Received: 8 November 2019 | Revised: 25 July 2020 | Accepted: 5 August 2020

DOI: 10.1111/mec.15848

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Asymmetric contributions of seed and pollen to gene dispersal in the marsh orchid *Dactylorhiza umbrosa* in Asia Minor

Mikael Hedrén[1] | Sven Birkedal[2] | Hugo de Boer[3,4,5] | Abdolbaset Ghorbani[4] | Barbara Gravendeel[5] | Sven Hansson[6] | Åke Svensson[7] | Shahin Zarre[8]

[1]Department of Biology, University of Lund, Lund, Sweden

[2]Lingonstigen 8, Åhus, Sweden

[3]Natural History Museum, University of Oslo, Oslo, Norway

[4]Department of Organismal Biology, Uppsala University, Uppsala, Sweden

[5]Endless Forms Group, Naturalis Biodiversity Center, Leiden, The Netherlands

[6]Koltrastvägen 2A, Täby, Sweden

[7]Department of Dermatology, Malmö University Hospital SUS, Malmö, Sweden

[8]Department of Plant Sciences, School of Biology, College of Science, University of Tehran, Tehran, Iran

Correspondence
Mikael Hedrén, Department of Biology, University of Lund, Sölvegatan 37, SE-223 62 Lund, Sweden.
Email: mikael.hedren@biol.lu.se

Funding information
Elly Olsson Foundation; Swedish Science Council – Swedish Research Links, Grant/Award Number: D0664801; Crafoordska Stiftelsen

Abstract

Orchids differ from other plants in their extremely small and partly air-filled seeds that can be transported long distances by wind. Seed dispersal in orchids is expected to contribute strongly to overall gene flow, and orchids generally express low levels of genetic differentiation between populations and low pollen to seed flow ratios. However, studies in orchids distributed in northern Europe have often found a poor geographic structuring of genetic variation. Here, we studied geographic differentiation in the marsh orchid *Dactylorhiza umbrosa*, which is widely distributed in upland regions from Asia Minor to Central Asia. These areas were less affected by Pleistocene ice ages than northern Europe and the orchid should have been able to survive the last ice age in local refugia. In the plastid genome, which is dispersed by seeds, populations at close distance were clearly divergent, but the differentiation still increased with geographic distance, and a significant phylogeographic structure had developed. In the nuclear genome, which is dispersed by both seeds and pollen, populations showed an even stronger correlation between genetic and geographic distance, but average levels of differentiation were lower than in the plastid genome, and no phylogeographic structure was evident. Combining plastid and nuclear data, we found that the ratio of pollen to seed dispersal (*mp*/*ms*) decreases with physical distance. Comparison with orchids that grow in parts of Europe that were glaciated during the last ice suggests that a balanced structure of genetic diversity develops only slowly in many terrestrial orchids, despite efficient seed dispersal.

KEYWORDS
Dactylorhiza umbrosa, genetic diversity, nuclear microsatellites, plastid DNA, pollen to seed dispersal ratio, spatial genetic structure

1 | INTRODUCTION

Gene dispersal in higher plants is the combined effect of pollen and seed dispersal. Together with colonization history, demography, selection and various environmental factors, gene dispersal patterns determine the genetic structure of a species (Hamrick & Godt, 1990, 1996). However, the magnitude of pollen and seed dispersal differ radically among species (Petit et al., 2005), and, moreover, the relative importance of seed and pollen dispersal to overall gene dispersal are expected to differ over temporal and spatial scales (Austerlitz et al., 2000; Hamrick & Trapnell, 2011; McCauley, 1997).

Seeds or other diaspores are required for dispersal and colonization of new sites, so the genetic structure of recently colonized regions of a species' distribution is primarily influenced by seed dispersal patterns (Hamrick & Trapnell, 2011; McCauley, 1995; Slatkin, 1977). In the early phase of establishment, individual populations founded by one or a few seeds are likely to be genetically homogeneous, whereas populations founded by several seeds could be genetically variable (Wade & McCauley, 1988). If the seeds founding new populations originate from afar, then adjacent sites are likely to be founded by seeds from different source areas and could be genetically divergent, but if seeds originate from close sources, then adjacent sites could be founded by related genotypes and the new populations would be less genetically differentiated (Hedrén & Nordström Olofsson, 2018). Early patterns of divergence may become less pronounced with time, as pollen and seeds from neighbouring populations contribute additional genes to the local populations. However, in most plant species, pollen dispersal seems to be more efficient than seed dispersal for transporting genes between populations (Petit et al., 2005), so differences in the organellar genomes dispersed by seeds takes longer to even out.

If conditions remain stable long enough, populations at close distance are expected to exchange genes with each other and become more genetically similar than population at long distance. As a result, a correlation between genetic and geographic distances between populations, i.e. an isolation by distance structure, IBD, will develop (Wright, 1943). Furthermore, if mutation rates are high in comparison to gene flow, a phylogeographic structure might eventually emerge (Pons & Petit, 1996), meaning that distinct genes in distant populations will be more diverged that distinct genes in close populations (or that distinct genes will be more divergent in different populations than distinct genes within populations; Hardy & Vekemans, 2002).

In contrast, if local populations are parts of a metapopulation system with frequent extinctions and recolonizations, the effects of the colonization process would become more prominent, and an isolation by distance structure might not appear (McCauley, 1995). In such a system, seed dispersal patterns will continue to have a strong influence on the structuring of genetic diversity (McCauley, 1997). Still, if a species is examined over a large enough range, extinction/recolonization processes can proceed at a regional scale, while more distant populations keep diverging from each other, and so an isolation by distance structure could still be observed within such a species.

Pollen and seeds are expected to show different dispersal distributions (McCauley, 1997). Most plants show strongly leptokurtic patterns of seed dispersal, and whereas the majority of seeds end up close to the mother individual, a small fraction could be dispersed very far (Cain et al., 2000). Although most of the seeds in the latter category are likely to end up in unsuitable habitats, some could nevertheless contribute to colonization of new sites and expansion of the species' range. In contrast, pollen dispersal ranges decline more gradually, and pollen dispersal seems to be more efficient than seed dispersal within populations and between adjacent populations. Accordingly, the relative contribution to gene flow by pollen and seeds is expected to be scale-dependent and to decrease with increasing geographic distance (Hamrick & Trapnell, 2011; McCauley, 1997). Moreover, as the dispersal distributions may extend over different ranges for pollen and seeds, depending on the mechanism of pollen and seed dispersal, the relative contributions of pollen and seeds for any particular distance range may differ substantially between species (Petit et al., 2005).

Orchids differ from other plants in having extremely small and dust-like seeds. The seeds are typically shorter than 1 mm, contain a minute embryo with no endosperm, and are mostly airfilled (Arditti & Ghani, 2000). Most species are wind-dispersed (Arditti & Ghani, 2000; Pridgeon et al., 1999), and there is evidence that the seeds can be transported more than a hundred kilometers to colonize new sites (Arditti & Ghani, 2000; Hedrén et al., 2018; Partomihardjo, 2003; Vanden Broeck et al., 2014). Seed dispersal is also considered as an important component to gene flow between established populations (Vanden Broeck et al., 2014).

It is generally assumed that orchid seeds contribute disproportionally much to overall gene dispersal, both in absolute and relative terms. On average, orchids express low levels of genetic differentiation between populations in comparison to other angiosperm families dominated by herbaceous species (Forrest et al., 2004; Hamrick & Godt, 1996; Phillips et al., 2012). The efficiency of seed dispersal in orchids is also manifested in low pollen to seed flow ratios, often being close to one (Cozzolino et al., 2003; Squirrel et al., 2001).

As a consequence of their putatively efficient gene dispersal, orchids are expected to rapidly erase patterns associated with early colonization and reach genetic equilibria manifested as e.g., high genetic similarity between neighbouring populations and isolation by distance patterns. However, these expectations are not always met. In two recent studies treating northern European members of the terrestrial orchid genus *Dactylorhiza*, we did not detect any significant isolation by distance structure but found a mosaic-like distribution of plastid haplotypes, implying that gene flow by seeds have not been very efficient (Hedrén & Nordström Olofsson, 2018; Hedrén et al., 2018). Likewise, other studies of terrestrial orchids in Europe have found that populations at the northern margins of their distribution have still not fully erased patterns imprinted from early colonization (Hedrén, 2009; Pfeifer et al., 2009). These studies were conducted in parts of Europe that were unavailable to orchids during the Weichselian ice age and that were colonized by long-distance dispersal within the last 10,000–15,000 years. It is possible therefore that species inhabiting these areas today could still be affected by colonization history and may not yet have reached genetic equilibrium (Hedrén & Nordström Olofsson, 2018; Hedrén et al., 2018).

Here, we have chosen to study genetic differentiation patterns in the terrestrial marsh orchid *Dactylorhiza umbrosa* (Kar. & Kir.) Nevski. This species has a wide distribution in upland regions of Asia Minor and Central Asia, allowing us to study genetic differentiation patterns at various geographic scales. The upland regions of Asia inhabited by *D. umbrosa* were less affected by glaciers than northern Europe during the last ice age (Sarikaya et al., 2011; van Zeist, 1967),

and the species should have been able to survive the last ice age within the present distribution area. Accordingly, the genetic structure in *D. umbrosa* is possibly closer to a long-term equilibrium state than in European species of *Dactylorhiza*.

We use molecular markers to test whether the genetic structure in *D. umbrosa* differs from that of *Dactylorhiza* species distributed in Europe. The contributions of seed and pollen dispersal to total gene dispersal could be determined by combining data from the plastid (chloroplast) genome with data from the nuclear genome (Ennos, 1994; McCauley, 1994; Petit et al., 2005). Since the plastid genome is exclusively inherited from the maternal parent in most angiosperms (Corriveau & Coleman, 1988), data from this genome is sufficient to describe gene flow patterns by seeds (McCauley, 1995). The nuclear genome is dispersed both by seeds and pollen, but by comparison with data from the plastid genome, the contribution by pollen dispersal to total gene dispersal could also be estimated (Ennos, 1994; Petit et al., 2005).

We predict that *D. umbrosa* should have experienced a long enough period of gene flow to develop a genetic structure reflecting that of geographic distances between populations, i.e. an isolation by distance structure, IBD (Wright, 1943). We also predict that the genetic differentiation will be stronger in the plastid genome, since this genome is dispersed by seeds only, whereas the nuclear genome is dispersed by both pollen and seeds. We combine data from the two genomes to estimate the ratio of pollen versus seed mediated gene dispersal. Third, we calculate genetic parameters separately for different geographic distance intervals and test whether pollen to seed dispersal ratios will decrease with increasing geographic distance separating populations. Lastly, we compare *D. umbrosa* with other members of the orchid family in an attempt to understand the importance of colonization history and habitat stability on the structuring of genetic diversity in orchids in general.

2 | MATERIALS AND METHODS

2.1 | Study species

Dactylorhiza umbrosa belongs to the genus *Dactylorhiza*, which comprises terrestrial orchids distributed in temperate parts of Eurasia (Pridgeon et al., 2001). Species of *Dactylorhiza* are erect plants reaching about half a metre and have a single leafy stem with a lax to dense cylindrical inflorescence of about 10–50 flowers. The plant develops from an underground tuber formed during the previous years' growth. After flowering, a new tuber will develop to continue the growth the next-coming year. *Dactylorhiza* species are thus perennial, but it is not possible to tell the age of a plant from its size. However, demographic studies in the European *D. majalis* ssp. *lapponica* (Øien & Moen, 2002) indicate that plants could have an average life-span of about 10 years. The flowers of *Dactylorhiza* are showy with a broad lip and a conical-cylindrical spur. The pollen is separated from the stigma within the flower, and the flowers are seemingly adapted for cross-pollination. The dominant group of pollinators is bumble-bees (Figure S1), but other insects including Diptera and Coleoptera have also been recorded among European species (Nilsson, 1981). The flowers still offer no reward to the visiting insects and act by deceit, a condition believed to result in longer flight distances between plants and more efficient pollen dispersal (Jersáková et al., 2006; Nilsson, 1992).

Dactylorhiza umbrosa is part of the *D. incarnata* complex which includes diploid species related to the European *D. incarnata* (L.) Soó. However, *D. umbrosa* usually grows taller, has larger flowers, and has a long, often curved spur quite different from the short conical spur in *D. incarnata*. Populations approaching *D. incarnata* are found in northern Turkey and in the Caucasus, but the exact relationships between members of the complex is not the focus here.

With a broad definition, *D. umbrosa* has a distribution that extends from the Anatolian plateau in Turkey into the Central Asian plateaus on the northern extensions of the Himalayas (Baumann et al., 2006). The typical habitat of *D. umbrosa*, is open wet fens and grasslands in upland regions at ca. 1000–2000 m asl. These wetlands can extend over large areas and populations sometimes contain thousands of flowering individuals, but small populations can also be found in small fens and along streams. Populations of *D. umbrosa* are often discrete and separated by dry habitats in which the orchid cannot grow.

2.2 | Plant material and sampling area

We obtained material of *D. umbrosa* from the Anatolian-Iranian Plateau connecting eastern Turkey and north-western Iran (Figure 1). The plateau lies between ca. 1500 and 2000 m asl and is mostly tree-less and covered by grasslands grazed by cows, sheep and goats. Fens with *D. umbrosa* are found in shallow depressions or along streams and water coarses. The climate was probably dryer and cooler during the Quaternary (Kehl, 2009; van Zeist & Bottema, 1988), when a steppe-like vegetation dominated by *Artemisia* species covered large parts of the plateau (van Zeist, 1967), although wetlands should still have been present. The plateau is enclosed by mountains ranging from ca. 3000 to 5000 m asl (Sarikaya et al., 2011). During last glacial maximum, the snowline lay between 1900 and 3500 m asl, but at present it lies between 3100 and 4300 m asl (Sarikaya et al., 2011). Suitable habitats for *D. umbrosa* are thus likely to have persisted throughout the Quaternary.

Whereas the sampling area covers a minor portion of the species' total distribution, we still consider the area large enough to include populations that have been isolated from each other for such a long time that they should have approached the maximum genetic divergence found between any pair of populations of the species.

The study included altogether 407 samples from 23 localities (Table 1, Figure 1). Between four and 32 plants (mean 17.7) per locality were sampled, depending on the extent of the population. To avoid sampling of close relatives within sites, we collected material from plants separated by a minimum distance of ca 4 m. The mean distance between localities was 284 km and the maximum distance was 697 km.

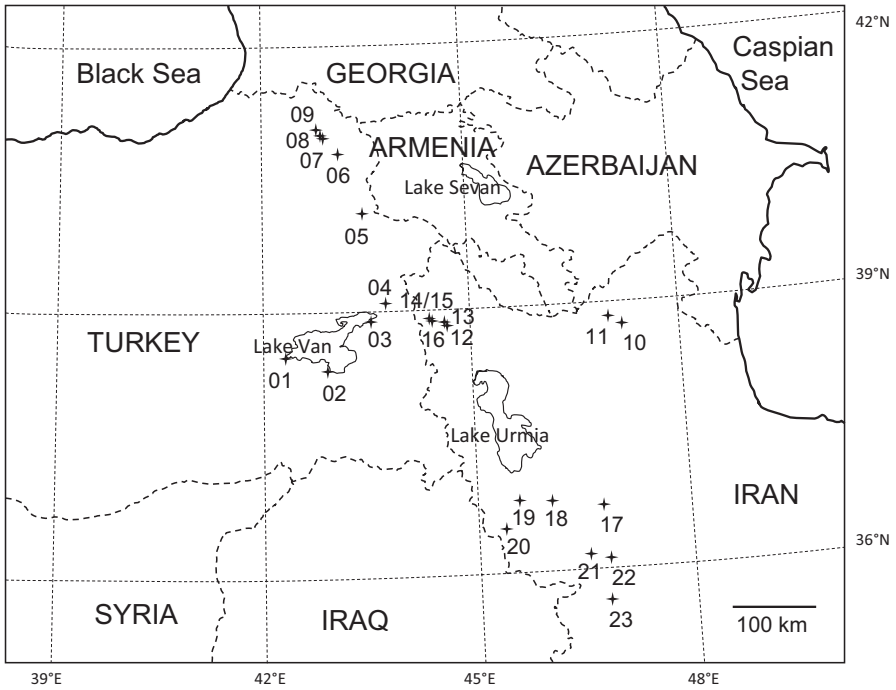

FIGURE 1 Geographic locations of sampling sites for *Dactylorhiza umbrosa*

At five of the 23 localities, we sampled material along transects and measured the relative position of all individuals. If the population was dense enough, we sampled plants at every 4 m and we included individuals up to 50 cm from the transect. This material was used to test for spatial genetic structure within populations, see below.

2.3 | Molecular methods

In the field, a few flowers with bracts were taken from each sampled plant and stored in plastic bags containing dry silica gel (Chase & Hills, 1991). In the laboratory, DNA was extracted from one or two flowers according to a modified 2 × CTAB procedure (Doyle & Doyle, 1990) adapted for extraction in 2 ml Eppendorf tubes.

Five non-coding regions of the plastid genome were amplified by means of specific primers developed for *Dactylorhiza* (Table S1). Loci L1, L10b, L11b and L19 are microsatellite repeats (Hedrén et al., 2008), whereas L3 is a nine base-pair insertion/deletion in the *trn*L intron (Hedrén, 2009). The combined variation patterns at all marker regions were recognized as haplotypes. Unlike in European *D. incarnata*, many samples amplified two fragments of different size when analysed for variation at plastid region L1, a polyA repeat in the *trn*T-*trn*L intergenic spacer (Soliva & Widmer, 1999). Although it would be desirable to map the exact positions of the two duplicated fragments, their variation gave no reason to doubt their location within the plastid genome. Many orchids contain duplications in intergenic spacers and introns within the plastid genome and these can be manifested as microsatellites, minisatellites, or more complex repeats (Bateman, 2001; Cafasso et al., 2001; Hedrén et al., 2008; Soliva & Widmer, 1999). To make use of the observed variation, we introduced a dummy region and sorted the observed fragments into two regions, one containing the shorter fragment and the other containing the longer fragment. This procedure was supported by the variation patterns observed within individual populations. Samples that gave rise to a single peak were treated as having fragments of the same length at both regions.

Variation in the nuclear genome was analysed by examining variation at four microsatellite loci (Table S2). Two of these, ms3 and ms11 are trinucleotide repeats, ms2 a dinucleotide repeat, and D2501 a mononucletide repeat. The first two loci were described in Nordström and Hedrén (2007), whereas the latter two were added in Hedrén et al., (2018).

We refer to the above-mentioned publications and Tables S1 and S2 for details on PCR primers and conditions. The PCR products from each reaction were mixed with 8.5 μl formamide containing appropriate size markers to enable exact size determination of the amplified fragments. The mixes were heated at 94°C for ca. 20 min before loading onto acrylamide gels and the DNA fragments were then separated by size on an ALFexpress II automated sequencer (Amersham Pharmacia Biotech). Fragment sizes were determined by using ALFWIN fragment analyser 1.03.01 software (Amersham Pharmacia Biotech).

2.4 | Data analysis

For each population, the following statistics were calculated for nuclear microsatellite data (Table 2): *A*, average number of alleles per locus; A_e, effective number of alleles per locus; *AR*, allelic richness adjusted for a population size of five samples (10 gene copies); H_e, expected diversity corrected for sample size; and F_{IS}, inbreeding

TABLE 1 Sampled localities of *Dactylorhiza umbrosa*

Pop	Country	Region	Locality	Lat.	Long.	Altitude	*N*
01	Turkey	Bitlis	Tatvan	38°28′N	42°19′E	1720 m	22
02	Turkey	Bitlis	Budakli	38°23′N	42°43′E	1750 m	9
03	Turkey	Van	Muradiye	38°54′N	43°35′E	1680 m	20
04	Turkey	Van	Çaldiran	39°08′N	43°45′E	2020 m	22
05	Turkey	Kars	Digor	40°22′N	43°25′E	1620 m	21
06	Turkey	Kars	Kars, 20 km N rd tw Susuz	40°45′N	43°09′E	1740 m	10
07	Turkey	Kars	Hasköy	40°59′N	42°54′E	1960 m	4
08	Turkey	Kars	Hasköy N	41°00′N	42°53′E	1940 m	21
09	Turkey	Kars	Ardahan-Hanak	41°08′N	42°52′E	1730 m	9
10*	Iran	East Azerbaijan	Ahar to Kaleybar, km 20	38°41′N	47°09′E	1720 m	19
11*	Iran	East Azerbaijan	Hejrandost to Shojaabad	38°51′N	46°58′E	1830 m	25
12*	Iran	West Azerbaijan	Khoy to Chalderan, km 30	38°41′N	44°40′E	2025 m	20
13	Iran	West Azerbaijan	Khoy to Chalderan, km 33	38°43′N	44°38′E	2020 m	25
14	Iran	West Azerbaijan	Khoy to Chalderan, after Zavieh village, Qezeljeh village Dam	38°50′N	44°28′E	2330 m	7
15	Iran	West Azerbaijan	Khoy to Chalderan, after Zavieh village, Qezeljeh village Dam	38°50′N	44°28′E	2330 m	18
16	Iran	West Azerbaijan	Chalderan to Khoy, 1 km before Zorava village	38°48′N	44°34′E	1700 m	32
17*	Iran	West Azerbaijan	Madoab to Bijar, 3.5 km N Menbar village	36°37′N	46°46′E	2110 m	25
18	Iran	West Azerbaijan	Mahabad to Bukan, ca. 3 km NE Darman village	36°45′N	45°54′E	1840 m	25
19	Iran	West Azerbaijan	Mahabad to Piranshahr road, ca 20 km W Mahabad	36°45′N	45°29′E	1830 m	8
20	Iran	West Azerbaijan	Piranshahr to Sardasht road, 5 km before Kubar village	36°33′N	45°13′E	1450 m	21
21	Iran	Kurdistan	Saqqez to Divandareh road, 59 km before Divandareh	36°11′N	46°33′E	1520 m	9
22*	Iran	Kurdistan	Saqqez to Divandareh road, 35 km before Divandareh	36°09′N	46°49′E	2060 m	22
23	Iran	Kurdistan	Divandarreh to Marivan, 1900 m	35°37′N	46°40′E	1970 m	13

Note: Populations used for examination of fine-scale spatial genetic patterns have been indicated by asterisks.

coefficients and their confidence intervals calculated according to Weir and Cockerham (1984).

For plastid haplotype data we calculated: *A*, number of haplotypes per population; A_e, effective number of haplotypes per population; *AR*, haplotype richness adjusted for a population of five samples (five haplotype copies); and H_e, expected diversity corrected for sample size (Table 2).

From five of the populations, we sampled material along transects and measured the relative position of all individuals. This information was used together with plastid haplotype data, and with nuclear microsatellite data, respectively, to estimate spatial genetic structure within each of the sites. Kinship was estimated according to Loiselle et al., (1995). We calculated mean kinship between individuals separated by 4, 8, 16, 32, 64 and 128 m, respectively, and analysed whether kinship decreased with distance within sites. All within-population statistics was calculated in SPAGeDi 1.4 (Hardy & Vekemans, 2002).

Differentiation between pairs of populations was estimated by N_{ST} and G_{ST} for plastid haplotype data, and by R_{ST} and F_{ST} for nuclear microsatellite data. N_{ST} is a differentiation measure that takes genetic distances between haplotypes into account (Pons & Petit, 1996). N_{ST} was calculated on basis of Euclidean distances between pairs of haplotypes, considering numbers of mutational steps at each plastid region separately. G_{ST}, along with 95% confidence intervals, were obtained by 10,000 rounds of permutation of Euclidean distances among haplotypes. R_{ST} is a differentiation measure that takes numbers of mutational steps between microsatellite alleles into account (Michalakis & Excoffier, 1996; Slatkin, 1995). F_{ST}, along with 95% confidence intervals, was estimated by 10,000 rounds of permutation of allele sizes among alleles within loci (see also Hardy

TABLE 2 Summary of nuclear microsatellite and plastid haplotype diversity statistics in *Dactylorhiza umbrosa*

	Nuclear microsatellites						Plastid DNA haplotypes				
Population	***N***	***A***	A_e	***AR***	H_e	F_{IS}	***N***	***A***	A_e	***AR***	H_e
All populations confounded	407	15.50	6.46	4.92	0.748	0.397***	402	96	32.42	4.71	0.969
01	22	6.75	3.24	3.87	0.617	0.364***	22	9	8.59	3.99	0.887
02	9	4.25	3.22	3.58	0.589	0.301**	9	3	2.84	2.54	0.667
03	20	7.00	4.95	4.45	0.705	0.343***	18	6	3.50	2.99	0.719
04	22	7.50	5.62	4.64	0.746	0.390***	21	11	11.10	4.23	0.914
05	21	6.25	4.31	4.14	0.719	0.382***	21	9	10.05	4.12	0.905
06	10	4.50	2.93	3.62	0.651	0.373***	10	6	7.64	3.97	0.889
07	4	4.25	4.75		0.812	0.422**	4	1	1.00		0.000
08	21	5.75	4.05	4.13	0.743	0.174*	21	7	5.15	3.50	0.810
09	9	4.25	4.20		0.724	0.178ns	9	4	4.10	3.21	0.778
10	19	5.25	3.09	3.45	0.566	0.425***	19	4	1.42	1.79	0.298
11	25	5.50	3.74	3.54	0.554	0.321***	25	5	3.81	3.00	0.740
12	20	6.25	4.48	4.27	0.682	0.313***	20	8	5.79	3.64	0.832
13	25	7.75	4.66	4.37	0.683	0.360***	25	7	4.93	3.40	0.800
14	7	5.25	5.67	4.69	0.758	0.455***	7	6	10.76	4.52	0.952
15	18	8.00	5.90	4.86	0.721	0.294***	16	6	5.04	3.43	0.808
16	32	9.00	7.14	5.10	0.786	0.410***	32	16	13.43	4.35	0.927
17	25	8.00	5.08	4.15	0.653	0.355***	25	12	9.12	4.07	0.893
18	25	8.00	5.91	4.16	0.651	0.291***	25	8	5.79	3.62	0.830
19	8	5.25	4.38	4.12	0.647	0.167ns	8	2	1.70	1.89	0.429
20	21	7.50	5.90	4.20	0.599	0.290***	21	8	8.44	3.97	0.886
21	9	4.00	3.00	3.39	0.546	0.069ns	9	2	1.28	1.56	0.222
22	22	5.75	4.89	3.81	0.643	0.230**	22	3	1.21	1.45	0.178
23	13	5.25	4.15	3.56	0.572	0.338***	13	2	1.18	1.38	0.154
Mean	17.7	6.14	4.58	4.10	0.668	0.315	17.5	6.3	5.56	3.21	0.675

Note: For nuclear data: *N*, number of plants sampled; *A*, average number of alleles per locus; A_e, effective number of alleles per locus; *AR*, allelic richness adjusted for a population of five samples (10 gene copies); H_e, expected diversity corrected for sample size; F_{IS}, inbreeding coefficient (departure from Hardy-Weinberg equilibrium: $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$).

For plastid haplotype data: *N*, number of plants sampled; *A*, number of different haplotypes per population; A_e, effective number of haplotypes per population; *AR*, haplotype richness adjusted for a population of 5 samples; H_e, expected diversity corrected for sample size.

& Vekemans, 2002). Calculations were performed in SPAGeDi 1.4 (Hardy & Vekemans, 2002).

The spatial structure of genetic differentiation was investigated by calculating mean pair-wise genetic differentiation between pairs of populations for a series of different geographic distance categories, and by investigating the strength of association between genetic divergence and spatial distance. Pairwise genetic distances were separated into 10 distance categories such that 25 or 26 comparisons were included in each category, which resulted in the distance categories reported in Table 3. Calculations were performed in SPAGeDi 1.4 (Hardy & Vekemans, 2002).

The presence of a phylogeographic structure was tested both by comparing global values for N_{ST} and G_{ST}, and R_{ST} and F_{ST}, for plastid haplotypes and for nuclear microsatellites, respectively, and by comparing the slopes of corresponding pairwise statistics on geographic distances (Table 3; see e.g., Hardy & Vekemans, 2002).

Tests for isolation by distance structure, IBD, were conducted separately for nuclear microsatellites and plastid haplotypes by means of Mantel tests. Pairwise genetic distances were given as pairwise $F_{ST}/(1 - F_{ST})$ for nuclear microsatellites, pairwise $G_{ST}/(1 - G_{ST})$ for plastid haplotypes, and geographic distances as the logarithm of geographic distance (Rousset, 1997). Pairwise genetic distances were calculated in SPAGeDi 1.4 (Hardy & Vekemans, 2002). Associations between distance matrices were tested by 10,000 rounds of permutation in NTSYS-pc 2.2 (Rohlf, 2005).

The distance matrices of pairwise F_{ST} for nuclear microsatellites and pairwise G_{ST} for plastid haplotypes were rotated in principal coordinates analyses (PCO) to extract the maximum of differentiation patterns possible by visual inspection. Pairwise genetic distances were obtained in SPAGeDi 1.4 (Hardy & Vekemans, 2002). PCOs were calculated in NTSYS-pc 2.2 (Rohlf, 2005).

TABLE 3 Genetic differentiation between pairs of populations of *D. umbrosa* at different geographic distances

Distance category	Global estimates	Mean pairwise estimates										Slope ln(dist)
		1	2	3	4	5	6	7	8	9	10	
Max distance (km)		53.0	122.4	203.2	242.1	280.0	300.1	347.5	418.8	530.6	696.6	
Mean distance (km)		25.9	89.4	170.3	226.1	258.2	291.3	320.4	385.7	463.4	598.5	
Estimates based on plastid haplotypes												
N_{ST}	0.504	0.252	0.341	0.320	0.459	0.494	0.542	0.502	0.570	0.595	0.633	0.105
G_{ST} (mean permuted value)	0.310	0.194	0.250	0.188	0.262	0.233	0.334	0.294	0.382	0.379	0.438	0.059
95% confidence interval for G_{ST}, superior value	0.374	0.230	0.293	0.232	0.327	0.309	0.411	0.360	0.466	0.448	0.524	0.076
95% confidence interval for G_{ST}, inferior value	0.252	0.159	0.210	0.149	0.205	0.174	0.271	0.239	0.309	0.318	0.369	0.045
p (one-sided test, H1: obs < exp)	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Estimates based on nuclear microsatellites												
R_{ST}	0.120	0.004	0.025	0.085	0.119	0.112	0.134	0.149	0.119	0.150	0.169	0.044
F_{ST} (mean permuted value)	0.082	0.027	0.028	0.054	0.081	0.072	0.081	0.088	0.118	0.129	0.132	0.029
95% confidence interval for F_{ST}, superior value	0.139	0.064	0.064	0.106	0.143	0.145	0.173	0.165	0.226	0.241	0.291	0.062
95% confidence interval for F_{ST}, inferior value	0.038	−0.004	−0.001	0.013	0.031	0.019	0.026	0.030	0.037	0.045	0.020	0.005
p (1-sided test, H1: obs < exp)	0.917	0.088	0.462	0.901	0.901	0.880	0.908	0.943	0.539	0.680	0.718	0.841
mp/ms	4.076	8.379	12.058	2.971	2.886	2.785	4.892	3.241	3.588	3.022	4.244	

Note: N_{ST} and G_{ST} report on differentiation in the plastid genome. N_{ST} includes information on amount of differentiation between haplotypes. G_{ST} is the mean obtained by 10,000 rounds of permutation of Euclidean distances among haplotypes, and corresponds to a G_{ST} obtained from haplotype frequencies. R_{ST} and F_{ST} report on differentiation in the nuclear genome by analysis of nuclear microsatellites. R_{ST} includes information on amount of differentiation between alleles, assuming a stepwise mutation model. F_{ST} is the mean obtained by 10,000 rounds of permutation of allele sizes among alleles within loci, and corresponds to an F_{ST} obtained from allele frequencies. The pollen to seed flow ratio, *mp/ms*, was calculated on basis of G_{ST}, F_{ST} and F_{IS} estimates obtained by permutation. Pairs of populations were separated into 10 distance categories such that 25–26 pairs were included in each distance category.

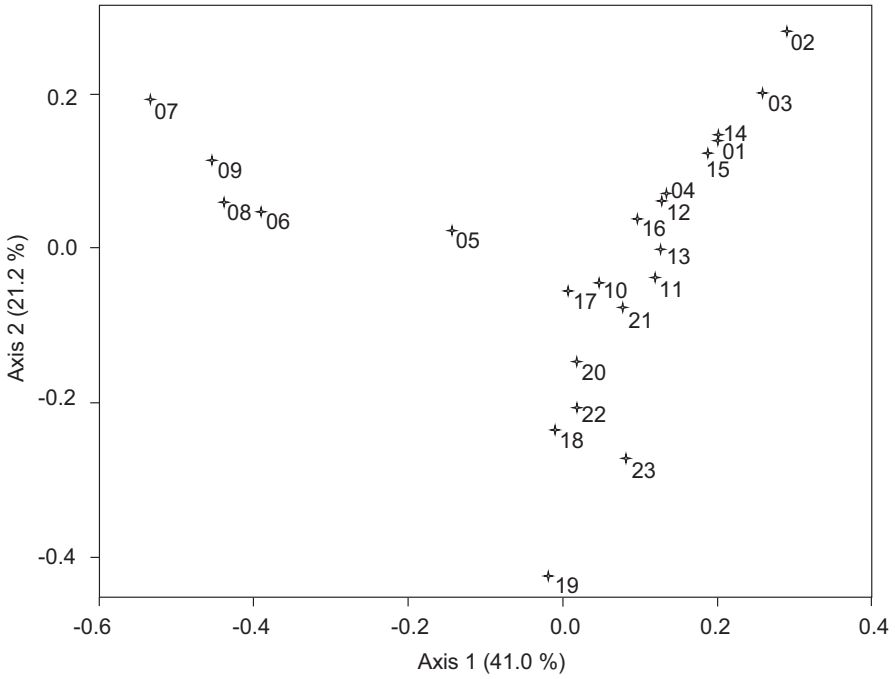

FIGURE 2 Principal coordinates analysis (PCO) summarizing differentiation between analysed populations of *Dactylorhiza umbrosa* on the basis of plastid data. The analysis was based on pairwise G_{ST} values. Population samples have been numbered as in Figure 1 and Table 1

Pairwise F_{ST} and the F_{IS} estimates from nuclear microsatellites and pairwise G_{ST} from plastid haplotypes were used to calculate the pollen to seed-flow ratio, *mp/ms* according to the formula given by Ennos (1994) and Petit et al., (2005):

$$\left\{\left[\left(1/F_{ST(B)}-1\right)\times\left(1+F_{IS(B)}\right)\right]-2\left(1/F_{ST(M)}-1\right)\right\}/\left(1/F_{ST(M)}-1\right),$$

where $F_{ST(B)}$ and $F_{IS(B)}$ are derived from biparentally inherited markers and $F_{ST(M)}$ is derived from maternal plastid markers. Separate calculations were performed for global data and for each of the 10 distance categories obtained in the analyses of spatial genetic structure outlined above.

Commonly adopted differentiation measures including F_{ST} and G_{ST} tend to underestimate levels of differentiation when applied to microsatellites and other marker systems with many alleles per locus (Hedrick, 1999), and a modified differentiation measure, G'_{ST}, has been designed for comparison across marker systems with different degrees of variability (Hedrick, 2005). To describe this effect on general population differentiation, we calculated both the standard G_{ST} according to Nei (1973) and the modified G'_{ST} for the nuclear microsatellite data. Calculations were performed in GenAlEx 6.5 (Peakall & Smouse, 2006, 2012). Still, standard metrics are regarded as better suited to describe population genetic processes in situations where mutation rates are high in comparison to migration rates (Kartzinel et al., 2013). Since our primary interest was to investigate the relative amounts of genetic differentiation in relation to geographic distance, the choice of metric should not impose any major effects on our results. Moreover, the use of standard metrics allow us to compare the outcomes of our analyses, including pollen to seed-flow ratios, with other studies of spatial genetic structure performed in the orchid family.

3 | RESULTS

Variation was found at all plastid regions except for L3, which was invariant in the material analysed. Altogether 96 different haplotypes were identified in the 402 samples that yielded complete data. The number of private haplotypes ranges from zero to 12 within populations. Most of the private haplotypes were confined to one or few samples. Populations 03, 07, 09, and 19 had no private haplotypes, and the highest numbers were found in populations 16 and 17 with 12 and 10 haplotypes, respectively. The nuclear microsatellite loci ms3, ms11, ms2, and msD2501 each yielded 13, six, 29 and 14 alleles, respectively. A total of 62 different alleles were recorded (Table S3). A few were restricted to single populations. Populations 15 and 20 had two private alleles each, whereas populations 01, 02, 04, 07, 08, 17 and 23 had one private allele each (Table S3).

Mean number of alleles per locus ranged from 4.0 to 9.0 within populations, and mean allelic richness calculated with $n = 5$ and excluding populations with less than five individuals varied between 3.39 and 5.10 (Table 2). Expected heterozygosity, H_e, ranged from 0.546 to 0.812, mean 0.668. Inbreeding coefficients were generally quite high (mean average over loci within populations $F_{IS} = 0.315$), and significantly higher than zero in the majority of populations (Table 2). The number of haplotypes per populations ranged from 1 to 16, and haplotype richness calculated with $n = 5$ and excluding populations with less than five individuals ranged from 1.38 to 4.52 (Table 2). Expected heterozygosity, H_e, for plastid haplotypes ranged from 0 (fixed in population 07) to 0.952, with mean 0.675.

We measured the exact relative positions of sampled plants in five of the studied populations to enable analysis of finescale spatial genetic structure. No significant correlation between kinship values

and geographic distance was discovered in any of the populations, neither for nuclear microsatellites, nor for plastid haplotypes.

There was a fairly strong subdivision among populations in the plastid haplotypes and there was also significant general phylogeographic structure in the haplotype data (Table 3; N_{ST} = 0.504 vs. G_{ST} = 0.310, p < .001). Subdivision among populations was lower in nuclear microsatellites and there was no significant general phylogeographic structure (R_{ST} = 0.120 vs. F_{ST} = 0.082, ns). The standardized measure of population differentiation, G'_{ST}, described a clearly higher value of differentiation than the standard G_{ST}, 0.345 and 0.108, respectively.

Plastid haplotype data and nuclear microsatellite data resulted in similar patterns of geographic differentiation when pairwise genetic differentiation between populations were used as input data for PCO ordinations (Figures 2 and 3). On the basis of plastid data, populations from the NW part of the sampling area (06–09) formed a separate cluster linked by pop 05 to the rest of the populations, which themselves formed an elongate cluster within which populations from the central part (01–04, 12–16), the eastern part (10–11, 17–23) and the southern part (17–23) were somewhat separated (Figure 2). Examination of the original data set revealed that samples from the NW part of the distribution usually amplified only a single fragment at region L1, and that no plants with a duplicated fragment were found in populations 06, 07, 08 and 09. In the PCO based on nuclear microsatellite data, populations from the northwestern part diverged towards the upper left, populations from the eastern and southern parts diverged towards the upper right and populations from the west-central part linked these groups together (Figure 3).

A significant isolation by distance structure was observed when pairwise G_{ST} derived from haplotype data was compared to geographic distances (r = .152, p [random $z \geq$ observed z] = .0006, t = 1.8808). Comparing population pairs in different distance categories (Table 3, Figure 4), we found that populations up to 53 km apart had an average N_{ST} = 0.252, whereas populations up to 696 km apart, had an average N_{ST} = 0.633. N_{ST} was rising more steeply than G_{ST} with increasing geographic distance for the plastid data (slope ln distance = 0.105 vs. 0.059, p < .0000), demonstrating a significant phylogeographic structure of plastid genomes.

An even stronger isolation by distance structure was evident in the nuclear microsatellite data when pairwise F_{ST} was compared to geographic distances (r = .591, p [random $z \geq$ observed z] = .0001, t = 8.0476). Populations up to 53 km apart were minimally differentiated from each other (average pairwise F_{ST} = 0.004; Table 3, Figure 4), whereas average differentiation in the category comprising the most distant pairs of populations was still moderate (F_{ST} = 0.169). However, the slope of R_{ST} to geographic distance did not differ from that of F_{ST} for nuclear microsatellites (slope ln (dist) = 0.044 vs. 0.029, p = .841), revealing no significant phylogeographic structure of nuclear microsatellite data.

The correlation between pairwise G_{ST} derived from haplotype data and pairwise F_{ST} derived from nuclear data was weak and nonsignificant (r = .177, p [random $z \geq$ observed z] = .071, t = 1.5448).

Pollen to seed flow ratios, *mp/ms*, were calculated for different distance categories on the basis of the corresponding G_{ST} from plastid data and F_{ST} from nuclear microsatellite data (Table 3, Figure 4). The ratio *mp/ms* was 8.4 and 12.8 for the distance categories of 0–26 and 26–89 km, respectively, whereas *mp/ms* ratios were lower and fluctuated between 2.8 and 4.9 for higher distance categories. The high ratio calculated for the second distance category (*mp/ms* = 12.8), was explained by strong divergence between population 23 and populations 21 and 22 in the plastid genome, which gave rise to a high average G_{ST} for this distance category. The overall pollen to seed flow ratio was 4.08 (Table 3).

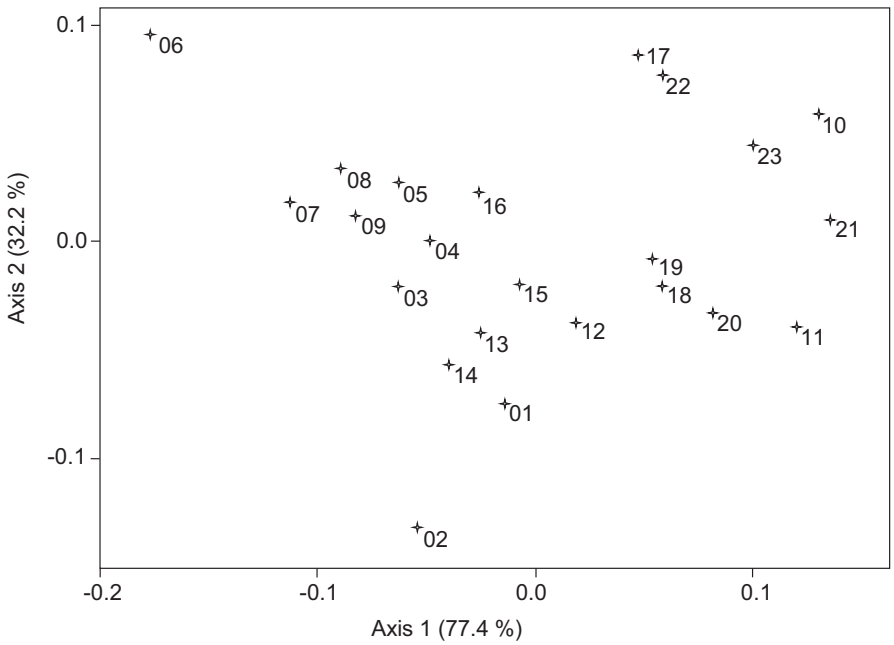

FIGURE 3 Principal coordinates analysis (PCO) summarizing differentation between analysed populations of *Dactylorhiza umbrosa* in nuclear microsatellites. The analysis was based on pairwise F_{ST} values. Population samples have been numbered as in Figure 1 and Table 1

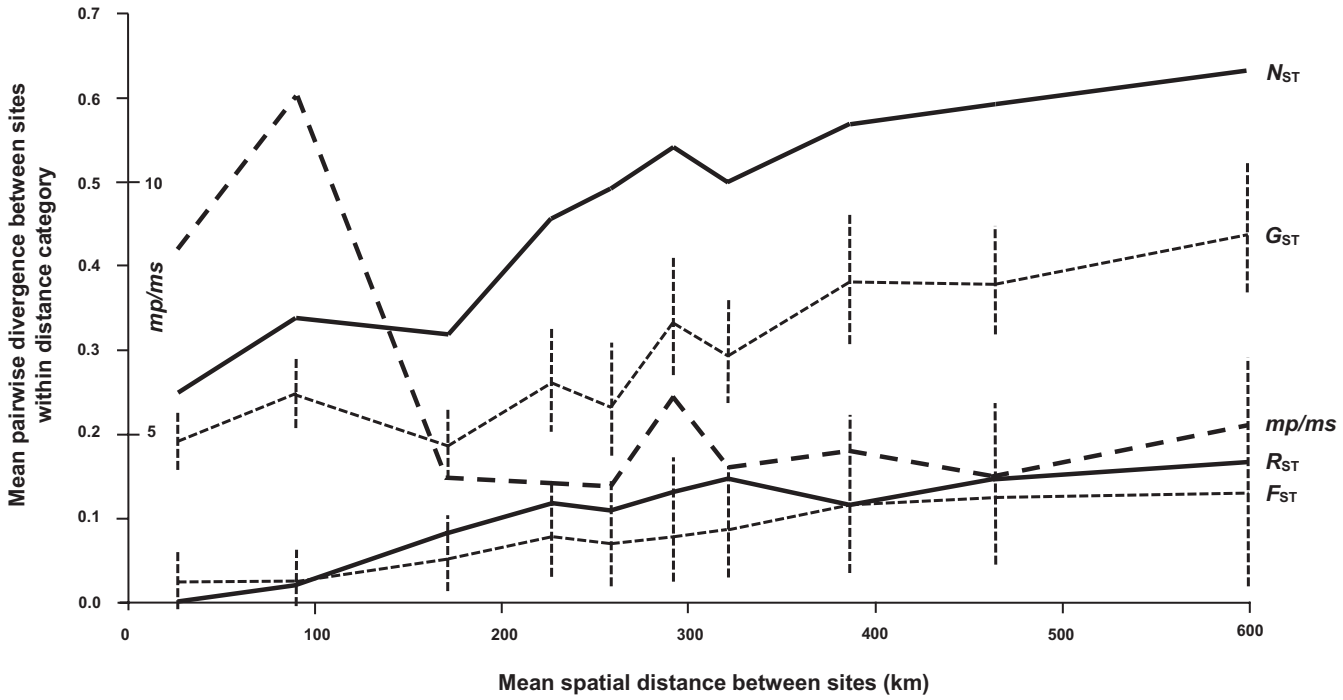

FIGURE 4 Correlogram reporting the relationship between spatial distance and various measures of population differentiation in *Dactylorhiza umbrosa*. N_{ST} and G_{ST} report on differentiation in the plastid genome. N_{ST} includes information on amount of differentiation between haplotypes, whereas G_{ST} only reflects differences in haplotype frequencies. 95% confidence intervals for G_{ST} were obtained by 10,000 rounds of permutation of Euclidean distances among haplotypes. R_{ST} and F_{ST} report on differentiation in the nuclear genome by analysis of nuclear microsatellites. R_{ST} includes information on amount of differentiation between alleles, assuming a stepwise mutation model. F_{ST} only reflects differences in allele frequencies. 95% confidence intervals for F_{ST} were obtained by 10,000 rounds of permutation of allele sizes among alleles within loci. The pollen to seed flow ratio, *mp/ms*, was calculated on basis of G_{ST}, F_{ST} and F_{IS} estimates obtained by permutation. Pairs of populations were separated into 10 distance categories such that 25–26 pairs were included in each distance category

4 | DISCUSSION

4.1 | Isolation by distance structure

We hypothesized that *D. umbrosa* might have been present within its current distribution area for a sufficiently long period to have developed a mature metapopulation structure, which would now be manifested as an isolation by distance structure (IBD) in neutral genetic marker systems. The structuring of nuclear microsatellite data was in agreement with this prediction. We found that pairwise F_{ST} estimates derived from these data were increasing with geographic distance between populations and that the association was significant ($r = .591$, $p < .001$).

Our result differs from that of two related marsh orchids studied in northern Europe, in which the correlations between F_{ST} and geographic distance was not as clear. In *D. majalis* ssp. *lapponica* on the island of Gotland (Hedrén et al., 2018), still a weak IBD ($r = .269$, $p = .0465$) was observed, but in *D. majalis* ssp. *majalis* studied over two geographic scales of 178 and 985 km, respectively (Hedrén & Nordström Olofsson, 2018), there was no significant IBD. At least two factors may explain the difference between *D. umbrosa* and the northern European members of *Dactylorhiza*.

First, Phillips et al., (2012), who summarized IBD statistics in terrestrial orchids studied by allozyme markers found that the significance of an IBD structure increased with the range of the sampling area. In studies covering less than 250 km only three out of seven studies showed a significant IBD, whereas in studies with wider sampling nine out of eleven studies showed a significant IBD. Here, we sampled populations of *D. umbrosa* at a maximum distance of 696 km, so the sampled range appears to be wide enough to disclose the presence of a significant IBD.

A second factor could be the age of the population system under study. Comparison of the genetic structure in *D. umbrosa* with that of *Dactylorhiza* species distributed in northern Europe suggests that it takes a long time for a metapopulation system to stabilize, such that local extinctions, recolonization and gene flow by pollen and seeds reach a state of equilibrium. Habitats in northern Europe were ice-covered or otherwise uninhabitable for orchids during the last ice age, which peaked at ca. 21 ka BP (Wohlfarth et al., 2008). Most of the genetic patterns we have observed in northern European *Dactylorhiza* seem to be linked to recolonization history, including immigration from multiple directions (Nordström & Hedrén, 2008; Ståhlberg & Hedrén, 2009), or marginal effects with lower amounts of diversity stored in populations of recent origins (Hedrén, 2009; Hedrén et al., 2018), but there is little structure in genetic differentiation patterns at the regional scale.

A long life-span may also contribute to a slow development of IBD in *Dactylorhiza* and related genera of terrestrial orchids. Although

the plants are herbaceous, the flowering stems arise from tubers produced during the previous growth season, and individual plants can become quite old. The exact generation time of *D. umbrosa* is not known, but in the more delicate congener *D. majalis* ssp. *lapponica*, examined at one site in central Norway, the half-life of flowering plants was estimated to 5.8 years (Øien & Moen, 2002). Adding a few years for development from seed to first flowering, individual plants observed in the field may on average be around 10 years. Furthermore, populations are often large and gene frequencies only change slowly following gene dispersal between sites. Changes in genetic patterns may also be retarded due to geitonogamy, but on the other hand, the deceit pollination system should promote gene dispersal of pollen over longer distances than in rewarding orchids (Jersáková et al., 2006).

Altogether, the clear geographic structuring of molecular data found in *D. umbrosa* seems to agree with the hypothesis that *D. umbrosa* has a long history within the sampled distribution area, probably stretching back much longer than the last ice age. Although individual populations of *D. umbrosa* may be as young as populations of other *Dactylorhiza* species in previously glaciated areas, vacant sites may still be colonized by seeds from neighbouring sites, and the regional populations should in any case be older than regional populations of *Dactylorhiza* in northern Europe. Given a long enough time, seed and pollen dispersal between populations may also have contributed to a stronger similarity between neighbouring populations than between distant populations, such that an isolation by distance structure could have developed.

We also observed a low level of differentiation between population of *D. umbrosa* (F_{ST} = 0.082). This value is in the lower range of overall estimates given for the orchid family in compilations by Hamrick and Godt (1996), Forrest et al., (2004) and Phillips et al., (2012), 0.087, 0.187, and 0.146, respectively, but still in agreement with orchids characterized by food deception. On basis of data compiled by Forrest et al., (2004), Cozzolino and Widmer (2005) calculated mean G_{ST} in deceptive orchids to be only half of that in rewarding orchids (G_{ST} = ca. 0.125 vs. G_{ST} = ca. 0.25). Other estimates of population differentiation in *Dactylorhiza* species range from G_{ST} = 0.0183 in *D. foliosa* to G_{ST} = 0.309 in *D. incarnata* (Table S4).

4.2 | Plastid differentiation and pollen to seed flow ratio

As expected, differentiation between populations of *D. umbrosa* as revealed by plastid markers was considerably higher than that revealed by nuclear markers (G_{ST} = 0.310 vs. F_{ST} = 0.082), but still lower than in many other orchids. Estimates obtained from other species of *Dactylorhiza* range from 0.137 in *D. foliosa* (Lowe) Soó to 0.707 in *D. majalis* (Rchb.) P.F. Hunt & Summerh. ssp. *majalis* (Table S4). High G_{ST} values for the plastid genome in orchids were also reported by Petit et al., (2005), ranging from 0.27 to 0.87 (mean 0.59). These values are similar to other plants with maternal inheritance of the plastid genome (mean G_{ST} = 0.67; Petit et al., 2005).

Combining the overall F_{IS} value and the F_{ST} for the nuclear genome with G_{ST} for the plastid genome, we calculated an overall pollen to seed dispersal ratio of *mp/ms* = 4.076 in *D. umbrosa*. This value is low compared to seed plants in general, in which a median value for the *mp/ms* ratio was estimated to ca 17 (Petit et al., 2005). However, *D. umbrosa* should also be compared to other orchids for which pollen to seed dispersal ratios are available (Table S4). In Figure 5 we have plotted corresponding estimates of population differentiation in the nuclear and plastid genomes against each other and also given the functions for *mp/ms* = 1, and *mp/ms* = 10. The ratio observed in *D. umbrosa* was slightly higher than that observed in some other temperate terrestrial orchids, but well within the range of ratios calculated for orchids growing in tropical and subtropical areas. Most temperate, terrestrial species express high levels of population differentiation in both the nuclear and the plastid genomes and have low pollen to seed dispersal ratios. *mp/ms* ratios calculated in other species of *Dactylorhiza* range from 0.96 in *D. majalis* ssp. *lapponica* (Hartm.) H. Sund. to 7.30 in the Madeiran *D. foliosa* (Table S4). Tropical epiphytic orchids express lower levels of differentiation in both the nuclear and the plastid genomes than temperate species, but higher pollen to seed dispersal ratios, often higher than 10. Also the subtropical terrestrial species *Epidendrum fulgens* Brongn. from southern Brazil expressed a high pollen to seed dispersal ratio, *mp/ms* = 14.25 (Pinheiro et al., 2011).

Whereas tropical epiphytes on average only express slightly lower levels of differentiation between populations than do temperate terrestrial orchids in the plastid genome, the main reason for their relatively high seed to pollen dispersal ratios is apparently the very low levels of differentiation in the nuclear genome (Table S4, Figure 5). In agreement with this observation, Phillips et al., (2012) found a lower level of population differentiation in epiphytic orchids (F_{ST} = 0.109) than in terrestrial orchids (F_{ST} = 0.161), although the difference was not statistically significant. Tropical orchids often grow in dispersed populations, but have elaborate pollination systems and some species have been shown to be strongly adapted to specific pollinators that are capable of transporting pollen over long distances (e.g., Janzen, 1971; Kroodsma, 1975). The three epiphytes included in this study for comparison are all adapted to be visited by large, strong-flying pollinators, *Laelia rubescens* Lindl. by hummingbirds (Trapnell & Hamrick, 2004), *Epidendrum firmum* Rchb.f. by strong-flying moths (Kartzinel et al., 2013) and *Brassavola nodosa* (L.) Lindl. by hawk-moths (Trapnell et al., 2013, 2019), apparently leading to efficient pollen dispersal and low levels of population differentiation in the nuclear genome. However, the efficiency of seed dispersal may not differ much between orchids in tropical and temperate areas, so a comparatively high pollen to seed flow ratio in tropical orchids is not surprising. On the other hand, average estimates for pollen to seed flow ratio might become lower when more data become available of especially tropical orchids pollinated by less strong-flying insects such as biting midges (Bogarin et al., 2018) and fruit flies (Karremans et al., 2015).

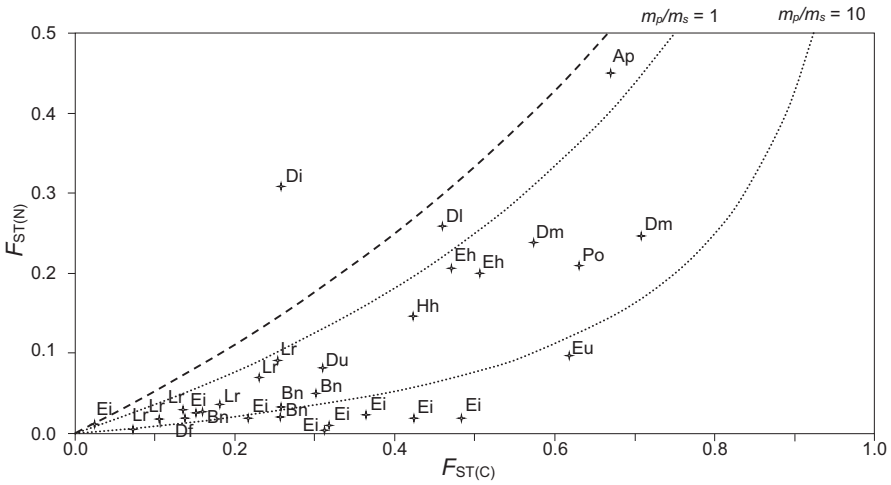

FIGURE 5 Plot of F_{ST} (or corresponding statistics describing differentiation between populations) derived from plastid markers against F_{ST} derived from nuclear markers in *D. umbrosa* and other orchids studied for differentiation in both genomes (Table S4). Some orchids have been reported for different geographic scales and are given as multiple entries. The dotted lines give *mp*/*ms* = 1 and *mp*/*ms* = 10, respectively. The broken line is the maximum possible theoretical value for nuclear F_{ST} given the corresponding F_{ST} of a maternally inherited plastid genome, assuming an island model at drift-migration equilibrium (see Petit et al., 2005). Abbreviations: Ap, *Anacamptis palustris*; Bn, *Brassavola nodosa*; Df, *Dactylorhiza maculata* ssp. *fuchsii*; Di, *D. incarnata*; Dl, *D. majalis* ssp. *lapponica*; Dm, *D. majalis* ssp. *majalis*; Du, *D. umbrosa*; Eh, *Epipactis helleborine*; Ei, *Epidendrum firmum*; Eu, *E. fulgens*; Hh, *Himantoglossum hircinum*; Lr, *Laelia rubescens*; Po, *Pogonia ophioglossoides*

4.3 | Pollen to seed flow ratio versus geographic distance

As expected from general observations of pollen and seed dispersal patterns (Hamrick & Trapnell, 2011; McCauley, 1997; Petit et al., 2005), we found that the pollen to seed dispersal ratio, *mp*/*ms*, in *D. umbrosa* decrease with geographic distance. A few other orchid species have also been analysed at different geographic scales, but in these cases, the correspondence between *mp*/*ms* and geographic distance has gone in different directions (Table S4). A weak indication of decreasing *mp*/*ms* ratios with distance was observed in epiphytic *Laelia rubescens* (Trapnell & Hamrick, 2004), but in *Brassavola nodosa* (Trapnell et al., 2013, 2019) and *Epidendrum firmum* (Kartzinel et al., 2013), even an opposite pattern was observed. In the present study, all samples were examined for plastid microsatellites with high variability, and we were able to describe geographic variation in the plastid genome to a fair amount of detail. Other studies have often used less variable markers than we did for analysis of the plastid genome, and have often analysed lower numbers of samples. As result, estimates of F_{ST}/G_{ST} from the plastid genome can be imprecise, and hence also pollen to seed flow ratios.

The finding that the pollen to seed flow ratio is high between populations at short distance, but drops at higher distances is compatible with the hypothesis that seed dispersal in orchids is highly leptokurtic (Hamrick & Trapnell, 2011). Most seeds fall near their source within populations, whereas pollen is on average dispersed more efficiently at medium range within populations, and between close populations. A small fraction of seeds can still be dispersed very far, which is obviously necessary for range expansion and colonization of vacant sites.

4.4 | Colonization history and metapopulation structure

During an early colonization phase, when a new population settles far from already existing populations, the seeds founding the new population may arrive from almost any distance and source. As long as the new population is small, other vacant sites in the neighbourhood are also likely to be colonized by seeds from long distance, and there is a distinct possibility that different populations are founded by seeds from different source areas (Hedrén & Nordström Olofsson, 2018; Hedrén et al., 2018). This pattern would initially result in high level of differentiation between populations. However, once multiple populations in a region have grown large in numbers of seed-producing individuals, further populations are likely to be founded by seeds from the same region (Trapnell et al., 2013). Given that local populations are connected in regional metapopulation systems with local extinctions and recolonizations (Wade & McCauley, 1988), neighbouring populations should successively share more of the regional haplotypes, and population differentiation should slowly decrease over time. In parallel, genetic divergence from other regions will increase due to mutation, selection and random processes, until an isolation by distance structure, IBD, has developed.

In this study, we nevertheless observed that populations at close distance were clearly differentiated in plastid haplotypes (N_{ST} = 0.252, G_{ST} = 0.194, Figure 4, Table 3). Members of *Dactylorhiza* are self-fertile (Neiland & Wilcock, 1995), and new populations can be established from single seeds. If the first plant starts to produce seed before any other seeds have arrived to the site, the first plant may fill the site with its own offspring (Trapnell et al., 2013). All plants would carry the same plastid haplotype, and seeds arriving

later should only have marginal impact on haplotype frequencies. The often relatively low number of recorded plastid haplotypes within populations, as compared to the total number of haplotypes (Table 2) is in agreement with such a process. Even if a new population is established from a seed with a haplotype that is present also in other populations in the same region, haplotype frequencies may still remain different, resulting in high estimates of pairwise N_{ST} or G_{ST} also between close sites (Trapnell et al., 2013).

Seed dispersal is necessary for founding new populations, but pollen dispersal seems to be more efficient for gene flow between already established populations of orchids (Helsen et al., 2015; Jacquemyn et al., 2009). Given a long enough time, allele frequencies at nuclear microsatellite loci should even out more efficiently between populations than plastid haplotype frequencies (Trapnell et al., 2013). As a consequence of pollen dispersal between neighbouring populations, nuclear alleles will also spread more efficiently within the species as a whole (McCauley, 1994). These processes explain why we found a stronger IBD structure in the nuclear genome than in the plastid genome, and also why there was no significant phylogeographic signal in the nuclear data, in contrast to the plastid data.

The dispersal and colonization processes in *D. umbrosa* seem to agree with those described for tropical epiphytes, such as *Brassavola nodosa* (Trapnell et al., 2013). In epiphytes, new populations (in separate trees) are typically founded by one or few seeds from surrounding populations. Gene flow between established populations proceeds by pollen dispersal, and if seeds arrive from long distance they should have negligible influence on already established metapopulations. However, whereas many tropical epiphytes form metapopulations with high turnover (Trapnell et al., 2013; Winkler et al., 2009), metapopulation processes in the terrestrial *D. umbrosa* are apparently much slower.

Whereas studies of northern European orchids have revealed only restricted geographic structuring of genetic variation in many cases, we found here a clear geographic structuring in *D. umbrosa*, expressed as an isolation by distance structure in the nuclear genome, a phylogeographic structure in the plastid genome, and a pattern of decreasing pollen to seed dispersal ratios at progressively larger geographic distances. We suggest that this discrepancy is a result of the much longer time period available for developing a genetic structure in *D. umbrosa*, as compared to its European counterparts. Although orchids are exceptional in having minute, dust-like seeds that are seemingly easily dispersed by wind and might contribute to efficient gene dispersal over long distances, we conclude that such a gene dispersal is still not efficient enough to overcome factors that promote genetic divergence in a longer time perspective. In these respects, orchids are no different from other plant species in which seed dispersal is more restricted.

ACKNOWLEDGEMENTS

We thank four anonymous reviewers and the editors of *Molecular Ecology* for constructive criticism and helpful comments to improve previous versions of this manuscript. The study was supported by grants from the Elly Olsson Foundation and the Crafoord Foundation to MH. Fieldwork in Iran was supported by Swedish Science Council - Swedish Research Links grant D0664801 to HB.

AUTHOR CONTRIBUTIONS

Sven Birkedal, Hugo de Boer, Abdolbaset Ghorbani, Barbara Gravendeel, Sven Hansson, Åke Svensson and Shahin Zarre organized the field work and took part in sampling. Hugo de Boer, Abdolbaset Ghorbani, Barbara Gravendeel and Shahin Zarre arranged permits. Mikael Hedrén extracted DNA, performed laboratory analyses, analysed the data and wrote the draft manuscript. All authors contributed to the manuscript.

DATA AVAILABILITY STATEMENT

The full data set including plastid microsatellite data, plastid haplotypes, and nuclear microsatellite genotypes for all analysed individuals has been deposited at Dryad, https://doi.org/10.5061/dryad.9p8cz8wdf.

ORCID

Mikael Hedrén https://orcid.org/0000-0001-8194-4617

REFERENCES

Arditti, J., & Ghani, A. K. A. (2000). Numerical and physical properties of orchid seeds and their biological implications. *New Phytologist*, *145*, 367–421.

Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H., & Godelle, B. (2000). Effects of colonization processes on genetic diversity, differences between annual plants and tree species. *Genetics*, *154*, 1309–1321.

Bateman, R. M. (2001). Evolution and classification of European orchids: Insights from molecular and morphological characters. *Journal Europäischer Orchideen*, *2001*, 33–119.

Baumann, H., Künkele, S., & Lorenz, R. (2006). *Orchideen Europas*. Eugen Ulmer.

Bogarin, D., Fernandez, M., Borkent, A., Heemskerk, A., Pupulin, F., Ramirez, S., Smets, E., & Gravendeel, B. (2018). Pollination of *Trichosalpinx* (Orchidaceae; Pleutothallidinae) by biting midges (Diptera: Ceratopogonidae). *Botanical Journal of the Linnean Society*, *186*, 510–543.

Cafasso, D., Pellegrino, G., Musacchio, A., Widmer, A., & Cozzolino, S. (2001). Characterization of a minisatellite repeat locus in the chloroplast genome of *Orchis palustris* (Orchidaceae). *Current Genetics*, *39*, 394–398. https://doi.org/10.1007/s002940100226.

Cain, M. L., Milligan, B. G., & Strand, A. E. (2000). Long-distance seed dispersal in plant populations. *American Journal of Botany*, *87*(9), 1217–1227.

Chase, M. W., & Hills, H. G. (1991). Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon*, *40*, 215–220.

Corriveau, J. L., & Coleman, A. W. (1988). Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American Journal of Botany*, *75*, 1443–1458.

Cozzolino, S., Cafasso, D., Pellegrino, G., Musacchio, A., & Widmer, A. (2003). Fine-scale phylogeographical analysis of Mediterranean *Anacamptis palustris* (Orchidaceae) populations based on chloroplast minisatellite and microsatellite variation. *Molecular Ecology*, *12*, 2783–2792.

Cozzolino, S., & Widmer, A. (2005). Orchid diversity: An evolutionary consequence of deception? *Trends in Ecology and Evolution*, *20*, 487–494.

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*, 13–15.

Ennos, R. A. (1994). Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, *72*, 250–259.

Forrest, A. D., Hollingsworth, M. L., Hollingsworth, P. M., Sydes, C., & Bateman, R. M. (2004). Population genetic structure in European populations of *Spiranthes romanzoffiana* set in the context of other genetic studies on orchids. *Heredity*, *92*, 218–227.

Hamrick, J. L., & Godt, M. J. W. (1990). Allozyme diversity in plant species. In A. D. Brown, M. T. Clegg, A. H. Kahler, & B. S. Weir (Eds.), *Plant population genetics, breeding and genetic resources* (pp. 43–63). Sinauer.

Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *351*, 1291–1298.

Hamrick, J. L., & Trapnell, D. W. (2011). Using population genetic analyses to understand seed dispersal patterns. *Acta Oecologica*, *37*, 641–649.

Hardy, O. J., & Vekemans, X. (2002). SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*, 618–620.

Hedrén, M. (2009). Plastid DNA haplotype variation in Scandinavian *Dactylorhiza incarnata* (Orchidaceae): Evidence for multiple independent colonization events. *Nordic Journal of Botany*, *27*, 69–80.

Hedrén, M., & Nordström Olofsson, S. (2018). High levels of genetic diversity in marginal populations of the marsh orchid *Dactylorhiza majalis* ssp. *majalis*. *Nordic Journal of Botany*, *2018*, e01747. https://doi.org/10.1111/njb.01747

Hedrén, M., Nordström Olofsson, S., & Paun, O. (2018). Orchid colonization: Multiple parallel dispersal events and mosaic genetic structure in *Dactylorhiza majalis* ssp. *lapponica* on the Baltic island of Gotland. *Annals of Botany*, *122*, 1019–1031. https://doi.org/10.1093/aob/mcy111.

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

Hedrick, P. W. (1999). Highly variable loci and their interpretation in evolution and conservation. *Evolution*, *53*, 313–318.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, *59*, 1633–1638.

Helsen, K., Meekers, T., Vranckx, G., Roldán-Ruiz, I., Vandepitte, K., & Honnay, O. (2015). A direct assessment of realized seed and pollen flow within and between two isolated populations of the food-deceptive orchid *Orchis mascula*. *Plant Biology*, *18*, 139–146.

Jacquemyn, H., Wiegand, T., Vandepitte, K., Brys, R., Roldán-Ruiz, I., & Honnay, O. (2009). Multigenerational analysis of spatial structure in the terrestrial, food-deceptive orchid *Orchis mascula*. *Journal of Ecology*, *97*, 206–216.

Janzen, D. H. (1971). Bees as long-distance pollinators of tropical plants. *Science*, *171*, 203–205.

Jersáková, J., Johnson, S. D., & Kindlmann, P. (2006). Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews*, *81*, 219–236.

Karremans, A. P., Pupulin, F., Grimaldi, D., Beentjes, K. K., Butot, R., Fazzi, G. E., Kaspers, K., Kruizinga, J., Roessingh, P., Smets, E. F., & Gravendeel, B. (2015). Pollination of *Specklinia* by nectar feeding *Drosophila*: First reported case of a deceptive syndrome employing aggregation pheromones in Orchidaceae. *Annals of Botany*, *116*, 437–455.

Kartzinel, T. R., Shefferson, R. P., & Trapnell, D. W. (2013). Relative importance of pollen and seed dispersal across a Neotropical mountain landscape for an epiphytic orchid. *Molecular Ecology*, *22*, 6048–6059.

Kehl, M. (2009). Quaternary climate change in Iran – The state of knowledge. *Erdkunde*, *63*, 1–17.

Kroodsma, D. E. (1975). Distances of male Euglossinae bees in orchid pollination. *Biotropica*, *7*, 71–72.

Loiselle, B. A., Sork, V. L., Nason, J. D., & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, *82*, 1420–1425.

McCauley, D. E. (1994). Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of in *Silene alba*: Implications for studies of gene flow in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 8127–8131.

McCauley, D. E. (1995). The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution*, *10*, 198–202.

McCauley, D. E. (1997). The relative contributions of seed and pollen movement to the local genetic structure of *Silene alba*. *Journal of Heredity*, *1997*, 257–263.

Michalakis, Y., & Excoffier, L. (1996). A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, *142*, 1061–1064.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*, 3321–3323.

Neiland, M. R. M., & Wilcock, C. C. (1995). Maximisation of reproductive success by European Orchidaceae under conditions of infrequent pollination. *Protoplasma*, *187*, 39–48.

Nilsson, L. A. (1981). Pollination ecology and evolutionary processes in six species of orchids. *Acta Universitatis Upsaliensis*, *593*, 1–40.

Nilsson, L. A. (1992). Orchid pollination biology. *Trends in Ecology and Evolution*, *7*, 225–259.

Nordström, S., & Hedrén, M. (2007). Development of polymorphic nuclear microsatellite markers for polyploid and diploid members of the orchid genus *Dactylorhiza*. *Molecular Ecology Notes*, *7*, 644–647.

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

Øien, D.-I., & Moen, A. (2002). Flowering and survival of *Dactylorhiza lapponica* and *Gymnadenia conopsea* in the Sølendet Nature Reserve, Central Norway. In P. Kindlmann, J. H. Willems, & D. F. Whigham (Eds.), *Trends and fluctuations and underlying mechanisms in terrestrial orchid populations* (pp. 3–22). Backhuys Publishers.

Partomihardjo, T. (2003). Colonisation of orchids on the Krakatau Islands. *Telopea*, *10*, 299–310.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – An update. *Bioinformatics*, *28*, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460.

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Vendramin, G. G. (2005). Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, *14*, 689–701.

Pfeifer, M., Schatz, B., Pico, F. X., Passalacqua, N. G., Fay, M. F., Carey, P. D., & Jeltsch, F. (2009). Phylogeography and genetic structure of the orchid *Himantoglossum hircinum* (L.) Spreng. across its European central–marginal gradient. *Journal of Biogeography*, *36*, 2353–2365.

Phillips, R., Dixon, K., & Peakall, R. (2012). Low population genetic differentiation in the Orchidaceae: Implications for the diversification of the family. *Molecular Ecology*, *21*, 5208–5220.

Pinheiro, F., de Barros, F., Palma-Silva, C., Fay, M. F., Lexer, C., & Cozzolino, S. (2011). Phylogeography and the genetic differentiation along the distributional range of the orchid *Epidendrum fulgens*: A Neotropical coastal species not restricted to glacial refugia. *Journal of Biogeography*, *38*, 1923–1935. https://doi.org/10.1111/j.1365-2699.201102539x

Pons, O., & Petit, R. J. (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, *144*, 1237–1245.

Pridgeon, A. M., Cribb, P. J., & Chase, M. W. (1999). *Genera Orchidacearum* (Vol. *1*). Oxford University Press.

Pridgeon, A. M., Cribb, P. J., Chase, M. W., & Rasmussen (2001). *Genera Orchidacearum* (Vol. *2*, part 1). Oxford University Press.

Rohlf, F. J. (2005). *NTSYSpc: Numerical taxonomy system*, *ver. 2.20*. Exeter Publishing.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, *145*, 1219–1228.

Sarikaya, M. A., Çiner, A., & Zreda, M. (2011). Quaternary glaciations of Turkey. In J. Ehlers, P. L. Gibbard, & P. D. Hughes (Eds.), *Quaternary glaciations – Extent and chronology. Developments in quaternary science* (Vol. *15*, pp. 393–403). : Elsevier.

Slatkin, M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology*, *12*, 253–262.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, *139*, 457–462.

Soliva, M., & Widmer, A. (1999). Genetic and floral divergence among sympatric populations of *Gymnadenia conopsea* s.l. (Orchidaceae) with different flowering phenology. *International Journal of Plant Science*, *160*, 897–905.

Squirrel, J., Hollingsworth, P. M., Bateman, R. M., Dickson, J. H., Light, M. H. S., MacConaill, M., & Tebbitt, M. C. (2001). Partitioning and diversity of nuclear and organelle markers in native and introduced populations of *Epipactis helleborine* (Orchidaceae). *American Journal of Botany*, *88*, 1409–1418.

Ståhlberg, D., & Hedrén, M. (2009). Systematics and phylogeography of the *Dactylorhiza maculata* complex (Orchidaceae) in Scandinavia: Insights from cytological, morphological and molecular data. *Plant Systematics and Evolution*, *273*, 107–132.

Trapnell, D. W., & Hamrick, J. L. (2004). Partitioning nuclear and chloroplast variation at multiple spatial scales in the neotropical epiphytic orchid, *Laelia rubescens*. *Molecular Ecology*, *13*, 2655–2666.

Trapnell, D. W., Hamrick, J. L., Ishibashi, C. D., & Kartzinel, T. R. (2013). Genetic inference of epiphytic orchid colonization; it may only take one. *Molecular Ecology*, *22*, 3680–3692.

Trapnell, D. W., Hamrick, J. L., Smallwood, P. A., Kartzinel, T. R., Ishibashi, C. D., & Quigley, T. C. (2019). Phylogeography of the neotropical epiphytic orchid, *Brassavola nodosa*: Evidence for a secondary contact zone in northwestern Costa Rica. *Heredity*, *123*, 662–674. https://doi.org/10.1038/s41437-019-0218-y.

van Zeist, W. (1967). Late Quaternary vegetation history of western Iran. *Review of Palaeobotany and Palynology*, *2*, 301–311.

van Zeist, W., & Bottema, S. (1988). Late Quaternary vegetational and climatic history of southwest Asia. *Proceedings of the Indian National Science Academy*, *54*(3), 461–480.

VandenBroeck, A., Van Landuyt, W., Cox, K., De Bruyn, L., Gyselings, R., Oostermeijer, G., Valentin, B., Bozic, G., Dolinar, B., Illyés, Z., & Mergeay, J. (2014). High levels of effective long-distance dispersal may blur ecotypic divergence in a rare terrestrial orchid. *BMC Ecology*, *14*, 20.

Wade, M. J., & McCauley, D. E. (1988). Extinction and recolonization: Their effects on the genetic differentiation of local populations. *Evolution*, *42*(5), 995–1005.

Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

Winkler, M., Hülber, K., & Hietz, P. (2009). Population dynamics of epiphytic orchids in a metapopulation context. *Annals of Botany*, *104*, 995–1004.

Wohlfarth, B., Björck, S., Funder, S., Houmark-Nielsen, M., Ingólfsson, Ó., Lunkka, J.-P., Mangerud, J., Saarnisto, M., & Vorren, T. (2008). Quaternary of Norden. *Episodes*, *31*, 73–81.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Hedrén M, Birkedal S, de Boer H, et al. Asymmetric contributions of seed and pollen to gene dispersal in the marsh orchid *Dactylorhiza umbrosa* in Asia Minor. *Mol Ecol*. 2021;30:1791–1805. https://doi.org/10.1111/mec.15848