Pollinator-relevant floral traits underlie bidirectional hybridisation in the orchid genus *Gymnadenia*

Mikhaela Neequaye[1], Roman T. Kellenberger[2,3], Rebecca Collier[1,4], Pirita Paajanen[1], Rea Antoniou-Kourounioti[5], Katherine E. Wenzell[6,7,8], Lionel Hill[1], Philipp M. Schlüter[3,9], Kelsey J.R.P. Byers[1,3]*

* to whom correspondence should be addressed: Kelsey.Byers@jic.ac.uk

1 John Innes Centre, Norwich, UK

2 Crop Genome Dynamics Group, Agroscope Changins, 1260, Nyon, Switzerland

3 Department of Systematic and Evolutionary Botany, University of Zürich, Zürich, Switzerland

4 Harper Adams University, Shropshire, UK

5 Plant Science Group, School of Molecular Biosciences, University of Glasgow, Glasgow G12 8QQ, UK

6 Department of Natural Resource Management, South Dakota State University, Brookings, SD 57007, USA,

7 Botany Department, California Academy of Sciences, San Francisco, CA 94118, USA

8 Department of Entomology, University of Maryland, College Park, MD 20742, USA

9 University of Hohenheim, Stuttgart, Germany

Abstract

Speciation through the emergence of reproductive isolation barriers is a fundamental concept of evolutionary biology. Here we use two closely related but phenotypically distinct *Gymnadenia* orchid species and their proposed hybrids to study the barriers associated with this phenomenon. We analysed the potential pre- and post-zygotic barriers to hybridisation and conducted the first comprehensive analysis of pollinator-influenced phenotypic traits in both parents and their hybrids. We also uncover the previously uncharacterised, but much hypothesised, genetic origins of the hybrids and use both existing literature and novel field observations to predict how the hybrids may have emerged in their respective populations. We demonstrate that all hybrids are F1s, despite their capacity to produce viable seed. We find the hybrids to be intermediate in most phenotypic traits except for their anthocyanin profile, which corresponds to one of the parent species. We use known anthocyanin biosynthesis genes in this system to study the origin of the hybrids' distinct pigment phenotype and model how the resulting floral colours may be perceived by different pollinator guilds. Interestingly, we find significant variation in the directionality of hybrid formation, in which the hybrid populations of two widely separated sample sites have inverse patterns of maternal parent species. This study progresses our understanding of the evolution and establishment of distinct pollinator-mediated floral phenotypes in groups of closely related species. Moreover, this study provides an insight into the influence of spatially distinct pollinator interactions in populations of species increasingly at risk as global climates continue to change and influence ecological processes.

Introduction

For many flowering plants, animal-mediated pollination acts to prevent the formation of hybrids between distinct species, providing a vital source of reproductive isolation (Ollerton et al., 2011; Ramsey et al., 2003). Pollinator isolation is a form of early acting pre-mating pre-zygotic reproductive isolation, therefore studying pollination helps us understand the origin of species (Coyne & Orr, 2004). Pollinator isolation is often achieved through the establishment of arrays of pollinator guild specific floral traits. For closely related species, specialisation of floral traits to distinct pollinator guilds can provide an essential means for pre-zygotic isolation (Bradshaw & Schemske, 2003; Jones, 1978; Katzer et al., 2019; Wessinger, 2021). Shape, size, texture, patterning, colour and scent all influence pollinator interactions, including pollinator attraction (Frachon et al., 2021; Ollerton et al., 2009; Wei et al., 2021) and mechanical fit between pollinator and flower (Romero-Bravo & Castellanos, 2024; Wilson, 1995).

Whilst extensive work has characterised the influence of floral traits involved in pollinator attraction, such as display dimensions and pigments, it is important to note that pre-zygotic barriers such as species-specific placement of pollen on the pollinators body can also form a barrier to hybridisation (Armbruster et al., 2014; Fenster et al., 2004; Grant, 1994). In orchids this may be achieved through the process of resupination, whereby the positioning of the generally most elaborate tepal - the labellum - is twisted during development to form a landing platform for pollinators (Cardoso et al., 2024; Yam et al., 2009). Consistency in floral traits can lead to flower constancy, the process in which pollinators will repeatedly visit the same flower type or single species in a pollination bout, even in the presence of other flower types (Waser, 1986; Yourstone et al., 2023). The ultimate evolutionary outcome of establishing these dynamics is pollinator-mediated assortative mating, which is a key pre-zygotic breeding barrier in flowering plants (Christie et al., 2022; Waser, 1993). However, these pre-zygotic barriers are not always upheld, , which can lead to "leaky" reproductive isolation and hybridization between distinct species (Jersáková et al., 2010).

Members of the genus *Gymnadenia*, including *Gymnadenia rhellicani, Gymnadenia odoratissima* and *Gymnadenia conopsea,* are fragrant Alpine orchids, found in European meadows and grasslands from Scandinavia (Olofsson, 2021) to the Alps and Mediterranean region (Claessens & Kleynen, 2011; Kellenberger et al., 2019). These orchids can be observed flowering from mid-May until mid-August (Delforge, 1994, 2006; Gross & Schiestl, 2015; Hackney, 2007) with flowering times becoming increasingly variable due to abiotic factors like temperature fluctuations, as a result of rapidly changing climate(Kolanowska et al., 2021). The *Gymnadenia* orchids of the European Alps provide a valuable model in which to study speciation events. They are singular in their floral morphology, closely related and produce hybrids. *G. conopsea* and *G. rhellicani* hybridise to form the purported nothospecies *G.* x *suaveolens* in zones of sympatry, e.g. Puflatsch, Seiser Alm, Südtirol, Italy (2125m) and Chandolin, Kt. Wallis, Switzerland (2225m). However, hybrids are rare (Hedrén et al., 2018) and have yet to have their parental origins confirmed genetically. Hybrid scarcity is suggested to be a result of the distinctive morphological traits and therefore pollination systems of the closely related *Gymnadenia* species (Pridgeon et al., 1997; Vöth, 2000). Hybrids therefore form a "natural laboratory" in which to understand both parental species' reproductive isolation - or lack thereof - and the biochemical and genetic mechanisms underlying traits and their integration.

Gymnadenia hybrids and their proposed progenitors possess distinctive morphological traits associated with insect-mediated pollination, most significantly in their floral displays. The relationship between morphology and phenology has been explored in one of the proposed parent species, *G. conopsea* (Chapurlat et al., 2015; Chapurlat et al., 2020) in addition to the influence pollinators have on these morphological characteristics (Sletvold & Ågren, 2010). Whilst there is some overlap in observed pollinator families, such as different families of Lepidoptera, effective pollinator composition differs between the proposed parent orchids (Claessens & Kleynen, 2011; Delforge, 1994, 2006; Hackney, 2007; Sun et al., 2014; Sun et al., 2015; Vöth, 2000). This is proposed to be due in part to their variation in spur length, which coincides with the difference in average proboscis length of their respective effective pollinators (Schiestl & Schlüter, 2009; Sletvold & Ågren, 2010). Variation in nectar spur length is compounded with longer spurs conferring higher nectar volumes (Brzosko & Mirski, 2021). Both *G. rhellicani* and *G. conopsea* are nectariferous perennials which rely on pollinators of various nectar requirements, for example larger Lepidopterans such as the hawkmoths, that have been observed pollinating *G. conopsea* (Vöth, 2000). It is possible this variation in nectar production, alongside spur length, contribute to effective pollinator separation (Jia & Huang, 2022). In the field, the hybrids are immediately distinguishable by their moderately dense, intense magenta inflorescences, whilst the proposed progenitor parent *G. conopsea* has more open lilac-coloured inflorescences, which are in stark contrast with that of the densely clustered dark red inflorescences of *G. rhellicani.* These floral colours vary intra-specifically, with *G. conopsea* displaying a range from light pink to deeper purple petals (Gross & Schiestl, 2015) and the red intensity of *G. rhellicani* varying from wild-type black-red to almost white in some populations (Kellenberger et al., 2019). The combined metabolic and morphological differences, including resupination, in floral phenotypes between the proposed parent species and their implications for pollinator differentiation are hypothesised as a source of hybrid scarcity in these populations.

Genetic regulation of traits such as pigment production are already known to be influenced by pollinator interactions in the proposed parent *G. rhellicani*, where pollinator-mediated overdominance provides the heterozygous and subsequently 'red' pigmented floral morphs with a greater fitness over the homozygous 'black' (dark red) and 'white' morphs in a polymorphic population in Northern Italy (Kellenberger et al., 2019). Earlier work on *G. conopsea* and the closely related species *G. odoratissima* has also looked at the relationship

between floral scent with regards to physiological responses in pollinators (Huber et al., 2005), however it is worth noting that this study was performed before *G. conopsea* and *G. densiflora* were considered separate species. Detailed characterisation of floral scent has not been performed for the *G.* x *suaveolens* hybrids or in *G. rhellicani*. Moreover, *G. conopsea* has become an important example of how pollinator interactions influence floral scent composition (Chapurlat et al., 2019; Chapurlat et al., 2018; D'Auria et al., 2020). Whilst attempts have been made to deconstruct the phylogenetic relationships between orchids of this group (Bateman, 2021; Brandrud et al., 2019; Piñeiro Fernández et al., 2019), characterisation of a full suite of phenotypic traits alongside whole genome sequencing has yet to be performed for these species and their proposed hybrids.

This study had three key aims; (1) Identify the potential barriers to the formation of the *G.* x *suaveolens* hybrids. (2) Characterise a suite of pollinator-relevant phenotypes in the proposed parents and hybrids. (3) Uncover the genetic origins and long-speculated parentage of known *G.* x *suaveolens* hybrids. Here we provide the first characterisation of the various floral traits associated with pollinator-mediated isolation in these *Gymnadenia* orchids, including floral morphology, scent and pigment. We integrate floral phenotypic analyses with genetics and fitness data, to uncover the genetic origins of these hybrids and their emergence in these natural laboratories.

Materials & Methods

Study species & Field Sites

The study species are *Gymnadenia conopsea* (L.) R.Br., *G. rhellicani* (Teppner & E. Klein) Teppner & E. Klein and their proposed hybrid, the nothospecies *G.* x *suaveolens* (Vill.) Rchb. f. Both *G. conopsea* and *G. rhellicani* are generally diploids with a chromosome number of 2n=40 (Rice et al., 2015). These fragrant Alpine orchids are perennials that do not reflower every year and produce a single inflorescence per plant per flowering year. These inflorescences are cylindrical in *G. conopsea* and conical to spherical in *G. rhellicani*, with variability observed in *G.* x *suaveolens*. All species are found in Alpine meadows and grasslands. *G. conopsea* can be found across Eurasia, whereas *G. rhellicani* is restricted to European mountain ranges including the Pyrenees and the Alps (Claessens & Kleynen, 2016; Delforge, 1994, 2006; Moen & Øien, 2002; Olofsson, 2021). The two species are only sympatric in the European Alps. For this reason we chose the following two field sites where both species grow in sympatry and *G.* x *suaveolens* has been reliably observed and documented (Claessens & Kleynen, 2016; Delforge, 1994, 2006). The first field site - denoted as (PUF) - was Puflatsch, Seiser Alm, Südtirol , Italy (2125m). The second field site - denoted as (CHA) - was Chandolin, Kt. Wallis, Switzerland (2225m). Field sites and their coordinates are shown in Figure S1.

Transect

A walked transect was performed in 2015 at PUF (RTK3 site) to quantify the number of individuals of each *Gymnadenia* species within the transect.

Phenology Quadrat

As a metric of phenology, we conducted four quadrats at PUF in 2021 towards the end of the flowering season, due to fluctuations in weather that year. We photographed every *Gymnadenia* individual in each quadrat and, using GIMP (GNU Image Manipulation Program) software (Team, 2019), and manually calculated the ratio of open flowers and buds to that of senesced flowers within an inflorescence. This provided a snapshot of inflorescence senescence at that point in time, for comparison between species.

Pollinator ecology

Floral visitors

In 2015, floral visitors at the Puflatsch site were haphazardly collected across a period of 2 days including morning, afternoon, and evening (dusk and night) collections. Floral visitors were noted as to type and time of day (diurnal or nocturnal) when observed landing and interacting with an orchid with open flowers. Where possible, those observed were captured using a net and moved into 50 ml falcon tubes with ethyl acetate killing fluid for later identification. A separate literature search was also undertaken to determine identities of historically recorded visitors to the two putative parent species. Data on pollinator and visitor proboscis lengths for both 2015 collections and the historical visitation data was taken from the literature for captured species where available (see Table S1). When species-level data were unavailable, an average value for the same genus or family was used. Proboscis lengths of Coleoptera and non-Empididae Diptera was inferred to be 0.01mm as the proboscis of these insects is too large to fit into the orchid spurs of either species.

Pan traps

To understand the diversity of potential floral visitors at sampling locations, 8 pan traps were placed at sites in Puflatsch and Chandolin in 2022 and this was repeated at Chandolin in 2023. To create the pan traps, 4 blue and 4 yellow bowls were filled ¾ full of tap water with 15 ml of unscented concentrated castile soap (Dr. Bronner's Magic Soaps, Vista, CA, USA). A blue and yellow bowl were placed together in various locations within the sampling sites of the parent species. Blue and yellow were chosen as they are most frequently used for pan trapping experiments of this nature (Saunders & Luck, 2013; Vrdoljak & Samways, 2012). The bowl pairs were placed ca. 10-15 metres from one another. Pan traps were left for 24 hours, so the collection is representative of the local diurnal and nocturnal insect communities but does not distinguish between the two. To remove the captured insects, the contents of the bowls were strained at the site and placed in gallon plastic bags, pooling all the yellow bowl collections and blue bowl collections separately. The contents of each bag were emptied, and each individual insect identified as one of the following groups: Lepidoptera, Hymenoptera, Diptera, and Other (mainly Coleoptera, Orthoptera and various spiders, with few potential pollinators in this category). Individuals were pooled in respective 50ml falcon tubes containing 100% ethanol for preservation.

Floral morphology

Morphological measurements

Plant height and inflorescence height were measured using a Stanley 035455 2m Wooden Folding Rule. Individual floral display height, display width and nectar spur length were measured using 150mm digital callipers (Linear Tools, RS Components).

Number of potential ovaries

The number of potential ovaries is the summation of buds, open flowers, and senesced flowers counted on each plant's single inflorescence. The images analysed were captured over a period of 6 years (2014-2016 and 2021-2023) intermittently, from both sampling locations. They were loaded into the GIMP (GNU Image Manipulation Program) software (Team, 2019), and manually scored (using different coloured markers) for buds, open flowers, and senesced flowers. This process was repeated three times per image to produce an average. The average number of buds, flowers and senesced flowers was then multiplied by a factor of 1.22, to encompass the parts of the inflorescence which were not captured in the photo, so that each number was representative of a whole plant (Kellenberger et al., 2019).

Resupination

The images were captured in the field as level as possible. In GIMP (Team, 2019), a line was drawn down the inflorescence on the image using the measurement tool and the image was adjusted using a rotation tool to an angle of 180°. The angle of resupination was then measured by drawing a line from the labellum to the centre of the flower. This was repeated for three flowers per inflorescence and a mean angle was then calculated.

Senescence

Whole inflorescences were photographed in the field in 2021. The length of each inflorescence was measured in pixels (px) using GIMP (Team, 2019), and then the proportion of senesced flowers was measured. The senesced proportion was subtracted from the total inflorescence height, consequently deducing the height of open flowers. Height measurements of senesced proportions (px) were converted to percentages, to represent the proportion of senescence for each individual inflorescence. Measurements were conducted three times, and an average was used for the subtractions.

Floral scent

Dynamic headspace capture (Raguso & Pellmyr, 1998) was performed on single individuals in the field as follows. Samples collected from 2014-2016 were collected, run, and analysed as in (Kellenberger et al., 2019).

For 2021-2023 samples, inflorescences were kept attached to the plant and enclosed in a PET oven bag (ca. 1L volume; Toppits, Minden, Germany) attached to a battery-driven pump (Spectrex PAS 500, Spectrex Corp., Redwood City, CA, USA) connected to a volatile trap consisting of a trimmed Pasteur pipette loaded with silanized glass wool and 100mg PorapakQ. Air was pulled from the bag through the trap at a flow rate of 100-200 ml min^{-1} for 40 minutes in earlier sampling seasons and increased 2 hours from 2023 for logistical purposes. The traps were kept cool in the field prior to elution from each column using 600 µl of 10% acetone in HPLC-grade hexane and stored in the freezer where possible, given field conditions. Background VOC levels were determined with simultaneous samples from the same time and location in field using empty bags. Samples were concentrated 3-fold from 150 µl to 50µl under ambient room air and 3µl was injected using a Gerstel MPS system into an Agilent 7890B GC with 5977A/B MS GC-MS in splitless mode with the following method. A Zebron ZB-5HT w/GUARDIAN column (Phenomenex, Macclesfield, UK; 35 m, 0.25 mm, 0.1 µm) was used, and helium was used as the carrier gas at a constant flow of 1 cc min−1. The initial oven temperature was 50°C for 4 min, followed by a heating gradient of 5°C min−1 to 230°C, which was then held isothermally for 4 min. The MS was operated in scan mode (50-600m/z). Initial identification and quantification of compounds was performed with Agilent Unknowns software (v10.1, Agilent, Santa Clara, CA, USA) based on a mass spectral library built on calibration curves using three to five different concentrations of authentic reference standards. This was followed by a custom analysis pathway for removal of contaminants identified in the blank columns collected and quantification of the remaining compounds (as described in Wenzell & Neequaye et al., 2025).

Floral pigment

Reflectance spectrophotometry

Samples collected from 2014-2016 were measured and analysed as in (Kellenberger et al., 2019). For 2021-2023, measurements were made on a spectrophotometry set up including an Ocean Insight FLAME-S-UV-VIS-ES Assembly with a Pulsed Xenon Lamp at 220 Hz, 220-750 nm with a Premium 600 um Reflectance Probe using Ocean View spectroscopy software, Ocean

Insight. The labellum of single, fresh flowers per individual inflorescence were used for reflectance spectrophotometry, with three technical replicates per flower. A white standard (Certified Reflectance Standard, Labsphere) was used to calibrate the spectrophotometer measurements prior to each set of measurements.

Anthocyanin analysis

Anthocyanin extraction was modified from (Butelli et al., 2008). Three whole flowers per individual were preserved in absolute methanol prior to extraction to facilitate transfer from the field. Flowers were ground in 80% methanol and 1% hydrochloric acid and shaken overnight at 4°C. Samples were spun down at maximum speed with an aliquot of supernatant taken for UHPLC-MS and spectrophotometric absorption at 525nm. Samples were run on an Agilent 1290 Infinity II UHPLC equipped with PDA and 6546 Q-Tof. Separation was on a 100×2.1mm 2.6µ Kinetex EVO column using the following gradient of acetonitrile versus 1% formic acid in water, run at 600µL.min-1 and 40°C. The diode array detector collected individual channels at 350nm (bw 4nm) and 525nm (bw 50nm), and also full spectra from 220-640nm, at 10 Hz.The Q-ToF collected positive electrospray MS using the JetStream source, together with data-dependent MS2. The MS was calibrated before use, but also had two lock-masses infused during the runs, at 121.05087 and 922.009798; components of Agilent's normal ESI calibration mix). Spray chamber conditions were 325°C gas, 10L.min-1 drying gas, 20psi nebulizer pressure, 400°C sheath gas, 12L.min-1 sheath gas, 4000V Vcap (spray voltage) with a nozzle voltage of 1000V. The fragmentor voltage was 180V.

Genetic analysis

Flow cytometry

Protocol modified from (Doležel et al., 2007; Gross & Schiestl, 2015; Kellenberger et al., 2019) . The internal standard was *Phaseolus coccineus* (Scarlet emperor) leaves, which contain a mix of diploid (2n) and tetraploid (4n) cells while the latter undergo DNA replication and cell division. Young leaves (for the internal standard) were chopped in a petri dish using a razor blade, and saturated in Baryani's solution (Otto I substitute, 0.1M citric acid monohydrate, 0.5% v/v Triton-100). The solution was then extracted, filtered through a 30µm CellTrics filter, and placed on ice. Pollinia samples (which contain a mix of haploid (n) and diploid (2n) cells) were ground in a 1.5ml Eppendorf using a small pestle with 40µl of Baryani's. 5 µl of the filtered internal standard solution were added to each pollinia sample, and the solution then filtered as above. Samples were spun for 5 minutes at 2000 rpm. The resulting pellet was resuspended in 40µl of Baryani's solution and placed on ice for 1-3 hours. Whilst chilling the pellet, 996µl Otto II solution (0.4M $Na_2HPO_4 \bullet 7H_2O$) was combined with 4µl of DAPI in an Eppendorf tube, and the tube was wrapped in aluminium foil to keep the DAPI in the dark. Samples from 2014-2016 were analysed on a Beckman Coulter Cell Lab Quanta flow cytometer (Beckman Coulter, USA) and those from 2021-2023 were analysed using a BD FACSMelody cell sorter (BD Biosciences, Wokingham, UK). The ploidy was then determined by the ratio between the median value for the internal standard (first, i.e. diploid, peak of two internal standard peaks) and pollinia (second, i.e. diploid, peak of two pollinia peaks) sample peaks displayed on each flow cytometric histogram.

Whole Genome Sequencing

DNA extractions were performed on silica gel-dried leaves collected from the field using the Cytiva Nucleon PhytoPure Genomic DNA Extraction Kit Code: RPN8511 (Cardiff, Wales). Samples were sent to an external sequencing provider for 150bp Paired-end Illumina sequencing using a NovaSeq 6000 following library preparation using NEBNext Ultra II FS DNA Library Prep. Parent species were sequenced with 20x coverage and hybrids in 10x. This

includes 12 parent species (5x G. conopsea, 5x G. rhellicani) and 21 hybrids from Chandolin, and 15 parents (5x G. conopsea, 5x G. rhellicani) and 42 hybrids from Puflatsch.

There are no reference genomes to *G. conopsea* or *G. rhellicani*, apart from the short fragments of MYB1 and ANS sequences from (Kellenberger et al., 2019). Thus, all the whole genome level work was done reference free using k-mer methods, using the sequenced fastq-files for each sample.

Counts of shared sequence content between hybrids and parents was performed with an in house pipeline incorporating yak (https://github.com/lh3/yak) software, which already had k-mer counting and parental share counting functions. Firstly the k-mers from raw fastq reads of each potential parent where counted using the 'yak count' function with the following parameters k-mer size 57 and Bloomfilter 37. Then the function 'yak triobin' was used with the k-mer counts of the potential parents and the hybrid as the offspring. Results were printed as percentages supporting each parent with an in-house script.

Orchid plastid genomes are derived from the maternal or "egg" parent (Cafasso et al., 2005; Corriveau & Coleman, 1988). The raw fastq reads from the hybrids were mapped to the plastid, MYB1 and ANS sequences using bwa v0.7.17. (Li & Durbin, 2009). Plastid sequence identity was determined using the number of trnL intronic repeats published in (Hedrén et al., 2018). MYB1 and ANS sequences were derived from (Kellenberger et al., 2019) and visualised in iGV (Robinson et al., 2011), where the identity of the sequences were visually assigned.

Seed volume and viability

Seed counts

Seeds were extracted from dry fruits, ranging in number from 3 to 5 fruits per orchid sample (only a small subset of fruits was collected for each plant to avoid population viability risks). Using paper, the seeds were carefully transferred to marked conical-bottom 2ml freezer vials. The vials were stored in a saturated lithium chloride (LiCl) atmosphere with open caps. The LiCl environment was created using a Tupperware box, filled 1cm deep with lithium chloride and 3cm of sterile water. After a month of drying, the seeds were spun down in a centrifuge at 3000 rpm. The seed vials were organised into groups of five and photographed from the side at bench height. Photographs were analysed using GIMP software (Team, 2019). The volume of seeds in each vial was calculated by measuring the height and width of the seed tube using the GIMP measurement tool (see Figure S2A). Using basic geometry, measurements were converted to cubic millimetres and from there to seed counts using the known seed volumes for *Gymnadenia* sp. (representing *G. conopsea*) and *"Nigritella"* (now *Gymnadenia* subgenus *Nigritella*, representing *G. rhellicani*) from (Arditti & Ghani, 2000). Potential ovary counts (see above) were used to scale up the number of seeds per collected fruit to the number of seeds per inflorescence.

Seed viability

Microscopy with a Zeiss Axio Zoom V16 (Zeiss, Cambourne, UK) was used to determine the fertility of individual plants. After calculating their volume (see above), seeds were examined for fertility. As orchid seeds contain no endosperm (Vinogradova & Andronova, 2002), the seeds are mainly transparent under light microscopy, with a solid embryo showing as an opaque rounded shape in the centre if fertile, and no embryo present if infertile (see Figure S2B). A petri dish was scored with grid lines 1cm apart using a black marker. The microscope was set to bright field, aperture 100%, magnification 10.5x. The Zeiss ZEN 3.1 software was paired to the microscope camera and was used to capture images of the seeds which were poured onto the scored petri dish and scattered evenly. Images were captured of each grid section, until

approximately 100 seeds per sample were captured. Seeds were marked for fertility using a custom R script ('CountSeeds_A_03.02') (Schlüter & Byers, 2015). They were manually marked in the script as fertile or infertile seeds, using left and right mouse clicks respectively with R saving the progress after each 100 seeds marked per sample. The number of fertile seeds per individual was calculated by multiplying the percentage of fertile seeds by the total number of seeds for that individual.

Data and code availability

Sample sizes for all analyses are given in Table S10. The 2014 *G. rhellicani* data from Chandolin and Puflatsch is already published in (Kellenberger et al., 2019). All genome data are available on ENA under accession number PRJEB76404. The R script for seed counting (written by Schlüter & Byers in 2015) is available on GitHub (https://github.com/plantpollinator/CountSeeds) and the shell scripts for sequence analysis are available on GitHub as well (https://github.com/plantpollinator/GymnadeniaGenomics).

Results

Assessment of potential reproductive isolating barriers

Hybrids are rare, even in large parental populations

When found in zones of sympatry, *Gymnadenia conopsea* and *G. rhellicani* can hybridise to form the purported nothospecies *G.* x *suaveolens.* We chose to focus our study on two of these sympatric zones (Figure S1); Puflatsch, Seiser Alm, Südtirol, Italy (2125m) (denoted as PUF in figures and tables) and Chandolin, Kt. Wallis, Switzerland (2225m) (denoted as CHA in figures and tables). We began with a walked transect, counting all *Gymnadenia* in PUF (RTK3 site) in 2015. Within this transect we found 1362 *G. conopsea*, 1062 *G. rhellicani* and 19 *G.* x *suaveolens* hybrids. Thus, the hybrids make up 0.789% of the total population of orchids. Rather than forming a "hybrid swarm" in this zone of sympatry, the hybrids appear to emerge in zones where the parent species are found in large abundance.

Hybrids and parents are diploid based on flow cytometry

To understand why the hybrids appear so infrequently in large parent populations, we first set out to uncover potential barriers to hybridisation. Variation in ploidy often functions as a barrier to hybrid formation (Ramsey & Schemske, 1998). Both *G. conopsea* and *G. rhellicani* are diploid with 2n=40 (Rice et al., 2015) . To confirm this we performed pollinia-based flow cytometry, which showed equivalent internal standard:pollinia nucleus size ratios between parents and hybrids. We confirmed a lack of interploidy between the proposed parent species with hybrids showing some variability (Figure 1B). We therefore conclude that gross-scale ploidy and chromosome count incompatibilities are likely not a reproductive isolating barrier in this system, although smaller incompatibilities (e.g. chromosomal inversions) may still play a role.

G. conopsea, G. rhellicani and hybrids all make fertile seeds

We asked whether the hybrids experience a penalty in fitness and whether this might provide a reproductive isolating barrier, leading to hybrid rarity. When considering total volume of seeds, *G. conopsea* was found to produce the greatest number of seeds (average 35,245 ± 33,840 seeds per plant), with *G. rhellicani* (average 3,458 ± 6,689 seeds per plant) and hybrids (average 9,359 ± 15,076 seeds per plant) producing similar amounts (Figure S2C). This correlates with *G. conopsea* producing more flowers per inflorescence and therefore more seeds as a result.

Whilst *G. conopsea* has a higher abundance of total seeds produced, it has as a similar proportion of seed fertility (57.08 ± 29.99% fertile seeds) when compared to that of *G. rhellicani* (71.38 ± 22.41% fertile seeds) (Figure 1C). Both parent species were found to have greater seed fertility than that of the hybrids (39.61 ± 20.35% fertile seeds) (Figure 1C). This level of seed fertility, alongside having the highest number of seeds, leads to *G. conopsea* having the greatest number of fertile seeds overall (Figure 1D). Initial hypotheses suggested that the hybrid seeds would be significantly less fertile as they amount to less than 1% of their respective populations. The hybrid samples show that many have a substantial proportion of fertile seeds, while some produce few if any fertile seeds. However, when compared against the abundance of seeds, the fertile hybrid seed set is relatively small (average 3,479 ± 5,725 fertile seeds) in comparison to *G. conopsea* (average 18,216 ± 14,237 fertile seeds) but more similar to that of *G. rhellicani* (average 1,944 ± 3,276 fertile seeds) (Figure 1D). See Table S5 for all statistical details of fitness and fertility.

Parent species overlap in flowering time

An additional potential barrier to hybridisation is discordance in flowering times (Levin, 1978; Michalski & Durka, 2015). Anecdotally, based on personal observations at PUF and CHA, parents strongly overlap in flowering time and experience peak flowering simultaneously. However, fluctuations in climate in recent years have contributed to disruption in historical flowering times of orchids of this group (Kolanowska et al., 2021). As a metric of phenology, we conducted quadrats at PUF in 2021 towards the end of in the flowering season, photographing every *Gymnadenia* spp. in each quadrat (in total across all four quadrats: 59 *G. conopsea* and 35 *G. rhellicani*) and calculating the ratio of open flowers and buds to that of senesced flowers within an inflorescence. We found *G. conopsea* inflorescences averaged 14.28% open flowers and buds with 85.72% of the inflorescence senesced. In *G. rhellicani* this was 37.45% open flowers and buds and 62.55% of the inflorescence had senesced (Figure 1E). These data thus demonstrate significant overlap in flowering time but also suggest that *G. conopsea* likely peaks earlier in the season than *G. rhellicani*, which may provide a small contribution to reproductive isolation.

G. conopsea and G. rhellicani have distinct pollinator guilds with some overlap in Lepidoptera

To assess whether pollinators might play a role in reproductive isolation between the parent species and thus hybrid rarity, we wanted to further interrogate the pollinator-mediated mechanisms for hybrid formation at the two sites. To do this we combined historical observation data from the literature with our own field observations (Figure 2, Table S1, Table S2). The historical literature suggests Lepidopteran species as focal pollinators for both parent species, especially *G. conopsea*, where Hymenoptera and Diptera are still frequently observed visitors (Claessens & Kleynen, 2011; Faegri & Van der Pijl, 1979; Proctor et al., 1996; Vöth, 1999; Vöth, 2000) (Figure 2B). Similarly, *G. rhellicani* also has Hymenopteran and Dipteran pollinators in addition to Lepidopterans (Claessens & Kleynen, 2011; Faegri & Van der Pijl, 1979; Vöth, 2000) (Figure 2C). *G. conopsea* is mostly pollinated by Lepidoptera with long probosces and *G. rhellicani* by Lepidoptera with medium-sized to short probosces (Schiestl & Schlüter, 2009) (Figure 2D), and this difference is significant in the historical literature data ($t = 6.60$, $df = 159.31$, $p = 5.7E-10$). Members of Lycaenidae, Hesperidae (day-active) and Sphingidae (day- and night-active) have been observed carrying pollinia from *G. conopsea* whereas Nymphalidae (day-active) and Pyralidae (night-active) have been observed carrying pollinia from *G. rhellicani* (Claessens & Kleynen, 2011, 2016; Faegri & Van der Pijl, 1979; Proctor et al., 1996; Vöth, 1999; Vöth, 2000). Both *G. conopsea* and *G. rhellicani* share five families of Lepidopterans known to carry pollinia, three day-active (Pieridae, Nymphalidae and

Zygaenidae) and two night-active (Noctuidae and Geometridae) (Claessens & Kleynen, 2011, 2016; Faegri & Van der Pijl, 1979; Proctor et al., 1996; Vöth, 1999; Vöth, 2000).

To get a wider ecological perspective of the diversity and abundance of potential pollinators, we set up pan traps over a 24-hour period in the location of plants during flowering in 2022 (Puflatsch and Chandolin) and 2023 (Chandolin only) (Figure S3). Pan trap data when controlling for bowl colour shows higher value for flies in Puflatsch and lower in Chandolin, for both blue and yellow pan traps. Overall, flies were the most frequently caught in the pan traps, and Lepidoptera were the least frequently caught. 'Other' organisms (mostly consisting of species not expected to pollinate, e.g. Orthoptera and spiders) were in the highest abundance in Puflatsch, and Hymenoptera were found more often in the blue bowls than the yellow at both sites. Overall, the pan trap data confirms that the same groups of potential floral visitors are present in the environment of both progenitors. However, the data also shows that there is variability in the abundance of certain pollinator groups, with significant variation by sample year at Chandolin, reflecting the substantial differences in climactic conditions and sampling time between sampling years (very dry and mostly senesced meadows in 2022, fresher meadows and abundant flowers in 2023). It is also possible that the pan traps missed some floral visitors which may not have been captured and accounted for, such as most butterflies and moths, including hawkmoths.

Pollinator-relevant floral phenotypes in parents and hybrids

Hybrid morphology is intermediate with a higher resemblance to the G. rhellicani parent.

First looking at gross morphological variation between the proposed parent species, *Gymnadenia conopsea* and *Gymnadenia rhellicani* and their hybrids, we found that *G. conopsea* and *G. rhellicani* differ in total plant height, with hybrids being intermediate (Figure S4A). This is also found in both absolute inflorescence height and the proportion of inflorescence height relative to total plant height (Figure S4B, Figure 3A). Looking closer at inflorescences, flower number (calculated as the total number of potential ovaries, i.e. buds plus open flowers plus senesced flowers) is found to be the same between *G. conopsea* and *G. rhellicani*, but slightly lower in hybrids (Figure 3B). Height and width of individual flower display is intermediate in hybrids, with *G. conopsea* producing larger flowers on both axes than those of *G. rhellicani* (Figure S4C, D). In addition to the relative sizes and arrangement of floral display, length of nectar spur can often be used as an indicator of likely pollinators, as *G. conopsea* is pollinated predominantly by Lepidoptera with longer probosces, whilst *G. rhellicani* favours Lepidopterans with shorter probosces, amongst other pollinator guilds (Figure 2D). In the case of these orchids, nectar spur length was also found to be intermediate in the hybrids, however closer in length to that of the *G. rhellicani* parent than that of *G. conopsea* (11mm in *G. conopsea*, 4.5mm in hybrids, 1.5mm in *G. rhellicani*) (Figure 3C). Resupination is an additional morphological marker often for specialised pollination syndromes and occurs in many orchids. It is a gravitropic mechanism regulated by auxin (Yam et al., 2009), where buds develop and twist, so the labellum is positioned at the bottom of the flower, typically around 180° to form a platform for floral visitors (Arditti, 2003; Zhang et al., 2018). The largest angle of rotation was found in *G. conopsea* (average 152.91° ± 17.95°) when compared to *G. rhellicani* (average 8.23° ± 4.37°) which effectively does not resupinate. In accordance with nectar spur length, floral resupination is intermediate but also highly variable in hybrids (average 44.36° ± 20.40°), with greater similarity to that of the *G. rhellicani* parent than that of *G. conopsea* but a similar variability to *G. conopsea* (Figure 3D). Both nectar spur length and floral resupination can affect placement of pollinia on pollinators therefore providing a potential mechanism for mechanical isolation between the two parent species, while the intermediate phenotype of the hybrid in both traits

483 may inhibit correct placement of pollinia and pollinator identity and thus effective pollination.
484 See Tables S3 and S4 for all statistical details of morphology.

485 *Hybrid scent is intermediate, but compound dependent.*

486 From studies in *Gymnadenia conopsea*, we know that floral scent is known to play a key role in
487 the attraction of pollinators of *Gymnadenia* orchids, with specific volatiles demonstrating both
488 physiological and behavioural activity (Chapurlat et al., 2019; Huber et al., 2005). The
489 *Gymnadenia* orchids in this study were found to produce up to 47 different volatile organic
490 compounds (VOCs), the majority of which being phenylalanine-derived aromatics (26) (Figure
491 4A). Hybrids do not display global up- or down-regulation of volatile pathways relative to the
492 parent species, but rather relative abundance volatile patterns in hybrids differ depending on
493 the specific volatile (Figure 4B). For example, whilst absolute phenylethylacetate production
494 (data from 2014-2016 only) is intermediate in hybrids, phenylethylalcohol production is more
495 similar to the *G. rhellicani* parent whereas phenylethylacetaldehyde production favours the *G.*
496 *conopsea* parent (Figure 4B). These volatiles are biochemically all within two reactions of each
497 other (Figure 4C) (Dudareva et al., 2013). Interestingly, many of these phenylalanine-derived
498 aromatics, including phenylacetaldehyde and phenylethylacetate, are known to elicit an
499 electrophysiological response in pollinators of orchids of this group (Huber et al., 2005). Whilst
500 the characterisation by Huber, Kaiser et al., 2005 was initially performed on closely related
501 *Gymnadenia odoratissima* and the previously synonymized *G. densiflora* and *G. conopsea*
502 combined (Huber, pers. comm.), it is important to note that both *G. rhellicani* and *G. conopsea*
503 share major pollinators. When taken altogether, VOC production is found to be intermediate in
504 hybrids, with significant overlap in patterns of production between species as well as between
505 species of the same sampling site (Figure 4D). See Tables S7, S8, and S9 for all statistical details
506 of floral volatile production analysis.

507 *Hybrid colour has intermediate perception by pollinators*

508 Perception of *Gymnadenia* spp. by pollinators is influenced by colour, in addition to scent
509 (Kellenberger et al., 2019). Reflectance spectrophotometry of live floral tissue reveals that
510 hybrids are intermediate in floral reflectance between the two proposed parent species, with
511 greater similarity to *G. conopsea* overall (Figure 5A, B). This is consistent in both the blue and
512 red wavelength peaks, which may be able to distinguish the orchids from the pollinators'
513 perception (Figure 5C, D). This trend is also found when modelling for pollinator perception
514 with greater overlap between hybrids and *G. conopsea* when considered through the visual
515 perception systems of both bees (using the honeybee *Apis mellifera* as a model) and flies (using
516 the fly *Musca domestica* as a model). When modelling this with perception by Hymenoptera,
517 such as bees, plants are perceived as "Bee UV-blue" to "Bee blue-green", with hybrids closer to
518 pure "Bee blue" than *G. rhellicani*, and hybrids favouring the *G. conopsea* parent overall (Figure
519 5E). For flies, which are generally attracted by opposing colours (Kellenberger et al., 2019), all *G.*
520 *conopsea* sampled are perceived the same (fall into the same quadrant); whereas *G. rhellicani*
521 sorted by colour due their higher variation in pigment in the different colour morphs, and thus
522 would be perceived differently from one another. In the fly perception model hybrids fall into
523 two quadrants (Figure 5F), suggesting they would be sorted into two different visual categories
524 by flies. When considering hawkmoths, a Lepidopteran species, none of the *Gymnadenia* species
525 produced peaks in the UV, which would act to repel hawkmoths (Figure 5A) (White et al., 1994).
526 Both hybrids and *G. conopsea* produce a high blue peak that is attractive to hawkmoths such as
527 *Macroglossum stellatarum* (hummingbird hawkmoth) which responds very strongly
528 behaviourally to blue wavelengths (Figure 5C) (Telles et al., 2014). We modelled general
529 Lepdiopteran perception as a generic tetrachromatic system (after León-Osper & Narbona
530 2022), demonstrating that both parents and hybrids fall primarily within the sensitivities of the

531 short- and long-wavelength receptors (Figure 5G), although we note that the long-wavelength
532 sensitive receptor is absent in *Macroglossum* (Telles et al., 2014); the short receptor of
533 *Macroglossum* peaks at 440nm, approximately the wavelength of the "blue peak" (Figure 5C).
534 Unfortunately, there are no data on the visual system of *Zygaena* spp., the major shared
535 pollinator of the two parent species. See Table S6 for all statistical details of reflectance at two
536 key wavelengths.

537 *Hybrid pigments reflect the G. conopsea parent.*

538 Having noted significant visual colour and reflectance differences between the parent species
539 and hybrids, we sought to characterise the anthocyanin production and profiles of these
540 orchids. The prospective parent species, *G. conopsea* and *G. rhellicani*, contain similar levels of
541 total anthocyanins per mg of floral tissue (absorbance at 525nm) to each other, with no
542 significant difference when compared with the hybrids (Figure 6A), despite their visually
543 different colours. This is consistent at different sampling locations, with individuals possessing a
544 range of pigment intensities when sampling. It is important to note that all three known
545 pigment types of *G. rhellicani* were sampled (Figure 2A) (Kellenberger et al., 2019). When
546 looking specifically at all untargeted anthocyanins and anthocyanin-intermediate metabolites,
547 hybrids produce an intermediate metabolite profile when compared to *G. conopsea* and *G.*
548 *rhellicani*, with slightly greater overlap with *G. conopsea* at the PUF field site compared to CHA
549 (Figure 6B, C). In contrast to this intermediate total metabolite phenotype, hybrids make a
550 similar final anthocyanin profile as *G. conopsea* as shown in their chromatograms at wavelength
551 525nm (Figure 6D). This is consistent regardless of sampling year or location. Preliminary
552 individual anthocyanin identification and example raw chromatograms of *G. rhellicani*, *G.*
553 *conopsea* and 3 hybrids of the same sampling year and location can be found in Figure S5.
554 Hybrid anthocyanin profiles at 525nm are not intermediate between *G. rhellicani* and *G.*
555 *conopsea* either globally or at the level of individual anthocyanins, which is inconsistent with the
556 other phenotypes measured.

557 *Volatile and UHPLC-MS modularity differs between parents and hybrids, as*
558 *well as by population*

559 Floral pigments are often coupled with floral scents, as similar precursor compounds and
560 biosynthesis pathways are involved in their production. For example, anthocyanin pigments are
561 related biosynthetically to benzenoid and phenylpropanoid volatiles (Li et al., 2024).
562 Interestingly in this system all species studied produce a high proportion of phenylpropanoid
563 and benzenoid volatiles (Figure 5A). These are derived from phenylalanine and are a result of
564 the shikimate pathway, alongside the anthocyanins discussed above, suggesting a potential
565 global upregulation of the shikimate pathway in both parental species and hybrids leading to
566 high levels of aromatic volatiles and anthocyanins.

567 Considering the relationship between anthocyanins and scent, we sought to uncover whether
568 there are any inter-trait correlations between these two categories of floral phenotypic traits.
569 Interestingly in VOCs, biochemical pathway distance in the aromatics (2014-2016 data only) is
570 correlated with inter-volatile correlation in *G. conopsea* ($r = -0.512$, $p = 10^{-10}$) and hybrids ($r = -$
571 0.581, $p = 10^{-13}$), but not in *G. rhellicani* ($r = 0.108$, $p = 0.64$). This supports the finding that
572 anthocyanin profiles of *G. conopsea* and hybrids being shared, suggesting an inherited
573 regulation of the Shikimate pathway in hybrids from the *G. conopsea* parent. However, this could
574 be an effect in *G. rhellicani* of smaller number of testable volatile comparisons. When looking at
575 overall absolute Pearson Correlation Coefficient (PCC) values across the parents and hybrids,
576 we found that these values differed between all three (Figure S6A), as did the proportion of
577 significant PCCs ($p < 0.05$: 9.14% for *G. conopsea*, 10.06% for *G. rhellicani*, and 16.87% for the

578 hybrids, Figure S6B, data points below the red line). The hybrids thus seemed to have slightly
579 stronger trait integration than either of the parent species within the volatile and anthocyanin
580 datasets. We did also find strong site-specific effects, with absolute PCCs being significantly
581 higher at PUF (mean absolute PCC 0.401) than at CHA (mean absolute PCC 0.311, t = -2.992, df =
582 1056, p = 0.0028). This suggests that plants at the PUF site overall have stronger integration
583 between volatiles and anthocyanins than plants at the CHA site, perhaps due to differences in
584 local pollinator communities.

585 *Hybrid identity and a demonstration case with a pollinator-relevant*
586 *floral phenotype*

587 *Hybrids are all F1s and show site-specific bidirectionality in parentage.*

588 Orchids plastid genomes are derived from the maternal or "egg" parent (Cafasso et al., 2005;
589 Corriveau & Coleman, 1988). Continuing from the work of (Hedrén et al., 2018) in which plastid
590 genome markers were used to determine maternal parentage, we used whole genome
591 sequencing to determine hybrid parentage using both genome balance proportions (55%
592 genome is maternal parent due to the additional plastid contribution, while the other 45%
593 comes from the pollen parent) and copy number variation of protoplastic SNPs in the same *trnL*
594 intron. Parent species analysed were site-specific, in that hybrid *trnL* intronic sequences were
595 compared to parents of that corresponding site. We were able to determine that all hybrids
596 sampled were F1s, resulting from a primary cross between the two parent species, rather than
597 advanced generation hybrids or backcrosses to either parent species, because there genome
598 balance was following 55/45% split between the parental species. We had earlier simulated
599 various advanced generation hybrids and backcrosses, but none of the hybrids showed genome
600 balances that would have been indicative of these scenarios. Our work shows bidirectionality in
601 F1 hybrid formation, in which both *G. rhellicani* and *G. conopsea* act as pollen-donor and egg-
602 donor. Interestingly, this bidirectionality is site-specific with majority of hybrids at Chandolin
603 being *G. conopsea* maternal and majority of hybrids at Puflatsch being *G. rhellicani* maternal
604 (Fishers Exact Test p=0.0015) (Figure 7A). This demonstrates that formation of the F1s occurs
605 at a higher frequency in different directions at the different sites.

606 Once the parental identities of hybrids were confirmed through sequencing, we used this subset
607 of hybrids to test whether the traits described above were influenced by parentage. For
608 example, we did not find any influence of parentage on hybrid seed set or fertility, with no
609 significant difference in either measurement (total seed set: t = -1.194, df = 31.838, p = 0.241;
610 fertility: t = 0.180, df = 30.493, p = 0.859; total fertile seeds: t = -0.954, df = 34.211, p = 0.347)
611 between *G. conopsea* maternal or *G. rhellicani* maternal hybrids. We also did not find any
612 evidence of parentage influencing VOC production, as individual and global VOC trends are not
613 influenced by the hybrid maternal parent (Figure S7).

614 *MYB1 and ANS genotypes reflect hybrid colour and parentage*

615 (Kellenberger et al., 2019) demonstrate the overdominance of a non-synonymous SNP at
616 position 663 with three allelic states in the R2R3 MYB transcription factor, *MYB1*, in the *G.*
617 *rhellicani* parent. The ancestral or 'wild-type' state is a C, with other states being either a G or an
618 A in the last exon of the coding region. This results in a premature stop codon, truncated protein
619 and downstream reduction in the expression of target-gene *ANTHOCYANIDIN SYNTHASE* (*ANS*)
620 in the non-wild-type alleles. This ultimately leads to a reduction in the production of cyanidins
621 in *Gymnadenia rhellicani*. The 'black' (very dark red) *G. rhellicani* morph is homozygous for the
622 ancestral state (C), the white morph never contains the wild-type (C) SNP, and the light red
623 morph is heterozygous, carrying one copy of the (C). We independently confirmed this in our
624 whole genome sequencing of white, light red and dark red *G. rhellicani* from Puflatsch. All

individuals of *G. conopsea* sequenced at both sites were homozygous for the C allele at position 663. This same SNP is homozygous C/C in most of the hybrids, as expected for F1s between the most common 'black' *G. rhellicani* and *G. conopsea* (Figure 7B, Figure S8). However, an additional SNP conferring a G to A transition at position 688 is found to be homozygous in *G. conopsea* and heterozygous in the magenta hybrids This is a non-synonymous SNP, conferring an aspartic acid to asparagine residue shift (Figure S8). This further supports the analyses described above confirming that the hybrids are in fact F1s. Interestingly, the single "pale pink" hybrid we sequenced (which had a notably paler floral colour in the field) is heterozygous (C/G) for the same SNP at position 663 derived from the white *G. rhellicani* sequenced and did not have this secondary heterozygous SNP at 688 found in the other hybrids (Figure 2B, Figure S8). Other paler hybrids were also observed at Puflatsch in 2015-2016 – these likely also result from a cross between the white *G. rhellicani* morph and *G. conopsea*, but we were unable to sequence them.

We also looked at the *ANTHOCYANIDIN SYNTHASE* (*ANS*) characterised by (Kellenberger et al., 2019). The *Gymnadenia conopsea* allele has a series of SNPs at the end of the sequence, in addition to a truncated sequence, when compared to the *G. rhellicani* sequence (Figure S9A). Once again, we found the hybrids to be heterozygous for the two parental species alleles, with single copy of the full length and truncated sequence inherited from the *G. rhellicani* and *G. conopsea* parent, respectively (Figure S9A). This once again confirms the identity of the hybrids as F1s. Some preliminary sequence analysis found the *G. conopsea* allele to indeed confer a truncated protein due to a frameshift mutation (Figure S9B), however, we cannot determine at this point whether this truncation confers a change in protein function.

Discussion

We tested four potential barriers to hybrid formation and found that pollination and, to a much lesser extent, phenology likely act as reproductive isolating barriers between *Gymnadenia conopsea* and *G. rhellicani*, while hybrid seed-set and ploidy likely play a lesser role. Specialised floral traits can act as a pollinator-filtering pre-zygotic barrier to hybrid formation. The proposed parent species of alpine orchid hybrids, *G. conopsea* and *G. rhellicani* differ in an array of floral traits known to be important factors in attracting pollinators – including their floral morphology, pigment and scent. Hybrids are largely intermediate between parents in morphology and perception of floral pigments; however, their anthocyanin profile is the only floral phenotype that replicates that of one parent, in this case the *G. conopsea* parent. This is reflected in the visual perception of orchids by potential pollinators, which is closer to that of *G. conopsea*. Floral scent in hybrids is more variable and not consistent even within closely related volatile organic compounds (VOCs). All hybrids sequenced were found to be F1s with bidirectional site-specific parental origins, in that hybrid populations significantly differed in the predominant maternal parent depending on their sampling site.

Here we have characterized the production and perception of the characteristic magenta-pink pigments of these hybrids and their aromatic-rich scent profile. Understanding that pigment biosynthesis and scent production are biochemically linked is important when considering that these orchids make both anthocyanin pigments and aromatic volatiles as the majority of their scent bouquet – both of which are derived from the common precursor amino acid phenylalanine (Dudareva et al., 2013; Li et al., 2024). *Gymnadenia* is a chemically diverse genus with 47 volatiles across both species, the majority of which by total amount are phenylalanine derived. Hybrids globally seem 'intermediate' in scent between the two parent species, with no clear pattern of individual volatile inheritance or dominance from either parent species, even

671 for very biosynthetically closely related volatiles such as phenylacetaldehyde, phenylethyl
672 acetate, and phenylethyl alcohol. Major pollinators of *Gymnadenia* are known to respond to
673 floral scent, including some of the volatiles characterised here, including benzaldehyde,
674 phenylacetaldehyde, benzyl acetate, phenylethyl acetate and eugenol (Huber et al., 2005).
675 Further behavioural response analysis through trapping found phenylacetaldehyde to be of the
676 most significantly attractive volatiles when tested alone (Huber et al., 2005). It is important to
677 note that these previous studies focused on *G. densiflora*, *G. odoratissima* , and to a lesser extent,
678 *G. conopsea* populations found in Sweden (Chapurlat et al., 2018). This study is the first to
679 characterise *G. rhellicani*, *G. conopsea* and their hybrids in their zones of sympatry.

680 The intermediate nature of hybrid scent is mirrored in their total phenylalanine-derived liquid
681 metabolites, reflecting their genetic composition as F1s or primary hybrids. The striking
682 exception to this pattern was the anthocyanin profile of the hybrids at 525nm, which reflect
683 only the *G. conopsea* parent, regardless of whether this parent is maternal or paternal. It
684 remains unclear as to why the hybrids produce the same anthocyanin profile as *G. conopsea* and
685 not *G. rhellicani* or an intermediate anthocyanin profile, as would be expected based on the
686 other metabolites and traits characterised in this study. It is possible that *G. conopsea* alleles are
687 dominant to *G. rhellicani* in the biosynthesis pathway and/or regulators leading to anthocyanin
688 production, or that the G. conopsea parent confers alleles to the hybrids that further modify the
689 profile through the inclusion of additional enzymatic steps. Alternatively, or in addition, there
690 may be linkage between the anthocyanin pathway genes and genes affecting survival in the
691 hybrids, such as those affecting mycorrhizal recruitment necessary for germination or other
692 survival factors (Li et al., 2021). It would be expected that hybrids may produce biochemically
693 distinct anthocyanins when accounting for their intermediate metabolites. Further genetic
694 analysis will elucidate the role of key metabolism genes and their regulators in the shikimate
695 pathway in both parents and hybrids. This is the pathway which produces both anthocyanins
696 (which are the most abundant pigment type in this system) and aromatic scent compounds
697 (which are the most abundant scent components in this system) (Li et al., 2024; Ma & Constabel,
698 2019; Narbona et al., 2021; Saigo et al., 2020).

699 We have independently confirmed that at least one of the "pale pink" hybrids is in fact derived
700 from hybridisation between *G. conopsea* and white-morph *G. rhellicani*. It is worth noting that
701 the white *G. rhellicani* has only been found in Puflatsch and Monte Bondone, to date
702 (Kellenberger et al., 2019). Moreover, this confirms that this bidirectionality in hybridisation
703 occurs regardless of the colour of the *G. rhellicani* parent, suggesting that pigment-mediated
704 pollinator attraction does not influence the production of these F1 hybrids. (Kellenberger et al.,
705 2019) also show that both bees and flies, despite opposing colour preferences, visited the 'red'
706 morph *G. rhellicani* in high frequency, resulting in their increased prevalence and fitness. It is
707 possible that fewer "pale" hybrids are found in Puflatsch simply because fewer white morph *G.*
708 *rhellicani* are present in the population (ca. 10% of the total *G. rhellicani* surveyed)
709 (Kellenberger et al., 2019). Further studies will interrogate the apparent site-specific
710 relationship between anthocyanin biosynthesis genes, in addition to the elucidation of the
711 complete anthocyanin biosynthesis pathway in these orchids, with pollinator perception and
712 observation studies. This will elucidate the role that this biochemical process, alongside the
713 production of aromatic VOCs, has in the attraction of pollinators and thus the evolution and
714 maintenance of these populations.

715 Consistent with all hybrids sequenced being F1s, we found the two anthocyanin biosynthesis
716 genes characterised by Kellenberger et al., (2019) to be heterozygous in hybrids. However, this
717 does not explain why hybrids only produce *G. conopsea* anthocyanins. Whilst this might suggest
718 that these genes do not play a role in the difference observed in hybrid anthocyanins, it is

important to note that at least one of the pale hybrids, only found in Puflatsch to date, has a white-morph *G. rhellicani* parent. This confirms that the white *G. rhellicani MYB1* allele does confer reduced anthocyanin production in hybrids as they become pale magenta. This implies that the differences in anthocyanin biosynthesis observed between *G. rhellicani* and *G. conopsea* happen at the tip of the pathway, whereby the pale pink hybrid receives less total anthocyanin influx due to the non-functional MYB1 allele inherited from *G. rhellicani* but ends up with the same profile due to inherited *G. conopsea* elements acting downstream of MYB1. The *ANS* allele in *G. conopsea* does confer an amino acid change leading to a premature stop and truncated protein sequence, when compared to *G. rhellicani*, but it is yet unconfirmed whether this confers a change in protein function. It is likely that both the transcriptional regulator and downstream biosynthesis genes play a significant role in mediating pollinator interactions and acting to prevent the formation of hybrids, as has been observed in *Phlox* (Hopkins & Rausher, 2012).

When considering the hybrid identity via genome proportion analysis and allelic heterozygosity, we were surprised to find all hybrids to be primary hybrids. Whilst making up a small proportion of the population, hybrids do produce viable seed. It is possible that F1 hybrids are simply rare enough (<1% of the population at Puflatsch) that we do not observe advanced generation hybrids or backcrosses; alternatively, there may be intrinsic factors rendering these alternatives inviable, or we may have overlooked these plants in the field. However, there have been rare photographic reports of putative backcrosses in the literature (Gerbaud et al., 1999), and their floral colour is different enough from the parent species that they are unlikely to have been overlooked in our comprehensive surveys, where we sampled every hybrid available. As demonstrated above, the hybrids produce a small but viable seed set. It is possible that backcrossed individuals do exist but are no longer phenotypically distinct in the way that *G. x sauveolens* is clearly identifiable in the field and may phenotypically resemble a single parent instead. However, backcrossed individuals were not found in the sequencing of parent species, which are genetically distinct, across the two sites. The distinct genetic make-up of the parent species and hybrids found in our sequencing, along with the scarcity of F1 hybrids found at the field sites makes the possibility of phenotypically indistinct backcrossed generations unlikely. It is also possible that factors involving the environment and soil microbiome required for germination of these orchids act as an additional post-zygotic barrier (Li et al., 2021; Lin et al., 2020). In addition to determining that all sequenced hybrids are F1s, we found the formation of hybrids to occur bidirectionally with a significant difference in pollinia transfer direction based on location, which contrasts with the prediction of only unidirectional formation from the literature based on differing parental nectar spur lengths (Claessens & Kleynen, 2011; Vöth, 2000). However, hybrid phenotypes were found to be intermediate and variable regardless of their parentage; as a result, egg or sperm donor identity cannot be identified based solely on the hybrid phenotype. This reflects the bidirectionality study conducted by (Hedrén et al., 2018). This indeterminable identification of hybrid parentage includes scent and pigmentation with one exception; we could predict that a pale hybrid has a white-morph *G. rhellicani* parent, although without genetic analysis we would not be able to conclude which parent (egg or sperm donor) was the white-morph *G. rhellicani*. Unfortunately, due to the low proportion of pale hybrids we still do not know whether this is consistent for maternal or paternal parents, but it is likely to follow the same trend as the other traits.

It is possible that at other sites *G. rhellicani* may hybridise with *G. odoratissima* instead of *G. conopsea*, a hybrid combination known as *G. x heufleri* that is observed occasionally in the zone of sympatry between all three species (Delforge, 2006). However, *G. odoratissima* was not found at either Chandolin or Puflatsch in any of the surveyed years (2014-2016, 2021-2023), removing it from consideration as a parent species. We were also able to rule out the sympatric, closely related *Pseudorchis albida* (previously *Gymnadenia albida*) as a parent based on our

genomic analyses, although this species was present at both of our study sites (Reinhammar et al., 2002). Further sequencing analysis would investigate any evidence of introgression between the *Gymnadenia* orchids, including *G. odoratissima* and *P. albida*, and whether this potential introgression is reflected by sampling location. Knowing that patterns of directional hybrid formation happen at different frequencies at the two sites would suggest that historical patterns of introgression may also be spatially distinct.

Based on their morphology, in particular the angles of resupination and spur length characterised here, it was proposed that pollen transfer would only occur in the direction of *G. rhellicani* to *G. conopsea*. This hypothesis of singular directionality would suggest that all hybrids would have a *G. conopsea* maternal parent and *G. rhellicani* paternal parent (Claessens & Kleynen, 2011; Vöth, 2000). However, our work aligns with that of (Hedrén et al., 2018) in which bidirectionality occurs in pollen transfer and both parent species act as paternal and maternal parents, although that study had a much smaller sample size. Interestingly, our study included two field sites in which the parent species occur in sympatry, 309 kilometres apart. Through whole genome sequencing we found that this bidirectionality of pollen transfer is site-specific, and statistically significant, with the majority of Puflatsch hybrids being *G. rhellicani* maternal and the majority of Chandolin hybrids being *G. conopsea* maternal. This suggests that pollination of the parent species varies between the two sites. It is possible that different pollinator groups differ in their effective pollination both temporally and spatially, giving rise to the hybrids of varying parentage at each sample site (Chapurlat et al., 2015; Meyer et al., 2007; Sletvold et al., 2012). Thus, producing a "geographic mosaic" of effective pollinators leading to a mosaic of maternal identities (Thompson, 2005).

In their study (Hedrén et al., 2018) suggest three possible mechanisms of interspecies pollen transfer. This first suggesting that long-tongued Lepidoptera may carry *G. conopsea* pollinaria at a distal position of their probosces to *G. rhellicani* infrequently. The second postulating that, rather than long-tongued Lepidoptera, short-tongued microlepidoptera and other insects occasionally visiting *G. conopsea* can carry pollen to the stigma of *G. rhellicani*. Finally, that another, unaccounted for, group of potentially randomised insects confer pollinia transfer completely at random. (Hedrén et al., 2018). One such pollinator we hypothesise has mediated the formation of hybrids is likely to be moths in the family Zygaenidae (particularly *Zygaena* spp.) as they have been observed on both species as well as having been observed in pairs. This which supports Hedrén's first hypothesis (1) regarding long-tongued Lepidopterans (the average proboscis length of *Zygaena* spp. is ca. 10.4mm, from observations in (Müller, 1881) which is close to the average spur length of *G. conopsea* at our sites, 11mm. Moths of this group are also found to be as threatened by changing climate as the orchids themselves (Kolanowska et al., 2021). It is also possible that other species have haphazardly carried pollinia on various points of their bodies and probed the densely packed inflorescences, as proposed by Hedrén et al.,'s third hypothesis (3). Pollinia have been observed on the bodies (e.g. under the forelegs or on the ventral metathorax) of two pollinators, a muscid fly (Muscidae) and a noctuid moth (*Cucullia umbratica*, Noctuidae) collected in 2015 on *G. rhellicani* and *G. conopsea* respectively. This would also provide some explanation as to why the hybrids are so scarce – there may simply be very few opportunities for pollinia to be transferred between the parent species due to their differences in resupination and spur length leading to differential pollinia placement on pollinators. This may also provide explanation as to why backcrosses or F2 generations have not been observed. Further work will elucidate exactly which pollinators have high specificity for a single *Gymnadenia* orchid and which can be observed transferring pollinia between the parent species. To do this, longer and more detailed observations are required, alongside the capture of any floral visitors observed interacting with both progenitors. In-depth analysis of pollinia placement on pollinators will tie together how hybrids are able to form in both

directions and at both field sites despite the pre-zygotic barrier that is the floral angle of resupination and spur length.

Recent work has proposed the concept of a "functional specialist and ecological generalist", in which a floral structure can be specialised to a pollinator guild but still be pollinated by a range of more generalist pollinators in order to increase fitness (Ohashi et al., 2021; Silló & Claßen-Bockhoff, 2024; Katherine E Wenzell et al., 2024). It is clear from their morphological and biochemical profiles that these closely related *Gymnadenia* species are functional specialists, however their historical pollinator observations and broad range of potential pollinators demonstrated through pan-trapping of the sites may allow for more generalist pollinia transfer, particularly in the more ecologically generalist *G. rhellicani*, which would be less common – hence the scarcity of the hybrids. Vital work going forward will also be to study how changing climate and the variation observed in temporality of both flowering time and potential pollinator emergence influences these highly coordinated ecological interactions. *G. rhellicani* is already under threat from rapid environmental change, with all climate change models predicting a reduction of available niches as well as pollinator habitats (Geppert et al., 2020; Kolanowska et al., 2021).

One of the major challenges of this work and exciting avenues for further research is in integrating abiotic factors, in particular climate. Sampling in 2022 was inhibited by extreme weather conditions leading to most plants being senesced or aborting their inflorescences prior to our visit at the typical flowering peak time. Further research will characterise what effect the trends currently being observed in changing climate have on this system. We are yet to fully understand how changing climate will also influence pollinator emergence. It is possible that with rapid changing climate that hybrids will increase or decrease in frequency, a phenomenon that would be interesting to model. It is apparent from this study that the emergence and maintenance of these hybrids varies geographically, likely as a result of their local pollinator ecology. Studies such as this will become essential when predicting how changing climate will affect systems such as these charismatic orchids, which are already under threat.

Acknowledgements

We thank the relevant authorities (the Dienststelle für Wald, Flussbau und Landschaft, Kanton Wallis, Switzerland and the Amt für Natur, Autonome Provinz Bozen-Südtirol, Italy) for permission to conduct this research and collect samples of protected species at Chandolin and Puflatsch, respectively. The authors would also like to thank Florian Schiestl, Mathias Kneubühler, Richard Lorenz, Barbara Baldan and Sebastiano Nigris and Benjamin Kellenberger for various efforts supporting the conducting of fieldwork from 2014 to 2023. We thank Barbara Keller, Franz Huber, Lena Stransfeld and Karin Gross for technical support. We additionally wish to thank the JIC metabolomics and Bioinformatics platforms for technical support, in particular Paul Brett, Saleha Bakht and Burkhard Steuernagel. We would like to thank Hannah Gunn for help in manuscript preparation. We would like to thank Florian Schiestl for support in funding acquisition. We would like to thank the following funding bodies for supporting this work; University of Zurich (UZH) Forschungskredit grant to R.T.K., an EU SP7 PLANT FELLOWS Postdoctoral Fellowship grant (GA-2010-267243) to K.J.R.P.B., and three travel grants from the Claraz Schenkung Zurich (two to R.T.K. and one to K.J.R.P.B.). P.M.S. acknowledges support by the Swiss National Science Foundation (SNF, project 31003A_155943)." This work was also supported by a Research Grant from the Royal Society to Byers (RGS\R1\211279) and start-up funding from the John Innes Centre, including funding

from the GEN (BB/P013511/1), BRiC (BB/X01102X/1), and HBio (BB/X01097X/1) Institute Strategic Programmes funded by the BBSRC.

References

Arditti, J. (2003). Resupination. *Lankesteriana*, 3(2).

Arditti, J., & Ghani, A. (2000). Numerical and physical properties of orchid seeds and their biological implications (vol 145, pg 367, 2000). *New Phytologist*, 146(3), 569-569.

Armbruster, W. S., Shi, X.-Q., & Huang, S.-Q. (2014). Do specialized flowers promote reproductive isolation? Realized pollination accuracy of three sympatric Pedicularis species. *Annals of botany*, 113(2), 331-340.

Bateman, R. M. (2021). Phenotypic versus genotypic disparity in the Eurasian orchid genus Gymnadenia: exploring the limits of phylogeny reconstruction. *Systematics and Biodiversity*, 19(4), 400-422.

Bradshaw, H., & Schemske, D. W. (2003). Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*, 426(6963), 176-178.

Brandrud, M. K., Paun, O., Lorenz, R., Baar, J., & Hedrén, M. (2019). Restriction-site associated DNA sequencing supports a sister group relationship of Nigritella and Gymnadenia (Orchidaceae). *Molecular phylogenetics and evolution*, 136, 21-28.

Brzosko, E., & Mirski, P. (2021). Floral nectar chemistry in orchids: A short review and meta-analysis. *Plants*, 10(11), 2315.

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., Schijlen, E. G., Hall, R. D., Bovy, A. G., & Luo, J. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature biotechnology*, 26(11), 1301-1308.

Cafasso, D., Widmer, A., & Cozzolino, S. (2005). Chloroplast DNA inheritance in the orchid Anacamptis palustris using single-seed polymerase chain reaction. *Journal of Heredity*, 96(1), 66-70.

Cardoso, J., Johnson, S., & Oliveira, P. (2024). Incomplete resupination during floral development leads to pollination failure in a slipper orchid. *Plant Biology*, 26(1), 34-40.

Chapurlat, E., Ågren, J., Anderson, J., Friberg, M., & Sletvold, N. (2019). Conflicting selection on floral scent emission in the orchid Gymnadenia conopsea. *New Phytologist*, 222(4), 2009-2022.

Chapurlat, E., Ågren, J., & Sletvold, N. (2015). Spatial variation in pollinator-mediated selection on phenology, floral display and spur length in the orchid Gymnadenia conopsea. *New Phytologist*, 208(4), 1264-1275.

Chapurlat, E., Anderson, J., Ågren, J., Friberg, M., & Sletvold, N. (2018). Diel pattern of floral scent emission matches the relative importance of diurnal and nocturnal pollinators in populations of Gymnadenia conopsea. *Annals of botany*, 121(4), 711-721.

Chapurlat, E., Le Roncé, I., Ågren, J., & Sletvold, N. (2020). Divergent selection on flowering phenology but not on floral morphology between two closely related orchids. *Ecology and evolution*, 10(12), 5737-5747.

Christie, K., Fraser, L. S., & Lowry, D. B. (2022). The strength of reproductive isolating barriers in seed plants: Insights from studies quantifying premating and postmating reproductive barriers over the past 15 years. *Evolution*, 76(10), 2228-2243.

Claessens, J., & Kleynen, J. (2011). The flower of the European orchid: form and function.

Claessens, J., & Kleynen, J. (2016). The Pollination of European Orchids Part 4: Goodyera and Spiranthes. *Journal of the hardy orchid society*, 13, 54-61.

Corriveau, J. L., & Coleman, A. W. (1988). Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American journal of botany*, 75(10), 1443-1458.

Coyne, J. A., & Orr, H. A. (2004). Speciation: a catalogue and critique of species concepts. Philosophy of biology: an anthology, 272-292.

D'Auria, M., Lorenz, R., Mecca, M., Racioppi, R., Romano, V. A., & Viggiani, L. (2020). Fragrance components of Gymnadenia conopsea and Gymnadenia odoratissima collected at several sites in Italy and Germany. *Natural Product Research*, 1-5.

Delforge, P. (1994). Guide des orchidées d'Europe, d'Afrique du Nord et du Proche-Orient. *Delachaux et Niestlé Lausanne*.

Delforge, P. (2006). Orchids of Europe, North Africa and the Middle East. *Timber Press*.

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature protocols*, 2(9), 2233-2244.

Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist*, 198(1), 16-32.

Faegri, K., & Van der Pijl, L. (1979). A short history of the study of pollination ecology. *The Principles of pollination ecology*, 3, 1-77.

Fenster, C. B., Armbruster, W. S., Wilson, P., Dudash, M. R., & Thomson, J. D. (2004). Pollination syndromes and floral specialization. *Annu. Rev. Ecol. Evol. Syst.*, 35, 375-403.

Frachon, L., Stirling, S. A., Schiestl, F. P., & Dudareva, N. (2021). Combining biotechnology and evolution for understanding the mechanisms of pollinator attraction. *Current Opinion in Biotechnology*, 70, 213-219.

Geppert, C., Perazza, G., Wilson, R. J., Bertolli, A., Prosser, F., Melchiori, G., & Marini, L. (2020). Consistent population declines but idiosyncratic range shifts in Alpine orchids under global change. *Nature communications*, 11(1), 5835.

Gerbaud, O., d'orchidophilie, S. f., Walter Schmid, B., & Schmid, W. (1999). Les hybrides des genres' Nigritella'et/ou'Pseudorchis'. *Société française d'orchidophilie*.

Grant, V. (1994). Modes and origins of mechanical and ethological isolation in angiosperms. *Proceedings of the National Academy of Sciences*, 91(1), 3-10.

Gross, K., & Schiestl, F. P. (2015). Are tetraploids more successful? Floral signals, reproductive success and floral isolation in mixed-ploidy populations of a terrestrial orchid. *Annals of botany*, 115(2), 263-273.

Hackney, P. (2007). Orchids of Europe, North Africa and the Middle East. *JSTOR*.

Hedrén, M., Lorenz, R., & Ståhlberg, D. (2018). Evidence for bidirectional hybridization between Gymnadenia and Nigritella. *Journal Europäischer Orchideen*, 50(1), 43-60.

Hopkins, R., & Rausher, M. D. (2012). Pollinator-mediated selection on flower color allele drives reinforcement. *Science*, 335(6072), 1090-1092.

Huber, F. K., Kaiser, R., Sauter, W., & Schiestl, F. P. (2005). Floral scent emission and pollinator attraction in two species of Gymnadenia (Orchidaceae). *Oecologia*, 142(4), 564-575.

Jersáková, J., Castro, S., Sonk, N., Milchreit, K., Schödelbauerová, I., Tolasch, T., & Dötterl, S. (2010). Absence of pollinator-mediated premating barriers in mixed-ploidy populations of Gymnadenia conopsea sl (Orchidaceae). *Evolutionary Ecology*, 24(5), 1199-1218.

Jia, L. B., & Huang, S. Q. (2022). An examination of nectar production in 34 species of Dendrobium indicates that deceptive pollination in the orchids is not popular. *Journal of Systematics and Evolution*, 60(6), 1371-1377.

Jones, C. E. (1978). Pollinator constancy as a pre-pollination isolating mechanism between sympatric species of Cercidium. *Evolution*, 189-198.

Katzer, A. M., Wessinger, C. A., & Hileman, L. C. (2019). Nectary size is a pollination syndrome trait in Penstemon. *New Phytologist*, 223(1), 377-384.

Kellenberger, R. T., Byers, K. J., Francisco, R. M. D. B., Staedler, Y. M., LaFountain, A. M., Schönenberger, J., Schiestl, F. P., & Schlüter, P. M. (2019). Emergence of a floral colour polymorphism by pollinator-mediated overdominance. *Nature communications*, 10(1), 1-11.

Kolanowska, M., Rewicz, A., & Nowak, S. (2021). Significant habitat loss of the black vanilla orchid (Nigritella nigra sl, Orchidaceae) and shifts in its pollinators availability as results of global warming. *Global Ecology and Conservation*, 27, e01560.

Levin, D. A. (1978). The origin of isolating mechanisms in flowering plants. In Evolutionary biology (pp. 185-317). *Springer*.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), 1754-1760.

Li, T., Wu, S., Yang, W., Selosse, M.-A., & Gao, J. (2021). How mycorrhizal associations influence orchid distribution and population dynamics. *Frontiers in Plant Science*, 12, 647114.

Li, Y., Bao, T., Zhang, J., Li, H., Shan, X., Yan, H., Kimani, S., Zhang, L., & Gao, X. (2024). The coordinated interaction or regulation between floral pigments and volatile organic compounds. *Horticultural Plant Journal*.

Lin, M., Xiong, H., Xiang, X., Zhou, Z., Liang, L., & Mei, Z. (2020). The effect of plant geographical location and developmental stage on root-associated microbiomes of Gymnadenia conopsea. *Frontiers in microbiology*, 11, 529358.

Ma, D., & Constabel, C. P. (2019). MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends in Plant Science*, 24(3), 275-289.

Meyer, B., Kröger, J., & Steffan-Dewenter, I. (2007). Contribution of Diurnal and Nocturnal Pollinators to the Reproductive Success of the Orchid species Gymnadenia conopsea.

Michalski, S. G., & Durka, W. (2015). Separation in flowering time contributes to the maintenance of sympatric cryptic plant lineages. *Ecology and evolution*, 5(11), 2172-2184.

Moen, A., & Øien, D. I. (2002). Ecology and survival of Nigritella nigra, a threatened orchid species in Scandinavia. *Nordic journal of botany*, 22(4), 435-461.

Müller, H. (1881). Alpenblumen, ihre Befruchtung durch Insekten und ihre Anpassungen an dieselben. *W. Engelmann*.

Narbona, E., del Valle, J. C., Arista, M., Buide, M. L., & Ortiz, P. L. (2021). Major flower pigments originate different colour signals to pollinators. *Frontiers in ecology and evolution*, 9, 743850.

Ohashi, K., Jürgens, A., & Thomson, J. D. (2021). Trade-off mitigation: a conceptual framework for understanding floral adaptation in multispecies interactions. *Biological Reviews*, 96(5), 2258-2280.

Ollerton, J., Alarcón, R., Waser, N. M., Price, M. V., Watts, S., Cranmer, L., Hingston, A., Peter, C. I., & Rotenberry, J. (2009). A global test of the pollination syndrome hypothesis. *Annals of botany*, 103(9), 1471-1480.

Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, 120(3), 321-326.

Olofsson, C. (2021). Pollination niches of Gymnadenia conopsea and G. densiflora in pure and mixed populations: evidence for character displacement? *DIVA*.

Piñeiro Fernández, L., Byers, K. J., Cai, J., Sedeek, K. E., Kellenberger, R. T., Russo, A., Qi, W., Aquino Fournier, C., & Schlüter, P. M. (2019). A phylogenomic analysis of the floral transcriptomes of sexually deceptive and rewarding European orchids, Ophrys and Gymnadenia. *Frontiers in Plant Science*, 10, 1553.

Pridgeon, A., Bateman, R., Cox, A., Hapeman, J., & Chase, M. (1997). Phylogenetics of subtribe Orchidinae (Orchidoideae, Orchidaceae) based on nuclear ITS sequences. 1. *Intergeneric relationships and polyphyly of Orchis sensu lato*.

Proctor, M., Yeo, P., & Lack, A. (1996). The natural history of pollination.

Raguso, R. A., & Pellmyr, O. (1998). Dynamic headspace analysis of floral volatiles: a comparison of methods. *Oikos*, 238-254.

Ramsey, J., & Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual review of ecology and systematics, 29(1), 467-501.

Ramsey, M., Seed, L., & Vaughton, G. (2003). Delayed selfing and low levels of inbreeding depression in Hibiscus trionum (Malvaceae). *Australian Journal of Botany*, 51(3), 275-281.

Reinhammar, L.-G., Olsson, E. G. A., & SØRMELAND, E. (2002). Conservation biology of an endangered grassland plant species, Pseudorchis albida, with some references to the closely related alpine P. straminea (Orchidaceae). *Botanical Journal of the Linnean Society*, 139(1), 47-66.

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB)–a community resource of plant chromosome numbers. *New Phytologist*, 206(1), 19-26.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature biotechnology*, 29(1), 24-26.

Romero-Bravo, A., & Castellanos, M. C. (2024). Nectar and floral morphology differ in evolutionary potential in novel pollination environments. *New Phytologist*.

Saigo, T., Wang, T., Watanabe, M., & Tohge, T. (2020). Diversity of anthocyanin and proanthocyanin biosynthesis in land plants. *Current Opinion in Plant Biology*, 55, 93-99.

Saunders, M. E., & Luck, G. W. (2013). Pan trap catches of pollinator insects vary with habitat. *Australian Journal of Entomology*, 52(2), 106-113.

Schiestl, F. P., & Schlüter, P. M. (2009). Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual review of entomology*, 54, 425-446.

Silló, N., & Claßen-Bockhoff, R. (2024). A multitude of bee pollinators in a phenotypic specialist-pollinator diversity from the plant's perspective. *Flora*, 152461.

Sletvold, N., & Ågren, J. (2010). Pollinator-mediated selection on floral display and spur length in the orchid Gymnadenia conopsea. *International Journal of Plant Sciences*, 171(9), 999-1009.

Sletvold, N., Trunschke, J., Wimmergren, C., & Ågren, J. (2012). Separating selection by diurnal and nocturnal pollinators on floral display and spur length in Gymnadenia conopsea. *Ecology*, 93(8), 1880-1891.

Sun, M., Gross, K., & Schiestl, F. P. (2014). Floral adaptation to local pollinator guilds in a terrestrial orchid. *Annals of botany*, 113(2), 289-300.

Sun, M., Schlüter, P. M., Gross, K., & Schiestl, F. P. (2015). Floral isolation is the major reproductive barrier between a pair of rewarding orchid sister species. *Journal of Evolutionary Biology*, 28(1), 117-129.

Swift, M. L. (1997). GraphPad prism, data analysis, and scientific graphing. *Journal of chemical information and computer sciences*, 37(2), 411-412.

Team, T. G. D. (2019). GIMP. See https://www. gimp. org.

Telles, F. J., Lind, O., Henze, M. J., Rodríguez-Gironés, M. A., Goyret, J., & Kelber, A. (2014). Out of the blue: the spectral sensitivity of hummingbird hawkmoths. *Journal of Comparative Physiology A*, 200, 537-546.

Thompson, J. N. (2005). Coevolution: the geographic mosaic of coevolutionary arms races. *Current biology*, 15(24), R992-R994.

Vinogradova, T. N., & Andronova, E. V. (2002). Development: Development of Orchid Seeds and Seedlings. *Orchid biology: reviews and perspectives*, VIII (pp. 167-234). Springer.

Vöth, W. (1999). Lebensgeschichte und Bestäuber der Orchideen am Beispiel von Niederösterreich (Vol. 65). Biologiezentrum des OÖ. *Landesmuseums*.

Vöth, W. (2000). Gymnadenia, Nigritella und ihre Bestäuber. *Journal Europäischer Orchideen*, 32(3/4), 547-573.

Vrdoljak, S. M., & Samways, M. J. (2012). Optimising coloured pan traps to survey flower visiting insects. *Journal of Insect Conservation*, 16, 345-354.

Waser, N. M. (1986). Flower constancy: definition, cause, and measurement. *The American Naturalist*, 127(5), 593-603.

Waser, N. M. (1993). Population structure, optimal outbreeding and assortative mating. The natural history of inbreeding and outbreeding: theoretical and empirical perspectives. *University of Chicago Press, Chicago*, 173-199.

Wei, N., Kaczorowski, R. L., Arceo-Gómez, G., O'Neill, E. M., Hayes, R. A., & Ashman, T.-L. (2021). Pollinators contribute to the maintenance of flowering plant diversity. *Nature*, 597(7878), 688-692.

Wenzell, K. E., Zhang, J. Y., Skogen, K. A., & Fant, J. B. (2024). Adaptive generalization in pollination systems: Hawkmoths increase fitness to long-tubed flowers, but secondary pollinators remain important. *Ecology and evolution*, 14(5), e11443.

Wenzell, K. E., Neequaye, M., Paajanen, P., Hill, L., Brett, P., & Byers, K. J. R. P. (2025). Within-species floral evolution reveals convergence in adaptive walks during incipient pollinator shift. *Nature communications*, 16(2721)

Wessinger, C. A. (2021). From pollen dispersal to plant diversification: genetic consequences of pollination mode. *New Phytologist*, 229(6), 3125-3132.

White, R. H., Stevenson, R. D., Bennett, R. R., Cutler, D. E., & Haber, W. A. (1994). Wavelength discrimination and the role of ultraviolet vision in the feeding behavior of hawkmoths. *Biotropica*, 427-435.

Wilson, P. (1995). Selection for pollination success and the mechanical fit of Impatiens flowers around bumblebee bodies. *Biological Journal of the Linnean Society*, 55(4), 355-383.

Yam, T. W., Arditti, J., & Cameron, K. M. (2009). "The orchids have been a splendid sport"—an alternative look at Charles Darwin's contribution to orchid biology. *American journal of botany*, 96(12), 2128-2154.

Yourstone, J., Varadarajan, V., & Olsson, O. (2023). Bumblebee flower constancy and pollen diversity over time. *Behavioral Ecology*, 34(4), 602-612.

Zhang, S., Yang, Y., Li, J., Qin, J., Zhang, W., Huang, W., & Hu, H. (2018). Physiological diversity of orchids. *Plant diversity*, 40(4), 196-208.

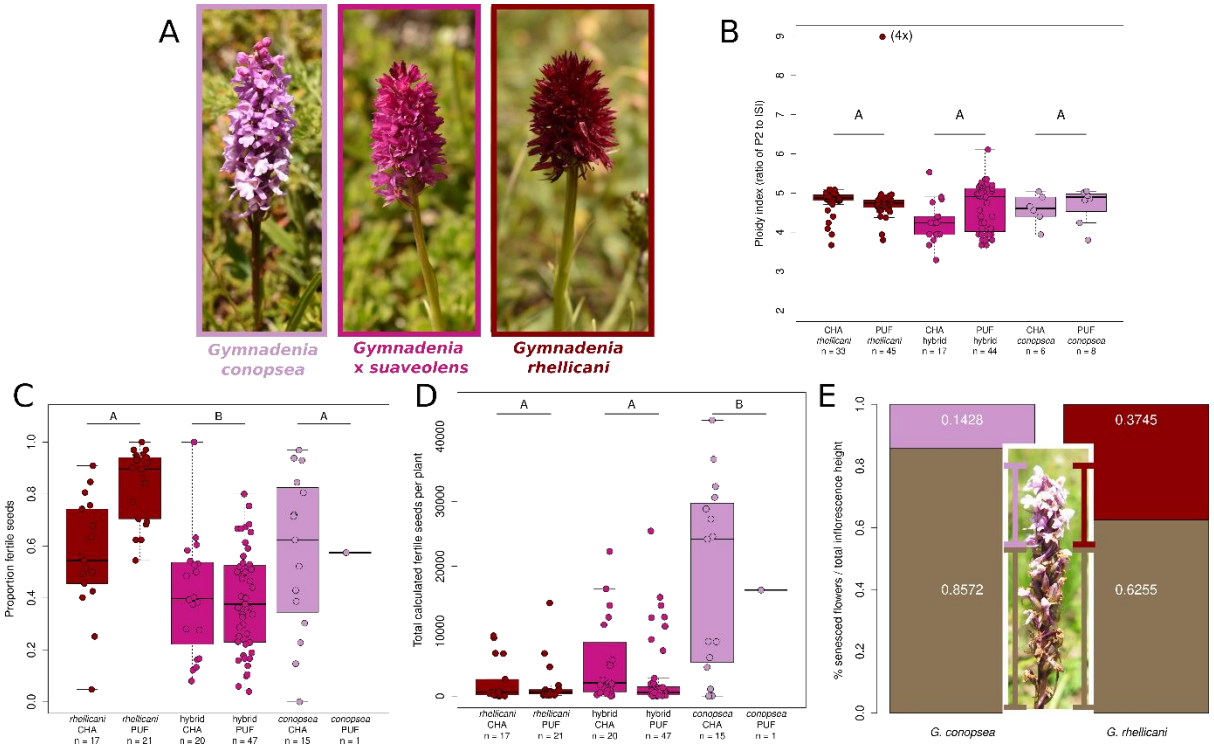


Figure 1: *Gymnadenia* and barriers to hybridization. (A) Focal species studied here: *G. conopsea, G. rhellicani,* and their putative hybrid *G.* x *suaveolens*. (B) Ploidy levels of parent and hybrid orchids from the two field sites expressed as the ratio of orchid pollinia tissue peak II (diploid nuclei) to internal standard (*Phaseolus coccinea*) leaf tissue peak I (diploid nuclei). (C) Proportion of fertile seeds per plant for each species at each field site. (D) Total fertile seeds per plant for each species at each field site. (E) Proportion of senesced (brown) versus open (purple or red) flowers along inflorescences for the two parent species photographed from quadrats at PUF in 2021.

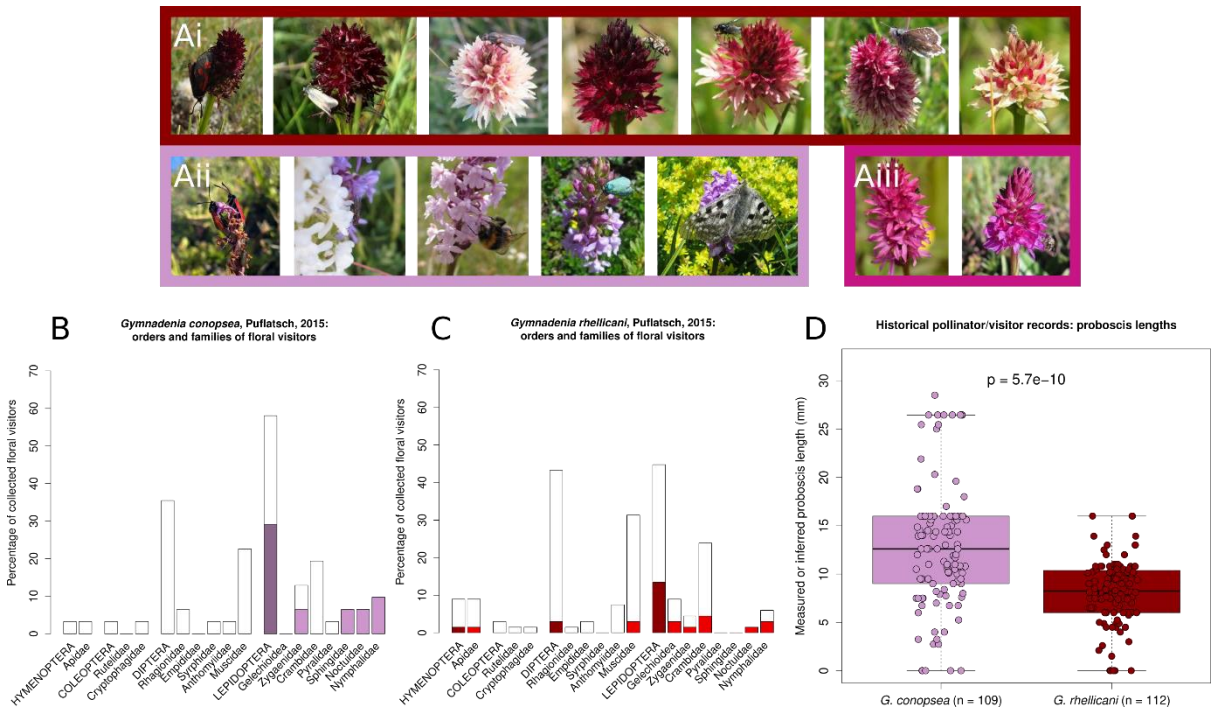


Figure 2: Pollinators and floral visitors to *Gymnadenia*. (Ai) Insects photographed on *G. rhellicani* at PUF and CHA. (Aii) Insects photographed on *G. conopsea* at PUF, CHA, and a site above Davos Dorf, Switzerland (not otherwise studied). (Aiii) Dipteran species photographed on *G.* x *suaveolens*. (B,C) Floral visitors captured on *G. conopsea* (B) and *G. rhellicani* (C) in 2015 at PUF via haphazard sampling across three days. Open bars: all visitors. Colored sub-bars: visitors carrying orchid pollinia. Darker bars represent order totals, paler bars represent family totals. (D) Differences in floral visitor proboscis lengths from the historical literature (excluding the present study).

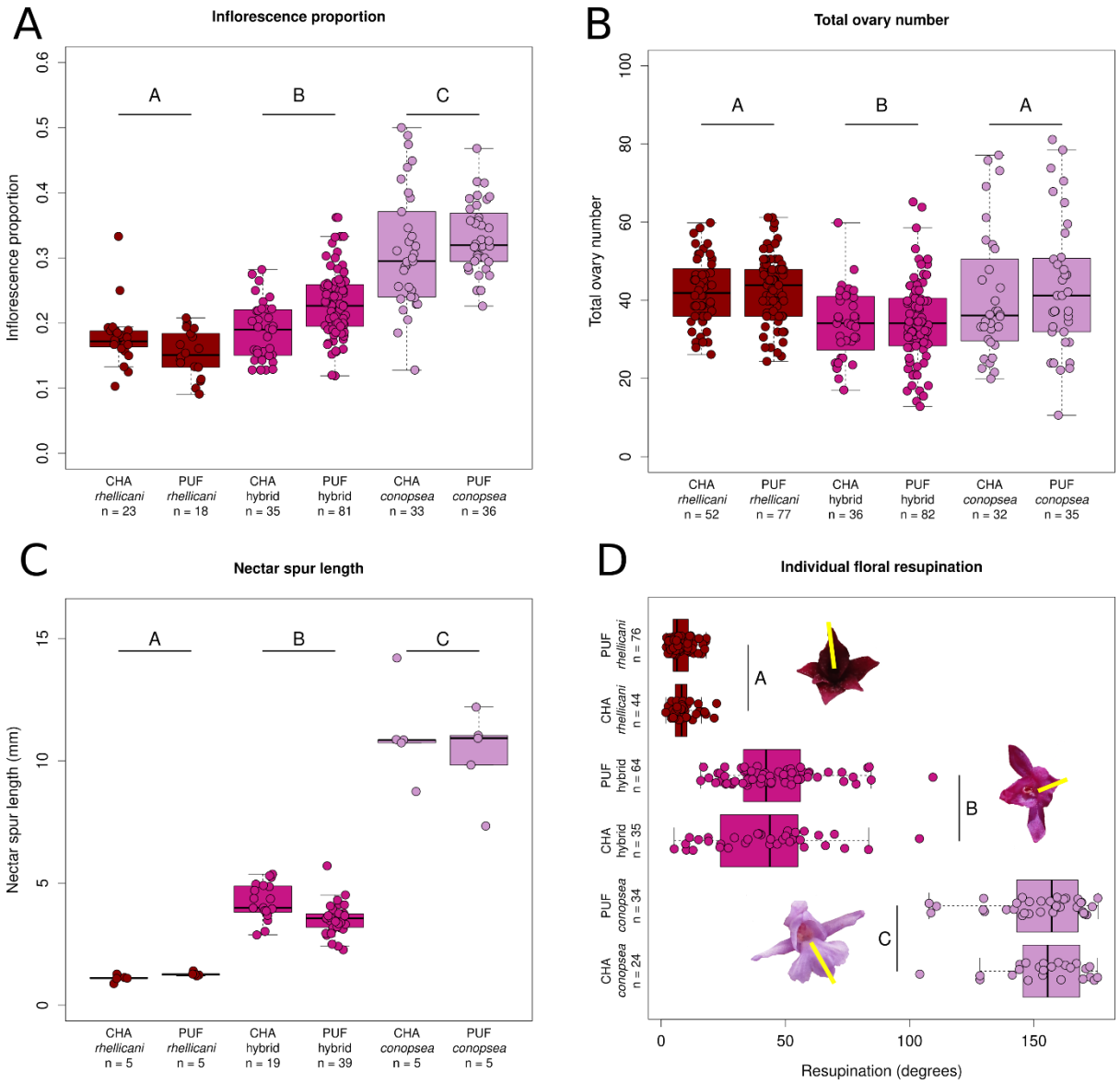


Figure 3: Morphological traits of parent and hybrid orchids. (A) Proportion of total plant height taken by the inflorescence for each species at each site. (B) Total ovary number (senesced flowers, open flowers, and buds) for each species at each site. (C) Nectar spur length in millimetres for each species at each site. (D) Individual floral resupination (rotation) in degrees for each species at each site. Inset images show the average degree of rotation for each species using the yellow line marked across the labellum (largest tepal). See Figure Sx for plots of more morphological traits.

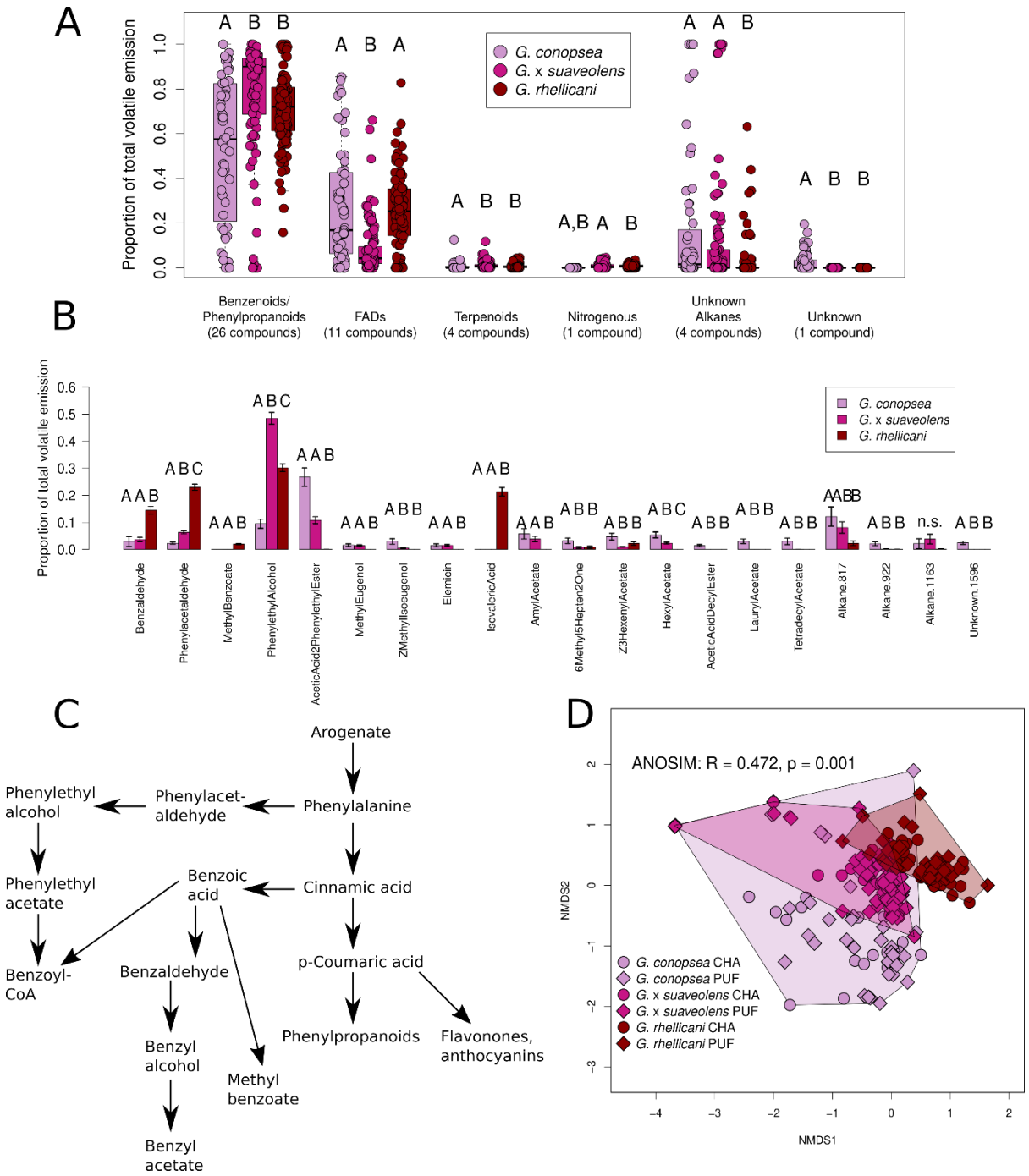


Figure 4: Floral scent of parent and hybrid *Gymnadenia*. (A) Relative (percentage of total) emissions for different volatile compound classes for parent and hybrid orchids. (B) Relative (percentage of total) emissions of major volatiles (those above 1% of total emission in any species) for parent and hybrid orchids. (C) Simplified aromatic volatile biosynthesis pathway (CITE) showing the close relationship of phenylacetaldehyde, phenylethyl alcohol, and phenylethyl acetate (upper left). (D) NMDS plot of relative (percentage of total) emissions of all floral volatiles for parent and hybrid orchids by site.

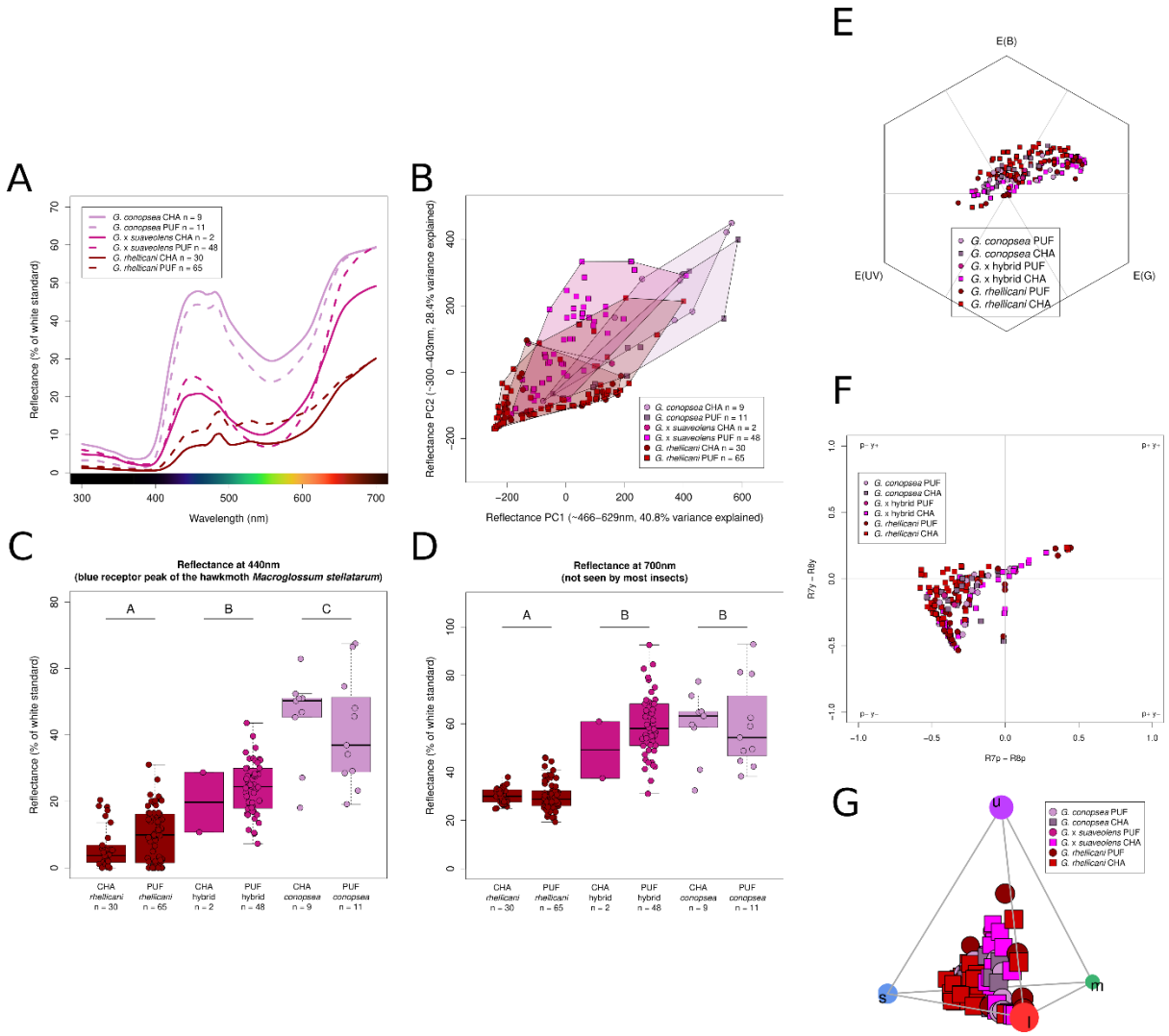


Figure 5: Floral colour and perception by insects. (A) Reflectance spectrophotometry curves in the UV-visible spectrum (300-700nm) for each species at each site. Only median values are shown. (B) PCA of the first two principal components of all reflectance spectrophotometry data for each species at each site. (C,D) Reflectance at two key wavelengths: 440nm (human blue, corresponding to the blue photoreceptor peak of the hawkmoth *Macroglossum stellatarum* (CITE)) and 700nm (human red, not visible to most insects (CITE)). (E) *Apis mellifera* (honey bee) hexagon visual model plot of reflectance spectrophotometry data for each species at each site; the centre point represents the achromatic centre. (F) *Musca domestica* (house fly) visual model plot of reflectance spectrophotometry data for each species at each site; only points in different quadrats are distinguishable. (G) Generic tetrachromat (average avian UV system) visual model plot of reflectance spectrophotometry data for each species at each site; corners represent the four receptors (u = UV; s/m/l = short, medium, and long wavelength receptors respectively).

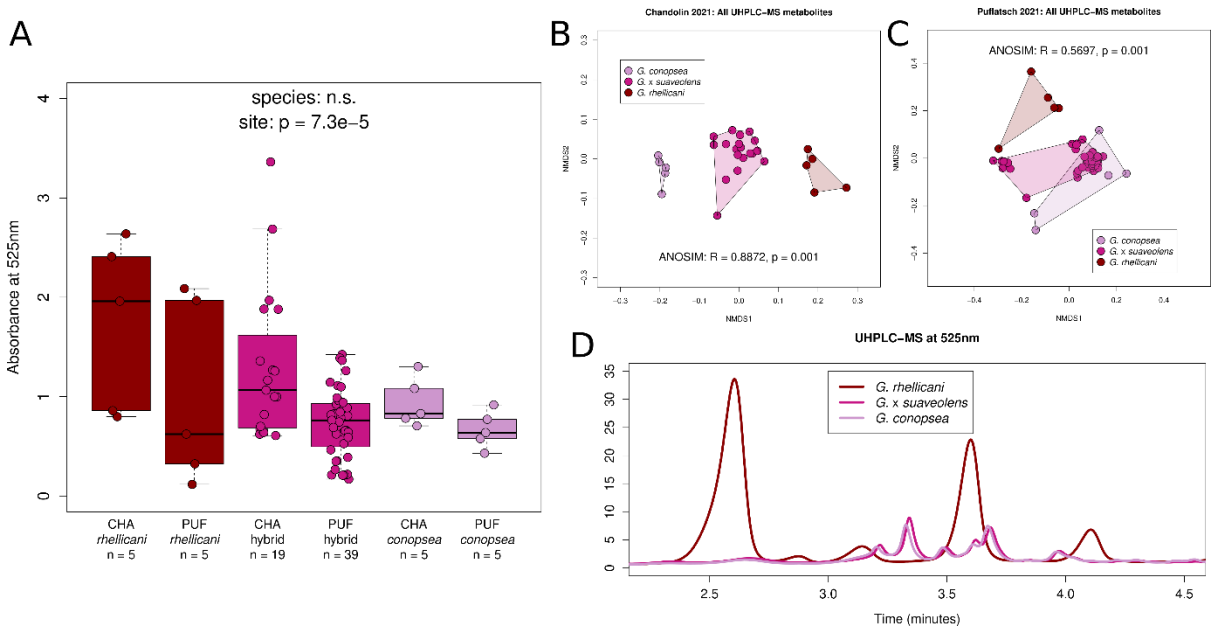
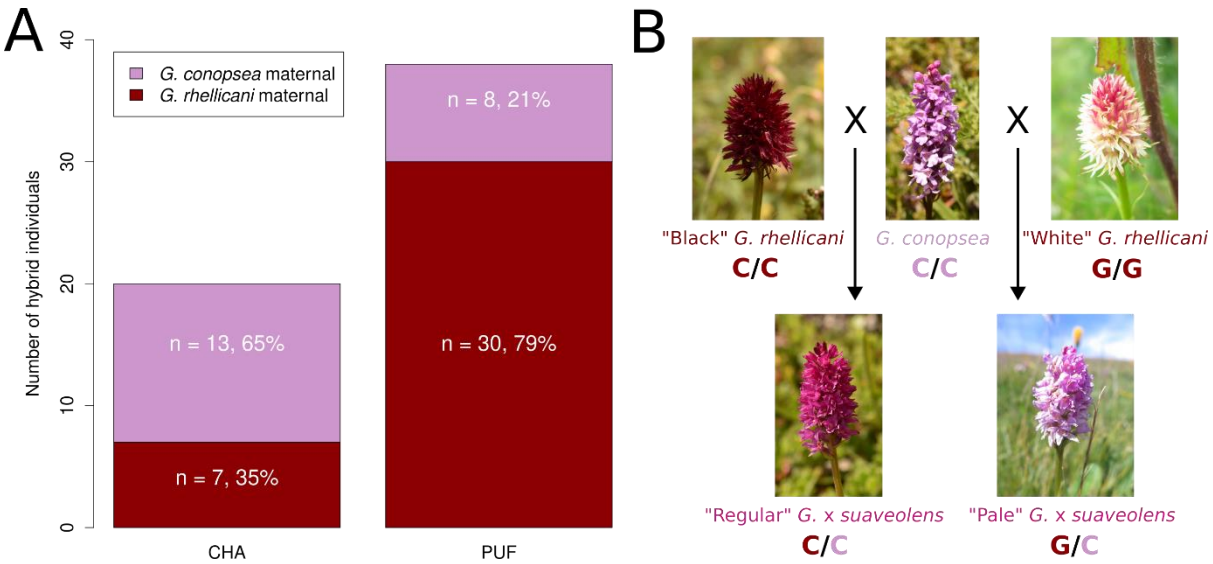


Figure 6: Floral anthocyanins of *Gymnadenia* orchids. (A) Total floral anthocyanins measured via absorbance at 525nm. (B,C) NMDS of all floral UHPLC-MS metabolites sampled at 525nm from CHA (B) and PUF (C) in 2021. (D) UHPLC-MS chromatograms of example parent and hybrid orchids sampled at 525nm.

1138



1139

1140 **Figure 7: Hybrid site-specific maternity and genetic analysis.** (A) Hybrid microsatellite
1141 maternal identity by site. (B) *MYB1* genotypes in hybrid and parent species reflect parental and
1142 hybrid visible floral colour (as described by (Kellenberger et al., 2019) for *G. rhellicani*).