ± 3.22	3.11 ± 0.19	7.60 ± 0.32	16.82 ± 0.98	1.35 ± 0.07	1.79 ± 0.04	0.93 ± 0.03
Stb. “Alter Steinbruch”	NG03	20	49.68 ± 2.18	12.80 ± 0.76	46.35 ± 3.58	3.66 ± 0.24	7.35 ± 0.30	16.29 ± 0.78	1.21 ± 0.06	1.65 ± 0.06	0.86 ± 0.03
Stb. Hehlen	NG04	20	48.35 ± 1.77	11.05 ± 0.63	45.60 ± 3.58	4.11 ± 0.22	8.05 ± 0.41	18.43 ± 1.03	1.37 ± 0.06	1.75 ± 0.05	0.94 ± 0.02
Stb. Bärenbrink	NG05	14	47.36 ± 2.37	10.93 ± 0.79	35.64 ± 2.32	3.40 ± 0.26	6.71 ± 0.35	17.84 ± 1.03	1.22 ± 0.05	-	-
Stb. Delligsen	NG06	20	43.43 ± 1.51	10.08 ± 0.49	28.85 ± 1.26	2.97 ± 0.18	7.20 ± 0.28	14.64 ± 0.73	0.98 ± 0.05	1.88 ± 0.05	1.00 ± 0.03
Burgberg	NG07[a]	20	55.00 ± 2.23	15.30 ± 0.98	54.75 ± 3.94	3.59 ± 0.16	8.15 ± 0.24	19.36 ± 0.64	1.24 ± 0.07	-	-
Rühle	NG08[a]	20	49.18 ± 1.92	11.53 ± 0.71	39.50 ± 3.39	3.38 ± 0.19	7.90 ± 0.31	17.25 ± 0.92	0.95 ± 0.05	1.82 ± 0.04	0.90 ± 0.03

Räuschenberg	NG09	20	54.68 ± 2.05	11.43 ± 0.65	42.75 ± 2.56	3.89 ± 0.26	9.25 ± 0.40	17.32 ± 0.60	1.28 ± 0.05	1.65 ± 0.04	0.87 ± 0.03
Poppenburg	NG10	20	49.30 ± 1.51	10.40 ± 0.43	40.15 ± 2.17	3.91 ± 0.22	7.45 ± 0.23	19.79 ± 1.21	1.13 ± 0.06	1.81 ± 0.04	0.90 ± 0.02
Bielenberg	NG11	20	54.10 ± 2.22	15.95 ± 0.81	42.75 ± 2.78	2.67 ± 0.10	8.35 ± 0.31	17.00 ± 0.48	1.08 ± 0.06	1.81 ± 0.04	0.97 ± 0.02
Bocksberg	NG12	20	46.28 ± 1.65	12.50 ± 0.80	38.70 ± 2.56	3.13 ± 0.13	7.85 ± 0.38	15.98 ± 0.52	0.95 ± 0.05	1.74 ± 0.05	0.96 ± 0.03
Holberg	NG13	20	49.28 ± 2.73	12.25 ± 0.66	48.80 ± 2.92	4.07 ± 0.23	7.80 ± 0.44	19.30 ± 0.84	1.35 ± 0.07	-	-

[a]Site code *sensu* Stark et al. (2009): EG16 = E1; EG17 = E2; EG20 = E3; NG07 = N1; NG01 = N2; NG08 = N3

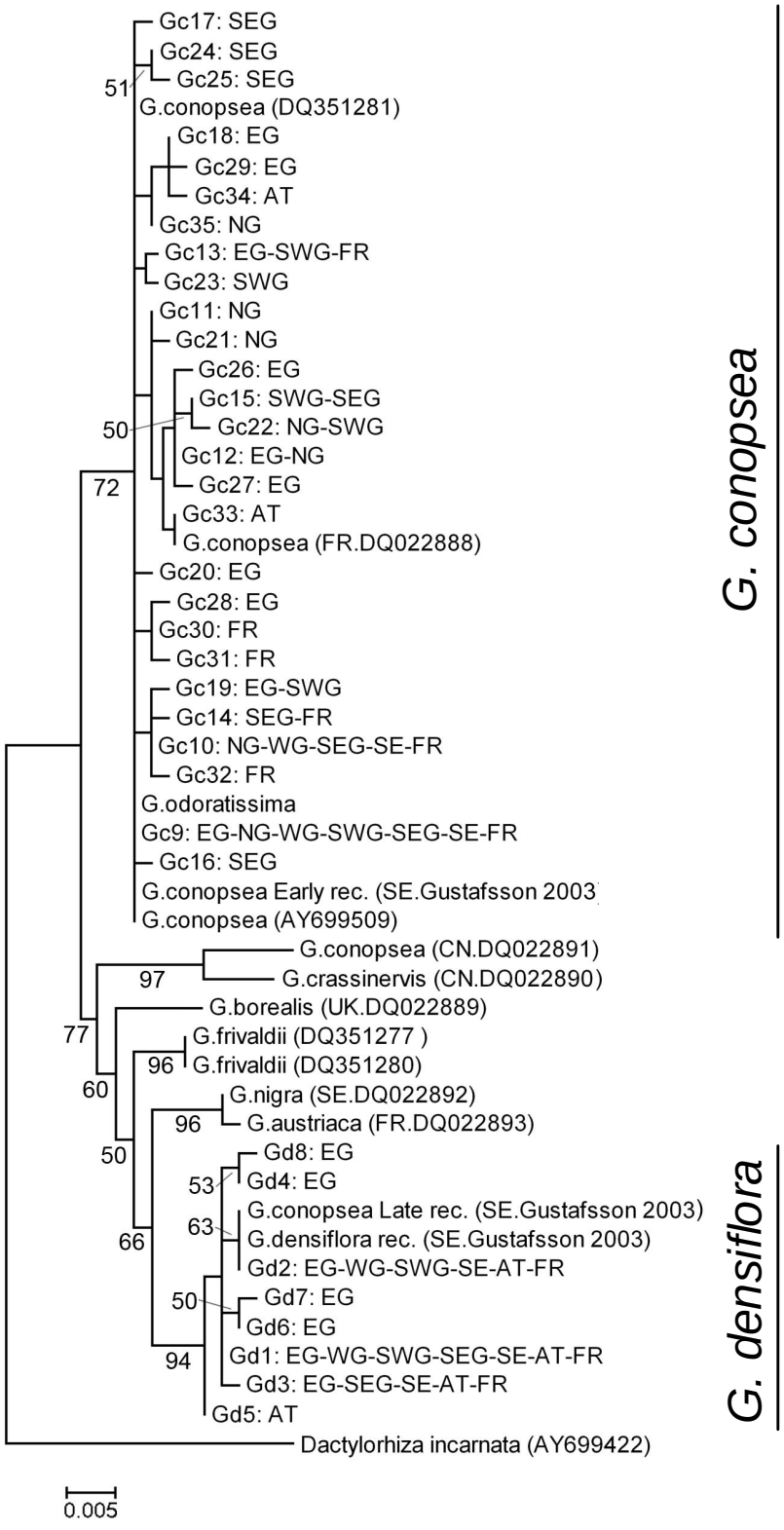

Fig. S1 Maximum likelihood phylogeny for unique *G. conopsea* (Gc9 - Gc35) and *G. densiflora* (Gd1 - Gd8) sequences of the ITS region (ITS1, 5.8s, ITS2) from this study and published *Gymnadenia* data. *Dactylorhiza incarnata* has been set as outgroup. Bootstrap support based on 500 replicates and > 50% is given. Localities where the sequence types occur are presented as EG = East Germany, NG = North Germany, WG = West Germany, SWG = South-West Germany, SEG = South-East Germany, SE = Sweden, FR = France, AT = Austria or by international abbreviation codes for GenBank samples.