Journal of Experimental Botany, Vol. 72, No. 2 pp. 654–681, 2021
doi:10.1093/jxb/eraa467 Advance Access Publication 15 January 2021

RESEARCH PAPER

Whole plastomes are not enough: phylogenomic and morphometric exploration at multiple demographic levels of the bee orchid clade *Ophrys* sect. *Sphegodes*

Richard M. Bateman*, Paula J. Rudall, Alexander R.M. Murphy, Robyn S. Cowan, Dion S. Devey and Oscar A. Peréz-Escobar

Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AB, UK

* Correspondence: r.bateman@kew.org

Received 3 June 2020; Editorial decision 23 September 2020; Accepted 15 December 2020

Editor: Rainer Melzer, University College Dublin, Ireland

Abstract

Plastid sequences have long dominated phylogeny reconstruction at all time depths, predicated on a usually untested assumption that they accurately represent the evolutionary histories of phenotypically circumscribed species. We combined detailed *in situ* morphometrics (124 plants) and whole-plastome sequencing through genome skimming (71 plants) in order to better understand species-level diversity and speciation in arguably the most challenging monophyletic group within the taxonomically controversial, pseudo-copulatory bee orchid genus *Ophrys*. Using trees and ordinations, we interpreted the data at four nested demographic levels—macrospecies, mesospecies, microspecies, and local population—seeking the optimal level for bona fide species. Neither morphological nor molecular discontinuities are evident at any level below macrospecies, the observed overlap among taxa suggesting that both mesospecies and microspecies reflect arbitrary division of a continuum of variation. Plastomes represent geographic location more strongly than taxonomic assignment and correlate poorly with morphology, suggesting widespread plastid capture and possibly post-glacial expansion from multiple southern refugia. As they are rarely directly involved in the speciation process, plastomes depend on extinction of intermediate lineages to provide phylogenetic signal and so cannot adequately document evolutionary radiations. The popular 'ethological' evolutionary model recognizes as numerous 'ecological species' (microspecies) lineages perceived as actively diverging as a result of density-dependent selection on very few features that immediately dictate extreme pollinator specificity. However, it is assumed rather than demonstrated that the many microspecies are genuinely diverging. We conversely envisage a complex four-dimensional reticulate network of lineages, generated locally and transiently through a wide spectrum of mechanisms, but each unlikely to maintain an independent evolutionary trajectory long enough to genuinely speciate by escaping ongoing gene flow. The frequent but localized microevolution that characterizes the *Ophrys sphegodes* complex is often convergent and rarely leads to macroevolution. Choosing between the contrasting 'discontinuity' and 'ethology' models will require next-generation sequencing of nuclear genomes plus ordination of corresponding morphometric matrices, seeking the crucial distinction between retained ancestral polymorphism—consistent with lineage divergence—and polymorphisms reflecting gene flow through 'hybridization'—more consistent with lineage convergence.

Keywords: Demographic level, genome skimming, morphometrics, next-generation sequencing, *Ophrys sphegodes*, phylogeny, plastome phylogenomics, sexual deceit, speciation, species circumscription.

Introduction

The last three decades have witnessed great change in the realm of molecular systematics. Early studies were typological in two important ways. Firstly, a single plant was required by authors to epitomize an entire species or, more often, an entire supraspecific taxon. Secondly, in most cases, a single genic region was required to represent both (animal) or all three (plant) genomes present in each study organism. The nuclear genome was most commonly represented by nuclear ribosomal regions [the comparatively slowly evolving 26S or 28S regions or the faster evolving internal transcribed spacer (ITS) or external transcribed spacer (ETS) regions], whereas the organelles were represented by the mitochondrial gene *cox1* (animals) or the plastid genes *rbcL* or *matK* (plants). These 'traditional' genic regions still figure prominently in recommendations for global-scale identification of organisms through more modern DNA 'barcoding' approaches (e.g. Li *et al.*, 2015; Bateman, 2016; Hollingsworth *et al.*, 2016).

Progressive advances in DNA sequencing technologies are now carrying us ever deeper into the genomics era. Increasingly automated analytical techniques of next-generation/second-generation sequencing are engaged in the process of unlocking the potential for almost limitless numbers of organisms to be characterized for almost limitless numbers of base pairs per plant (Harrison and Kidner, 2011; Lemmon and Lemmon, 2013; McCormack *et al.*, 2013; Reuter *et al.*, 2015; Olson *et al.*, 2016). For centuries it has been possible, at least in theory, to reveal every detail of the phenotype of an organism, given sufficient time and effort. Today, it is also theoretically possible to reveal every detail of the genotype of an organism—a goal that arguably is now considerably easier to achieve than thoroughly exploring its phenotype. It is tempting to assume that comprehensive molecular datasets obtained from vastly increased numbers of organisms should yield greatly improved circumscription of species and, through the generation of molecular trees, greatly improved circumscription of statistically robust (and ideally monophyletic) higher taxa. Determining whether such radical progress is indeed occurring is hampered by the fact that (other than in the case of briefly repeatedly divided and irradiated subpopulations of a cultivated bacteriophage; Hillis *et al.*, 1992; Bull *et al.*, 1993) we do not have access to the absolute truth offered by the yardstick of a known genealogy. All other factors being equal, we naturally assume—usually with little if any accompanying discussion—that the larger the underlying matrix the greater is the probability that it will yield an accurate phylogeny.

This assumption raises a further practical consideration of great importance, specifically the degree to which concatenation of contrasting categories of data constitutes the optimal approach to data analysis when attempting to reconstruct phylogenies. Different regions within the same genome are now routinely concatenated, and there is an increasing trend of also combining multiple genomes (for plants: nuclear plus plastid, and occasionally also mitochondrial). In the increasingly rare phylogenetic studies in which morphology has been scored and analysed cladistically (rather than simply mapped across molecular trees as so-called 'traits'), that data category is also sometimes concatenated with—perhaps better described as buried within—the inevitably far larger spectrum of molecular characters. Only comparatively rarely has the argument been advanced that greater insights, particularly at the level of process rather than pattern, may be gained by analysing contrasting categories of data separately, in order to develop process-based explanations of the almost inevitable topological incongruences that emerge between the taxonomic circumscriptions and/or phylogenies that are generated (e.g. Bateman, 2020; topic reviewed by Smith *et al.*, 2020).

Here, we apply genomic and morphometric techniques to an especially recalcitrant clade of putatively radiating temperate terrestrial orchids, with the joint aims of better understanding speciation in the group and of evaluating the relative merits of genome skimming (syn. ultra-barcoding; Elshire *et al.*, 2011; Kane *et al.*, 2012; Berger *et al.*, 2017; Wickland *et al.*, 2017) and *in situ* morphometrics (*sensu* Bateman, 2001; Bateman *et al.*, 2017) for taxonomic circumscription and phylogeny reconstruction in higher plants. Naturally, we focus on attempting to identify optimally the all-important, and theoretically self-defining, species level. Our study also adds to the growing body of evidence (epitomized by Cozzolino *et al.*, 2020) that even whole-organelle sequences are inadequate for circumscribing recently speciated or actively speciating lineages.

Study group

Framework phylogenies have been generated for bee orchids of the genus *Ophrys* L. (Orchidoideae: Orchidaceae) based on both morphological (Devillers and Devillers-Terschuren, 1994; Bateman *et al.*, 2018*a*) and molecular (Soliva *et al.*, 2001; Bateman *et al.*, 2003, 2018a; Devey *et al.*, 2008; Breitkopf *et al.*, 2015) data, the molecular work reliably circumscribing three subgenera but failing to adequately resolve their topological relationship (Bateman *et al.*, 2018*a*).

However, this phylogenetic research has simply fuelled the controversy that has for so long dogged species circumscription within *Ophrys*. This highly charismatic genus is strikingly attractive and is the archetypal model system for the unusual pollination mechanism of pseudo-copulation; naïve male insects are duped by olfactory, visual, and tactile cues into transferring pollen masses between flowers on the rare occasions when they repeatedly attempt to mate with them (e.g. Kullenberg, 1961; Paulus and Gack, 1990; Vereecken, 2009; Vereecken *et al.*, 2012; Paulus, 2019). Consequently, *Ophrys* has attracted much attention from both evolutionary biologists (e.g. Schiestl *et al.*, 1999; Vereecken and Schiestl, 2008; Breitkopf *et al.*, 2013,

2015; Sedeek *et al*., 2014; Baguette *et al*., 2020) and taxonomists (e.g. Devillers and Devillers-Terschuren, 1994; Kreutz, 2004; Pedersen and Faurholdt, 2007; Hennecke, 2013; Véla *et al*., 2015; Delforge, 2016). Recent published estimates of the number of species in the genus range from nine (Bateman *et al*., 2018*a*) to 353 (Delforge, 2016); if accepted, the higher figure would constitute approximately half of all orchid species native to Europe and Asia Minor! Moreover, new *Ophrys* 'species' continue to be formally described through traditional taxonomy (often following claims of pollinator specificity) at the rate of ~10 per annum, along with an ever-more confusing panoply of infraspecific taxa of even more questionable value. Obviously, species concepts applied (all too often implicitly rather than explicitly) have differed radically among the many authors who have studied *Ophrys* and impact greatly on the perceived and actual diagnosability of the resulting taxa. Attempts by some authors to side-step this perennial controversy by arguing that they are studying the speciation process rather than species delimitation (e.g. Baguette *et al*., 2020) fall at the first logical hurdle; by definition, only bona fide species can have undergone speciation successfully.

In order to discuss the products of contrasting species concepts applied to the genus, Bateman (2018; also Bateman *et al*., 2018*a*) established a pragmatic working terminology of three levels of 'species': microspecies recognized by ethologically oriented taxonomists such as Delforge (2016) on the basis of intuited morphological difference; mesospecies defined pragmatically as the 23 categories into which Delforge (2016) grouped these microspecies on supposedly reliable morphological distinctions; and nine macrospecies that could be distinguished by Devey *et al*. (2008) on the basis of a minimum of one or two fairly reliable (though imperfectly differentiated) base pair differences within the nuclear ribosomal ITS (nrITS) 'barcode' region. Throughout this paper, microspecies epithets only are given in uniformly lower case letters, whereas mesospecies and macrospecies begin with an upper case letter.

Among the nine macrospecies of *Ophrys*, the richest in both Delforgean mesospecies (nine) and microspecies (113) is the *Sphegodes* clade. Even Pedersen and Faurholdt (2007) felt obliged to divide the *Sphegodes* clade into seven species, despite advocating a 'lumpers' classification in their monograph that recognized only 19 species across the entire genus *Ophrys*. Highly evolutionarily derived and closely related sister to the similarly microspecies-rich *Fuciflora* clade (Devey *et al*., 2008; Bateman *et al*., 2018*a*; G. Sramkó *et al*., unpublished), the *Sphegodes* clade spans almost the entire geographic range of the genus, extending north–south from southern England to southernmost Spain and west–east from Portugal to Iran. Although the clade encompasses only a few near-identical ribotypes and so certainly qualifies as a single macrospecies, it demonstrably possesses an unusually wide range of subtle phenotypic variation, not only in (micro)morphology (Bradshaw *et al*., 2010) but also in biochemistry, phenology, pollinator spectrum, and habitat preference (e.g. Ayasse *et al*., 2011; Xu *et al*., 2011; Breitkopf *et al*., 2013; Sedeek *et al*., 2014). Admittedly, examination of extreme phenotypes such as those illustrated in Fig. 1 may appear to challenge the counter-argument that they constitute elements of a phenotypic and genotypic continuum (as claimed by Bateman *et al*., 2011; Bateman, 2018, 2021)—a viewpoint that has been heavily criticized by some other observers (e.g. Vereecken *et al*., 2011; Paulus, 2018, 2019; Baguette *et al*., 2020).

Here, we use a sampling strategy designed to explore every demographic level (from macrospecies>mesospecies>microspecies>population>individuals within population) in order to test Bateman's (2018, 2020) continuum hypothesis of variation within macrospecies *Sphegodes*. Separate analyses of morphology based on *in situ* morphometrics and plastid sequences based on genome skimming provide a context within which a representative sample set of plants is used as the basis of a more focused analysis. We explore the relationship between detailed genotype and detailed phenotype to compare the relative strengths of signal they provide for both species circumscription and phylogeny reconstruction, seeking more general conclusions that could help to optimize approaches to the study of evolution in other taxonomically challenging groups. Most notably, we reappraise the reliability of plastome data for circumscribing, and determining the relationships among, recently diversified groups of higher plants, particularly in the light of the study by Cozzolino *et al*. (2020) that not only demonstrated but also cogently explained incongruity between nuclear and organellar genomes when used to explore the relationships between four sympatric microspecies within the *O. sphegodes* complex.

Materials and methods

Plant materials

Fieldwork for our broader research programme targeted specifically at the genus *Ophrys* began in 2004 and is ongoing; populations have been sampled across the geographic range of the genus other than the Levant and the Caucasus. Our standard procedure is to randomly select study plants within field populations and to measure their vegetative characters *in situ*, also obtaining 1:1-scale images of a representative flower taken from each plant (Bateman, 2001, 2011). Sampling is thus confined to excising one flower for mounting and morphometric study later in the same day and placing a second flower in a sachet of fine-ground silica-gel for subsequent DNA analyses (though note that the supplementary silica-gel samples obtained for us by other collectors inevitably lacked corresponding morphometric datasets). The number of morphometric datasets accumulated thus far exceeds 600, one-third of which represent the *Sphegodes* group that is the focus of this study.

Morphometrics

The 51 morphometric characters employed in the present study are described in greater detail in Table 1. In retrospect, our study would have benefited from the inclusion of an additional phase of data collection conducted under a binocular microscope, in order to better detail features of the gynostemium, stigmatic surface, and labellar 'neck'. The 53 characters initially measured did include two microscopic characters

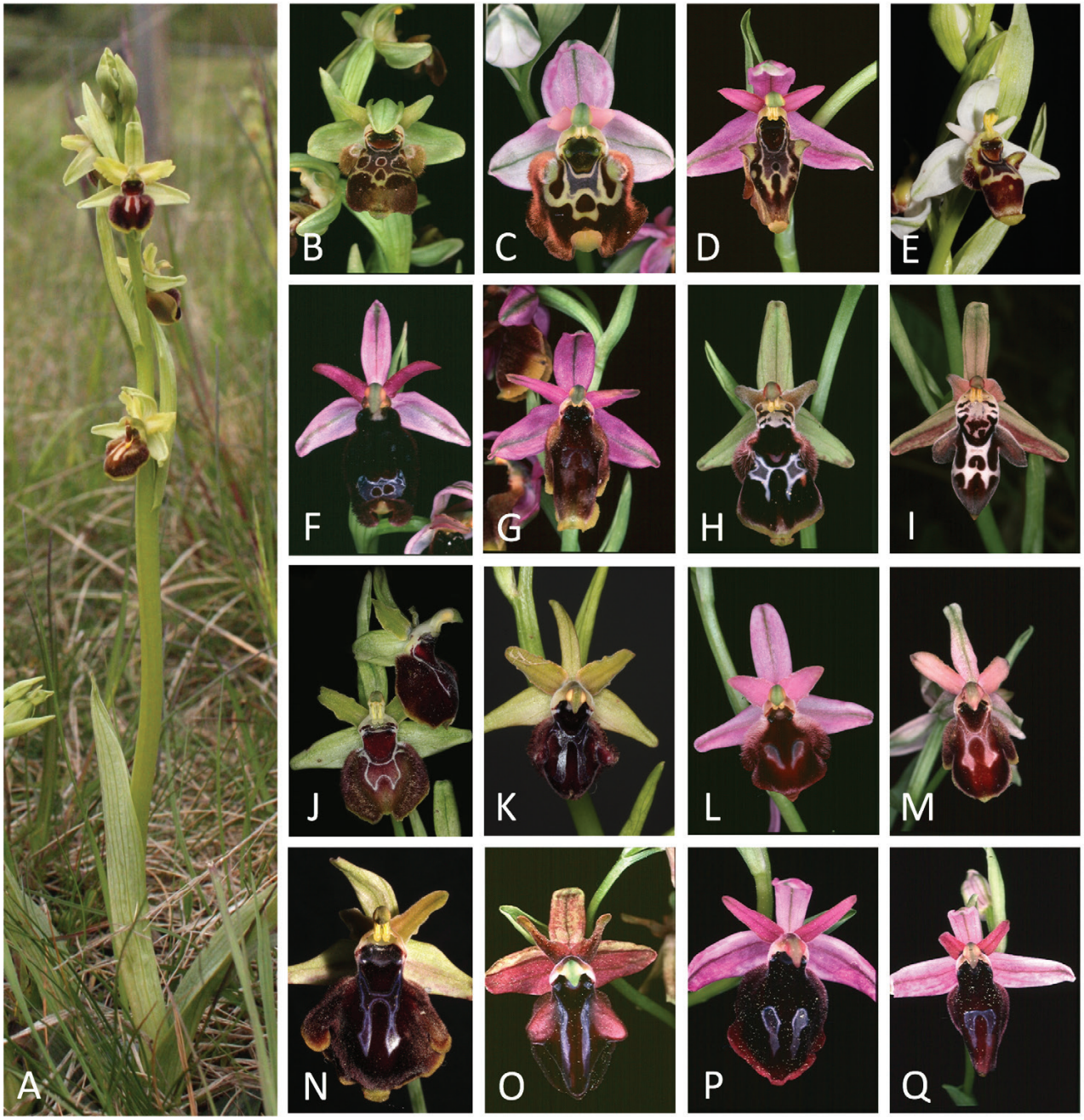

Fig. 1. Plate illustrating the typical shared vegetative phenotype (A) and a representative diversity of floral phenotypes (B–Q) exhibited by the *Ophrys sphegodes* clade and its two closest outgroup macrospecies. Names are given as Mesospecies–microspecies followed by source locality. Entire plant. (A) *Ophrys sphegodes s.s.*, N France. Flowers. Outgroups: Macrospecies *Umbilicata*: (B) *Umbilicata–attica*, S Greece. Macrospecies *Fuciflora*: (C) *'Bornmuelleri'!–episcopalis*, Crete; (D) *Heldreichii–homeri*, Chios; (E) *Scolopax–picta*, S Spain. Ingroup: Macrospecies *Sphegodes*: (F) *Bertolonii–bertolonii*, Sicily; (G) *Lunulata–lunulata*, Sicily; (H) *Reinholdii–reinholdii*, S Greece; (I) *Reinholdii–cretica*, Crete; (J) *Provincialis–provincialis*, S France; (K) *Incubacea–incubacea*, Sicily; (L) *Argolica–argolica*, S Greece; (M) *Exaltata–exaltata*, Sicily; (N) *Mammosa–grigoriana*, Crete; (O) *Mammosa–spruneri*, Chios; (P) *Mammosa–ferrum-equinum* (*ferrum-equinum* morph), Chios; (Q) *Mammosa–ferrum-equinum* (*labiosa* morph), Chios. Plant (A) is ~22 cm tall; the horizontal dimension of images (B)–(Q) is 22 mm. Images: Richard Bateman.

describing marginal bract cells, but these characters rapidly proved to be insufficiently informative relative to the time consumed in recording them and hence they were soon discarded. The remaining 51 characters contributing to the statistical analyses describe the stem and inflorescence (5), leaves and bracts (7), gynostemium and ovary (3), labellum (20), and lateral petals and sepals (16). They can alternatively be categorized as metric (33), meristic (3), multistate-scalar (13), and bistate (2). Flower colour was recorded by matching the colour of the lower half of the labellum (excluding the speculum), the sepals, and the lateral petals to the closest colour block(s) of the Royal Horticultural Society Colour Chart,

Table 1. List of 51 morphometric characters measured from 124 individuals of the *Ophrys sphegodes* group plus three representative outgroup members

(A) Labellum (20 characters)
1. Maximum width
2. Maximum length (excluding appendix)
3. Depth of indentation [if present], from the maximum extent of the lateral portion of the central lobe inward to the base of the notch containing the appendix
4. Maximum length of speculum
5. Maximum width of speculum
6. Speculum position relative to stigma (scale 1–3) (grades into stigma: connected to stigma: not connected to stigma)
7. Pale zone along lower half to entire margin of speculum (scale 0–2) (absent: subdued: prominent)
8. Speculum shape (scale 1–4) (entire + U + W: I I + o o: H: single ring with radiating projections+three rings)

X. Base colour immediately below speculum
9. Colour (x)
10. Colour (y)
11. Colour (Y, %)
12. Width of pale-coloured marginal zone [if present] of labellum
13. Pilosity of central lobe margin of labellum 1 mm inside the margin and immediately above the appendix (scale 0–2) (none/negligible: short: long)
14. Pilosity of central lobe margin of labellum 1 mm inside the margin and at 45° to the vertical (scale 0–2) (none/negligible: short: long)
15. Pilosity of 'shoulders'/lateral lobes of labellum 1 mm inside the margin (scale 0–2 (none/negligible: short: long)
16. Appendix length [if present]
17. Appendix width [if present]
18. Length of 'horns' [if present]
19. Maximum length of lateral lobes [if present] following mounting
20. Degree of curvature of labellum viewed transversely from base (scale 1–3) (±flat: gently convex: lateral lobes strongly recurved)

(B) Lateral petals and sepals (16 characters)
21. Length of lateral petals
22. Maximum width of lateral petals
23. Basal lateral teeth on lateral petals (scale 0–2) (absent: subdued: prominent)

X. Base colour of lateral petals
24. Colour (x)
25. Colour (y)
26. Colour (Y, %)
27. Degree of curvature of lateral petals (scale 1–5) (strongly deflexed: deflexed: ±flat: recurved: strongly recurved)
28. Length of lateral sepals
29. Maximum width of lateral sepals

X. Base colour of upper half of lateral sepals
30. Colour (x)
31. Colour (y)
32. Colour (Y, %)
33. Degree of curvature of lateral sepals (scale 1–5) (strongly deflexed: deflexed: ±flat: recurved: strongly recurved)
34. Degree of curvature of median sepal (scale 1–5) (strongly deflexed: deflexed: ±flat: recurved: strongly recurved)
35. Outline shape of median sepal (scale 1–3) (basally expanded obovate: ovate: apically expanded obovate)
36. Suffusion of dark pigment in lower half of lateral sepal (0/1=absent/present)

(C) Column and ovary (3 characters)
37. Length of ovary
38. Length of column
39. Maximum width of column

(D) Stem and inflorescence (5 characters)
40. Stem height
41. Stem diameter immediately above leaves
42. Inflorescence length
43. Number of flowers/buds
44. Angle subtended by labellum relative to stem (scale 1–3) (0–30°=parallel: 31–60°: 61–90°=perpendicular)

(E) Leaves and bracts (7 characters)
45. Number of basal (spreading) leaves
46. Number of sheathing (±upright) leaves
47. Length of longest basal leaf
48. Maximum width of longest basal leaf
49. Position of maximum with relative to position along length from base (scale 1–2) (<50%: >50%,=ovate-lanceolate: obovate leaf shapes)
50. Length of basal bract
51. Maximum width of basal bract

Numbers of the 15 characters measured in the field are italicized; colours were matched to the RHS colour chart before conversion to CIE coordinates.

for subsequent conversion into three quantified variables long recognized by the Commission Internationale de l'Eclairage.

Data for individual plants were summarized in an Excel v15.4 spreadsheet. Two rounds of multivariate data analysis were performed. The first, stand-alone analysis involved the complete matrix of 124 individuals (Supplementary Table S1 at *JXB* online), together encompassing all nine mesospecies and 31 of the 113 microspecies listed by Delforge (2016). The second analysis was confined to the 24 plants that also yielded corresponding whole-plastome datasets: 21 plants from the *Sphegodes* clade plus two 'inner outgroups' of the *Fuciflora* clade and one 'outer outgroup' from the *Umbilicata* clade, as resolved by Bateman *et al.* (2018*a*).

The larger morphometric matrix of 124 plants×51 characters (total 6324 cells) contained 286 (4.5%) missing values that largely reflected two contrasting influences: premature desiccation of the leaves of plants growing in unusually arid environments, and the addition of four characters (C7, C14, C23, and C35) to our original character list only after the first 11 plants had already been measured. A smaller matrix of 24 plants×51 characters (total 1224 cells) was developed for explicit comparison of the inferred levels of phenotypic divergence with levels of genotypic divergence assessed through genome skimming. This much-reduced matrix contained 29 (2.4%) missing values: of these, all represented vegetative measurements, the most severely affected being two plants of the mesospecies *Mammosa* that suffered premature desiccation of vegetative organs in the comparatively arid climate of Cyprus. Two of the 51 scored characters (C6 and C23) proved to be invariant within the spectrum of taxa analysed in this smaller matrix.

For both matrices, the assembled data were analysed via multivariate methods using Genstat v14 (Payne *et al.*, 2011). They were employed to compute a symmetrical matrix that quantified the similarities of pairs of datasets (i.e. plants) using the Gower similarity coefficient (Gower, 1971) on unweighted datasets scaled to unit variance. The matrix was in turn used to construct a dendrogram and a minimum spanning tree (Gower and Ross, 1969) and subsequently to calculate principal coordinates (Gower, 1966, 1985)—compound vectors that incorporate positively or negatively correlated characters that are most variable and therefore potentially diagnostic of putative taxa. Principal coordinates are especially effective for simultaneously analysing heterogeneous suites of morphological characters and can comfortably accommodate missing values. They have proven invaluable for assessing relationships among orchid species and populations throughout the last three decades (reviewed by Bateman 2001, 2011, 2020) and are the crux of the morphometric element of the present study.

For each multivariate analysis, the first four principal coordinates (PC1–PC4) were plotted together in pairwise combinations to assess the degree of morphological separation of individuals (and thereby of populations and taxa) in these dimensions, and pseudo-*F* statistics were obtained to indicate the relative contributions to each coordinate of each of the original variables. The resulting ordinations were presented using Deltagraph v7.1 (SPSS/Red Rock software).

Genome skimming

Near-complete plastome sequences were obtained via next-generation sequencing (NGS) genome skimming from 71 accessions of *Ophrys* (Supplementary Table S2): seven outgroups plus 64 plants of the *Sphegodes* group (sampled from 57 localities across Europe), together representing 40 of the 113 microspecies, and all nine of the mesospecies, listed for the molecularly circumscribed *Sphegodes* group by Delforge (2016).

Data generation

We extracted genomic DNA from silica-dried flowers using the Quiagen DNeasy Plant kit, following the manufacturer's protocol. High-throughput sequencing involved preparing a genomic Illumina paired-end library using the NEBNext Ultra II library preparation kit, following the manufacturer's protocol and setting an average insert size of 300 bp. Library sequencing was performed on a MiSeq platform, yielding ~15 Gb and a total of 55 million paired-end reads.

The Illumina raw reads were quality filtered using Trim Galore v0.4 (Krueger, 2015), discarding sequences with an averaged Phred33 score <20. Pre- and post-trimming read quality was assessed using FASTQC v0.1 (Andrews *et al.*, 2015). We mapped trimmed reads of all sequenced accessions to plastid target 'genomes of reference' belonging to *O. sphegodes sensu stricto* (*s.s.*) (Roma *et al.*, 2018: GenBank accession no. AP018717.1) and the very molecularly similar *O. aveyronensis* (Bertrand *et al.*, 2019: GenBank accession no. MN120441), checked against additional data released by Cozzolino *et al.* (2020). Read mapping, alignment, and DNA damage analyses were implemented through the pipeline PALEOMIX v1.2.13 (Schubert *et al.*, 2014). The trimmed read data were mapped using BowTie v2.3.4.1 with a Phred score quality-filter value of 20, followed by a realigning step around indels and filtering of duplicated reads with GATK v3.8.1 (McKenna *et al.*, 2010) and Picard Tools v1.137 (Broad Institute, accessed October 2019). Read mapping, PCR duplicate, and average coverage statistics for each accession analysed in this study are provided in Supplementary Table S3.

Data analysis

We produced consensus plastome sequences of *Ophrys* accessions from the BAM files produced by PALEOMIX by following a modified statistical base-calling approach of Li *et al.* (2008), requiring a minimum depth coverage of 10 and bases matching at least 50% of the reference sequence. The whole-plastome consensus sequences were produced in Geneious v8.0 and aligned using Mauve, employing a progressive algorithm and assuming collinearity (Darling *et al.*, 2004). We recovered ~147 000 bp out of an estimated whole-plastome total of ~150 000 bp. For phylogenomic analysis, the resulting alignment was first trimmed to exclude misaligned regions and positions with >90% missing data and then subjected to maximum likelihood (ML) tree inference in RAxML v8.0 (Stamatakis, 2014), using the GTR substitution model, 25 gamma categories, and 1000 bootstrap replicates.

To ordinate plastomes of the *O. sphegodes* complex, we relied on genotype likelihoods (GLs) to assess the relatedness of each sequenced accession in a plastid genomic context. We computed plastid GLs using the software ANGSD v0.929 (Korneliussen *et al.*, 2014) by implementing the GATK GL model, inferring the minor and major alleles, and retaining polymorphic sites with a minimum *P*-value of $1\text{-}e^{6}$. A covariance matrix was derived from plastome GL values using PCAngds (Meisner and Albrechsten, 2018), and principal components were computed from the matrix using the function prcomp of the R package STATS. The first two components were plotted as a dotplot using the package ggplot2 (Wickham, 2016) before final versions were generated in Deltagraph.

Morphological versus molecular distances

Degrees of contrast between the morphometric matrix and the molecular single nucleotide polymorphism (SNP) matrix were compared for the reduced sample set of 24 plants in all pairwise combinations to generate two corresponding symmetrical matrices. Morphometric similarities were calculated via Genstat as Gower similarity values (Gower, 1971), whereas for the plastome data percentage dissimilarities in SNPs (excluding any position coded as a gap or missing in any of the 24 sequences) were calculated via MEGA v7.0 (Kumar *et al.*, 2016). The resulting 276 pairwise comparisons allowed us to estimate the degree of correlation between phenotypic and genotypic disparity within a single sample set, albeit a sample set much smaller than the 70+ accessions originally intended.

Results

Morphometrics

The first four principal coordinates for the larger morphometric matrix, including 124 plants of macrospecies *Sphegodes* but no outgroups, are plotted in pairwise combinations in Fig. 2 and Supplementary Fig. S1, individual plants being labelled according to mesospecies *sensu* Delforge (2016). Although weak, the first two coordinates are considerably stronger than the remainder. The first coordinate is dominated by the contrast of plants with green (left) versus pink (right) lateral petals, which divides the plants into two ill-defined clusters (Supplementary Fig. S2A). Subordinate contributory characters show plants located toward the right of the plot to typically possess labellar appendixes, have longer columns, larger sepals, and labella that are comparatively hairy but tend to lack forward-pointing 'horns'. The second coordinate represents several more evenly weighted characters that together reflect comparatively large plant size and, to a lesser degree, large flower size—effectively constituting a vigour coordinate. The third coordinate combines other vegetative dimensions and leaf number with a mixed bag of labellum characters (Supplementary Fig. S1). All coordinates of fourth order and below reflect very few characters and offer little if any taxonomic discrimination.

Viewed at the level of mesospecies, the impact of the first coordinate (Fig. 2A) is overly dependent on whether the

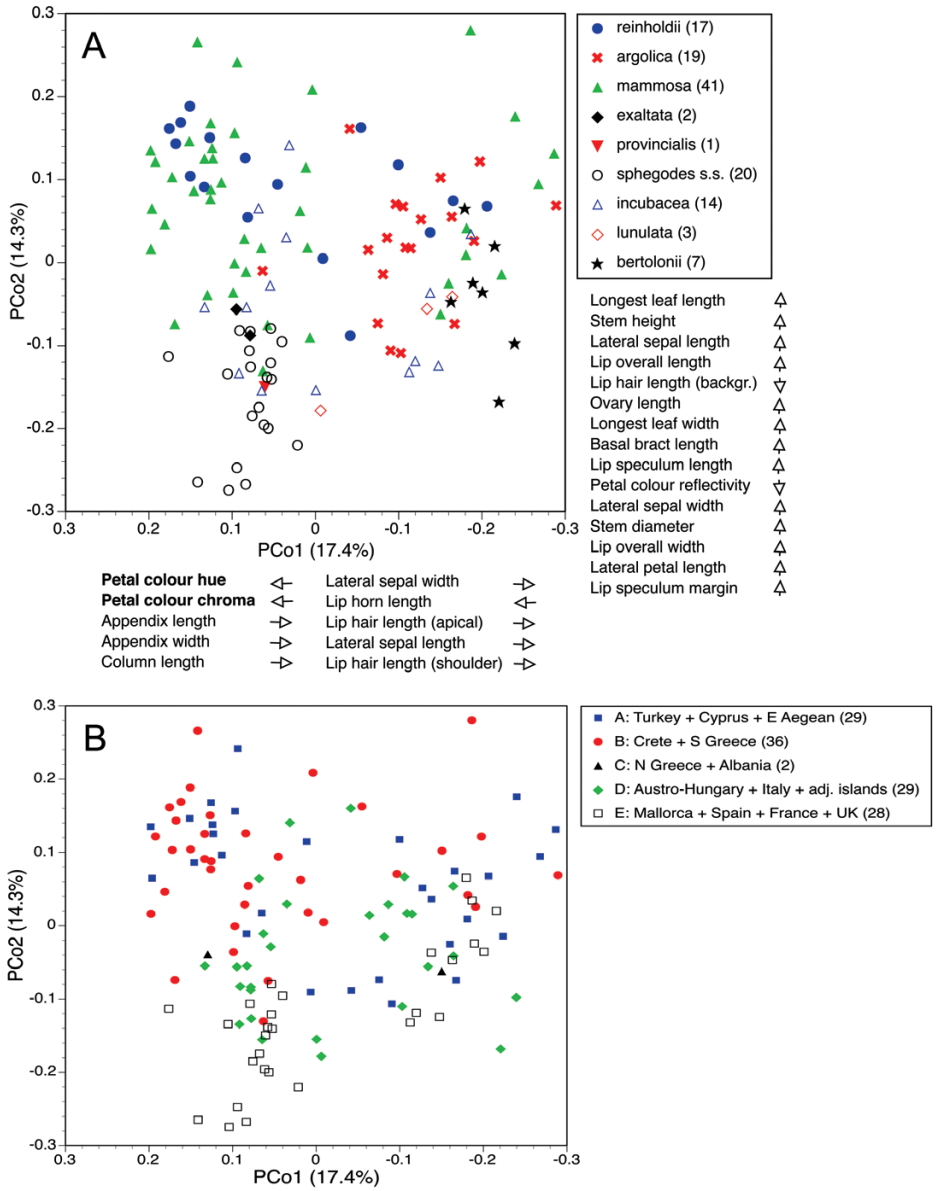

Fig. 2. Plot of principal coordinates 1 and 2 for 51 morphometric characters and 124 individuals of the *Ophrys sphegodes* clade. (A) Labelled for mesospecies groups. (B) Labelled for geographic region of origin.

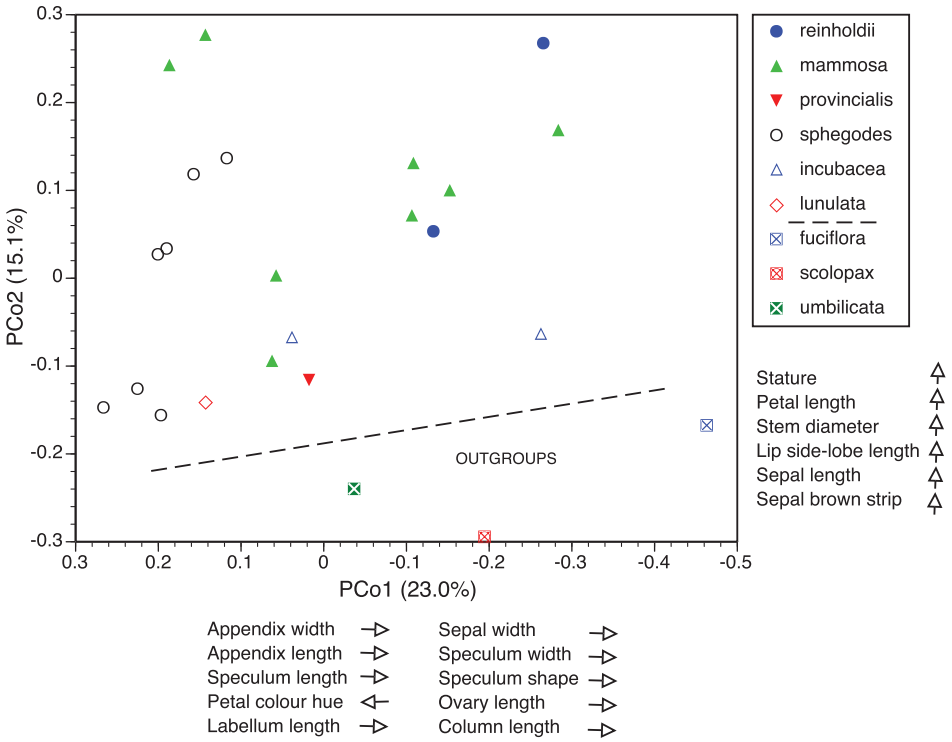

Fig. 3. Plot of principal coordinates 1 and 2 for 49 morphometric characters and 24 *Ophrys* individuals for which whole-plastome data are also available: 21 belong to the macrospecies *Sphegodes*, two to the macrospecies *Fuciflora*, and one to the macrospecies *Umbilicata*.

mesospecies in question encompasses a mixture of green- and pink-petaled plants; consequently, those mesospecies considered capable of exhibiting both colours (*Mammosa*, *Incubacea*, and *Reinholdii*) are spread more widely along PC1 than are those that are either reliably green (all lack anthocyanins, e.g. *Sphegodes*) or reliably pink (all possess anthocyanins, e.g. *Bertolonii*). The instability of this character is emphasized by the fact that a single self-pollinated flower can produce both green-flowered and pink-flowered progeny (Malmgren, 2008). The second coordinate gives almost complete separation of the large-bodied, large-flowered mesospecies *Mammosa* and *Reinholdii* from the more modestly sized mesospecies *Sphegodes* and *Provincialis*. The third coordinate serves only to partially separate mesospecies *Incubacea* from the remainder (Supplementary Fig. S1).

Morphological variation within populations of a single microspecies was estimated through analysis of 15 pairs of con-microspecific plants that together encompassed the full morphological range exhibited by macrospecies *Sphegodes*. Distances separating these paired individuals in Fig. 2 varied greatly from <0.01 to 0.21, averaging 0.050±0.046 for PC1 and 0.079±0.061 for PC2. The comparatively large mean value for PC2 relative to PC1 is readily explained by contrasts in plant size that are likely to reflect differences in development (ontogeny) and environment of growth (ecophenotypy) at least as much as any direct genetic influence. When this ordination is labelled according to geographic region rather than mesospecies identity, it becomes clear that only the second coordinate shows a geographically correlated trend—one of diminishing sizes of flowers and especially of vegetative organs from east to west (Fig. 2B).

Unsurprisingly, analysis of the smaller morphometric matrix (24 plants, including three outgroups) allows the first two coordinates to capture a greater proportion of the total variance (Fig. 3). Once again, appendix and column dimensions and petal colour strongly influence the first coordinate, though here they are combined with the size and shape of the speculum, presumably reflecting the inclusion of outgroups that deviate somewhat in these characters. This coordinate correlates poorly with assignment of plants to mesospecies, other than in partially separating mesospecies *Sphegodes* from the remainder. The second coordinate provides no discrimination among mesospecies within the *Sphegodes* clade. As in the larger analysis, coordinate 2 once again involves stem dimensions, but here it also features the lengths of both lateral petals and sepals and the presence of lateral lobes—characters that serve primarily to separate the macrospecies *Sphegodes* from the outgroups chosen to represent macrospecies *Fuciflora* and *Umbilicata* (Fig. 3).

Genome skimming: trees

Phylogenies derived from the 'whole-plastome' tree (Fig. 4) support previous studies in distinguishing three macrospecies

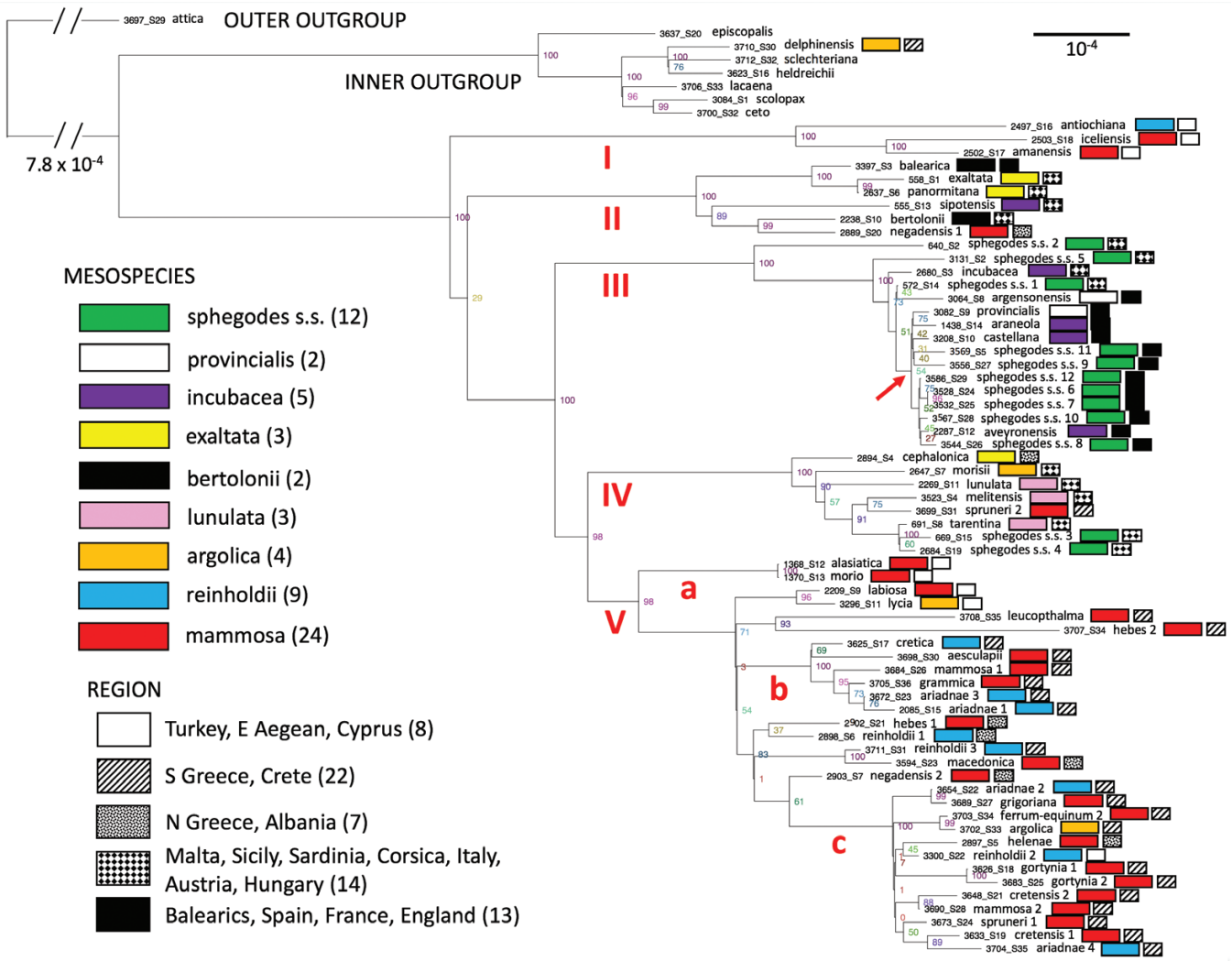

Fig. 4. Maximum likelihood phylogeny of whole-plastomes obtained via genome skimming from 64 individuals of the *Sphegodes* macrospecies, six individuals of the *Fuciflora* macrospecies ('inner outgroup'), and one individual of the *Umbilicata* macrospecies (functional 'outer outgroup'). Plants are labelled according to both mesospecies attribution and geographic region sampled. Figures supporting nodes are bootstrap percentages.

among the analysed accessions and in placing *O. umbilicata sensu lato* (*s.l.*) as sister to *O. fuciflora s.l.* plus *O. sphegodes s.l.*, each macrospecies being subtended by a comparatively long branch and receiving 100% bootstrap support. However, a single accession of the localized Greek microspecies *O. delphinensis* (traditionally ascribed to mesospecies *Argolica* of macrospecies *Sphegodes*) occupies a derived position within the macrospecies *Fuciflora* rather than the morphologically predicted placement within the macrospecies *Sphegodes*. The *Fuciflora* clade resembles the *Sphegodes* clade in many properties, including the facts that most internal relationships are weakly supported and that microspecies show little correlation with either mesospecies assignment or geographic origin (G. Sramkó *et al.*, unpublished).

The 64 samples of the *Sphegodes* clade resolved into five main clades (denoted by roman numerals in Fig. 4). All but the largest [clade V, bootstrap support (BS) 98%] received 100% bootstrap support, though the node separating clade I from clade II attracted considerably less statistical support. Each Delforgean mesospecies was represented by between two and 24 plants, but even with this limited sampling, all nine mesospecies proved to be polyphyletic as perceived through their plastome sequences (Fig. 4A). The three least well-sampled mesospecies were confined to single plastome clades (*Bertolonii*=II, *Provincialis*=III, and *Lunulata*=IV), but a further five mesospecies were each divided between two of the clades. Moreover, the 24 analysed plants of the most intensively sampled mesospecies, *Mammosa*, were scattered across four of the five clades, being absent only from clade (III) dominated by samples of *O. sphegodes s.s.*

Moving down one demographic level to consider mesospecies, even multiple samples of the same microspecies typically failed to resolve as monophyletic groups. Nine of the microspecies were each analysed for between two and four samples (Fig. 4). The pair of samples of *O. negadensis* (mesospecies *Mammosa*) were divided between plastome clades,

the Albanian sample being placed in clade II but the northern Greek sample occupying clade V. Similarly, the Peloponnesian sample of *O. spruneri* (mesospecies *Mammosa*) was placed in clade IV but the Cretan specimen appeared in clade Vc. Most of the remaining multiply sampled microspecies were also assignable to mesospecies *Mammosa* (*O. mammosa*, *O. ferrum-equinum*, *O. gortynia*, *O. cretensis*, and *O. hebes*, each two samples), though two were assignable to mesospecies *Reinholdii* (*O. ariadnae*, four samples; *O. reinholdii*, three samples). All seven of these multiply sampled microspecies were confined to the largest plastome clade, V, but within this clade multiple accessions of a single microspecies were typically widely separated. Only one pair of samples representing the same microspecies were resolved as sisters (two *O. gortynia* samples from Crete, in clade Vc), implying an overall frequency of sisterhood so low that it could reasonably be explained by chance. Of the 12 accessions of *O. sphegodes s.s.* sequenced by us, 10 spanned the full range of sequence variation within clade III but nonetheless two of the five central European samples of *sphegodes s.s.* included in our analysis were instead nested within clade IV.

Only when we move down a further demographic level to consider accessions of the same microspecies sampled from the same local geographic area do plants begin to cluster together on comparatively short terminal branches. The best example is the more derived portion of clade III (BS=54%; indicated by an arrow in Fig. 4), where near-identical plastomes characterized all six samples of the microspecies *O. sphegodes s.s.* collected from four populations scattered along the Normandy coast of northern France, along with one plant of *O. sphegodes s.s.* sampled across the Channel in south-east England. Admittedly, these plants are intermingled with plastomes of five other plants: from southern France single plants of *O. provincialis*, *O. argensonensis* (both in mesospecies *Provincialis*), *O. araneola*, and *O. aveyronensis*, plus from central Spain *O. castellana* (all three ascribed to mesospecies *Incubacea*). The possession of this particular plastome by the spectacularly divergent phenotype *O. aveyronensis* is especially striking. The effectiveness of genome skimming for identifying gene flow is indicated by results obtained at the lowest demographic level. Two pairs of plants were sampled from within the same two populations of *O. sphegodes s.s.* in Normandy; plastomes of one pair (3567 and 3569), obtained from a well-established and comparatively extensive population, were very similar but non-identical, whereas the other pair (3528 and 3532), collected from a small, highly localized and possibly only recently established population, were identical.

There is, of course, an inevitable degree of positive correlation between taxonomy and biogeography. For example, the distributions of the morphologically delimited western Mediterranean mesospecies *Incubacea* and eastern Mediterranean mesospecies *Mammosa* are almost mutually exclusive geographically, supposedly overlapping only in Albania and the former Yugoslavia (Delforge, 2016). Also, the inclusion within plastome clade IV of all three microspecies representing the mesospecies *Lunulata* (*O. lunulata s.s.* from Sicily, *O. tarentina* from southern Italy, and *O. melitensis* from Malta) might be taken as evidence of taxonomic clustering. However, their three regions of origin all constitute the same biogeographic realm, allowing equally justifiable claims of geographic rather than taxonomic clustering. If we ignore mesospecies assignment and instead concentrate entirely on geographic region of origin, the plastome tree exhibits stronger clustering according to geography than according to taxonomy (Fig. 4), implying widespread gene flow independent of taxonomic identity. Of the five major clades of plastomes, those within the large clade V and much smaller clade I all originate from east of the Adriatic–Carpathian Divide, clade I being exclusively Turkish. In contrast, most samples resolved in clades II–IV were obtained dominantly west of the Adriatic–Carpathian line, though clade II also contains a single accession from Albania and clade IV contains single accessions from both Albania and southern Greece.

Figure 4 also suggests the existence of three groups of potential interest within clade V, each attracting a BS value of 100%. Within clade V, clade Va consists of a pair of Cypriot plastomes, whereas the six samples of clade Vb all originated from Crete and the Peloponnese, as did the great majority of those forming clade Vc (Fig. 4). Within clade III, a monophyletic group of samples from France, Spain, and England is nested within a somewhat more sequence-divergent grade of samples from Hungary, mainland Italy, and Sardinia.

The unrooted uncorrected P-distances network (Supplementary Fig. S2) also clearly separated outgroups from ingroup, and within the ingroup equally clearly yielded groups I–V. However, relationships among some accessions that were poorly supported in Fig. 4 differed subtly in Supplementary Fig. S2, where two Greek accessions from mesospecies *Mammosa* (3707 and 3708) drifted outside group V.

Genome skimming: ordinations

The strong first principal component of the full plastome matrix (Fig. 5A) clearly separated the 'outer' outgroup, 'inner' outgroup, and ingroup, though the 'outer' outgroup was placed midway between the 'inner' outgroup and the ingroup. The much weaker second coordinate divided the ingroup into two apparently distinct groups. The tighter cluster consists of samples of mesospecies *Sphegodes*, *Provincialis*, and *Incubacea* from geographic areas E (79%) and D (21%); it corresponds to clades II and III in Fig. 4. The larger, more diffuse cluster consisting of most of the remaining samples contains only one sample from geographic area E (2%); it terminates in a tight cluster of most of the plants sampled in area B, and corresponds to clades I, IV, and V in Fig. 4. Midway between the clusters are placed two seemingly intermediate plants: a plant of *O. sphegodes s.s.* from central Italy (basal to, but divergent from, clade III in Fig. 4) and a plant of *O. antiochiana* (mesospecies *Reinholdii*) from Turkey (basal to, but divergent from, clade I in Fig. 4). We speculate that these two plants are misplaced in Fig. 4,

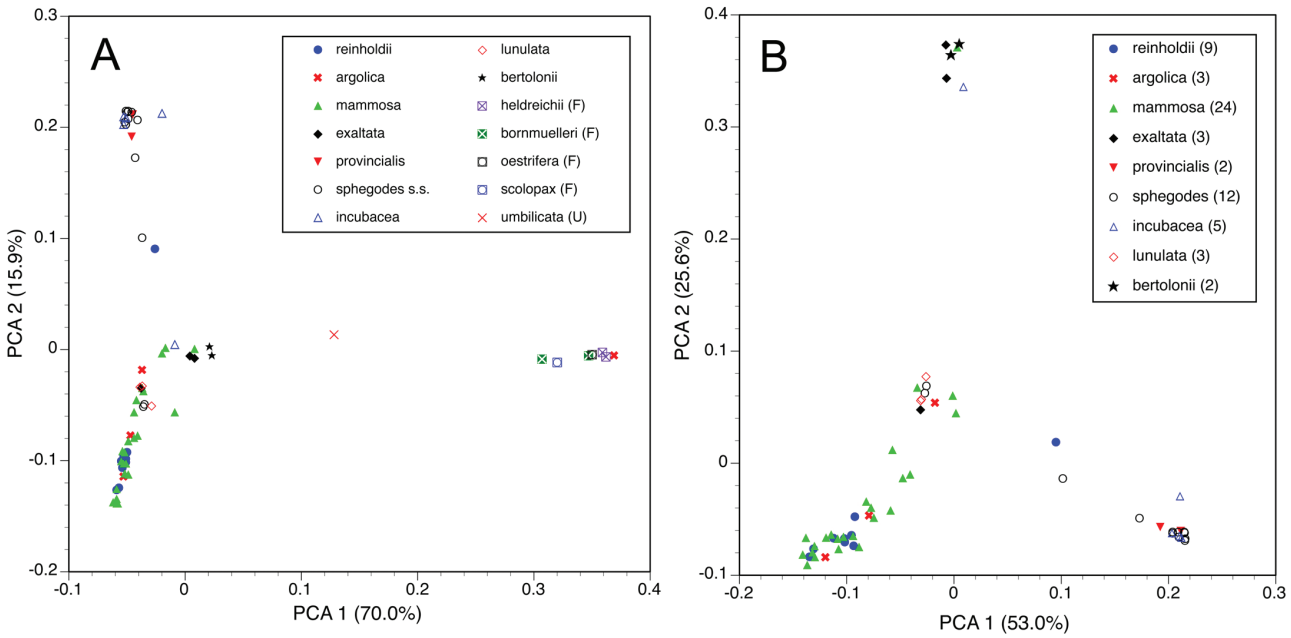

Fig. 5. Principal components ordination of plastome data with plants labelled according to mesospecies assignment. (A) Eight outgroup samples included. (B) Eight outgroup samples excluded. In the key to (A), F indicates outgroups of the *fuciflora* group and U indicates outgroups of the *umbilicata* group.

each having combined elements of a clade III plastome with a member of the larger plastome cluster collectively characterizing clades I, IV, and V.

Reanalysis after eliminating the eight outgroup samples (Fig. 5B) preserved the above clustering. However, the second component became stronger and separated out six taxonomically heterogeneous samples, unified only in lacking any representative from region A; four of the six plants were derived from geographic region C, and correspond to clade II in Fig. 4. All four clusters evident on the PCA contain representatives from region C, where a comparatively high degree of haplotype admixing may have occurred.

The two most structurally disparate plastomes are illustrated in Supplementary Fig. S3. No major length variations or structural rearrangements were detected by us among plastomes within macrospecies *Sphegodes* (cf. Roma *et al.*, 2018; Bertrand *et al.*, 2019). Molecular dating of an ultrametric conversion of the plastome tree shown in Fig. 4 inferred a likely origin of macrospecies *Sphegodes* between 125 ka and 80 ka, though error bars extend the potential date range to between 650 ka and 35 ka; moreover, the conceptual underpinnings of node dating *per se* have recently incurred heavy criticism.

Discussion

Tenuous monophyly of the Sphegodes *clade*

We will now reappraise molecular evidence accumulated during the last two decades of apparent gene flow involving the macrospecies *Sphegodes* at three demographic levels: (i) with the remaining eight macrospecies recognized by Bateman *et al.* (2018); (ii) among *Sphegodes* mesospecies; and (iii) among *Sphegodes* microspecies *sensu* Delforge (2016).

nrITS

The phylogenetically broad ITS phylogeny constructed by Bateman *et al.* (2003) showed the ribotypes of the 14 *Sphegodes* group microspecies analysed by them as collectively being reliably monophyletic. Devey *et al.* (2008) subsequently obtained ITS data from 22 microspecies of the *Sphegodes* group that together encompassed eight of the nine *Sphegodes* mesospecies. When cloned, six of these plants yielded multiple ribotypes (up to eight in the case of *O. garganica*). A single ribotype formed at least half of the total ribotypes found in each mesospecies other than *Sphegodes s.s.* (which was represented by only one plant), and most of the cloned ribotypes deviated from the 'core' ribotype by only a single SNP. It is therefore theoretically possible that all plants in the *Sphegodes* clade possess the same functional ribotype, though two plants of the *Fuciflora* macrospecies were placed within the *Sphegodes* clade of Devey *et al.* (2008), complicating its apparent monophyly. In contrast, no clustering of ribotypes was evident at the lower mesospecies level, and the lack of duplicate samples in this essentially typological study precluded analysis at the microspecies level (cf. Bateman, 2020). Sampling conducted by Tyteca and Baguette (2017), which was considerably richer in both plants and microspecies, similarly revealed very small levels of ITS divergence among accessions of the *Sphegodes* clade; however, it also suggested more extensive intermixing of their ribotypes with those of the closely related *Fuciflora* clade.

Plastid sequences

Compared with nrITS results, the plastid intergenic spacers *trnH–psbA* and *trnD–trnT* sequenced by Devey *et al.* (2008) yielded stronger separation of the *Umbilicata* clade (the 'outer' outgroup of the present study) but offered similarly tenuous separation of the *Sphegodes* clade from the *Fuciflora* clade. The tree of Soliva *et al.* (2001), which combined the plastid region *trnL–F* with nrITS, found the majority of its 18 *Sphegodes* group samples to be monophyletic, but two *Sphegodes* group microspecies (*O. aveyronensis* and *O. delphinensis*) were placed among the five *Fuciflora* group plants analysed. Employing a tree based on six concatenated low-copy nuclear genes, Breitkopf *et al.* (2015) similarly found interdigitated within their five analysed microspecies of the *Fuciflora* clade 18 *Sphegodes* group microspecies, together with the poorly statistically supported macrospecies *Umbilicata*—here employed as an 'outer outgroup' following Devey *et al.* (2008) and Bateman *et al.* (2018).

RAD-seq (NGS)

Aiming to build on previous Sanger-based molecular studies using NGS, Bateman *et al.* (2018) obtained RAD-seq data on 32 accessions that together encompassed all nine of the macrospecies that comprise the genus *Ophrys*. Their trees supported the circumscription of nine macrospecies and strengthened statistical probabilities underpinning relationships inferred among those nine macrospecies. The *Sphegodes* group (sampled for seven microspecies) was strongly supported as sister to the *Fuciflora* group (sampled for nine microspecies), and the two macrospecies were perceived as mutually monophyletic with respective bootstrap values of 97% and 86%. In these features, their data-rich trees, based on 4159 nuclear SNPs, contrasted with the admixtures of the two apparently polyphyletic groups (perhaps more accurately interpreted as a single group) depicted in previous candidate gene-based phylogenies. However, within each of the two macrospecies, branches were short and statistical support was generally weak. In addition, the Bateman *et al.* (2018) tree revealed that a much larger molecular distance separating *Sphegodes* and *Fuciflora* from the macrospecies *Umbilicata* was strongly supported, despite similarities in floral morphology between members of the *Fuciflora* and *Umbilicata* groups (Fig. 1).

It was the RAD-seq study of Bateman *et al.* (2018*a*) that established the experimental parameters for the current study of the *Sphegodes* group by firmly identifying the *Fuciflora* group as the most appropriate 'inner outgroup' and the *Umbilicata* group as the most appropriate 'outer outgroup'. More importantly, it established a higher level topology that could justifiably be employed as a 'relative truth'—or at least a valuable yardstick—by which to measure the phylogenetic accuracy of the whole-plastome sequences that lie at the core of the present study of the *Sphegodes* group.

Seeking circumscribable entities within the plastome data

Macrospecies

Both the ML tree (Fig. 4) and ordinations (Fig. 5) readily distinguish the outgroups of macrospecies *Umbilicata* and *Fuciflora* from the ingroup, macrospecies *Sphegodes*, supporting earlier molecular studies distinguishing these groups. The sole intriguing exception to this rule, *O. delphinensis*, is explicable (see below). Indeed, it usefully demonstrates that, at the macrospecies level, plastomes still have a useful role to play in at least crudely identifying the ovule parent of hybrid plants and perhaps also exploring the deeper history of the maternal lineage.

Otherwise, plastome sequences inform us that our prior field knowledge of the morphology of these plants evidently allowed us to succeed in our goal of eliminating from our sample set F_1 hybrids between macrospecies. Moreover, comparison with the data-rich nuclear SNP tree of Bateman *et al.* (2018*a*) strongly suggests that we can legitimately continue to view plastomes as potentially competent to reconstruct phylogeny at the macrospecies level (and above), though they are inherently less reliable than their nuclear equivalents for inferring species trees.

Extraordinary placement of *Ophrys delphinensis*

Hybrids between members of different *Ophrys* macrospecies can usually be identified with reasonable confidence on the basis of their morphology, though, admittedly, hybrids among the particular three macrospecies analysed in the present study are the most difficult to recognize. Every effort was made by us to exclude macrospecies-level hybrids from the present study, yet one plastome undermines the otherwise perfect mutual monophyly of macrospecies *Sphegodes* and macrospecies *Fuciflora* in Fig. 4. Specifically, the representative plant of the microspecies *O. delphinensis* (mesospecies *Argolica*, macrospecies *Sphegodes*) yielded a plastome that nested deeply within macrospecies *Fuciflora*, emerging as sister to *O. heldreichii* and its segregate *O. schlechteriana* (mesospecies *Heldreichii*, macrospecies *Fuciflora*). A similar placement for *O. delphinensis* was obtained in the candidate-gene tree of Soliva *et al.* (2001), its sequences being identical to those of *O. episcopalis* and an Italian plant of *O.* '*cornuta*' (now, following further taxonomic division, attributed to *O. rhodostephane*: mesospecies *Oestrifera*, macrospecies *Fuciflora*). Most of these microspecies are either confined to, or centred on, Crete and southern Greece (Delforge, 2016).

This placement is an intriguing outcome because *O. delphinensis* was originally described by Danesch and Danesch (1972) as a putative hybrid between *O. argolica* and *O.* '*cornuta*' that had spread to become localized in the region of the Gulf of Corinth in southern Greece; these populations were only later raised to the status of a putatively hybridogenic species, and were ascribed taxonomically to mesospecies

Argolica (macrospecies *Sphegodes*) despite being hypothesized to also have a member of macrospecies *Fuciflora* as their ancestor (e.g. Devillers and Devillers-Terschuren, 1994; Delforge, 2006; Pedersen and Faurholdt, 2007). Although this possibility was recently firmly discounted by Paulus (2018, p. 288), the results of Soliva *et al*. (2001) and the present study are consistent with the theory that *O. delphinensis* arose through hybridization between members of the *Sphegodes* and *Fuciflora* clades. Moreover, by donating its plastome to its progeny, the *Fuciflora* parent revealed itself to be the 'mother' (ovule parent). Artificial crosses made between the two suspected parental microspecies, *O. argolica* and *O. oestrifera*, have yielded phenotypes strikingly similar to those exhibited by *O. delphinensis* (S. Malmgren, personal communication, 2008).

Populations and microspecies

If we apply the spotlight to the opposite end of the demographic hierarchy from macrospecies, the comparatively small number of samples analysed from the same population of the same microspecies reassuringly yielded very similar plastomes, as did the linear suite of four populations of *O. sphegodes s.s.* sampled along the Normandy coast of France (near-identical within clade III of Fig. 4). In terms of assessing the precision of our sequencing through genome skimming, it is even more heartening that the pair of *O. sphegodes s.s.* samples that are most likely to represent a very recently established population (and so are likely to have experienced a genetic bottleneck) yielded identical plastomes. However, any belief that the plastomes could be population specific or microspecies specific is belied by the fact that plants of the same microspecies are always either intermixed with, or closely similar to, plants of other microspecies. Admittedly, seven of the nine multiply sampled microspecies are each confined to just one of the five plastome clades within the *Sphegodes* group (labelled I–V in Fig. 4), but the majority are comparatively divergent within their chosen clade. On average, proximities of samples of the same microspecies appear closer on the ordination of ingroup plastomes (Fig. 5B), but again, they are reliably intermingled with other microspecies.

In summary, although evidence suggests that our plastome sequencing is accurate, we are nonetheless unable to resolve either populations or especially microspecies as anything close to monophyletic (*ergo*, nor can we reliably identify them) on the basis of their plastomes. The most positive conclusion to be drawn is that, in at least some cases, it may be possible to use plastomes to narrow down the range of possible taxonomic assignments of an 'unknown' plant.

Mesospecies

The demographic level where the plastomes might have been expected (or at least hoped) to achieve greater resolution and taxonomic clarity within macrospecies *Sphegodes* is the nine mesospecies established on morphological grounds by Delforge (2016). If the macrospecies *Sphegodes* genuinely contains within it self-circumscribing aggregates of populations that might themselves have credible claims for species status, and plastids accurately represent species trees, we would anticipate from first principles that sufficient sequence divergence in plastomes would have occurred to yield entities that are at least tentatively monophyletic and show both phenotypic and geographic cohesion. This is clearly not the case; the average mesospecies spans 1.7 of the five main plastome clades in Fig. 4 and is interdigitated phylogenetically with representatives of 4.4 other mesospecies (figures for overlap between mesospecies range from 2 for mesospecies *Reinholdii* to the theoretical maximum of 8 for the comparatively well-sampled mesospecies *Mammosa*). Moreover, had we sampled more extensively, these disappointingly high figures for intermixing could only have increased further.

The challenge of interpreting these results taxonomically is further compounded by short, weakly supported branches within much of clades Vc and III. In the case of clade III, the lack of resolution and extensive interdigitation of representatives of three mesospecies within the clade argue particularly strongly against separating the *Provincialis* and *Incubacea* mesospecies from mesospecies *Sphegodes*, a conclusion reinforced by the tight cluster formed by these taxa in the corresponding ordination (Fig. 5B).

We can only conclude that, on present evidence, none of the nine S*phegodes* mesospecies recognized by Delforge (2016)—the primary demographic level addressed by this study—even approximates monophyly, and that they can be neither circumscribed nor identified on the basis of their plastome sequences. We will return later to consider the implications of this case study for the broader disciplines of plant systematics and plant phylogenetics.

Biogeography of Sphegodes *plastomes*

The most confident prediction that can be made from our whole-plastome sequences (with ~95% confidence) is that a plant of the macrospecies *Sphegodes* from west of the Carpathian Divide will have a plastome of clades II, III, or IV, whereas a plant from east of the Divide will have a plastome of clade I or, more probably, clade V (Fig. 4). In this feature, macrospecies *Sphegodes* parallels macrospecies *Fuciflora*, which was demonstrated via candidate gene studies to show a similar genetic threshold approximating the Carpathians (Devey *et al*., 2009; G. Sramkó *et al*., unpublished). However, all four clusters evident on the ordination of plastome data lacking the outgroups (Fig. 5B) contain representatives from geographic region C, immediately north and west of the Adriatic–Carpathian Divide, where we suspect that a greater degree of haplotype admixing may have occurred. It is tempting to view the three main clusters in Fig. 5B as representing the derivatives of the three main post-glacial migration routes northward via Morocco, Tunisia, and the Levant (cf. Ferris *et al*., 1999; Widmer and Lexer, 2001; Médail and Diadema, 2009).

Within plastome clade III, a monophyletic group of molecularly near-identical samples from France, Spain, and England is nested within a somewhat more sequence-divergent grade of samples from Hungary, Italy, and Sardinia (Fig. 4). This topology is consistent with a hypothesis of comparatively recent migration toward the western and north-western margins of Europe of plants bearing this haplotype category. Similar patterns have been detected within other Eurasian orchid genera. In addition, population-level sequencing of nrITS also demonstrated that macrospecies *Sphegodes* mirrors macrospecies *Fuciflora* in bizarrely maintaining greater ribotype diversity in England, where it is rare and at the absolute margin of its overall distribution, than in adjacent areas of continental Europe, where it is far more frequent (cf. Devey *et al.*, 2009).

Seeking circumscribable entities within the morphometric data

Macrospecies

The outgroup macrospecies were included only in the reduced data matrix ordinated as Fig. 3. Although the second coordinate separated all three outgroup members from the ingroup, they could be viewed as part of a morphological continuum, thereby contrasting with the clear morphological discontinuities evident in the tree (Fig. 4) and ordinations (Fig. 5) derived from the plastome data.

Populations and microspecies

Differences in ontogenetic status ('vigour' *s.l.*) ensured that most plants of the same microspecies measured in the same local population did not plot as close together on the morphometric ordinations as they did on the plastome ordinations. Plants of the same microspecies measured in different populations were often considerable distances apart on the plots (Figs 3, 4), presumably because ontogenetic and ecophenotypic influences on phenotype were operating in addition to genetic differences.

Mesospecies

The most effective method of exploring degrees of morphological differentiation at the mesospecies level is to compare the areal extent of individual mesospecies on the plot of the first two principal coordinates relative to the areal extent that the nine mesospecies encompass in aggregate (Fig. 2A). Unsurprisingly, the number of plants measured per mesospecies has a strong influence on perceived morphological diversity. Among the six mesospecies that yielded statistically acceptable sample sizes, the analysed individuals of mesospecies *Bertolonii* (n=7) form a tight cluster that occupies only 7% of the total morphospace. Individuals of mesospecies *Sphegodes* are also comparatively tightly clustered (15% of morphospace, n=20), perhaps partly reflecting the fact that they represent only a single microspecies, whereas the similarly sized samples of *Incubacea* (40%, n=14), *Reinholdii* (33%, n=17), and *Argolica* (36%, n=19) each exceed one-third of the total morphospace. However, most remarkably, the most intensively sampled mesospecies, *Mammosa*, encompassed almost the entire total morphospace summarized in the plot (94%, n=41), failing only to overlap the bulk of the region of the plot that is occupied by mesospecies *Sphegodes*. The perceived exceptional morphological diversity of mesospecies *Mammosa* was presumably aided by the fact that we analysed samples of as many as 13 of its microspecies—nine more than for any other mesospecies.

There are few areas of Fig. 2 where less than three mesospecies overlap morphologically, and two areas where as many as six of the nine sampled mesospecies overlap. If one measured the 51 characters on a random plant of macrospecies *Sphegodes* and re-ran the analysis to include that plant, the likelihood of 'correct' assignment of the 'unknown' plant to the most appropriate mesospecies would therefore be low—arguably even lower than the probability of success offered by obtaining a corresponding whole-plastome sequence. This morphological ambiguity is perhaps understandable, as the weak differentiation of genes determining morphology is further compounded by ontogenetic and ecophenotypic impacts on epigenetic expression of the underlying genetic 'programming.' Nonetheless, the morphometric analyses give a clear impression that the nine mesospecies are no more than arbitrary subdivisions of what is actually a morphological continuum—a continuum that encompasses the whole of macrospecies *Sphegodes*. Even the influence of geography appears weak; it is unclear whether the concentration of somewhat smaller western plants toward the negative end of the second coordinate (Fig. 2B) simply reflects the preponderance there of mesospecies *Sphegodes* and *Incubacea*, or even the preference of mesospecies *Sphegodes* for exposed open habitats that tend to reduce average plant size.

Overall, these results demonstrate conclusively that morphological variation within macrospecies *Sphegodes* is massively multidimensional and that any distinctions competent to induce unequivocal clustering were too subtle to be captured by our character suite, despite the large number of characters used to describe each plant in its entirety.

Correlation of phenotype and genotype

We further explored the relationship between phenotype and genotype by conducting additional analyses on the 24 samples (21 plants of macrospecies *Sphegodes*, two plants of macrospecies *Fuciflora*, and one plant of macrospecies *Umbilicata*) that yielded both plastome sequences and morphometric data (Fig. 3). As expected, when plotted along the SNP axis the pairwise comparisons yielded three clusters, depending on whether the comparison involved the 'outer outgroup' (left), the 'inner outgroup' (central), or neither (right) (Fig. 6A). In contrast, morphometric similarities failed to show any discontinuities, and there is considerable overlap between the clusters circumscribed by the plastomes. The 'inner outgroup' versus ingroup cluster has a skewed tail of especially low morphometric

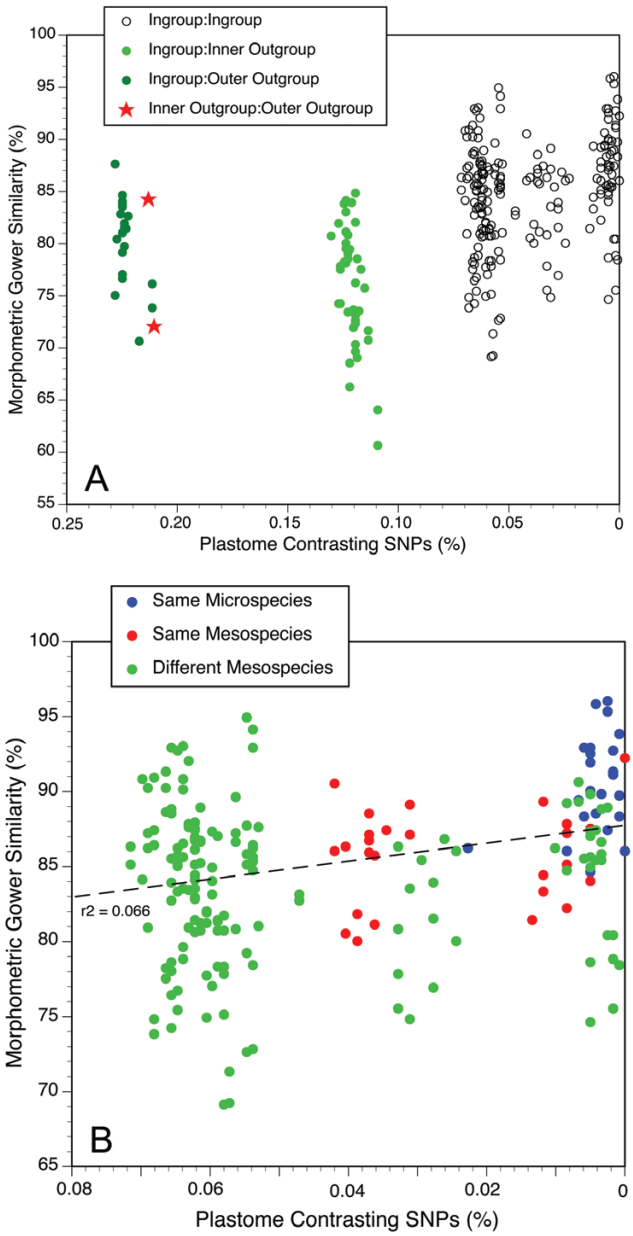

Fig. 6. (A) Comparison of genotypic and phenotypic distances for 21 plants belonging to the macrospecies *Sphegodes* plus two belonging to the macrospecies *Fuciflora* and one to the macrospecies *Umbilicata*. (B) Magnification of the top right portion of (A) to show relationships between the six analysed mesospecies of macrospecies *Sphegodes*.

similarities whereas, as expected, the ingroup versus ingroup comparisons are reliably capable of achieving greater morphometric similarities.

Focusing on relationships among ingroup members (Fig. 6B) suggests the presence of three less cohesive plastome groups, each exhibiting a continuum of morphometric similarities, the left-hand and right-hand clusters differing by only 3% in mean morphometric similarity. In other words, although positive, the correlation between similarity of genotype and phenotype is extremely poor. The more molecularly divergent cluster consists only of pairwise comparisons between mesospecies, but such comparisons are distributed among all three clusters. In contrast, all but one of the pairwise comparisons between members of the same microspecies are confined not only to the least molecularly divergent cluster but also to the upper half of the range of morphometric similarities evident within that plastome cluster.

Unfortunately, the close geographic juxtaposition of each pair of samples that belong to the same microspecies means that we cannot determine whether this strong similarity in phenotype and especially genotype reflects potentially meaningful low-level taxonomy or merely geographic proximity aiding plastid capture. This is an issue that is better explored through co-occurring populations of different microspecies; we now briefly summarize and reappraise one such study.

Comparison with a more taxonomically focused study within macrospecies Sphegodes

Our study applied genome skimming and character-based morphometrics to, respectively, 40 and 34 microspecies of macrospecies *Sphegodes*, plus two outgroup macrospecies, sampled widely across the entire distribution of the macrospecies. In contrast, Sedeek *et al.* (2014; genetic data reappraised by Cozzolino *et al.*, 2020) operated at a much lower demographic level, intensively sampling just four *Sphegodes* group microspecies (*O. sphegodes s.s.*, mesospecies *Sphegodes*; *O. exaltata*, mesospecies *Exaltata*; and *O. incubacea* and *O. garganica*, both mesospecies *Incubacea*) in populations collectively confined to five localities along a 60 km west–east transect through the Gargano Peninsula of east-central Italy. Each local population included in their study maintained in sympatry between two and four of the microspecies under investigation, and a wide range of analytical approaches were deployed.

Of the non-genetic properties explored by Sedeek *et al.* (2014), peak spring flowering differed by no more than a week among the four microspecies. Insect-mediated pollination frequencies were remarkably low (averaging 4%) but most pollen deposition events did usually involve pollinaria of the 'correct' donor microspecies. Artificial crosses between microspecies revealed negligible post-mating isolation; any barriers among microspecies were predominantly pre-mating, as predicted from previous studies (e.g. Scopece *et al.*, 2007). Differences in flower size and shape (determined primarily by three-dimensional scanning) were modest, their apparent significance being exaggerated by presentation of the data as canonical variates plots—an algorithm explicitly designed to maximize apparent distances between organisms allocated *a priori* to expected categories (in this case, to the four microspecies). Microspecies also differed little in flower colour and labellar markings. The biochemistries of floral odour cocktails differed mainly quantitatively rather than qualitatively,

but nonetheless the observed differences were regarded by Sedeek *et al.* (2014) as the primary (and earliest, most rapidly evolving) factor determining pollinator preference, thus reinforcing similar conclusions drawn from several earlier studies (e.g. Ayasse *et al.*, 2000, 2010; Mant *et al.*, 2005; Schiestl and Cozzolino, 2008; Stökl *et al.*, 2008; Xu *et al.*, 2011).

With regard to the genetic data that are of particular interest in the present context, the initial genotyping by sequencing (GBS) analysis (Elshire *et al.*, 2011) conducted by Sedeek *et al.* (2014) included 127 individuals, though the subsequent experimental re-analysis of the molecular data by Cozzolino *et al.* (2020) reduced this number to a subset of 54 plants that had yielded the most complete molecular data. No outgroups were included in either analysis, leading to questionable rooting of the trees generated by Cozzolino *et al.* (2020). Excessive variation in sequence coverage inevitably led to extensive missing values in Structure files, such that 'best *K* estimates were highly inconsistent among analyses and showed no apparent patterns.' This observation led Sedeek *et al.* (2014, p. 6198) to draw the understandable conclusion that 'although genetic structure was broadly consistent with [micro]species groups, [micro]species were very similar and only weakly differentiated.' Indeed, 'only a very small portion of the genome (<0.05%) … is interpreted as being associated with [micro]species divergence' (Sedeek *et al.*, 2014, p. 6201). We will return later to address the key question of whether *Sphegodes* microspecies are indeed actively 'diverging.'

The subsequent, more nuanced analysis of the same GBS data by Cozzolino *et al.* (2020) usefully distinguished among probable nuclear, plastid, and mitochondrial SNPs. It vividly demonstrated the profound impact on inferences of sample relationships made by filtering the data at contrasting percentages of shared loci. An analysis based on loci shared by at least 30% of the individuals analysed (59 435 informative SNPs) rendered three of the four microspecies tentatively monophyletic, whereas an analysis based on a far less patchy matrix confined to at least 70% shared loci and selectively filtering for homozygous/organellar loci (thus reducing informative SNPs to just 253) became dictated by plastid haplotypes that did not equate with assignment of individuals to microspecies. The resulting topology collapsed perceptions of not only monophyletic microspecies but also monophyletic mesospecies, given that the four microspecies collectively represent three mesospecies *sensu* Delforge (2016).

This example elegantly illustrates the metastability of not only *Ophrys* microspecies but also NGS matrices, and emphasizes the crucial effects of the many filters typically applied to NGS data prior to generating the trees and/or ordinations that routinely constitute the basis of scientific interpretation. The thought-provoking microspecies-focused studies of Sedeek *et al.* (2014) and Cozzolino *et al.* (2020) are summarized here because they provide the ideal companion to the present, taxonomically broader mesospecies-focused investigation.

Macrospecies trump mesospecies

In practice, if the taxonomy of genus *Ophrys* is ever to stabilize at anything close to an optimal outcome, only one of the three demographic levels of macrospecies, mesospecies, or microspecies (or perhaps, some as yet unspecified alternative concept) can ultimately be chosen as reflecting the preferred species concept (Bateman, 2018, 2021).

Macrospecies

Here, the macrospecies level in *Ophrys* is summarized simply as nine monophyletic entities that can be confidently identified using morphology by even inexperienced field botanists and that give rise to hybrids that can be identified with justifiable confidence by experienced field botanists. Both the macrospecies and their hybrids can similarly be readily identified through candidate gene sequencing of a few well-chosen, readily analysed genic regions such as nrITS, and each of the nine entities is geographically widespread, spanning the majority of the geographic region occupied by the genus as a whole. They are biologically cohesive and immune to accusations of transient existence; they are species with a guaranteed future.

Mesospecies

The status of the mesospecies level cannot be summarized so easily or simply. As delimited by specialists such as Devillers and Devillers-Terschuren (1994), Pedersen and Faurholdt (2007), and Delforge (2016), several of the 19–26 mesospecies recognized as representing an intermediate level within these taxonomies fail the test of monophyly because they combine microspecies of two or more macrospecies. In the case of Delforge (2016), both mesospecies *Bornmuelleri* and mesospecies *Heldreichii* admix microspecies belonging to both of the readily molecularly differentiable macrospecies *Fuciflora* and *Umbilicata* (Fig. 4). Assigning at least some unknown plants to mesospecies categories on a morphological basis is challenging even for orchid experts, and 'hybrids' between these categories cannot be identified with confidence because the majority of mesospecies are too similar in morphology to any inter-mesospecies progeny. By definition, in a phenotypic continuum, 'hybrids' between adjacent 'species' cannot be identified using morphology (Bateman, 2020, 2021). DNA-based assignments require NGS of the nucleus rather than the more basic candidate gene sequencing and, even given such data, they are not wholly reliable. In addition, some microspecies placed in different mesospecies have proven to be more similar, both phenotypically and genotypically, than microspecies placed in the same mesospecies. For example, Sedeek *et al.* (2014) showed that, in the Gargano Peninsula, *O. incubacea* and *O. garganica* (both assigned to mesospecies *Incubacea*) were less similar in every measured category of data than *O. incubacea* was relative to *O. sphegodes s.s.* (mesospecies *Sphegodes*); indeed, the

supposed genetic separation of *O. incubacea* from *O. sphegodes s.s.* failed to survive intact a Structure analysis.

Biogeography within mesospecies *Argolica*

The biogeographic clustering of plastomes is well illustrated in Fig. 4 by the four analysed samples of mesospecies *Argolica*, each of which represents a highly geographically restricted microspecies. *Ophrys argolica s.s.*, sampled in southern Greece, occurs in plastome clade Vc alongside other mesospecies from southern Greece and Crete. *Ophrys lycia*, an exceptionally rare microspecies from south-west Turkey, occurs in clade V alongside other mesospecies from Turkey and Cyprus. *Ophrys morisii*, sampled in Corsica, is placed in clade IV alongside other mesospecies from the central Mediterranean region around Italy. Also, *O. delphinensis* is placed with the outgroup macrospecies *Fuciflora*, reflecting its true nature as a recent hybrid between members of two macrospecies.

Each of the 13 microspecies of mesospecies *Argolica* recognized by Delforge (2016) has an extremely limited distribution within the Mediterranean Basin; none exceeds 300 km in diameter. As mapped by Hennecke and Münzinger (2014, fig. 1), only *O. argolica s.s.* and *O. delphinensis* have partially overlapping distributions, an observation that is easily explained by *O. delphinensis* actually having a hybrid origin between *O. argolica s.s.* and a member of macrospecies *Fuciflora*. Even more striking is the fact that the majority of the remaining microspecies in the *Argolica* group do not even have adjacent distributions; rather, they are depicted as distributional 'islands' scattered across the central and eastern Mediterranean. Pedersen and Faurholdt (2007) accused *O. morisii* of hybridity, while Hennecke and Münzinger (2014) argued that all of the microspecies in mesospecies *Argolica* other than *O. argolica s.s.* have a hybridogenic origin—one that, in each case, involved *O. argolica s.s.* However, for this hypothesis to be true, *O. argolica s.s.* would have once had to be widely distributed across the Mediterranean Basin before retreating to southern Greece, leaving in its wake only scattered hybridogenic populations separated by zones where mesospecies *Argolica* is absent, even though those same zones readily support other mesospecies of the *Sphegodes* group. In the context of the young, radiating genus that *Ophrys* is widely envisaged to represent (Soliva *et al.*, 2001; Paulus, 2006; Stökl, 2007; Ayasse *et al.*, 2010; Breitkopf *et al.*, 2015; Pineiro-Fernandez *et al.*, 2019; Baguette *et al.*, 2020), this scenario appears highly improbable.

A more likely scenario is that mesospecies *Argolica* lacks biological cohesion and is instead an artificial construct of authoritarian taxonomy. It is more probable that populations possessing the supposed diagnostic features of mesospecies *Argolica* (broad stigmatic cavity, more or less circular labellar outline, speculum often more or less detached from the stigmatic area) emerged several times at different locations within the overall distribution of macrospecies *Sphegodes*. The iterative origins of broadly similar phenotypes could reflect either drift in small, potentially hybridogenic populations, or selectively driven adaptation to specific pollinating insects (reputedly different species of the bee genus *Anthophora*). Selective origins are made less likely by the absence of sympatry among the microspecies of the *Argolica* group, since facilitating sympatry with potential competitors is the usual explanation given for such supposed fidelity to a single species of pollinating insect (Cortis *et al.*, 2009; Gögler *et al.*, 2009, 2015; Vereecken *et al.*, 2010; Sedeek *et al.*, 2014). Are these *Argolica* microspecies really bona fide species, or are they merely recently originated and potentially transient local morphs? Does each supposedly subtly distinct morph simply feature the plastome that is characteristic of all *Sphegodes* group members in its particular geographic region, rather than showing clear evidence of genetic isolation?

Viewed objectively, the mesospecies are clearly categories of convenience, but unlike the macrospecies and microspecies, there is no particular underlying species concept available to justify their recognition. Thus, the well-informed battleground for conceptual supremacy remains that between the two extremes, macrospecies and microspecies. Despite much effort, we have been unable to identity the conceptual underpinnings needed to support the many protagonists in the debate (including ourselves!) who desired the compromise solution of an accepted number of species in the genus *Ophrys* that was greater than nine but far fewer than 350+ (e.g. Pedersen and Faurholdt, 2007; Vereecken *et al.*, 2011; Tyteca and Baguette, 2017; Fateryga *et al.*, 2018; Kühn *et al.*, 2019).

Macrospecies versus microspecies: two radically contrasting evolutionary models

All of the research applied to this remarkable orchid genus can ultimately be boiled down to two models that, despite much debate, remain radically different in both evolutionary and taxonomic interpretation (cf. Bateman *et al.*, 2011; Vereecken *et al.*, 2011): the ethological model favours microspecies whereas the intrinsic discontinuity model favours macrospecies.

Areas of broad agreement

Proponents of both models have increasingly accepted that both genotypic and phenotypic differences among closely related microspecies are extremely limited, and that any genuine differences are most likely to impact strongly on successful pollination via the remarkable pseudo-copulatory pollination system about which so much has been written. Both models recognize fragrance—specifically, the precise biochemical composition of pseudo-pheromone cocktails, particular regarding relative proportions of various alkenes—as probably the most important of the three categories of cue that help to attract pollinators (e.g. Schiestl *et al.*, 1999; Schlüter *et al.*, 2011; Sedeek *et al.*, 2013, 2016). Although complex and seemingly well adapted, the shape and texture of the flower, and its range

and spatial distribution of colours, are of secondary importance (cf. Vereecken and Schiestl, 2009; Streinzer *et al.*, 2010; Rakosy *et al.*, 2012). Indeed, self-pollinated flowers with anthocyanin-rich pink sepals can yield some progeny with anthocyanin-less green sepals, and vice versa (Malmgren, 2008), helping to explain the frequent polymorphism in sepal colour that is evident in Fig. 1 and is a major contributor to Fig. 2A. It is therefore unsurprising that recent studies of putative speciation events in *Ophrys* have emphasized any small genetic differences likely to influence pseudo-pheromone composition (e.g. Sedeek *et al.*, 2013, 2014, 2016). Also, both models agree that, despite the appealing sophistication of the many floral adaptations to pseudo-copulatory pollination, the mechanism is strikingly inefficient compared with other orchid species that operate through food deception and especially through nectar reward (reviewed by Cozzolino and Widmer, 2005; Claessens and Kleynen, 2011). Setting aside the routinely autogamous macrospecies *Apifera*, half of the 102 pollination studies of *Ophrys* microspecies summarized by Claessens and Kleynen (2011, appendix 2) yielded frequencies of <10% successful fertilization. This appears to us a high price to pay for a mechanism whose reputed strength is increasing the probability of avoiding self-pollination through geitonogamy—a questionable evolutionary goal in any case, given that most terrestrial orchids remain highly successful despite routinely experiencing high levels of geitonogamy (e.g. Maad and Reinhammar, 2004; Kropf and Renner, 2008; Sramkó *et al.*, 2019; Bateman, 2020, 2021).

The two opposing perspectives on *Ophrys* speciation differ primarily in four main criteria: (i) the amount of gene flow that is believed to occur between *Ophrys* microspecies; (ii) the minimum amount of gene flow that is regarded as acceptable between bona fide species; (iii) the degree of pollinator fidelity enjoyed by a typical *Ophrys* microspecies; and (iv) the degree to which pollinator fidelity, whatever its actual degree, has been fine-tuned into the *Ophrys* microspecies by strong directional or disruptive selection reflecting pollinator choice.

Ethological model

The 'desktop radiation' of *Ophrys* into not tens but rather hundreds of microspecies that began in the 1990s (Delforge, 1994; Devillers and Devillers-Terschuren, 1994) was prompted by targeted exploration of the (in)famous pseudo-copulatory pollination mechanism, which bizarrely allowed the putatively dominant species of insect pollinator to transcend traditional morphology as the primary criterion for species recognition (cf. Bateman *et al.*, 2011; Vereecken *et al.*, 2011; Bateman, 2018; Paulus, 2018). The subtlest of perceived morphological differences were then advanced as secondary justifications of species-level distinction already awarded to the populations in question.

The ethological perspective views the *Ophrys* microspecies as bona fide species generated through what has been widely termed ecological speciation, typically defined as 'the process by which ecologically based divergent selection between different environments leads to the creation of reproductive barriers between populations.' The inefficiency of the high-risk pseudo-copulatory pollination mechanism is considered to induce strong competition within the population. At least two flowers must make physical contact with an insect before cross-pollination is possible; the first contact allows flower 1 to deposit the pollinaria on the visiting insect and the second—considerably rarer—contact allows flower 2 to recapture the pollinaria from that insect. Thus, the probability of successful pollination will be enhanced if the orchid can increase the period elapsed before the visiting naïve male bee learns to avoid the 'false female' represented by the flower (e.g. Schiestl, 2005; Paulus, 2006; Vereecken and Schiestl, 2009). This phenomenon is hypothesized to encourage negative frequency-dependent selection; within a particular population, rarer phenotypes are more likely to induce a slower learning response in the prospective pollinator than do more common phenotypes. The rarer phenotypes are thus selectively favoured, thereby increasing in relative (and presumably often absolute) numbers within the population.

However, this theory could equally be viewed as a recipe for constraining rather than encouraging variation in pseudo-pheromone composition, because it implies that as the novel phenotype increases in frequency it will be decreasingly selected for within the population—a classic negative feedback loop. The only potential opportunity for speciation within this scenario is if the phenotype shifts sufficiently strongly to become substantially more attractive to one or more novel pollinators at the expense of its ancestral pollinator(s), or if the novel morph rapidly migrates to an area that lacks the ancestral orchid morph(s) but nonetheless possesses suitable pollinator(s). As described by Baguette *et al.* (2020): p. 1659, frequency-dependent 'shifts in pollinator species are due to the random crossing of peaks in the olfactory landscape of the pollinator guild that is syntopic [non-competitively sympatric] to each particular *Ophrys* population. This selective process on individual, random variation in pseudo-pheromone bouquets is followed by directional selection on flower phenotypes [presumably meaning visual and textural cues, given that the olfactory cue has already been modified] that will reinforce the attraction of the new pollinator.' This hypothesis necessitates directional selection through numerous generations of a strength that is sufficient to over-ride the effects of incoming gene flow and/or within-population drift. Also relevant are the many other stochastic effects that are likely to be enhanced by the typically small effective populations that characterize *Ophrys* microspecies and the frequent disturbance that characterizes many of their preferred habitats. Indeed, to be effective, any modest divergence achieved through selection would also probably need to be enhanced through 'divergence hitchhiking' (i.e. neutral genes increase in frequency simply as a result of their close proximity to a strongly selected gene; Feder *et al.*, 2012).

Admittedly, there exist two camps among proponents of pollinator-mediated ecological speciation. Hard-core ethologists persist in arguing for a perfect relationship between a single *Ophrys* microspecies and typically a single dedicated pollinator species that, more than any other single criterion, dictates species identity. Secondary pollinators are considered both uncommon and woefully ineffective (e.g. Paulus, 2018; Baguette *et al*., 2020). This stance is often accompanied by the argument that reports of gene flow among *Ophrys* microspecies have been greatly exaggerated, despite ample evidence gathered to the contrary by many authors (Devey *et al*., 2008; Stökl *et al*., 2008; Cortis *et al*., 2009; Vereecken *et al*., 2010; Xu *et al*., 2011). In contrast, soft-core ethologists are prepared to allow a meaningful role for secondary pollinators—a position supported by the stronger pieces of ethological evidence currently available. For example, Sedeek *et al*. (2014) reported huge annual fluctuations in pollinator success for *Sphegodes* group microspecies, and Breitkopf *et al*. (2013) documented multiple pollinators operating at single sites for both of their studied *Sphegodes* group microspecies, as well as noting contrasting dominant pollinating bee species of those microspecies at sites on the west and east coasts of central Italy. As described by Breitkopf *et al*. (2013, p. 2198), 'local selection on plants imposed by a variable geographical and temporal mosaic of potential pollinators could lead to multiple pollination ecotypes.' We agree that this is a credible scenario, though we are also inclined to ascribe modest local divergences in phenotype to drift and inward gene flow (e.g. Tremblay *et al*., 2005), enhanced by various epigenetic phenomena (e.g. Paun *et al*., 2010; Hirsch *et al*., 2012; Bateman *et al*., 2018*b*). We therefore regret that the relative contributions of these process to population-level differentiation within and among *Ophrys* microspecies remain unexplored.

Intrinsic discontinuity

We retain our long-held concerns, both theoretical and practical, regarding the validity of the assumption-laden ethological species concept. We seek species that are essentially self-circumscribing through being separated from all other species by discontinuities in intrinsic properties of the plants and their populations, rather than the extrinsic and potentially transient property of pollinator identity. Moreover, those discontinuities need to have stood the test of time, and should in addition be sufficiently practical to enable reliable field identification. In practice, these stipulations require a phenotypic discontinuity that is demonstrably underpinned by a genotypic discontinuity. Until the presence of such a discontinuity has been thoroughly demonstrated, a putative species remains no more than an untested hypothesis (Bateman, 2001, 2020).

When our results are viewed collectively, it is difficult to perceive the myriad *Ophrys* microspecies as anything other than arbitrary divisions of not only a morphological continuum (Fig. 2) but also a genetic continuum (Fig. 4; Bateman *et al*., 2018*a*). Sedeek *et al*. (2014, p. 6202) eloquently described the microspecies as 'showing genic rather than genome-wide differences', but the genic differences that they detected are at best both remarkably few and remarkably modest. Also, the occurrence of such genic synapomorphies within microspecies has not yet been demonstrated beyond a few geographically restricted populations (unfortunately, even effective, well-integrated population-level studies are inevitably constrained in sampling strategies as a result of being resource intensive: cf. Stökl *et al*., 2008; Xu *et al*., 2011; Breitkopf *et al*., 2013; Sedeek *et al*., 2014; Gögler *et al*., 2016).

How much morphological differentiation is sufficient?

The microspecies are not viable from a practical perspective. Insect visits to *Ophrys* flowers are few, and successful pollinations are rarely observed in nature, precluding their use as an identification tool. Similarly, the subtle differences in pseudo-pheromone cocktails that are hypothesized to be the strongest intrinsic property relating microspecies to preferred pollinator(s) would require a sophisticated mass spectrometer if they are to be quantified—currently an unlikely piece of equipment for field use.

We do not consider it possible to rely on the 'secondary sex' characteristics of flower shape, colour, and texture to identify microspecies. Quite apart from the discouraging morphometric results documented here, even the experts who have established these microspecies encounter difficulties when attempting to identify them in the field. The plate of same-scale illustrations presented here as Fig. 7 has been shown to experienced European orchid specialists as representing a range of plants encountered across Continental Europe. Once given this misinformation, they typically assigned these plants to not only between three and five microspecies but also to multiple mesospecies. In truth, all 12 images were captured in just three typical *Ophrys* populations distributed along the southern coast of England, a country universally acknowledged by European orchid specialists as supporting only one (rare) native microspecies of macrospecies *Sphegodes*, specifically *O. sphegodes s.s.* The human mind is inherently better attuned to seeking differences than to seeking similarities.

In addition, the four flowers of *O. argolica s.s.* (mesospecies *Argolica*, macrospecies *Sphegodes*) shown in Fig. 8, which exhibit subtle morphological differences equivalent to those that distinguish closely related microspecies, were in fact grown from seed extracted from a single artificially self-pollinated capsule (Malmgren, 2008). Given that the four plants were grown under identical controlled conditions, thus precluding ecophenotypic variation, all of the morphological variation evident among these flowers can only be attributed to epigenetic phenomena plus any subtle genetic rearrangements incurred during meiosis. We conclude that no SNP-based differences are required to generate microspecies-level morphological variation.

Given these observations, how is the field botanist expected to successfully distinguish subtle morphological differences

Fig. 7. Representative plants photographed in three populations of a single microspecies, *Ophrys sphegodes s.s.*, located along the south coast of England. These plants collectively show sufficient morphological variation for attribution to several microspecies recognized in continental Europe. The horizontal dimension of each image is 22 mm. Images: Richard Bateman.

that reflect underlying genetic differences from those merely reflecting epigenetics, ontogeny, and/or ecophenotypy? Also, accepting that there will be at least minimal genetic differences detectable between any two populations, how much of a difference is needed to declare them separate species?

We will now consider whether, having failed at a practical level, evolutionary biology can rescue the microspecies at a conceptual level.

Placing Ophrys *in the broader context of species concepts*

'Incipient species': life between a rock and a hard place

Of particular interest to us was the repeated assumption by several other authors of having caught *Sphegodes* microspecies (and indeed, by our estimate, 97% of all *Ophrys* microspecies) midway through a divergence process; one that will continue

Fig. 8. Four plants of *O. argolica* (mesospecies *Argolica*) grown by Svante Malmgren from seed collected from a single artificially self-pollinated flower. Although genetically identical and grown in the same controlled conditions, the flowers differ appreciably in size, shape, and labellar patterning. Scale=approximately 10 mm. Image: Svante Malmgren.

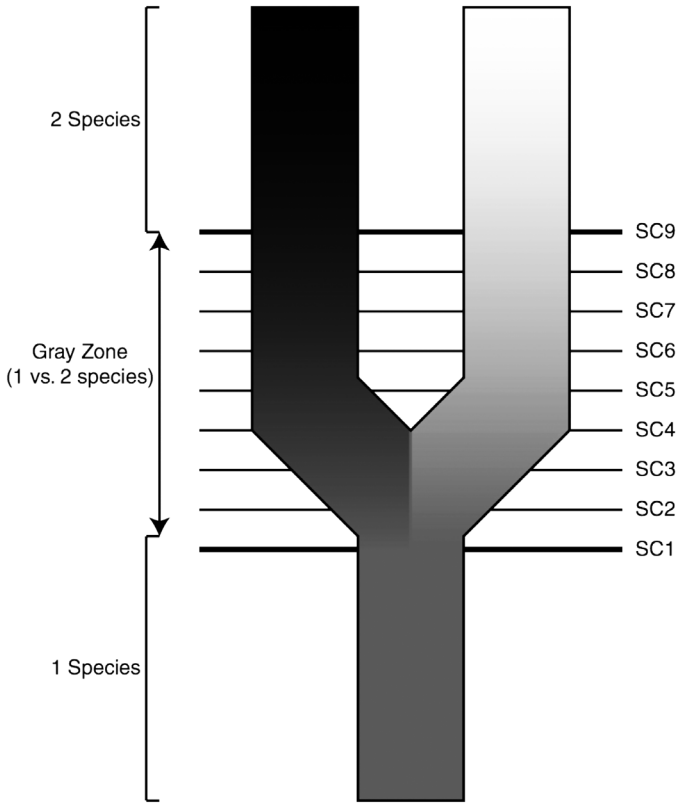

Fig. 9. Representation of the classic 'tuning fork' model of speciation, depicted as gradual but irreversible divergence. The 'grey zone' of ambiguity encompasses a wide range of conceptual, evidence-based thresholds ('species criteria', SC), any one of which could in theory be selected as denoting successful speciation (fig. 1 of de Queiroz, 2007). Species concepts and species delimitation. Systematic Biology 56, 879–886, by permission of Oxford University Press.

into the future and thereby cement their respective fates as bona fide species. For example, Sedeek *et al.* (2014, p. 6202) reasonably concluded that '*Ophrys* orchids have only reached an early stage in the speciation continuum' (*sensu* Feder *et al.*, 2012). A remarkable spectrum of terms has been used to describe *Ophrys* microspecies, including 'extremely young species', 'incipient species', and 'lineages at the interface between incipient species and divergent ecotypes' (Breitkopf *et al.*, 2013, 2015; Sedeek *et al.*, 2014; Cozzolino *et al.*, 2020). All of these phrases may have earned their places in the pantheon of evolutionary biology theory, but in practice they do little to assist a taxonomist seeking to better classify these troublesome plants. Indeed, these terms merit serious reconsideration.

For example, does an 'incipient species' routinely precede a 'divergent ecotype'? So many of the terms used in attempts to describe speciation processes at least implicitly invoke temporal directionality; previous authors argued that *Ophrys* microspecies 'have only reached an early stage in the speciation continuum' and may lie 'at the interface between incipient species and divergent ecotypes.' A typical definition of an ecotype is 'a genetically distinct geographic variety, population or race *within a species* [our italics], which is genotypically adapted to specific environmental conditions … Although ecotypes exhibit phenotypic differences (such as in morphology or physiology) stemming from environmental heterogeneity, they are capable of interbreeding with other geographically adjacent ecotypes without loss of fertility or vigour.' This definition seems to us to be an accurate description of a typical *Ophrys* microspecies. Also, if an 'incipient species' is by definition an earlier stage of the genesis of a species than a 'divergent ecotype' (to us these terms are broadly synonymous: e.g. Sramkó *et al.*, 2019; Bateman, 2021), it has an even weaker claim to recognition as a bona fide species. In theory, it allows any local population (or, in the case of evolutionary processes such as polyploidy, any individual organism) to legitimately be regarded as an 'incipient species'.

Two fallacies of the tuning fork model

Most discussions of species concepts are rooted in the classic 'tuning fork' model of one biological lineage dividing gradually but inexorably into two lineages, a sharp point of dichotomy being both preceded and succeeded by a series of definable events, any one of which might be chosen by the observer as the key point at which one species officially became two (Fig. 9). This popular, frequently illustrated model raises two particularly important questions. Firstly, which of the available evolutionary events (and which category of evidence revealing that event) should be viewed as the critical threshold for species recognition, and secondly, is gradual, irreversible divergence even actually occurring?

Selecting a criterion as a threshold for successful speciation

Attempting to avoid having to select a particular step as the basis for recognizing successful speciation, de Queiroz (2005, 2007) formulated the process-free and step-free Unified Species Concept, which was used by Baguette *et al.* (2020) to underpin their arguments regarding species recognition in *Ophrys*. According to de Queiroz (2005: 1263), 'All [species concepts] either explicitly or implicitly equate species with separately evolving (segments of) metapopulation lineages, where a *metapopulation* is an inclusive population made up of a set of connected subpopulations, and a *lineage* (at the population level) is a population extended through time or an ancestral–descendant series of time-limited (instantaneous) populations' [our italics].

This definition obliges the reader to attempt to differentiate between 'population' and 'metapopulation', to define 'inclusivity', and then to determine how in practice the time element inherent in 'lineage' can be assessed—all without reference to particular data categories or evolutionary processes! Moreover, if not only 'metapopulations' but also 'segments of metapopulations' can be viewed as 'separately evolving' and therefore potentially represent legitimate species, the Unified Species Concept is evidently so laissez faire that in practice it could perhaps allow *Ophrys* macrospecies, mesospecies, and microspecies to simultaneously co-exist as legitimate species. Such arguments actually constitute an evasion from, rather than a solution to, the long-standing species problem; although challenging, it is essential to define a threshold and to state the evidence needed to judge whether that threshold has been transcended in any particular case.

We have already made clear our preference for seeking coincident phenotypic and genotypic boundaries detected via numerous population-level surveys. From a genetic viewpoint, gene flow and lineage divergence are the yin and yang of speciation. However, if complete absence of gene flow was a pre-requisite for species status, the genus *Ophrys* would contain only a single highly morphologically heterogeneous (and highly unpopular) species. This conclusion throws into sharp relief the desirability of quantifying rates of gene flow.

Despite the valiant ongoing protestations of some observers (e.g. Paulus, 2018), there is ample and incontrovertible genetic evidence of considerable natural gene flow among *Ophrys* microspecies (e.g. Soliva and Widmer, 2003; Devey *et al.*, 2008, 2009; Stökl *et al.*, 2008; Ayasse *et al.*, 2010, 2011; Schlüter *et al.*, 2011; Xu *et al.*, 2011; Cotrim *et al.*, 2016; Gögler *et al.*, 2016; Cozzolino *et al.*, 2020). Clearly, it is the level of ongoing gene flow that is critical, preferably standardized to typical population sizes and frequencies of successful reproduction per unit time. One obvious (if simplistic) threshold for declaring successful speciation would be a gene flow level of less than one migrant per generation (i.e. F_{ST}>0.2), a level that from a theoretical viewpoint precludes allele fixation through drift. In this context, Forrest *et al.* (2004) collated F_{ST} (as G_{ST}) data for 76 orchid-based studies but reported values exceeding 0.2 in only 23 (30%) of them. Another possible approach is to emphasize the presence or absence of fixed SNP differences in multicopy nuclear regions such as nrITS; subject to concerted evolution, such regions have a superior ability to coalesce (Donnelley and Tavaré, 1995; Kane *et al.*, 2012; Bateman, 2018; Pérez-Escobar *et al.*, 2020).

The perplexing challenge of demonstrating divergence

However detailed and thorough, studies of one or at most few populations are, in effect, observations confined to a single point in both space and time. In the absence of a discontinuity irrevocably separating a set of populations from all others (or of a strong fossil record that has captured in time extinct morphologies intermediate to their extant descendants—'missing links', which are singularly absent from the genus *Ophrys*), anyone interpreting the data is obliged to guess at the trajectory being taken by their study populations: are their respective paths divergent, parallel, or convergent? Genetic data are stronger than other categories of data in being able to provide at least a sketchy account of the recent history of a lineage (e.g. Bateman, 2018), but even genetics cannot realistically predict the future of a lineage—in particular, whether it will become more, or less, cohesive.

Based on the evidence we and others have gathered, we are unable to envisage the 113 microspecies of the *Sphegodes* group recognized by Delforge (2016) as a matched set of 113 'tuning forks'. Rather, we perceive the macrospecies *Sphegodes* as representing a four-dimensional reticulate network of (meta) population-level relationships. They are a mélange of local populations capable of limited diversification through a wide range of processes but transient when viewed on an ecological time scale, either rapidly becoming extinct or being drawn back into the genetic mainstream through ongoing gene flow. Reliance on pre-zygotic isolation is a recipe for continual phenotypic fluctuations that may appear striking in the short term but are typically directionless in the longer term; in other words, microevolutionary but not macroevolutionary. Populations seemingly matching the expected tuning fork trajectory at one moment in time will soon either disappear or reticulate, without ever having achieved the level of intrinsic independence that we regard as essential for a bona fide species.

We further argue that the scale of both genotypic and phenotypic differentiation demonstrated among populations within macrospecies *Sphegodes* would, in other genera of higher plants, be viewed as that routinely expected among conspecific populations. This single macrospecies is given sufficient cohesion through periodic gene flow among populations that further macrospecies—required by us to be self-delimiting through both genotypic and phenotypic discontinuities—are unlikely to emerge from within the ancestral macrospecies for the foreseeable future.

Reticulation precludes radiation

Our simple requirement for long-term cohesion in a bona fide species challenges almost every statement ever written about *Ophrys* microspecies. It adds to mathematically based criticisms of increasingly popular attempts to infer differential speciation (and extinction) rates from molecular phylogenies (cf. Louca and Pennell, 2020; Pagel, 2020). If the microspecies are not bona fide species then, by definition, they do not represent examples of successful speciation, nor collectively do they constitute a spectacular textbook example of an adaptive radiation (*contra* Soliva *et al*., 2001; Paulus, 2006; Ayasse *et al*., 2010; Breitkopf *et al*., 2015; Pineiro-Fernandez *et al*., 2019; Baguette *et al*., 2020). Qualifying as a genuine radiation, as traditionally defined, would require a collective long-term pattern of extensive divergences of fully reproductively isolated lineages—a pattern that simply has not been successfully demonstrated in *Ophrys* at any demographic level below the nine macrospecies. Also, rejection of species status means that the transfer of genes between individuals within each of the nine macrospecies can only legitimately be described as gene flow; strictly, the terms 'hybridization' and 'introgression' should be reserved for successful fertilization following pollen transfer between genuine species. Thus, the nine *Ophrys* macrospecies would limit the maximum theoretical number to 36 potential primary hybrids, negating the concept of books that exist only to document innumerable supposed hybrid combinations (e.g. Danesch and Danesch, 1972; Souche, 2008).

It would be helpful if more authors studying evolutionary processes in *Ophrys* taxa were also to express and justify conceptually their own taxonomic conclusions, rather than evading that particular responsibility. If the species category is going to live up to its universal reputation as being the most fundamental entity in systematic biology, it is essential that its systematic application should transcend arbitrary divisions that have been imposed rather than exposed through study in nature. Also, if we as systematists are to claim that we have successfully circumscribed a particular species, by definition it must be reliably separated from all other comparative entities by some kind of intrinsic phenotypic discontinuity that is underpinned by a corresponding genotypic discontinuity. Until the presence of such a discontinuity has been thoroughly demonstrated, a putative species essentially remains an untested hypothesis.

Implications for systematic applications of plastomes

Arguably the single most important facet of this study is that it adds to the increasingly large body of evidence demonstrating that not only traditional barcoding plastid regions such as *rbcL*, *matK*, and *trnL* (e.g. Li *et al*., 2015; Hollingsworth *et al*., 2016; Wang *et al*., 2018) but also whole-plastome sequences are incapable of accurately resolving clades that may (or may not!) be undergoing active macroevolution (i.e. speciation). This conclusion has been drawn not just from orchids (Cozzolino *et al*., 2020) but also from a wide variety of other plant groups. Indeed, hard incongruence in plastome trees has already become an accepted phenomenon in groups such as Asteraceae (Walker *et al*., 2019), Cucurbitaceae (Bellot *et al*., 2020), and legumes (Zhang *et al*., 2020). Plastomes are increasingly viewed as rich in regions offering little if any phylogenetic signal and, where signals are found, they are often conflicting (e.g. Dong *et al*., 2012; Goncalves *et al*., 2019; Bellot *et al*., 2020). Walker *et al*. (2019) found that they could recover their whole-plastome topology when using only *rpoC2*, an ~4100 bp gene within the large single-copy region that, ironically, had hitherto been little used for phylogeny reconstruction or advocated as a suitable DNA barcode (see also Logacheva *et al*., 2007).

A further intriguing potential complication in this case is provided by the reputed occurrence of plastid dimorphism in *O. sphegodes* (Lux and Hudák, 1987), suggesting that maintenance of multiple plastid micromorphs is feasible within individual *Ophrys* plants. At the genomic level, Wang and Lanfear (2019) surveyed the extent among angiosperms of structural heteroplasmy, demonstrating the presence within single plants of both 'A' and 'B' structural haplotypes (consistently in numbers judged statistically equal) in all 58 species analysed. The opposing orientations of the single-copy regions are attributed to frequent intramolecular 'flip flop' recombination mediated by the pair of inverted repeats (Wolfe and Randle, 2004; Sullivan *et al*., 2017; Sancho *et al*., 2018; Walker *et al*., 2019; Wang and Lanfear, 2019). Although we did not seek multiple haplotypes within single individuals, we recovered approximately equal numbers of 'A' and 'B' haplotypes across the *O. sphegodes* clade (Supplementary Fig. S3), suggesting that orchids too probably exhibit heteroplasmy. This observation raises a more pertinent question of how frequently interplastome recombination may occur (Sullivan *et al*., 2017; Sancho *et al*., 2018; Walker *et al*., 2019), potentially 'hybridizing' (and thus imposing a reticulate history upon) plastome lineages.

The comparative strength of geographic relative to taxonomic clustering of plastomes in groups suspected of undergoing active macroevolution can only credibly be explained through ongoing plastid capture that is sufficiently frequent to mask transient local diversification toward putative adaptive optima. Current evidence suggests that nuclear genomes are able to become distinct from each other more rapidly and reliably than plastid or mitochondrial genomes and are therefore far more informative about evolutionary relationships and processes. Plastomes appear unable to unify (and thus cannot permit circumscription of) potential species until the reproductive isolation of the relevant lineages is (almost) complete (Bateman, 2021). Given such evidence, the continuing popularity of plastid sequences as a tool (often the primary tool) to address such questions is regrettable.

Studies of clades that are suspected of actively undergoing macroevolution routinely cite—typically in the same sentence—both incomplete lineage sorting (shared/retained ancestral polymorphism) and 'hybridization' (gene flow through secondary contact) as confounding factors (e.g. Blanco-Pastor

et al., 2012; Luo *et al.*, 2014; Pérez-Escobar *et al.*, 2016, 2020; Cozzolino *et al.*, 2020). However, few such studies suggest how these two processes, contrasting in nature but similar in overall genetic effect, can be distinguished with any confidence. Unfortunately, making that distinction is crucial to understanding whether or not the two or more lineages in question are likely to have been better differentiated (genotypically and/or phenotypically) in the past than they are today. In short, not all reticulation events are equal, but collectively, they represent a genuine threat to our ability to identify the threshold beyond which a lineage has acquired sufficient genotypic and phenotypic cohesion to constitute a bona fide species. Moreover, they seriously challenge the role long held by the repeatedly dichotomous tree as being the most appropriate motif for representing such relationships (reviewed by Quammen, 2018).

Conclusions

(i) Several authors have argued that, intuitively, the true number of species within the Mediterranean orchid genus *Ophrys* should be more than the nine molecularly distinct macrospecies recognized by Bateman *et al.* (2018) but fewer than the 360+ supposedly reproductively isolated microspecies recognized by Delforge (2016) and other ethologists. This project was devised primarily to seek intermediate 'mesospecies' within the most taxonomically controversial of all of the macrospecies, *O. sphegodes s.l.*, by means of combining whole-plastome sequencing with detailed *in situ* morphometrics.

(ii) Our morphometric results demonstrate conclusively that morphological variation within macrospecies *Sphegodes* is massively multidimensional and that any distinctions competent to induce unequivocal clustering into potential morphologically distinct species were too subtle to be captured by our character suite, despite the large number of characters scored. Rather, there is considerable overlap between not only mesospecies but also microspecies; the overlap undoubtedly partly reflects ontogenetic and ecophenotypic variation, but these factors are insufficient to account for such extensive overlap. Both mesospecies and microspecies therefore appear to be arbitrary subdivisions of a morphological continuum.

(iii) Far from being a panacea for species circumscription, plastome haplotypes actually reflect geographic location more strongly than assignment to mesospecies or microspecies. Five haplotype groups were recognized within macrospecies *Sphegodes* in the ML trees, most mesospecies occupying at least two of the five groups. These results imply extensive regional plastid capture reflecting considerable gene flow among mesospecies.

(iv) Compared with the plastome trees, ordination of the plastome data provided stronger clustering that suggests only three main groups, approximating the west, central, and eastern Mediterranean, respectively. The most distinct haplotype transition coincides with the Carpathian Divide. The three groups are hypothesized to reflect the three classic northward migration routes following the end-Pleistocene deglaciation of Europe, implying that macrospecies *Sphegodes* already spanned the Mediterranean Basin prior to the onset of the Weichselian glaciation at ~120 ka BP.

(v) Molecular similarities reliably exceed morphological similarities. Although positive, the correlation between morphological similarity and molecular similarity within macrospecies *Sphegodes* is poor. Nonetheless, both morphological and especially molecular similarities are on average greater for multiple plants of the same microspecies, particularly if they are sampled in the same geographic region.

(vi) The popular 'ethological' evolutionary model recognizes as numerous 'ecological species' (microspecies) those lineages that are perceived as actively diverging. The primary cause of the supposed divergence in inferred to be density-dependent selection on very few features that immediately dictate extreme pollinator specificity. However, proponents of this model have assumed, rather than actually demonstrated, that the many microspecies are genuinely diverging and will continue to do so. It seems to us remarkable that ~97% of all *Ophrys* species would currently be caught in an intermediate stage of classic 'tuning fork' divergence rather than having completely broken the reproductive link with their respective ancestral populations.

(vii) We conversely envisage macrospecies *Sphegodes* as a complex four-dimensional reticulate network of lineages, generated locally and transiently through a wide spectrum of mechanisms that include directional selection but also include drift and epigenesis, and especially gene flow with other transient lineages. Each crudely and transiently adapted local 'metapopulation' is unlikely to maintain an independent evolutionary trajectory long enough to genuinely speciate, particularly in the face of rapid environmental change. Speciation requires near-complete escape from ongoing gene flow, whereas in *Ophrys* gene flow occurs among not only microspecies but also mesospecies and macrospecies. The frequent but localized and poorly structured microevolution that characterizes the *Ophrys sphegodes* complex and other *Ophrys* macrospecies is often convergent and rarely leads to macroevolution. Active evolution does not necessarily equate with speciation and without speciation there can be no speciation rate or radiation, adaptive or otherwise. Also, related terms such as 'incipient species' and 'divergent ecotype' need to be carefully defined, and their subsequent use in particular cases requires detailed justification.

(viii) Choosing between these contrasting models of divergent versus reticulate evolution will require NGS of nuclear genomes (preferably supported by an assembled genome of reference) plus ordination of corresponding morphometric and biochemical matrices, as well as conceptual advances in data analysis that will help to make the crucial distinction between retained ancestral polymorphism—consistent with lineage divergence—and polymorphisms reflecting gene flow through 'hybridization'—more consistent with lineage convergence.

(ix) Sufficient data have now accumulated from contrasting taxonomic groups to suggest that the phylogenetic value of initially plastid regions and latterly whole plastomes has been greatly exaggerated. Any resolution achieved using plastomes depends primarily on prior extinction of intermediate lineages, meaning that plastomes cannot adequately document either reticulation occurring early in a putative speciation event or cases of more advanced, classically dichotomous speciation if they occur frequently from the same lineage within a short period of time (i.e. are *de facto* radiations).

Supplementary data

The following supplementary data are available at *JXB* online.

Table S1. Details of *Ophrys* accessions subjected to morphometric analysis.

Table S2. Details of *Ophrys* accessions subjected to genome skimming.

Fig. S1. Plot of principal coordinates 3 and 4 for 51 morphometric characters and 124 individuals of the *Ophrys sphegodes* clade.

Fig. S2. Unrooted uncorrected P-distances network of whole-plastomes obtained via genome skimming from 64 individuals of the *Sphegodes* macrospecies, six individuals of the *Fuciflora* macrospecies ('inner outgroup'), and one individual of the *Umbilicata* macrospecies (functional 'outer outgroup').

Fig. S3. Structural comparison of the two most sequence-divergent ingroup plastomes analysed.

Acknowledgements

The samples of *Ophrys sphegodes s.s.* from Normandy were collected by RMB and PJR when they were in receipt of modest but greatly appreciated field support from the Botanical Society of Britain and Ireland. We thank profusely Jess Stone, Kath Barrett, and Barry Tattersall for providing silica-gel samples of reliably identified *Ophrys* plants that not only filled critical gaps in our broader microspecies coverage but also proved invaluable in the wake of unfortunate failures during laboratory preparation of samples initially prioritized by us because they had yielded corresponding morphometric data. Lastly, we thank Svante Malmgren for permission to reproduce the image in Fig. 8.

Author contributions

The project was conceived by RMB and PJR, who also together collected the research materials. Morphometric data were collected by RMB and PJR, analysed and presented by RMB. Genome skimming data were collected by ARMM under the supervision of RSC and DSD, analysed and curated by OAP, and presented by RMB and OAP. The manuscript was written and revised by RMB, with assistance from PJR and OAP.

Data availability

Genome skimming data have been deposited as aligned sequences in GitHub. Morphometric data are available from the corresponding author (RB) upon reasonable request.

References

Andrews SC. 2015. FastQC: a quality control tool for high throughput sequence data. Cambridge, UK: Babraham Institute.

Ayasse M, Gögler J, Stökl J. 2010. Pollinator-driven speciation in sexually deceptive orchids of the genus *Ophrys*. In: Glaubrecht M, ed. Evolution in action. Berlin: Springer, 101–118.

Ayasse M, Schiestl FP, Paulus HF, Löfstedt C, Hansson B, Ibarra F, Francke W. 2000. Evolution of reproductive strategies in the sexually deceptive orchid *Ophrys sphegodes*: how does flower-specific variation of odor signals influence reproductive success? Evolution **54**, 1995–2006.

Ayasse M, Stökl J, Francke W. 2011. Chemical ecology and pollinator-driven speciation in sexually deceptive orchids. Phytochemistry **72**, 1667–1677.

Baguette M, Betrand J, Stevens VM, Schatz B. 2020. Why are there so many bee-orchid species? Adaptive radiation by intraspecific competition for mnemonic pollinators. Biological Reviews of the Cambridge Philosophical Society **95**, 1630–1663.

Bateman RM. 2001. Evolution and classification of European orchids: insights from molecular and morphological characters. Journal Europäischer Orchideen **33**, 33–119.

Bateman RM. 2016. Après le déluge: ubiquitous field barcoding should drive 21st century taxonomy. In: Olson PD, Hughes J, Cotton JA, eds. Next generation systematics. Systematics Association Special Volume 85. Cambridge: Cambridge University Press, 123–153.

Bateman RM. 2018. Two bees or not two bees? An overview of *Ophrys* systematics. Berichte aus den Arbeitskreisen Heimische Orchideen **35**, 5–46.

Bateman RM. 2020. Implications of next-generation sequencing for the systematics and evolution of the terrestrial orchid genus *Epipactis*, with particular reference to the British Isles. Kew Bulletin **75**, 4.

Bateman RM. 2021. Species circumscription within 'cryptic' clades: a nihilist's view. In: Monro A, ed. Cryptic species. Systematics Association Special Volume 88. Cambridge: Cambridge University Press (in press).

Bateman RM, Bradshaw E, Devey DS, Glover BJ, Malmgren S, Sramko G, Thomas MM, Rudall PJ. 2011. Species arguments: clarifying concepts of species delimitation in the pseudo-copulatory orchid genus *Ophrys.* Botanical Journal of the Linnean Society **165**, 336–347.

Bateman RM, Guy JJ, Rudall PJ, Leitch IJ, Pellicer J, Leitch AR. 2018*b*. Evolutionary and functional potential of ploidy increase within individual plants: somatic ploidy mapping of the complex labellum of sexually deceptive bee orchids. Annals of Botany **122**, 133–150.

Bateman RM, Hollingsworth PM, Preston J, Luo Y-B, Pridgeon AM, Chase MW. 2003. Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). Botanical Journal of the Linnean Society **142**, 1–40.

Bateman RM, Molnár VA, Sramkó G. 2017. In situ morphometric survey elucidates the evolutionary systematics of the Eurasian *Himantoglossum* clade (Orchidaceae: Orchidinae). PeerJ **5**, e2893.

Bateman RM, Sramkó G, Paun O. 2018*a*. Integrating restriction site-associated DNA sequencing (RAD-seq) with morphological cladistic analysis clarifies evolutionary relationships among major species groups of bee orchids. Annals of Botany **121**, 85–105.

Bellot S, Mitchell TC, Schaefer H. 2020. Phylogenetic informativeness analyses to clarify past diversification processes in Cucurbitaceae. Scientific Reports **10**, 488.

Berger BA, Han J, Sessa EB, Gardner AG, Shepherd KA, Ricigliano VA, Jabaily RS, Howarth DG. 2017. The unexpected depths of genome-skimming data: a case study examining Goodeniaceae floral symmetry genes. Applications in Plant Sciences **5**, 1700042.

Bertrand JAM, Gibert A, Llauro C, Panaud O. 2019. Characterization of the complete plastome of *Ophrys aveyronensis*, a Euro-Mediterranean orchid with an intriguing disjunct geographic distribution. Mitochondrial DNA B: Resources **4**, 3256–3257.

Blanco-Pastor JL, Vargas P, Pfeil BE. 2012. Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean *Linaria*. PLoS One **7**, e39089.

Bradshaw E, Rudall PJ, Devey DS, Thomas MM, Glover BJ, Bateman RM. 2010. Comparative labellum micromorphology in the sexually deceptive temperate orchid genus *Ophrys*: diverse epidermal cell types and multiple origins of structural colour. Botanical Journal of the Linnean Society **162**, 502–540.

Breitkopf H, Onstein RE, Cafasso D, Schlüter PM, Cozzolino S. 2015. Multiple shifts to different pollinators fuelled rapid diversification in sexually deceptive *Ophrys* orchids. New Phytologist **207**, 377–389.

Breitkopf H, Schlüter PM, Xu S, Schiestl FP, Cozzolino S, Scopece G. 2013. Pollinator shifts between *Ophrys sphegodes* populations: might adaptation to different pollinators drive population divergence? Journal of Evolutionary Biology **26**, 2197–2208.

Bull JJ, Cunningham CW, Molineux IJ, Badgett MR, Hillis DM. 1993. Experimental molecular evolution of Bacteriophage T7. Evolution **47**, 993–1007.

Claessens J, Kleynen J. 2011. The flower of the European orchid: form and function. Voerendaal, The Netherlands: Published by the authors.

Cortis P, Vereecken NJ, Schiestl FP, Barone Lumaga MR, Scrugli A, Cozzolino S. 2009. Pollinator convergence and the nature of species' boundaries in sympatric Sardinian *Ophrys* (Orchidaceae). Annals of Botany **104**, 497–506.

Cotrim H, Monteiro F, Sousa E, Pinto MJ, Fay MF. 2016. Marked hybridization and introgression in *Ophrys* sect. *Pseudophrys* in the western Iberian Peninsula. American Journal of Botany **103**, 677–691.

Cozzolino S, Scopece G, Roma L, Schlüter PM. 2020. Different filtering strategies of genotyping-by-sequencing data provide complementary resolutions of species boundaries and relationships in a clade of sexually deceptive orchids. Journal of Systematics and Evolution **58**, 133–144.

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

Danesch E, Danesch O. 1972. Orchideen Europas 3: *Ophrys* Hybriden. Bern: Hallwag.

Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research **14**, 1394–1403.

de Queiroz K. 2005. Different species problems and their resolution. BioEssays **27**, 1263–1269.

de Queiroz K. 2007. Species concepts and species delimitation. Systematic Biology **56**, 879–886.

Delforge P. 1994. Orchidés d'Europe, d'Afrique du Nord et du Proche-Orient. Paris: Delachaux et Niestle.

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

Delforge P. 2016. Orchidés d'Europe, d'Afrique du Nord et do Proche-Orient, 4th edn. Paris: Delachaux et Niestle.

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

Devey DS, Bateman RM, Fay MF, Hawkins JA. 2009. Genetic structure and systematic relationships within the *Ophrys fuciflora* aggregate (Orchidaceae: Orchidinae): high diversity in Kent and a wind-induced discontinuity bisecting the Adriatic. Annals of Botany **104**, 483–495.

Devillers P, Devillers-Terschuren J. 1994. Essai d'analyse systématique du genre *Ophrys*. Naturalistes Belges **75**, 273–400.

Dong W, Liu J, Wang L, Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS One **7**, e35071.

Donnelly P, Tavaré S. 1995. Coalescents and genealogical structure under neutrality. Annual Review of Genetics **29**, 401–421.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One **6**, e19379.

Fateryga AV, Efimov VA, Fateryga VV. 2018. Taxonomic notes on the genus *Ophrys* L. (Orchidaceae) in the Crimea and the North Caucasus. Turczaninowia **21**, 9–18.

Feder JL, Egan SP, Nosil P. 2012. The genomics of speciation-with-gene-flow. Trends in Genetics **28**, 342–350.

Ferris C, King RA, Hewitt GM. 1999. Isolation within species and the history of glacial refugia. In: Hollingsworth PM, Bateman RM, Gornall RJ, eds. Molecular systematics and plant evolution. London: Taylor & Francis, 20–34.

Forrest AD, Hollingsworth ML, Hollingsworth PM, Sydes C, Bateman RM. 2004. Population genetic structure in European populations of *Spiranthes romanzoffiana* set in the context of other genetic studies on orchids. Heredity **92**, 218–227.

Gögler J, Stökl J, Cortis P, Beyrle H, Barone Lumaga MR, Cozzolino S, Ayasse M. 2015. Increased divergence in floral morphology strongly reduces gene flow in sympatric sexually deceptive orchids with the same pollinator. Evolutionary Ecology **29**, 703–717.

Gögler J, Stökl J, Sramkova A, Twele R, Francke W, Cozzolino S, Cortis P, Scrugli A, Ayasse M. 2009. Ménage a trois—two endemic species of deceptive orchids and one pollinator species. Evolution **63**, 2222–2234.

Gögler J, Zitari A, Paulus H, Cozzolino S, Ayasse M. 2016. Species boundaries in the *Ophrys iricolor* group in Tunisia: do local endemics always matter? Plant Systematics and Evolution **302**, 481–489.

Gonçalves DJP, Simpson BB, Ortiz EM, Shimizu GH, Jansen RK. 2019. Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. Molecular Phylogenetics and Evolution **138**, 219–232.

Gower JC. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **52**, 325–338.

Gower JC. 1971. A general coefficient of similarity and some of its properties. Biometrics **27**, 857–872.

Gower JC. 1985. Measures of similarity, dissimilarity and distance. In: Encyclopedia of statistical sciences 5. New York: Wiley, 397–405.

Gower JC, Ross GJS. 1969. Minimum spanning trees and single linkage cluster analysis. Journal of the Royal Statistical Society C **18**, 54–64.

Harrison N, Kidner CA. 2011. Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? Taxon **60**, 1552–1556.

Hennecke M. 2013. Morphologisches Dendrogramm der *Ophrys*-Sektionen. Berichte aus den Arbeitskreisen Heimische Orchideen **30**, 90–108.

Hennecke M, Munzinger S. 2014. *Ophrys* subgen. *Fuciflorae* sect. *Araniferae* subsect. *Argolicae*. Berichte aus den Arbeitskreisen Heimische Orchideen **31**, 232–238.

Hillis DM, Bull JJ, White ME, Badgett MR, Molineux IJ. 1992. Experimental phylogenetics: generation of a known phylogeny. Science **255**, 589–592.

Hirsch S, Baumberger R, Grossniklaus U. 2012. Epigenetic variation, inheritance, and selection in plant populations. Cold Spring Harbor Symposia on Quantitative Biology **77**, 97–104.

Hollingsworth P, De-Zhu L, Van der Bank M, Twyford A. 2016. Telling plant species apart with DNA: from barcodes to genomes. Philosophical Transactions of the Royal Society B: Biological Sciences **371**, e20150338.

Kane N, Sveinsson S, Dempewolf H, Yang JY, Zhang D, Engels JM, Cronk Q. 2012. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. American Journal of Botany **99**, 320–329.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics **15**, 356.

Kreutz CAJ. 2004. Kompendium der Europäischen Orchideen. Landgraaf: Published by the author.

Kropf M, Renner SS. 2008. Pollinator-mediated selfing in two deceptive orchids and a review of pollinium tracking studies addressing geitonogamy. Oecologia **155**, 497–508.

Krueger F. 2015. TrimGalore!: a wrapper tool around Cutadept and FastQC to consistently apply quality and adapter trimming to FastQ files. Cambridge, UK: Babraham Bioinformatics.

Kühn R, Pedersen HA, Cribb P. 2019. Field guide to the orchids of Europe and the Mediterranean. Kew: Royal Botanic Gardens Kew.

Kullenberg B. 1961. Studies in *Ophrys* pollination. Zoologiska Bidrag från Uppsala **34**, 1–340.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution **33**, 1870–1874.

Lemmon EM, Lemmon AR. 2013. High-throughput genomic data in systematics and phylogenetics. Annual Review of Ecology, Evolution and Systematics **44**, 99–121.

Li H, Ruan J, Durbin R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research **18**, 1851–1858.

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. 2015. Plant DNA barcoding: from gene to genome. Biological Reviews of the Cambridge Philosophical Society **90**, 157–166.

Logacheva MD, Penin AA, Samigullin TH, Vallejo-Roman CM, Antonov AS. 2007. Phylogeny of flowering plants by the chloroplast genome sequences: in search of a lucky gene. Biochemistry **72**, 1324–1330.

Louca S, Pennell MW. 2020. Extant timetrees are consistent with a myriad of diversification histories. Nature **580**, 502–505.

Luo J, Hou BW, Niu ZT, Xue QY, Ding XY. 2014. Comparative chloroplast genomes of photosynthetic orchids: insights into evolution of the Orchidaceae and development of molecular markers for phylogenetic applications. PLoS One **9**, e99016.

Lux A, Hudák J. 1987. Plastid dimorphism in leaves of the terrestrial orchid *Ophrys sphegodes*. New Phytologist **107**, 47–51.

Maad J, Reinhammar LG. 2004. Incidence of geitonogamy differs between two populations in the hawkmoth-pollinated *Platanthera bifolia* (Orchidaceae). Canadian Journal of Botany **82**, 1586–1593.

Malmgren S. 2008. Are there 25 or 250 *Ophrys* species? Journal of the Hardy Orchid Society **5**, 95–100.

Mant J, Peakall R, Schiestl FP. 2005. Does selection on floral odor promote differentiation among populations and species of the sexually deceptive orchid genus *Ophrys*? Evolution **59**, 1449–1463.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution **66**, 526–538.

McKenna A, Hanna M, Banks E, *et al.* 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research **20**, 1297–1303.

Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. Journal of Biogeography **36**, 1333–1345.

Meisner J, Albrechtsen A. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics **210**, 719–731.

Olson PD, Hughes J, Cotton JA, **eds.** 2016. Next generation systematics. Systematics Association Special Volume 85. Cambridge: Cambridge University Press.

Pagel M. 2020. Evolutionary trees can't reveal speciation and extinction rates. Nature **580**, 461–462.

Paulus HF. 2006. Deceived males—pollination biology of the Mediterranean orchid genus Ophrys (*Orchidaceae*). Journal Europäischer Orchideen **38**, 303–353.

Paulus HF. 2018. Pollinators as isolation mechanisms: field observations and field experiments regarding specificity of pollinator attraction in the genus *Ophrys* (Orchidaceae und Insecta, Hymenoptera, Apoidea). Entomologia Generalis **37**, 261–316.

Paulus HF. 2019. Speciation, pattern recognition and the maximization of pollination: general questions and answers given by the reproductive biology of the orchid genus *Ophrys*. Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology **205**, 285–300.

Paulus HF, Gack C. 1990. Pollinators as prepollinating isolating factors: evolution and speciation in *Ophrys* (Orchidaceae). Israel Journal of Botany **39**, 43–97.

Paun O, Bateman RM, Fay MF, Hedrén M, Civeyrel L, Chase MW. 2010. Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza*: Orchidaceae). Molecular Biology and Evolution **27**, 2465–2473.

Payne RW, Harding SA, Murray DA, Souter DM, Baird DB, Glaser AI, Welham SJ, Gilmour AR, Thompson R, Webster R, **eds.** 2011. Genstat v14. Hemel Hempstead, UK: VSN International.

Pedersen H, Faurholdt N. 2007. *Ophrys*: the bee orchids of Europe. Kew: Kew Publishing.

Pérez-Escobar OA, Balbuena JA, Gottschling M. 2016. Rumbling orchids: how to assess divergent evolution between chloroplast endosymbionts and the nuclear host. Systematic Biology **65**, 51–65.

Pérez-Escobar OA, Bogarín D, Schley R, *et al.* 2020. Resolving relationships in an exceedingly young Neotropical orchid lineage using Genotyping-by-sequencing data. Molecular Phylogenetics and Evolution **144**, 106672.

Piñeiro Fernández L, Byers KJRP, Cai J, Sedeek KEM, Kellenberger RT, Russo A, Qi W, Aquino Fournier C, Schlüter PM. 2019. A phylogenomic analysis of the floral transcriptomes of sexually deceptive and rewarding European orchids, *Ophrys* and *Gymnadenia*. Frontiers in Plant Science **10**, 1553.

Quammen D. 2018. The tangled tree: a radical new history of life. London: HarperCollins.

Rakosy D, Streinzer M, Paulus HF, Spaethe J. 2012. Floral visual signal increases reproductive success in a sexually deceptive orchid. Arthropod–Plant Interactions **6**, 671–681.

Reuter JA, Spacek DV, Snyder MP. 2015. High-throughput sequencing technologies. Molecular Cell **58**, 586–597.

Roma L, Cozzolino S, Schlüter PM, Scopece G, Cafasso D. 2018. The complete plastid genomes of *Ophrys iricolor* and *O. sphegodes* (Orchidaceae) and comparative analyses with other orchids. PLoS One **13**, e0204174.

Sancho R, Cantalapiedra CP, López-Alvarez D, Gordon SP, Vogel JP, Catalán P, Contreras-Moreira B. 2018. Comparative plastome genomics and phylogenomics of *Brachypodium*: flowering time signatures, introgression and recombination in recently diverged ecotypes. New Phytologist **218**, 1631–1644.

Schiestl FP. 2005. On the success of a swindle: pollination by deception in orchids. Die Naturwissenschaften **92**, 255–264.

Schiestl FP, Ayasse M, Paulus HF, Lofstedt C, Hansson BS, Ibarra F, Francke W. 1999. Orchid pollination by sexual swindle. Nature **399**, 421–422.

Schiestl FP, Cozzolino S. 2008. Evolution of sexual mimicry in the orchid subtribe orchidinae: the role of preadaptations in the attraction of male bees as pollinators. BMC Evolutionary Biology **8**, 27.

Schlüter PM, Ruas PM, Kohl G, Ruas CF, Stuessy TF, Paulus HF. 2011. Evidence for progenitor-derivative speciation in sexually deceptive orchids. Annals of Botany **108**, 895–906.

Schubert M, Ermini L, Der Sarkissian C, *et al.* 2014. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nature Protocols **9**, 1056–1082.

Scopece G, Musacchio A, Widmer A, Cozzolino S. 2007. Patterns of reproductive isolation in Mediterranean deceptive orchids. Evolution **61**, 2623–2642.

Sedeek KE, Scopece G, Staedler YM, Schönenberger J, Cozzolino S, Schiestl FP, Schlüter PM. 2014. Genic rather than genome-wide differences between sexually deceptive *Ophrys* orchids with different pollinators. Molecular Ecology **23**, 6192–6205.

Sedeek KE, Whittle E, Guthörl D, Grossniklaus U, Shanklin J, Schlüter PM. 2016. Amino acid change in an orchid desaturase enables mimicry of the pollinator's sex pheromone. Current Biology **26**, 1505–1511.

Sedeek KEM, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu Z-J, Grossniklaus U, Schiestl FP, Schlüter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. PLoS One **8**, e64621.

Smith SA, Walker-Hale N, Walker JF, Brown JW. 2020. Phylogenetic conflicts, combinability, and deep phylogenomics in plants. Systematic Biology **69**, 579–592.

Soliva M, Kocyan A, Widmer A. 2001. Molecular phylogenetics of the sexually deceptive orchid genus *Ophrys* (Orchidaceae) based on nuclear and chloroplast DNA sequences. Molecular Phylogenetics and Evolution **20**, 78–88.

Soliva M, Widmer A. 2003. Gene flow across species boundaries in sympatric, sexually deceptive *Ophrys* (Orchidaceae) species. Evolution **57**, 2252–2261.

Souche R. 2008. Hybrides d'*Ophrys* du Bassin Méditerranéen Occidental. Societé Occitane d'Orchidologie.

Sramkó G, Paun O, Brandrud MK, Laczkó L, Molnár A, Bateman RM. 2019. Iterative allogamy–autogamy transitions drive actual and incipient speciation during the ongoing evolutionary radiation within the orchid genus *Epipactis* (Orchidaceae). Annals of Botany **124**, 481–497.

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

Stökl J. 2007. Pollinator driven radiation in sexually deceptive orchids of the genus *Ophrys.* Doctoral dissertation, University of Ulm, Germany.

Stökl J, Schlüter PM, Stuessy TF, Paulus HF, Assum G, Ayasse M. 2008. Scent variation and hybridization cause the displacement of a sexually deceptive orchid species. American Journal of Botany **95**, 472–481.

Streinzer M, Ellis T, Paulus HF, Spaethe J. 2010. Visual discrimination between two sexually deceptive *Ophrys* species by a bee pollinator. Arthropod-Plant Interactions **4**, 141–148.

Sullivan AR, Schiffthaler B, Thompson SL, Street NR, Wang XR. 2017. Interspecific plastome recombination reflects ancient reticulate evolution in *Picea* (Pinaceae). Molecular Biology and Evolution **34**, 1689–1701.

Tremblay RL, Ackerman JD, Zimmerman JK, Calvo RN. 2005. Variation in sexual reproduction in orchids and its evolutionary consequences: a spasmodic journey to diversification. Biological Journal of the Linnean Society **84**, 1–54.

Tyteca D, Baguette M. 2017. *Ophrys* (Orchidaceae) systematics: when molecular phylogenetics, morphology and biology reconcile. Berichte aus den Arbeitskreisen Heimische Orchideen **34**, 37–103.

Véla E, Rebbas K, Martin R, de Premorel G, Tison J-M. 2015. Waiting for integrative taxonomy: morphospecies as an operational proxy for the radiative and reticulate genus *Ophrys* L. (Orchidaceae)? European Journal of Environmental Sciences **5**, 153–157.

Vereecken NJ. 2009. Deceptive behaviour in plants. I. Pollination by sexual deception in orchids: a host–parasite perspective. In: Baluska F, ed. Plant–environment interactions. Berlin: Springer, 203–222.

Vereecken NJ, Cozzolino S, Schiestl FP. 2010. Hybrid floral scent novelty drives pollinator shift in sexually deceptive orchids. BMC Evolutionary Biology **10**, 103.

Vereecken NJ, Schiestl FP. 2008. The evolution of imperfect floral mimicry. Proceedings of the National Academy of Sciences, USA **105**, 7484–7488.

Vereecken NJ, Schiestl FP. 2009. On the roles of colour and scent in a specialized floral mimicry system. Annals of Botany **104**, 1077–1084.

Vereecken NJ, Streinzer M, Ayasse M, Spaethe J, Paulus HF, Stökl J, Cortis P, Schiestl FP. 2011. Integrating past and present studies on *Ophrys* pollination—a comment on Bradshaw *et al*. Botanical Journal of the Linnean Society **165**, 329–335.

Vereecken NJ, Wilson CA, Hötling S, Schulz S, Banketov SA, Mardulyn P. 2012. Pre-adaptations and the evolution of pollination by sexual deception: Cope's rule of specialization revisited. Proceedings of the Royal Society B: Biological Sciences **279**, 4786–4794.

Walker JF, Walker-Hale N, Vargas OM, Larson DA, Stull GW. 2019. Characterizing gene tree conflict in plastome-inferred phylogenies. PeerJ **7**, e7747.

Wang W, Lanfear R. 2019. Long-reads reveal that the chloroplast genome exists in two distinct versions in most plants. Genome Biology and Evolution **11**, 3372–3381.

Wang X, Gussarova G, Ruhsam M, de Vere N, Metherell C, Hollingsworth PM, Twyford AD. 2018. DNA barcoding a taxonomically complex hemiparasitic genus reveals deep divergence between ploidy levels but lack of species-level resolution. AoB Plants **10**, ply026.

Wickham H. 2016. Ggplot2: elegant graphics for data analysis, 2nd edn. Springer.

Wickland DP, Battu G, Hudson KA, Diers BW, Hudson ME. 2017. A comparison of genotyping-by-sequencing analysis methods on low-coverage crop datasets shows advantages of a new workflow, GB-eaSy. BMC Bioinformatics **18**, 586.

Widmer A, Lexer C. 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. Trends in Ecology & Evolution **16**, 267–269.

Wolfe AD, Randle CP. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics. Systematic Botany **29**, 1011–1020.

Xu S, Schlüter PM, Scopece G, Breitkopf H, Gross K, Cozzolino S, Schiestl FP. 2011. Floral isolation is the main reproductive barrier among closely related sexually deceptive orchids. Evolution **65**, 2606–2620.

Zhang R, Wang YH, Jin JJ, *et al.* 2020. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. Systematic Biology **69**, 613–622.