

# Evolution of *Dactylorhiza baltica* (Orchidaceae) in European Russia: evidence from molecular markers and morphology

ALEXEY B. SHIPUNOV\*, MICHAEL F. FAY and MARK W. CHASE

Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, UK

Received November 2003; accepted for publication September 2004

Four plastid markers, four nuclear markers and 14 morphometric characters were used in this study to investigate the evolution of *Dactylorhiza baltica* (Orchidaceae) in European Russia. In total, 98, 214 and 775 samples from 85, 112 and 121 populations were involved in the combined and separate molecular and morphometric analyses, respectively. In most cases, morphometric measures were done on exactly the same plants that were used for DNA studies. *Dactylorhiza baltica* plants from European Russia are most probably the products of several recent and mostly local hybridization events between the diploids *D. fuchsii* and *D. incarnata*, which have each been the maternal parent on different occasions. Considerable introgression into the parental diploids via the allopolyploid *D. baltica* is also hypothesized. Several morphological characters, such as length of the lip lateral lobe and the length of longest leaf, were found to be robust and could be useful in identification of *D. baltica*. This study demonstrates the advantage of 'combined' techniques, especially in the case of taxonomically complex taxa. © 2005 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2005, 147, 257–274.

ADDITIONAL KEYWORDS: allotetraploids – microsatellites – morphometrics – systematics.

## INTRODUCTION

*Dactylorhiza* Necker ex Nevski (Orchidaceae) is, along with *Epipactis* and *Ophrys*, one of the most taxonomically controversial orchid genera in Europe. There is great instability in the accepted number of species and infraspecific taxa. The borders between many species are unclear, and there are considerable difficulties in the determination of single plants (Averyanov, 1990; Reinhard, 1990; Delforge, 1995; Stace, 1997; Bateman, 2001).

Many of the most problematic taxa are allotetraploids, most of which are believed to be the result of multiple hybridization events between two broadly defined parental species, *D. fuchsii* (Druce) Soó and *D. incarnata* (L.) Soó (Heslop-Harrison, 1968; Hedrén, 2002; Devos *et al.*, 2003; Y. Pillon, unpubl. data). Multiple lines of evidence indicate that this complex is an '... unusually dynamic system of polyploid speciation

and extinction in which polyploids evolve continuously from the same set of broadly defined parental lineages' (Hedrén, 2003: 2678). Furthermore, the limits of the diploid parental taxa are sometimes made less clear by the exchange of genetic material, hypothesized to be via allotetraploids (Hedrén, Fay & Chase, 2001; Hedrén, 2003), in spite of the differences in their ploidy.

One good example of such polyploid species and at the same time a less well known member of this complex is *D. baltica* (Klinge) Orlova, for which the distribution, unlike other named allotetraploids of this complex (which occur principally in western Europe), is restricted to the eastern part of Germany, Poland, the Baltic countries, southern Finland and Russia. The eastern parts of its distribution are less definite; some authors (Nevski, 1935; Smoljaninova, 1976) have argued that it is restricted to the western parts of European Russia (Pskov and Leningrad regions) together with some localities in the northern Urals and southern Siberia, whereas others expand the European portion across all of European Russia

\*Corresponding author. E-mail: a.shipunov@kew.org

(between the Arctic Circle and 50°N latitude) to the Urals (Soó, 1980; Averyanov, 1990). The most recent evidence is, however, that '*D. baltica*' populations in the southern Urals have been misidentified and should be assigned to *D. fuchsii* (Kulikov & Filippov, 1999a). Thus, current opinion limits the distribution of *D. baltica* in European Russia to between 50 and 60°N latitude (with two exceptions in the northern Urals) and west of 60° longitude (Fig. 1).

The epithet *baltica* was first used by Klinge (1895, 1898) for a subspecies of '*Orchis*' *latifolia* L., nom. illeg. [= *Dactylorhiza majalis* (Reichb.) P.F.Hunt & Summerhayes], a species with a western European distribution, long believed to be another member of the polyploid complex (Averyanov, 1990). This subspe-



**Figure 1.** The putative European distribution of *Dactylorhiza baltica* (Averyanov, 1990; Kulikov & Filippov, 1999a). Each collection site is labelled with an abbreviated region name (see Appendix 1).

cies was later upgraded to species rank by Nevski (1935) because there are some obvious morphological differences between *D. majalis* s.s. and *D. baltica*, especially in leaf form and lip shape (Table 1). Most authors now accept *D. baltica* as a separate species, some suggesting that *D. praetermissa* (Druce) Soó and *D. purpurella* (T. & T.A. Stephenson) Soó are its closest relatives (e.g. Vermeulen, 1947; Senghas, 1968; Averyanov, 1990). Morphological characters advocated for distinguishing *D. baltica* vary among authors (Table 1), although most descriptions mention the long, pointed leaves, short inflorescence and relatively wide lip with small lateral lobes.

Recent morphometric investigations have shown that *Dactylorhiza* allotetraploids often have morphological character states that are generally intermediate between *D. incarnata* and *D. fuchsii* (e.g. Tyteca & Gathoye, 1993). Biochemical and molecular methods can highlight molecular markers that are able to reveal inheritance, parentage and possible introgression between taxa. Studies of allozyme markers (Hedrén, 2002), AFLPs (Hedrén *et al.*, 2001), plastid markers and ITS alleles (Y. Pillon, unpubl. data) and plastid RFLPs (Hedrén, 2003) showed that: (1) most allotetraploids have indeed originated from hybridization between *D. incarnata* and *D. fuchsii* or *D. maculata* (L.) Soó; (2) they have originated several times (and are most likely still being generated); (3) most western European allotetraploids are easily distinguished by molecular characters from their parental species; (4) most allotetraploids have inherited plastid markers from either *D. fuchsii* or *D. maculata* rather than from *D. incarnata*, indicating that *D. fuchsii* and *D. maculata* are more often maternal parents; and (5) some allotetraploids have acquired markers thus far not found in parental taxa (Hedrén *et al.* 2001; Hedrén, 2003; Devos *et al.*, 2003; Y. Pillon, unpubl. data; Shipunov *et al.*, 2004).

**Table 1.** The most diagnostic morphological characters of *Dactylorhiza baltica* in this study compared with those from three previous studies (all measurements in millimetres)

Characters	Klinge (1898)	Nevski (1935)	Delforge (1995)	This study
Plant height	250–700	300–600	250–700	250–700
Leaf length	100–250	90–200	100–250	90–250
Leaf width	15–35	20–32	15–40	15–40
Leaf spots (1 light, 2 heavy)	1	1	1	1–2
Length of inflorescence	20–80	30–95	30–100	20–100
Length of lowest bract	–	20–30	–	>20
Spur length	6–9	7.5–9	6–9	6–9
Lip length	6–7	7–8.5	6–9	6–9
Lip width	8–12	9–10	8–13	7–13
Length of lip middle lobe (from the base of sinuses)	<3	2.5–3.5	–	<4

However, no molecular analysis has yet been performed on *D. baltica*, which is unique among other *Dactylorhiza* allotetraploids due its eastern distribution and relative isolation from other allotetraploids. Moreover, there are few morphometric studies of Russian dactylorchids (Kulikov & Filippov, 1999a, b). Our recent investigation of European Russian *Dactylorhiza* showed good agreement between morphometric characters and molecular markers such as plastid microsatellites and ITS alleles (Shipunov *et al.*, 2004). Plastid microsatellites had previously been shown to be useful for revealing geographical patterns, the maternal parentage of hybrids, and even some relationships among populations (Y. Pillon, unpubl. data; Shipunov *et al.*, 2004). ITS alleles, on the other hand, generate clear phylogenetic patterns and thereby help to distinguish species (Bateman *et al.*, 2003), but their biparental inheritance and especially ITS conversion following allopolyploid events (Chase *et al.*, 2003) can blur species boundaries. Pillon (Y. Pillon, unpubl. data) studied the *D. maculata* complex (primarily *D. maculata* s.s. and *D. fuchsii*), the *D. incarnata* complex and their allotetraploid derivatives throughout their ranges, but particularly in western and northern Europe. We have made use of the markers (plastid microsatellites and ITS sequences) identified in this study. Shipunov *et al.* (2004) used the same markers to study general patterns of these same species and allotetraploid complexes in Russian Europe. The goal of this study is to explore diversity in detail in one of the Russian allotetraploid taxa, *D. baltica*, via morphological and molecular markers in the context of its likely origin via hybridization between the *D. fuchsii* and *D. incarnata* aggregates (both are treated as broadly defined species for simplicity). To the two sets of markers developed in Shipunov *et al.* (2004) and Y. Pillon (unpubl. data), we have added a set of two nuclear microsatellite markers, which we hope will be more variable than ITS and thus reveal more structure among populations of the putative parental taxa. We chose to focus on this allotetraploid taxon because it appeared to us that it was likely to be operating locally as a 'bridge' between the diploid taxa and would therefore make an appropriate subject for a more detailed study to determine whether we could detect evidence of this phenomenon though the study of both morphological and molecular markers.

## MATERIAL AND METHODS

Some of the samples were used in a previous study (Shipunov *et al.*, 2004), but many samples from European Russia (mostly from central and north-western regions) and Britain were newly collected for this

investigation (see Appendix). All incoming samples were initially identified and assigned to a priori species by experts in regional floras (G. Konechnaja and I. Kucherov in Botanical Institute, Saint-Petersburg; N. Reshetnikova in Main Botanical Garden, Moscow and M. Vakhrameeva in Moscow University). In total, 98, 214 and 775 samples from 85, 112 and 121 populations were involved in simultaneous combined and molecular and morphometric analyses, respectively. For most analyses, we used a subset of samples that consisted of the allotetraploid species together with the putative parental species, *D. fuchsii* s.l. and *D. incarnata* s.l. One herbarium sample of *D. baltica* (Smolensk region, A. Averyanov, 2000, LE) and one of *D. traunsteineri* (Saut. ex Reichb.) Soó (Karelia, I. Kucherov, 1999, LE) were used as yardsticks for morphological comparison and also for DNA extraction. One sample of *D. baltica* from Estonia in the RBG Kew DNA Bank (Chase 9485) was also used for sequencing.

## MOLECULAR MARKERS

Samples for DNA extraction were dried in silica gel (Chase & Hills, 1991). DNA was extracted by the 2 × CTAB protocol (Doyle & Doyle, 1987 but without an RNA treatment). PCR was performed with a set of primers designed by Y. Pillon & M. F. Fay (unpubl. data) to amplify four polymorphic plastid loci: Orch1, Msf, Ms1 and Ms2, located in three plastid DNA regions: the *trnS-trnG* spacer, *trnL* intron and *trnL-trnF* spacer. Two pairs of specific primers were also used to amplify length-variable regions of ITS ribosomal DNA that, taken together, indicate which ITS alleles are found in each sample (Shipunov *et al.*, 2004; Y. Pillon, unpubl. data).

To identify other molecular markers that are sufficiently polymorphic to reveal interpopulational structure, we have developed several nuclear microsatellites, two of which proved useful for this study. To develop these markers, we used a strategy proposed by Fisher, Gardner & Richardson (1996), which employs a degenerate primer PCT4 (Brachet *et al.*, 1999) that contains a (CT)<sub>6</sub> repeat at its 3' end. The conditions for PCR amplification were those of Fisher *et al.* (1996). Several PCR products were cloned using the Promega pGEM-T Easy Vector System. These were reamplified from transformed bacterial colonies by touching them with a sterile toothpick and using that sample as the template in a further round of PCR. Primers for this PCR were located on the vector. Amplified DNA fragments were purified using QIAquick PCR mini-columns (QIAGEN, Inc.), following the manufacturer's protocols, and sequenced on a 3100 genetic analyser (Applied Biosystems Inc.), following the manufacturer's protocols (we again used the primers that annealed to sites on the vector).

**Table 2.** Nuclear microsatellite loci used

Locus	Da963_1-2	Ds3978_2-1
Repeat	(CAG) <sub>5</sub>	(TTA) <sub>6</sub>
Annealing temperature, °C	52	52
Size range in base pairs (bp)	103–120	126–153
No. of alleles	5	10
Primer sequence (5'–3')		
Forward	TCCATATCCCCCTTCCTCAA	GAGATATATAGAGTGGTGGT
Reverse	CTCTCTCTCTTGTCTTTA	TATGCGTTGGTATTGGGAGT

Sequence editing and assembly of the two complementary strands used SequenceNavigator and Auto-Assembler (Applied Biosystems Inc.) software. Several pairs of specific primers were subsequently designed to amplify the most promising microsatellite loci. The resulting fragments were checked to determine whether they revealed any polymorphisms, and two loci were then chosen for this investigation (Tables 2, 3). The size of each fragment was determined using GeneScan and Genotyper software (Applied Biosystems Inc.). The subsequent unweighted pair-group (UPGMA) tree construction used PAUP\* version 4.0b10 (Swofford, 2000). For most statistical analyses, the calculation of distances between samples was based on the proportion of shared alleles.

#### MORPHOLOGY

We used the set of 14 morphological characters, slightly modified from previous work (Shipunov *et al.*, 2004). These characters were measured in nature on either the same plants that were used for DNA extractions or, on a few occasions (e.g. for *D. praetermissa* and several populations of *D. baltica*), we measured neighbouring plants in the same population. We used principal component analysis (PCA) and multidimensional scaling (MDS) of individual and population data. In the latter case, population medians (because these are usually more robust than means; Fowler, Cohen & Jarvis, 1999) were used. The analysis of population data was wider than the analysis of individual data because we included some species and populations for which DNA sampling and morphometric measurements were made on different plants. We have also analysed correlation from individual data (all species included) for all morphological measurements and nuclear DNA markers, and used recursive partitioning analysis, which is the model-based version of discriminant analysis, describing which character values best predict the existing classification (Breiman *et al.*, 1984). Statistical calculations used the R program, version 1.8 for

**Table 3.** Nuclear microsatellites alleles most typical for some *Dactylorhiza* species

Locus	Typical lengths (base pairs) and names for alleles		
	<i>D. incarnata</i>	<i>D. fuchsii</i>	<i>D. maculata</i>
Da963–1-2	110 (da110)	114 (da114),	103 (da103)
Ds3978–2-1	126 (ds126)	118 (da118), 130 (ds130), 153 (ds153)	143 (ds143), 147 (ds147)

Windows (Venables, Smith & R Development Core Team, 2002).

## RESULTS

### DNA MARKERS

Most *D. baltica* plants have the A haplotype (the unique combinations of plastid fragment lengths, in this case four, from different regions of the plastid genome, henceforth termed 'haplotypes'), which is typical of *D. fuchsii* (see Appendix), but several populations also contained the E and H haplotypes from *D. incarnata*. Most plants have more than one ITS allele, with the *D. incarnata* and *D. fuchsii* alleles being most common (several samples also have the *D. maculata* ITS allele). The nuclear microsatellite alleles are also mostly those of *D. incarnata* (da110, sometimes ds126) and/or *D. fuchsii* (da114, da118, ds130, ds153). For most population samples, multiple alleles were amplified (see Appendix); putative diploids displayed 1–2 alleles, and several of the allotetraploids had up to four alleles. Some samples collected as *D. fuchsii* (a diploid) have 3–4 alleles, but these plants are in fact 'northern tetraploids' (Shipunov *et al.*, 2004) and are morphologically intermediate between *D. maculata* and *D. fuchsii*. Several *D. incarnata* samples also displayed three alleles; most of them belong to populations 242 and 215, and

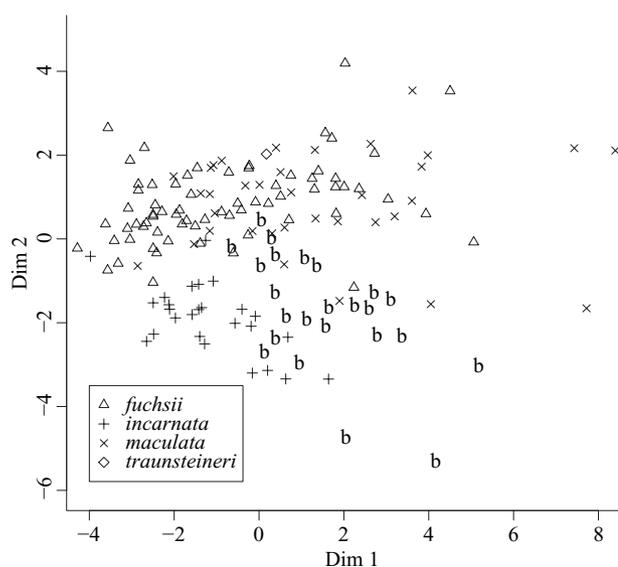
bear *D. fuchsii* alleles ds153, and da114–da118, respectively. In addition, plants from population 215 have the A plastid haplotype (see below). We can guess that these samples belong to triploids of likely hybrid origin.

Although the repeats selected were all in triplets (Table 2), the lengths of the fragments produced indicates that some of the variation we detected occurred in the flanking regions rather than just in the repeats; we did not verify this hypothesis by sequencing some of the variants. Nonetheless, some of the length variants (alleles) detected were diagnostic of the taxa studied, which made them useful markers. Analysis reveals that for nuclear markers overall polymorphism in *D. baltica* and *D. fuchsii* is higher than in *D. incarnata* ( $F = 2.89$  and  $2.39$ , respectively;  $P < 0.05$ ). The distribution of the nuclear microsatellite alleles is complex, and it is difficult to say which, if any, are typical for the parental taxa. Some are predominant in one species, e.g. da110 in *D. incarnata*, but then these also show up in other species occasionally (*D. fuchsii*) and commonly in *D. baltica*, which is expected since *D. baltica* has *D. incarnata* as one of its parents. When they occur in the other parent, they are often associated with the ITS allele of the other species as well, which thus provides two lines of evidence for introgression. Other microsatellite alleles are found in only some populations of one of the parental taxa (e.g. ds153 in some *D. fuchsii*) and then in some *D. incarnata* (often again with the *D. fuchsii* ITS allele) and some *D. baltica*. Such patterns show that local populations of all species, parental and allotetraploids, are likely to have similar microsatellite alleles. This pattern emerges particularly clearly in the UPGMA analysis (see below).

#### MULTIVARIATE ANALYSES

Both PCA and MDS of the morphological data revealed similar patterns. There are three overlapping groups (Fig. 2) consistent with haplotype and ITS allele distributions and species descriptions (*D. fuchsii* + *D. maculata*; *D. baltica* and *D. incarnata*) in general. Plants of the presumed autotetraploid *D. maculata* were not clearly separated from those of *D. fuchsii*. *Dactylorhiza baltica* plants are located not between the two putative parents but rather, are offset to the bottom right-hand corner of the graph. Several plants of *D. baltica*, however, overlap with the two parental groups. Similar results were obtained by Tyteca & Gathoye (1993) for *D. majalis* s.s., whereas *D. praetermissa* samples in their multivariate plot were located directly between the presumed parental species (again, *D. fuchsii* and *D. incarnata*).

*Dactylorhiza baltica* plants with haplotypes A and E are distributed closer to *D. fuchsii* (A haplotype) and



**Figure 2.** Multidimensional scaling of morphological data from individuals (species highlighted). All points marked with the letter 'b' correspond to *Dactylorhiza baltica*.

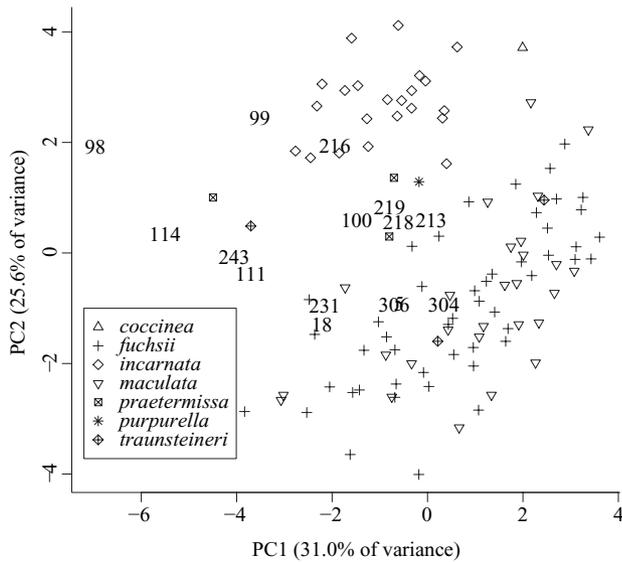
*D. incarnata* (E haplotype), respectively (not shown). The percentages of the *D. incarnata* ITS allele vary among *D. baltica* individuals; plants possessing more copies of this allele are usually located closer to *D. incarnata*. Many plants of *D. maculata* have other haplotypes, B, N or X, which is typical of this species throughout its range.

Population-level analysis revealed similar groups (Fig. 3), but for *D. baltica*, in this case, the offset was less and the diversity greater. We were able to include some additional species in this analysis so it could be seen that British populations of *D. praetermissa* and *D. purpurella* were located near the *D. baltica* points. Addition of other morphological characters would be likely to improve the separation of these allotetraploid taxa; the characters we selected were those that appeared to be good for separating *D. fuchsii* from *D. incarnata* and *D. baltica*. A population of *D. incarnata* ssp. *coccinea* (Pugsley) Soó from Wales was marginal to the *D. incarnata* group.

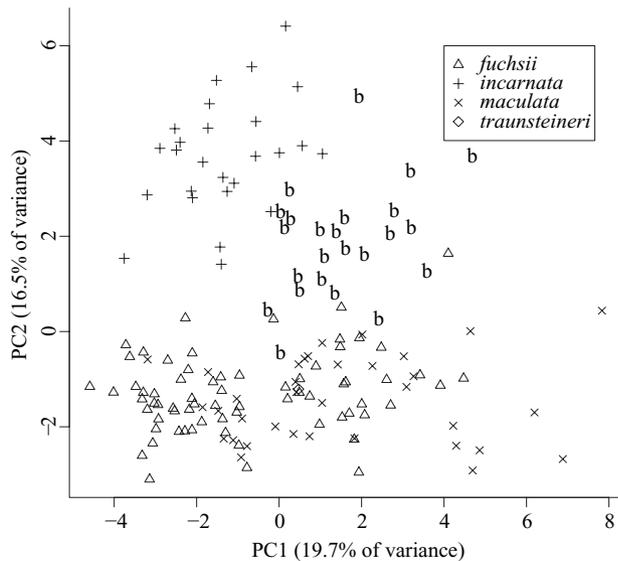
In both cases the most important characters (which have relatively high loadings in the first component, PC1) are for individuals, plant heights, all leaf characters and inflorescence lengths, and for populations, bract lengths, stem diameters and leaf lengths.

Simultaneous analysis of morphology, ITS alleles and nuclear microsatellites produced a less ambiguous structure (Fig. 4), both for individuals and populations (the latter not presented), demonstrating that there is agreement between these kinds of data.

An analysis of individuals combining all characters, including the uniparentally inherited plastid

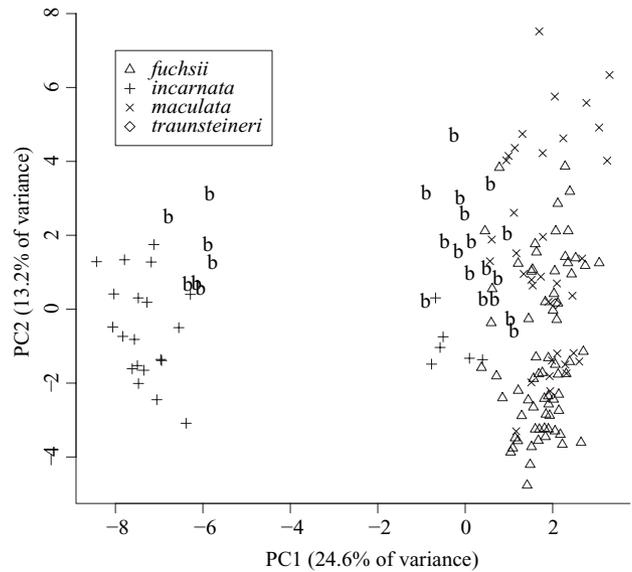


**Figure 3.** Principal component analysis of the morphological data for populations (species highlighted). All points marked with population numbers correspond to *Dactylorhiza baltica*.



**Figure 4.** Simultaneous principal component analysis of characters with biparental origin (morphology, ITS alleles and nuclear microsatellites) from individuals. All points marked with the letter 'b' correspond to *Dactylorhiza baltica*.

sequences, changed the picture completely (Fig. 5). *Dactylorhiza baltica* specimens were divided into two groups, each corresponding to their contrasting haplotypes and consequently, to their putative maternal parents. Some *D. incarnata* samples appeared close to



**Figure 5.** Simultaneous principal component analysis of all characters (morphology, plastid DNA markers, ITS alleles and nuclear microsatellites) from individuals. All points marked with the letter 'b' correspond to *Dactylorhiza baltica*.

*D. fuchsii*; these individuals, from population 215, have a *D. incarnata* morphology, but most have the A haplotype typical of *D. fuchsii* with a low frequency of the *D. incarnata* ITS allele. The corresponding analysis of populations generated a similar result (not presented).

To represent better some interpopulation relationships and possible geographical patterns, we constructed a UPGMA tree in PAUP for the nuclear DNA data (presence/absence of molecular markers) for populations of *D. baltica* and its putative parental species. This analysis showed that most *D. baltica* populations have clear relationships with their putative parents, either *D. fuchsii*, *D. incarnata* or even, in some cases, both. The tree (Fig. 6) demonstrates that most *D. baltica* populations share the same terminal clusters with nearby populations of *D. incarnata* and/or *D. fuchsii* (i.e. from adjacent regions, the same regions, or even the same collection sites). The tree cannot clearly distinguish between *D. fuchsii* and *D. incarnata*, but these species have been distinguishable in other molecular analyses (Bateman *et al.*, 2003; Y. Pillon, unpubl. data).

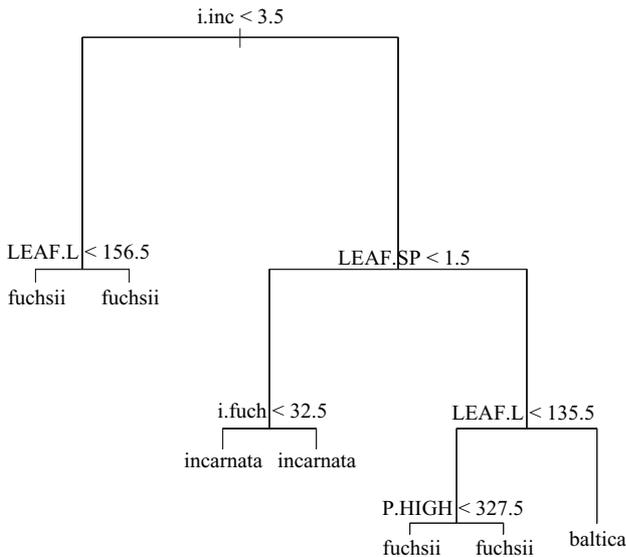
The morphological characters formed three correlation groups: (1) most of the vegetative characters, including bract and inflorescence length but not leaf spots, (2) leaf spots and (3) floral characters in which the largest significant correlation is between lateral lobe and mid-lobe lengths ( $r = 0.83$ ,  $P < 0.05$ ). The DNA characters most correlated with species parti-



tioning are percentage of *D. incarnata* ITS alleles, which formed a correlation group (4) with da110 and da118 nuclear microsatellite alleles. This group also has a significant correlation with groups (2) and (3).

Recursive partitioning revealed that the three characters most important in the analysis of data for *D. fuchsii*, *D. incarnata* and *D. baltica* (Fig. 7; misclassification error rate = 5.2%) were: (a) percentage of

*D. incarnata* ITS allele (the plants with less than 3.5% are *D. fuchsii*), (b) leaf spots (plants with unspotted leaves belong to *D. incarnata*) and (c) leaf length (plants with leaves shorter than 136 mm are *D. baltica*). Analysis of the morphological characters (misclassification error rate = 8.2%) alone added to this set another character (d): length of the lateral lobes of the lip (plants with lateral lobes less than 2 mm are mostly *D. incarnata*).



**Figure 7.** Tree of binary recursive partitioning analysis (a dichotomous key) of morphological and nuclear DNA characters in the model. The nodes are marked with character codes (see Table 4). Analysis performed on data from individuals of three species (*Dactylorhiza incarnata*, *D. fuchsii* and *D. baltica*).

## DISCUSSION

The position of *D. baltica* plants between *D. fuchsii* and *D. incarnata* in all analyses provides general support for the hypothesis of origin of *D. baltica* by hybridization between these two species. Although *D. maculata* and *D. fuchsii* are difficult to separate morphometrically, they are clearly distinct on a molecular basis, particularly plastid haplotypes, and it is clear that the markers in *D. baltica* are those of *D. fuchsii* and *D. incarnata*, with occasional markers (ITS alleles and nuclear microsatellites) from *D. maculata*. Some populations of *D. maculata* have the A haplotype of *D. fuchsii* rather than B and other haplotypes related to B that are typical of *D. maculata* throughout its range (Shipunov *et al.*, 2004; Y. Pillon, unpubl. data). In our sampled plants of *D. maculata*, the A haplotype is more common than the B haplotype, whereas no sample of *D. fuchsii* has the B haplotype; so, based purely on the results of our study, we could not clearly state that A is the haplotype of *D. fuchsii* and B that of *D. maculata*, although we know this to be the case from other studies.

The UPGMA results also demonstrate that *D. baltica* is related to *D. incarnata* and *D. fuchsii* (not *D. maculata*), but it is clear that in this region the two

**Table 4.** Morphological characters used (all measurements in millimetres)

Label	Description
P.HIGH	Plant height, from the ground to the top of inflorescence
LEAF.L	Length of longest leaf
LEAF.W	Width of longest leaf
L.WPOS	Position of maximal width (the distance from leaf base to the place of maximal width)
LEAF.SP	Leaf spots (0 none, 1 light, 2 heavy)
ST.DIAM	Stem diameter (measured just above the node of longest leaf)
INFL.L	Length from the lowest bract to the top of inflorescence
SPUR.L	Spur length, measured from lower side of spur
LIP.L	Lip length, from the base to the top of middle lobe
LIP.W	Lip maximum width
MIDD.L	Length of middle lobe of the lip, from the base of the sinus to the top apex of lobe
LATER.L	Length of lateral lobe of the lip, from the base of the sinus to the top apex of lobe
BR.L	Length of lowermost bract
LIP.COL	Lip colour (1 white or nearly white, 2 pink, 3 dark pink)

species of the *D. maculata* complex are difficult to distinguish, and pure populations of either are rare. At the same time, the *D. baltica* points are offset in all graphs, as if some independent evolution of this hybrid has occurred and is reflected in a morphological bias. Hedrén *et al.* (2001) also noticed in their AFLP study the same type of offset pattern for *Dactylorhiza* allotetraploids relative to their putative parents. It is also possible that this offset pattern could be generated by introgression from another (or other) species, but we favour the former explanation because we do not find markers of any other species of *Dactylorhiza* in *D. baltica*; Y. Pillon (unpubl. data) examined haplotype and ITS alleles for a complete set of European species, and we know which markers these other species exhibit. In the investigation of Tyteca & Gathoye (1993), allotetraploid *D. majalis* s.s. has the same position relative to its presumed parents in the graph, but in this case, we have evidence of some unique markers (notably, the C haplotype) in this species (Y. Pillon, unpubl. data). On other hand, Kulikov & Filippov (1999a) observed that a southern Siberian ( $2n = 40$ ) diploid (and probably recent) hybrid between *D. incarnata* and *D. fuchsii* (*D. × intermedia* (Serg.) Kulikov et Filippov) has a distinct position in multivariate morphological analyses of these species, despite many morphological similarities to *D. baltica*. Additionally, the position of some long-recorded *D. baltica* populations (213 and 218, both homogeneous in nuclear markers) are placed exactly between the parents (Fig. 3). Therefore, the presence or absence of the offset cannot distinguish between primary hybrids and stable hybrid species. The uniparentally inherited plastid markers are able to distinguish among 'ordinary' nonhybrid species (Shipunov *et al.*, 2004), and they also reveal multiple origins of allotetraploids such as *D. baltica* because most populations investigated contain at least two haplotypes (see Appendix).

Several *D. baltica* plants and populations are embedded in the *D. fuchsii* or (less often) *D. incarnata* groups in the multivariate analyses (Figs 2–5), which could be evidence of backcrossing of *D. baltica* and introgression. Occurrence of introgression between *D. incarnata* and *D. fuchsii* is supported by finding the *D. incarnata* ITS allele in some true *D. fuchsii* plants (and sometimes vice versa) from neighbouring populations. Introgression from *D. maculata* into *D. baltica* is less likely to occur, but in a few cases we found *D. maculata* ITS and microsatellite alleles in individuals of *D. baltica* (see Appendix). It is likely that these cases occur because of hybridization between *D. maculata* and *D. fuchsii*, which then subsequently permits the *D. maculata* markers to enter *D. baltica*; in no case did we find any unique *D. maculata* haplotypes (B, N or X) in *D. baltica*, and the frequency of other typical

*D. maculata* DNA markers was low compared with those of *D. incarnata* and *D. fuchsii*. Although *D. maculata* and *D. fuchsii* are often considered morphologically similar, in other studies (especially in western and north-western Europe) they have been shown to be separable with morphometric and molecular approaches (Bateman & Denholm, 2003; Y. Pillon, unpubl. data). Here, the selected morphometric measures that separate *D. fuchsii*, *D. incarnata* and *D. baltica* are less effective in separating *D. maculata* from *D. fuchsii*. This observation could also be explained by hybridization and likely introgression between these species, especially in northern Russia (Shipunov *et al.*, 2004) where most of the overlapping samples originated and allotetraploids formed by these two species occur in large populations.

Exchange of alleles/haplotypes between the parental diploids is consistent with the allotetraploids forming a 'bridge' for gene flow between the two diploids. There is also some correlation between the proportion of *D. incarnata* ITS allele (determined from the peak height in the PCR of the length-variable ITS fragments) in *D. baltica* plants and their morphology, which could be explained by backcrossing with *D. incarnata*. Some *D. baltica* samples (especially from north-western populations) lack or have only a small proportion of *D. fuchsii* (da114 or da118) nuclear microsatellite alleles, which again could be evidence of backcrossing. The *D. incarnata* plants from populations 210 and 215 with the A haplotype fall into the *D. fuchsii* group in the combined analysis (Fig. 5); this situation could result from a solitary introgression event, leading in particular to 'plastid capture' (Jackson *et al.*, 1999) of the *D. fuchsii* plastid genome. It is also possible that some of the *D. baltica* individuals are primary, diploid hybrids between *D. fuchsii* and *D. incarnata*, which would mean that exchange is taking place directly rather than through allotetraploids. It is easier to imagine that *D. baltica* is acting as a bridge because the flowering periods of *D. fuchsii* (late) and *D. incarnata* (early) have little overlap whereas *D. baltica* overlaps both because it is intermediate. We do not know the frequencies of diploids, triploids and tetraploids in *D. baltica* populations, so we cannot say that it is the allotetraploids that are the conduit for the introgression that we detected, but this seems most likely. If triploids formed by crosses between *D. baltica* and either *D. fuchsii* or *D. incarnata* produce some diploid gametes, then these could lead to the introgression we detected in both *D. fuchsii* and *D. incarnata*. Some of the plants identified as *D. incarnata* had three microsatellite alleles, which means that these are likely to be triploids. Some of these individuals were also morphologically more similar to *D. baltica*, which also supports their status as backcrosses between

*D. baltica* and *D. incarnata*. Such individuals will have two doses of the *D. incarnata* genome, which means that they should be morphologically more similar to *D. incarnata* rather than to *D. baltica*, resulting in their identification as *D. incarnata*.

Orchids in general, including *Dactylorhiza*, produce thousands of ovules, and fertilization involves deposition of large parts, often whole masses of pollen, so even though diploid gametes might be rare in triploids, they could be common enough for this route to be effective in producing introgression. Further evidence of introgression involving the allotetraploids is found in the UPGMA results, which show geographical clusters of all taxa sharing nuclear microsatellite markers unique to that region (Fig. 6).

There is some evidence of the association between *D. baltica*, *D. praetermissa* and *D. purpurella* that supports Averyanov's (1990) classification of *Dactylorhiza*, but this conclusion needs clearer support from more thorough sampling of these taxa. On the other hand, *D. praetermissa* (north-western and western Europe) differs from *D. baltica* in its high frequency of an ITS allele never found in *D. fuchsii* or *D. incarnata* but sometimes found in *D. saccifera* (Brongn.) Soó from south-eastern Europe (Italy, Greece) through to Turkey (Y. Pillon, unpubl. data). Like *D. baltica*, *D. purpurella* also has multiple (and probably recent) origins from *D. incarnata* and *D. fuchsii* (Y. Pillon, unpubl. data). Another allotetraploid found in eastern Europe is *D. traunsteineri*, but our data are insufficient for a genetic comparison with *D. baltica*.

If the various allotetraploids that have been consistently recognized are the products of repeated crosses between the same two parental species, it becomes difficult to understand how their morphological and sometimes ecological differences have arisen. For example, we see in this study that the Russian *D. fuchsii* plants bear the same haplotype and ITS alleles as those in western Europe (Y. Pillon, unpubl. data). Nonetheless, the differences observed between allotetraploid species with same species-level parentage could be based on geographical variability of the parents, most importantly *D. incarnata*. This species has the E haplotype in western Europe, whereas many Russian samples of *D. incarnata* have the H haplotype (Shipunov *et al.*, 2004). Despite clear morphological variability, *D. incarnata* has previously been found to be highly genetically uniform in most allozymes and AFLPs studies (Hedrén *et al.*, 2001; Hedrén, 2002, 2003), whereas *D. fuchsii* has been more variable locally, although a similar spread of variation occurs throughout the range of the latter species. On the other hand, *D. incarnata* has a considerable range of morphological variation and more named varieties than *D. fuchsii*; thus, we suspect that the distinctiveness of the allotetraploids is a result of this variation

across its range. We hope that by developing more microsatellite loci we can dissect the variation in *D. incarnata* and *D. fuchsii* and thus understand better the intriguing variation in morphology and ecology observed in the allotetraploids.

The geographical patterns reflected in the UPGMA tree of nuclear DNA data (Fig. 6) provide additional evidence for the hypothesis of multiple origins of these tetraploids (the probability of these plants being triploids is low because most of our *D. baltica* samples represent homogeneous populations). Plants of *D. baltica* associate with nearby populations of one of its two parental taxa, and these could be similar to (or even be) the parents of the *D. baltica* populations. These origins might thus be local, and perhaps even restricted to the area of the collection sites (typically around 100–500 km<sup>2</sup>). The alternative explanation for these patterns is that all plants from the localities cluster together because they share alleles due to local hybridization and introgression. Distinguishing between these alternate explanations is difficult.

The analyses of character loading, correlation and regression trees (Fig. 7) support the main morphological characters used in *Dactylorhiza* classification, especially in recent morphometric studies (Bateman & Denholm, 1983, 1985, 1989, 2003; Reinhard, 1990; Tyteca & Gathoye, 1993). The result of the character analysis led us to propose that identification of *D. baltica* is possible from single, unmodified characters such as leaf spots, length of longest leaf and length of lateral lip lobes; all of these characters are mentioned as diagnostic in Table 1 and yield the highest loadings in our PCA (Figs 3–5). The correlation groups also demonstrate the relative independence of floral characters, leaf spots and the rest of the vegetative characters measured. Our study displays the good agreement between the results of simultaneous molecular and morphometric analyses, which demonstrates the advantage of 'combined' techniques, especially in cases of complex taxa with frequent hybridization and/or introgression.

#### ACKNOWLEDGEMENTS

We are grateful to all the colleagues and friends who helped us in collection of *Dactylorhiza* plant material and morphometric data: L. Abramova, Zh. Altshuler, I. Blinova, T. Braslavskaja, S. Glagolev, N. Guryanova, A. Karjakin, G. Konechnaja, J. Kosenko, E. Kost, I. Kucherov, A. Kvashenko, M. Logacheva, K. Markvicheva, S. Nazarova, A. Pegova, E. Peskova, S. Polevova, N. Reshetnikova, N. Rimskaja-Korsakova, N. Sajtanova, A. Shipunova, A. Skorobogatov, D. Sokoloff, S. Suhov, D. Suhova, M. Vakhrameeva, T. Vinogradova, P. Volkova, T. Vorozhbieva and I. Yufryakov. We also thank R. Cowan, L. Csiba, J. Clarkson, E.

Kapinos and Y. Pillon for their valuable help and advice with the laboratory and computer analyses, and R. Bateman for his many comments on the manuscript during review. This work was supported by the Royal Society of London and sponsored by NATO and the British Foreign and Commonwealth Office.

## REFERENCES

- Averyanov LV. 1990.** A review of genus *Dactylorhiza* Neck. ex Nevski (Orchidaceae), 3 [in Russian]. *Novosti Sistematiki Vysshikh Rastenij* **27**: 32–62.
- Bateman RM. 2001.** Evolution and classification of European orchids: insights from molecular and morphological characters. *Journal Europäischer Orchideen* **33**: 33–119.
- Bateman RM, Denholm I. 1983.** A reappraisal of the British and Irish dactylorchids, 1. The tetraploid marsh-orchids. *Watsonia* **14**: 347–376.
- Bateman RM, Denholm I. 1985.** A reappraisal of the British and Irish dactylorchids, 2. The diploid marsh-orchids. *Watsonia* **15**: 321–355.
- Bateman RM, Denholm I. 1989.** A reappraisal of the British and Irish dactylorchids, 3. The spotted-orchids. *Watsonia* **17**: 319–349.
- Bateman RM, Denholm I. 2003.** The heath spotted-orchid (*Dactylorhiza maculata* (L.) (Soó) in the British Isles: a cautionary case-study in delimiting infraspecific taxa and inferring their evolutionary relationships. *Journal Europäischer Orchideen* **35**: 3–36.
- Bateman RM, Hollingsworth PM, Preston J, Luo YB, Pridgeon AM, Chase MW. 2003.** Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Botanical Journal of the Linnean Society* **142**: 1–40.
- Brachet S, Jubier MF, Richard M, Jung-Muller B, Frascaria-Lacoste N. 1999.** Rapid identification of microsatellite loci using 5'-anchored PCR in the common ash *Fraxinus excelsior*. *Molecular Ecology* **8**: 157–168.
- Breiman L, Friedman JH, Olshen RA, Stone CJ. 1984.** *Classification and regression trees*. New York: Chapman & Hall.
- Chase MW, Hills HG. 1991.** Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* **40**: 215–220.
- Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, Joseph J, Savolainen V, Parokonny AS. 2003.** Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). *Annals of Botany* **92**: 107–127.
- Delforge P. 1995.** *Orchids of Britain and Europe*. London: Harper Collins.
- Devos N, Tyteca D, Raspé O, Wesselingh RA, Jacquemart AL. 2003.** Patterns of chloroplast diversity among western European *Dactylorhiza* species (Orchidaceae). *Plant Systematics and Evolution* **243**: 85–97.
- Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of leaf tissue. *Phytochemical Bulletin* **19**: 11–15.
- Fisher PJ, Gardner RC, Richardson TE. 1996.** Single locus microsatellites isolated using 5'-anchored PCR. *Nucleic Acids Research* **24**: 4369–4371.
- Fowler J, Cohen L, Jarvis P. 1999.** *Practical statistics for field biology*. 2nd edn. Chichester: John Wiley & Sons.
- Hedré M. 2002.** Speciation patterns in the *Dactylorhiza incarnata/maculata* polyploid complex (Orchidaceae): evidence from molecular markers. *Journal Europäischer Orchideen* **34**: 707–731.
- Hedré M. 2003.** Plastid DNA variation in the *Dactylorhiza incarnata/maculata* polyploid complex and the origin of allotetraploid *D. sphagnicola* (Orchidaceae). *Molecular Ecology* **12**: 2669–2680.
- Hedré 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.
- Heslop-Harrison J. 1968.** Genetic system and ecological habit as factors in dactylorchid variation. *Jahresberichte des Naturwissenschaftlichen Vereins in Wuppertal* **21–22**: 20–27.
- Jackson HD, Steane DA, Potts BM, Vaillancourt RE. 1999.** Chloroplast DNA evidence for reticulate evolution in *Eucalyptus* (Myrtaceae). *Molecular Ecology* **8**: 739–751.
- Klinge J. 1895.** *Orchis latifolia* subsp. *baltica*. In: Lehmann E, ed. *Flora von Polnisch-Livland*. Jurjew: C. Mattiensen, 188.
- Klinge J. 1898.** *Dactylorchidis*, orchidis subgeneris, monographiae prodromus. *Acta Horti Petropolitani* **17**: 56.
- Kulikov PV, Filippov EG. 1999a.** *Dactylorhiza baltica* (Klinge) Orlova complex (Orchidaceae) in the Urals and western Siberia. *Bulletin of the Moscow Society of Naturalists, Biological Series* **104**: 29–33.
- Kulikov PV, Filippov EG. 1999b.** *Dactylorhiza* aggr. *traunsteineri* (Saut.) Soó complex (Orchidaceae) in the Urals: taxonomic structure and distribution. *Bulletin of the Moscow Society of Naturalists, Biological Series* **104**: 61–65.
- Nevski SA. 1935.** Orchidaceae Lindl. In: Komarov VL, ed. *Flora USSR*. Vol. 4. Leningrad: Academy of Science Publisher, 589–730.
- Reinhard HR. 1990.** Kritische Anmerkungen zu Einigen *Dactylorhiza*-Arten (Orchidaceae) Europas. *Mitteilungsblatt AHO Baden-Württemberg* **22**: 1–72.
- Senghas K. 1968.** Taxonomische Übersicht der Gattung *Dactylorhiza* Necker ex Nevski. *Jahresberichte des Naturwissenschaftlichen Vereins in Wuppertal* **21–22**: 32–67.
- Shipunov AB, Fay MF, Pillon Y, Bateman RM, Chase MW. in press.** *Dactylorhiza* (Orchidaceae) in European Russia: combined molecular and morphological analysis. *American Journal of Botany* in press.
- Smoljaninova LA. 1976.** *Dactylorhiza* Nevski [in Russian]. *Flora Evropejskoj Chasti SSSR* **2**: 49–57.
- Soó R. 1980.** *Dactylorhiza* Necker ex Nevski. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA, eds. *Flora Europaea*. Vol. 5. Cambridge: Cambridge University Press, 333–337.
- Stace CA. 1997.** *New flora of the British Isles* 2nd edn. Cambridge: Cambridge University Press.

**Swofford DL. 2000.** *PAUP\*: phylogenetic analysis using parsimony*, Version 4.0. Washington, D.C.: Smithsonian Institution

**Tyteca D, Gathoye JL. 1993.** On the morphological variability of *Dactylorhiza praetermissa* (Druce) Soó (Orchidaceae). *Belgian Journal of Botany* **126**: 81–99.

**Venables WN, Smith DN, the R Development Core Team. 2002.** *An introduction to R*. Bristol: Network Theory Ltd.

**Vermeulen P. 1947.** *Studies in dactylorchids*. Utrecht: Schotanus and Jens.

## APPENDIX 1

Dactylorhizid populations sampled for use in at least one type of analysis

Species	Population(s)	Collection site	Locality
<i>baltica</i>	5	Middle Russia, Tver' region (Tv)	Udomlja District, Lake Moldino environs
<i>baltica</i>	18	Middle Russia, Smolensk region (Sm)	Przheval'sk District, Lake Chistik environs
<i>baltica</i>	98–100	Middle Russia, Orel region (Or)	Znamenskoe District, Vytebet' River environs
<i>baltica</i>	111	Middle Russia, Smolensk region (Sm)	Demidov District, Lake Rzhavez environs
<i>baltica</i>	114, 213	North-west Russia, St Petersburg region (Sa)	St Petersburg (in the city)
<i>baltica</i>	216–219	North-west Russia, Pskov region (Ps)	Sebez District
<i>baltica</i>	231	North-west Russia, St Petersburg region (Sa)	St. Petersburg, Strel'na settlement
<i>baltica</i>	243	Middle Russia, Kaluga region (Ka)	Kozel'sk district, Zhizdra River environs
<i>baltica</i>	304, 306	Middle Russia, Moscow region (Mo)	Volokolamsk District
<i>fuchsii</i>	9, 10, 12, 325–331	Tv	Udomlja District, Lake Moldino environs
<i>fuchsii</i>	14, 334	Tv	Vyshnij Volochek District, Lake Ol'shevo environs
<i>fuchsii</i>	19	Sm	Przheval'sk District, Lake Chistik environs
<i>fuchsii</i>	22–24	Northern Russia, Archangelsk region	Ust'jansk District
<i>fuchsii</i>	30, 31, 33–35, 60	Northern Russia, Northern Karelia	Chupa District
<i>fuchsii</i>	73	Mo	Odintsovo District
<i>fuchsii</i>	76, 241	Ka	Kozel'sk district, Zhizdra River environs
<i>fuchsii</i>	77	Mo	Ruza District
<i>fuchsii</i>	101	Or	Znamenskoe District, Vytebet' River environs
<i>fuchsii</i>	133	Northern Russia, Murmansk region	Kirovsk District
<i>fuchsii</i>	206	England	Avon, Bristol
<i>fuchsii</i>	212	Sa	Primorsk District
<i>fuchsii</i>	221, 222	Sa	Volhov District
<i>fuchsii</i>	302, 303, 305, 307, 308, 310–314, 321, 322	Mo	Volokolamsk District
<i>fuchsii</i>	2011	Wales	South Glamorgan
<i>incarnata</i>	4, 6, 7	Tv	Udomlja District, Lake Moldino environs
<i>incarnata</i>	8, 15, 16, 323, 332, 333	Tv	Vyshnij Volochek District, Lake Ol'shevo environs
<i>incarnata</i>	17	Sm	Przheval'sk District, Lake Chistik environs
<i>incarnata</i>	25	Northern Russia, Archangelsk region	Ust'jansk District
<i>incarnata</i>	70	Middle Russia, Tambov region	Pervomajsk District
<i>incarnata</i>	72, 209–211	Mo	Odintsovo District
<i>incarnata</i>	71	Mo	Taldom District
<i>incarnata</i>	81	Mo	Ruza District
<i>incarnata</i>	215	Ps	Sebez District
<i>incarnata</i>	232	Sa	Gatchina District
<i>incarnata</i>	242	Ka	Kozel'sk district, Zhizdra River environs
<i>incarnata</i>	251, 252	Northern Russia, Murmansk region	Kandalaksha District
<i>incarnata</i>	309	Mo	Volokolamsk District

APPENDIX 1 *Continued*

Species	Population(s)	Collection site	Locality
<i>incarnata</i>			
<i>ssp. coccinea</i>	2042	Wales	Dyfed, Ynyslas
<i>maculata</i>	13, 324	Ty	Udomlja District, Lake Moldino environs
<i>maculata</i>	20	Sm	Przheval'sk District, Lake Chistik environs
<i>maculata</i>	21		Ustjansk District
<i>maculata</i>	32, 36, 37, 39, 47–50, 257, 258	Northern Russia, Archangelsk region	Chupa District
<i>maculata</i>	40–46	Northern Russia, Northern Karelia	Kandalaksha District
<i>maculata</i>	90–94	Northern Russia, Murmansk region	Chupa District
<i>maculata</i>	74, 75	Mo	Odintsovo District
<i>maculata</i>	79, 60, 130, 132	Northern Russia, Murmansk region	Kirovsk District
<i>maculata</i>	197	England	Cornwall, Traboe Moors
<i>maculata</i>	198	England	Cornwall, The Lizard
<i>maculata</i>	200	England	Cumbria, Duddon Valley
<i>maculata</i>	203	Wales	Dyfed, Cors Fochno
<i>maculata</i>	214	Ps	Sebezh District
<i>maculata</i>	223, 224	Sa	Sosnovyj Bor District
<i>maculata</i>	233	Sa	Pavlovsk
<i>maculata</i>	253, 254, 256	Northern Russia, Murmansk region	Kandalaksha District
<i>praetermissa</i>	201	Wales	South Glamorgan
<i>praetermissa</i>	202	Wales	Carmarthenshire
<i>praetermissa</i>	204, 205	Wales	Dyfed, Ynyslas
<i>purpurella</i>	204, 205	Wales	Dyfed, Ynyslas
<i>traunsteineri</i>	78, 131	Northern Russia, Murmansk region	Kirovsk District
<i>traunsteineri</i>	110	Northern Russia, Karelia	Onego Lake District







