MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 6192–6205 doi: 10.1111/mec.12992

Genic rather than genome-wide differences between sexually deceptive *Ophrys* orchids with different pollinators

KHALID E. M. SEDEEK,* GIOVANNI SCOPECE,† YANNICK M. STAEDLER,‡ JÜRG SCHÖNENBERGER,‡ SALVATORE COZZOLINO,† FLORIAN P. SCHIESTL* and PHILIPP M. SCHLÜTER*

*Institute of Systematic Botany, University of Zurich, Zollikerstr. 107, CH-8008 Zurich, Switzerland, †Department of Biology, University of Naples Federico II, Complesso Universitario MSA, Via Cinthia, I-80126 Naples, Italy, ‡Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, A-1030 Vienna, Austria

Abstract

High pollinator specificity and the potential for simple genetic changes to affect pollinator attraction make sexually deceptive orchids an ideal system for the study of ecological speciation, in which change of flower odour is likely important. This study surveys reproductive barriers and differences in floral phenotypes in a group of four closely related, coflowering sympatric *Ophrys* species and uses a genotyping-by-sequencing (GBS) approach to obtain information on the proportion of the genome that is differentiated between species. *Ophrys* species were found to effectively lack postpollination barriers, but are strongly isolated by their different pollinators (floral isolation) and, to a smaller extent, by shifts in flowering time (temporal isolation). Although flower morphology and perhaps labellum coloration may contribute to floral isolation, reproductive barriers may largely be due to differences in flower odour chemistry. GBS revealed shared polymorphism throughout the *Ophrys* genome, with very little population structure between species. Genome scans for F_{ST} outliers identified few markers that are highly differentiated between species and repeatable in several populations. These genome scans also revealed highly differentiated polymorphisms in genes with putative involvement in floral odour production, including a previously identified candidate gene thought to be involved in the biosynthesis of pseudo-pheromones by the orchid flowers. Taken together, these data suggest that ecological speciation associated with different pollinators in sexually deceptive orchids has a genic rather than a genomic basis, placing these species at an early phase of genomic divergence within the 'speciation continuum'.

Keywords: ecological speciation, floral reproductive isolation, genome scan, genotyping by sequencing, pollination, sexual deception

Received 11 August 2014; revision received 25 October 2014; accepted 29 October 2014

Introduction

Speciation is characterized by the formation of barriers to gene flow between populations that will ultimately result in genetic divergence (Coyne & Orr 2004; Lowry *et al*. 2008). Whereas speciation research historically focused upon the geographic mode of speciation (e.g. sympatric vs. allopatric), focus has recently shifted on the processes of divergence (Rundle & Nosil 2005; Wolf *et al*. 2010; Smadja & Butlin 2011; Feder *et al*. 2012; Nosil 2012). Barriers to gene flow can be mediated by many genes of small effect that are distributed throughout the genome, or by few genes with large effect on reproductive isolation (RI); this latter situation is conceptualized in the genic view of speciation (Wu 2001; Wu & Ting

Correspondence: Philipp M. Schlüter, Fax: +41 44 63 48403; E-mail: philipp.schlueter@systbot.uzh.ch

2004; Lexer & Widmer 2008). In genic speciation, only few loci under divergent selection are responsible for species differences, whereas the rest of the genome may be 'porous' to gene flow (i.e. subject to allele exchange between diverging species) (Wu 2001; Wu & Ting 2004; Lexer & Widmer 2008). This conceptually overlaps with ecological speciation, in which divergent selection by the (biotic or abiotic) environment drives species divergence, a process which may in principle proceed in the presence of gene flow (Nosil 2012; Rundle & Nosil 2005; Smadja & Butlin 2011; but see Cruickshank & Hahn 2014). It has been proposed that the overall accumulation of genomic divergence between incipient species best be seen as a 'speciation continuum', which proceeds from a genic stage with direct selection on few loci, via 'divergence hitchhiking' (Via & West 2008) and 'genome hitchhiking' to completed speciation and postspeciation genome divergence (Feder *et al.* 2012; Nosil 2012). However, the question of the relative importance of these postulated phases of speciation is debated, and it is therefore important to assess, in as many cases as possible, the phase of genomic divergence that has been reached in any given case of ecological speciation (Feder *et al.* 2012; Nosil 2012).

Early stages of speciation may allow the identification of the initial barriers that caused species divergence, along with their genomic architecture (Rundle & Nosil 2005). Likewise, incipient speciation may shed light on the order in which different reproductive barriers evolve (Rundle & Nosil 2005; Lexer & Widmer 2008). It has been postulated that the evolution of premating barriers, such as pollinator-mediated RI, often precedes the emergence of postmating barriers (Coyne & Orr 2004; Lowry *et al.* 2008). Moreover, such early evolving barriers probably involve alleles of large effect, the action of which can later be modified by minor-effect alleles (epistasis) as species divergence continues (Coyne & Orr 2004; Widmer *et al.* 2009). Putative cases of pollinator-driven speciation (Johnson 2006; Schiestl 2012), in which divergence of plant species is due to strong pollinator-mediated RI (floral isolation; Schiestl & Schlüter 2009), may therefore be prime candidates for genic ecological speciation processes, especially where the genetic control of pollinator attraction is simple (Schlüter *et al.* 2011b; Peakall & Whitehead 2014).

Sexually deceptive plants achieve high pollinator specificity by mimicry of pollinator females (Schiestl 2005) and provide good study systems for ecological speciation and the evolution of RI (Peakall *et al.* 2010; Schlüter *et al.* 2011a; Xu *et al.* 2012b; Peakall & Whitehead 2014). In the best-studied cases of sexual deception by plants, namely the orchid genera *Ophrys* in the Mediterranean and *Chiloglottis* in Australia, speciation is often associated with pollinator shifts (Paulus & Gack 1990; Peakall *et al.* 2010; Ayasse *et al.* 2011; Xu *et al.* 2012b). In particular, *Ophrys* bears all the hallmarks of ecological speciation (*sensu* Nosil 2012), namely (i) divergent selection and (ii) RI due to pollinators, and (iii) a genetic mechanism linking these two (Schlüter & Schiestl 2008; Xu *et al.* 2012b). Both *Ophrys* and *Chiloglottis* attract pollinators by chemical mimicry of insect sex pheromones (Schiestl 2005; Schlüter & Schiestl 2008; Ayasse *et al.* 2011). In *Ophrys*, a blend of cuticular hydrocarbons, especially alkanes and alkenes, is responsible for pollinator attraction (Schlüter & Schiestl 2008; Ayasse *et al.* 2011), and simple genetic changes affecting hydrocarbon biosynthesis can have a drastic effect on pollination; simple mutations may therefore lead to pollinator shifts, floral isolation and speciation (Schlüter *et al.* 2011b; Xu *et al.* 2012a,b).

This study investigates four closely related *Ophrys* species (*O. exaltata*, *O. garganica*, *O. incubacea* and *O. sphegodes*), among which the published literature indicates a lack of pollinator sharing that is expected to result in floral isolation (Paulus & Gack 1990; Gaskett 2011). Previous work suggests low levels of genetic differentiation, but odour differentiation and strong floral isolation (Soliva & Widmer 2003; Mant *et al.* 2005; Xu *et al.* 2011; Breitkopf *et al.* 2013), and has revealed candidate genes for pollinator attraction (Schlüter *et al.* 2011b; Xu *et al.* 2012a; Sedeek *et al.* 2013). However, as those studies examined different subsets of species, a complete picture of (genetic and phenotypic) species differences and reproductive barriers in this species group is still lacking. Moreover, previous population genetic studies (Soliva & Widmer 2003; Mant *et al.* 2005; Xu *et al.* 2011; Breitkopf *et al.* 2013) were limited by a small number of genetic markers and hence do not provide detailed information on the genomics of species divergence. This study therefore aims to provide a comprehensive picture of species and speciation in *Ophrys* by addressing the following questions: (i) are the study species reproductively isolated and if so, due to which reproductive barriers; (ii) to what extent are the species phenotypically differentiated and distinct; (iii) what proportion of the genome is associated with species differences (i.e. which stage of divergence in the speciation continuum has been reached?); and (iv) can genes associated with species divergence be identified?

Materials and methods

Study species and plant material

This study examines sympatric populations of closely related *Ophrys* species (Fig. S1, Supporting information in Appendix S3), which are pollinated by sexual

deception of male solitary bees: *O. exaltata* subsp. *archipelagi* (Gölz & H.R. Reinhard) Del Prete, *O. garganica* Nelson ex O. & E. Danesch, *O. incubacea* Bianca and *O. sphegodes* Miller (abbreviated Exa, Gar, Inc and Sph, respectively). They are pollinated by *Colletes cunicularius* (Linnaeus 1761), *Andrena pilipes* Fabricius 1781 (syn. A. carbonaria), *Andrena morio* Brullé 1832 and *Andrena nigroaenea* (Kirby 1802), respectively (Paulus & Gack 1990; Xu *et al.* 2011). Initial species identification in the field was based upon floral morphology, and photographs and odour samples were taken for later re-assessment wherever possible. As a previous study (Xu *et al.* 2011) had collected a considerable amount of data (particularly on RI) on Exa, Gar and Sph, the sampling of this study prioritized data collection that enabled incorporation of Inc into the data set. All study species co-occur in the Gargano area of southern Italy and coflower in spring (March–May); plants of all species occur in close proximity to each other and well within the expected distance of pollinator-mediated pollen transfer (see Peakall & Schiestl 2004; Xu *et al.* 2011). Field experiments were performed at populations near Marina di Lesina (MDL; previously used by Xu *et al.* 2011) and Mattinata (MTT; N41.7349°, E16.1055°; chosen because *O. incubacea* appeared commoner there). To achieve a more comprehensive genetic sampling, additional populations in the area were sampled, namely San Nicandro Garganico (SNG; N41.8063°, E15.5075°), Cagnano Varano (CGV; N41.8476°, E15.6955°) and Capoiale (CAP; see Xu *et al.* 2011). The four study species occur in mosaic sympatry (*sensu* Mallet *et al.* 2009) at these sites (map in Fig. S2, Supporting information), each being sympatric for at least three of the four study species; in particular, Inc was not observed at populations CAP and SNG, and Exa was absent from CGV and MTT.

Reproductive isolation

Different components of premating and postmating RI between *Ophrys* species were quantified as detailed in Appendix S3 (sample sizes: Table S1, Supporting information). Briefly, for premating RI, flower phenology and pollinator-mediated floral isolation (and pollination success) were assessed; floral isolation was measured by direct tracking of pollen transfer (see Xu *et al.* 2011). Assuming absence of other prezygotic barriers, the frequency of interspecific pollen in the pollen pool will depend exclusively upon the relative frequency of flowers of different species. Therefore, following Martin & Willis (2007), pairwise temporal RI indices for the three studied species were calculated based on the frequencies of flowers of the different species during their flowering seasons. For this calculation, we used the spreadsheet provided in Table S3 of Lowry *et al.* (2008). To evaluate potential postmating barriers, ploidy was measured and crossing experiments conducted to determine seed capsule formation (fruit set) and the proportion of developed (potentially viable) embryos in seeds (as in Scopece *et al.* 2007; Xu *et al.* 2011). As orchid female gametophyte development is triggered by pollination (Zhang & O'Neill 1993), fruit set is considered to be an estimate of prezygotic postmating RI. These data were used to (i) estimate temporal and floral RI, and pre- and postzygotic postmating RI and (ii) to calculate total pre-/postmating RI and (iii) the contribution of different barriers to overall RI, as described previously (Scopece *et al.* 2007).

Phenotypic analysis of floral traits

Floral traits potentially contributing to pollinator-mediated RI, particularly flower size, morphology, colour, odour and speculum shape, were evaluated as described in detail in Appendix S3 (sample sizes in Table S1, Supporting information). In brief, flower size, colour and speculum shape were measured from photographs, and spectral reflectance data were collected and mapped into the honeybee colour-space (as in Chittka & Kevan 2005); three-dimensional (3D) flower morphology was determined by micro-computed tomography (μCT) (see Staedler *et al.* 2013; Table S2, Supporting information); floral odour was analysed by gas chromatography as described previously (Xu *et al.* 2011).

Population genomic data generation and analysis

Population genomic data for the study species were obtained using a previously published genotyping-by-sequencing (GBS) protocol (Elshire *et al.* 2011) with modifications as detailed in Appendix S3 (Supporting information). Paired-end Illumina HiSeq 2000 sequence data (two lanes; 32 + 96 samples; Tables S1, S3 and S4, Supporting information) from 127 plant individuals and a replicate were analysed as described in Supporting information. Briefly, after demultiplexing, single nucleotide polymorphisms (SNPs) were identified using STACKS (Catchen *et al.* 2011), population structure analysed by STRUCTURE (Falush *et al.* 2003) and principal coordinate analysis (PCoA), and genome scans for F_{ST} outlier loci conducted using FDIST2 and BAYESCAN (Beaumont & Nichols 1996; Foll & Gaggiotti 2008). The latter analysis was conducted both as a global analysis and between species pairs, identifying outlier loci that were repeatedly detected in either two or three sympatric populations. Linkage disequilibrium (LD) was compared between the entire data set and the outlier loci.

Results

Flowering phenology and temporal isolation

Flowering time was recorded in 2012 using a total of 1061 flower observations in the field, the lower number of observations for Inc (Tables S1 and S5, Supporting information) reflecting the lower abundance of this species at the monitored populations. Phenology data from four patches were combined and showed a similar trend as (qualitative) data from 2011. Exa and Sph flowering peaked a week before Gar and Inc, the number of flowers declining thereafter, with the exception of Exa which displayed a 2-week flowering peak. Exa and Sph flowering time overlapped by 81%, Gar and Inc by 94%, although ≥46% of flowers of all species were open simultaneously (Table S6, Supporting information). Corresponding mean $RI_{phenology}$ ranged from 0.025 to 0.5 (Table 1), with an average of 0.26. Temporal RI was weakest between Gar/Inc and Exa/Sph and strongest between Sph/Gar and Sph/Inc.

Floral isolation and pollination success

Average pollination success was markedly different for the species between the years (Fig. S3, Supporting information); pollination success for Sph was higher than for Gar or Inc in 2011, whereas the opposite was true in 2012. In total, 2500 stained flowers were used in pollen tracking plot experiments. Of 108 pollinated flowers (4.32%), the majority received only unstained massulae (from newly opened flowers on the experimental plants or from the surrounding naturally occurring plants). In total, 46 flowers received stained massulae (Table S7, Supporting information), 45 of which indicated within-species transfers (15 Sph, 12 Gar, 18 Inc), and one interspecies pollination event (Table S7, Supporting information). Specifically, one Gar flower had received massulae from Inc. The mean floral isolation index (RI_{floral}) was 0.96 between Gar and Inc and one for all other tested species combinations (Table 1). Likewise, complete floral isolation ($RI_{floral} = 1$) was previously observed for all species combinations involving Exa, Gar and Sph (Xu *et al.* 2011).

Ploidy level

All *Ophrys* species were clearly diploid, consistent with previous reports (D'Emerico *et al.* 2005; Xu *et al.* 2011); only a single Sph individual (population CAP) appeared to be triploid (Fig. S4A, Supporting information). A small but significant difference in nuclear DNA A/T content between Inc and Gar was observed, consistent with karyological data (D'Emerico *et al.* 2005).

Postmating isolation

After hand pollination, 90% of intraspecies (Inc × Inc) and 64–94% of interspecies crosses developed capsules (Table S8, Fig. S4B, Supporting information). Similarly, 68% of intra- and 63–76% of interspecies seeds contained embryos (therefore considered viable; Table S8, Fig. S4C, Supporting information). Although mean fruit set for Inc × Gar and Inc × Exa appeared comparatively low, this finding was not significant. Neither was there a significant difference in the proportion of viable seeds in any cross. Fruit set and viable seed proportion between intra- and interspecies crosses were used to estimate the strength of postmating prezygotic and postzygotic barriers as described previously (Scopece *et al.* 2007); negative values were assumed to indicate zero. Postmating isolation was largely weak or absent, although the $RI_{prezygotic}$ values for the aforementioned Inc × Gar and Inc × Exa crosses were also comparatively high (Table 1). Weak postmating isolation is in line with previous data for other species combinations (Xu *et al.* 2011), which were re-analysed and included in Table 1 for comparison.

Overall RI

In nature, species may be isolated by several barriers, which work in a sequential manner; each barrier therefore also affects the action of subsequent barriers. Data from this study and Xu *et al.* (2011) were used to estimate the absolute strength and the contribution of each barrier to total RI (Table 1). The early acting barriers (phenology and floral isolation) were strong and responsible for the majority of total isolation. Combined premating RI was 1 (i.e. complete) for all species comparisons except for Gar and Inc ($RI_{premating} = 0.96$). Combined postmating RI was generally weak (max. 0.19 for Exa/Gar); however, in all cases, postmating barriers effectively did not contribute to overall RI.

Flower morphology and size

Comparison of flower size revealed significant interspecies differences (Fig. S5, Supporting information): Exa and Gar differed from Inc in total flower width, width of the stigmatic cavity and the position of the widest point of the labellum. Gar and Inc had a wider labellum than Sph and Exa; Exa had the longest flowers; and Sph had a smaller labellum area than Gar and Inc. The largest overall differences were in labellum 'slenderness' (length/width ratio; Fig. S5I, Supporting information), which was highest for Exa, medium for Sph and Inc, and lowest for Gar.

Table 1 Overview of reproductive barriers, showing RI indices (where 0 indicates no RI and 1 complete RI), combined pre-/postmating RI indices and contributions to total RI. RI indices are shown for premating barriers, namely temporal isolation ($RI_{phenology}$) and floral isolation (RI_{floral}), and pre- and postzygotic postmating RI ($RI_{prezygotic}$ and $RI_{postzygotic}$). The combined pre- and postmating barriers are listed under $RI_{premating}$ and $RI_{postmating}$, respectively. Additionally, the contributions of floral isolation and combined postmating isolation to total RI (RI_{total}) are presented. Negative values are interpreted to be zero, that is absence of RI. Exa, *Ophrys exaltata;* Gar, *Ophrys garganica*; Inc, *Ophrys incubacea*; and Sph, *Ophrys sphegodes*

Comparison				Premating barriers		Postmating barriers		Combined barriers		Contribution to total RI		Total RI
Species A	Species B	Fruit set %	Seed viability %	$RI_{phenology}$	RI_{floral}[1]	$RI_{prezygotic}$	$RI_{postzygotic}$[2]	$RI_{premating}$	$RI_{postmating}$	RI_{floral}	$RI_{postmating}$	RI_{total}
Inc	Sph	86.7	75.9	0.53	1	−0.05	−0.45	1	−0.53	0.47	0	1
Sph	Inc	93.8	63.9	0.25	1	−0.14	−0.22	1	−0.39	0.75	0	1
Average Inc/Sph		90.3	69.9	0.39	1	−0.09	−0.34	1	−0.46	0.61	0	1
Inc	Gar	64.3	71.6	0.04	1	0.29	−0.18	1	0.16	0.96	0	1
Gar	Inc	88.9	65.4	0.01	0.917	0.01	−0.08	0.927	−0.06	0.91	0.001	0.928
Average Inc/Gar		76.6	68.5	0.03	0.959	0.15	−0.13	0.964	0.05	0.93	0.001	0.964
Inc	Exa	66.7	63.3	0.44	n.d.	0.30	−0.12	n/a	0.22	n/a	n/a	n/a
Exa	Inc	86.7	74.1	0.13	n.d.	0.09	−0.31	n/a	−0.20	n/a	n/a	n/a
Average Inc/Exa		76.7	68.7	0.29	n.d.	0.19	−0.21	n/a	0.01	n/a	n/a	n/a
Sph	Gar	100.0	55.4	0.55	1	−0.21	−0.22	1	−0.48	0.45	0	1
Gar	Sph	100.0	40.2	0.45	1	−0.21	0.11	1	−0.08	0.55	0	1
Average Sph/Gar		100.0	47.8	0.50	1	−0.21	−0.05	1	−0.28	0.50	0	1
Sph	Exa	100.0	49.3	0.14	1	−0.14	−0.20	1	−0.37	0.86	0	1
Exa	Sph	100.0	50.6	0.11	1	−0.14	−0.23	1	−0.41	0.89	0	1
Average Sph/Exa		100.0	50.0	0.13	1	−0.14	−0.21	1	−0.39	0.88	0	1
Gar	Exa	90.0	39.7	0.32	1	0.05	0.20	1	0.24	0.68	0	1
Exa	Gar	90.0	45.1	0.35	1	0.05	0.09	1	0.14	0.65	0	1
Average Gar/Exa		90.0	42.4	0.34	1	0.05	0.14	1	0.19	0.67	0	1

[1]n.d., not determined; RI, [2]Calculated as $RI_{postmating} = RI_{prezygotic} + (1 - RI_{prezygotic}) RI_{postzygotic}$; reproductive isolation.

A more detailed view of floral morphology was obtained by μCT scanning of 56 flowers (Figs 1 and S6 to S8; movies: Files S1 to S4, Supporting information). Figure S6 (Supporting information) presents average and 'extreme' reconstructed 3D flowers for each species; the geometric landmarks used for analysis are depicted in Fig. S7A (Supporting information). Canonical variate analysis (CVA) based on these landmark data showed the four species to be clearly morphologically distinct (Fig. 1). Within the reconstructed morphometric shape space of the pooled four-species data set, the first two principal components (PCs) of shape variation both represent relative changes of position of the viscidia with respect to the labellum. Along PC1 (36% variance explained), the viscidia move more or less in parallel to the plane of the labellum, whereas along PC2 (30% variance explained), the angle between column and labellum varies, thereby bringing the viscidia closer to, or farther from the labellum (Fig. S7B, Supporting information). These two relatively simple deformations allow a large flexibility in the positioning of the viscidia with respect to the labellum. A striking detail that differs between Inc and the other species are *Ophrys incubacea*'s longer trichomes on the side of the labellum (Fig. S8, Supporting information; cf. Cortis *et al*. 2009).

Flower colour

Differences in colour, as perceived by a human observer, were quantified from colour-adjusted flower photographs. In the field, flower labella of Gar and Inc appear darker than those of Sph, and this was reflected in a significant difference in labellum R[ed], G[reen] and B[lue] values (Fig. S9, Supporting information). Likewise, the whitish perigon (see Fig. S1, Supporting information) of Exa, unsurprisingly, constituted a significant difference; other differences were minor. Spectral reflectances of Sph and Inc flower parts were collected and mapped into a colour hexagon representing the bee visual space (Fig. S10, Supporting information). This analysis suggests that while the dark labella of both species are uncoloured (achromatic) to their pollinators, their specula provide a UV/UV-blue colour signal; perigon parts appeared green.

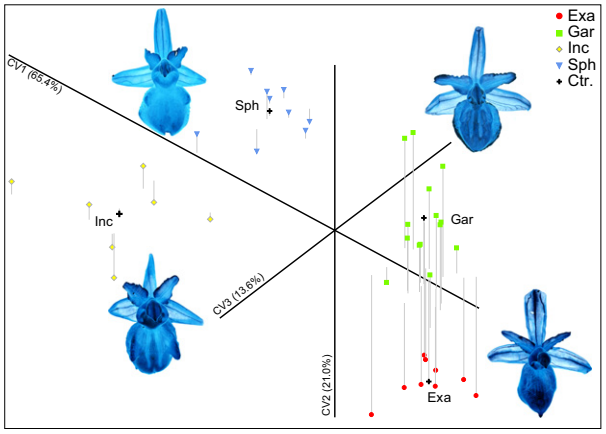

Fig. 1 CVA scatter plot of 3D morphological data, showing separation of species and example μCT flower scans. Exa, *Ophrys exaltata;* Gar, *Ophrys garganica*; Inc, *Ophrys incubacea*; Sph, *Ophrys sphegodes*; and Ctr., group centroid for each species. The three axes explain 100% of morphological variation. The four species appear morphologically distinct. μCT, micro-computed tomography; CVA, canonical variate analysis.

Speculum shape

The colour difference between speculum and the rest of the labellum implies that speculum shape should be detectable by pollinators. Speculum shape is highly variable (Fig. S11, Supporting information), but elliptic Fourier-descriptor-based PC analysis (EFD-PCA) revealed no clear shape differences between species (Fig. S11B, Supporting information). However, species differed significantly in the proportion of labellum area taken up by the speculum (Fig. S11F, Supporting information; highest: Exa, medium: Gar, lowest: Inc). Speculum pattern complexity (number reflective patches/nonreflective holes) differed between species, with the largest difference between Sph and Inc (Fig. S11G, H, Supporting information).

Floral odour

Ophrys floral odour has been well-studied because of its importance for pollinator attraction. Floral odours of Exa, Gar and Sph in the study area had previously been investigated (Mant *et al*. 2005; Xu *et al*. 2011); our data are in agreement with previous findings. Here, we analysed 24 compounds, comprising six *n*-alkanes with different chain lengths (C_{21}-C_{31}), 16 alkenes with different chain lengths (C_{21}-C_{31}) and double-bond positions [(*Z*)7-, (*Z*)9-, and (*Z*)12-alkenes], and two esters (2-nonyl palmitate and 2-nonyl oleate) (Fig. S12, Supporting information). Amongst these, esters were only detectable in Gar and, at low levels, in Inc (Fig. S12F, Supporting information). Sph had the lowest overall alkene level. Gar, Sph and Inc had higher levels of 9- and 12-alkenes than Exa, and conversely, Exa had the highest proportion of 7-alkenes, consistent with previous findings (Schlüter *et al*. 2011b; Xu *et al*. 2012a), Inc also containing a high proportion of 7-alkenes as compared to Gar/Sph (Fig. 2A). The latter two species had a similar alkene double-bond composition and primarily differed in carbon chain length, Gar featuring the longest hydrocarbons on average and Exa the shortest (Fig. 2B). Linear discriminant analysis (LDA) based on all identified

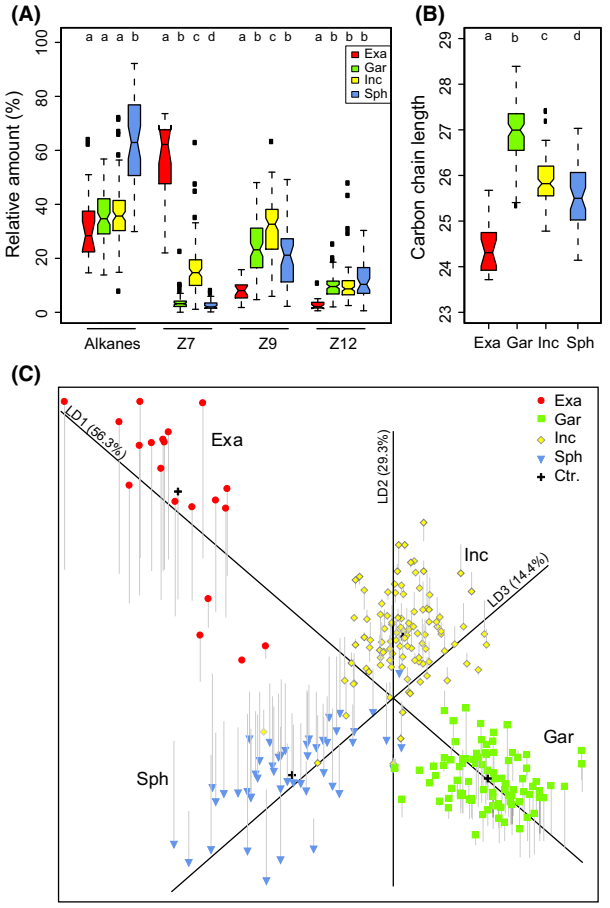

Fig. 2 Species differences in flower odour. Differences in (A) hydrocarbon double-bond composition and (B) chain length. Panel (A) shows the relative proportion of alkanes and alkene double-bond classes in floral odour extracts and panel (B) mean chain length for all alkanes and alkenes. Different letters indicate significant differences ($P < 0.01$; Wilcoxon rank-sum test). (C) LDA scatter plot of floral odour, indicating percentage of trace explained by the axes (total: 100%). Exa, *Ophrys exaltata;* Gar, *Ophrys garganica*; Inc, *Ophrys incubacea*; Sph, *Ophrys sphegodes*; and Ctr., group centroid for each species. LDA, linear discriminant analysis.

compounds clearly separated the four species into four corresponding clusters (Fig. 2C).

Population genomic data

Genotyping-by-sequencing data were collected for 127 biological individuals, one of which was repeated to assess reproducibility. In total, 141 million assignable paired-end reads were obtained, and 67 752 (raw: 2 538 827) unique sequence tags with 158 871 (raw: 707 059) SNPs identified. Unfortunately, the number of reads per sample was highly uneven (especially in the 96-plex run), and ranged from 43 345 to 14 738 130, with SNPs/sample ranging from 550 to 78 729 (Table S4, Supporting information). Nevertheless, genotyping repeatability was high for loci sampled twice among technical replicates (99.2%; 63 conflicting genotypes out of 3037 SNPs at coverage ≥10). Overall, 2 823 364 genotype calls were observed at ≥10 read coverage (157–68 306 calls/individual; mean 22 058). Uneven sequence coverage implied a potentially large amount of missing data in STRUCTURE input files (88–119 individuals at 585–4807 loci), the effect of which was investigated by varying the stringency of data filtering. However, best K estimates were highly inconsistent among analyses (K = 2–15) and showed no apparent patterns; the analyses generally suggested a lack of strong genetic structure in the study species. Often, when clusters corresponding to the study species were identified (e.g. Fig. S13B, Supporting information), they only accounted for a small proportion of the genome, whereas the majority of markers were typically shared among species. Exa and Gar were most frequently separated from other species; Inc was frequently split into two groups.

Principal coordinate analysis revealed a weak but discernible clustering according to species identity (Fig. 3), the first three axes however only explaining 19.6% of variation. This analysis qualitatively confirmed STRUCTURE analysis, including the suggestion of genetic structure within Inc and the presence of a few Sph/Inc individuals that were not clearly assignable to a species. Although improbable, we note that sample misidentification cannot be completely excluded. Like STRUCTURE, mean F_{ST}-based analysis also suggested Inc and Sph to be the most similar, and Exa the most dissimilar study species (Fig. S13, Supporting information), although no in-group relationship had any bootstrap support. Overall, although genetic structure was broadly consistent with species groups, species were very similar and only weakly differentiated.

Outlier analysis (FDIST2) on 95 079 SNP markers typically revealed 9–13 tags that both contained global F_{ST} outliers and repeated outliers in pairwise species comparisons in at least three sympatric populations (21–197 replicated in ≥2 populations) (Figs 4A and S13F, Supporting information). The stringent global BAYESCAN analysis of the same 95 079 markers revealed 19 (0.02%) F_{ST} outlier SNPs (in 17 tags) at FDR <0.05 (Fig. 4B). A significantly elevated extent of LD was found for all sets of outlier loci (all $P < 0.0005$; Fig. 4C), and LD in pairwise species outliers increased with divergence (Fig. S14, Supporting information). As expected for GBS markers sampled from a large genome, BLAST searches of tags identified by threefold population replication and/or BAYESCAN analysis (Appendices S1 and S2, Supporting information) did not identify confirmed (i.e. not 'predicted') or function-

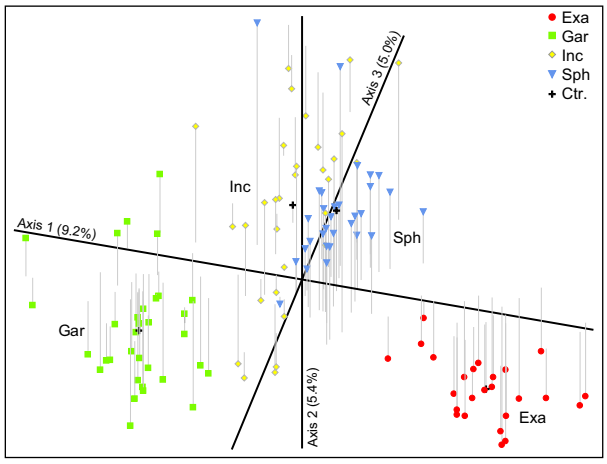

Fig. 3 PCoA scatter plot showing genetic separation of study species, based on pairwise genetic distances calculated from genome-wide high-throughput sequencing data. The percentage of variation along the first three PCoA axes (totalling 19.6%) is indicated. Exa, *Ophrys exaltata;* Gar, *Ophrys garganica*; Inc, *Ophrys incubacea*; Sph, *Ophrys sphegodes*; and Ctr., group centroid for each species.

ally annotated gene sequences, except for three sequence tags. Two independent sequence tags (tags 12 662 and 33 591) matched the *Arabidopsis thaliana Vacuolar Protein Sorting 45* (*VPS45*; AT1G77140.1) gene (Fig. S13G, Supporting information), and one tag (33 833) matched the *A. thaliana ECERIFERUM 1* (*CER1*; AT1G02205.3) gene (Fig. 4D). Both outlier SNPs found in *VPS45*, like the outlier SNP in *CER1*, were synonymous. However, the SNP in *CER1* was placed at a predicted exon/intron junction (Fig. 4D).

Discussion

Reproductive barriers

The emergence of RI between populations is critical for speciation. All *Ophrys* study species were found to be strongly reproductively isolated from each other, as was expected given their different pollinators. Our survey of reproductive barriers demonstrates the importance of premating RI (Table 1), that is temporal and floral RI. Flower phenology differences between early (Exa/Sph) and late-flowering species (Gar/Inc) accounted for considerable temporal RI (Table 1), for instance between the most closely related species pair (Sph/Inc). As efficient gamete transfer requires the presence of pollinators, one may expect pollinator-imposed selection to synchronize flower phenology with pollinator flying times. Interestingly, pollination success was higher in 2011 for the early flowering spe-

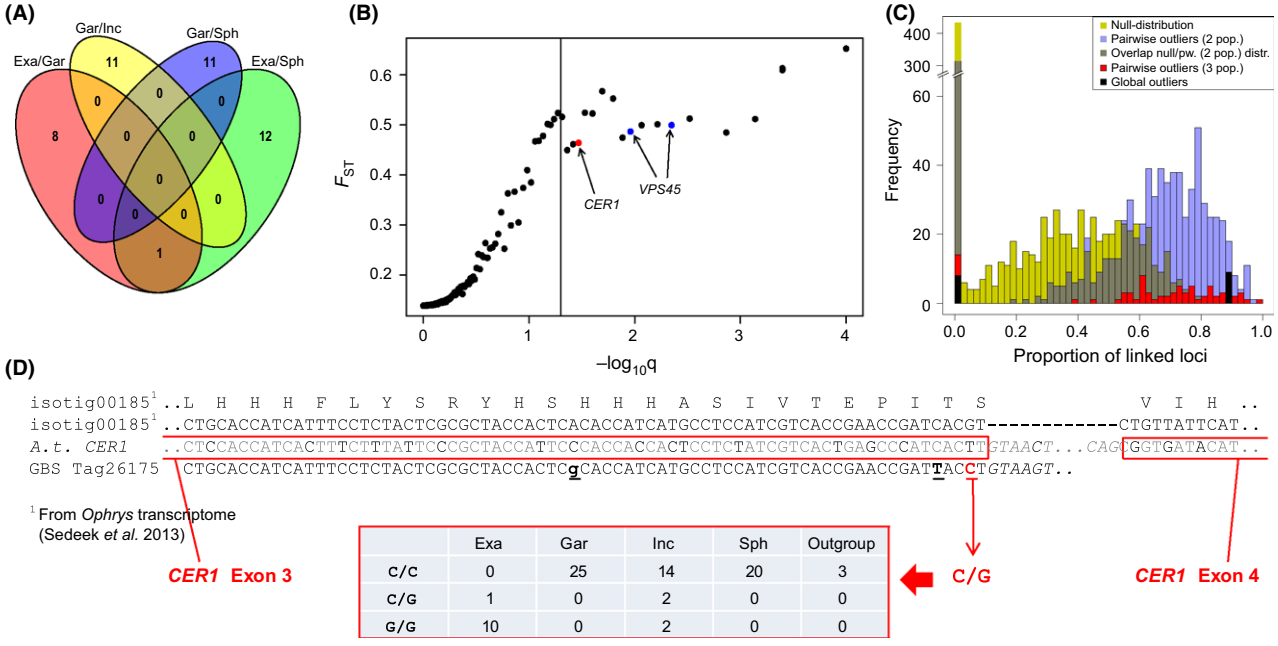

Fig. 4 Genome scan results. (A) Venn diagram showing overlap in of F_{ST} outlier tags, replicated in three populations, between pairwise species comparisons. (B) Global BayeScan genome scan (threshold at FDR = 0.05) indicating outlier SNPs in *CER1* and *VPS45*; F_{ST} is plotted against $\log_{10} q$, which is a *P*-value analogue that takes multiple testing into account. (C) Comparison of LD distributions, showing the expected LD null distribution among all tags vs. the LD distributions among different set of outliers. (D) Detail of the outlier F_{ST} SNP in the *CER1* locus, placed at a predicted intron/exon boundary (5′ splice site). The distribution of genotypes at this SNP is indicated. Exa, *Ophrys exaltata;* Gar, *Ophrys garganica*; Inc, *Ophrys incubacea*; Sph, *Ophrys sphegodes*; and Outgroup, *Ophrys insectifera s.l.*

cies (Sph), whereas in 2012, the late-flowering species (Gar, Inc) had higher success (Fig. S3, Supporting information), suggesting a difference in the abundances of the respective pollinators. This is consistent with the notion of a spatiotemporal pollinator mosaic (Johnson 2006; Breitkopf *et al.* 2013).

In absolute terms, the strongest barrier was floral isolation, which was complete ($RI_{floral} = 1$) in all measured cases (including those published by Xu *et al.* 2011) except for Gar/Inc, where a single interspecies pollen transfer was observed. Despite this, floral isolation was also very high ($RI_{floral} = 0.96$) in this case. Floral trait-based RI is mediated by pollinator behaviour, which in *Ophrys* has been shown to be largely determined by floral odour that acts as a pseudo-pheromone (Schiestl *et al.* 2000; Schlüter & Schiestl 2008). The finding of strong floral isolation is entirely consistent with previous data on sexually deceptive orchids (Schlüter *et al.* 2007, 2009; Peakall *et al.* 2010; Xu *et al.* 2011; Peakall & Whitehead 2014; Whitehead & Peakall 2014), but we can now conclude that strong floral isolation is consistent over several seasons (Sph/Gar: 4 years) and locations (cf. Xu *et al.* 2011). Strong, constant floral isolation supports the role of pollinators in species divergence, and notably, such strong floral isolation may also be found in more generalized pollination systems (Schiestl & Schlüter 2009 and references therein). Although floral isolation is clearly the most effective barrier, the finding of considerable temporal RI—which acts before floral RI—suggests that flower phenology may play an important role in speciation in sexually deceptive orchids that merits closer attention (cf. Xu *et al.* 2012b).

There were only small genome size differences between Gar and Inc, and postmating barriers were effectively absent between species (Table 1, Fig. S4, Supporting information and Xu *et al.* 2011). However, even with stronger postmating RI, the large premating RI indices would certainly have rendered the contribution of postmating to total RI negligible when species are sympatric. Nonetheless, this also indicates that rare interspecies pollination events are expected to result in gene flow between the species. Overall, our data suggest that early acting premating barriers, but not later-acting barriers, have evolved between the four study species, similar to findings in the distantly related Australian *Chiloglottis* (Whitehead & Peakall 2014) and other plant groups (Rieseberg & Willis 2007; Lowry *et al.* 2008; Widmer *et al.* 2009). As less closely related *Ophrys* species have been shown to build up postmating barriers (Cortis *et al.* 2009), this supports the idea that premating barriers may generally evolve earlier (Coyne & Orr 1989; Widmer *et al.* 2009).

Floral trait differences

Floral isolation, the strongest barrier between the study species, is mediated by floral traits. We investigated a suite of floral traits that may potentially be involved in floral isolation for differences between species. Given the apparent variability of *Ophrys* flowers in the field (Fig. S1, Supporting information), previous investigations of our study species (e.g. Mant *et al.* 2005; Xu *et al.* 2011) largely ignored variation in floral shape and morphology. Study species showed only minor floral size differences; labellum length may contribute to floral isolation because of the need to match pollinator body size (Paulus & Gack 1990; Paulus 2006; Vereecken 2009); however, no significant labellum length differences were found between species. Nonetheless, 3D morphometric analysis revealed a clear separation of species in overall flower shape and suggests that an important component of shape variation is the positioning of the viscidia with respect to the labellum, thereby potentially affecting placement of pollinia on the pollinators (Fig. S7B, Supporting information). While in *Ophrys incubacea* the lateral protrusions of the labellum are conspicuously developed and trichomes are longer than in other species (Fig. S8, Supporting information), the functional significance of these features remains unclear. Although trichome direction has been implied in mechanical floral isolation between *Ophrys* groups pollinated by the same insects by pollinaria attached to either the head or the abdomen (e.g. Ågren *et al.* 1984), this may constitute only a relatively weak barrier (Cortis *et al.* 2009). Likewise, although potentially preventing interspecies pollination, slight differences in pollinarium placement would be expected to contribute only modestly to overall floral isolation.

Colour cues may aid specific pollinator attraction and thereby contribute to floral isolation. One may expect flowers (i) to share features of the mimicked female insect and (ii) to be distinguishable from other species by their pollinators. Among the three *Andrena*-pollinated study species (Gar, Inc, Sph), labellum coloration of Inc and Gar was markedly blacker than of Sph. This intriguingly mirrors their pollinators' black (Inc, Gar) and brownish (Sph) body colour, suggesting that this trait may be under divergent selection by pollinators. As labella of both Inc and Sph appear achromatic to pollinators (Fig. S10E, Supporting information), we hypothesize that pollinator-mediated selection on labellum coloration acts on the level of brightness rather than colour (cf. Renoult *et al.* 2013). A coloured perigon has been shown to increase pollinator visitation in *Ophrys heldreichii* (e.g. Spaethe *et al.* 2007; Streinzer *et al.* 2010). However, this is less likely in the species studied here, given that no effect of perigon coloration on attractiveness to pollinators was

found in (the more closely related) *Ophrys arachnitiformis* (Vereecken & Schiestl 2009). In contrast to overall labellum coloration, the speculum likely constitutes a visual colour signal (Fig. S10E, Supporting information), implying that its highly variable shape is detectable by pollinators against the uncoloured labellum background. No clear interspecies differences in speculum shape, however, were found between species, suggesting that speculum shape does not aid the between-species discrimination of flowers by pollinators. However, large between-individual/within-species shape variation may well help pollinators to learn to avoid individual 'false females' they have already visited, thereby promoting outcrossing (Ayasse *et al.* 2000). Although species differences in speculum complexity and relative area were detected, their functional significance remains unclear. Taken together, labellum coloration may well reflect selection by pollinators, but overall evidence for the contribution of colour signals to floral isolation is limited.

Floral odour underlies the chemical mimicry of pollinators' sex pheromones in sexually deceptive orchids and thus is a key contributor to floral isolation (Schiestl *et al.* 2000; Schiestl & Schlüter 2009; Stökl *et al.* 2009; Ayasse *et al.* 2011; Xu *et al.* 2011, 2012b; Peakall & Whitehead 2014). Clear differences in floral odour were observed, the study species differing in carbon chain length and alkene double-bond position (Fig. 2). This is consistent with previous data documenting floral odour differentiation between Exa, Gar and Sph (Mant *et al.* 2005; Xu *et al.* 2011). Similarly, *O. incubacea* produces a distinct odour blend, which interestingly includes 7-alkenes (only present in small amounts in Gar and Sph), albeit at lower levels than Exa. Such alkene double-bond differences have previously been linked to changes at desaturase-encoding genes (Schlüter *et al.* 2011b; Xu *et al.* 2012a). Besides hydrocarbons, two esters were included in our odour analysis because of their high levels in Gar. These two esters were also present at low levels in Inc but undetectable in Sph and Exa. Unfortunately, the sex pheromone of *Ophrys garganica*'s pollinator, *Andrena pilipes*, has not yet been determined and the function of these esters requires experimental testing.

Overall, several traits may contribute to floral isolation, but phenotypic differences suggest a role of floral odour, morphology and perhaps even coloration. Among these, based upon previous data from our study group (e.g. Schiestl *et al.* 2000; Mant *et al.* 2005; Vereecken & Schiestl 2009; Xu *et al.* 2012a), floral odour is certainly the trait with the largest contribution and with pronounced species differences. This is consistent with findings in *Chiloglottis*, where, however, no clear morphological differences were found (Peakall & Whitehead 2014). Hence, one may hypothesize that changes in floral odour precede changes in other floral traits during species divergence. Likewise, one may therefore expect genetic changes affecting floral isolation (odour) and/or temporal isolation (flowering time) to initiate species divergence. This prediction is not restricted to the study species but may more broadly be applied to other plants which diverge via the evolution of strong premating barriers.

Genomics of species divergence

A genome-wide set of SNP markers revealed all four study species to be only weakly differentiated, with the majority of allelic variation shared among all species. Weak genetic interspecific differentiation mirrors earlier population genetic reports that utilized a small number of microsatellite or AFLP markers (Soliva & Widmer 2003; Mant *et al.* 2005; Xu *et al.* 2011; Breitkopf *et al.* 2013). Although initially interpreted as high levels of gene flow among distinct species in secondary contact (Soliva & Widmer 2003), those genetic results may also be explained by shared ancestral polymorphism and recent species divergence (Mant *et al.* 2005; Schlüter *et al.* 2011a,b; Xu *et al.* 2011; Peakall & Whitehead 2014). More evidence now points in the direction of recent and perhaps ongoing speciation, specifically: (i) phylogenetic dating analysis suggests that the species group has indeed only recently diversified in the Pleistocene (Inda *et al.* 2012; Breitkopf *et al.* in press), (ii) the measurement of consistently strong RI in the field makes rampant gene flow unlikely, (iii) ancestral polymorphism has been documented in *Ophrys* (Schlüter *et al.* 2011a) and (iv) large effective population sizes, together with plant longevity (including clonal vegetative reproduction), make the maintenance of low nuclear differentiation plausible (Mant *et al.* 2005; Peakall & Whitehead 2014). Despite the large amount of shared polymorphism, the study species could largely be distinguished genetically, the separation of Inc/Sph being least clear. The latter is inferred to be the most closely related species pair, with the caveat that there was essentially no statistical support for the inference of species relationships.

Only a very small proportion of the genome (<0.05%) was repeatedly identified as more strongly differentiated between pairs of species than expected and is therefore interpreted as being associated with species divergence. These outlier loci may be expected to be linked to targets of divergent selection, for example by pollinators. A caveat to this analysis is that the GBS data only cover a small fraction of the ~10 Gbp (Leitch *et al.* 2009) genome (on average, one analysed tag every <1 kbp); however, it should be noted that the GBS method biases against sampling of methylated, highly repetitive parts of the genome (Elshire *et al.* 2011). A second caveat is that moderately divergent parts of the

genome might not be detected as outliers and that we may therefore underestimate the true extent of genomic differentiation that is linked to species divergence. It is, however, surprising that differentiation between species is so low, given that they differ in a number of phenotypic traits and consequently would be expected to differ at a number of genetic loci involved in specifying these traits. Moreover, considerable genetic structure has previously been reported for other closely related *Ophrys* species based on fewer markers (Schlüter *et al*. 2011a), making it unlikely that significant genome-wide differentiation has gone undetected in our study. It is currently unclear whether several genes involved in phenotypic differences are closely linked in the genome, for instance due to variation in chromosome structure or divergence hitchhiking. Alternatively, it is also conceivable that indeed only few genetic polymorphisms differ between species if change in several traits (many of which concern features of epidermal cells) is affected by variation in, for example, a key transcriptional regulator controlling phenotypic differences at the level of gene expression. Nonetheless, the small proportion of F_{ST} outliers suggests that the *Ophrys* species studied here have only reached an early phase of the speciation continuum, that is, either direct selection on genic targets or at best the stage of divergence hitchhiking. Moreover, elevated LD among F_{ST} outlier loci indicates that the few outlier loci identified probably stem from even fewer divergent genomic regions, among-outlier LD increasing with species divergence. The small number of outlier regions therefore certainly suggests genic rather than genome-wide patterns of divergence between *Ophrys* species and notably fits the model of species divergence proposed for other sexually deceptive orchids (Peakall & Whitehead 2014).

Genic targets of pollinator-mediated selection would be expected to include genes involved in specifying floral odour-based RI. Such odour changes may only require few genetic changes, as exemplified by the gene *SAD2* (not covered by GBS data) involved in alkene differences between Exa and Sph (Schlüter *et al.* 2011b; Xu *et al*. 2012a). The product of this gene catalyses specific double-bond insertion into precursors of alkenes, and *SAD2* alleles with different enzymatic activities are differentially expressed between the two species; *SAD2* thereby contributes to a major difference in hydrocarbon composition between Exa and Sph/Gar that demonstrably affects pollinator attraction and thus RI (Schlüter *et al*. 2011b; Xu *et al*. 2012a). In principle, other genes affecting pseudo-pheromone production (see Sedeek *et al.* 2013) may also be expected to respond to divergent selection by pollinators. Here, two annotated genes, *VPS45* and *CER1*, were identified among the GBS tags containing F_{ST} outliers. *Arabidopsis* VPS45 is an SM protein family member involved in vesicle trafficking and regulation of a SNARE complex localized on the trans-Golgi network (TGN) that is required for membrane fusion (Zouhar *et al*. 2009). Interestingly, secretion of alkanes to the plant surface by epidermal cells has recently been found to require Golgi- and TGN-mediated vesicle trafficking (McFarlane *et al.* 2014), thereby raising the possibility that VPS45 is involved in producing the adaptive, pollinator-attractive odour phenotype. Although the outlier SNPs in *VPS45* were synonymous, evidence for selection on, and functional relevance of, synonymous sites (e.g. for miRNA binding) is mounting (Gu *et al.* 2012; Lawrie *et al.* 2013). It may, however, be more likely that these SNPs are linked to other polymorphisms of functional importance. The finding of an F_{ST} outlier in *CER1* is especially striking, considering that this gene has been considered an a priori candidate for affecting pollinator attraction (Sedeek *et al.* 2013). *Arabidopsis* CER1 is an aldehyde decarbonylase acting in the last step of hydrocarbon biosynthesis (Bernard *et al*. 2012); the outlier SNP at the *Ophrys CER1* exon/intron junction may well have functional consequences by affecting pre-mRNA splicing, which may in turn result in pollination-relevant phenotypic change. It must be borne in mind that hypotheses on gene function based upon homology to *Arabidopsis* are inherently uncertain and, even if correct, that the outlier SNPs identified here may simply be linked to other, more important, polymorphisms. Although intriguing, the involvements of *VPS45* and *CER1* in pollinator-mediated species differences therefore require testing. Nevertheless, the fact that one (or perhaps even both) of the two outlier-containing genic regions has putative functions linked to odour production suggests that the outliers identified are not spurious and that these genic regions might represent direct targets of selection, as would be expected in the early genic phase of species divergence.

Conclusion

This study examines pollinator-driven ecological speciation between sexually deceptive orchids with a specialized pollination system. Species are isolated by premating barriers, namely temporal isolation and, especially, strong floral isolation. Floral odour and, to a lesser extent, flower geometry and perhaps labellum coloration probably underlie floral isolation. Population genomic data revealed the majority of polymorphism to be shared across species. Only few repeatable F_{ST} outliers were found in a genome scan, two genes with potential involvement in odour production being among the outliers. Overall, these data suggest that these species are defined mostly by *genic* rather than genome-wide differences, suggesting that *Ophrys* orchids have only reached an early stage in the speciation

continuum. We hypothesize that this also applies to other sexually deceptive orchids, where odour genes may be among the first to diverge. More generally, genic speciation may well be common in systems where subtle trait changes strongly affect prezygotic RI.

Acknowledgements

The authors are grateful for financial support by the Swiss National Science Foundation (grant 31003A_130796 to PMS) and the Georges & Antoine Claraz Foundation. We thank W. Francke for ester identification/synthesis; C. Aquino for Illumina sequencing; K. Gross, M. Sun and S. Xu for help in the field; S. Sontag for μCT sample mounting; S. Pamperl for help with landmarking; B.I. Rosc-Schlüter for assistance with image analysis; K. Gross for flow cytometry training; T. Aleksiev and S. Maffioletti for grid computing support; J. Spaethe and M. Streinzer for advice on colour/pattern analysis; J. Cai for sharing GBS data; and C. Lexer, K. Gross, C. Klingenberg and S. Xu for discussion.

References

Ågren L, Kullenberg B, Sensenbaugh T (1984) Congruences in pilosity between three species of *Ophrys* (Orchidaceae) and their hymenopteran pollinators. *Nova Acta Regiae Societatis Scientiarum Upsaliensis Ser. V:C*, **3**, 15–25.

Ayasse M, Schiestl FP, Paulus HF *et al.* (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.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B*, **263**, 1619–1626.

Bernard A, Domergue F, Pascal S *et al.* (2012) Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell*, **24**, 3106–3118.

Breitkopf H, Schlüter PM, Xu S *et al.* (2013) Pollinator shifts between *Ophrys sphegodes* populations: might adaptation to different pollinators drive population divergence? *Journal of Evolutionary Biology*, **26**, 2197–2208.

Breitkopf H, Onstein RE, Cafasso D, Schlüter PM, Cozzolino S (in press) Multiple shifts to different pollinators fuelled rapid diversification in sexually deceptive *Ophrys* orchids. *New Phytologist*.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171–182.

Chittka L, Kevan P (2005) Flower colour as advertisement. In: *Practical Pollination Biology* (eds Dafni AKP, Husband BC), pp. 157–196. Enviroquest, Cambridge, Ontario.

Cortis P, Vereecken NJ, Schiestl FP *et al.* (2009) Pollinator convergence and the nature of species' boundaries in sympatric Sardinian *Ophrys* (Orchidaceae). *Annals of Botany*, **104**, 497–506.

Coyne JA, Orr HA (1989) Patterns of speciation in *Drosophila*. *Evolution*, **43**, 362–381.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, **23**, 3133–3157.

D'Emerico S, Pignone D, Bartolo G *et al.* (2005) Karyomorphology, heterochromatin patterns and evolution in the genus *Ophrys* (Orchidaceae). *Botanical Journal of the Linnean Society*, **148**, 87–99.

Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

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

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Gaskett AC (2011) Orchid pollination by sexual deception: pollinator perspectives. *Biological Reviews*, **86**, 33–75.

Gu W, Wang X, Zhai C, Xie X, Zhou T (2012) Selection on synonymous sites for increased accessibility around miRNA binding sites in plants. *Molecular Biology and Evolution*, **29**, 3037–3044.

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

Johnson SD (2006) Pollinator driven speciation in plants. In: *The Ecology and Evolution of Flowers* (eds Harder LD, Barrett SCH), pp. 295–310. Oxford University Press, Oxford.

Lawrie DS, Messer PW, Hershberg R, Petrov DA (2013) Strong purifying selection at synonymous sites in *D. melanogaster*. *PLoS Genetics*, **9**, e1003527.

Leitch IJ, Kahandawala I, Suda J *et al.* (2009) Genome size diversity in orchids: consequences and evolution. *Annals of Botany*, **104**, 469–481.

Lexer C, Widmer A (2008) The genic view of plant speciation: recent progress and emerging questions. *Philosophical Transactions of the Royal Society B*, **363**, 3023–3036.

Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH (2008) The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B*, **363**, 3009–3021.

Mallet J, Meyer A, Nosil P, Feder JL (2009) Space, sympatry and speciation. *Journal of Evolutionary Biology*, **22**, 2332–2341.

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.

Martin NH, Willis JH (2007) Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric Mimulus species. *Evolution*, **61**, 68–82.

McFarlane HE, Watanabe Y, Yang W *et al.* (2014) Golgi- and trans-Golgi network-mediated vesicle trafficking is required

for wax secretion from epidermal cells. *Plant Physiology*, **164**, 1250–1260.
Nosil P (2012) *Ecological Speciation*. Oxford University Press, New York.
Paulus H (2006) Deceived males – pollination biology of the Mediterranean orchid genus *Ophrys* (Orchidaceae). *Journal Europäischer Orchideen*, **38**, 303–353.
Paulus HF, Gack C (1990) Pollinators as prepollinating isolation factors: evolution and speciation in *Ophrys* (Orchidaceae). *Israel Journal of Botany*, **39**, 43–79.
Peakall R, Schiestl FP (2004) A mark-recapture study of male *Colletes cunicularius* bees: implications for pollination by sexual deception. *Behavioral Ecology and Sociobiology*, **56**, 579–584.
Peakall R, Whitehead MR (2014) Floral odour chemistry defines species boundaries and underpins strong reproductive isolation in sexually deceptive orchids. *Annals of Botany*, **113**, 341–355.
Peakall R, Ebert D, Poldy J *et al.* (2010) Pollinator specificity, floral odour chemistry and the phylogeny of Australian sexually deceptive *Chiloglottis* orchids: implications for pollinator-driven speciation. *New Phytologist*, **188**, 437–450.
Renoult JP, Thomann M, Schaefer HM, Cheptou PO (2013) Selection on quantitative colour variation in *Centaurea cyanus*: the role of the pollinator's visual system. *Journal of Evolutionary Biology*, **26**, 2415–2427.
Rieseberg LH, Willis JH (2007) Plant speciation. *Science*, **317**, 910–914.
Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.
Schiestl FP (2005) On the success of a swindle: pollination by deception in orchids. *Naturwissenschaften*, **92**, 255–264.
Schiestl FP (2012) Animal pollination and speciation in plants: general mechanisms and examples from the orchids. In: *Evolution of Plant–Pollinator Relationships* (ed. Patiny S), pp. 263–278. Cambridge University Press, Cambridge, UK.
Schiestl FP, Schlüter PM (2009) Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology*, **54**, 425–446.
Schiestl FP, Ayasse M, Paulus HF *et al.* (2000) Sex pheromone mimicry in the early spider orchid (*Ophrys sphegodes*): patterns of hydrocarbons as the key mechanism for pollination by sexual deception. *Journal of Comparative Physiology A*, **186**, 567–574.
Schlüter PM, Schiestl FP (2008) Molecular mechanisms of floral mimicry in orchids. *Trends in Plant Science*, **13**, 228–235.
Schlüter PM, Ruas PM, Kohl G *et al.* (2007) Reproductive isolation in the Aegean *Ophrys omegaifera* complex (Orchidaceae). *Plant Systematics and Evolution*, **267**, 105–119.
Schlüter PM, Ruas PM, Kohl G *et al.* (2009) Genetic patterns and pollination in *Ophrys iricolor* and *O. mesaritica* (Orchidaceae): sympatric evolution by pollinator shift. *Botanical Journal of the Linnean Society*, **159**, 583–598.
Schlüter PM, Ruas PM, Kohl G *et al.* (2011a) Evidence for progenitor-derivative speciation in sexually deceptive orchids. *Annals of Botany*, **108**, 895–906.
Schlüter PM, Xu S, Gagliardini V *et al.* (2011b) Stearoyl-acyl carrier protein desaturases are associated with floral isolation in sexually deceptive orchids. *Proceedings of the National Academy of Sciences of the USA*, **108**, 5696–5701.
Scopece G, Musacchio A, Widmer A, Cozzolino S (2007) Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution*, **61**, 2623–2642.
Sedeek KEM, Qi W, Schauer MA *et al.* (2013) Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE*, **8**, e64621.
Smadja CM, Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, **20**, 5123–5140.
Soliva M, Widmer A (2003) Gene flow across species boundaries in sympatric, sexually deceptive *Ophrys* (Orchidaceae) species. *Evolution*, **57**, 2252–2261.
Spaethe J, Moser WH, Paulus HF (2007) Increase of pollinator attraction by means of a visual signal in the sexually deceptive orchid, *Ophrys heldreichii* (Orchidaceae). *Plant Systematics and Evolution*, **264**, 31–40.
Staedler YM, Masson D, Schönenberger J (2013) Plant tissues in 3D *via* X-ray tomography: simple contrasting methods allow high resolution imaging. *PLoS ONE*, **8**, e75295.
Stökl J, Schlüter PM, Stuessy TF *et al.* (2009) Speciation in sexually deceptive orchids: pollinator-driven selection maintains discrete odour phenotypes in hybridizing species. *Biological Journal of the Linnean Society*, **98**, 439–451.
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.
Vereecken NJ (2009) Deceptive behavior in plants: I. Pollination by sexual deception in orchids: a host-parasite perspective. In: *Plant–Environment Interactions* (ed. Baluska F), pp. 203–222. Springer Verlag, Heidelberg.
Vereecken NJ, Schiestl FP (2009) On the roles of colour and scent in a specialized floral mimicry system. *Annals of Botany*, **104**, 1077–1084.
Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.
Whitehead MR, Peakall R (2014) Pollinator specificity drives strong prepollination reproductive isolation in sympatric sexually deceptive orchids. *Evolution*, **68**, 1561–1575.
Widmer A, Lexer C, Cozzolino S (2009) Evolution of reproductive isolation in plants. *Heredity*, **102**, 31–38.
Wolf JBW, Lindell J, Backström N (2010) Speciation genetics: current status and evolving approaches. *Philosophical Transactions of the Royal Society B*, **365**, 1717–1733.
Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.
Wu C-I, Ting C-T (2004) Genes and speciation. *Nature Reviews – Genetics*, **5**, 114–122.
Xu S, Schlüter PM, Scopece G *et al.* (2011) Floral isolation is the main reproductive barrier among closely related sexually deceptive orchids. *Evolution*, **65**, 2606–2620.
Xu S, Schlüter PM, Grossniklaus U, Schiestl FP (2012a) The genetic basis of pollinator adaptation in a sexually deceptive orchid. *PLoS Genetics*, **8**, e1002889.
Xu S, Schlüter PM, Schiestl FP (2012b) Pollinator-driven speciation in sexually deceptive orchids. *International Journal of Ecology*, **2012**, Article ID 285081.
Zhang XS, O'Neill SD (1993) Ovary and gametophyte development are coordinately regulated by auxin and ethylene following pollination. *Plant Cell*, **5**, 403–418.
Zouhar J, Rojo E, Bassham DC (2009) AtVPS45 is a positive regulator of the SYP41/SYP61/VTI12 SNARE complex involved in trafficking of vacuolar cargo. *Plant Physiology*, **149**, 1668–1678.

S.C., F.P.S., P.M.S. designed research; K.E.M.S., G.S., Y.M.S., P.M.S. performed research; Y.M.S., J.S. contributed new techniques; K.E.M.S., G.S., Y.M.S., J.S., S.C., P.M.S. analysed data; K.E.M.S., P.M.S. wrote the manuscript. Genomic data: K.E.M.S., S.C., P.M.S.; reproductive barriers: K.E.M.S., G.S., S.C., P.M.S.; 3D morphology: Y.M.S., J.S.; and colour/odour data: K.E.M.S., F.P.S., P.M.S.

Data accessibility

Illumina sequences: NCBI sequence read archive (SRA), Study: SRP045216 (BioProject PRJNA257331), Experiment: SRX670659–SRX670786, Run: SRR1537566–SRR1537693, Sample: SRS673287, SRS673989, SRS675258–SRS675340, SRS675342–SRS675383.

μCT data: http://phaidra.univie.ac.at/o:363510.

Data/input files: DRYAD, doi:10.5061/dryad.05db4.

Software/code: BIOP/BIOP4R: https://sourceforge.net/projects/biop/; FAMD/GBSPRO: http://www.famd.me.uk/famd.html; modified FDIST2: https://sourceforge.net/projects/fdist3/.

Supporting information

Additional supporting information may be found in the online version of this article.

File S1 MPEG format movie of μCT 3D scanned *O. exaltata* subsp. *archipelagi* flower (sample 491D) (movie file).

File S2 MPEG format movie of μCT 3D scanned *O. garganica* flower (sample 495F) (movie file).

File S3 MPEG format movie of μCT 3D scanned *O. incubacea* flower (sample 498B) (movie file).

File S4 MPEG format movie of μCT 3D scanned *O. sphegodes* flower (sample 484) (movie file).

Appendix S1 FASTA file of tag sequences containing the best outlier SNPs in the global BAYESCAN F_{ST} outlier analysis.

Appendix S2 FASTA file of tag sequences containing the best outlier SNPs from pairwise FDIST2 F_{ST} outlier analyses, replicated in three populations.

Appendix S3 Methods , Figures (S1–S14) and Tables (S1–S8).

Fig. S1 Photographs of sets of randomly sampled flowers of the study species, illustrating floral variation.

Fig. S2 Map of the Gargano region of southern Italy, showing the location of populations used for DNA sample collection (sample locality), and sites at which plot experiments were conducted (experimental sites).

Fig. S3 Pollination success per species during plot experiments 2011 and 2012 at the two experimental sites.

Fig. S4 Potential postmating barriers: (A) ploidy level, (B) fruit set and (C) seed viability in crosses.

Fig. S5 Measurement of different flower parts from photographs.

Fig. S6 μCT-based 3D reconstructions of *Ophrys* flower surfaces, showing average and extreme flowers for each species (see text).

Fig. S7 Morphological landmarks used for analysis of flower shape and shape deformation along the first two principal component axes.

Fig. S8 Side views of labella and zoom-in on details of trichomes based upon high-resolution μCT scans for each species.

Fig. S9 RGB (red, green, blue) colour measurements from standardized, colour-adjusted flower photographs.

Fig. S10 Flower reflectance spectra measurements and mapping into honeybee colourspace, for Inc and Sph.

Fig. S11 Analysis of speculum shape and complexity for the four species.

Fig. S12 Differences in floral odour for the study species, grouped by chemical class or property, as relative proportions of floral odour.

Fig. S13 Results of genetic and population genomic analyses.

Fig. S14 Comparison of the distributions of (A) F_{ST} and (B) LD in pairwise F_{ST} outliers (FDIST2; replicated in at least two populations) as a function of relatedness (as in Fig. S13A).

Table S1 This table lists sample sizes per species as used for different experiments and analyses.

Table S2 Samples and scanning conditions used in the μCT study.

Table S3 Oligonucleotides used for GBS library construction.

Table S4 Summary of GBS experimental setup and overview statistics of sequencing data generated.

Table S5 Flowering time data, showing the number of open flowers per species on a given date.

Table S6 Pairwise flowering time overlap, calculated as in Table S3 of Lowry *et al.* (2008), and as the area under the normalized, extrapolated phenology curve.

Table S7 Summary of pollen tracking experiments to measure floral isolation.

Table S8 Overview of data from crossing experiments.