

Taxonomic complexity in eyebrights (*Euphrasia* L.,
Orobanchaceae) and the British flora

Max Brown

2020

Contents

1 General introduction	5
1.1 Taxonomic complexity in plants	5
1.2 The role of hybridisation in taxonomic complexity	6
1.3 The role of polyploidy in taxonomic complexity	7
1.4 Novel features of TCGs; parasitism and plasticity	10
1.5 The genus <i>Euphrasia</i>	12
1.6 The British flora	18
1.7 New methods for investigating taxonomic complexity	18
1.8 Thesis aims	19
2 The emerging importance of cross-ploidy hybridisation and introgression	21
2.1 Abstract	21
2.2 Introduction	21
2.3 Occurrence of natural cross-ploidy hybrids	24
2.4 Biology of cross-ploidy hybrids: general features	35
2.5 Future perspectives	41
3 The genetic landscape of hybridisation in UK flora	43
3.1 Abstract	43
3.2 Main	43
3.3 Methods	51
4 Is there evidence of diploid-tetraploid hybridisation in a <i>Euphrasia</i> contact zone?	55
4.1 Abstract	55
4.2 Introduction	55
4.3 Materials and methods	57
4.3.1 Population sampling and DNA extraction	57
4.3.2 Sanger sequencing and sequence analysis	58
4.3.3 GBS & SNP discovery	59
4.3.4 Identifying hybridisation between ploidy levels	59
4.3.5 Quantification of genetic variability within and between ploidy levels	60
4.3.6 Demographic inference with $\delta\alpha\delta I$	60
4.4 Results	61

4.4.1	Patterns of genetic structure in the ITS1	61
4.4.2	Patterns of genetic structure from GBS data	62
4.4.3	Detecting genomic hybridisation	64
4.4.4	Demographic modelling of gene flow	65
4.5	Discussion	67
5	Horticultural protocols for experimental studies of eyebrights (<i>Euphrasia</i>, Orobanchaceae)	69
5.1	Introduction	69
5.2	General considerations for cultivating <i>Euphrasia</i>	70
5.3	Laboratory conditions	71
5.4	Pot trials	71
5.5	Field trials	75
5.6	Reciprocal transplant experiments	76
5.7	Conclusions and future directions	79
6	Life history evolution, species differences, and phenotypic plasticity in hemiparasitic <i>Euphrasia</i>	80
6.1	Abstract	80
6.2	Introduction	80
6.3	Materials and Methods	83
6.3.1	Experimental design and plant cultivation	83
6.3.2	Species differences experiment	83
6.3.3	Phenotypic plasticity experiment	83
6.3.4	Cultivation protocol	84
6.3.5	Common garden trait measurements and statistical analyses	84
6.3.6	Trait variation in the wild	85
6.4	Results	86
6.4.1	Species differences	86
6.4.2	Phenotypic plasticity	90
6.4.3	Variation in the wild	90
6.5	Discussion	91
6.5.1	Life history variation in a generalist hemiparasitic plant	92
6.5.2	Phenotypic plasticity in response to host	94

6.6	Conclusions	95
7	Conserved and host-specific interactions in a multi-host parasite system	96
7.1	Abstract	96
7.2	Introduction	96
7.3	Methods	98
7.3.1	Plant material, cultivation and trait measurements	98
7.3.2	Statistical analyses	99
7.4	Results	101
7.4.1	Hemiparasite performance across host species	101
7.4.2	Hemiparasite-host interactions	105
7.5	Discussion	106
8	General discussion	109
8.1	The interaction and importance of ploidy and hybridisation	109
8.2	Evidence of parasitism playing a role in taxonomic complexity	109
8.3	Critique and further study	110
8.4	Thesis conclusions	112
	Bibliography	113

1 General introduction

1.1 Taxonomic complexity in plants

Much of biology is built upon the idea that the diversity we see in nature can be categorised into discrete units, which at its heart is the concept of a species (Simpson, 1951). One of the most important species concepts is the biological species concept (BSC), where species are recognised as a community of interfertile individuals (Mayr, 1963). The BSC fits many organisms, and is the foundation of evolutionary research, especially in plants and animals (Butlin and Stankowski, 2020). Although plants contain some notoriously complex groups, most taxonomic species and phenotypic clusters represent reproductively isolated lineages (70%+; Rieseberg et al., 2006). The concept of species in plants is therefore biologically meaningful and useful in the majority, with phenomena such as asexuality and polyploidy at the forefront of species delimitation issues.

Ever since this task of discriminating and categorising species has been embarked upon, difficulties have emerged as to how to delimit certain species, mainly because the processes that give rise to species are in constant motion (Luo et al., 2018). Arguably some of the most difficult groups to classify are collectively placed under the umbrella term of ‘taxonomically complex groups’ (TCGs; Ennos et al., 2005). A TCG is here defined as one where it is difficult to categorise the biodiversity present due to underlying processes which blur species boundaries. TCGs are problematic, not only in terms of classification but also as to how to best conserve them (Ennos et al., 2012; Federici et al., 2013). TCGs are present across the tree of life, from fungi (Leavitt et al., 2011), to fish (Garcia-Melo et al., 2019), and arthropods (Stenberg et al., 2003). TCGs are remarkably frequent in plants however, and particularly common in certain families such as Poaceae (Roodt and Spies, 2003), Rosaceae (Dickinson et al., 2007), and Asteraceae (Czapik, 1996). TCGs can be the result of plant populations that no longer undergo sexual, random mating (Hollingsworth et al., 2006), but other TCGs are the result of phenotypic plasticity and recent speciation blurring species boundaries (Belton et al., 2014; Wang et al., 2018).

Factors commonly contributing to TCGs include selfing, apomixis (or agamospermy), hybridisation, and polyploidy. Selfing and apomixis disturb random mating and restrict gene flow between populations, and often interact with hybridisation and polyploidy (Hollingsworth et al., 2006). For example, apomixis coupled with rare hybridisation can produce arrays of microspecies, characteristically seen in genera such as *Rubus* (330+ microspecies in Britain), *Hieracium* (400+) and *Taraxacum* (240+; Stace, 2019). Mating system shifts from outcrossing to selfing or apomixis, drive rapid reproductive isolation from progenitor lineages through founder effects, genetic drift, and selection of advantageous recessive alleles (Hollingsworth et al., 2006). This

leads to strong population structure, characterised by many and varied subtle morphological changes between populations. Similar phenomena result from ploidy level variation (Spaniel et al., 2011; Raggi et al., 2015), however hybridisation can assist in moving genetic material between populations, which can produce yet more morphological variation (Alix et al., 2017). Species concepts usually break down in TCGs, as the evolutionary processes creating new variation overwhelm any stable pattern of species, and the relationships between species become subtle, finely divided, and overlap.

1.2 The role of hybridisation in taxonomic complexity

Hybridisation, defined here as the mating between different species, is an important factor driving taxonomic complexity (Stebbins, 1959; Campbell and Wright, 1996; Ennos et al., 2005). The role of hybridisation is not simple however, as many different outcomes are possible. Hybridisation can be destructive, where rare species may lose their genetic integrity, resulting in populations of entirely hybrid genotypes, and can eventually lead to the extinction of the rarer species (Brochmann, 1984; Rhymey and Simberloff, 1996). On the other hand hybridisation can be creative, by allowing adaptive traits to move between species (Chapman and Abbott, 2010). Hybridisation can also lead to introgression, where there is the incorporation of genetic material from one species in the genetic background of another (Twyford and Ennos, 2012). In the extreme, new species can be formed in a process known as hybrid speciation (Mallet, 2007). There are two main pathways to speciation involving hybridisation: polyploid hybrid speciation where the hybrid species has duplicated its chromosome complement (allopolyploidy; Rieseberg and Willis, 2007), and homoploid hybrid speciation where the parental species and the hybrid remain at the same ploidy level (Rieseberg, 1997). Many TCGs involve hybridisation which blurs species boundaries, coupled with processes that may allow the hybrid derivatives to persist (Ennos et al., 2005). Three categories in which TCGs can be placed include agamic complexes, in which hybridisation is combined with a mode of asexual reproduction (e.g. apomixis) to propagate lineages (Hersh et al., 2016), polyploid complexes, where the hybrid derivatives are sexual polyploids (Zohary and Nur, 1959), and homogamic complexes (or homoploid complexes, e.g. *Helianthus* Rieseberg et al., 1995), where hybrid derivatives are mainly diploid and isolated from parental progenitors ecologically.

Some TCGs defy these three broad categories, by combining properties of different species complexes. A good example of this is the genus *Sorbus* in the British Isles, where there are 45 taxa, plus seven more which have been introduced (Nelson-Jones et al., 2002; Pellicer et al., 2012). In *Sorbus*, hybridisation, polyploidy, and apomixis have interacted to form numerous endemic species in England, Scotland, and Wales (Ludwig et al., 2013). In England alone, this process happens because there are few sexual diploid species (*S. torminalis*, *S.*

aria, and *S. aucuparia*, but not *S. domestica*) which hybridise with apomictic polyploid derivatives of the species *S. aria* (at least 20 taxa; Pellicer et al., 2012; Robertson et al., 2010). Apomictic *Sorbus* species require pollen to achieve successful asexual reproduction, however at low frequencies the pollen can fertilise the maternal embryo and this leads to a stable, new, polyploid, apomictic taxon (Ludwig et al., 2013). In all, more than 31 apomictic species have arisen in this way (Stace et al., 2015). The apomictic condition ensures that any new hybrid genotype is frozen in stasis, and can lead to complex reticulate evolutionary histories, which is seen in many other apomictic plant systems (Wittzell, 1999; Sochor et al., 2015). Not only is hybridisation important in generating biological diversity and complexity, it is also a common phenomenon both geographically, and phylogenetically (Ellstrand et al., 1996). The clear and widespread abundance of hybridisation means it has had, and continues to have, a profound effect on the evolution of plants – especially in conjunction with selfing, apomixis, and polyploidy.

1.3 The role of polyploidy in taxonomic complexity

Polyplody is the condition where a cell contains more than two sets of chromosomes as a result of whole genome duplication (WGD), and is featured in almost all TCGs in plants (Ennos et al., 2005; Brysting et al., 2007). The two major routes to polyploidy are either through WGD of a single species chromosome complement, known as autopolyploidy, or through hybridisation between two species followed by WGD, known as allopolyploidy (Figure 1; however, there are other mechanisms; see Ramsey and Schemske, 1998). As polyploid individuals tend to be larger, more vigorous, and quicker growing, it is no surprise that many crop species are polyploid (e.g. wheat, rice, potato, maize Renny-Byfield and Wendel, 2014). Although the majority of extant plant species are diploid (~67%, Rice et al., 2019), extensive variability in ploidy levels exist across flowering plants, at all taxonomic levels (Soltis et al., 2010; Kolar et al., 2017). Most if not all plant species have experienced historical WGD a number of times (Clark and Donoghue, 2018). Over evolutionary time the process of diploidisation reduces the size of these previously polyploid genomes through loss of repetitive sequence, chromosomal rearrangements, and reductions in chromosome number (Wendel, 2015). Both the spatial and phylogenetic distribution of ploidy variation are unlikely to be uniform due to climatic and clade specific effects on unreduced gamete formation, which is the main driver in the creation of polyploid organisms (Bretagnolle and Thompson, 1995; Kreiner et al., 2017a; Rice et al., 2019).

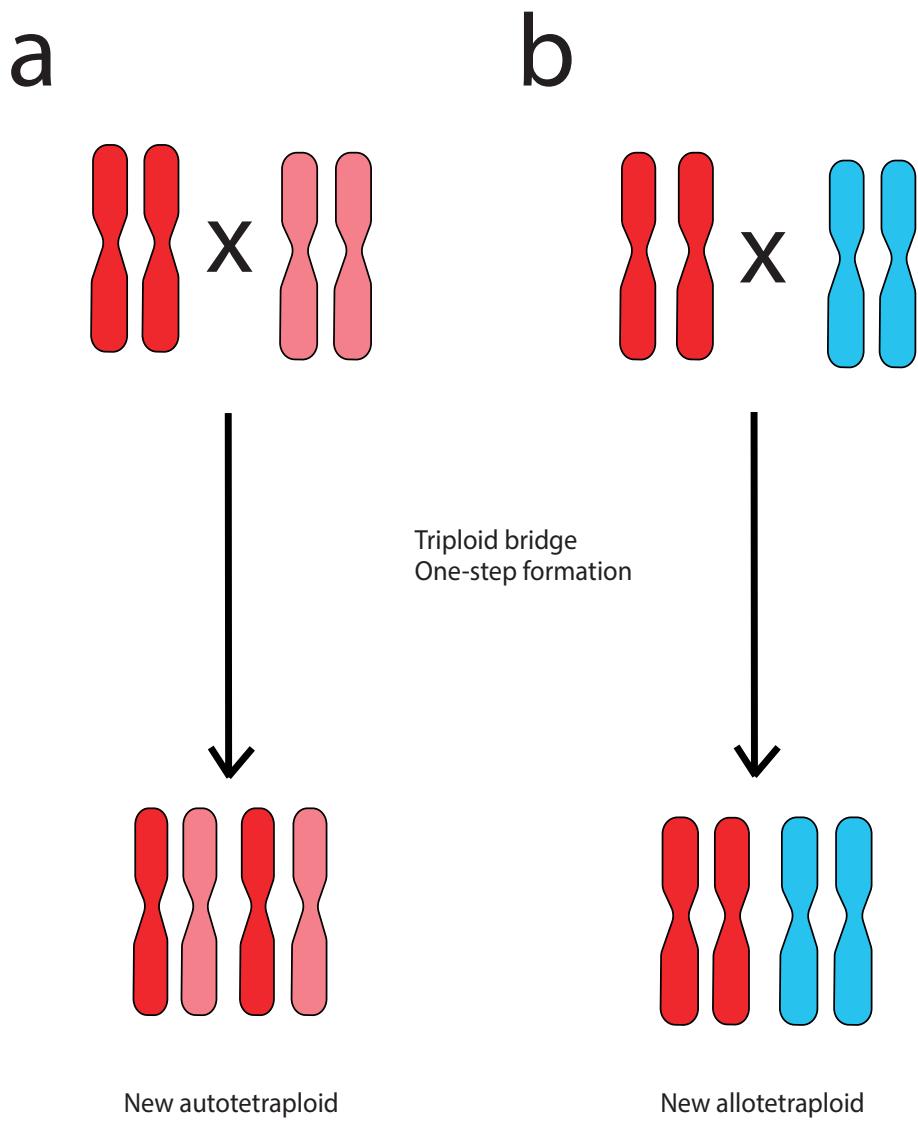


Figure 1: A simplified diagram of the formation of tetraploids from diploid progenitors. a) the formation of an autotetraploid from two diploid related individuals (red and pink) and b) the formation of an allotetraploid from two divergent diploid species (red and blue). Here, triploid intermediates can be formed through the combination of unreduced and reduced gametes. These triploids can then either self-fertilise, or backcross with parental species to produce tetraploids. Alternatively the union of unreduced gametes from both parents can produce a tetraploid in one step. Note: the distinction between auto and allopolyploids are blurred as the parental species may only be partially differentiated (Spoelhof et al., 2017), and there are other rarer mechanisms to produce auto and allopolyploids involving genome doubling.

Polyplody is a contributing factor to taxonomic complexity, with recurrent polyplloid formation followed by hybridisation within and between ploidy levels being the main driver of some TCGs (Brochmann et al., 2004). Further, there is evidence to suggest that polyplody facilitates hybridisation as ploidy level increases, i.e. tetraploid x tetraploid crosses are more successful than diploid x diploid crosses (Monnahan et al., 2019; Novikova et al., 2020). Shifts in mating system from outcrossing to selfing or apomixis are also correlated with polyplloid formation, due to the breakdown of self-incompatibility, selection against minority cytotype exclusion, and the capability of polyploids to alleviate inbreeding depression (Barringer, 2007). Recurrent polyploidisation has been shown to generate taxa of ever higher ploidy levels, some of which may persist to form new species, or backcross with parental species to form a complex reticulate group (Brochmann, Soltis, et al., 1992). Polyploid type (auto or allopolyploid) is important in determining their role in taxonomic complexity. Autopolyploids and allopolyploids are both frequent in nature; current estimates for both kinds of polyploids are around 10% (Barker et al., 2016). Autopolyploids form complexes within species where ecology can shape the distribution of cytotypes (e.g. Wilson et al., 2020), however these cytotypes are rarely considered species (Ramsey and Ramsey, 2014). Allopolyploids contribute to taxonomic complexity by combining two diverged genotypes to form a usually distinct and isolated taxon from the parents (Qiu et al., 2020). Both kinds of polyploid are present within some TCGs (e.g. *Sorbus* and *Cochlearia*; Ludwig et al., 2013; Gill et al., 1978).

Unusual cytogenetic features in some plant groups have generated TCGs that have defied classification for over a hundred years. Two examples are the dog roses in *Rosa* section Caninae and the evening primroses in *Oenothera* section Oenothera (Cleland, 1944; Lim et al., 2005). The dog roses present an unusual breeding system where the female parent contributes 3-5 copies of the genome and the male parent contributes only one (section Caninae species are usually pentaploid; Rowley, 1967). This results in the hybrids containing a genome which is mostly maternal, and has the added complication that reciprocal crosses entirely change the genomic constitution of the hybrid. Species are difficult to tell apart when this breeding system is combined with hybridisation between rose species, as hybrids are fertile and readily backcross, forming swarms of individuals that may be impossible to identify morphologically (Stace et al., 2015). Another example of strange cytogenetic behaviour, but at the diploid level, are the evening primroses. Here, the species exhibit a breeding system where translocation hybrids and balanced lethals produce chains of chromosomes which are inherited as single units (Cleland, 1944). While evening primroses breed true when selfing, different combinations of chromosome chains generated through hybridisation can produce completely new morphological taxa (Cleland, 1972). This introduces taxonomic complexity as, like the roses, when two species come into contact they form hybrid swarms of intermediate genotypes, each of which would breed true in isolation (Stace et al., 2015).

1.4 Novel features of TCGs; parasitism and plasticity

Approximately 1% of all angiosperm species are parasitic, with some genera being particularly speciose such as *Pedicularis* (c.a. 650 sp.), *Euphrasia* (c.a. 260 sp.), and *Thesium* (c.a. 300 sp. Nickrent, 2020; Moore et al., 2010; Twyford, 2018). Parasitic plants are defined by the formation of a structure called the haustorium (Figure 2), which is used to extract water and soluble nutrients from the host plant (Twyford, 2018). The haustorium can attach either to roots (e.g. *Orobanche*; Musselman, 1980) or shoots (e.g. *Viscum*; Becker, 1986), or rarely the haustorium is present inside the host plant itself (e.g. endoparasitic *Pilostyles*; Fernandes et al., 1998). There are two types of parasitic plants – hemiparasites which retain photosynthetic competency, and holoparasites which are devoid of chlorophyll and entirely dependent on host plants (Joel et al., 2013). Host species range and identity varies widely between parasitic plants, and some generalists can parasitise more than 100 host plant species (e.g. *Amyema miquelii*; Clark et al., 2020), while others specialise on a single or few host plants (e.g. *Epifagus virginiana* on *Fagus grandiflora*; Tsai and Manos, 2010).



Figure 2: A variety of parasitic plants, showing haustorial connections. a) *Cuscuta europaea* (dodder), showing flowers and twining red stems. b) *Erianthemum ngamicum* showing its large terminal haustorium. c) *Orobanche hederae* haustorial connection to the host root (light) d) *Hydnora visseri* forming multiple haustoria on host root (light). e) *Cassytha pubescens* twining around host forming many haustorial connections. f) *Viscum album* showing self-parasitism where two younger *V. album* individuals have established on an internode of an older specimen. g) Terminal (asterisk) and lateral (arrows) haustoria of *Plicosepalus kalachariensis*. h) flowers of *Agelanthus gracilis*. Figure and text adapted from Joel et al. (2013).

How different host species may influence parasitic plant phenotypes however, has been little explored in relation to taxonomic complexity. Yet it is well known that different host species can dramatically impact the

growth, development, and evolution of parasitic plant individuals (Rowntree et al., 2014; Matthies, 2017). There are two main ways in which the parasitic lifestyle may contribute to taxonomic complexity. Firstly, parasitic plant species may show phenotypic plasticity when utilising different host species, which may be substantial enough to confuse species identification (Wilkins, 1963). For example, a suite of traits used in species discrimination in the Orobanchaceae, including corolla length, node to first flower, and plant height, all vary in relation to host quality (Jonstrup et al., 2016; Matthies, 2017). Secondly, differential host use can drive the evolution of cryptic taxa, which has been seen in the plant genus *Orobanche*, which specialise on different host species (Thorogood, Rumsey and Hiscock, 2009).

1.5 The genus *Euphrasia*

The genus *Euphrasia* (eyebrights) are a large group of 260, mainly annual but sometimes perennial, hemiparasitic plant species in the Orobanchaceae (Yeo, 1978; Nickrent, 2020). A study of global *Euphrasia* species has established its bipolar distribution, and estimated the age of the genus to be around 20-30 Mya (Gussarova et al., 2008). *Euphrasia* species are found in Chile, Australia and New Zealand, parts of South East Asia, and widely across the northern hemisphere. It is here, in the northern hemisphere that *Euphrasia* species are most diverse and where the genus is a notorious TCG (Yeo, 1978). Within Europe, Britain and Ireland contain the highest concentration of species, where 21 have been described (Metherell and Rumsey, 2018). The number of described species in Britain may be an exaggeration however due to over-splitting. Here, *Euphrasia* is widespread and occupies a range of habitats, from coastal scrub to heather moorland, and damp grassland to mountain tops (Metherell and Rumsey, 2018). While the split between diploid species and tetraploid species is unequivocal, within each ploidy level there is much uncertainty in species limits due to rapid recent divergence (Wang et al., 2018). Within tetraploids, some species groups are more distinct than others, for example *E. micrantha* and *E. scottica* form a natural group, as do the widespread outcrossing species *E. arctica* and *E. nemorosa* (French et al., 2008). Omitting these two species groups, it leaves a nebulous pool of widespread and more localised selfing species that may be genetically partitioned more by geography than by species limits. Within the diploid species, there is evidence to suggest that the endemic diploid species *E. vigursii* and *E. rivularis* are distinct from the other diploids (*E. anglica* and *E. rostkoviana*; French et al., 2008).

The taxonomic complexity of *Euphrasia* in Britain is reflected in the unstable taxonomy over the past hundred years, in the great diversity of morphologies found in the field, and in the complex genetic structure of the species (Wettstein, 1896; Metherell and Rumsey, 2018). The main drivers of this diversity are recent postglacial divergence, the tendency to self-fertilise, and rampant hybridisation – 71 hybrid combinations

have been reported to date (Figure 3; Wang et al., 2018; Stace et al., 2015; Metherell and Rumsey, 2018). At least five species of the British species of *Euphrasia* have a putative hybrid origin, with two species having arisen from diploid-tetraploid hybridisation events (Yeo, 1956; Silverside, 1990). The evidence of cross-ploidy hybridisation began with early cytological work on a triploid hybrid which was intermediate between *E. micrantha* and *E. anglica* (= *E. vigursii*) in morphology (Yeo, 1956). Further evidence was found by French et al. (2008) when tetraploid specific AFLP bands were seen in diploid samples. Lastly, Becher et al. (2020) in a genomic analysis of British *Euphrasia*, found some evidence for gene flow between *E. arctica* and the sampled diploid species. These cross-ploidy hybrid species are of particular interest, as they combine three important features of TCGs – hybridisation, polyploidy, and mating system variability.

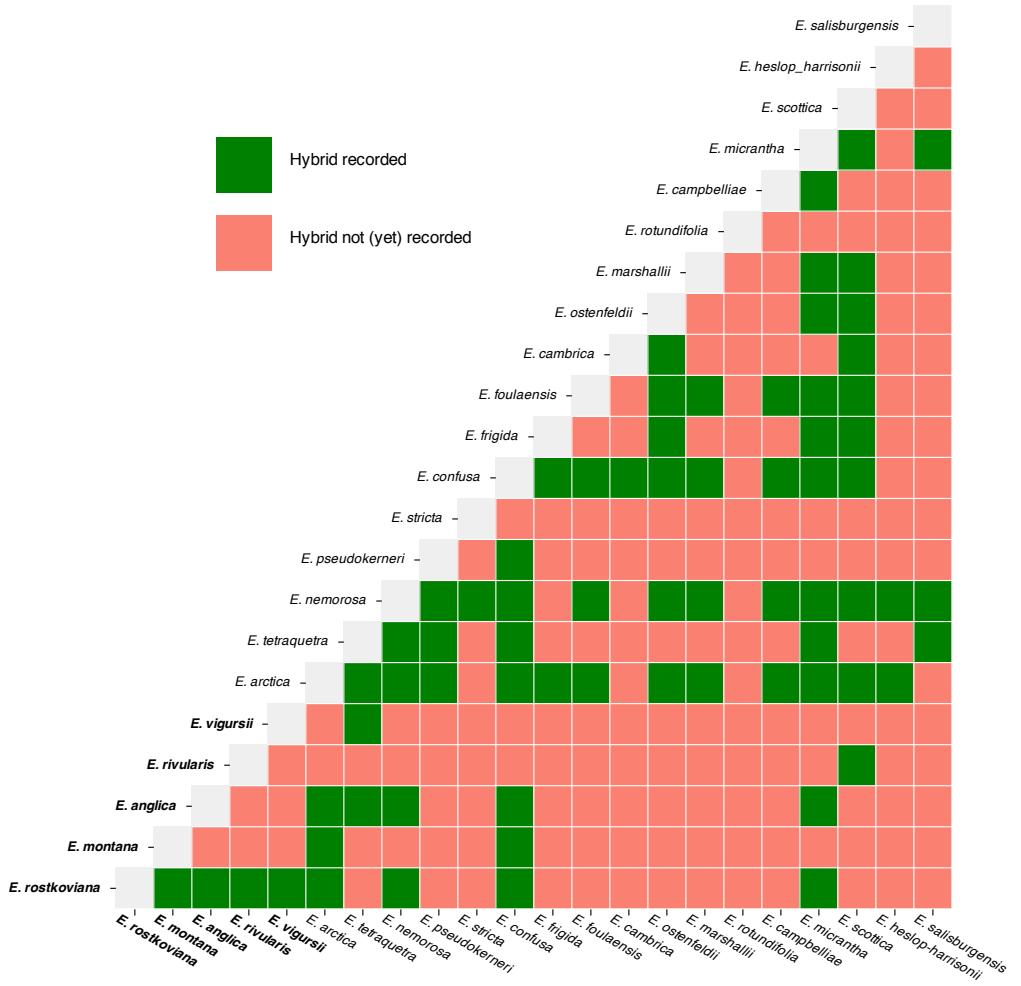


Figure 3: A matrix showing interspecific crossing barriers in British *Euphrasia* species. Grey squares indicate intraspecific crosses, pink squares show crosses not known to produce hybrids in the wild, and green squares show those crosses to have produced hybrids in the wild. Diploid species are emphasised in bold type. Data taken from Metherell and Rumsey (2018).

Euphrasia in Britain and Ireland have two different ploidy levels, with diploids ($2n = 2x = 22$) that have a southern distribution and tetraploids ($2n = 4x = 44$) more predominant in the north (Yeo, 1978). The tetraploid species are allotetraploids, containing one subgenome which is closely related to extant diploids (0.2% divergent; Becher et al., 2020). This low divergence has led to the hypothesis that pairing can occur between the chromosomes of diploid species and the diploid-like subgenomes of tetraploids, which could explain diploid-tetraploid hybridisation in the genus (see Figure 4; Yeo, 1956). The divergence between diploids and tetraploids is around 5% based on both ITS sequencing, and genome wide data, which corresponds to a

split time of around 8 Mya (Wang et al., 2018; Becher et al., 2020). High divergence between diploids and tetraploids points to ploidy being an effective barrier to gene exchange, however there is mounting evidence that gene flow between ploidy levels is present, but rare (French et al., 2008; Becher et al., 2020). Hand crosses have generally failed to produce diploid-tetraploid hybrids, and only a single triploid hybrid has been found in the field (Yeo, 1954, 1956). Controlled crosses may yet yield useful insights into the biology of diploid-tetraploid hybrids in *Euphrasia*.

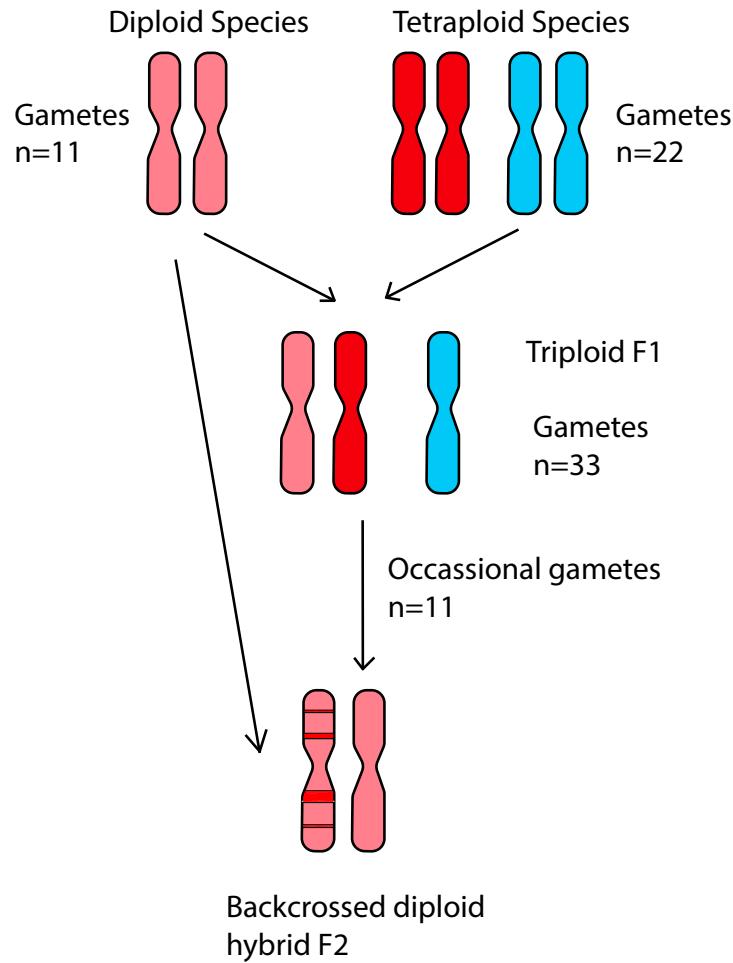


Figure 4: Schematic of diploid-tetraploid hybridisation in *Euphrasia*. Here, a diploid species hybridises with a tetraploid species to form a triploid intermediate. Note that pink and red colours indicate the low (0.2%; Becher et al., 2020) divergence between diploid (pink) and diploid-like (red) tetraploid subgenomes. This triploid F1 can then backcross to the diploid parent through rare haploid gamete segregation to produce a backcrossed F2 individual. Note the striped chromosomes in the F2 backcross indicate recombination between the diploid (pink) and the diploid-like (red) subgenome of the tetraploid. It is thought this process has given rise to the hybrid species *Euphrasia vigursii* and *E. rivularis* (Yeo, 1956; Silverside, 1990).

The mating system in British *Euphrasia* is highly variable, and ranges from outcrossing to highly selfing (note apomixis has not been found in the genus). Outcrossing rate is correlated with flower size in *Euphrasia* ($r = -0.89$; French et al., 2005), with smaller flowered species more likely to self than larger flowered species with showy corollas. For example, the tetraploid mountain specialist *E. cambrica* which is endemic to Wales has a corolla length of 4mm, while the diploid *E. montana* of wet grasslands has a corolla length of up to 12mm

(and can be larger in cultivation; Figure 5; Metherell and Rumsey, 2018). Mating system may also impact the directionality of introgression in some cases, for example between tetraploid *E. micrantha* and diploid *E. anglica* which are the putative parental species for the diploid hybrid species *E. vigursii*. *E. anglica* has large flowers and is mainly outcrossing, therefore it is likely to be the pollen donor as its flowers are visited first with greater probability (Yeo, 1968). Therefore the diploid species pollen will be more competitive on the stigma of the tetraploid *E. micrantha* (Ruhsam et al., 2011). As pollen fitness is low in the triploid F1 hybrid, selfing is unlikely to occur and the diploid *E. anglica* is likely to fertilise the F1 (Ruhsam et al., 2013). If this happens over many generations, introgression will occur from the tetraploid to the diploid species (as shown in Figure 5).



Figure 5: corolla size extremes in British *Euphrasia* species. a) shows diploid *E. montana* from the Lake District, England, which has a corolla size of ~12mm. b) shows the Welsh endemic *E. cambrica* from Snowdon, with small flowers ~4mm which rarely open fully (Metherell and Rumsey, 2018). Photo credit: author.

Being generalist parasites, *Euphrasia* can gain benefit from a wide variety of plant species which in turn impact the morphology, growth, and fitness of the parasitic *Euphrasia* plants (Svensson and Carlsson, 2004). Common garden experiments have allowed researchers to investigate these factors in different hemiparasitic plant systems, especially in the Orobanchaceae. For example, it has been shown in *Rhinanthus* that different host species employ different resistance mechanisms to haustorial attack, and this underlies the performance

of hemiparasitic *Rhinanthus* (Cameron et al., 2006). Common garden experiments have also established that host species have strong effects on the biomass and morphology of *Melampyrum* (Matthies, 2017). As shown from early work by Yeo (1964), Euphrasia can easily be brought into cultivation, and their annual life histories (like many other members of the hemiparasitic Orobanchaceae) facilitate experimental work within the timeframe of a PhD.

1.6 The British flora

To create a synthesis of taxonomic complexity, it is useful to be able to frame it in a broad, comparative context. The British flora is an ideal study system, as it contains a manageable number of native plant species (~1,400), but with around 20% of all familial flowering plant diversity (Stace, 2019). Alien species increase the total number of species over two-fold, and most of these are very well characterised (Stace and Crawley, 2015). The British flora represents the most comprehensively studied flora to date, across a variety of disciplines due to the collaboration between amateur and professional botanists for over a century (Allen, 1986). The Botanical Society of Britain and Ireland (BSBI) have played a leading role, holding large databases of plant distributions. The result of this, is that plants can be found and told apart easily in the field, and has led to a huge growth in knowledge of the system. Now for almost all native plants, there is detailed information on plant identification (Stace, 2019), alien species (Stace and Crawley, 2015), ecology and life history (Fitter and Peat, 1994), chromosome numbers and ploidy level (BSBI Cytology Database), genome sizes (Kew C-value Database), hybridisation (Stace et al., 2015), and most recently DNA barcoding (Jones et al., In Review). This wealth of knowledge has been leveraged in parts of this thesis to gain a broader perspective on the topic at hand.

1.7 New methods for investigating taxonomic complexity

In recent years, new methods have emerged that can help us to understand taxonomic complexity, and to resolve major questions. Older marker types such as allozymes, AFLPs, RAPD, and microsatellites had many limitations, including needing large amounts of DNA, problems with homology, detection of few polymorphisms, and lack of reproducibility (Lowe et al., 2004). Next generation sequencing (NGS) is one tool to aid resolution of complex taxonomic relationships, for example - genotyping by sequencing (GBS) (Anderson et al., 2017), restriction associated digestion (RAD) sequencing (Zhou et al., 2020), target capture (Carter et al., 2019), and whole genome sequencing (Dupuis et al., 2017) are becoming common. These datasets can be used to infer relationships across the entire genome, to correlate morphology and genetics to

reveal the genes underlying certain traits, and to compare large and diverse sample sets. DNA barcoding can now be deployed at scale across diverse taxa (Hollingsworth et al., 2009; Kuzmina et al., 2017), which means large scale phylogenetic analyses that incorporate diverse information are now possible. New analytical models now provide a framework for accounting for different sources of variation, and are applicable to use for many different questions across evolutionary biology (e.g. (Generalised) Linear Mixed Models; Bolker et al., 2009; Hadfield, 2010). These methods are used throughout this thesis to gain novel insights into taxonomic complexity.

1.8 Thesis aims

The main aim of my thesis is to investigate taxonomic complexity in *Euphrasia*, and the role of hybridisation across the British flora, with a particular focus polyplody, parasitism, and their interactions. It is not currently known what the extent and prevalence of such interactions between hybridisation and plants of different ploidy level are, and is important to explore because of the potentially large impact they can have on plant evolution. In addition, the parasitic habit of some plants and how this affects their phenotype has been little explored and opens up a new avenue of explanations for taxonomic complexity.

The thesis is split into two parts which target different aspects of taxonomic complexity, and draw on broader themes in evolutionary biology. The first part concentrates on the contributions of hybridisation to taxonomic complexity in plants (and to a lesser degree, animals), with a case study in the genus *Euphrasia*. Here, I firstly ask how prevalent hybridisation is between plants that differ in ploidy level, and whether it is a significant evolutionary phenomenon (Chapter 2). The interaction between ploidy level and hybridisation is a poorly explored topic and scattered across the literature. I bring together a comprehensive list of examples, and synthesise current knowledge on this topic. After understanding this global variation, I concentrate on the British flora, and model the probability of hybridisation across the flora (Chapter 3). Significantly, this model includes phylogenetic relationships, and genetic distances based on the first complete DNA barcode dataset across a flora, which previous studies have not been able to comprehensively address (e.g. Ellstrand et al., 1996; Mitchell et al., 2019). Then, I focus on the promiscuous genus *Euphrasia* and ask whether we see evidence of hybridisation between divergent species of different ploidy at a fine spatial scale. Using reduced representation sequencing of genome wide markers (genotyping by sequencing; GBS), I use a combination of classical population genetic tools, genome sequencing, and demographic simulation, to understand the pattern of hybridisation in a *Euphrasia* contact zone (Chapter 4).

In the second part of the thesis, two novel features of taxonomic complexity - phenotypic plasticity and

parasitism - are explored in the parasitic plant genus *Euphrasia*. Developing *Euphrasia* as a horticultural system to explore these ideas is discussed first (Chapter 5). I then use common garden experiments to ask how different host species affect the morphology of *Euphrasia* and the ability to discriminate between *Euphrasia* species (Chapter 6). This was investigated using a single species of *Euphrasia* grown across eight different host species and multiple species of *Euphrasia* on a single host, where various morphological traits of *Euphrasia* were quantified. Then I look at the role of host parasite interactions to understand host specialisation in the genus, by growing multiple *Euphrasia* and multiple host species together (Chapter 7). Finally, I synthesise the findings of this thesis in the last chapter, and highlight areas of future research.

2 The emerging importance of cross-ploidy hybridisation and introgression

2.1 Abstract

Ploidy differences are often considered a strong barrier to hybridisation and introgression because viable hybrid zygotes are rarely formed, and if they are formed, they are usually produce sterile organisms. The increasing accumulation of cytogenetic, genetic, and genomic datasets however, show that ploidy differences may be a more permeable barrier than previously thought. Cross-ploidy hybridisation is most common in plants, with a third of all hybrids in the British flora forming from parental species of differing ploidy level, and dozens of studies have confirmed the existence of cross-ploidy hybridisation across plants with molecular markers. Cross-ploidy hybridisation is also present in certain animal lineages that are characterised by relatively high levels of polyploidy, and where hybrid derivatives are hybridogenic and do not recombine. Groups with cross-ploidy hybrids often involve allopolyploidy, where homologous chromosomes may pair regularly with a related diploid, and show directional introgression towards the higher ploidy parent. Cross-ploidy hybridisation may lead to introgression of adaptive loci, and be an important driver of diversification and speciation in plants and animals.

2.2 Introduction

Climate change, habitat disturbance and large-scale translocations of species resulting from human activities are increasing contacts between species previously isolated by geographical and ecological barriers, thus raising their potential to hybridise (Crispo et al., 2011; Brennan et al., 2014; Larson et al., 2019). Closely related species isolated by prezygotic barriers are more likely to hybridise (Vallejo-Marin and Hiscock, 2016); however, even species isolated by very strong postzygotic barriers do hybridise in some instances. Polyploidy (see Glossary; Box 1), which is particularly common in plants, creates a very strong postzygotic barrier between species that differ in ploidy (Box 2). For example, many crosses between diploid and tetraploid species either fail to produce a viable zygote or produce poorly formed ones, depending on the direction of the cross. This phenomenon is known as ‘triploid block’ (Ramsey and Schemske, 1998). Should a triploid hybrid form, it is normally either completely or partially sterile, due to formation of malfunctioning gametes containing unbalanced chromosome numbers. On occasion, however, some species differing in ploidy do produce hybrid offspring, triggering gene exchange or possibly the origin of new species via allopolyploidy (Box 3). The

importance of such events is not to be underestimated; for example, they have led to the origin of some very recently originated plant species, which are now models for the study of polyploid speciation (Vallejo-Marin and Hiscock, 2016), and also to the origin of some of our most important crop plants, including wheat, sweet potato and sugar cane. Nonetheless, the frequency of cross-ploidy (or interploidy) hybridisation in the wild is a neglected topic, with information related to it scattered through the literature. Here, we bring this information together and emphasise its biological significance.

Box 1: Polyploidy

Polyploidy is the condition where a cell contains more than two sets of chromosomes as a result of whole genome duplication (WGD). The two major routes to polyploidy are either through WGD of a single species chromosome complement, known as autopolyploidy or through hybridisation between two species followed by WGD, known as allopolyploidy (Ramsey and Schemske, 1998). It is driven especially by the production of unreduced gametes in diploid species (Moghe and Shiu, 2014) and this is affected by a range of factors including specific genes (Ravi et al., 2008) and environmental stresses (Rice et al., 2019).

Although worldwide, the majority of plant species are diploid (~67%, Rice et al., 2019), extensive variability in ploidy levels exist at all taxonomic levels and scales (Soltis et al., 2010; Kolar et al., 2017). Both the spatial and phylogenetic distribution of ploidy variation are unlikely to be uniform however, due to climatic and clade specific effects on unreduced gamete formation (Bretagnolle and Thompson, 1995; Kreiner et al., 2017a; Rice et al., 2019). Historically, autopolyploidy has been regarded as both less frequent and less important in an evolutionary context, than allopolyploidy (Soltis et al., 2010). The current wealth of cytological data suggests, however, that at least 10% of species are autopolyploids, with allopolyploids estimated to be at least as frequent (Soltis et al., 2010; Barker et al., 2016; Kolar et al., 2017). Allopolyploids have received more attention as they are mainly distinctive morphological taxa described as species, while autopolyploids are often morphologically cryptic and lumped into species complexes (Ramsey and Ramsey, 2014; Barker et al., 2016). Polyploidy has been important over evolutionary time in the genesis of new plant and animal lineages (Otto and Whitton, 2000), and its signature is imprinted several times over in the genome of every flowering plant (Wendel et al., 2016). Both polyploid speciation conferring immediate reproductive isolation (Whitton, 2004) and polytopic origins of polyploids can lead to new lineage formation (Thompson and Merg, 2008). There is also increasing evidence that polyploidy facilitates lineage diversification, though this remains a controversial topic (Wood et al., 2009; Ren et al., 2018; Han et al., 2020). Polyploidy is also associated with major shifts in ecology and morphology across a wide variety of plant species (Husband and Sabara, 2004; Parisod et al., 2010; Paule et al., 2017).

The first known artificial hybrid from crossing two parents of differing ploidy level was created by Kölreuter in 1761 between diploid *Nicotiana paniculata* and allotetraploid *N. rustica*. This hybrid was known as the first “botanical mule” due to its shrivelled anthers and malformed ovaries, indicative of high sterility (Roberts, 1929). Further artificial crosses demonstrated the formation of other interploidy hybrids that were partially or completely sterile, but nothing was discovered of the frequency or importance of the phenomenon in the wild until much later (Lawrence, 1936). Beginning around the mid C20th, cytogenetic studies became more frequent and revealed extensive ploidy variation both within and between species that varied with geography, and which could be used to explain evolutionary relationships (Love and Love, 1943; Stebbins, 1956). However, it was with the availability of multiple nuclear markers in the 1990s that researchers reliably detected hybridisation and introgression between species of differing ploidy (Nason et al., 1992; Abbott et al., 1992). Now, by examining many thousands of genetic markers across the genomes of target species, there is potential to detect cases of adaptive introgression (Suarez-Gonzalez et al., 2018). Moreover, by focusing on specific genes, examples are now known of cross-ploidy introgression resulting in the transfer of particular traits that markedly affect the biology and fitness of recipient species (Kim et al., 2008; Chapman and Abbott, 2010; Baduel et al., 2018; Monnahan et al., 2019).

While there have been many recent reviews on the mechanisms that underlie polyploidy and the prevalence of polyploids in nature (e.g. Alix et al., 2017; Soltis et al., 2004; Chen, 2010; Kohler et al., 2010; Marques et al., 2018), and on the importance of natural hybridisation (e.g. Abbott et al., 2013; Soltis and Soltis, 2009; Todesco et al., 2016; Suarez-Gonzalez et al., 2018), our aim is to reconcile early work on cytological variation with recent work on genomics, to consider whether cross-ploidy hybridisation between species may be more prevalent and important than previously known. We first review the presence of cross-ploidy hybridisation in the British and Irish flora, the most well-studied, large-scale flora examined to date. Next, we review the prevalence of cross-ploidy hybridisation inferred with genetic markers that has been reported in the literature, and highlight some general patterns. Lastly, we explore the biology of cross-ploidy hybrids, and discuss how advances in sequencing technology may aid hybrid detection to assess more accurately the state of interploidy hybridisation in nature. We emphasise case studies in flowering plants, where hybridisation and polyploidy are particularly prevalent and well-documented, but also consider other organismal groups where cross-ploidy hybridisation may occur.

2.3 Occurrence of natural cross-ploidy hybrids

Of major interest is how common cross-ploidy hybrids are in nature given the varied constraints of both pre and postzygotic isolation in their generation (Box 2). The evidence required to prove interploidy hybridisation is confirmation of parental ploidy differences, which may come from chromosome counts (Rice et al., 2015), genome size estimates (Plant C-value Database) or genomic information (Ranallo-Benavidez et al., 2020), and evidence of hybridisation, which may be from genetic data or from other sources such as morphology (Rieseberg and Ellstrand, 1993); though see issues with using morphological data to detect hybrids below). Data on both ploidy and hybridisation are patchy, and this limits our current understanding of the frequency of cross-ploidy hybrids in nature. To illustrate the extent of cross-ploidy hybridisation, we consider the case of the British and Irish flora, which contains a manageable number of native species (~1500, excluding large taxonomically complex groups; Stace, 2019), and is exceptional in having near complete information on species chromosome counts (BSBI Cytology Database), and the extent of natural hybridity (Stace et al., 2015). Most of the 1295 species for which there is detailed ploidy information are diploids (56%), with higher ploidy levels becoming exponentially less common (Figure 6). Between families, however, the distribution of ploidy levels changes significantly, which alters the raw material for cross-ploidy hybridisation to act on (Appendix 1 Figure 1). In terms of hybridisation, there are 909 known hybrids present in the flora (Stace et al., 2015). Of the 588 hybrids that contain ploidy information (321 hybrids lack appropriate data), 203 interploidy hybrids have formed in Britain and Ireland (35%; Appendix 1 Table 1), in comparison to 385 intraploidy hybrids (65%). Cross-ploidy hybrids occur in 67 genera, with over a quarter present in the genera *Rumex* (Polygonaceae, 24), *Salix* (Salicaceae, 19) and *Euphrasia* (Orobanchaceae, 13; Figure 8). The majority (55%) of cross-ploidy hybrids involve diploid-tetraploid crosses, with higher order ploidy crosses closely following (43%), and diploid-triploid crosses in the minority (2%). Same ploidy parental species are 22% more likely to form hybrids than parental species of different ploidies (Chapter 3), indicating that ploidy level represents a considerable barrier to hybrid formation, but is far from complete. In addition for flowering plants surveyed above, cross-ploidy hybridisation is likely to be prevalent in other plant groups, such as ferns and fern allies, due to highly variable ploidies and abundant hybridisation. One dramatic example, inferred based on morphology and habitat, is from the lycopod genus *Isoetes*, where the diploid *Isoetes echinospora* ($2n = 22$) hybridises with the decaploid *I. lacustris* ($2n = 110$) to produce a hexaploid hybrid ($2n = 66$).

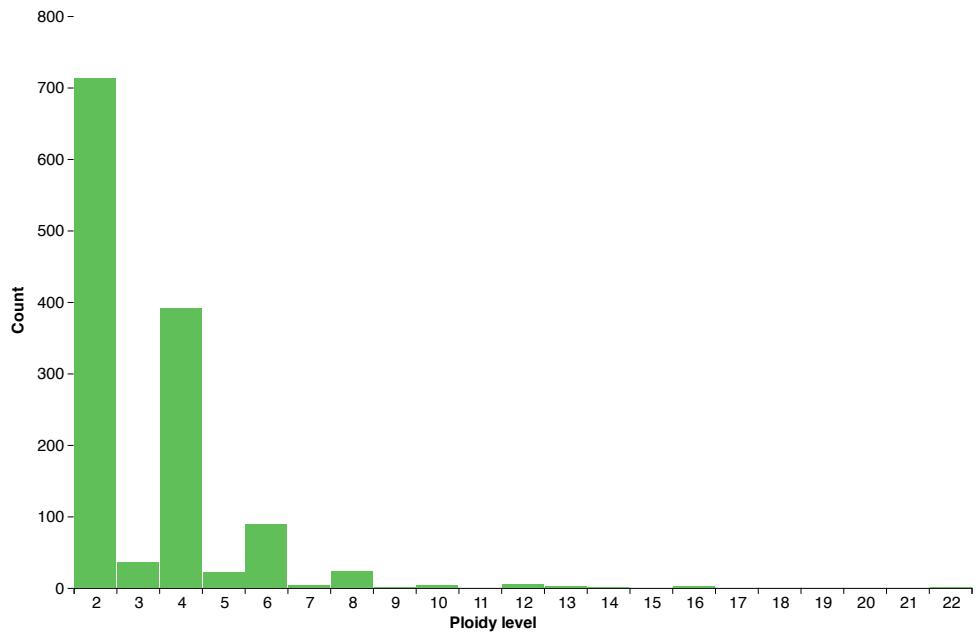


Figure 6: Distribution of ploidy levels across the British and Irish flora. Shown are the number of species at each ploidy level which are not known to have multiple cytotypes. Odd ploidies are less frequent than even ploidies, resulting in a ‘saw tooth’ pattern. The most highly polyploid species is *Leucanthemum maximum* at 22-ploid.

Box 2: Ploidy differences as a reproductive barrier

Ploidy differences have often been cited as strong reproductive barriers to hybridisation in plants (Husband and Sabara, 2004; Sutherland and Galloway, 2017). Cross-ploidy hybridisation is therefore usually considered rare because hybrids will have unbalanced chromosome content and therefore irregular pairing of chromosomes, rendering the hybrid infertile. This infertility prevents or limits the formation of backcross hybrids and the potential for introgression. In cross-ploidy hybridisation the usual reproductive barriers to cross species mating apply, along with specific factors associated with ploidy level difference between parental species. In addition to reproductive barriers caused by differences in geography, phenology, morphology and mating system etc. (Kay, 2006; Martin and Willis, 2007; Laport et al., 2016), the ploidy ratio of the pollen:style is important (Watkins, 1932; Stace, 1975), and following fertilisation is a period where endosperm development and (epi)genetic compatibilities are critical (Bomblies and Weigel, 2007; Lafon-Placette and Kohler, 2016).

Box 2: Ploidy differences as a reproductive barrier

There are two main pathways to creation of cross-ploidy hybrids; either through reduced or unreduced gametes. Reduced (“normal”) gametes of the both parental species results in the generation of a hybrid with intermediate ploidy. These hybrids, usually triploids derived from diploid-tetraploid crosses, are common and found in a variety of taxa where congeners co-occur, for example *Aconitum*, *Ficaria*, *Dactylorhiza* and *Senecio* (Irwin and Abbott, 1992; De Hert et al., 2012; Sutkowska et al., 2017; Popelka et al., 2019). A barrier to the creation of these hybrids through reduced gametes is known under the umbrella term ‘triploid block’ (Ramsey and Schemske, 2002; Kolar et al., 2017). Early work on experimental diploid-autopolyploid crosses established the presence of a triploid block and that direction of crosses was important (Thompson, 1930; Valentine and Woodell, 1960; Stebbins, 1971). The major cause of triploid block is attributed to genomic conflict in the maternal endosperm, which is usually triploid and composed of a ratio of two maternal and one paternal genomes (Lafon-Placette and Kohler, 2016). Deviations from this ratio cause the endosperm to malfunction in development and function (Kohler et al., 2010). Reciprocal crosses differ in their likelihood of success, and it is a general phenomenon that crosses where the higher ploidy parent is female are more likely to produce viable offspring, due to endosperm ratios which are better tolerated (Burton and Husband (2000); Figure 7 panels a and b). Triploid block may also be caused by the action of allelic incompatibilities at an early stage in development, although this topic is little explored (Scott and Bolbol, 2013). A second possibility in the creation of cross-ploidy hybrids is where the lower ploidy parent produces unreduced (“polyploid”) gametes. Unreduced gamete production is on average 0.1-2%, with rare individuals and hybrids that produce considerably higher frequencies (>85%; Kreiner et al., 2017a, 2017b; Mason and Pires, 2015). In addition, many different taxa produce unreduced gametes, and their production also varies with environmental variables (Baduel et al., 2018; Rice et al., 2019). Successful crosses occur more readily when unreduced gametes are produced by the diploid parent, thus restoring the gamete ploidy to that of the higher ploidy parent (Figure 7 panel c) Ramsey and Schemske, 2002).

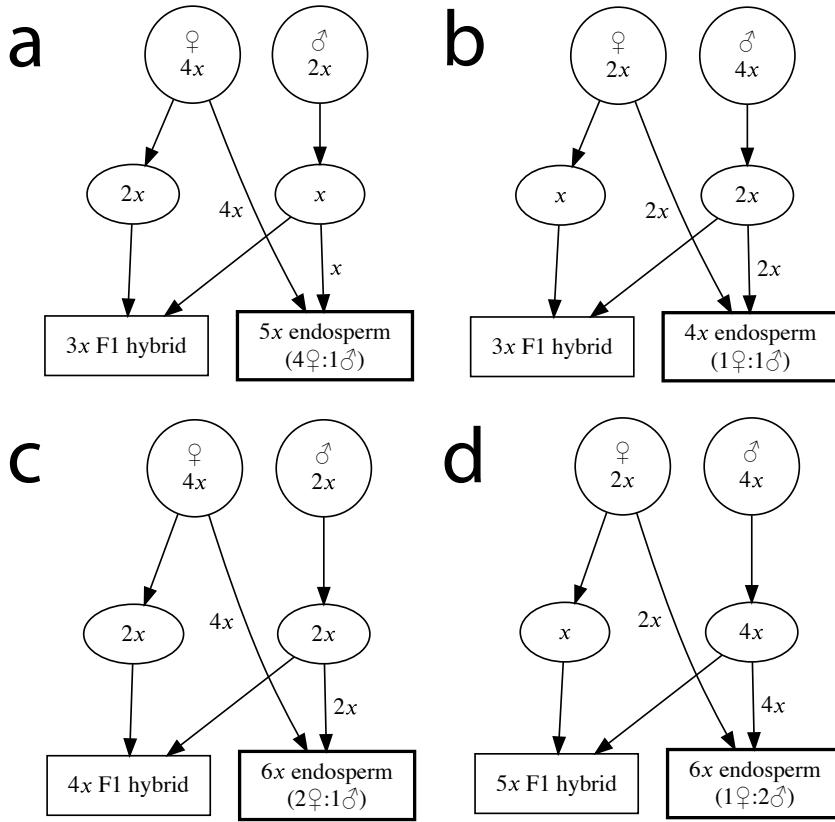


Figure 7: Potential outcomes of hybridisation between diploid and tetraploid species. In each panel, the top two circles refer to the parental species, the middle two ellipses to the gametes produced from each parent, the bottom left box to the F1 hybrid and the bottom right box to the endosperm. Panels a and b consider hybridisation with reduced gametes and therefore generate triploid hybrids, while panels c and d consider hybridisation where one parent produces unreduced gametes. In particular, panel c illustrates that a fertile polyploid can be generated in a single generation. Figure generated with graphviz (Ellson et al., 2002).

Inferring hybridisation from morphology, geography, cytology and limited genetic data, as is the case with many hybrids in the British and Irish flora, will overlook cryptic hybridisation and introgression that can be detected with multiple genetic markers. Moreover, the extent of cross-ploidy hybridisation in this flora is likely to be affected by extensive habitat disturbance and the prevalence of alien taxa. A wider survey of published studies of hybridisation based on multiple genetic markers or strong cytogenetic evidence revealed 43 different parental species combinations from 48 studies resulting in cross-ploidy hybridisation, with such hybrids present in 33 genera from 16 angiosperm families, three fern families, and three animal families (Table 3). Diploid-tetraploid crosses are found in 32 of the 43 parental crosses, with the rest being higher ploidy crosses. This confirms that interploidy hybridisation is likely to be much more widespread than is

currently appreciated. The taxonomic spread of interploidy hybridisation is especially broad in angiosperms, as evidenced by data both from the British and Irish flora and the wider literature. For example, monocots are well represented (Liliaceae, Orchidaceae, Poaceae), as are basal eudicots (Ranunculaceae, Papaveraceae) and throughout the rest of the phylogenetic tree scattered in the Fabids, Malvids and Superastrids. This distribution indicates interploidy hybridisation is very widespread and potentially abundant throughout the flowering plant phylogeny (Figure 8). On the other hand, the conspicuous absence of records from large, diverse families with variable ploidy, such as Rubiaceae potentially indicate a phylogenetic skew in interploidy hybridisation. Cases of such hybridisation are not just phylogenetically but also geographically widespread, with examples reported from across four continents, though tropical regions are poorly represented and most studies report hybridisation in large temperate or cosmopolitan plant families (e.g. Asteraceae and Orchidaceae). In terms of life form, most well-documented cross-ploidy hybrids (with the notable exception of *Euphrasia*) are perennial, a factor which correlates strongly with hybridisation regardless of parental ploidy level (Mitchell et al., 2019).

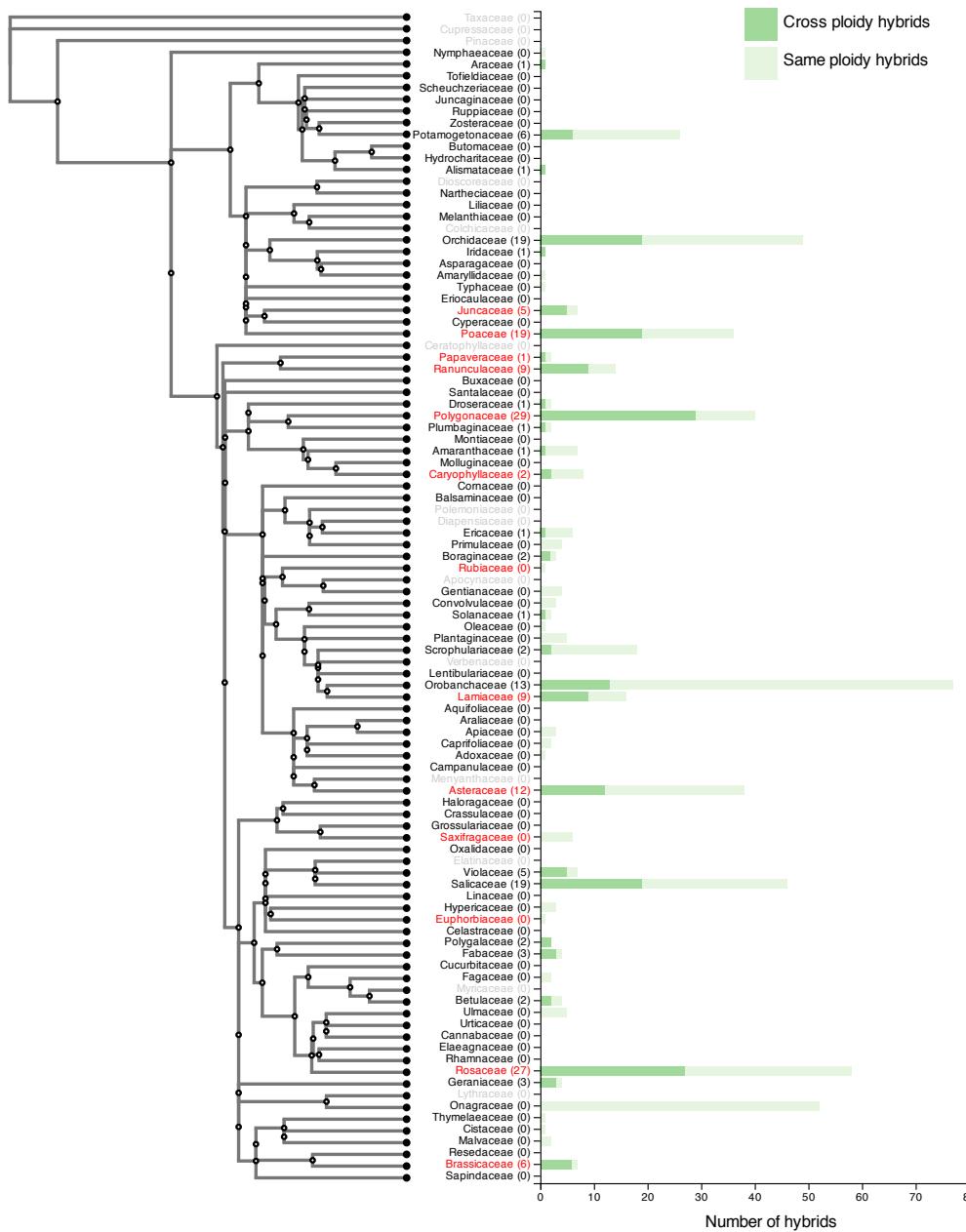


Figure 8: Distribution of cross-ploidy hybrids across the British and Irish flora. The number of cross-ploidy (dark bar) and intra-ploidy (light bar) hybrids are shown per family, in the context of family-level phylogenetic relationships from DNA Barcode UK (Jones et al., In Review). Faded family names indicate missing ploidy data, and red family names highlight those families which contain five or more different ploidy levels. Numbers in parentheses are the number of cross-ploidy hybrids formed per family. Phylogeny generated using *lwPhylo* (Brown, 2020).

In contrast to flowering plants, polyploidy in animals and fungi is thought to be rare, famously so in mammals and birds, though many examples are known in certain lineages of amphibians, teleost fish and reptiles (Spoelhof et al., 2020). In animal groups where diploids and polyploids are both present there may be interploidy hybridisation and subsequent introgression, though based on the published literature this is very uncommon, with only three well-studied examples (Table 3). In many other cases where taxa with contrasting ploidies mate introgression is limited, as the hybrid derivatives are hybridogenetic taxa which lack recombination. For example, the edible frog *Pelophylax esculentus* is an extremely ecological successful and widespread hybrid species formed between the diploid taxa *P. ridibundus* and *P. lessonae*. *P. esculentus* includes two cytotypes, a diploid and a triploid, with the triploid being formed and maintained by haploid sperm fertilising the unreduced eggs from a diploid hybrid female (Hoffman et al., 2015). However, this taxon appears to be in a state of flux, with no documented all-triploid populations, and tetraploids are extremely rare. Opportunities for novel allelic combinations and introgression are limited as the parental genomes rarely recombine.

Table 3: Studies reporting cross-ploidy hybrids based on cytological and/or molecular genetic analyses. Details are provided of plant family, hybridising species, broad geographic locality, and the direction of introgression (if known). Superscripts indicate whether the polyploids are allopolyploid (^{allo}) or autopolyploid (^{auto}).

Family	Hybridising species	Location	Direction to	Reference
Animals				
Bufonidae	<i>Bufo turanensis</i> (2n = 2x = 22) x <i>Bufo pezwowi</i> (2n = 4x = 44) ^{allo}	Kyrgyzstan	Diploid	(Stöck et al., 2010)
Cyprinidae	<i>Squalius alburnoides</i> (2n = 2x = 50; 3n = 75; 4n = 100) x <i>S. pyrenaicus</i> (2n = 2x = 50) ^{allo}	Iberia	-	(Alves et al., 2001; Crespo-López et al., 2007)
Myobatrachidae	<i>Neobatrachus sutor</i> (2n = 2x = 24) x <i>N. kunapalari</i> (2n = 4x = 48) ^{auto}	Australia	Tetraploid	(Novikova et al., 2020)
Plants				
Aspleniaceae	<i>Asplenium scolopendrium</i> (2n = 2x = 72) x <i>A. adiantum-nigrum</i> (2n = 4x = 144)	Britain	-	(Stace et al., 2015)
Cyatheaceae	<i>Gymnosphaera denticulata</i> (2n = 2x = 138) x <i>G. metteniana</i> (2n = 4x = 274) ^{allo}	China	Tetraploid	(Wang et al., 2020)
Dryopteridaceae	<i>Polystichum setiferum</i> (2n = 2x = 82) x <i>P. aculeatum</i> (2n = 4x = 164)	Britain	-	(Manton, 1950)
Asteraceae	<i>Achillea clypeolata</i> (2n = 2x = 18) x <i>A. collina</i> (2n = 4x = 36) ^{allo}	Bulgaria	Tetraploid	(Guo et al., 2005)
Asteraceae	<i>Centaurea pseudophrygia</i> (2n = 2x = 22) x <i>C. jacea</i> (2n = 4x = 44)	Czech Republic	-	(Koutecký et al., 2011)
Asteraceae	<i>Cirsium carniolicum</i> ssp. <i>rufescens</i> (2n = 2x = 16) x <i>C. palustre</i> (2n = 4x = 34)	France	Tetraploid	(Segarra-Moragues et al., 2007)

Family	Hybridising species	Location	Direction to	Reference
Asteraceae	<i>Ixeris repens</i> (2n = 2x = 16) x <i>I. debilis</i> (2n = 6x = 48) ^{auto}	Japan	Hexaploid(?)	(Denda and Yokota, 2003)
Asteraceae	<i>Packera paupercula</i> (2n = 4x = 44) x <i>P. indecora</i> (2n = 8x = 88)	USA; Michigan	-	(Kowal et al., 2011)
Asteraceae	<i>Senecio madagascariensis</i> (2n = 2x) x <i>S. pinnatifolius</i> (2n = 4x)	Australia	-	(Prentis et al., 2007)
Asteraceae	<i>Senecio squalidus</i> (2n = 2x = 20) x <i>S. vulgaris</i> (2n = 4x = 40) ^{allo}	Britain	Tetraploid; chromo- some doubling	(Irwin and Abbott, 1992; Abbott et al., 1992, 2007; Abbott and Lowe, 2004; Chapman and Abbott, 2010)
Betulaceae	<i>Betula nana</i> (2n = 2x = 28) x <i>B. pubescens</i> (2n = 4x = 56) ^{allo}	Britain	Tetraploid; both(?)	(Palme et al., 2004; Thorsson et al., 2007; Wang et al., 2014)
Betulaceae	<i>Betula pendula</i> (2n = 2x = 28) x <i>B. pubescens</i> (2n = 4x = 56) ^{allo}	Britain	Tetraploid	(Zohren et al., 2016)
Betulaceae	<i>Betula × purpusii</i> (2n = 5x = 70) x <i>B. alleghaniensis</i> (2n = 6x = 84) ^{allo}	Michigan; USA	Hexaploid	(Barnes and Dancik, 1985)
Brassicaceae	<i>Cardamine apennina</i> (2n = 2x = 16) x <i>C. amporitana</i> (2n = 4x = 32)	Italy	Tetraploid	(Lihova et al., 2004)
Brassicaceae	<i>Cardamine × insueta</i> (2n = 3x = 24) x <i>C. pratensis</i> (2n = 4x = 32)	Switzerland	-	(Mandakova et al., 2013)
Brassicaceae	<i>Cochlearia officinalis</i> (2n = 4x = 24) x <i>C. danica</i> (2n = 6x = 42)	Britain	Tetraploid	(Fearn, 1977)

Family	Hybridising species	Location	Direction to	Reference
Brassicaceae	<i>Draba incana</i> ($2n = 4x = 32$) x <i>D. norvegica</i> ($2n = 6x = 48$) ^{allo}	Scandinavia	-	(Brochmann, Stedje, et al., 1992)
Brassicaceae	<i>Draba nivalis</i> ($2n = 2x = 16$) x <i>D. daurica</i> ($2n = 8x = 64$)	Scandinavia	-	(Brochmann, Stedje, et al., 1992)
Brassicaceae	<i>Draba arctica</i> ($2n = 10x = 80$) x <i>D. corymbosa</i> ($2n = 16x = 128$) ^{allo}	Scandinavia	-	(Brochmann, Stedje, et al., 1992)
Brassicaceae	<i>Rorippa australica</i> ($2n = 2x = 16$) x <i>R. sylvestris</i> ($2n = 4x/6x = 32/48$)	Germany	Both	(Bleeker, 2003); see also Bleeker (2007)
Fabaceae	<i>Lotus stepposus</i> ($2n = 2x = 12$) x <i>L. × ucrainicus</i> ($2n = 4x = 24$) ^{allo}	Ukraine, Turkmenistan, Kazakhstan, Mongolia	-	(Kramina et al., 2018)
Liliaceae	<i>Erythronium mesochoreum</i> ($2n = 2x = 22$) x <i>E. albidum</i> ($2n = 4x = 44$)	Nebraska; USA	-	(Roccaforte et al., 2015)
Orchidaceae	<i>Dactylorhiza fuchsii</i> ($2n = 2x = 40$) x <i>D. praetermissa</i> ($2n = 4x = 80$) ^{allo}	Belgium	-	(De Hert et al., 2012)
Orchidaceae	<i>Dactylorhiza incarnata</i> ($2n = 2x = 40$) x <i>D. praetermissa</i> ($2n = 4x = 80$) ^{allo}	Belgium	-	(De Hert et al., 2011, 2012)
Orchidaceae	<i>Dactylorhiza incarnata</i> subsp. <i>cruenta</i> ($2n = 2x = 40$) x <i>D. lapponica</i> ($2n = 4x = 80$) ^{allo}	Norway	Tetraploid	(Aagaard et al., 2005)

Family	Hybridising species	Location	Direction to	Reference
Orchidaceae	<i>Dactylorhiza incarnata</i> (2n = 2x = 40) x <i>D. traunsteineri</i> (2n = 4x = 80) ^{allo}	Sweden	Tetraploid	(Hedren, 2003); see also Balao et al. (2016)
Orchidaceae	<i>Dactylorhiza fuchsii</i> (2n = 2x = 40) x <i>D. maculata</i> (2n = 4x = 80) ^{auto}	Europe to Caucasus	-	(Shipunov et al., 2004)
Orchidaceae	<i>Epidendrum fulgens</i> (2n = 2x = 24) x <i>E. puniceoluteum</i> (2n = 4x = 52)	Brazil	Tetraploid	(Pinheiro et al., 2010)
Orobanchaceae	<i>Euphrasia anglica</i> (2n = 2x = 22) x <i>E. micrantha</i> (2n = 4x = 44) ^{allo}	Britain	Diploid(?)	(Yeo, 1956; French et al., 2008)
Phrymaceae	<i>Mimulus guttatus</i> (2n = 2x = 28) x <i>M. luteus</i> (2n = 4x = 60-2) ^{allo}	Britain	Chromosome doubling	(Vallejo-Marin, 2012)
Plantaginaceae	<i>Callitrichie cophocarpa</i> (2n = 2x = 10) x <i>C. platycarpa</i> (2n = 4x = 20) ^{allo}	Europe	-	(Prancl et al., 2014)
Poaceae	<i>Vulpia fasciculata</i> (2n = 4x = 28) x <i>Festuca rubra</i> (2n = 6x = 42)	Britain	Hexaploid(?)	(Bailey et al., 1993)
Polygalaceae	<i>Polygala calcarea</i> (2n = 2x = 34) x <i>P. vulgaris</i> (2n = 4x = 68)	Britain	Tetraploid	(Lack, 1995)
Polygonaceae	<i>Fallopica sachaliensis</i> (2n = 4x = 44) x <i>F. japonica</i> var <i>japonica</i> (2n = 8x = 88)	Britain	-	(Bailey, 2013); see also Bailey and Wisskirchen (2004) and Hollingsworth et al. (1999)
Polygonaceae	<i>Rumex obtusifolius</i> (2n = 4x = 40) x <i>R. aquaticus</i> (2n = 20x = 200)	Britain	20-ploid	(Ruhsam et al., 2015)
Primulaceae	<i>Dodecatheon frenchii</i> (2n = 2x = 44) x <i>D. meadia</i> (2n = 4x = 88)	Illinois; USA	Tetraploid	(Oberle et al., 2012)

Family	Hybridising species	Location	Direction to	Reference
Ranunculaceae	<i>Aconitum variegatum</i> ($2n = 2x = 16$) x <i>A. firmum</i> ($2n = 4x = 32$) ^{allo}	Europe	Diploid?	(Sutkowska et al., 2017)
Ranunculaceae	<i>Ficaria verna</i> subsp. <i>verna</i> ($2n = 4x = 32$)	Europe	-	(Popelka et al., 2019)
Rosaceae	<i>Rosa rugosa</i> ($2n = 2x = 14$) x <i>R. mollis</i> ($2n = 4x = 28$)	Europe	Tetraploid	(Kellner et al., 2012)
Violaceae	<i>Viola reichenbachiana</i> ($2n = 2x = 20$) x <i>V. riviniana</i> ($2n = 4x = 40$) ^{allo}	Germany	-	(Neuffer et al., 1999); see also Migdalek et al. (2017)

2.4 Biology of cross-ploidy hybrids: general features

Cross-ploidy hybrids can arise in a variety of situations. Many, but not all, examples occur in contact zones between parental species, where hybrid zones and hybrid swarms may form. Some of these hybrid zones have shifted over time (e.g. *Betula*, Wang et al., 2014), or are mosaic in structure (Popelka et al., 2019). In addition, there are notable differences in genetic structure between contact zones, with some comprising a swarm of F1, F2 and backcrossed hybrids (Fearn, 1977), indicating low genetic divergence between parental species (Edmands, 2002), while others contain only a few early generation hybrids, suggesting that parental species are more distantly related, and show higher levels of pre and post-zygotic isolation (Koutecky et al., 2011). Moreover, the direction of introgression is overwhelmingly towards the higher ploidy parent (21 out of 26 studies in Table 3 that reported directionality). This is unsurprising as the union of an unreduced $2n = 2x$ gamete of a diploid and a reduced $n = 2x$ gamete of a tetraploid provides a direct pathway for introgression in this direction, whereas the alternative direction is a two-step process via the triploid bridge (Stebbins, 1971; Baduel et al., 2018). As such, only two plant studies and one animal study report the opposite scenario (*Aconitum* and *Euphrasia*, *Neobatrachus*; Sutkowska et al., 2017; Yeo, 1956; Novikova et al., 2020), and a further two studies report bidirectional introgression (in *Betula* and *Rorippa*, Thorsson et al., 2007; Bleeker, 2003). However, other factors may still pose limits for introgression in the direction of the higher ploidy parent. Polyploids evolve meiotic stability to ensure reliable segregation of additional chromosomes at meiosis, with loci underlying tetraploid meiotic stability shown to be under selection in natural populations of autotetraploid

Arabidopsis arenosa (Hollister et al., 2012). Cytogenetic evidence in *Arabidopsis* suggests introgression from diploids to tetraploids may introduce genetic variants that disrupt regular meiosis in tetraploids (Morgan et al., 2020).

Hybrids may also occur in the absence of one or both parents, normally where greater lifespans allow persistence long after hybrid formation (Bailey, 2013; Preston and Pearman, 2015). Where cross-ploidy hybrids are present without their parents, they may represent stable lineages that survive through asexual reproduction (e.g. vegetative reproduction or apomixis), and are therefore different to some ephemeral forms present in hybrid zones. On occasion, cross-ploidy hybridisation has led to recent speciation (<200 years). This has occurred in the plant genera *Senecio* (Lowe and Abbott, 2004; Abbott and Lowe, 2004) and *Mimulus* (Vallejo-Marin, 2012). These hybrids are also notable in the context of the British Isles, as they involve alien species as either one, or both parental species. Further examples where cross-ploidy hybridisation involves aliens species are in *Rosa rugosa* and both parental species of *Fallopia* (Table 3). Human mediated translocations of species therefore continue to have a profound effect on hybridisation. Older hybrid species (10,000+ years) have also originated in a similar way to *Senecio* and *Mimulus* hybrid species, with this inferred either through morphology and cytogenetic analysis, or through sequence analysis showing ‘ghost’ subgenomes of allopolyploid species (e.g. *Euphrasia*, *Packera*, Yeo, 1956; Kowal et al., 2011).

A key determinant of genetic variation in cross-ploidy hybrids will be whether the polyploid parent(s) are auto or allopolyploids. In allotetraploid parents characterised by disomic inheritance, preferential chromosome pairing between the most similar, homeologous subgenomes, may lead to a subset of polyploid variation introgressing. In contrast, in autotetraploids with tetrasomic inheritance, free recombination between chromosomes may allow any region of the tetraploid to introgress. In our literature survey, we reported what kind of polyploid the higher ploidy parent in the cross was. 20 out of the 23 studies which contained information on polyploid type reported allopolyploids. While allopolyploids garner more research interest than autoployploids in studies of hybridisation (Spoelhof et al., 2017), the higher number of studies reporting allopolyploids may be biologically significant. For example, chromosome pairing of an allotetraploid subgenome more related to the diploid parent could lead to higher probabilities of successful hybridisation than in diploid-autotetraploid hybridisation, where chromosome pairing would be disrupted.

In addition to interploidy hybridisation between species, much early work, both theoretical and empirical, has explored crosses within mixed-ploidy species complexes (Levin, 1975; Fowler and Levin, 1984; Lumaret and Barrientos, 1990). The outcomes of crosses within (diploid x autoployploid) or between species (diploid x autoployploid/allopolyploid; Box 3) are similar in many cases; with triploid hybrids still formed (Vandijk et al., 1992; De Hert et al., 2012), unreduced gametes remaining an important driver of hybridisation (Lihova et

al., 2004; Baduel et al., 2018), and the direction of introgression is usually towards the higher ploidy parent (Table 3; Stebbins, 1956; Pinheiro et al., 2010). On the other hand, between species hybridisation can lead to higher levels of genetic variation through fixed heterozygosity in hybrids, and backcrossing to parental species, resulting in higher fitness (Ramsey and Schemske, 2002). In addition, the higher the divergence between species, the higher the likelihood of whole genome duplication post hybridisation, and therefore the generation of novel polyploid species (Paun et al., 2009).

Box 3: Outcomes of cross-ploidy hybridisation

The evolutionary outcomes once a hybrid has been generated are diverse and depend upon factors relating to hybrid creation frequency, population sizes of parental species, niche separation of hybrid and parental species (Fowler and Levin, 2016), the direction of introgression (Stebbins, 1971), hybrid fitness (Milne et al., 2003), and hybrid fertility (Petit et al., 1999). Taken together, these myriad barriers pose problems not only to the formation, but also to the establishment of cross-ploidy hybrid lineages.

After a cross-ploidy hybrid has formed, three outcomes may occur. The hybrid individual or population may either die before reaching maturity or go extinct, act as a conduit to gene flow between ploidy levels, or persist and establish to form a new hybrid entity or species. Firstly, extinction of the hybrid is highly likely if it is formed at low frequencies and parental species are rare (i.e. low propagule pressure; Fowler and Levin, 2016). The growth and development of the hybrid can be affected by bringing together incompatible parental allelic combinations, causing the hybrid to be unfit (e.g. hybrid necrosis; Bomblies and Weigel, 2007). Ultimately, fertility of an F1 hybrid will determine its persistence in a population.

Triploid F1 hybrids that overcome triploid block often display very low fertility (Figure 9 panels a and b) due to irregularities at meiosis which form aneuploid gametes (Tate et al., 2005). Tetraploid hybrids formed from unreduced gametes (Figure 7 panel c) have higher fertility than triploids (Petit et al., 1999); however there is no evidence to suggest that newly formed allotetraploids have higher fertility than autotetraploids, which may be expected if pairing behaviour is more regular in allotetraploids (Ramsey and Schemske, 2002).

Given that an F1 hybrid can produce (even rare) fertile gametes, low levels of outcrossing can promote gene flow between ploidy levels through backcrossing with parental species. For a triploid F1 hybrid, there are two pathways to generate a backcross of equivalent ploidy to one of the parental species. Firstly, the triploid F1 may produce reduced pollen which combines with reduced pollen from the diploid male parent (Figure 9 panel a) which has been hypothesised to occur in *Euphrasia* and *Aconitum* (Yeo, 1956; Sutkowska et al., 2017). Secondly, the triploid F1 hybrid can produce unreduced gametes that can either combine with reduced gametes from the tetraploid parent or unreduced gametes from the diploid parent (Figure 9 panel b; e.g. *Senecio eboracensis*; Lowe and Abbott, 2004). Tetraploids therefore are much more readily produced, as in addition to the two pathways mentioned, tetraploids can be produced in a single generation following cross-ploidy hybridisation (Figure 7 panel c). The bias towards tetraploid production has been known since Stebbins in the 1950s (Stebbins, 1956) and is the reason why introgression in the direction of the tetraploid is more common (Baduel et al., 2018).

Box 3: Outcomes of cross-ploidy hybridisation

For persistence of a hybrid lineage to occur, reproductive isolation between the newly formed hybrid and the parental progenitors is paramount. Unlike cases of polyploid hybrid speciation where the hybrid is of differing ploidy level to both parents, backcrossed F1 hybrids derived from cross-ploidy hybridisation will match one parental ploidy and therefore lack the strong reproductive barrier that polyploidy confers. In this case, other factors contribute to reproductive isolation, including ecological selection, niche differentiation, selfing, and chromosomal or genetic sterility barriers (Grant, 1981; Rieseberg, 1997; Gross and Rieseberg, 2005). Lastly, reproductive isolation of a cross-ploidy hybrid can occur by the doubling of the triploid F1 chromosome complement to produce a fertile hexaploid that is isolated by ploidy level from the parental species. This scenario has been recorded twice in recent history and has given rise to two neallohexaploid species, *Senecio camrensis* (Abbott and Lowe, 2004) and *Mimulus peregrinus* (Vallejo-Marin, 2012).

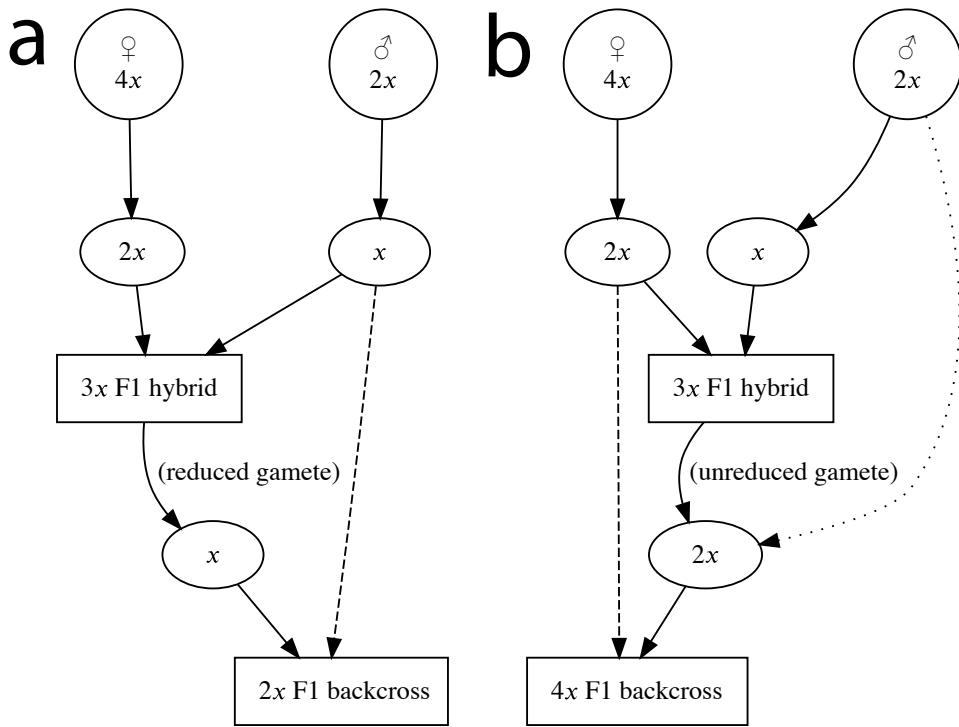


Figure 9: Potential outcomes of a triploid F1 backcrossing to the parental species. In both panels, the schematic follows that of Figure 7 panel a. Interrupted lines indicate backcrosses to parental species. In panel a the triploid F1 hybrid produces reduced gametes that combine with reduced gametes from the diploid male parent. In panel b there are two pathways to produce a tetraploid F1 backcross: firstly the unreduced gametes from the triploid F1 can combine with reduced gametes from the female tetraploid parent, secondly the unreduced gametes from the triploid F1 can combine with unreduced gametes from the diploid male parent. Figure generated with graphviz (Ellson et al., 2002).

It has been proposed that within polyploid complexes a widespread tetraploid could acquire genes via unilateral introgression from ecogeographically isolated diploid taxa occurring sympatrically with it in different parts of its range (Stebbins, 1956). In this way, several different forms of a tetraploid might originate, with each one bearing a close resemblance to the local diploid it hybridised with. Based on cytotaxonomic evidence, Stebbins (1956); Stebbins (1971) suggested this has occurred in numerous polyploid complexes of a number of plant genera, including *Dactylis*, *Knautia*, *Grindelia*, *Phacelia* and *Campanula*. Recently, genomic evidence has been obtained to provide support for Stebbins' proposal from work conducted on a polyploid complex comprising diploid and tetraploid forms of *Arabidopsis arenosa* in Europe (Arnold et al., 2015). Genomic analysis indicates that autotetraploid *A. arenosa* arose once and then split into five major lineages as it spread into different parts of Central Europe (Arnold et al., 2015). For two of the lineages, there is evidence that particular haplotypes, not found in any other tetraploid lineage, are shared with proximal diploid forms of

A. arenosa, indicating these haplotypes were acquired from the local diploid type and are adaptive (Arnold et al., 2015). In addition, one of the five tetraploid lineages is a ruderal form, widely distributed along the railways of Central and Northern Europe. Subsequent analysis indicates that the widespread lowland form of this early flowering and rapid cycling “railroad ecotype” likely originated as a result of introgression of genes from diploid *A. arenosa* occurring on the Baltic Coast of Germany and Poland into local populations of the tetraploid (Baduel et al., 2018; Monnahan et al., 2019).

2.5 Future perspectives

While cross-ploidy hybridisation is likely more common than previously thought, particularly in plants, there is still much uncertainty in our understanding of the phenomenon. To better determine the frequency of cross-ploidy hybridisation, we need to broaden the taxonomic scope under study. There is currently a dearth of information on animal examples, even though polyploid incidence can be high in some groups (e.g. insects, decapods, fish, and amphibians; Otto and Whitton, 2000). Further, while we found many angiosperm examples, half were derived from the large families Asteraceae and Orchidaceae. A broader scope will also determine more readily whether there is a phylogenetic signal to the phenomenon, and which attributes, from ecological to genetic factors, facilitate cross-ploidy hybridisation and introgression. Further, more detailed mechanistic research across a wide variety of taxa will reveal the underlying genomic variants that allow chromosomes to pair in newly formed polyploid hybrids (Morgan et al., 2020), which is important in establishment and persistence of hybrids. Most research on cross-ploidy hybridisation so far has focused on either contact zones or cryptic introgression; studying stabilised hybrids outside of these situations will provide a more detailed picture of how these lineages persist, and under which conditions (e.g. see Abbott et al., 1998 and references within). Next generation sequencing promises to reveal cross-ploidy hybridisation more easily (Wang et al., 2020), quantify the directionality of introgression accurately (Zohren et al., 2016), and determine parental genomic contributions to cross-ploidy hybrids (Bertioli et al., 2016). The latter point is particularly important, as hybrids may be introgressed at only a few loci in the genome. Detecting these few loci requires a good polyploid genome assembly, preferably with phase information, and new sequencing methods and software are beginning to address these problems (Zhang et al., 2019). Specifically, long-read Oxford Nanopore Technologies and Pacific BioSciences sequencing, as well as Hi-C and BioNano for scaffolding will produce highly improved, contiguous genome assemblies, and population long-read sequencing will be able to detect fine level introgression more easily. In addition, sequencing of diploid relatives, and haploid tissue in ferns will allow us to distinguish between subgenomes and work out phase. Given the extensive ploidy variation throughout plants and animals, and the high degree to hybridisation is detected in these groups, cross-ploidy

hybridisation may be more frequent and important in plant and animal evolution than is currently thought.

3 The genetic landscape of hybridisation in UK flora

I acknowledge that Natasha De Vere and Laura Jones provided unpublished DNA barcoding data for the British flora.

3.1 Abstract

Hybridisation has a profound impact on the evolution of plants, with consequences including the generation of novel phenotypes (Lexer et al., 2003), introgression of adaptive alleles (Chapman and Abbott, 2010), and the origin of new species (Ainouche et al., 2009), as well as the blurring of species boundaries (Bardy et al., 2011), and the extinction of rare taxa. While natural hybridisation is common and widespread both geographically and phylogenetically (Ellstrand et al., 1996; Whitney et al., 2010; Beddows and Rose, 2018), the ecological factors and species traits that promote or prevent hybridisation in natural plant communities are poorly understood. Many important traits related to hybridisation, such as parental genetic distance and ploidy level differences have not been systematically quantified across diverse species in a flora, while estimates for phylogenetic effects of hybridisation in plant communities are highly variable (Whitney et al., 2010; Beddows and Rose, 2018; Mitchell et al., 2019). Here, we combine use phylogenetic mixed effect models to combine estimates of hybridisation from extensive field observations (Stace et al., 2015) with DNA barcoding data and ecological attributes for over 1,000 native species of flowering plant. Our results quantify the influence of different predictors, and the genetic distances over which hybridisation is most likely. We also quantify the importance of phylogenetic relatedness and parental ploidy differences in shaping the likelihood of hybrid formation. Life history of parental species does not impact hybrid formation despite perennial life history being pervasive in both parental species and hybrids. Although the effect of genetic distance requires careful interpretation as low parental divergence may either allow hybrids to form or be a consequence of genetic homogenisation, taken together with ploidy and phylogenetic effects, genetic factors are key predictors of hybridisation across diverse taxa in the UK flora.

3.2 Main

Natural hybridisation plays an important role in plant evolution by facilitating adaptation and promoting species survival, generating novel variation, or alternatively in some cases leading to a reduction in fitness and even causing extinction (Rhymer and Simberloff, 1996; Rieseberg et al., 1999; Chapman and Abbott, 2010;

Becker et al., 2013). Focused studies on plant evolutionary model systems such as *Helianthus* (Lexer et al., 2003), *Senecio* (Abbott et al., 2009), and *Tragopogon* (Novak et al., 1991) have revealed how hybridisation may contribute to range expansion, invasiveness and phenotypic and genomic novelty. These focused genetic studies have selected species and study systems for their noteworthy hybrid outcomes, and it is currently unclear how these results generalise to natural plant communities and entire floristic assemblages. A relatively small body of work has looked more broadly at hybridisation between diverse species in floras, with the aim of understanding the likelihood of hybridisation in the context of species attributes, ecology, and phylogeny (Ellstrand et al., 1996; Whitney et al., 2010; Beddows and Rose, 2018; Marques et al., 2018; Mitchell et al., 2019). A limitation of these studies has been the reliance on phylogenies where the tips represent higher taxonomic units such as genera, families, or orders (Whitney et al., 2010; Beddows and Rose, 2018; Mitchell et al., 2019). Hybridisation is an outcome of interactions between species, and a resolved species-level phylogeny is required to calculate the genetic distances of hybridising taxa and to accurately estimate the phylogenetic signal of hybridisation.

DNA barcoding is proving increasingly useful for the genetic characterisation of species assemblages and complex communities (Gomez-Rodriguez et al., 2015; Young et al., 2019). By amplifying a small and standardised set of target loci, DNA barcoding is inherently scalable to deployment across large sample sets and is well-suited to comparative analyses of diverse taxa. As DNA data is generated at pace and at scale it is becoming feasible to generate comprehensive DNA barcoding datasets for all plant species in a country. This provides the opportunity for the integration with ecological data collected at a national level to test major ecological and evolutionary questions at broad geographic scales.

In this study we characterise the genetic landscape of natural hybridisation across flowering plant species in the UK. The UK flora is an ideal study system for investigating hybridisation as it contains ~1400 species and is therefore manageable for genetic characterisation, but has sufficient diversity to include ~20% of all known angiosperm plant families and numerous hybrid combinations (Stace et al., 2015). There is also a hybrid flora — a unique resource describing all known vascular plant inter-specific hybrids present across the British Isles, of which 616 represent hybridisation between native UK flowering plant species or archaeophytes (introduced pre-1500 or potentially natives) (Stace, 1975; Stace et al., 2015; Preston and Pearman, 2015). The hybrid flora also summarises their ecology, distribution, cytology and parentage (Stace et al., 2015). We use the Barcode UK dataset (Jones et al. In Review.), a new DNA barcoding resource that includes a three locus DNA barcode of *rbcL*, *matK*, and *ITS2* for the native and archaeophyte seed plants of the UK. We integrate this DNA barcode data with information on hybridisation from the hybrid flora, as well as other species level traits and ecological information. We assess the determinants of hybridisation across the flora in the context

of parental species range overlap, genus size, life history (annual vs perennial), ploidy differences, parental genetic distance and phylogeny.

Out of the 6117 possible unique pairwise congeneric combinations between 1100 species suitable for analysis in the UK flora (see Methods), only 7.8% produce hybrids. From 244 genera containing multiple species, 96 contain hybrids, and the 480 recorded hybrids are disproportionately concentrated in just five genera, with 45.8% of hybrids found in *Euphrasia* ($n = 62$), *Carex* ($n = 50$), *Rosa* ($n = 40$), *Epilobium* ($n = 35$), and *Salix* ($n = 33$). Some genera have many species that prolifically hybridise (e.g. *Euphrasia* has 4 species that are each parents to 14+ hybrids), while others include widespread species that disproportionately contribute to the number of hybrids (e.g. *Rumex crispus*, involved in 12 hybrid combinations in our data). We then explored whether the number of hybrids is a simple function of genus size using phylogenetic mixed models. Genus size is predicted to affect hybridisation as larger genera contain more possible pairwise combinations of species, and thus greater opportunity for hybridising with congeners (Johnson, 2018). Our models show that the probability congeneric species hybridising is independent of genus size (pMCMC = 0.92, Appendix 2 Table 2). Although hybridisation tends to occur in species rich genera, not all species rich genera form hybrids (see also Preston and Pearman (2015)). For example, genera such as *Trifolium* (19 species; usually self-incompatible) and *Alchemilla* (12; all apomictic) are relatively species rich but form no naturally occurring hybrids, whilst *Veronica* (15) and *Galium* (13) form only one naturally occurring hybrid each. This shows that the number of hybrid taxa deviates from a simple model determined by the number of congeneric taxa, with other factors such as phylogenetic relatedness, life history, genetic distance, and ploidy level (explored below) interacting to determine hybridity.

We then investigate hybrid formation in the context of a newly generated 3-locus phylogeny of the British Flora from the Barcode UK data (see Methods). Hybridisation propensity, here defined as the number of hybrid combinations a species produces weighted by genus size, is highly uneven across the plant phylogeny (Figure 10). The phylogenetic signal of hybridisation is high at 0.62 (0.32-0.77 CI – 95% Credible Intervals) meaning that closely related lineages are likely to have similar levels of hybridisation (Appendix 2 Table 3). Inspection of the species level Best Linear Unbiased Predictors (BLUPs; see Methods) from the phylogenetic model shows that the monocot Asparagales clade including orchids (Orchidaceae, Iridaceae, Asparagaceae and Amaryllidaceae) are most likely to hybridise after accounting for other model factors, whilst legumes (Fabaceae) are the least likely (Appendix 2 Figure 2). These results are consistent with previous work; many legumes, such as clovers (*Trifolium*) and related genera (e.g. peas, *Lathyrus*) are known to hybridise very little due to strong between species incompatibilities (Evans, 1962), while orchids have been shown to hybridise rampantly, even across ploidy levels (De Hert et al., 2012). Overall, the pattern of hybridisation is highly

heterogeneous across the phylogeny for the British flora, with phylogenetic position a good predictor of hybrid formation.

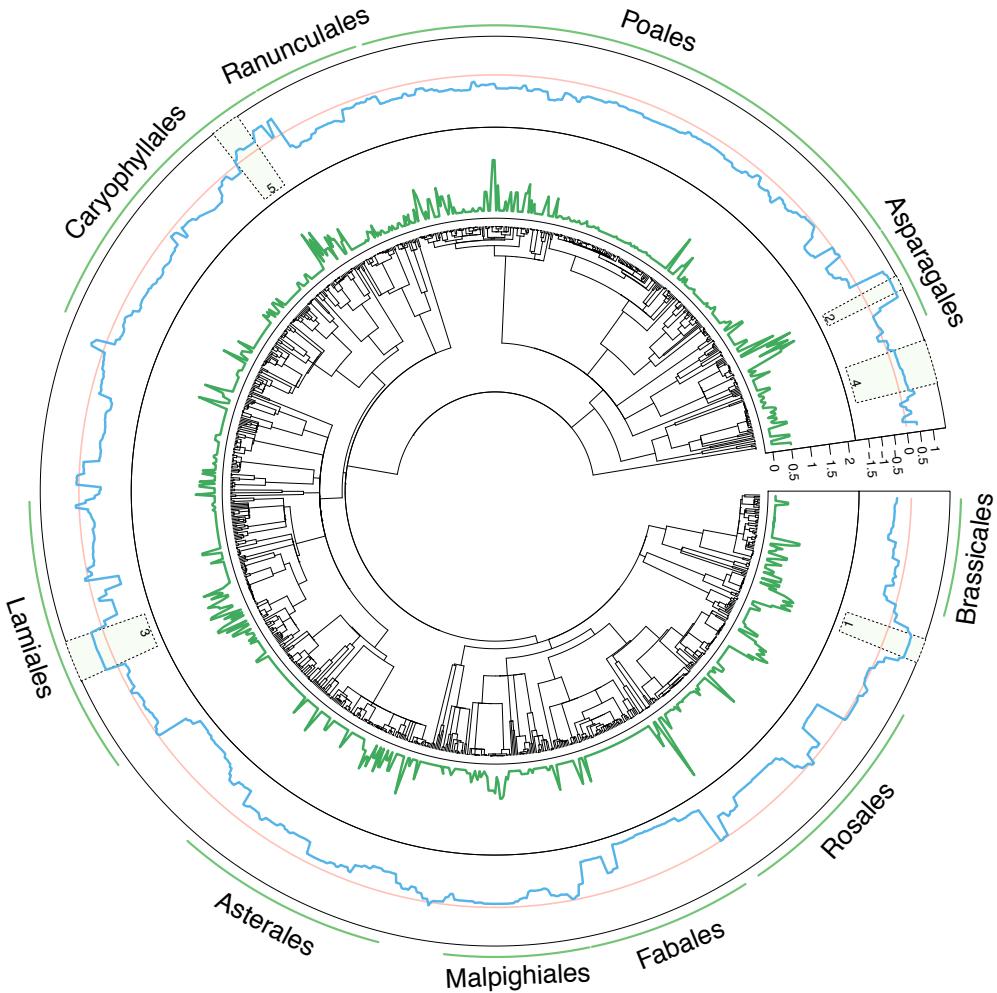


Figure 10: Distribution of hybrids across the phylogeny of the British flora. Innermost-ring shows phylogenetic relationships of 1098 British native species from an alignment of ITS2, *matK* and *rbcL*. Phylogenetic reconstructions used maximum likelihood implemented in IQ-TREE. The middle ring (green line) shows species-level hybrid propensity weighted by size of genus. The outer ring shows the probit scale posterior mean of the probability of a particular species hybridising (blue line). The zero line is represented in pale red and positive probit values indicate higher probabilities of hybridisation. The figure is annotated with the five genera with highest probabilities of hybridisation, given variation in model fixed effects, indicated from the sum of the species level posterior means from the phylogenetic model (1. *Epilobium*, 2. *Euphrasia*, 3. *Dactylorhiza*, 4. *Potamogeton*, and 5. *Rumex*), and the 10 largest plant orders around the outside.

Previous studies have explored the relationship between life history and hybridisation in plant species (Stace, 1975; Ellstrand et al., 1996; Preston and Pearman, 2015; Beddows and Rose, 2018; Mitchell et al., 2019). It

has been documented that most hybridising species are perennial, at least in temperate floras (e.g. 97% in the Michigan flora Beddows and Rose (2018), 68% in our data). Out of all potential congeneric species pairs of the same life history in the British flora however, perennial only parental combinations form proportionally fewer hybrids (7.7%, n = 4725) than annual only parental combinations (15.6%, n = 588). Perennial plant species are thought to participate in hybridisation events more frequently because (a) perennials tend to outcross more than annuals, which tend to be more highly selfing (Morgan, 2001) and (b) perennial plants are longer lived and therefore more gametes are produced over a longer period of time (Ellstrand et al., 1996). Despite the high frequency of perennial parent plant species we find no evidence for a significant effect of life history on the probability of hybrid formation ($\chi^2 = 4.04$, df=2, P=0.13; Appendix 2 Table 4). Life history may be still be important in the persistence of hybrid lineages (Ellstrand and Schierenbeck, 2000), however our results show that the probability of forming a hybrid is independent of parental species life history.

Increasing overlap in parental species distribution is expected to increase the probability of hybridisation due to the greater opportunity for crossing and therefore hybrid formation. Whilst this is an important variable when considering hybridisation, few studies have critically looked at range overlap with respect to hybridisation in multi-species systems (Preston and Pearman, 2015; Mitchell et al., 2019). We leverage accurate information on the distribution of British plant species to infer the influence of species range overlap measured as a count of the number of 10x10km grid (hectad) overlaps. As expected, mean overlap in congeneric parental species distribution is higher for pairs of species known to give rise to hybrids (739 hectads ± 27 SE, Standard Error), compared to those that have not been recorded to successfully hybridise (353 hectads ($=10 \times 10 \text{ km}^2$) ± 6 SE). Our models predict significantly higher probabilities of hybrid formation when there is larger overlap in parental species distribution (pMCMC < 0.001, Appendix 2 Table 2). Although significant, the variability in the effect of parental distribution overlap is very low (posterior SD: 0.0001) compared to that of genetic distance between parental species, which is five orders of magnitude more variable (posterior SD: 3.77; Appendix 2 Figure 3). This suggests that range overlap and correlated attributes such as species abundance may be secondary to intrinsic genetic factors in determining hybridisation at a broad-spatial scale (Brown, 1984), though they may be key factors at a local scale (Heinze, 2011). In sum, there are opportunities for hybrids to occur even in areas where closely related parental species overlap little or not at all, with hybrid presence potentially affected by historical range overlap, long distance cross-pollination when parental species have coincident phenologies, or independent dispersal of hybrids (Lamont et al., 2003; Preston and Pearman, 2015).

Ploidy level variation is frequent, both within and between species in the same genus (Husband et al., 2013). Hybrids formed from parental species of differing ploidy level (cross ploidy hybrids) appear in the British flora,

with 131 detected (38% of hybridising species pairs with ploidy information Stace et al. (2015)), but their relative importance has not yet been investigated across a flora. Hybrid formation is expected to decrease when the parental species have contrasting ploidy levels due to endosperm imbalance in the fertilised embryo (Tate et al., 2005). We used ploidy data for 684 species across the British Flora and determined for each pairwise comparison of species whether they were of the same or different ploidy. Our model shows that parental species with the same ploidy are 35% more likely to form hybrids than parents of differing ploidy levels, when fixed at mean overlap in geographical distribution, mean branch length between species pairs, and accounting for phylogenetic effects ($pMCMC < 0.001$, Appendix 2 Table 4; Appendix 2 Figures 4 and 5). Cross ploidy hybridisation has been reported in many plant genera and has led to the generation of new species (Elkington, 1984) and introgression of genes affecting fitness (Chapman and Abbott, 2010), highlighting the importance of rare hybridisation between ploidy levels.

We then investigated the impact of parental genetic distance on hybrid formation across taxa in the flora. The likelihood of hybrid formation is expected to decrease with parental genetic distance due to a greater number of genetic incompatibilities (Edmands, 2002). However, low genetic distances may also be a consequence of genetic homogenisation from hybridisation. We observe a ten-fold variation in mean congeneric ITS distance (see Methods for definition) across the 35 genera containing hybrids with more than five taxa, from low mean pairwise distance in *Agrostis*, *Cochlearia* and *Rosa*, to high distance in *Geranium*, *Juncus* and *Saxifraga*. Overall, hybridising congeneric species showed a significantly lower pairwise genetic distance (mean ITS distance = 0.097, SE = 0.004) than non-hybridising congeneric species pairs (mean ITS distance = 0.215, SE = 0.001, Wilcoxon Test, $P < 0.001$; Figure 11). In our tree based phylogenetic models, the probability of forming a hybrid strongly decreases as branch length between parental species increases ($pMCMC < 0.001$, Appendix 2 Table 2 and Appendix 2 Figure 6) and shows a greater standardised effect size than both pairwise overlap in distribution and size of genus. The stronger effect of genetic distance is evident from the joint probability distribution of hybridisation and geographical distance (Appendix 2 Figure 2), which are predicted by theory to be correlated (Felsenstein, 1976). For species which hybridise, the highest average parental genetic distance is seen in the genus *Saxifraga* (ITS2 distance: 0.28), with other divergent hybridising taxa seen in the genera *Poa* (0.22), *Cardamine* (0.19), *Potamogeton* (0.17) and *Fumaria* (0.17), showing rare examples of hybrid formation between divergent taxa. In each of these five genera, hybrids that do form tend to be sterile and therefore introgression and genetic homogenisation are unlikely (Hegde et al., 2006). This contrasts with genera characterised by low mean parental genetic distance, such as *Salicornia* (ITS2 distance: 0.00), *Prunus* (0.00), *Rosa* (0.02), *Epipactis* (0.05), and *Atriplex* (0.05), where hybrids that form tend to be fertile (Stace et al., 2015; Martin and Mendelson, 2018). Overall these results show parental genetic divergence is a good predictor of hybrid formation, but that low parental genetic divergence may either be a cause of

hybrid formation, or a consequence of genetic homogenisation. In cases where hybridisation does occur, there may be long lasting evolutionary consequences such as phenotypic novelty (Stelkens and Seehausen, 2009) and polyploid hybrid speciation, especially when there is high parental divergence (Chapman and Burke, 2007).

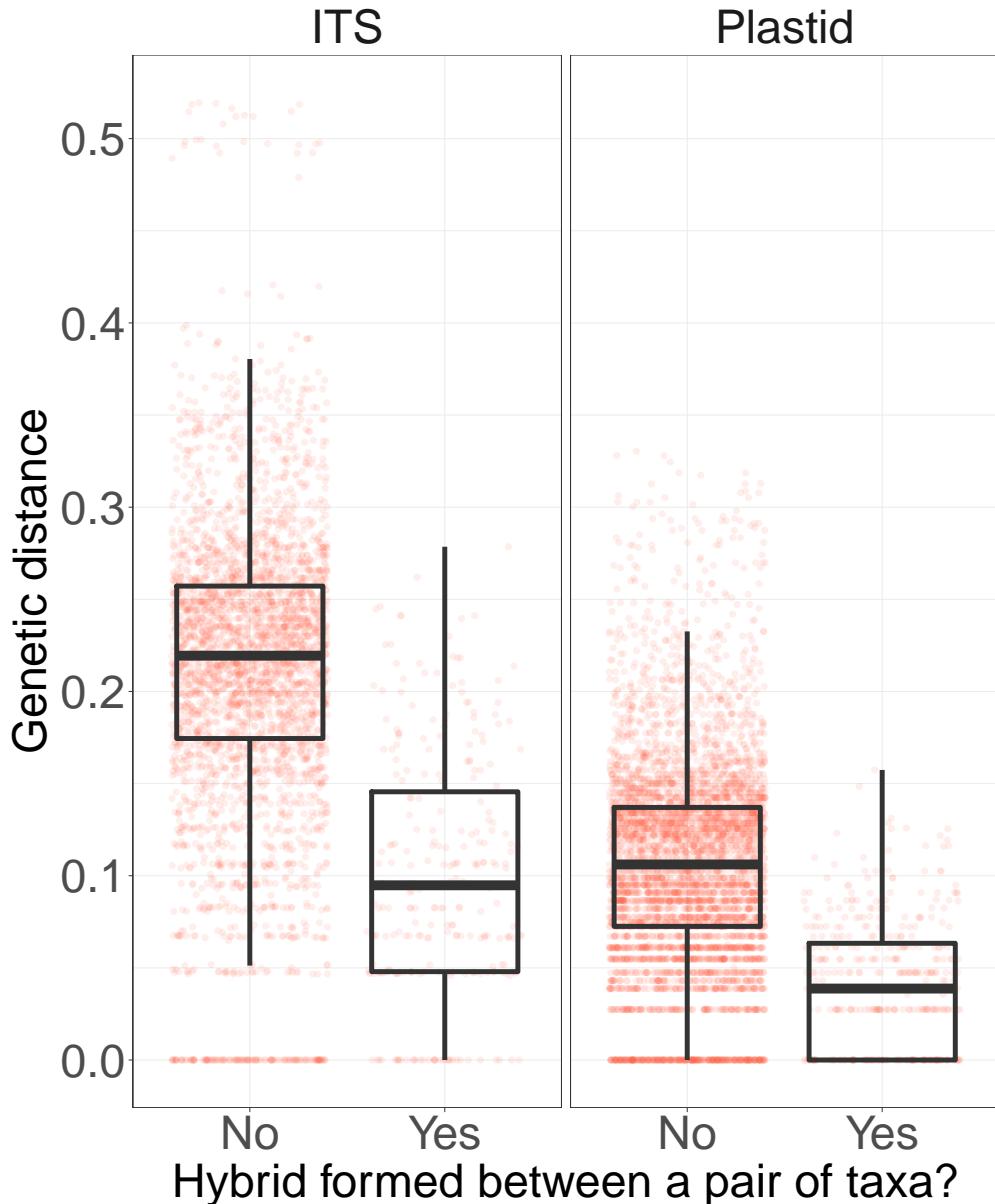


Figure 11: Hybrid formation in the context of genetic divergence in the British flora. Jittered points represent genetic distances between pairs of congeneric taxa, grouped by whether a pair of taxa produced a recorded hybrid or not. Both ITS and plastid loci are shown.

This study is the first to integrate a species level phylogeny, species traits, and ecological information for an entire flora, to understand the relative effects of genetic and ecological factors that affect hybridisation in plants. We report a high phylogenetic signal of hybridisation in flowering plants, supporting some (Whitney et al., 2010) but not all previous studies (Beddows and Rose, 2018; Mitchell et al., 2019), and highlighting the importance of taking species relationships into account. Phylogenetic effects are complex, and may be attributed to multiple unmeasured traits of parental species, such as habitat preferences, chromosomal stability, or mating systems (Ramsey et al., 2003; Brys et al., 2014; Bittencourt, 2019). Genetic distance between parental species emerges as the strongest predictor of hybridisation, emphasising the importance of this variable known to be important in pre and postzygotic isolation (Edmands, 2002; Moyle et al., 2004). Causality is not absolute however, as low genetic distance between parental species could also be driven by genetic homogenisation reducing sequence variability. Recently diverged species within genera such as *Euphrasia*, *Sorbus* and *Epipactis* (Ennos et al., 2012) are examples of genera with closely related species which hybridise extensively and where plastid and ITS introgression is likely. Hybridisation is also affected by ploidy level of parental species, with parents of differing ploidy less likely to form hybrids. Whilst ploidy differences are expected to lower the probability of hybridisation due to prezygotic barriers such as meiotic irregularities (Tate et al., 2005), the probability of forming a hybrid in this situation is still well above zero. Ploidy level can therefore be considered a leaky barrier to hybridisation (Abbott and Lowe, 2004). Overall, genetic factors at all levels have a profound impact on hybridisation.

Other, non-genetic, factors have been reported to play an important role in the prezygotic isolation of plant species and therefore in hybrid formation (Widmer et al., 2009). Factors such as genus size and life history are expected to predict hybridisation (Whitney et al., 2010), however we find no evidence of this. In the case of genus size we attribute the lack of a clear cut effect to phylogenetic signal and genetic distance between species outweighing increased opportunity for hybridisation due to more congeners alone. Previous work has emphasised the importance of the perennial life history in hybridisation (Ellstrand et al., 1996; Preston and Pearman, 2015; Beddows and Rose, 2018; Mitchell et al., 2019). We detect no effect of life history in the probability of hybrid formation, and attribute this lack of effect to a few different factors. High recorder effort in the British Isles means that even ephemeral annual hybrids are routinely found (Preston and Pearman, 2015), and we accounted for fewer annual than perennial plant species in the British flora by modelling all pairwise intrageneric species pairs. We also disentangled hybrid formation from hybrid persistence by using observed hybrid occurrences. Lastly, it may be that the annual genus *Euphrasia* is having a marked effect, as it is one of the few solely annual genera with many hybrids recorded (Preston and Pearman, 2015; Metherell and Rumsey, 2018). Hybridisation is also constrained and shaped by the biogeography of parental species (Lowry et al., 2008). We reveal that a hybrid is more likely to be formed when parental species distributions

are highly overlapping due to increased opportunity for crossing events; a factor not yet statistically modelled in any flora-wide study to date. Parental species overlap remains a crude estimate due to lack of resolution on fine scale co-occurrence, and does not take into account habitat change through space, or levels of habitat disturbance. Britain has a postglacial flora with high levels of disturbance which is known to change the landscape of hybridisation (Abbott, 1992; Guo, 2014) and more studies are needed in undisturbed habitats to understand hybridisation in a comparative context.

3.3 Methods

We extracted information on hybrid taxa, their parental progenitors, and ploidy from the Hybrid Flora of the British Isles (Stace et al., 2015) and used the latest plant taxonomy according to the New Flora of the British Isles (Stace, 2019). The Hybrid Flora excludes complex, apomictic groups like *Hieracium*, *Taraxacum* and *Rubus*. Further ploidy information was extracted from the BSBI Cytology database and the Kew C-value database using custom python scripts (see https://github.com/Euphrasiologist/web_mining). Ploidy levels were inferred using R scripts (https://github.com/Euphrasiologist/Floristic_DNA_Barcoding) and each was checked manually, using only chromosome counts based on native UK material. Species with multiple ploidy levels were excluded unless it was exactly known which cytotype contributed to the hybrid. We excluded: (a) hybrids known not to have formed in the British Isles (e.g. taxa introduced as hybrids), (b) triple hybrids, (c) dubious or doubtful hybrids, (d) crosses at below specific rank (subspecies, varieties) and (e) hybrids where at least one parent was a recently introduced non-native species (however archaeophytes, introduced pre-1500, were included). We also removed the rare cases of intergeneric hybridisation (some Rosaceae, Poaceae and Orchidaceae) due to model scaling issues associated with looking at all possible species combinations across the flora. Downstream, hybrids were excluded if there was no barcode data associated with the parental species. Hybrid propensity was calculated by counting the total number of hybrid taxa that a particular parental taxa had participated in. We did not scale the hybrid propensity by genus size (sensu Whitney et al. (2010)) as the data was only used for visual interpretation.

We estimated phylogenetic relationships from the Barcode UK dataset (Jones et al., In Review), which includes a three locus DNA barcode for British native flowering plant species. Complex or apomictic groups, as omitted from the Hybrid Flora, were not sequenced (except for *Sorbus*, as there was well sampled DNA barcodes for the group). Due to the different diversities and alignment success of plastid and nuclear ribosomal DNA we used a single alignment of plastid sequences to infer relationships between all taxa, while ITS was aligned separately for each genus and thus only used to infer congeneric relationships. Plastid DNA was

aligned for all taxa using the R package DECIPHER (W., 2016), while ITS sequences were aligned per genus using MAFFT (Katoh and Standley, 2013). Between genera, sequences were gapped using the program catfasta2phym (https://github.com/nylander/catfasta2phym) using padding N's. Phylogenetic inferences were made using IQ-TREE (Nguyen et al., 2015) in an analysis with three partitions allowing models of molecular evolution to differ between loci, and including a multifurcating constraint tree based on APGIV relationships generated with Phylomatic (Webb and Donoghue, 2005). Tree support was estimated using 1000 ultrafast bootstraps (Hoang et al., 2018). Tree-based genetic distances were inferred using the R function cophenetic.phylo() from the package ape (Paradis et al., 2004) while separate pairwise distances for ITS and plastid DNA were calculated with the R function dist.alignment() from the seqinr package (Charif and Lobry, 2007). The resulting values were the square root of pairwise distances. Tree manipulation took place in R, with the circular plot made with the R package circlize (Gu et al., 2014); the phylogeny was coerced into a circular dendrogram for visualisation. Other plots were generated with the R package ggplot2 (Wickham, 2016) and lattice (Sarkar, 2008). All other data manipulation took place in R version 3.6.1 using base R, and packages data.table (Dowle and Srinivasan, 2019) and dplyr (Wickham et al., 2020).

Our final dataset for the phylogenetic analysis contained a maximum of 466 hybrid combinations in 6103 unique congeneric pairwise combinations. In the analysis of genetic distance, the plastid genetic distance dataset contained 413 hybrid combinations, and the ITS genetic distance dataset 279. We used phylogenetic generalised linear mixed models implemented in the R package MCMCglmm as it allows for addition of a phylogeny with flexible variance structures for the random effects (Hadfield, 2010). The response variable was a binary response of whether two congeneric species produced recorded hybrids or not, and was assumed to have residuals approximated by a probit distribution. We used parameter expanded priors for better mixing and fixed the residual variance at 1. The models were run for 1.3 million iterations with a thinning interval of 1000 and a burn-in of 300000. We used five fixed effect covariates to understand their contribution to explaining the variation in hybrid formation. Firstly, pairwise branch length between parental species calculated from the phylogeny (above) was added to understand the contribution of intrageneric relatedness. Pairwise overlap in geographical distribution (number of 10x10km²; generated from data at https://database.bsbi.org/) accounted for extent of overlap of parental species. Genus size for each species was calculated from species present in the phylogeny. Lastly, whether the parental species were of the same ploidy level or not, was added as a two level categorical factor; same ploidy level (homoploid) or not (heteroploid). We ran two models which differed in their fixed effect structure only, with the addition of the ploidy level and one without. This was because ploidy data was limited either due to missing counts or if ploidy was difficult to estimate; the number of species analysed decreased from 1098 to 684 upon addition of ploidy level. 148 out of 244 families contained missing ploidy data, with a few genera containing high amounts of missing data, e.g. *Sorbus*, *Alchemilla*,

Juncus, and especially *Carex*. The phylogenetic (species level) best linear unbiased predictors (BLUP's; means of the posterior distribution) were extracted from the model excluding ploidy level and are equivalent to the per species point estimates of the probability of hybridisation.

The inverse relatedness matrix (unscaled phylogeny) and species were fitted as random effects in a multi-membership model structure, as each hybrid event is the outcome of two parental species. The effect of phylogeny was added to the fixed effect predictions by calculating:

$$\sigma_u^2 + 2\sigma_s^2 + (\sigma_p^2)\bar{B} \quad (1)$$

Where u is the residual variance, s is the species variance and p is the phylogenetic variance. \bar{B} is defined as the average tree of all species pairs:

$$\bar{B} = 2\bar{d} + 2\bar{o} \quad (2)$$

Where \bar{d} is the average species phylogenetic variance and \bar{o} is the sum of the pairwise species phylogenetic covariances divided by the number of possible combinations of species multiplied by two. As we did not allow intergeneric hybridisation, we calculated \bar{B} at the species level constrained to genus (\bar{B}_S):

$$\bar{B}_S = \frac{2 \sum o_i n c_i + 2 \sum d_i n_i}{\sum n c_i} \quad (3)$$

i represents the i th genus. Variables o and d are defined as above but calculated for each genus, n_i is the number of individuals in a genus and $n c_i$ is the number of possible pairwise combinations within genus. Using this method, we were able to account for the size of genus in our phylogenetic variance estimates. We implement this algorithm in the R package VCVglmm (Brown, 2019). Lastly, phylogenetic signal was calculated using:

$$\frac{\sigma_p^2 \bar{B}_S}{\sigma_u^2 + 2\sigma_s^2 + 2\sigma_p^2 \bar{B}_S} \quad (4)$$

All parameters are sampled from the posterior distribution of each coefficient, and distributions are summarised using modes and highest posterior density intervals at the 95% level. P-values were taken directly from model

output for continuous covariates or categorical covariates with only two levels, otherwise (for life history effects) Wald Tests jointly tested all factor levels (Brown, 2019). Genetic distance comparisons were made using Wilcoxon tests on the pairwise genetic distances, which does not assume that the distances approximate any distribution.

4 Is there evidence of diploid-tetraploid hybridisation in a *Euphrasia* contact zone?

I acknowledge that Hannes Becher carried out demographic simulations and assisted in the interpretation of the genetic data, and Sebastian Williams generated all ITS1 sequence data.

4.1 Abstract

Cross-ploidy hybridisation is expected to be rare due to the strong ploidal prezygotic barrier. Hybridisation between species that differ in ploidy level however, has been reported across flowering plants. British eyebrights (*Euphrasia*) represent a good study system to investigate cross-ploidy hybridisation, as 71 different hybrid combinations have been reported and at least two putative hybrid species (*E. vigursii* and *E. rivularis*) have formed from progenitor species of different ploidy levels. In this study, we analysed a contact zone between the diploid species *E. rostkoviana* and the tetraploid species *E. arctica* in Wales. We sequenced the internal transcribed spacer region (ITS1), and used Genotyping by Sequencing (GBS) to look for evidence of cross-ploidy hybridisation and introgression. All sites in the ITS1 region were fixed between diploids and tetraploids, indicating a strong barrier to hybridisation. Further, analysis of the GBS data using PCA, STRUCTURE, and AMOVA across 270 SNPs, indicated clear separation between the ploidy levels. While the global F_{ST} between species was high at 0.44, the distribution across all SNPs was bimodal, indicating potential differential selection on loci between diploids and tetraploids. Only using demographic inference with *δαδI* did we find evidence of limited gene flow – around one or fewer migrants per generation. Overall, our results are consistent with cross-ploidy hybridisation being rare or absent. Overall, secondary contact between species of different ploidies can result in a mix of outcomes - from hybrid speciation, to rare hybridisation as reported here.

4.2 Introduction

Natural hybridisation is an important evolutionary phenomenon with wide ranging consequences, from extinction (Rhymer and Simberloff, 1996), to hybrid speciation (Hegarty and Hiscock, 2005). Most studies to date have investigated hybridisation between diploid species, while hybridisation between species that differ in their ploidy level (cross-ploidy hybridisation) has generally received less attention (though see Petit et al., 1999). While this may in part be due to technical issues with inferring homology relationships between diploids

and polyploids, there are also clear biological reasons. Contrasting ploidy levels represent a known, highly effective barrier to hybridisation (Husband and Sabara, 2004). The main barriers are abnormal endosperm ratios of maternal:paternal genomes at fertilisation which prevent hybrid seed formation (Johnston et al., 1980), and later hybrid sterility caused by irregularities in chromosome pairing at meiosis leading to aneuploid gametes (Tate et al., 2005). Both of these factors prevent hybridisation and introgression between species with contrasting ploidy levels. These barriers can be overcome through unreduced gamete production in the lower ploidy parent, or where a triploid (or other intermediate ploidy) F1 hybrid is formed, by either backcrossing to one of the parental species (Ramsey and Schemske, 1998), or by whole genome duplication to restore fertility (Abbott and Lowe, 2004). Cross ploidy hybridisation may be an important mechanism for maintaining genetic variation in polyploid species (although distinguishing this from recurrent polyploidisation can be difficult Shimizu-Inatsugi et al., 2009), exchanging adaptive alleles between species (Chapman and Abbott, 2010), and has been shown to generate new polyploid cytotypes or species (Abbott and Lowe, 2004).

A number of natural cross-ploidy hybridisation examples exist in the literature and these cover a variety of phylogenetically distinct taxa. For example, there is cross ploidy hybridisation reported in *Dactylorhiza* (De Hert et al., 2012), *Mercurialis* (Buggs and Pannell, 2007), and *Epidendrum* (Pinheiro et al., 2010), the latter two of which often form hybrid zones. Where hybridisation is particularly common, hybrid swarms can develop, as seen in co-occurring diploid and tetraploid species of *Cochlearia* (Fearn, 1977). Cross ploidy hybridisation in the genera *Senecio* and *Mimulus* have led to the creation of three hybrid species endemic to Britain (Abbott and Lowe, 2004; Vallejo-Marin, 2012). Two of these hybrid species have resulted from whole genome duplication of initial triploid F1 hybrids (*Senecio camrensis* and *Mimulus peregrinus*), whilst the other species was created through introgression to the tetraploid parent (*Senecio eboracensis*). Although far from exhaustive, these examples highlight that cross-ploidy hybridisation involves mostly diploid and tetraploid species and introgression, when it occurs, is usually in the direction of the higher ploidy parental species. Here, the higher likelihood of unreduced gametes produced by the lower ploidy parent may allow it to form fertile lineages with the higher ploidy parent (Stebbins, 1971; Baduel et al., 2018).

Euphrasia (Orobanchaceae) is a large temperate genus of hemiparasitic plants, with around 263 species worldwide (Nickrent pers. comm.). In Britain and Ireland, there are 21 species of *Euphrasia*, which are considered a taxonomically complex group characterised by recent postglacial divergence (Wang et al., 2018), plastic phenotypes (Brown et al., 2020), and the widespread occurrence of natural hybridisation (Metherell and Rumsey, 2018). The genus in Britain and Ireland consists of five diploid and sixteen tetraploid species. This ploidy difference is associated with contrasting mating systems where the tetraploids are mixed maters, or highly selfing, while the diploids outcross more extensively (French et al., 2005). The tetraploid species are

allotetraploids, with two subgenomes that are around 5% divergent. One subgenome of the tetraploid species is only 0.2% divergent from the diploid species genome, making it likely that the British diploid species are one of the parents of British tetraploid species (Becher et al., 2020). Out of 72 hybrid *Euphrasia* combinations reported in the British flora, 13 are reported to be diploid-tetraploid hybrids based on morphology (Stace et al., 2015). Unusually, these cross ploidy hybrids are purported to be diploids derived from triploid F1s that backcross to the diploid parent (Yeo, 1956). This contrasts with most prior predictions as to the directionality of cross ploidy hybridisation, where the hybrids are tetraploids. In the proposed scenario, genetic material from the tetraploid species is expected to introgress into the diploid species through backcrossing to the diploid parent. The diploid-like subgenome of the tetraploid species is homologous to extant diploid species and therefore expected to successfully pair and recombine with diploid chromosomes.

Here, we present a genetic analysis of a contact zone between diploid *E. rostkoviana* and tetraploid *E. arctica*. Both of these species have large corollas suggestive of higher rates outcrossing (Metherell and Rumsey, 2018), and hybrids have been identified widely across Britain (Stace et al., 2015), however putative F1 triploid hybrids are yet to be found in the wild. We look for any evidence of hybridisation and introgression between ploidy levels. First, we used Sanger sequencing of a locus that shows diagnostic differences between diploids and tetraploids to see if there is evidence of hybridisation. We then used Genotyping by Sequencing (GBS) to look for evidence of introgression. We also used demographic modelling to investigate the most likely scenario of historical gene flow to explain the observed genetic structure. The results enable us to discuss the processes governing reproductive isolation in a diploid-tetraploid contact zone. We predict that as reproductive barriers in *Euphrasia* are low, and as cross ploidy hybridisation is being increasingly found in other plant groups using genomic data, that there will be cross ploidy hybrids and evidence for introgression.

4.3 Materials and methods

4.3.1 Population sampling and DNA extraction

Population sampling took place in July 2017, at the managed hay meadow at Cae Trawscoed in the National Botanical Garden Wales (lat/long: 51.8447/-4.14531) where there is a mixed population of diploid *Euphrasia rostkoviana* and tetraploid *Euphrasia arctica*. These species can be easily separated by morphology based on the presence or absence of long-stalked flexuous glandular hairs. Putative hybrids show a low density of long glandular hairs and are intermediate for a range of other traits. Both species were abundant and distributed amongst each other, with *E. rostkoviana* present amongst taller vegetation, and *E. arctica* more

dominant in shorter cropped vegetation. A total of 95 individuals were sampled from the mixed population, 45 being identified based on morphology as diploid and 50 being tetraploid. Plants were sampled evenly along transect of approximately 12m where plants were highly intermixed. Specimens were transferred into silica bags to dry until DNA extraction. DNA was extracted using the DNeasy Plant Mini kit (Qiagen), following manufacturer's protocols.

4.3.2 Sanger sequencing and sequence analysis

ITS sequencing was used as there is a known diagnostic difference between diploid and tetraploid species (Wang et al., 2018), while plastid sequencing used the most variable locus that is widely used in population genetic studies of *Euphrasia*. The sample size for the ITS sequencing was 70 individuals, while for plastid sequencing eight individuals were used. PCRs were performed in 25 μ L reactions; DNA amplification protocols and conditions for PCRs are given in Appendix 3 Tables 5 and 6. To check the quality of PCR products, the DNA was visualised on a 1% agarose gel. PCR products were then cleaned with Exo-SAP (Affometrix) using standard protocols and submitted to Edinburgh Genomics for sequencing reactions using BigDye Terminator Cycle Sequencing chemistry and Sanger Sequencing on an ABI 3730. ITS1 PCR products were sequenced in the forwards direction only (with the ITS4 primer), whilst *rpL32-trnL_{UAG}* was sequenced in both directions.

ITS1 and *trnL* spacer chromatograms were aligned and edited in Geneious (version 9.0.5). Low quality bases were trimmed at the beginning and ends of sequences; six ITS1 sequenced were excluded due to poor sequence quality. The 558bp *trnL* spacer alignment of eight sequences showed no variable sites and was therefore not analysed further. The final ITS1 alignment included 62 individuals and was 658 bp in length with 58 variable sites in total. 26 diploids and 36 tetraploids were present in the alignment. An outgroup species *E. transmorrisonensis* (diploid; from Taiwan), was added from NCBI (GenBank accession number: AY165615) to polarise the ITS1 phylogeny for visualisation but was not used in downstream analyses. We constructed a Maximum Likelihood phylogeny using IQ-TREE (version 1.5.5; Nguyen et al., 2015) using ModelFinder to find the most suitable substitution model (using the TESTNEWERGE model flag) with 1000 ultra-fast bootstraps (Hoang et al., 2018). The resulting newick file was visualised in ggtree (version 2.1.2; Yu et al., 2017). To further characterise sequence variation, this alignment was read into R version 3.6.1 (R Core Team, 2019) using the function read.dna() from the package ape (version 5.4; Paradis and Schliep, 2019) and converted to a genind object with adegenet (version 2.1.2; Jombart, 2008). We tested how much genetic variation was partitioned between ploidy levels using Analysis of Molecular Variance (AMOVA) implemented in the poppr (version 2.8.4; Kamvar et al., 2014) function poppr.amova().

4.3.3 GBS & SNP discovery

GBS library preparation was performed following the protocol in (Elshire et al., 2011). We used the enzyme ApeKI to fragment the genome and create cut sizes for adapter ligation. Samples were then pooled and cleaned before PCR amplification and sequencing. A single well was used as a water control. The pooled library was submitted to the University of Oregon for 100bp single end sequencing on the Illumina HiSeq 4000, generating 17,397,350 reads. We then used the TASSEL 5 pipeline version 2 to discover SNPs using default settings, except the k-mer length was increased to 75 (Glaubitz et al. (2014); see https://github.com/Euphrasiologist/GBS_V2_Tassel_5). Master sequence tags ($n = 1,001,272$) were aligned to a *Euphrasia arctica* reference genome (Becher et al., 2020) using BWA with default settings (version 0.7.17; Li and Durbin, 2009). All scaffolds in the reference genome greater than 1000bp were merged into a single scaffold to reduce computational time. The VCF file was filtered for variants with >50% missing data and individuals with >75% missing data using vcftools (version 0.1.16; Danecek et al., 2011).

We focused our analyses of genetic structure, hybridisation and introgression on regions that are homologous between diploids and tetraploids, and not in the other more divergent tetraploid subgenome where gene exchange with diploids is unlikely. As such, variants were excluded if they were not located on the ‘conserved’ set of scaffolds of presumed disomic inheritance that were identified in genome-wide sequence comparisons of diploid and tetraploid *Euphrasia* by Becher et al. (2020). The conserved set comprises 46 Mbp ($n = 3454$) of sequence and is likely to represent (a part of) the conserved subgenome. A minor allele frequency filter was applied to remove invariant sites using vcftools. One variant was kept per scaffold to ensure variants were not tightly linked, using PLINK (version 1.9; Purcell et al., 2007). This left 92 individuals (42 diploids and 50 tetraploids) and 270 variable sites for analyses.

4.3.4 Identifying hybridisation between ploidy levels

To investigate hybridisation between the two species of *Euphrasia* present in the contact zone, we first conducted a Principal Component Analysis (PCA) of our genomic GBS dataset. First, the VCF was loaded into R using the package vcfR (version 1.10; Knaus and Grunwald, 2017) and the PCA was carried out using the adegenet (Jombart, 2008) function dudi.pca() where missing values were substituted by the mean allele frequencies. The PCA was visualised using ggplot2 (version 3.2.1; Wickham, 2016). Second, we performed a Bayesian admixture analysis in the program STRUCTURE (version 2.3.4; Pritchard et al., 2000) using the same GBS dataset. The VCF was converted to a STRUCTURE file format using PGDSpider (version 2.1.1.5; Lischer and Excoffier, 2012). We set the K-value to be 2 as there were two divergent species of differing ploidies, and the run was

set with a burn-in of 100,000 for 1,000,000 iterations on the ‘admixture’ option. In addition, 90% probability intervals were stored using the ANCESTPINT option. Samples where probability intervals overlapped zero or one were considered non-hybrids. The Q-matrix, and probability intervals, were extracted from the output and plotted using a custom R script (<https://github.com/Euphrasiologist/StructuRe>). Third, we explicitly attempted to identify hybrid individuals using the program NEWHYBRIDS (version 1.1; Anderson and Thompson, 2002), which classifies individuals into one of six potential categories (parent A, parent B, F1, F2, backcross (BC)1 to parent A, BC1 to parent B) based on their SNP genotypes. The model was run with a burn in of 100,000 iterations and a run length of 100,000 sweeps.

4.3.5 Quantification of genetic variability within and between ploidy levels

We computed several population genetic statistics on the GBS dataset to understand population structure in the contact zone, and used AMOVA (as above) to detect regions of the genome that may have introgressed. Weir and Cockerham’s estimator of F_{ST} was calculated for each SNP in the GBS data using vcftools (Danecek et al., 2011) and visualised in ggplot2 (Wickham, 2016). The average F_{ST} across all SNPs was reported as the global F_{ST} . An AMOVA was run for both the ITS1 and GBS datasets using the function poppr.amova() from the R package poppr (Kamvar et al., 2014), which was used to understand the partitioning of genetic variability both within and between ploidy levels. P-values were then derived from the output using the randtest() function from ade4 (version 1.7-15; Bougeard and Dray, 2018), randomly permuting sample matrices 9999 times.

4.3.6 Demographic inference with $\delta a\delta I$

We carried out demographic model fitting using the package $\delta a\delta I$ to see what the best fitting model of hybridisation was for the disomic GBS data (Gutenkunst et al., 2009). Missing data was handled by scaling down the size of the joint site-frequency spectra to 24 (haploid) genomes per species. We implemented four models, each involving one ancestral population splitting into two sub-populations corresponding to diploids and tetraploids, which could differ in size. Model parameters corresponding to the diploid and tetraploid sub-populations are denoted with the subscripts D and T. The models differ by the amount of gene flow allowed between the sub-populations: (1) constant gene flow with five parameters (two population sizes NeD and NeT , two migration rates MDT and MTD , and the time of the population split $T0$), (2) historic gene flow only with six parameters (two population sizes, two migration rates, the time when gene flow ceased $T1$, and the time difference between $T1$ and the time of the split denoted $T0$), (3) secondary contact as (2) but with

gene flow in T0 and not in T1, and (4) no gene flow with three parameters (two population sizes and T0, see Figure 14 A for schematics). We fixed F at 0.75 and 0.81 for *E. arctica* and *E. rostkoviana* according to the empirical estimates. To assess the uncertainty of the model fitting, we used 99 individual sub-samplings of our data set. Because we found the model fitting results to depend strongly on the initial conditions, we ran 99 replicates with randomly perturbed starting values for each model and down-sampled data set resulting in 39,204 optimisations. From each set of 99 replicates, we selected the one with the best log likelihood. In order to compare these nested models with different numbers of parameters, we computed the Akaike Information Criterion (AIC) of each fit, and we plotted the results with matplotlib (Hunter, 2007).

4.4 Results

4.4.1 Patterns of genetic structure in the ITS1

The 658bp alignment of the ITS1 sequence data included 26 diploid individuals and 36 tetraploid individuals. Eight sequences were removed due to low sequence quality. The chromatograms revealed no evidence of sequence additivity or double peaks, which might have indicated hybrid individuals or retained duplicate copies in the polyploids. The maximum likelihood phylogenetic tree showed two distinct clades of *Euphrasia* that were highly supported (Figure 12 A). All 58 SNPs in the alignment were fixed between the ploidy groups. Accordingly, an analysis of molecular variance (AMOVA) showed that 99.5% of the ITS1 variation in the samples was explained by ploidy ($p < 0.001$; Table 5 A). Limited sampling of four individuals from each species for the *trnL* spacer showed no differences between the species.

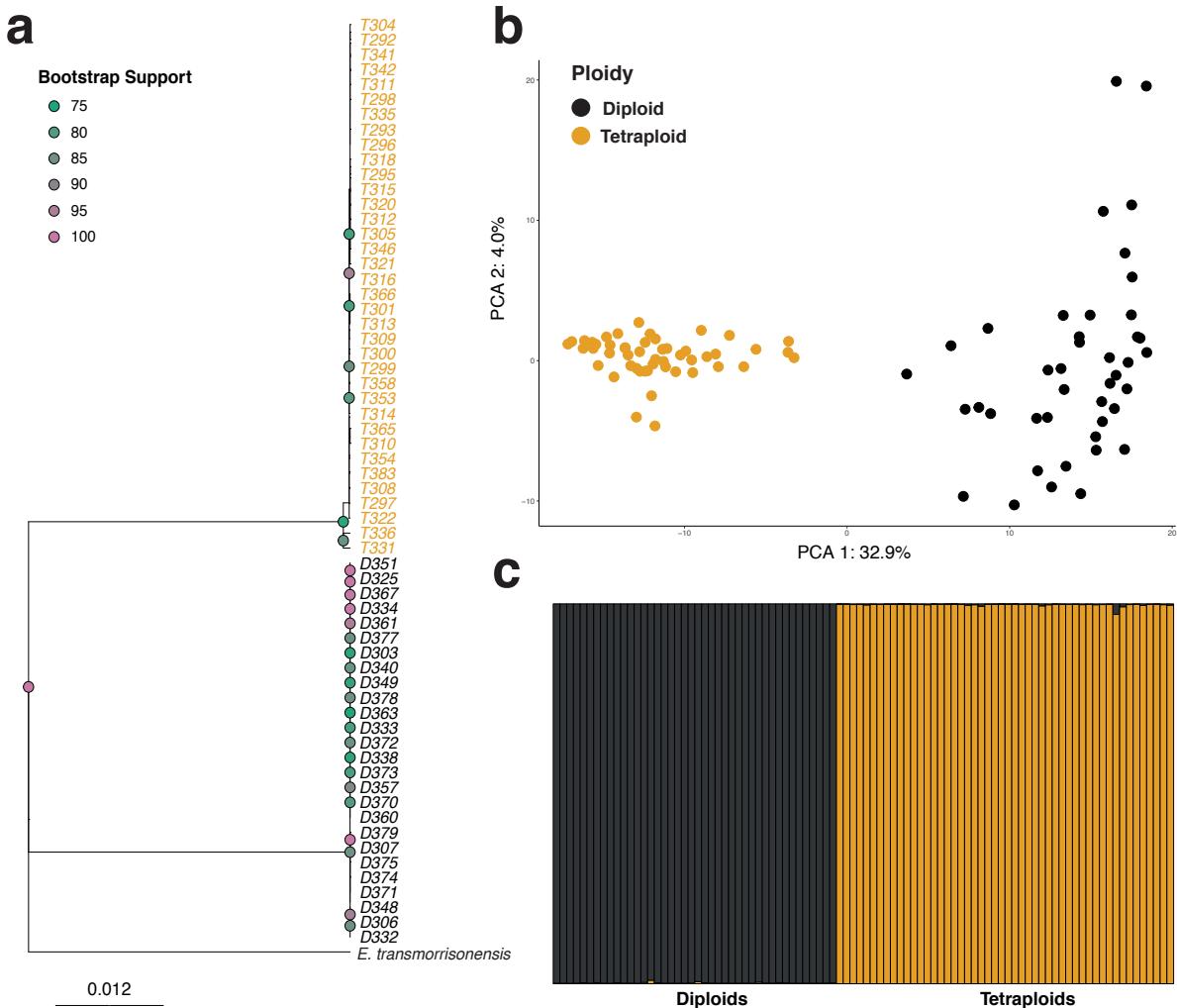


Figure 12: ITS1 marker and GBS analysis suggest strong barriers to gene flow between diploid (black) and tetraploid (orange) species of *Euphrasia*. A) Phylogenetic tree of the ITS1 marker generated from IQ-TREE; *E. transmorrisonensis* is used as an outgroup species. Bootstrap support is visualised on a colour scaled point at each node and only support values above 75 are shown. B) A Principal Component Analysis of 92 *Euphrasia* individuals across 270 loci located on different scaffolds of the reference *E. arctica* genome. The first two Principal Components are plotted against each other. C) Admixture analysis performed in STRUCTURE on the same 92 individuals, where K = 2.

4.4.2 Patterns of genetic structure from GBS data

Analysis of SNP data generated using GBS corroborated the findings from the ITS1 dataset. In total, 270 SNPs were analysed representing putatively disomically inherited scaffolds in the *Euphrasia arctica* reference

genome shared between diploids and tetraploids (see Methods). Principal Component analysis (PCA) showed that the first principal component explained 32.9% of the genomic variation and clearly separated individuals by ploidy, with two separate clusters and no evidence of intermediate genotypes (Figure 12 B). The same pattern was also present in the STRUCTURE analysis with a K value of 2, which reported Q values all above 0.98 (Figure 12 C) that assigned individuals to clusters consistent with their morphological identification. The probability interval of each Q value overlapped either zero or one, indicating no hybrid genotypes. An AMOVA on the GBS data showed 78.4% of genomic variation was explained by ploidy ($p < 0.001$; Table 5 B). The remainder of the genomic variation (21.6%) was due to differences within ploidy level.

Table 5: Hierarchical analysis of molecular variance (AMOVA) for both ITS1 and GBS data sets. The total variation is partitioned between and within ploidy level of individuals in the analysis. Degrees of freedom (df), the variance of each of the observations (Sum Sq), and percentage of variation explained by each level of variation (% Var) are reported. The significance of the components of variance are reported as p-values.

Data set	Variation	DF	Sum Sq	% Var	p-Monte Carlo
A) ITS1	Between ploidy	1	454.9	99.5	0.0001
	Within ploidy	60	4.9	0.5	
	Total	61	459.8	100	
B) GBS	Between ploidy	1	56	78.4	0.0001
	Within ploidy	90	30.3	21.6	
	Total	91	86.3	100	

4.4.3 Detecting genomic hybridisation

We calculated population genetic parameters from GBS data to further investigate potential hybridisation between the two species differing in ploidy level. The global F_{ST} between species was high at 0.44, indicating that diploid *E. rostkoviana* and tetraploid *E. arctica* were highly differentiated. The distribution of F_{ST} however showed a bimodal distribution with high count of SNPs that were either mostly shared or private (Figure 13). The NEWHYBRIDS analysis assigned each individual as 100% either parental species, with no evidence of F1 hybrids or backcrossed individuals. Altogether, the AMOVA, PCA, STRUCTURE, and NEWHYBRIDS results showed a clear signal of strong differentiation between the two species.

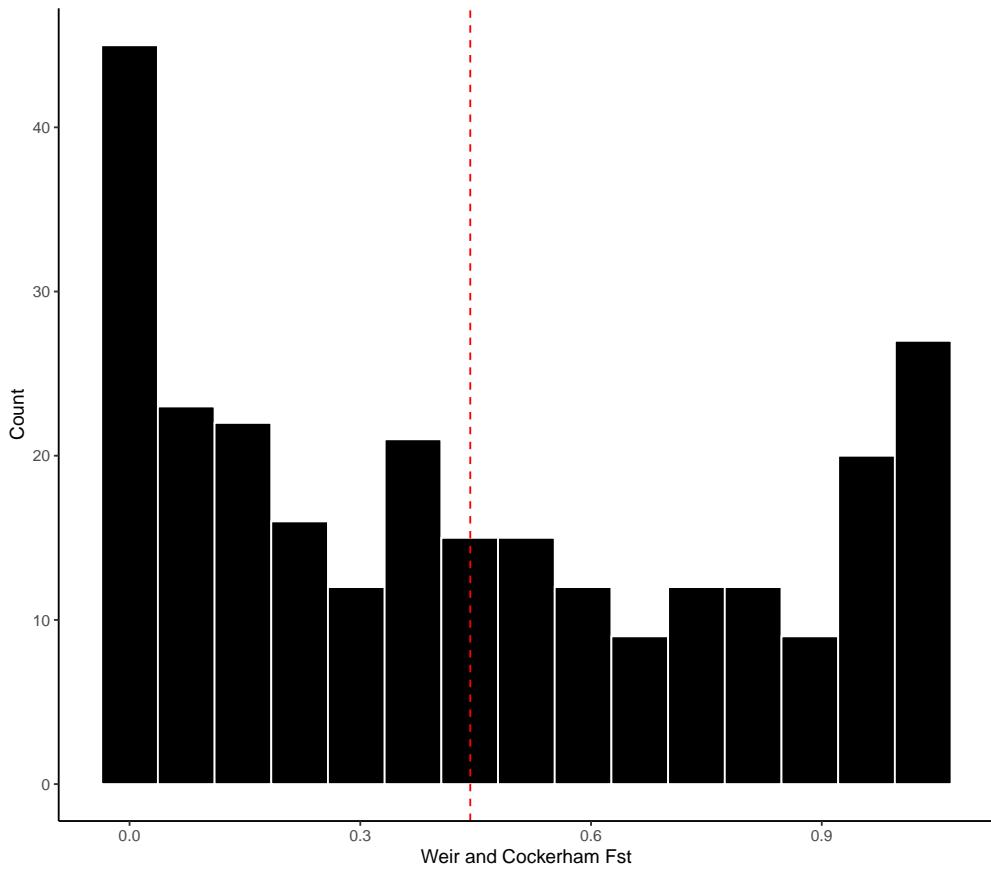


Figure 13: The distribution of F_{ST} for each of the 270 SNPs in the GBS dataset across all 92 individuals between the two ploidy levels. The red dashed vertical line indicates the mean global F_{ST} across all SNPs (0.44).

4.4.4 Demographic modelling of gene flow

Demographic model fitting with $\delta a \delta I$ resulted in the highest overall support for models with constant gene flow and with secondary contact (Figure 14 B). These models consistently scored low AIC values. The difference in AIC between the best model (constant gene flow, median AIC 425.77) and the runner-up (secondary contact, median AIC 425.89) was not significant (t-test $t = -1.38$, $p = 0.17$). The other two models had significantly higher median AICs than the best model (historic gene flow: 462.83, $t = -7.29$, $p < 0.001$; no gene flow: 452.05, $t = -52.31$, $p < 0.001$). However, for some realisations of the data re-sampling, the model with no gene flow scored the best AIC coinciding with generally low estimates of the age of the population divergence time (T_0 tends to be lower without gene flow than in alternative models, see Appendix 3 Figure 7).

The levels of gene flow fitted to our models tend to be low with M , the number of migrants per generation, of the order of 1 or less. The mean values over all replicates for the constant gene flow model were 0.3 (diploid \rightarrow tetraploid, $sd = 0.09$) and 0.4 (tetraploid \rightarrow diploid, $sd = 0.13$). For the model with secondary contact the migration rates were fitted to be 1.3 (diploid \rightarrow tetraploid, $sd = 1.02$) and 0.3 (tetraploid \rightarrow diploid, $sd = 0.09$). In all models, the effective population size tends to be slightly higher in the diploids than in the tetraploids (Appendix 3 Figure 7, left-hand panels). While our data suggest the presence of gene flow, the similar AIC values we obtained under different models show how different demographic scenarios may produce similar genetic patterns.

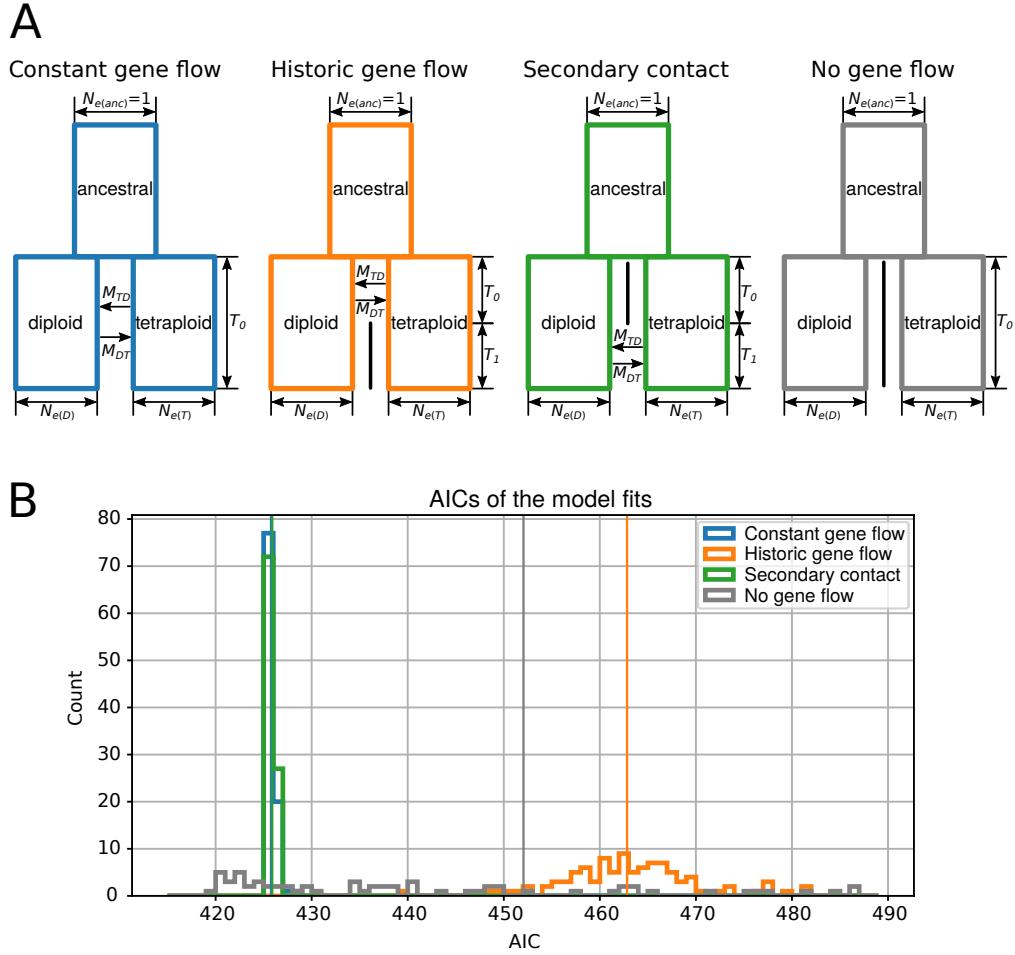


Figure 14: A) Schematics of the demographic models fitted indicating their parameters. N_e , the effective population size, is scaled to be 1 in the ancestral population. $N_{e(D)}$ and $N_{e(T)}$ indicate the ratios of the present-day effective population sizes of the diploid and tetraploid populations respectively, relative to the ancestral one. T is the number of generations to coalescence, which is subdivided in two epochs in the models with ancestral gene flow and secondary contact. M is the number of migrants per generation with subscripts indicating the direction of gene flow. B) The distributions of AIC values for model fits to 99 sub-sampled data sets. The models with constant gene flow and secondary contact are not distinguishable. Vertical lines indicate the distribution medians.

4.5 Discussion

In this study, we investigated hybridisation between diploid *E. rostkoviana* and tetraploid *E. arctica* in a contact zone in south Wales. Neither ITS1, which is known to have diagnostic differences between species, nor the GBS dataset, provide evidence for recent hybridisation or introgression between these two species of *Euphrasia*. This contrasts with the extensive hybridisation known to occur between *Euphrasia* species

of the same ploidy level (Stace et al., 2015; Metherell and Rumsey, 2018). For example, using GBS data, Zlonis and Gross (2018) showed that there is extensive gene flow from native to invasive tetraploid species of *Euphrasia* in North America. Although we expected to find cross ploidy hybrids, to date there has only been a single wild triploid hybrid *Euphrasia* individual that has been found (Yeo, 1956), and previous attempts to synthesise cross ploidy hybrids artificially have been unsuccessful (Yeo, 1968). This result is consistent with ploidy representing a strong reproductive barrier between species (Husband and Sabara, 2004). While we are confident that recent hybridisation is rare (or absent), based on the small sample of individuals we cannot exclude the possibility that early generation hybrids may have been overlooked if they are rare.

Although conventional population genetic analyses show little evidence of recent gene flow between the two *Euphrasia* species in this study, demographic modelling using $\delta a\delta I$, and inspection of the distribution of F_{ST} values across all SNPs indicates that there may be limited ongoing gene flow. $\delta a\delta I$ gives low estimates for bidirectional gene flow, which is known to occur in other cross ploidy hybrid systems (Bleeker, 2003). In particular, the higher estimate for tetraploid to diploid gene flow in the constant gene flow model may give support to the hypothesis that Yeo outlined more than 60 years ago (Yeo, 1956), and corroborates a recent study on British *Euphrasia* which suggested possible diploid-tetraploid gene flow (Becher et al., 2020). Although the signature of gene flow is low this may be because it is limited to specific regions of the genome. Here, selection may be operating differentially in the diploid and tetraploid species, which may explain the bimodal F_{ST} distribution seen in our dataset (Whitlock, 2008).

The results from this study are consistent with some cross-ploidy hybrid systems where hybridisation is very rare, or where hybrids are strongly selected against. For example, in diploid *Centaurea pseudophrygia* and tetraploid *C. jacea*, only targeted sampling was able to reveal cross ploidy hybrids which were otherwise not found by random sampling in 12 different contact zones (Koutecky et al., 2011). In a contact zone of diploid *Senecio madagascariensis* and tetraploid *S. pinnatifolius* no hybrids have been detected in the field, however genetic analysis of the seeds revealed hybrid genotypes (Prentis et al., 2007). It is possible in the *Euphrasia* contact zone that hybrid seed is being formed but either does not germinate, or hybrid seedlings do not survive to maturity. Our results stand in strong contrast to those found in *Dactylorhiza*, where triploid F1 cross ploidy hybrids, and backcrossed individuals are found frequently (De Hert et al., 2012; Balao et al., 2016). The mechanisms underlying these differences in the frequency of cross ploidy hybrids in divergent species are yet to be fully established, and are likely an area for fruitful research.

5 Horticultural protocols for experimental studies of eyebrights (*Euphrasia*, Orobanchaceae)

5.1 Introduction

Parasitic plants are a diverse group of 4,500 species that are defined by possessing a parasitic feeding organ called a haustorium that can attach and steal nutrients from a host plant (Nickrent and Musselman, 2017). Some of the most familiar parasitic plants include the crop pest witchweed (*Striga*, Orobanchaceae), mistletoes such as *Viscum album* (Santalaceae), the common grassland wildflower yellow rattle (*Rhinanthus*, Orobanchaceae), and the species possessing the largest flower in nature, *Rafflesia arnoldii* (Rafflesiaceae; Twyford, 2018). The diversity of parasitic plants, with parasitism described from 12 plant families (Westwood et al., 2010), is matched by the diversity of growing conditions necessary to succeed in cultivating these plants. Even related parasitic plant species can be found in contrasting conditions and it is important that these are matched in cultivation (e.g. the Orobanchaceae; Joel et al., 2013). There are also a number of specific horticultural issues associated with growing parasitic plants that must be overcome (e.g. germination stimulants Yoneyama et al., 2010 and host specificity @RN1557), though a wide variety of parasitic plants have been and continue to be cultivated (Pignone and Hammer, 2016). Parasitic plants are a particular challenge for cultivation as they require conditions suitable for the parasite, the host, and their interaction. Many parasites require host-specific cues in order to germinate, and in many cases the parasite must be added at an appropriate time for the parasite to establish (e.g. Yoder, 1999).

In this article, we describe our experience of optimising horticultural protocols for eyebrights (*Euphrasia*, Orobanchaceae). The genus *Euphrasia* contains approximately 263 species (Daniel Nickrent, pers. comms.) distributed throughout temperate areas of the northern and southern hemisphere, and in montane regions of tropical South East Asia (Gussarova et al., 2008). It includes both perennial and annual species (e.g. Yeo, 1973). *Euphrasia* are generalist hemiparasites, meaning they are photosynthetically competent and can grow without a host, but perform much better when grown with one of many potential hosts (grasses, forbs, legumes; Yeo, 1964). We are currently developing British *Euphrasia* as a study system for understanding plant parasitism, and for investigating evolutionary questions related to hybridisation, polyploidy, and self-fertilisation. There are 21 British native species and these show rich variation in habitat, associated species, ploidy (there are diploids and tetraploids; Yeo, 1956) and mating system (there are selfing and outcrossing species; French et al., 2005; Metherell and Rumsey, 2018). They are known to hybridise extensively in the field and produce a diverse array of local hybrids as well as stable hybrid species (Stace and Crawley,

2015; Metherell and Rumsey, 2018). Here, we focus on protocols for experimental studies under laboratory conditions, in pot trials, and under field conditions.

5.2 General considerations for cultivating *Euphrasia*

Most experimental work on plants focuses on species that are simple to grow, are small in size at maturity, and rapidly complete their life cycle, such as the thale-cress *Arabidopsis thaliana*, which is a widely used model system (Weinkoop et al., 2010). While British native *Euphrasia* are also small annuals, their cultivation is made difficult due to specific germination requirements and the need for a host. Firstly, *Euphrasia* species require a period of cold before they are able to germinate (Liebst and Schneller, 2008). In cultivation, seeds therefore need to be stratified. In our experimental work we have achieved this in two ways: leaving seeds outside over winter, and forcing germination through placing seeds in the fridge. Other than requiring stratification, *Euphrasia* species require no special methods for germination, with wild collected seeds planted without removing abortive seeds usually giving 40-50% germination success (chapters 6 and 7). The probability of germination depends on the condition of the seed. In a seed quality experiment in February 2019, we sowed 20 seeds of *E. arctica* from each of three categories (shriveled, intermediate, and plump, see Figure 15) onto moist soil in seed trays that were kept in cold frames. The seeds had been collected in September of the preceding year, kept dry, and stored at 4°C overwinter. Germination differed between categories. By mid-June none of the shriveled seeds had germinated, but six intermediate ones, and ten plump ones. It is likely that the shriveled seeds have been aborted by the parental plant, either because of developmental problems, genetic abnormalities, or limited resources (Stephenson, 1981).



Figure 15: Three categories of seed quality used to assess the germination probability: shrivelled (left), intermediate (centre), and plump (right).

Additionally, *Euphrasia* species require a host for acceptable growth and this requirement must be met within

a few days of germination to ensure growth is not hindered (Yeo, 1961). In pots, a host can be transplanted carefully with minimal mortality, otherwise *Euphrasia* can be sown with host seed at the same time. Sowing *Euphrasia* into pre-existing vegetation is possible, however *Euphrasia* is a poor competitor and the vegetation must be sufficiently low or sparse for successful establishment. *Euphrasia* seeds can be stored by routine drying (e.g. at room temperature for one week) and then refrigerating or freezing, with frozen seeds surviving for at least three years (Chapman et al., 2019).

5.3 Laboratory conditions

For detailed studies of plant development, it is necessary to grow plants under controlled laboratory conditions. *Euphrasia* seeds will readily germinate on moist filter paper under sterile conditions. Ethanol should be used to sterilise the plates and seeds, then the filter paper placed inside with the seeds on top. The lid should then be placed on top and the plate sealed to avoid contamination. Seeds should be maintained in a fridge at ~4°C until germination (no supplemental light required). We have tested two seed collections this way, an unidentified diploid species sourced from Wales by commercial supplier Emorsgate, and tetraploid *E. arctica* wild-collected from Inverkeithing in Scotland. Only the second genotype germinated, after six weeks. This method is suitable for growing young seedlings, such as those required for cytology, but further refinement is necessary to make this suitable for growing plants to a larger growth stage. For more detailed developmental studies and for genetic manipulation (such as virus-induced gene silencing) plants can be grown on sterilised plates. Seeds will germinate after ~10 days at 4°C on 1/4 Hoagland media, a widely used hydroponic nutrient solution used in other parasitic plant research (Delavault et al., 1998).

5.4 Pot trials

Growing *Euphrasia* in pots has the benefit of plants being in a soil-like substrate where they can form more natural host interactions than in the laboratory. Such conditions are preferred when studying differences between species, because less natural conditions used in the lab may trigger growth that is not normally observed. When different species are grown under common conditions, phenotypic differences between individuals and between species have to be due to genetic differences (e.g. Riihimäki and Savolainen, 2004). Reduced differentiation under common conditions relative to the wild, however, shows that the phenotypic differences observed in nature were mainly due to environmental differences. This is the basis for common garden experiments, which have been extensively used in ecology, evolution and genetics to understand ‘nature vs nurture’.

We have performed four experimental common garden studies with *Euphrasia* plants grown in pots. Our initial experiment took place in a glasshouse at the nursery at RBGE in 2016, and aimed to study the different growth patterns of five different *Euphrasia* species (and multiple hybrids), and individuals of *E. arctica* grown on eight different hosts (and without a host; see chapter 6). We found that host species impacted on the phenotype of *E. arctica* for some traits (e.g. height) but not for others (e.g. nodes to flower) and that certain *Euphrasia* species overlapped in many traits (e.g. *E. arctica/confusa/nemorosa*), while others were relatively distinct (e.g. *E. micrantha*). While successful, the most notable challenge was the relatively high mortality of individuals, and issues with high temperatures under glass on hot summer days which caused flower buds to abort. In the second and third experiments (see chapter 7), which took place over two years (2017/18), used a new, better-ventilated and climatically controlled glasshouse for the *Euphrasia* post germination. In the first year, we measured the same species, *E. arctica*, but this time on 45 different hosts, and found that survival and fitness varies greatly between *Euphrasia* on different host species. In the second year, we measured fitness across four different *Euphrasia* species on thirteen different hosts, which revealed host-parasite interactions in specific *Euphrasia*-host combinations. The fourth experiment (Becher et al., 2020) looked to investigate different morphologies of three tetraploid *Euphrasia* species collected from Fair Isle (*E. arctica*, *E. foulensis*, and *E. micrantha*; Becher et al., 2020).

The Fair Isle experiment used seeds from two populations of each of the three *Euphrasia* species, grown with twelve host species, including wild-collected heather (*Calluna vulgaris*) and juniper (*Juniperus communis*) from Fair Isle. This experiment was performed in an outside seed frame (seen in Figure 16a), which is a well-ventilated structure that protects from damage by animals or wind. While there are benefits to growing plants outside (such as less-vigorous growth more similar to that seen in the wild), these plants grew relatively poorly, due to the partial shade of the side of the seed frames. We also found the soil mix, RBGE 1, has a tendency to become waterlogged, which creates problems after heavy or extended wet periods, but can also dry out rapidly in hot weather. This was less of a problem in the glasshouse where watering can be more easily controlled.

Each of these four experiments followed a similar protocol. All experiments aimed for at least 30 pots of each population-host combination and allowed for approximately 50% germination success. A single seed is planted in the centre of the pot, with this placement helping to identify it from any contaminant weed seeds. Planting pots of 9x9x9.5 cm are filled with a loose bark-based substrate (RBGE1). After sowing plants are lightly top dressed with sieved soil, and the pots moved to a seed frame where they remained until germination (Figure 16a). Host seeds are sown into trays with RBGE1 in February and introduced after the *Euphrasia* germinates.

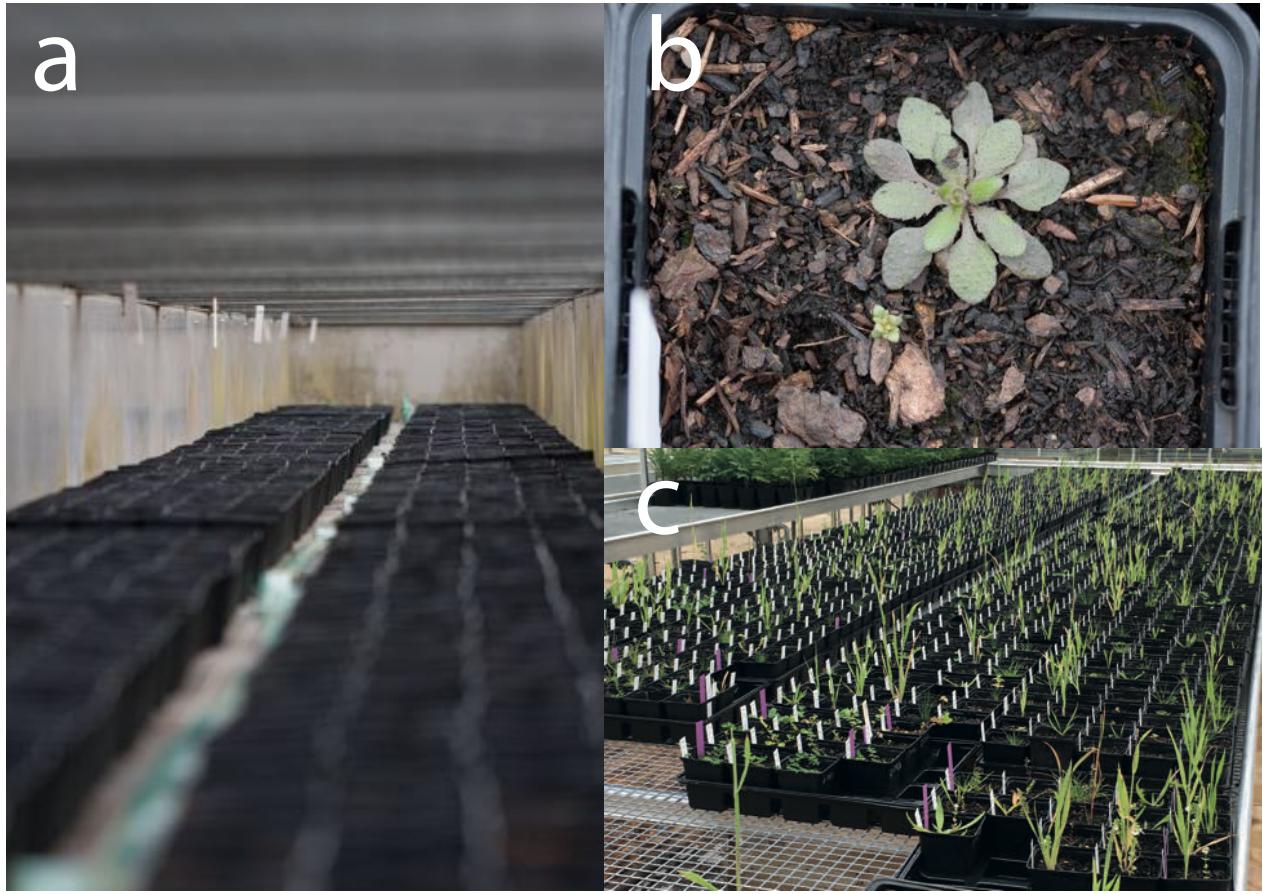


Figure 16: Pot trials in *Euphrasia* plants. a – thousands of *Euphrasia* seeds placed in pots, which are located in an outside cold frame throughout the winter (2016) to stimulate germination the following spring (2017). b – close-up of a pot with a *Euphrasia* individual (bottom left) parasitising an *Arabidopsis thaliana* individual (2017). c – a full scale growth experiment in progress with hundreds of mature *Euphrasia* individuals parasitising a multitude of host plants (2018).

Proper introduction of a host plant is critical in the establishment of a connection between *Euphrasia* and the host. We developed a protocol where the *Euphrasia* did not have to be transplanted, as early trials indicated this resulted in high mortality. Instead, we transplant young host plants (< 2 weeks post germination) into a pot containing a *Euphrasia* individual in the centre (Figure 16b). The transplanted individual is placed equidistant between the *Euphrasia* individual and one of the four corners of the pot. In common garden trials with many plants, this allows us to keep a consistent distance between the *Euphrasia* and the host plant so there is no effect of distance to host. Placing a host too close to the *Euphrasia* could lead to either very early attachment (beneficial) or high levels of competition (detrimental), and vice versa.

In April, each germinant pot received one host plant, the survival of which was generally high. Plants are

then moved to their final growing conditions (the glasshouse, or outside; Figure 16c). During the course of the summer, host plants had to be trimmed to avoid shading of neighbouring pots. Pots were randomised monthly to minimise block effects (e.g. plants on one bench growing better than others). Every day, plants were checked to measure traits on the day of first flowering, which represented a standardised time point to measure traits between individuals for use in statistical analyses.

Watering is necessary to prevent pots drying out. We have watered when required rather than as a matter of routine. We have found watering by hand to be more reliable for experimental work than automated irrigation, which can be patchy and may lead to uneven growth. *Euphrasia* favours drier over damper growing conditions, although this sometimes resulted in sub-optimal conditions for the hosts.

Supplementary feeding has proven necessary for vigorous *Euphrasia* growth in small pots. Our feeding regime begins in May before the transplantation of hosts had been completed. When *Euphrasia* flower, feeding is increased from fortnightly to weekly. Liquid feed was diluted at 1.5% by a Dosatron when watering with a fine rose. While the vigour of some of the hosts visibly improved after feeding, there seemed to be less of a direct correlation between feeding and increased *Euphrasia* vigour.

In addition to the previously mentioned trimming of the hosts, the pots were regularly weeded. Special attention must be paid at the start of the process, before *Euphrasia* seeds have apparently germinated, so that a *Euphrasia* individual does not begin to parasitise a weed. Moss and liverworts also need to be periodically removed when they threaten to smother the *Euphrasia*. Great care has to be taken to not disturb the roots of either the *Euphrasia* or host when weeding, therefore all weeding is best done with a pair of tweezers.

There are a few species of pest which attack *Euphrasia* in cultivation. The most serious pests are aphids (Aphididae), which attack the upper stem and leaves. In heavy infestations, leaves can fall off the plant, and in some cases aphid damage can be fatal. The effect of aphids can be alleviated by spraying a soapy solution on the plants to suffocate the aphids, or if *Euphrasia* individuals are kept in glasshouses, to keep air movement and ventilation. Another common pest is a species of rust (*Coleosporium*; likely alternate host of pine; Ellis and Ellis, 1985), which although an alarming orange colour, has an unknown effect on *Euphrasia*. Lastly, some caterpillar species in the Lepidopteran genus *Perizoma* attack *Euphrasia* by spinning and eating the leaves (Fitter and Peat, 1994).

As *Euphrasia* are annual plants, all individuals die at the end of the season, which may last until late September. After germination, if a *Euphrasia* individual is unable to find a host then it either remains small in stature for months before dying, or dies quickly - usually within two weeks. Host species can influence survival of *Euphrasia* dramatically. Fast growing hosts which *Euphrasia* cannot attach to compete with the

Euphrasia seedling for light, and increase the probability of *Euphrasia* mortality. In general, leguminous or grass host species confer higher probabilities of survival for *Euphrasia* than forbs or woody plants.

5.5 Field trials

Yellow rattle (*Rhinanthus minor*) has widely been exploited in meadows for its properties of parasitising and reducing the vigour of surrounding plants, thus reducing the need for mowing and for maintenance (Ameloot et al., 2005a; Westbury and Dunnett, 2007). *Euphrasia*, as a related hemiparasite, could be used for similar purposes, with the wide-range of habitats of different *Euphrasia* species making it potentially useful in habitats where *Rhinanthus* does not survive. To produce seeds for large-scale planting requires *Euphrasia* to be cultivated on a field scale. In a collaboration with Scotia Seeds (<http://www.scotiaseeds.co.uk/>) we set up some field plots with a view to understand how feasible it would be to cultivate *Euphrasia* at scale. In late October 2018, a 200m² plot was cleared to allow the sowing of host and *Euphrasia* species, with 24 2m² plots used for planting. Two different populations of *Euphrasia arctica* were trialled, one from North Berwick collected in 2016 and one from Inverkeithing collected in 2017. Sowing densities for *Euphrasia* were at 500 seeds per square metre (0.625g) and host densities were around 2g per metre square. We used four different host treatments in combination with *Euphrasia*: *Lotus corniculatus*, *Cynosurus cristatus*, *Plantago lanceolata* and Mavisbank Meadow Mix (see details of mix on website: (<https://www.scotiaseeds.co.uk/shop/mavisbank-mix/>)). In early September 2019 the mature plants were hand collected from each plot. Ten random individuals from each plot were sampled in order to determine vigour of the *Euphrasia* plants in each plot.

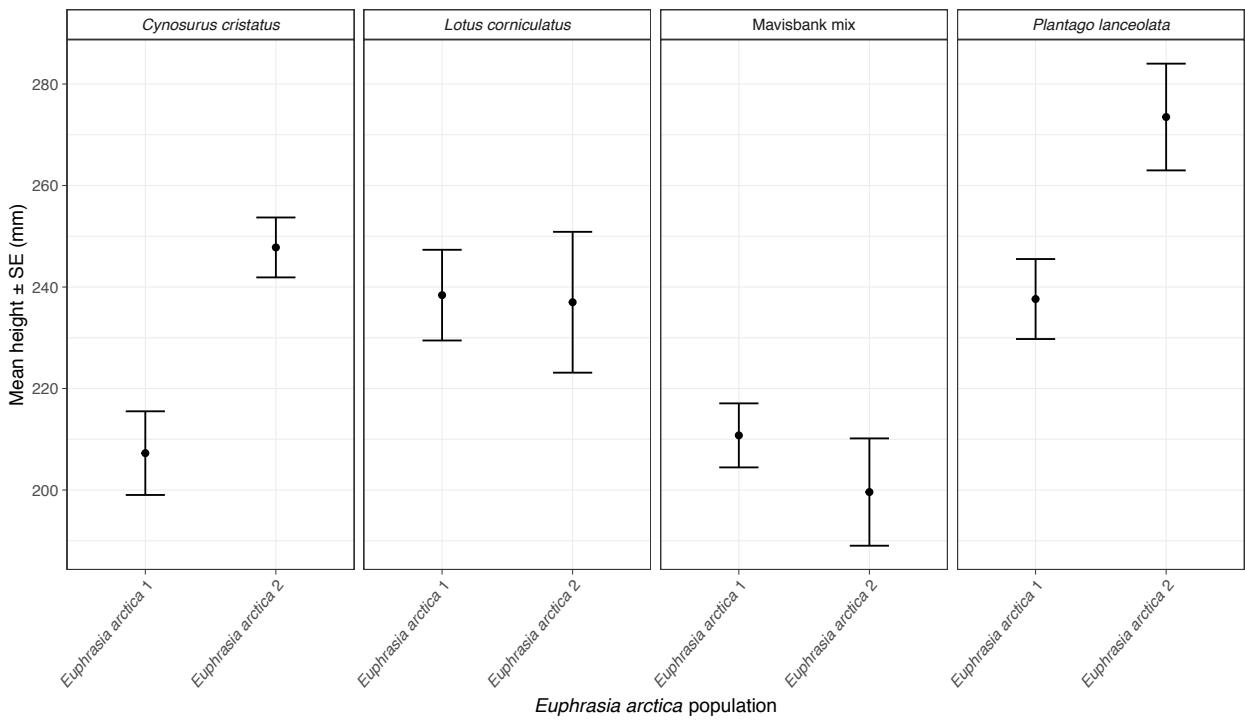


Figure 17: Mean heights and standard errors of both *Euphrasia arctica* populations from plots with each dominant host. Measurements were pooled from three replicate plots for *E. arctica* genotype 1 and one plot for *E. arctica* genotype 2.

Preliminary results indicate that cultivation of *Euphrasia arctica* on a field scale yields vigorous plants in excess of 20cm in height (Figure 17). From experimental data from pot trials, which used the same species of *Euphrasia*, plants around 20cm tall produced ~150 reproductive nodes and therefore around 1800 seeds (extrapolated from models in chapter 7, and assuming six seeds per capsule). For harvesting efficiency it is recommended that either *L. corniculatus* or *P. lanceolata* is used as grass species cause *Euphrasia* to form flexuous stem bases, which are difficult to harvest. *Euphrasia arctica* therefore could be managed on a field scale and give suitable yields for Scotia Seeds.

5.6 Reciprocal transplant experiments

While the protocols outlined above emulate aspects of natural conditions, planting *Euphrasia* in the wild is the ‘gold standard’ for experimental studies as it is the only way to include all biotic and abiotic stressors found in nature. Reciprocal transplant experiments involve translocating individuals between natural environments, and can be used to distinguish between genetic and environmental effects. If after the transplant, each

transplanted sample shows the same traits as at their ‘home’ site, it can be assumed that the trait differences have a genetic base. If on the contrary, the growing site determines some trait of interest, then the plants react plastically to their environment, and their genes do not determine the trait in question. It may well be that the result of a reciprocal transplant experiment is intermediate – many traits are affected by both genetic and environmental factors. Reciprocal transplants are particularly useful for investigating local adaptation, such as whether plants produce more seeds in their home site.

Conducting a transplant with hemiparasitic *Euphrasia* bears unique challenges. It is well-established now that the performance of *Euphrasia* plants depends critically on whether there are host plants available, which species they are, and on the timing of haustorium attachment. This means that, unlike in other large transplant experiments, clearing the local vegetation is not an option. We have therefore tested two approaches to transplants in *Euphrasia*.

To test the suitability of different substrates and planting containers, a preliminary field trial was carried out near Inverkeithing, Scotland, in 2018. We buried in the ground compostable plant pots (filled with local soil) and inflated “Jiffy” peat pellets (Figure 4). In each, we placed one *Euphrasia* seed. On our next visit, approximately four months later, we found that the compostable pots had persisted virtually unchanged and the outer mesh of the “Jiffy” pellets had not decomposed. Some pellets had lost their mesh, which we found in the nearby vegetation. While the surrounding vegetation was in good condition, our planting containers appeared dry and we recorded hardly any germination. 2018 was a very dry year, which may have affected the decomposition process, but the materials tested did not seem appropriate to be used in the wild in larger experiments.



Figure 18: A small *Euphrasia* plant germinated inside a “Jiffy” pellet. After four months in the ground, the pellet’s outer mesh remained virtually unchanged. Photo: Hannes Becher.

On our second attempt, this time in field sites on Fair Isle, we decided to make planting holes and to fill them with sterile substrate. We used commercial compost (John Innes). Small planting holes were filled with the sterile compost in order to exclude local plants. These holes are small enough for surrounding plants to root through them in a short time, but they reduced the possibility of a local *Euphrasia* seeding directly onto the planting spot. Planting was carried out in mid-September. Sites were chosen in three habitat types: heathland, coastal turf, and grassland. Woody vegetation, chiefly *Calluna vulgaris* and *Empetrum nigrum*, was cut back to allow access to the ground. Planting holes were dug with a bulb planter with the planting holes arranged in grids for ease of recoverability. Each hole was filled with compost to soil level. A single *Euphrasia* seed was then placed in the centre of each circular compost patch and a light top dressing of composts was applied. Each planting hole was individually labelled with a pencil-marked plastic tag. The

label positions were recorded so that finding one label allowed to locate all other planting holes in a grid. Each label was sunk into the ground completely. While in many respects this type of experiment under wild conditions provides the most natural settings, it comes with extensive challenges. The ease of digging into substrate depended on the location; heather bog was especially challenging to produce neat holes to sow *Euphrasia* into. Sowing is also very difficult in windy conditions, even with the help of tweezers. Finding the holes months later required good mapped locations of each grid, as some were remarkably well hidden. Some birds also removed the labels and sheep were liable to trample on the holes made, both of which resulted in missing data. One of the biggest challenges was the low percent germination, which although observed in common garden experiments, is even lower in the wild. Despite these challenges, it is still possible to carry out these kinds of experiments in natural settings.

5.7 Conclusions and future directions

The genus *Euphrasia* represents an excellent study system to investigate a wide variety of topics, from the evolution of parasitism, to the role of polyploidy, hybridisation, and mating system in taxonomic complexity. *Euphrasia* can easily be brought into cultivation in petri dishes, and in pots, both in the glasshouse and outside. In addition, *Euphrasia* can be grown on many different species of host plant, making it an ideal system to investigate host effects on *Euphrasia* and vice versa. Field trials have been met with success on a commercial scale, however reciprocal transplant experiments in the wild are difficult and require many thousands of replicates for statistically robust inference.

There are many possibilities for future research building on these protocols. The laboratory protocols for fine scale developmental, genetic and host-parasite interface work should be developed to understand the nature of haustoria formation in *Euphrasia* on different host plants (as has been done in *Rhinanthus* Rümer et al., 2007). Further, transcriptome sequencing of haustoria may reveal the genes underlying parasitism in *Euphrasia* (Yang et al., 2015). Investigation of below ground host-*Euphrasia* interactions in pots will open up an area of research into the number of haustorial connections made to host plants, and may reveal host preferences. Further large scale reciprocal transplant experiments should be made to understand host preferences and the extent of local adaptation in the genus.

6 Life history evolution, species differences, and phenotypic plasticity in hemiparasitic *Euphrasia*

This manuscript has been published as:

Brown, M., Frachon, N., Wong, E. L. Y., Metherell, C. and Twyford, D. (2020) Life history evolution, species differences, and phenotypic plasticity in hemiparasitic eyebrights (*Euphrasia*) *American Journal of Botany* 107(3), 1–10

6.1 Abstract

Species delimitation in parasitic organisms is challenging because traits used to identify species are often plastic and vary depending on the host. Here, we use species from a recent radiation of generalist hemiparasitic *Euphrasia* to investigate trait variation and trait plasticity. We tested whether *Euphrasia* species show reliable trait differences, investigated whether these differences correspond to life history trade-offs between growth and reproduction, and quantified plasticity in response to host species. We used common garden experiments to evaluate trait differences between 11 *Euphrasia* taxa grown on a common host, document phenotypic plasticity when a single *Euphrasia* species is grown on eight different hosts, and relate observations to trait differences recorded in the wild. *Euphrasia* exhibited variation in life history strategies; some individuals transitioned rapidly to flowering at the expense of early season growth, while others invested in vegetative growth and delayed flowering. Life history differences were present between some species, though many related taxa lacked clear trait differences. Species differences were further blurred by phenotypic plasticity — many traits were plastic and changed with host type or between environments. Trait differences present between some species and populations demonstrate the rapid evolution of distinct life history strategies in response to local ecological conditions.

6.2 Introduction

Parasitism is a ubiquitous feature of the natural world, with parasitic organisms found living in every ecosystem and exploiting all free-living organisms (Price, 1980; Windsor, 1998). Parasitic plants are a group of ca. 4500 species of 12 separate evolutionary origins that have evolved a modified feeding organ, the haustorium, which allows them to attach to a host plant and extract nutrients and other compounds (Westwood et al.,

2010; Nickrent and Musselman, 2017; Twyford, 2018). Parasitic plants are morphologically diverse and present a broad range of life history strategies and host interactions (Schneeweiss, 2007; Tesitel et al., 2010). Hemiparasitic plants, i.e., those that are parasitic but also photosynthesise, can often attach to a broad range of hosts; the well-studied grassland parasite *Rhinanthus* has been found to attach to over 50 co-occurring grass and herbaceous species (Cameron et al., 2006). All hemiparasitic plants are exoparasites; leaves, stems, roots, and flowers grow outside the host, and only the haustorium invades and grows within the host (Twyford, 2017).

Research to date has largely focused on three aspects of life history variation in parasitic plants. First, a body of work has looked to understand variation for specific traits between populations and related species. For example, work on natural populations of the hemiparasite *Pedicularis* has shown how investment in male reproductive organs depends on extrinsic environmental conditions (Guo et al., 2010a), while seed mass is primarily determined by intrinsic factors such as plant size rather than extrinsic factors such as elevation (Guo et al., 2010b). Second, researchers have investigated how parasite life history traits are affected by interactions with their host. In the widespread and weedy obligate holoparasite *Phelipanche ramosa*, the duration of the life cycle differs depending on the host (Gibot-Leclerc et al., 2013), with a more rapid life cycle on local rather than non-local hosts. In hemiparasitic *Rhinanthus minor*, biomass depends on the host species and the number of haustorial connections (Rowntree et al., 2014). Finally, a number of studies have looked at life history variation between species studied in a phylogenetic context (Schneeweiss, 2007; Těšitel et al., 2010). For example, broad-scale analyses of the Rhinantheae clade in the Orobanchaceae has shown a shift from a perennial ancestor to annuality, with correlated shifts to a reduced seed size (Tesitel et al., 2010). Despite the diversity of this research, there are still considerable gaps in our knowledge as to how life history trait variation is maintained (e.g., how common are trade-offs between life history traits), how much of this variation is genetic and how much is plastic, and which traits are the targets of natural selection.

In this study, we explore trait variation in generalist hemiparasitic eyebrights (*Euphrasia*, Orobanchaceae). *Euphrasia* is one of the largest genera of parasitic plants and is characterized by recent transoceanic dispersal and rapid species radiations (Gussarova et al., 2008). In the United Kingdom, there are 21 *Euphrasia* species, which are mostly indistinguishable at DNA barcoding loci (Wang et al., 2018), show complex morphological variation (Yeo, 1968; Metherell and Rumsey, 2018), and readily hybridise (Liebst, 2008; Stace et al., 2015). Despite shallow species differences due to postglacial divergence, *Euphrasia* species demonstrate substantial ecological divergence, with many taxa restricted to specific habitats such as coastal turf, mountain scree, heathland, or open grassland. Habitat differences would be expected to exert strong selection on life history traits, and this may include selection on growth to match seasonal water availability and to exploit local hosts,

or selection on flowering time in response to local competition from surrounding plants, or in response to mowing or grazing (Hellstrom et al., 2004).

Our research builds on a large body of experimental work, with *Euphrasia* used in common garden studies for over 125 years (Koch, 1959). The first experimental work on *Euphrasia* revealed that phenotypic differences between two related species, *E. rostkoviana* and *E. montana*, are maintained in a common garden environment (Wettstein, 1895). Experimental work in the 1960s showed the growth of various *Euphrasia* species differs depending on the host species (Wilkins, 1963; Yeo, 1964). More recent experiments using large sample sizes in common gardens (Matthies, 1998; Zopfi, 1998; Lammi et al., 1999: @RN947) or in experimental field sites (Seel and Press, 1994; Hellstrom et al., 2004) have shown the effect of commonly encountered hosts such as grasses and legumes on hemiparasite biomass, mineral accumulation, plant architecture and reproductive output. Despite this extensive experimental work, studies in *Euphrasia* have yet to compare life history strategies of different species and the extent of phenotypic plasticity in life history traits. This work is critical for improving our knowledge of hemiparasite evolution and for understanding the nature of species differences in a taxonomically complex group. It is also unclear whether *Euphrasia* are restricted to growing on hosts such as grasses and herbaceous species or can parasitize a broad range of taxa including novel hosts rarely encountered in the wild. To address these questions requires simultaneously investigating the growth of multiple *Euphrasia* species and multiple host species with sufficient replication to enable suitable statistical comparisons.

Here, we used a series of common garden experiments, in conjunction with field observations, to understand life history trait evolution, species differences, and phenotypic plasticity in hemiparasitic *Euphrasia*. Our first experiment assessed the morphological distinctiveness among several *Euphrasia* species and their hybrids when grown on a single host species in standardised common garden conditions. This experiment also addressed whether there is life history trait divergence among recently diverged hemiparasite species and whether these trait differences correspond to life history trade-offs. We then inspected the plasticity of a single focal *Euphrasia* population grown on many different hosts. This experiment quantified the magnitude of trait change when *Euphrasia* are grown on different hosts. It also tested whether they are truly generalist parasites by observing their growth on a wide range of hosts and without a host. Finally, we related our trait observations in a common garden to records of herbarium specimens collected in the wild. This comparison will help us understand whether life history traits and species' morphological differences are consistent between the common garden and the wild. Overall, our joint observations of phenotypic variation between closely related taxa and the extent of host-induced plasticity within a species, in an experiment and in the wild, provide new insights into variation in life history strategies in these hemiparasitic plants.

6.3 Materials and Methods

6.3.1 Experimental design and plant cultivation

We performed two common garden experiments to investigate phenotypic variation in *Euphrasia*. Both common garden experiments took place in parallel in 2016. The experiments used wild-collected, open-pollinated *Euphrasia* seeds that were pooled across individuals in a population. Seeds were contributed by plant recorders as part of the Eye for Eyebrights (E4E) public engagement project and as such included a scattered geographic sample across Great Britain (Appendix 4 Table 7). All *Euphrasia* species were identified from the herbarium specimens of field collections, and from living material grown in the glasshouse, by *Euphrasia* referee Chris Metherell. Host seeds were sourced from commercial suppliers and from field collections (Appendix 4 Table 8).

6.3.2 Species differences experiment

We observed trait differences of 24 populations from five *Euphrasia* species and six natural *Euphrasia* hybrids when grown on clover (*Trifolium repens*). This experiment included sampling multiple populations of three widespread and closely related grassland species, *E. arctica*, *E. confusa*, and *E. nemorosa*, and sparse population sampling of the moorland specialist *E. micrantha* (one population) and calcareous grassland specialist *E. pseudokernerii* (two populations). We chose clover as a host because it usually supports vigorous hemiparasitic growth and confers high survival (Zopfi, 1998).

6.3.3 Phenotypic plasticity experiment

We measured traits of a focal *Euphrasia* taxon, *E. arctica*, when grown with eight potential hosts (*Arabidopsis thaliana*, *Equisetum arvense*, *Festuca rubra*, *Holcus lanatus*, *Marchantia polymorpha*, *Pinus sylvestris*, *Plantago lanceolata*, and *Trifolium repens*) and without a host. These hosts were chosen to include a broad representation of functional groups and phylogenetic diversity, with species encountered in the wild and with novel hosts (full details in Appendix 4 Table 8). The novel hosts were included to see the limits to which parasitic *Euphrasia* can associate, namely with a tree (*Pinus*), a pteridophyte that produces adventitious roots (*Equisetum*), and a liverwort that produces rhizoids (*Marchantia*).

6.3.4 Cultivation protocol

Reliable cultivation of *Euphrasia* can be challenging due to low seed germination, variation in time to establishment, the requirement of seed stratification, and high seedling mortality when transplanted (Yeo, 1961; Zopfi, 1998). We developed cultivation protocols that combine winter germination cues that improve germination and mimic nature, but also used highly standardised and replicated pot conditions that avoid transplanting *Euphrasia* and thus maximise survivorship. We filled 9-cm plastic pots with Melcourt Sylvamix Special growing medium (Tetbury, Gloucestershire, UK) in December, placed one *Euphrasia* seed per pot, and left pots outside over winter at the Royal Botanic Garden Edinburgh (RBGE) for seeds to experience natural seed stratification. Hosts were planted in seed trays in April. *Euphrasia* plants were moved to an unheated and well-ventilated greenhouse in the spring once the cotyledons were fully expanded, and a single seedling from each host (or a 1-cm² clump of *Marchantia*) was transplanted into the pot containing *Euphrasia*. Hosts that died within 10 days of planting were replaced. Twenty or more replicates were grown for each host-parasite combination. Plants were subsequently grown to flowering with regular watering, the locations of pots randomised at weekly intervals, and foreign weed seedlings removed.

6.3.5 Common garden trait measurements and statistical analyses

We measured seven morphological traits at first flowering related to life history variation, indicators of plant vigor, or characters used in taxonomy. In addition to date of first flowering, we recorded corolla length, the ratio of cauline leaf length to internode length below the measured leaf (“internode ratio”), number of leaf teeth on the lower floral leaf (bract), number of nodes to flower, number of branches, and plant height. All lengths were measured to the nearest millimeter as done by Metherell and Rumsey (2018). For the phenotypic plasticity experiment, we also recorded early season growth (height 6 weeks after transplantation of potential host) and height at the end of season after senescence. We did not directly observe host attachment, as preliminary investigations revealed a fine root structure where haustoria were difficult to observe. Instead, we inferred that attachment is likely to have taken place based on observations of height according to Yeo (1964). In his study, *Euphrasia* that attached to a “good” host tended to grow tall with elongated internodes, while *Euphrasia* that did not attach or attached to a “bad” host were much smaller (see discussion for more details).

We used a combination of fixed effect and mixed models to gain insights into the differences in means and the magnitude of variability in our data. In all models, response variables were analysed as either Gaussian (and log-transformed if necessary) or Poisson. If the response variable was analyzed as Poisson, the model was checked for overdispersion and if it was overdispersed, an observation-level random effect was

fitted. All correlations between variables were Pearson's correlations. Multiple correlation comparisons were corrected using Holm's correction method. Phenotypic clustering was inspected using principal component analysis (PCA). All analyses were done in R version 3.4.3, with the packages lme4 (Bates et al., 2015) and MCMCglmm (Hadfield, 2010) for generalised linear mixed effects models, base R for linear models, RcmdrMisc for correlations (Fox, 2020) and ggplot2 for data visualisation (Wickham, 2016). MCMCglmm models were run for a minimum of 70,000 iterations using either inverse Wishart or parameter-expanded priors with a minimum burn-in period of 30,000 iterations. Model convergence was assessed visually by plotting the posterior distributions and Markov chains.

In the species differences experiment, species of *Euphrasia* was fitted as a fixed effect, and population of *Euphrasia* was treated as a random effect. We excluded hybrids from these analyses because we were interested in testing differences between species. In the case of height and cauline to internode ratio, the traits were log transformed. Likelihood ratio tests calculated the overall significance of species, where this was not possible, deviance information criteria were used to test better model fit. We calculated proportion variance explained by population of *Euphrasia* (after accounting for fixed effects) by dividing the population random effect variance by the total variance in the model. Tukey post hoc tests were performed on each pairwise comparison of *Euphrasia* species and adjusted p-values calculated, using Tukey honestly significant difference (HSD) test and correcting for family-wise error rate in the emmeans R package (Lenth, 2020). For the phenotypic plasticity experiment, host species was fitted as a fixed effect. The models were re-levelled so that "no host" was the baseline. Analysis of variance was used to determine overall significance of host species. Tukey post hoc tests were then performed on each pairwise comparison of host species, with adjusted p-values calculated in base R and the multcomp package (Hothorn et al., 2008).

6.3.6 Trait variation in the wild

We tested how phenotypes in the experiments related to those in nature by comparing results from the species differences experiment to phenotypic measurements of herbarium specimens of the same population sampled in the wild. Three individuals were measured from each collection sheet for a given population for each trait. Pressed plants submitted by collectors varied in quality, and therefore, we were unable to measure the height of these plants, nor was it possible to infer date of first flowering. We analysed the data using generalised linear mixed effect models with where individuals were grown (i.e., common garden or wild-collected) as a fixed effect, with each of five traits as the response variable. We treated species and population of *Euphrasia* as random effects to understand the relative contributions of each to the overall variability in a given trait.

Response variables that were considered count data were analysed with a Poisson distribution, in all other cases a Gaussian distribution was used. R-values were calculated using Pearson's correlations of the population level means between the common garden and the wild samples.

6.4 Results

6.4.1 Species differences

Our species differences experiment revealed extensive morphological trait variation across *Euphrasia* species when compared at first flowering. From the 222 *Euphrasia* individuals that survived to flower on their clover host, the greatest variation was seen in number of branches (9-fold difference between species), internode ratio (2.7-fold) and height (2.5-fold), while traits such as node to flower (1.6-fold) and corolla length (1.6-fold) proved less variable (Figure 19 A–D; Appendix 4 Table 9). A large degree of this variation was separated by species and by population (Table 6). The species with the most distinct life history strategy was *E. micrantha*, which flowered from a low node on the plant (8.3 ± 0.2 nodes) while it was short (70 ± 8 mm; Appendix 4 Table 9). It also formed a partly distinct cluster in the PCA (Appendix 4 Figure 8). *Euphrasia pseudokerneri* was relatively distinct, flowered once it had grown tall (176 ± 16 mm) and from a high node on the plant (13.2 ± 0.4 nodes), but showed little separation in the PCA. The morphologically similar *E. arctica*, *E. confusa*, and *E. nemorosa* differed for some traits, with *E. nemorosa* initiating flowering 14 days later and from 3.3 nodes higher than *E. arctica*, but overlapped in many other traits and in overall multi-trait phenotype (Appendix 4 Figure 8 and Appendix 4 Table 10). Despite species being a significant factor in the models, and some notable differences in specific traits, there were few significant pairwise Tukey comparisons due to substantial within-species variation (Appendix 4 Figure 9). Of the seven significant pairwise trait differences, three were for node to flower and three for number of leaf teeth, with four of the seven significant comparisons involving *E. micrantha*. In most cases, hybrids combined morphological characters of their parental progenitors. For example, hybrids involving *E. nemorosa* flowered later in the season and initiated flowering from a higher node than *E. arctica* 6 hybrids (Figure 19 A–D).

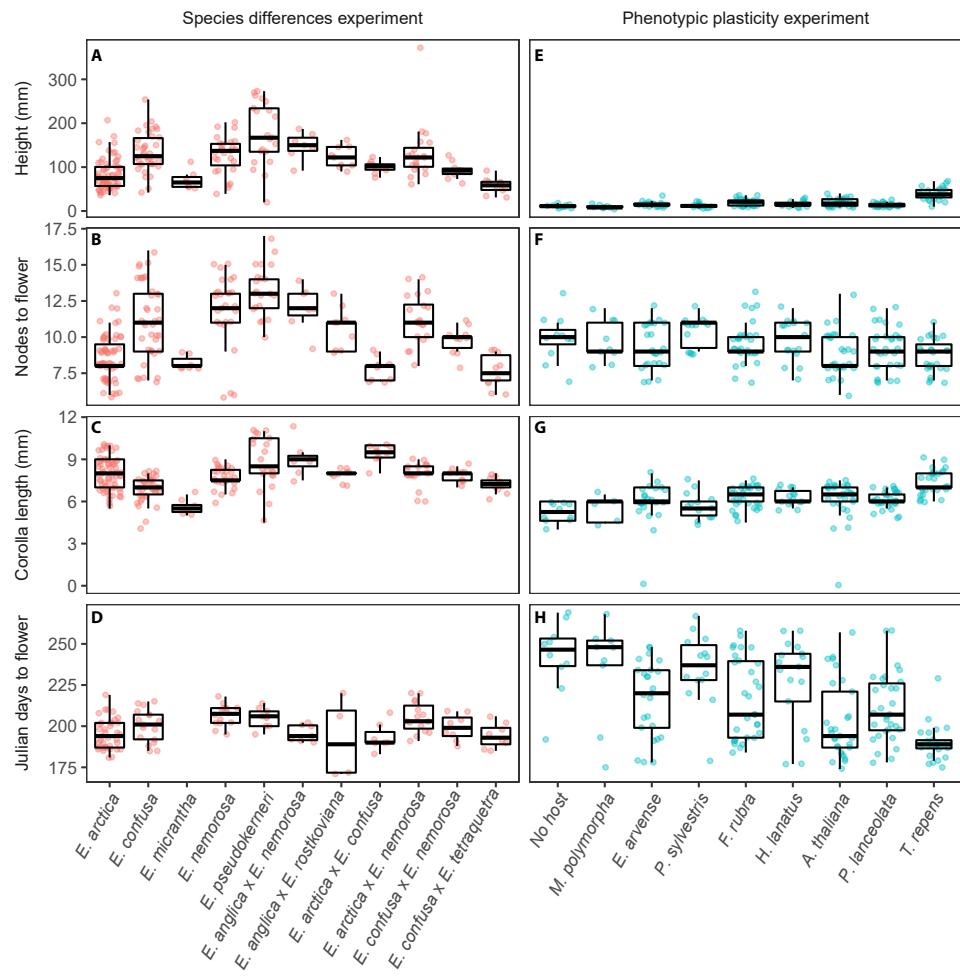


Figure 19: Trait variation in a common garden experiment of diverse *Euphrasia* species and hybrids grown on clover (A–D); *Euphrasia arctica* grown on many different hosts (E–H). The edges of the box plots show the first and third quartiles, the solid lines the median, the whiskers the highest and lowest values within 1.5-fold of the inter-quartile range, and the jittered dots each individual measurement. Length measurements were recorded in millimeters.

Table 6: Summary of generalized linear mixed effects models for *Euphrasia* trait values measured in a common garden environment. Model outputs are summarized for five *Euphrasia* species grown with clover in the species differences experiment and for *E. arctica* grown with eight hosts and without a host in the phenotypic plasticity experiment. For the phenotypic plasticity experiment, we report model outputs with all potential hosts, as well as models excluding *Pinus* and *Marchantia* when there was no evidence of attachment or interactions (reported in square brackets). The percentage variance explained by random effects are reported in parentheses along with the 95% credibility interval. ***p < 0.001, **p < 0.01, *p < 0.05.

Trait	Species differences		
	Species	Population	Phenotypic plasticity
Branches	DIC _{full} = 676.08; DIC _{spp} = 679.48	DIC _{full} = 676.08; DIC _{pop} = 714.87 (25.9%, 13.4–57.4%)	NA
Corolla length	$\chi^2(4) = 11.91^*$	$\chi^2(1) = 41.38^{***}$ (54.1%, 24.4–69.2%)	$(F_{8,173} = 9.85)^{***}$ [($F_{6,157} = 11.38$)***]
Height	$\chi^2(4) = 11.67^*$	$\chi^2(1) = 57.13^{***}$ (61.2%, 35.0–79.7%)	$(F_{8,185} = 23.14)^{***}$ [($F_{6,164} = 24.39$)***]
Internode ratio	$\chi^2(4) = 13.00^*$	$\chi^2(1) = 34.38^{***}$	$(F_{8,184} = 3.36)^{**}$ [($F_{6,163} = 4.11$)***]
Julian days to flower	$\chi^2(3) = 2.26$	$\chi^2(1) = 1.42 E-14$ (58.7%, 28.6–80.9%)	$\chi^2(8) = 192.39^{***}$ [$\chi^2(6) = 141.67^{***}$]
Node to flower	$\chi^2(4) = 15.42^{**}$	$\chi^2(1) = 2.87$ (14.1%, 1.0–33.5%)	$\chi^2(8) = 5.02$ [$\chi^2(6) = 3.04$]
Number of leaf teeth	$\chi^2(3) = 12.45^{**}$	$\chi^2(1) = 0.0059$ (0.12%, 2.8%–23%)	$\chi^2(8) = 26.79^{***}$ [$\chi^2(6) = 17.04^{**}$]

Correlation analyses across species revealed clear suites of traits that are related. Significant correlations were found between 12 of the 21 pairwise comparisons after correcting for multiple tests, with five of these correlations with R > 0.6 (Table 7 A). Plants flowering at a late node are more likely to be tall, more highly branched, and have many teeth on the lower floral leaf. The relationship of traits is also supported in the PCA, with many traits contributing to multiple principal components (Appendix 4 Table 10). Traits related to height and flowering node were largely uncorrelated with internode ratio and corolla length.

Table 7: Pearson's correlation coefficients for seven phenotypic traits measured in a common garden experiment for (a) five *Euphrasia* species and six hybrids, (b) *Euphrasia arctica* grown with eight hosts and without a host. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Asymptotic p-values values are reported from the Hmisc package in R using the rcorr() function.

(A) Species differences						
	experiment					
Trait	Corolla length (mm)	Height (mm)	Internode ratio	Julian days to flower	Leaf teeth	Nodes to flower
Branches	0.260**	0.609***	-0.116	0.057	0.658***	0.775***
Corolla length (mm)		0.319***	-0.161	-0.127	0.197*	0.049
Height (mm)			0.246**	0.292*	0.563***	0.628***
Internode ratio				0.204	-0.120	0.076
Julian days to flower					0.053	0.249*
Leaf teeth						0.651***

(B) Phenotypic plasticity						
	experiment					
Trait	Corolla length (mm)	Height (mm)	Internode ratio	Julian days to flower	Leaf teeth	Nodes to flower
Branches	0.524***	0.834***	-0.299***	-0.572***	0.694***	-0.572**
Corolla length (mm)		0.503***	0.098	-0.406***	0.536***	-0.166
Height (mm)			0.477***	-0.481***	0.692***	-0.186
Internode ratio				-0.034	0.168	-0.009
Julian days to flower					-0.691***	0.530***
Leaf teeth						-0.239**

6.4.2 Phenotypic plasticity

Our phenotypic plasticity experiment showed substantial morphological variation across 194 *E. arctica* plants grown with eight different potential host species and the 22 plants grown without a host. Plants growing on clover transitioned to flower quickly (189.8 ± 2.0 Julian days), grew tall by the time of first flowering (39 ± 3 mm), and produced large flowers (7.4 ± 0.2 mm; Figure 19 E–H, Appendix 4 Table 11). These results contrast with *Euphrasia* with no host, which flowered on average 52 days later (241.3 ± 7.9 Julian days), were extremely short at first flowering (11 ± 1 mm), and produced small flowers (5.3 ± 0.2 mm). *Euphrasia arctica* grown on *Arabidopsis*, *Equisetum*, *Festuca*, *Holcus*, or *Plantago* were all statistically significantly different from no host for at least one trait (Tukey comparisons, $p < 0.05$), while *E. arctica* on *Marchantia* or *Pinus* was not significantly different from no host for any trait ($p > 0.05$; Appendix 4 Table 12). While the overall effect of host was significant for all traits except nodes to flower (Appendix 4 Tables 13–15), three traits showed relatively little plasticity, with few statistically significant pairwise Tukey comparisons for nodes to flower (0 significant comparisons), number of leaf teeth (3), and internode length (4), while the other three traits showed many pairwise differences (days to flower, 21 significant comparisons; height, 16; corolla length, 12; Appendix Table 12). Our comparison of growth traits across host treatments measured through the year showed that height at the end of the season was weakly predicted from height 6 weeks after introducing a host ($R = 0.47$), but strongly correlated with height at first flowering ($R = 0.82$; Appendix Figure 10). Plants that flowered early were more likely to grow larger by the end of season ($R = -0.55$) and become more highly branched ($R = -0.57$; Appendix Figure 10).

Across host treatments, there was a significant negative correlation between Julian days to flower and most other traits (Table 7 B). We find that late flowering individuals are likely to be smaller at first flowering and have fewer branches, leaves with fewer teeth, and smaller flowers. While these traits were strongly correlated, there were substantial differences in the magnitude of response. For example, days to flower differed considerably depending on host, with a 3.8-fold greater difference than seen between means for different *Euphrasia* species grown on the same host (Figure 19 D,H). In contrast, corolla length and node to flower proved less variable depending on host, with a 1.4-fold and 1.2-fold change between means, respectively.

6.4.3 Variation in the wild

The comparison between the species differences common garden experiment and wild-collected herbarium specimens revealed population means of a single trait, nodes to flower, are strongly correlated ($R = 0.79$), and trait values are not significantly different ($p_{MCMC} = 0.71$) between environments (Figure 20; Appendix

Figure 11 and Table 16). All other traits did differ significantly between environments ($pMCMC < 0.05$), with *Euphrasia* plants in the common garden having corollas on average 1.4 mm longer, with 0.2 more teeth on the lower floral leaves, an increase in internode ratio of 1.0 mm, and 4 more pairs of branches. Despite these differences, there were correlations between the common garden and the wild-collected specimens for corolla length ($R = 0.93$, $pMCMC < 0.001$), internode ratio ($R = 0.65$, $pMCMC < 0.001$) and number of branches ($R = 0.29$, $pMCMC < 0.001$), but not for number of leaf teeth ($R = 0.07$, $pMCMC = 0.034$).

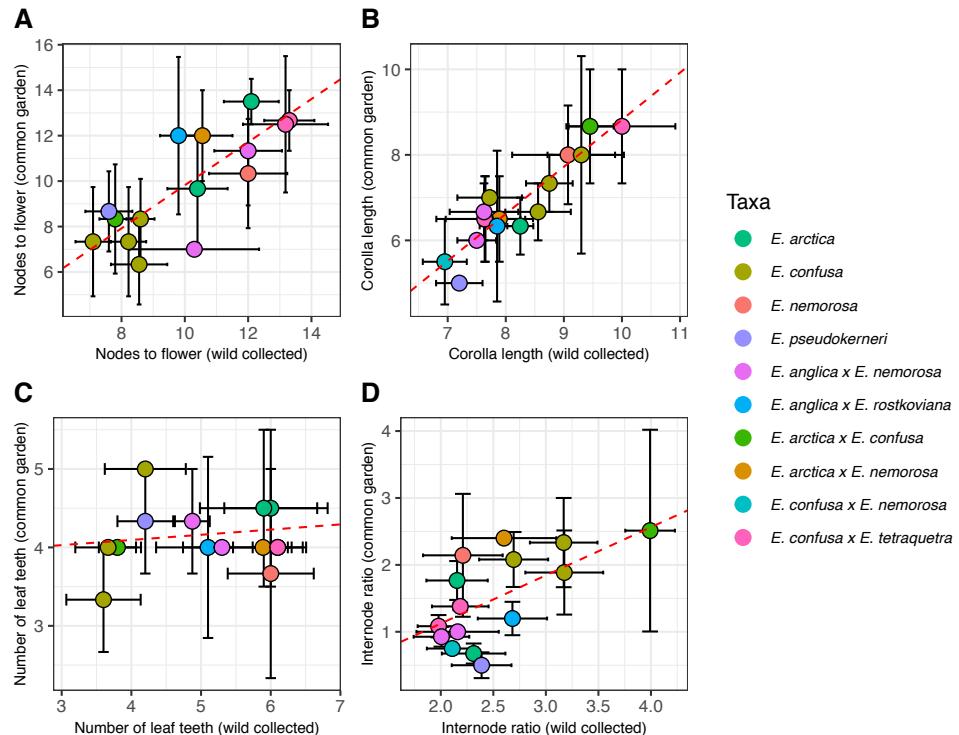


Figure 20: Relationship between morphological trait measurements made in the common garden and on wild-collected herbarium specimens for diverse *Euphrasia* species. Points are *Euphrasia* population means; bars represent the standard error of measurements. The line of best fit was calculated using coefficients from linear regression models on the means of each *Euphrasia* population. Length measurements are reported in millimeters. For an alternative representation of pairwise comparisons, see Appendix Figure 11.

6.5 Discussion

Our study sheds light on species differences, life history evolution and phenotypic plasticity of the generalist parasitic plant *Euphrasia*. We found different life history strategies between recently diverged species, with

some species rapidly transitioning to flower at the expense of growth-related traits, while others delay flowering and invest in early-season vegetative growth. However, many traits are phenotypically plastic and change in response to the host. While plants in benign common garden conditions grew vigorously, the correlation between life history traits in a common garden and in the wild suggests our experimental observations are indicative of patterns observed in nature. Morphological differences between species in the common garden also suggest that the currently delimited *Euphrasia* species are, at least in part, distinct. Overall, our study highlights the value of integrating trait data from multiple common garden experiments and field collections to study life history strategies in parasitic plants and demonstrates the rapid evolution of life history differences in a postglacial radiation of hemiparasites.

6.5.1 Life history variation in a generalist hemiparasitic plant

We found evidence for different life history strategies in British *Euphrasia*. *Euphrasia arctica*, *E. micrantha*, and hybrids such as *E. arctica* × *E. confusa*, transition rapidly to flower, flower while they are short, and produce their first flower from a low node on the plant. These rapid flowering species contrast with *E. pseudokernerii*, *E. nemorosa* and hybrids involving *E. nemorosa* that delay flowering until later in the season, grow tall before flowering, and produce their first flower from a late node on the main axis. These different life history strategies correspond to the known ecology of these species, with *E. nemorosa* flowering late in tall mixed grassland, while *E. micrantha* flowers early in patchy heathland (Metherell and Rumsey, 2018). While species show some general differences in life history strategies, there is also significant variation between populations within species. A relationship between internode number and habitat has previously been observed within *Euphrasia* species, with populations of *E. rostkoviana* in Sweden flowering at a lower node in a common garden if they have been collected from intensely grazed pasture (Zopfi, 1998). Overall, these observations within and between species are consistent with the classic life history trade-off between growth and reproduction (Stearns, 1992; Roff, 2002). For *Euphrasia* growing in the wild, early reproduction allows the plants to reliably complete their life cycle before summer competition, herbivory, mowing, summer drought, and other seasonal abiotic and biotic stresses. However, early flowering involves reproducing at the expense of early-season growth and at a time when the resource budget may be constrained by relatively few haustorial connections. These trait trade-offs pose an interesting comparison to the well-studied *Mimulus guttatus* (syn. *Erythranthe guttata*), a nonparasitic relative in the Lamiales that shares the same basic plant architecture. In *M. guttatus*, multiple traits related to growth and reproduction are correlated, both within and between populations, due to genetic trade-offs between time to flower and fecundity (Mojica et al., 2012; Friedman et al., 2015). In *Euphrasia*, the genetics underpinning this life history trade-off have yet to be characterized and may be a consequence of

multiple independent loci or trade-offs at individual loci (Hall et al., 2010).

While much life history variation is captured by differences in time to flower and growth-related traits, we also see evidence for flower size representing a separate axis of variation across *Euphrasia* species. In our common garden, *E. micrantha* has small corollas, while *E. arctica* and *E. nemorosa* have larger corollas, and corolla size is not strongly correlated with other traits. *Euphrasia* species are well known to have flower size variation, with a continuum between small-flowered species that are highly selfing (e.g., *E. micrantha*, corolla size = 4.5–6.5 mm, inbreeding coefficient $F_{IS} > 0.88$; Stone, 2013) and large-flowered species that are highly outcrossing (e.g., *E. rostkoviana*, flower size 8–12 mm, $F_{IS} = 0.17$ –0.25; French et al., 2005). Such wide variation in outcrossing rate has been documented in species of *Datura* (Motten and Stone, 2000), *Mimulus* (Karron et al., 1997), and *Nicotiana* (Breese, 1959). Small flowers have shorter anther–stigma separation and thus increased potential for autogamous selfing (Karron et al., 1997), while also having reduced attractiveness to pollinators and thus receiving less outcross pollen (Mitchell et al., 2004). In addition to differences in corolla size between *Euphrasia* species, corolla size also shows a change of up to 2 mm in response to host species. This change in flower size is of a magnitude that may potentially affect the mating system (Luo and Widmer, 2013) and suggests host species represents a previously unaccounted factor affecting the mating system of parasitic plants.

Our comparisons of *Euphrasia* species in a common garden also shed light on the distinctiveness of these recently diverged species and can be used to refine the suite of traits that are reliable in telling *Euphrasia* species apart. *Euphrasia* is a taxonomically complex plant genus, with the 21 currently described British species presenting complex and often overlapping morphological variation (French et al., 2008; Metherell and Rumsey, 2018; Wang et al., 2018). Our study suggests varying degrees of morphological distinctiveness of *Euphrasia* species. We see *E. micrantha* is morphologically distinct in the common garden and *E. pseudokerneri* somewhat distinct, while the closely related species *E. arctica*, *E. confusa*, and *E. nemorosa* differ in life history traits such as nodes to flower, but overlap in many other traits and are not clearly separated in the PCA. The morphological trait differences between species observed under standardised conditions are correlated with values from field-collected herbarium specimens where plants have associated with diverse hosts, been exposed to different ecological conditions, and were collected at different life-stages. These correlations suggest that our common garden results generalize to observations in nature. However, our study is likely to overestimate the distinctiveness of taxa by only including a subset of UK species and by choosing populations that could be identified to species level in the field. We suspect adaptive divergence between closely related *E. arctica*, *E. confusa*, and *E. nemorosa* is a consequence of differential natural selection for local ecological conditions such as soil water availability or mowing. Selection appears to be operating at a fine spatial scale, with significant

life history trait differences evident between populations within species. *Euphrasia* taxa may be genetically cohesive, either showing genome-wide divergence or divergence in genomic regions underlying life history differences (Twyford and Friedman, 2015), or alternatively these taxa may be polytopic and not genetically cohesive (Hollingsworth et al., 2017). Genomic sequencing of natural populations will help resolve the nature of species differences in *Euphrasia*.

6.5.2 Phenotypic plasticity in response to host

Our phenotypic plasticity experiment shows *Euphrasia* are affected by growing with a range of different hosts. Specifically, *E. arctica* with a host such as clover rapidly transitions to flowering. At the other extreme, *Euphrasia* grown without a host are small and flower late. These differences in growth are established early in the season, and early-flowering plants go on to grow the tallest and are more highly branched. Most other hosts result in a continuum of *Euphrasia* phenotypes between these extremes. Two surprising results were that *E. arctica* parasitizing *Arabidopsis* grew relatively tall despite the host senescing early in the growth season and that growth of *Euphrasia* with *Equisetum* was similar to growth on the commonly encountered grass *Holcus lanatus*. This result suggests that it attached to *Equisetum*, which would need to be confirmed by excavating root systems and observing haustoria, or it indirectly benefits without attachment through association with *Equisetum* fungal symbionts (Bouwmeester et al., 2007). Less surprising was the poor growth of *E. arctica* with *Pinus*. However, an association between *Melampyrum pratense* and *Pinus sylvestris* suggests at least some hemiparasitic Orobanchaceae benefit from attachment to woody host species or from interactions with their associated ectomycorrhizal fungi (Salonen et al., 2000).

The diverse effects of host on parasite growth are complex, but the variation we saw in our experiments may be attributed to host root architecture, germination time, and resource availability, as well as the presence of mechanisms to defend against parasite attack, such as cell wall thickening, localized host dieback, and chemical defence (Cameron et al., 2006; Twyford, 2018). While *Euphrasia* is generally thought to have low reliance on host resources, deriving only ~30% of carbon heterotrophically (Těšitel et al., 2010), at least under our experimental conditions *Euphrasia* only produced multiple flowers on certain hosts. Overall, our results point to *E. arctica* being a true generalist hemiparasite, but one where vigorous growth is only observed with a subset of potential hosts.

In terms of specific traits, only three pairs of trait correlations show consistent correlation coefficients in both *Euphrasia* common garden experiments (between height, number of branches, and leaf teeth), with other correlations between species breaking down when *Euphrasia* are grown on different hosts. The most notable

plasticity is seen in flowering time, with plants on clover rapidly transitioning to flower within ~100 days of germination, while plants with a more typical host (e.g., *Holcus lanatus*) flower a month later. Phenotypic plasticity in flowering time in response to resource availability is well documented in many plant groups, particularly *Arabidopsis* (e.g. Zhang and Lechowicz, 1994), but has received less attention in studies of parasitic plants, which are more likely to look at growth-related traits such as biomass (Ahonen et al., 2006; Matthies, 2017). However, date of first flowering has been shown to differ by up to 10 weeks in populations of *Rhinanthus glacialis* across Switzerland (Zopfi, 1995). Overall, we expect date of first flowering to be critical for the lifetime reproductive success of parasitic plants in the wild.

In contrast to seeing traits with extensive plasticity, we also saw evidence of developmental constraint in number of nodes to flower. For *E. arctica*, this trait showed the least plasticity with different hosts, is consistent between populations within species, and between the common garden and the field. Thus, the developmental event of transitioning to flower may be genetically determined, with changes in flowering time altered by plasticity in internode length and not nodes to flower. This developmental constraint may explain why nodes to flower is such an important diagnostic trait for species identification in *Euphrasia* and related species in the Rhinantheae (Jonstrup et al., 2016). Despite nodes to flower changing little in response to host species, our overall impression is that *Euphrasia* show considerable plasticity and little developmental constraint in many aspects of growth. In particular, differences between individuals on a given host also suggests other sources of variation, such as genetic background in host and parasite, as well as the timing of attachment, may be crucial in determining performance.

6.6 Conclusions

Despite over a century of experimental studies in parasitic plants, our understanding of the evolution of life history strategies in these diverse organisms is extremely limited. Our results with *Euphrasia* provide strong support for the rapid evolution of distinct life history strategies in response to local ecological conditions, with phenotypic plasticity further altering plant growth in response to host availability. We anticipate that future studies that test lifetime reproductive success of many parasitic plant species grown on many different host species will give further insight into the complex nature of host–parasite interactions.

7 Conserved and host-specific interactions in a multi-host parasite system

7.1 Abstract

Generalist parasites may infect multiple host species and experience complex host-parasite interactions. However, understanding the dynamics of host-parasite interactions remains a major challenge in generalist parasite systems. In this study, we use an experimentally tractable generalist parasitic plant group to understand host-parasite interactions in a multi-host system. Using common garden experiments, we show that extensive variation in the performance of hemiparasitic eyebrights (*Euphrasia*) across its host range can be attributed to both host life history and host phylogenetic relationships, but not host functional group as widely expected. While host-dependent parasite performance is generally conserved between eyebright species, with a restricted and phylogenetically divergent subset of hosts being highly beneficial, some eyebrights have more specialised host-parasite interactions. These host-parasite interactions show that a generalist parasite can respond to individual host selection pressures and may adapt to local host communities.

7.2 Introduction

The fitness of parasites depend on many aspects of the host to which they attach and feed, including their condition (Tscharren et al., 2007), defences (including immunity; Cameron et al. (2006); Bize et al. (2008)), growth rates (Barber, 2005), biomass (Matthies, 2017) and genotype (Vale and Little, 2009). These host attributes have widely been investigated using single-parasite single-host experiments in economically important pathogens and evolutionary model systems, for example malaria in humans (Becker et al., 2004; Mackinnon and Marsh, 2010) and nematodes in *Apodemus* wood mice (Meyer-Lucht and Sommer, 2005). However, increasing evidence shows that most natural systems have complex parasite dynamics where parasites affect multiple host species (Ameloot et al., 2005b), hosts are infected by multiple parasite species (Pedersen and Fenton, 2007), and parasite and host abundance differ in space and time (Mudrak and Leps, 2010). As such, the single-host single-parasite species doctrine is increasingly seen as a poor fit to natural systems (Poulin and Forbes, 2012), and studies of parasitism should seek to understand the more complex parasite-host dynamics present in nature.

Parasitic plants are a diverse group of 4,500 species of 12 separate origins that obtain water and nutrients from other plants using a specialised feeding organ called a haustorium (Westwood et al., 2010; Twyford,

2018). The majority of parasitic plant species are hemiparasites, which feed directly from other plants, but maintain their green habit and photosynthetic competency (Twyford, 2018). These hemiparasitic plants are ideal for investigating host-parasite interactions as they are experimentally tractable and can be grown in different combinations and with a wide host range (Brown et al., 2020). Moreover, hemiparasitic plants include ecosystem engineers that reduce the growth of competitively dominant taxa in grassland communities (Pywell et al., 2004), and species that threaten food security and cause billions of dollars' worth of crop losses in agricultural systems every year (Spallek et al., 2013), making the study of host range of crucial importance.

Generalist hemiparasitic plants that grow in mixed grassland communities may attach to any of the diverse range of co-occurring plant species. The dominant paradigm is that generalist hemiparasite performance is associated with host plant functional groups such as legumes, grasses, or forbs, with legumes often the best hosts (Rowntree et al., 2014; Matthies, 2017). However, substantial variation in host quality within functional groups suggests functional group alone may not be a good predictor of host quality (Rowntree et al., 2014). Moreover, some functional groups are monophyletic clades such as grasses (Poaceae), while some are paraphyletic groups such as forbs. As such, the performance of hemiparasites may be better predicted by host phylogeny rather than functional group, with some host clades possessing attributes such as defence against parasites (Cameron et al., 2006) or root architecture (Roumet et al., 2006) that confer higher growth to the parasite. Hemiparasite performance may also be expected to be affected by other aspects of the host, for example different host life history strategies (annual or perennial) that may have different resource accessibility (Garnier, 1992) or relative carbon and nitrogen content (Garnier and Vancaeyzeele, 1994). Finally, theoretical models of parasitism predict more complex host-parasite interactions will arise in heterogeneous environments with variable host abundance and/or a mix of different host genotypes (Gandon, 2002). Such host-parasite interactions are not known in facultative generalist hemiparasitic plants, but have been observed in the obligate hemiparasitic plant *Striga*, where specific host-genotype parasite-population interactions underlie the success of *Striga* parasite development (Huang et al., 2012).

Here, we use facultative generalist hemiparasitic eyebrights (*Euphrasia*, Orobanchaceae) to investigate the host attributes that determine parasite performance. This genus is an ideal model for studying host-parasite interactions as they are small in size and easy to cultivate with a rapid annual lifecycle (Brown et al., 2020), and species co-occur with diverse hosts in different habitats (Metherell and Rumsey, 2018). We consider multiple aspects of *Euphrasia* performance, including survival and reproduction through the year, and aim to quantify how hemiparasite performance responds to many host species. Specifically, we ask: (1) how does *Euphrasia* perform across its host range and on non-hosts? (2) Do common host characteristics such as functional group, life history, or relatedness (phylogeny) impact on the survival and performance of hemiparasitic *Euphrasia*?

(3) Do different *Euphrasia* species perform similarly with a given host species, or is reproductive success determined by host-parasite interactions? Our aim is to understand the potentially complex responses of a generalist parasite to diverse host attributes.

7.3 Methods

7.3.1 Plant material, cultivation and trait measurements

We addressed parasite-dependent host performance in two common garden experiments. Experimental 1 aimed to understand the performance of *Euphrasia* across a range of hosts, and to link performance with host characteristics. For this experiment, we grew a single species, *Euphrasia arctica*, on forty-five diverse host species (Appendix 5 Table 17). *E. arctica* was chosen as the focal species due to its widespread nature in Britain, where it mainly occupies lowland grassland habitats. Experiment 2 was designed to detect parasite-host interactions, and we used four different species of *Euphrasia* and thirteen species of hosts (Appendix 5 Table 18). Two diploid species (*E. anglica*, *E. vigursii*) and two tetraploid species (*E. micrantha*, *E. tetraquetra*) of *Euphrasia* were chosen to represent the diversity of the genus in Britain.

For both experiments, we used wild-collected open-pollinated seeds of *Euphrasia* (Appendix 5 Table 19). *Euphrasia* seeds were sown in individual pots filled with Sylvamix 1 compost and placed in an outside array at the Royal Botanical Garden Edinburgh (RBGE). Details of soil mixes and sowing protocols followed (Brown et al., 2020). In Experiment 1, a total of 3000 *Euphrasia* seeds were sown in winter 2016 of which 1308 germinated. In Experiment 2, a total of 2880 *Euphrasia* seeds were sown in winter 2017 of which 988 germinated. In both experiments after *Euphrasia* germination, plants were grown in an unheated, well ventilated greenhouse. Host plants were then transplanted into *Euphrasia* pots using tweezers which carefully extracted the host plant and a dibber to create space in the *Euphrasia* pot. Host plants were then placed half way between the central *Euphrasia* germinant and a pot vertex. In Experiment 1, all 45 hosts were transplanted in spring 2017 into *E. arctica* pots (Appendix 5 Table 17) and in Experiment 2, all 13 hosts were placed in all four *Euphrasia* host species pots in spring 2018 (Appendix 5 Table 18). Host plants were replaced if mortality occurred within two weeks of the transplant date. All pots containing *Euphrasia*-host combinations were randomized weekly.

The traits measured on *Euphrasia* plants were proxies of fitness, to understand how *Euphrasia* fitness is affected by host plant species (Experiment 1) and whether specialised interactions occur between *Euphrasia* and particular host species. For Experiment 1 we measured date of first flowering, and then both the number of reproductive nodes and whether an individual *Euphrasia* was alive or dead every 30 days beginning on

the 30.05.17 until the 30.09.17. These are referred to as time points one (May) to five (September) in the rest of the chapter. For Experiment 2, we measured reproductive nodes at the end of the season. In both experiments, germination date and date of host introduction were recorded. Here, reproductive nodes are defined as whether a node on a *Euphrasia* plant contained either a flower or fruit or not (binary), summed for all nodes on a plant. Cumulative across the season, the number of reproductive nodes is the reproductive output of a *Euphrasia* individual. The traits of host species which were included in Experiment 1 were the functional group of host (woody, fern, forb, grass, or legume), and the life history of the host (whether annual or perennial).

7.3.2 Statistical analyses

We integrated a phylogenetic tree into the analysis of Experiment 1 to understand if host plant relatedness impacted reproductive output of *Euphrasia*. The phylogeny used in this analysis was based on the two gene alignment of *rbcL* and *matK* from (Lim et al., 2014). Six sequences from three species (*Zea mays*, *Hordeum vulgare* and *Lagurus ovatus*) were added from NCBI, as they were not present in the original dataset. The maximum likelihood phylogeny was generated using IQ-TREE (Nguyen et al., 2015) with branch support estimated using 1000 ultrafast bootstrap replicate, and using the TESTNEWERGE flag for model selection. A constraint tree was created using the phylomatic function in the R package brranching (Chamberlain, 2019) and used to topologically constrain the phylogeny based on the APG IV phylogeny. The tree was then made ultrametric prior to model-based analyses.

The models for Experiment 1 are designed to assess the impact of host species and their attributes on the performance and fitness of *Euphrasia arctica*. All subsequent statistical analyses were conducted in R version 3.6.1 with all data manipulation in base R or data.table (Dowle and Srinivasan, 2019). In Experiment 1, the three *Euphrasia* traits of interest – survival, number of days to flower, and reproductive nodes of *Euphrasia* – were estimated using a Bayesian generalized linear mixed effect model approach in the MCMCglmm package (Hadfield, 2010). *Euphrasia* survival was modelled using the “threshold” option in MCMCglmm (i.e. the CDF of the Gaussian distribution) which is also known as an event history analysis model (EHA) as it can take into account time-varying covariates. The number of days to flower and reproductive nodes (both at the end of the season, and over time) were modelled using a Poisson distribution.

Functional group and life history of host, as well as normalized transplant date (time lag between germination and receiving a host, scaled to difference in first transplant date), were added as fixed effects, whilst host species and phylogenetic effects were treated as random effects. In the EHA, time point was also added as a

fixed effect to model the effect of time itself on *Euphrasia* survival. Time point five was removed from the EHA, as all but two individuals were dead at this time. We parameterized the reproductive output over time model differently. Time point and its interaction with host life history were additional fixed effects and time points one and five were removed due to lack of reproduction at these time points. We included a random effect variance structure of an interaction of time point and host species using the us() variance function in MCMCglmm which allows covariance between host and time point:

$$V_{HE} = \begin{matrix} & T_{2,2} & T_{2,3} & T_{2,4} \\ T_{2,3} & & & \\ & T_{3,3} & T_{3,4} \\ T_{2,4} & & & \\ & T_{3,4} & T_{4,4} \end{matrix} \quad (5)$$

Where V_{HE} is the variance in host effect and T is the time point. The residual (V_e) variance-covariance matrix allowed no covariance between time points using the MCMCglmm function idh():

$$V_e = \begin{matrix} & T_{2,2} & 0 & 0 \\ 0 & & T_{3,3} & 0 \\ & 0 & 0 & T_{4,4} \end{matrix} \quad (6)$$

Models were run for a minimum of 130000 iterations, following a burn-in of 30000 iterations, and a thinning interval of 100. Parameter expanded priors were used to improve convergence, and effective sample sizes of focal parameters were in excess of 500 and mostly approaching 1000. Significance of categorical covariates with more than one level were determined using Wald Tests (Brown, 2019), otherwise the pMCMC value of the covariates were reported. Variance explained by random effects including phylogeny (phylogenetic signal) were calculated as ratios of the variance of the parameter of interest to the residual variance in the model. For joint phylogenetic estimates, the posterior distributions of the phylogenetic and host species effects were summed. Significance of random effects were determined using likelihood ratio tests in the package lme4 where appropriate (Bates et al., 2015). Convergence and autocorrelation of models was assessed visually by plotting the posterior distributions of the estimated parameters.

The model in Experiment 2 differed significantly from the models in Experiment 1, as phylogenetic random effects and host trait fixed effects (functional group and life history) were not included. Only one response variable was modelled, which was the cumulative number of reproductive nodes at the end of the season, using a Poisson distribution. The fixed effects included the *Euphrasia* species themselves, which population a *Euphrasia* species came from (Appendix 5 Table 19), and the normalised transplant date (as above). Host

species and the host species interaction with *Euphrasia* species were added as single parameter random effects. The correlation in hosts effects was calculated as the ratio of host effects to the host effects and host species interaction with *Euphrasia* species combined. Models were run as specified for Experiment 1.

7.4 Results

7.4.1 Hemiparasite performance across host species

We first investigated survival of *Euphrasia arctica* across the 45 different host species, where 1252 *Euphrasia* plants survived to be measured at the first time point. Survival was not significantly affected by host functional group ($\chi^2 = 3.38$, df = 4, p = 0.50; Figure 21 shows legumes and grasses as examples) or host life history ($\chi^2 = 0.40$, df = 1, p = 0.53; Appendix 5 Table 20) in an event history analysis. Instead, between-host effects explained 24.6% of variation in survival when accounting for phylogeny (13.4-55.4% CI, 95% Credible Intervals), with the probability of survival ranging from 0.31 when grown on heather (*Erica tetralix*) to 0.75 on cleavers (*Galium aparine*). The overriding effect of host on *Euphrasia* survival was also evident from the standard deviation of the host effects (0.57, 0.39-1.11 CI) being of a greater magnitude than the fixed effects of life history (0.14, -0.25-0.61 CI) and functional group (-0.19, -1.42-0.67 CI; Appendix 5 Table 20). Taken together, these results indicate host species impacts hemiparasite survival in a common garden environment, with survival being species specific rather than being influenced by host plant group (i.e. phylogenetic clade, or life history).

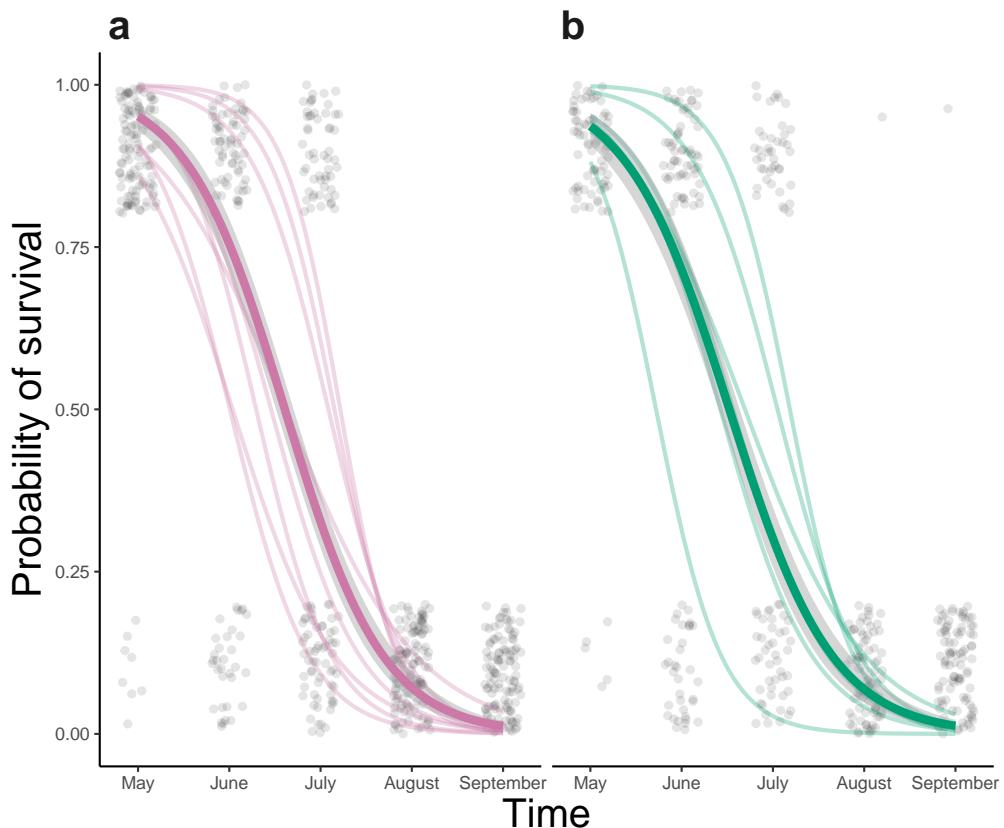


Figure 21: Probability of *Euphrasia arctica* surviving in a common garden experiment on 14 host species from two representative families, the Fabaceae (a) and Poaceae (b), using host species binomial regressions. Pale regressions represent individual species and bold regressions represent family level regressions. Pale grey dots are jittered raw values of an individual's living status (binary) at each time point from earliest census in May to the latest in August.

To understand how host species impacts on reproduction, we then tracked first flowering and reproductive output of *Euphrasia* individuals in the common garden through the growing season. The date of first flowering differed 3.5-fold across *Euphrasia* plants, with *Euphrasia* on good hosts flowering earlier (e.g. Bird's foot trefoil, *Lotus corniculatus* = $78.0 \text{ days} \pm 3.5 \text{ SE}$, Standard Error) than those on poor hosts (e.g. maize, *Zea mays* = $129.2 \text{ days} \pm 5.1 \text{ SE}$). The difference in days to flower could not be explained by host functional group ($\chi^2 = 2.00$, $df = 4$, $p = 0.73$) and instead between-host effects explained 35.1% (20.0-83.5% CI) of the variation when accounting for phylogeny. Life history was marginally significant ($\chi^2 = 3.88$, $df = 1$, $p = 0.05$; Appendix 5 Table 21), although highly variable in its effect (77.4-101.9 days to flower CI). We found *Euphrasia* flowered earlier on annual hosts, which may be expected as annuals are a more ephemeral resource. To investigate performance over time we observed reproductive output at five time points (May-September) throughout

the season. Over time, the effect of host functional group was non-significant ($\chi^2 = 7.37$, df = 4, p = 0.12), however host life history interacted with the September census point, with 4.7 times fewer reproductive nodes in *E. arctica* on annual hosts than perennial hosts (0.14-127 times CI; $\chi^2 = 103$, df=2, P<0.001), Appendix 5 Table 22). While *Euphrasia* flowered earlier on annual hosts, and therefore had the potential for a longer reproductive period, these same hosts were more likely to die earlier in the season. *Euphrasia* had consistently high reproductive success on some hosts (e.g. *L. corniculatus* and *Trifolium pratense*; Appendix 5 Figure 12), however other hosts (e.g. *Cynosurus cristatus*) conferred high reproduction for *Euphrasia* earlier in the season and gradually declined to zero. Overall, this shows the trajectory of hemiparasite reproductive success depend on the specific host species, and their life history (Appendix 5 Figure 12).

By the end of the season, *Euphrasia* produced on average more than one reproductive node on 28 out of the 45 hosts. Total reproductive output could not be explained by host functional group ($\chi^2 = 6.83$, df = 4, p = 0.14, Appendix 5 Table 23) or host life history ($\chi^2 = 0.08$, df = 1, p = 0.78). Instead, host species explained 81.8% (65.9-95.6% CI) of the variability in end of season reproductive nodes accounting for phylogeny and phylogenetic signal was high for this trait (0.82, 0.17-1.00 CI; Appendix 5 Figure 13). *Euphrasia* produced a large numbers of reproductive nodes only with few host species such *Lotus corniculatus* (104.5 ± 19.1 SE reproductive nodes), *Cynosurus cristatus* (53.6 ± 8.4) and the plantain *Plantago lanceolata* (35.5 ± 3.7 ; Figure 22). These results highlight the importance of phylogenetic relatedness of host plant species in predicting hemiparasite performance, above host species functional group.

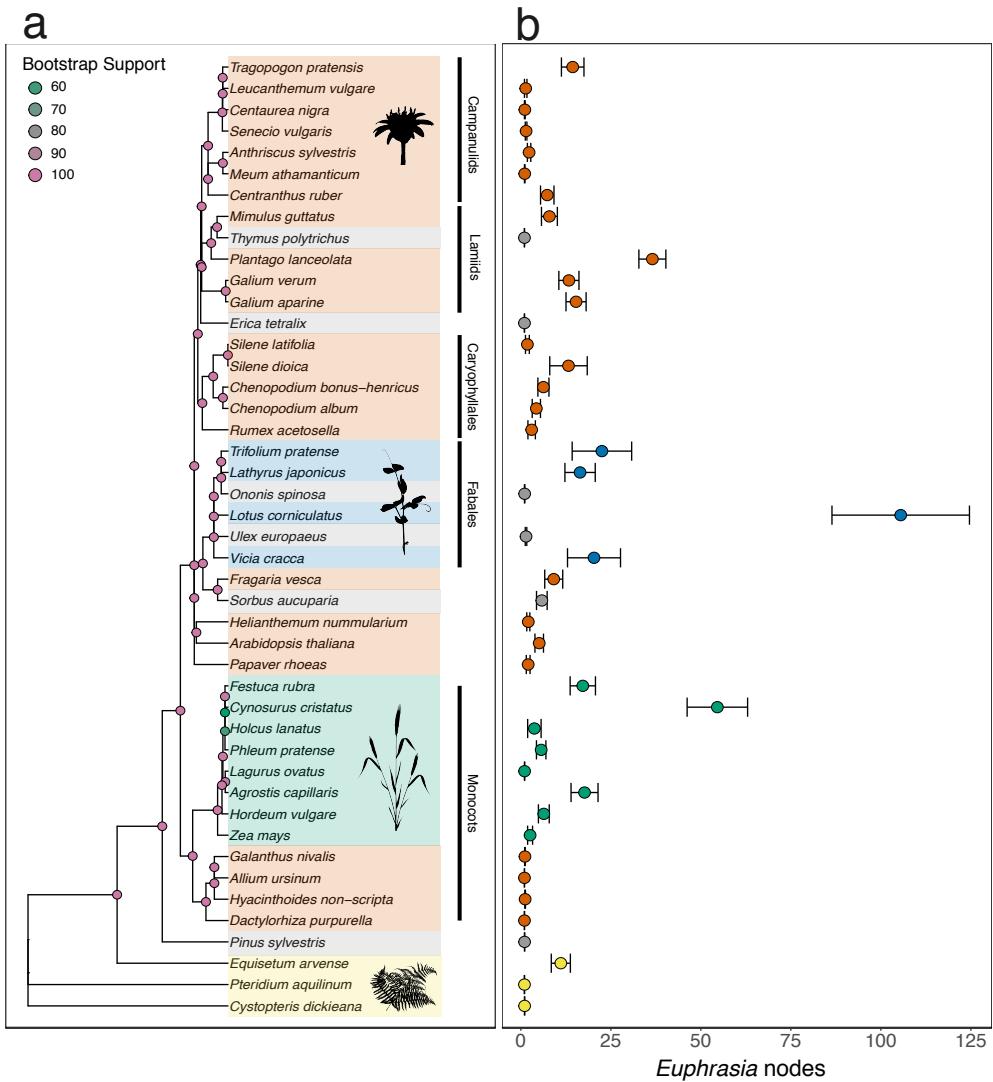


Figure 22: Performance of hemiparasitic *Euphrasia arctica* measured as cumulative reproductive nodes, in the context of host species and host phylogeny. (a) Maximum likelihood phylogeny of 45 hosts based on *rbcL* and *matK*. Bootstrap values are shown for each node on the phylogeny. Monocots, the two largest orders and two superorders are labelled. Host species are coloured by functional group, orange = forbs, grey = woody plants, blue = legumes, green = grasses and yellow = ferns. (b) Values are mean cumulative reproductive nodes of *Euphrasia* per species with colours corresponding to functional group of host \pm one standard error. Silhouetted pictures are from phylropic.org.

7.4.2 Hemiparasite-host interactions

We then tested for complex hemiparasite-host interactions, by measuring the reproductive success of six populations from four divergent species of *Euphrasia* in a common garden using 13 hosts from different habitats (Appendix 5 Tables 18,19). A total of 635 *Euphrasia* plants survived to the end of the season and were measured. After taking into account differences between *Euphrasia* species and populations in their reproductive output ($\chi^2 = 4.40$, df = 6, p < 0.001; Appendix 5 Table 24), there was evidence for both consistent host driven differences in parasite performance, and specific parasite-host interactions (Figure 23). Host species accounted for most of the variation in reproductive nodes at the end of the season (26%; $\chi^2 = 15.6$, df = 1, p < 0.001), followed by host interacting with *Euphrasia* species (12.3%; $\chi^2 = 27.1$, df = 1, p < 0.001; Appendix 5 Figure 14). *Euphrasia* species tended to react similarly to a given host, with a 0.76 (0.37-0.93 CI) correlation in reproductive output when two hosts are picked at random. By investigating model best linear unbiased predictors (BLUPs), we find differences in host effect are driven by *L. corniculatus*, the speedwell *Veronica chamaedrys*, and sea plantain *Plantago maritima*, each of which have antagonistic interactions with different *Euphrasia* species. Moreover, two divergent species of *Euphrasia* from the same geographic location, diploid *E. vigursii* and tetraploid *E. tetraquetra*, show similar responses to the same set of hosts, with no significant interactions detected in these two species (Appendix 5 Figure 15; $\chi^2 = 0.22$, df = 1, p = 0.64). Although the dominant signal is that of conservatism of performance across *Euphrasia* species on the same host, parasite-host interactions explain a significant proportion of the variation in performance.

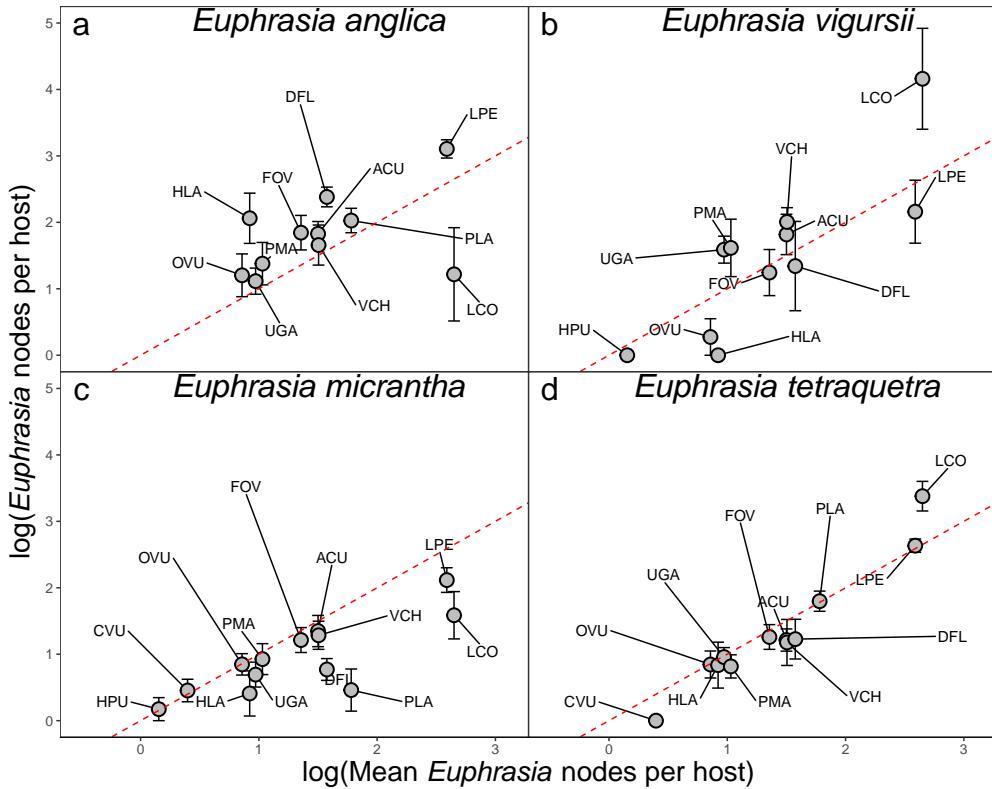


Figure 23: Performance of four species of *Euphrasia* on thirteen different species of host plants measured as cumulative reproductive nodes at the end of the season. Each panel represents a unique *Euphrasia* species. The x-axis represents the number of reproductive nodes of *Euphrasia* for each host averaged across all *Euphrasia* species, while the y-axis shows reproductive nodes per *Euphrasia* species \pm one standard error. Both axes are log transformed. The red dashed line graphs $y=x$; points above the line indicate elevated response to a host beyond the average, while points below the line indicate the opposite. Host species are ranked by average performance conferred to a *Euphrasia* species, where HPU = *Hypericum pulchrum*, CVU = *Calluna vulgaris*, HLA = *Holcus lanatus*, OVU = *Origanum vulgare*, UGA = *Ulex gallii*, PMA = *Plantago maritima*, PLA = *Plantago lanceolata*, VCH = *Veronica chamaedrys*, FOV = *Festuca ovina*, DFL = *Deschampsia flexuosa*, ACU = *Agrostis curtisii*, LPE = *Lolium perenne* and LCO = *Lotus corniculatus*.

7.5 Discussion

We have shown that the performance of a hemiparasitic plant is determined by a complex suite of host attributes that impact on different aspects of survival, the initiation of reproduction, and reproductive success through time. Our results support the emerging view that host functional group alone is not a good single predictor of hemiparasitic plant performance (Rowntree et al., 2014; Matthies, 2017) and that other aspects of the host biology must be considered. For example, we find host life history has a notable effect

on hemiparasite reproduction through time, which may be of significance in natural communities due to the restricted availability of many annual host plants later in the season (Kelly et al., 1988; Zopfi, 1993). Moreover, there is a phylogenetic signal in how host impacts hemiparasite growth and reproduction, indicating that host traits such as defences to parasitism (Cameron et al., 2006), root architecture (Roumet et al., 2006), nutrient availability and the uptake of secondary compounds (Adler, 2000), and competitive ability (Keith et al., 2004) vary in a predictable way across the plant phylogeny.

The few other studies to date to incorporate host phylogeny and species traits in multi-host parasite systems have made similar findings. For example, a study of apicomplexan parasites that infect diverse bird hosts found that host phylogeny was important in explaining variation in infection status on top of environmental and host species traits (Barrow et al., 2019). In *Euphrasia*, the quality of hosts can partly be explained by phylogeny and host life history, though there are also idiosyncratic patterns of specific hosts that require further study. Generalist parasite species are often thought to have intermediate fitness across several hosts (Leggett et al., 2013). Our experiments however, show that there are a restricted set of highly phylogenetically divergent host species which confer high benefit to *Euphrasia* (especially *L. corniculatus*, *C. cristatus* and *P. lanceolata*). Overall, while *Euphrasia* is a true generalist able to benefit from parasitising plants throughout the vascular plant phylogeny, it only gains major benefit from attaching to a subset of taxa. *Euphrasia* species may therefore lie in a ‘grey zone’ in between generalist and specialist parasites, as has been observed in other systems (Lievens et al., 2018).

Our finding that hosts beneficial to one *Euphrasia* species are generally beneficial across all *Euphrasia* species reveals conserved host-parasite interactions. This is perhaps unsurprising as hemiparasites are likely to respond in a similar way to high quality resources, for example perennial hosts that are large, nitrogen rich with few defences (Seel et al., 1993; Cameron et al., 2006; Krasnov et al., 2006). In our multi-parasite experiment, the hosts that emerged as most consistently advantageous across all four *Euphrasia* species were *Lolium perenne* and *L. corniculatus*, which fulfil many of the above criteria (Beddows, 1967; Jones and Turkington, 1986). These conserved parasite responses to the same host species are notable in the context of substantial divergence between the two diploid and two tetraploid species examined (~5% divergence Wang et al., 2018). In contrast, host conservation in holoparasitic Orobanchaceae is uncommon, with host specific ecotypes found even within the same parasite species (Thorogood, Rumsey, Harris, et al., 2009). This is likely to be due to the critical need for holoparasites to respond to the correct host to ensure establishment.

We do however find significant host-parasite interactions and species-specific responses to some hosts, suggesting weak differential host adaptation. Host species are spatially heterogeneous in their distribution and vary in abundance by habitat and geographic area, creating conditions that may allow local host adaptation. The

low migration rate between *Euphrasia* populations, particularly in small flowered selfing taxa (French et al., 2005), may cause differentiation and promote local adaptation. While the drivers and tempo of local host adaptation are not fully understood, further investigations with dense host sampling will shed light on the nature of these interactions, and reveal if host species should be considered in conservation translocations.

8 General discussion

8.1 The interaction and importance of ploidy and hybridisation

The first part of my thesis investigated the interaction between ploidy variation and hybridisation at different phylogenetic and geographical scales. While it is generally thought that ploidy level acts as a significant barrier to hybridisation (Husband and Sabara, 2004), a literature review (Chapter 2) showed strong evidence that hybridisation between ploidy levels was common, particularly in many groups of plants. Such interactions between ploidy level of the parental species and hybridisation can have long lasting evolutionary consequences disproportional in effect to the frequency of their occurrence, such as new lineage or species formation, and may be being amplified by global change (Mandakova et al., 2013; Vallejo-Marin and Hiscock, 2016). At a regional floristic scale, the British flora provided an excellent framework for studying hybridisation and polyploidy due to the wealth of morphological, ecological, and genetic data. At this flora wide level, I found (phylo)genetic factors and ploidy level to be important in predicting whether a species pair will successfully form a hybrid (Chapter 3). Species of differing ploidy level were less likely to form hybrids, but my model indicated that the barrier is porous. At a finer scale still, I used the genus *Euphrasia* to investigate a contact zone between two species of differing ploidy level (Chapter 4). Although I found a lack of evidence for hybridisation and introgression using conventional population genetic tools, demographic modelling indicated there may have been limited historical or ongoing gene flow between these two species. Combining all present evidence in the thesis, hybridisation between ploidy levels was common globally, but may be rare in particular study systems. In the TCG *Euphrasia*, even rare hybridisation between ploidy levels may be sufficient to generate new species. My thesis adds to the emerging view that polyploidisation, in conjunction with hybridisation plays an important role in creating and maintaining diversity across different taxonomic groups both historically (e.g. *Galium* Kolar et al. (2015)) and presently (e.g. *Senecio* Abbott and Lowe (2004)).

8.2 Evidence of parasitism playing a role in taxonomic complexity

My results using common garden experiments with *Euphrasia* showed that the hemiparasitic habit can contribute to taxonomic complexity. In the first experiment (Chapter 6), the results suggested that although some species were consistently morphologically distinct (e.g. *Euphrasia micrantha*), other closely related species overlapped in many traits and sometimes could not be reliably told apart in a common garden setting. I suggested that there may be differential natural selection for local ecological conditions in the wild that drive life history differences between species. There was also considerable phenotypic plasticity in relation to the

host species being parasitised, with only a few traits (e.g. node to first flower) that showed consistency when a *Euphrasia* species was grown on a particular host species. Next, I investigated host-parasite interactions across a range of host-*Euphrasia* species combinations to see if there was evidence of host specificity (Chapter 7). I found evidence of host parasite interactions, which could be responsible for local host adaptation, and in turn may drive the evolution of cryptic specialisation. Although this hypothesis is consistent with my results, more work would need to prove local host adaptation in the wild. In sum, host species influenced both the morphology and fitness of *Euphrasia* individuals, creating the potential for species confusion through phenotypic plasticity, and cryptic specialisation of *Euphrasia* on certain host species. Little work has been done on taxonomic complexity through the lens of parasitism in other hemiparasitic plant genera. A good candidate for investigation would be *Pedicularis*, which is a large genus, containing species complexes (Garg, 2010). Due to the relatedness to *Euphrasia* both phylogenetically and in terms of the presence of much recent speciation, I would postulate that *Pedicularis* has similar processes that may be driving taxonomic confusion and diversity in the genus.

8.3 Critique and further study

Given the broad aims of the thesis, there are areas which could benefit from further study. In particular, an increased taxonomic scope would help to generalise more across plants. A wealth of high quality data is available for the British flora (Chapter 3; e.g. chromosome counts (BSBI Cytology Database), DNA barcoding data, ecological data Fitter and Peat, 1994, hybrid identities @RN1007), but this is not so for many other floras, especially in tropical regions of the world. Regions with genetic, ecological and hybrid identity data available (i.e. temperate regions) may also not be representative of the processes operating elsewhere, as these same regions are for the most part highly degraded habitats with disturbed ecological processes and altered evolutionary trajectories. Degradation and disturbance would affect findings, as we expect hybrids to be more common, and more alien species to be present in these situations (Vallejo-Marin and Hiscock, 2016). Another area which would benefit from increased scope is the GBS study (Chapter 4). Here, I investigated only one contact zone and a limited number of individuals, where hybrids may have been missed due to the small sample sizes. A single contact zone does not represent the whole spectrum of possibilities, for example some contact zones may be younger than others, or present asymmetrical numbers of parental plants and hybrid genotypes (Twyford et al., 2015). Therefore sampling multiple contact zones across multiple species which do and do not differ in their ploidy level would have been ideal. On the other hand, Yeo (1954) found only a single triploid in his large cytogenetic survey of *Euphrasia*, indicating hybridisation between ploidy levels is rare. In the latter growth experiment (Chapter 7), relating host preferences to *Euphrasia* growing in the wild

would give a clearer indication of host preferences. This could be either done by using quadrats to relate nearby hosts, or more accurately by looking for haustorial connections to host roots (Gibson and Watkinson, 1989). This technique of looking at haustorial connections by uprooting *Euphrasia* individuals would give a better proxy for fitness and host suitability in a common garden setting.

There are many avenues for potential future research and I will highlight a few here. For further investigation into the genetics of *Euphrasia*, a complete and contiguous (potentially phased) whole genome assembly of both a diploid and a tetraploid species is essential. Using these complete genomes it will be possible to accurately characterise regions underlying adaptive introgression, and detect structural changes between the hybrid species and their parental progenitors (Chapman and Abbott, 2010; Jay et al., 2018). In Chapter 4, I used genotyping by sequencing to generate SNPs across the genome, however it would be useful to use whole genome data to resolve fine level introgression. For example, this could show recent or historic introgression between the diploid species and homologous regions of the tetraploids. This is difficult, as it requires the identification of ploidy-specific diagnostic sites that must be shown in putative hybrids. Whole genome data across a wider range of species would also yield powerful comparative insights of the extent of genomic introgression across *Euphrasia* in the UK. I did generate whole genome sequence data and draft assemblies for two species of particular interest, *Euphrasia micrantha* and *E. vigursii*. A previously sequenced *E. anglica* genome meant that now both putative parental species for *E. vigursii* were available for analysis. With postdoc Hannes Becher, we mapped putatively disomic scaffolds of *E. anglica* and *E. vigursii* to the *E. micrantha* reference and found slightly more sequence of *E. vigursii* mapped (342Mb) than *E. anglica* (327Mb). This result is consistent with *E. vigursii* being a hybrid species, however the results are far from conclusive and work is ongoing to resolve the relationships of these three taxa. Better assemblies, and more individuals of each species along with the parental progenitors would allow us to understand whether these are hybrid species (and if so, what are the parental contributions to the hybrid genome), or simply derived diploid populations diverged in allopathy. Indeed, finding and sequencing hybrid *Euphrasia* individuals or populations would allow major insight into the nature of hybridisation in *Euphrasia*.

In Chapters 6 and 7 I used common garden experiments to understand how host species impact the morphology and fitness of *Euphrasia* individuals. It would be ideal to characterise the hemiparasitic habit of *Euphrasia* further by using field experiments in the wild. Many different populations of *Euphrasia* could be studied to yield information about host association in the wild to understand the correlation between host species and morphology. In an attempt to begin this process, I measured the number of potential host species present around 20 randomly sampled *Euphrasia* individuals at each population that was sampled for use in the growth experiment in Chapter 7. I related these host species occurrences in the wild to the growth of *Euphrasia* in

cultivation to see if host species that were more frequently encountered in the wild led to higher *Euphrasia* fitness. Although I found no significant association, there were obvious limitations - I used few populations and my statistical power was low. Another potential field experiment includes excavating *Euphrasia* plants to relate the number of haustorial connections to neighbouring host plant species, to the morphology and fitness of *Euphrasia* plants. This is a more realistic and rigorous approach to the problem, but made difficult due to the fine haustorial connections which can be broken easily. Lastly, to place my results in a more comparative context, it would be good to understand host preferences in other genera in the Orobanchaceae/Rhinanthae – are they the same as in *Euphrasia*? Do host shifts between genera occur? If so, why? Is there phylogenetic signal in host preference? I would expect to find differences between genera, as there appears to be some host specialisation at this level (e.g. *Melampyrum* on *Medicago sativa/Achillea millefolium* Matthies (2017); *Rhinanthus* on *Festuca ovina/Cynosurus cristatus* Cameron et al. (2006)), but more comprehensive datasets are needed to address this rigorously.

8.4 Thesis conclusions

The main aims of this thesis were to understand both the interaction between hybridisation and polyploidy in both *Euphrasia* and the British flora, and the role of hemiparasitism in driving taxonomic complexity in *Euphrasia*. In Chapter 2, I found that cross ploidy hybridisation is a common phenomenon across plants, which had previously been little explored. Chapter 3 revealed that (phylo)genetic factors and ploidy level were critical in explaining hybridisation across the British flora. Chapter 4 used GBS data to show that hybridisation in a cross ploidy *Euphrasia* contact zone is rare. Chapter 5 described the horticultural protocols developed so far in *Euphrasia*, some of which are used in the final two chapters. In Chapter 6, I showed that *Euphrasia* species can overlap in morphology on a single clover host species, and different host species drive phenotypic plasticity in a common garden. Lastly, in Chapter 7 I used another common garden experiment which revealed that *Euphrasia* responses to host species were mainly conserved across species, however host-parasite interactions were also present. Ploidy variation in conjunction with hybridisation, and parasitism, continue to shape the evolution of plants in profound ways, and warrant further study to understand the mechanisms underlying these phenomena.

Bibliography

- Aagaard, S. M. D., S. M. Sastad, J. Greilhuber, and A. Moen. 2005. A secondary hybrid zone between diploid *Dactylorhiza incarnata* ssp *cruenta* and allotetraploid *D. lapponica* (Orchidaceae). *Heredity* 94: 488–496.
- Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. W. Boughman, et al. 2013. Hybridization and speciation. *Journal of Evolutionary Biology* 26: 229–246.
- Abbott, R. J. 1992. Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology & Evolution* 7: 401–405.
- Abbott, R. J., A. C. Brennan, J. K. James, D. G. Forbes, M. J. Hegarty, and S. J. Hiscock. 2009. Recent hybrid origin and invasion of the British Isles by a self-incompatible species, Oxford ragwort (*Senecio squalidus* L., Asteraceae). *Biological Invasions* 11: 1145–1158.
- Abbott, R. J., F. C. Bretagnolle, and C. Thebaud. 1998. Evolution of a polymorphism for outcrossing rate in *Senecio vulgaris*: Influence of germination behavior. *Evolution* 52: 1593–1601.
- Abbott, R. J., H. E. Ireland, and H. J. Rogers. 2007. Population decline despite high genetic diversity in the new allopolyploid species *Senecio camrenensis* (Asteraceae). *Molecular Ecology* 16: 1023–1033.
- Abbott, R. J., J. A. Irwin, and P. A. Ashton. 1992. Genetic diversity for esterases in the recently evolved stabilized introgressant, *Senecio vulgaris* L. var *hibernicus* Syme, and its parental taxa *S. vulgaris* L. var *vulgaris* and *S. squalidus* L. *Heredity* 68: 547–556.
- Abbott, R. J., and A. J. Lowe. 2004. Origins, establishment and evolution of new polyploid species: *Senecio camrenensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society* 82: 467–474.
- Adler, L. S. 2000. Alkaloid uptake increases fitness in a hemiparasitic plant via reduced herbivory and increased pollination. *American Naturalist* 156: 92–99.
- Ahonen, R., S. Puustinen, and P. Mutikainen. 2006. Host use of a hemiparasitic plant: No trade-offs in performance on different hosts. *Journal of Evolutionary Biology* 19: 513–521.
- Ainouche, M. L., P. M. Fortune, A. Salmon, C. Parisod, M. A. Grandbastien, K. Fukunaga, M. Ricou, and M. T. Misset. 2009. Hybridization, polyploidy and invasion: Lessons from *Spartina* (Poaceae). *Biological Invasions* 11: 1159–1173.

- Alix, K., P. R. Gerard, T. Schwarzacher, and J. S. Heslop-Harrison. 2017. Polyploidy and interspecific hybridization: Partners for adaptation, speciation and evolution in plants. *Annals of Botany* 120: 183–194.
- Allen, D. E. 1986. The botanists : A history of the Botanical Society of the British Isles through a hundred and fifty years. St. Paul's Bibliographies.
- Alves, M. J., M. M. Coelho, and M. J. Collares-Pereira. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: A genetic review. *Genetica* 111: 375–385.
- Ameloot, E., K. Verheyen, and M. Hermy. 2005a. Meta-analysis of standing crop reduction by *Rhinanthus* spp. And its effect on vegetation structure. *Folia Geobotanica* 40: 289–310.
- Ameloot, E., K. Verheyen, and M. Hermy. 2005b. Meta-analysis of standing crop reduction by *Rhinanthus* spp. And its effect on vegetation structure. *Folia Geobotanica* 40: 289–310.
- Anderson, B. M., K. R. Thiele, S. L. Krauss, and M. D. Barrett. 2017. Genotyping-by-Sequencing in a species complex of Australian hummock grasses (*Triodia*): Methodological insights and phylogenetic resolution. *Plos One* 12.
- Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160: 1217–1229.
- Arnold, B., K. Sang-Tae, and K. Bomblies. 2015. Single geographic origin of a widespread autotetraploid *Arabidopsis arenosa* lineage followed by interploidy admixture. *Molecular Biology and Evolution* 36: 1382–1395.
- Baduel, P., S. Bray, M. Vallejo-Marin, F. Kolar, and L. Yant. 2018. The "Polyploid Hop": Shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers in Ecology and Evolution* 6: 19.
- Bailey, J. 2013. The Japanese knotweed invasion viewed as a vast unintentional hybridisation experiment. *Heredity* 110: 105–110.
- Bailey, J. P., S. T. Bennett, M. D. Bennett, and C. A. Stace. 1993. Genomic in-situ hybridisation identifies parental chromosomes in the wild grass hybrid *x Festulpia hubbardii*. *Heredity* 71: 413–420.
- Bailey, J., and R. Wisskirchen. 2004. The distribution and origins of *Fallopia x bohemica* (Polygonaceae) in Europe. *Nordic Journal of Botany* 24: 173–199.
- Balao, F., M. Tannhauser, M. T. Lorenzo, M. Hedren, and O. Paun. 2016. Genetic differentiation and

- admixture between sibling allopolyploids in the *Dactylorhiza majalis* complex. *Heredity* 116: 351–361.
- Barber, I. 2005. Parasites grow larger in faster growing fish hosts. *International Journal for Parasitology* 35: 137–143.
- Bardy, K. E., P. Schonswetter, G. M. Schneeweiss, M. A. Fischer, and D. C. Albach. 2011. Extensive gene flow blurs species boundaries among *Veronica barrelieri*, *V. orchidea* and *V. spicata* (Plantaginaceae) in southeastern europe. *Taxon* 60: 108–121.
- Barker, M. S., N. Arrigo, A. E. Baniaga, Z. Li, and D. A. Levin. 2016. On the relative abundance of autoployploids and allopolyploids. *New Phytologist* 210: 391–398.
- Barnes, B. V., and B. P. Dancik. 1985. Characteristics and origin of a new birch species, *Betula murrayana* from southeastern Michigan. *Canadian Journal of Botany-Revue Canadienne De Botanique* 63: 223–226.
- Barringer, B. C. 2007. Polyploidy and self-fertilization in flowering plants. *American Journal of Botany* 94: 1527–1533.
- Barrow, L. N., S. M. McNew, N. Mitchell, S. C. Galen, H. L. Lutz, H. Skeen, T. Valqui, et al. 2019. Deeply conserved susceptibility in a multi-host, multi-parasite system. *Ecology Letters* 22: 987–998.
- Bates, D., M. Machler, B. M. Bolker, and S. C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.
- Becher, H., M. Brown, G. Powell, C. Methereill, N. Riddiford, and A. Twyford. 2020. Maintenance of species differences in closely related tetraploid parasitic *Euphrasia* (Orobanchaceae) on an isolated island. *Plant Communications*.
- Becker, H. 1986. Botany of European mistletoe (*Viscum album* L.). *Oncology* 43: 2–7.
- Becker, K., L. Tilley, J. L. Vennerstrom, D. Roberts, S. Rogerson, and H. Ginsburg. 2004. Oxidative stress in malaria parasite-infected erythrocytes: Host-parasite interactions. *International Journal for Parasitology* 34: 163–189.
- Becker, M., N. Gruenheit, M. Steel, C. Voelckel, O. Deusch, P. B. Heenan, P. A. McLenaghan, et al. 2013. Hybridization may facilitate in situ survival of endemic species through periods of climate change. *Nature Climate Change* 3: 1039–1043.
- Beddows, A. R. 1967. *Lolium perenne* l. *Journal of Ecology* 55: 567–587.

- Beddows, I., and L. E. Rose. 2018. Factors determining hybridization rate in plants: A case study in Michigan. *Perspectives in Plant Ecology Evolution and Systematics* 34: 51–60.
- Belton, G. S., W. F. P. van Reine, J. M. Huisman, S. G. A. Draisma, and C. F. D. Gurgel. 2014. Resolving phenotypic plasticity and species designation in the morphologically challenging *Caulerpa racemosa/peltata* complex (Caulerpaceae, Chlorophyta). *Journal of Phycology* 50: 32–54.
- Bertioli, D. J., S. B. Cannon, L. Froenicke, G. D. Huang, A. D. Farmer, E. K. S. Cannon, X. Liu, et al. 2016. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics* 48: 438–446.
- Bittencourt, N. S. 2019. Reproductive systems and low outbreeding barriers between *Jacaranda cuspidifolia* and *J. mimosifolia* (Jacarandeae, Bignoniaceae). *Nordic Journal of Botany* 37: 12.
- Bize, P., C. Jeanneret, A. Klopfenstein, and A. Roulin. 2008. What makes a host profitable? Parasites balance host nutritive resources against immunity. *American Naturalist* 171: 107–118.
- Bleeker, W. 2003. Hybridization and *Rorippa austriaca* (Brassicaceae) invasion in Germany. *Molecular Ecology* 12: 1831–1841.
- Bleeker, W. 2007. Interspecific hybridization in *Rorippa* (Brassicaceae): Patterns and processes. *Systematics and Biodiversity* 5: 311–319.
- Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J. S. S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution* 24: 127–135.
- Bomblies, K., and D. Weigel. 2007. Hybrid necrosis: Autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics* 8: 382–393.
- Bougeard, S., and S. Dray. 2018. Supervised multiblock analysis in R with the ade4 package. *Journal of Statistical Software* 86: 1–17.
- Bouwmeester, H. J., C. Roux, J. A. Lopez-Raez, and G. Becard. 2007. Rhizosphere communication of plants, parasitic plants and AM fungi. *Trends in Plant Science* 12: 224–230.
- Breese, E. 1959. Selection for differing degrees of out-breeding in *Nicotiana rustica*. *Annals of Botany* 23: 331–344.

- Brennan, A. C., G. Woodward, O. Seehausen, V. Munoz-Fuentes, C. Moritz, A. Guelmami, R. J. Abbott, and P. Edelaar. 2014. Hybridization due to changing species distributions: Adding problems or solutions to conservation of biodiversity during global change? *Evolutionary Ecology Research* 16: 475–491.
- Bretagnolle, F., and J. D. Thompson. 1995. Tansley Review No. 78 - Gametes with the somatic chromosome number, mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist* 129: 1–22.
- Brochmann, C. 1984. Hybridization and distribution of *Argyranthemum coronopifolium* (Asteraceae, Anthemideae) in the canary islands. *Nordic Journal of Botany* 4: 729–736.
- Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A. C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* 82: 521–536.
- Brochmann, C., P. S. Soltis, and D. E. Soltis. 1992. Recurrent formation and polyphyly of nordic polyploids in *Draba* (Brassicaceae). *American Journal of Botany* 79: 673–688.
- Brochmann, C., B. Stedje, and L. Borgen. 1992. Gene flow across ploidal levels in *Draba* (Brassicaceae). *Evolutionary Trends in Plants* 6: 125–134.
- Brown, J. H. 1984. On the relationship between abundance and distribution of species. *American Naturalist* 124: 255–279.
- Brown, M. 2020. LwPhylo: A lightweight, low level javascript library to plot phylogenies from a newick file, version 1.0.4.
- Brown, M. 2019. VCVglmm: Manipulating model outputs from MCMCglmm and lme4 objects, usually involving variance-covariance matrices.
- Brown, M., N. Frachon, E. Wong, C. Metherell, and A. Twyford. 2020. Life history evolution, species differences, and phenotypic plasticity in hemiparasitic eyebrights (*Euphrasia*). *American Journal of Botany* 107: 1–10.
- Brys, R., A. Vanden Broeck, J. Mergeay, and H. Jacquemyn. 2014. The contribution of mating system variation to reproductive isolation in two closely related *Centaurium* species (Gentianaceae) with a generalized flower morphology. *Evolution* 68: 1281–1293.
- Brysting, A. K., B. Oxelman, K. T. Huber, V. Moulton, and C. Brochmann. 2007. Untangling complex histories of genome mergings in high polyploids. *Systematic Biology* 56: 467–476.

- Buggs, R. J. A., and J. R. Pannell. 2007. Ecological differentiation and diploid superiority across a moving ploidy contact zone. *Evolution* 61: 125–140.
- Burton, T. L., and B. C. Husband. 2000. Fitness differences among diploids, tetraploids, and their triploid progeny in *Chamerion angustifolium*: Mechanisms of inviability and implications for polyploid evolution. *Evolution* 54: 1182–1191.
- Butlin, R. K., and S. Stankowski. 2020. Is it time to abandon the biological species concept? No. *National Science Review* 7: 1400–1401.
- Cameron, D. D., A. M. Coats, and W. E. Seel. 2006. Differential resistance among host and non-host species underlies the variable success of the hemi-parasitic plant *Rhinanthus minor*. *Annals of Botany* 98: 1289–1299.
- Campbell, C. S., and W. A. Wright. 1996. Apomixis, hybridization, and taxonomic complexity in eastern North American *Amelanchier* (Rosaceae). *Folia Geobotanica & Phytotaxonomica* 31: 345–354.
- Carter, K. A., A. Liston, N. V. Bassil, L. A. Alice, J. M. Bushakra, B. L. Sutherland, T. C. Mockler, et al. 2019. Target capture sequencing unravels *Rubus* evolution. *Frontiers in Plant Science* 10.
- Chamberlain, S. 2019. Brranching: Fetch 'phylogenies' from many sources.
- Chapman, M. A., and R. J. Abbott. 2010. Introgression of fitness genes across a ploidy barrier. *New Phytologist* 186: 63–71.
- Chapman, M. A., and J. M. Burke. 2007. Genetic divergence and hybrid speciation. *Evolution* 61: 1773–1780.
- Chapman, T., S. Miles, and C. Trivedi. 2019. Capturing, protecting and restoring plant diversity in the UK: RBG Kew and the Millennium Seed Bank. *Plant Diversity* 41: 124–131.
- Charif, D., and J. Lobry. 2007. SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In U. Bastolla, M. Porto, H. Roman, and M. Vendruscolo [eds.], Structural approaches to sequence evolution, 207–232. Springer Verlag.
- Chen, Z. J. 2010. Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science* 15: 57–71.
- Clark, J. W., and P. C. J. Donoghue. 2018. Whole-genome duplication and plant macroevolution. *Trends in Plant Science* 23: 933–945.
- Clark, N. F., J. A. McComb, and A. W. Taylor-Robinson. 2020. Host species of mistletoes (Loranthaceae and Viscaceae) in Australia. *Australian Journal of Botany* 68: 1–13.

- Cleland, R. E. 1972. *Oenothera* : Cytogenetics and evolution. Academic Press, London.
- Cleland, R. E. 1944. The problem of species in *Oenothera*. *American Naturalist* 78: 5–28.
- Crespo-López, M. E., I. Pala, T. L. Duarte, T. E. Dowling, and M. M. Coelho. 2007. Genetic structure of the diploid–polyploid fish *Squalius alburnoides* in southern Iberian basins Tejo and Guadiana, based on microsatellites. *Journal of Fish Biology* 71: 423–436.
- Crispo, E., J. S. Moore, J. A. Lee-Yaw, S. M. Gray, and B. C. Haller. 2011. Broken barriers: Human-induced changes to gene flow and introgression in animals. *Bioessays* 33: 508–518.
- Czapik, R. 1996. Problems of apomictic reproduction in the families Compositae and Rosaceae. *Folia Geobotanica & Phytotaxonomica* 31: 381–387.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.
- De Hert, K., H. Jacquemyn, S. Van Glabeke, I. Roldan-Ruiz, K. Vandepitte, L. Leus, and O. Honnay. 2011. Patterns of hybridisation between diploid and derived allotetraploid species of *Dactylorhiza* (Orchidaceae) co-occurring in Belgium. *American Journal of Botany* 98: 946–955.
- De Hert, K., H. Jacquemyn, S. Van Glabeke, I. Roldan-Ruiz, K. Vandepitte, L. Leus, and O. Honnay. 2012. Reproductive isolation and hybridization in sympatric populations of three *Dactylorhiza* species (Orchidaceae) with different ploidy levels. *Annals of Botany* 109: 709–720.
- Delavault, P., E. Estabrook, H. Albrecht, R. Wrobel, and J. I. Yoder. 1998. Host-root exudates increase gene expression of asparagine synthetase in the roots of a hemiparasitic plant *Triphysaria versicolor* (Scrophulariaceae). *Gene* 222: 155–162.
- Denda, T., and M. Yokota. 2003. Hybrid origins of *Ixeris nakazonei* (Asteraceae, Lactuceae) in the Ryukyu Archipelago, Japan: Evidence from molecular data. *Botanical Journal of the Linnean Society* 141: 379–387.
- Dickinson, T. A., E. Lo, and N. Talent. 2007. Polyploidy, reproductive biology, and rosaceae: Understanding evolution and making classifications. *Plant Systematics and Evolution* 266: 59–78.
- Dowle, M., and A. Srinivasan. 2019. Data.table: Extension of ‘data.frame’.
- Dupuis, J. R., B. M. T. Brunet, H. M. Bird, L. M. Lumley, G. Fagua, B. Boyle, R. Levesque, et al. 2017. Genome-wide SNPs resolve phylogenetic relationships in the North American spruce budworm (*Choristoneura*

- fumiferana*) species complex. *Molecular Phylogenetics and Evolution* 111: 158–168.
- Edmands, S. 2002. Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution* 17: 520–527.
- Elkington, T. T. 1984. Cytogenetic variation in the British flora - origins and significance. *New Phytologist* 98: 101–118.
- Ellis, M. B., and J. P. Ellis. 1985. Microfungi on land plants. An identification handbook. Croom Helm Ltd.
- Ellson, J., E. Gansner, L. Koutsofios, S. C. North, and G. Woodhull. 2002. Graphviz - open source graph drawing tools. *Graph Drawing* 2265: 483–484.
- Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* 97: 7043–7050.
- Ellstrand, N. C., R. Whitkus, and L. H. Rieseberg. 1996. Distribution of spontaneous plant hybrids. *Proceedings of the National Academy of Sciences of the United States of America* 93: 5090–5093.
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *Plos One* 6.
- Ennos, R. A., G. C. French, and P. M. Hollingsworth. 2005. Conserving taxonomic complexity. *Trends in Ecology & Evolution* 20: 164–168.
- Ennos, R. A., R. Whitlock, M. F. Fay, B. Jones, L. E. Neaves, R. Payne, I. Taylor, et al. 2012. Process-based species action plans: An approach to conserve contemporary evolutionary processes that sustain diversity in taxonomically complex groups. *Botanical Journal of the Linnean Society* 168: 194–203.
- Evans, A. M. 1962. Species hybridization in *Trifolium* 1. Methods of overcoming species incompatibility. *Euphytica* 11: 164 –&.
- Fearn, G. M. 1977. Morphological and cytological investigation of *Cochlearia populations* on the Gower peninsula, Glamorgan. *New Phytologist* 79: 455–458.
- Federici, S., A. Galimberti, F. Bartolucci, I. Bruni, F. De Mattia, P. Cortis, and M. Labra. 2013. DNA barcoding to analyse taxonomically complex groups in plants: The case of *Thymus* (Lamiaceae). *Botanical Journal of the Linnean Society* 171: 687–699.

- Felsenstein, J. 1976. Theoretical population genetics of variable selection and migration. *Annual Review of Genetics* 10: 253–280.
- Fernandes, G. W., E. A. de Mattos, A. C. Franco, U. Luttge, and H. Ziegler. 1998. Influence of the parasite *Pilostyles ingae* (Rafflesiaceae) on some physiological parameters of the host plant, *Mimosa naguirei* (Mimosaceae). *Botanica Acta* 111: 51–54.
- Fitter, A. H., and H. J. Peat. 1994. The ecological flora database. *Journal of Ecology* 82: 415–425.
- Fowler, N. L., and D. A. Levin. 2016. Critical factors in the establishment of allopolyploids. *American Journal of Botany* 103: 1236–1251.
- Fowler, N. L., and D. A. Levin. 1984. Ecological constraints on the establishment of a novel polyploid in competition with its diploid progenitors. *American Naturalist* 124: 703–711.
- Fox, J. 2020. RcmdrMisc: R commander miscellaneous functions.
- French, G. C., R. A. Ennos, A. J. Silverside, and P. M. Hollingsworth. 2005. The relationship between flower size, inbreeding coefficient and inferred selfing rate in British *Euphrasia* species. *Heredity* 94: 44–51.
- French, G. C., P. M. Hollingsworth, A. J. Silverside, and R. A. Ennos. 2008. Genetics, taxonomy and the conservation of British *Euphrasia*. *Conservation Genetics* 9: 1547–1562.
- Friedman, J., A. D. Twyford, J. H. Willis, and B. K. Blackman. 2015. The extent and genetic basis of phenotypic divergence in life history traits in *Mimulus guttatus*. *Molecular Ecology* 24: 111–122.
- Gandon, S. 2002. Local adaptation and the geometry of host-parasite coevolution. *Ecology Letters* 5: 246–256.
- Garcia-Melo, J. E., C. Oliveira, G. J. D. Silva, L. E. Ochoa-Orrego, L. H. G. Pereira, and J. A. Maldonado-Ocampo. 2019. Species delimitation of neotropical Characins (Stevardiinae): Implications for taxonomy of complex groups. *Plos One* 14.
- Garg, A. 2010. Adaptive responses underpinning speciation in *Pedicularis* L. (Scrophulariaceae). *Current Science* 99: 733–736.
- Garnier, E. 1992. Growth analysis of congeneric annual and perennial grass species. *Journal of Ecology* 80: 665–675.
- Garnier, E., and S. Vancaeyzeele. 1994. Carbon and nitrogen content of congeneric annual and perennial grass species - relationships with growth. *Plant Cell and Environment* 17: 399–407.

- Gibot-Leclerc, S., F. Dessaint, C. Reibel, and V. Le Corre. 2013. *Phelipanche ramosa* (L.) Pomel populations differ in life-history and infection response to hosts. *Flora* 208: 247–252.
- Gibson, C. C., and A. R. Watkinson. 1989. The host range and selectivity of a parasitic plant: *Rhinanthus minor* L. *Oecologica* 78: 401–406.
- Gill, J. J. B., H. A. McAllister, and G. M. Fearn. 1978. Cytotaxonomic studies on the *Cochlearia officinalis* L. group from inland stations in Britain. *Watsonia* 12: 15–21.
- Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire, Q. Sun, and E. S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *Plos One* 9: 11.
- Gomez-Rodriguez, C., A. Crampton-Platt, M. Timmermans, A. Baselga, and A. P. Vogler. 2015. Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution* 6: 883–894.
- Grant, V. 1981. Plant speciation. Columbia University Press.
- Gross, B. L., and L. H. Rieseberg. 2005. The ecological genetics of homoploid hybrid speciation. *Journal of Heredity* 96: 241–252.
- Gu, Z. G., L. Gu, R. Eils, M. Schlesner, and B. Brors. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics* 30: 2811–2812.
- Guo, H., S. J. Mazer, and G. Z. Du. 2010a. Geographic variation in primary sex allocation per flower within and among 12 species of *Pedicularis* (Orobanchaceae): Proportional male investment increases with elevation. *American Journal of Botany* 97: 1334–1341.
- Guo, H., S. J. Mazer, and G. Z. Du. 2010b. Geographic variation in seed mass within and among nine species of *Pedicularis* (Orobanchaceae): Effects of elevation, plant size and seed number per fruit. *Journal of Ecology* 98: 1232–1242.
- Guo, Q. F. 2014. Plant hybridization: The role of human disturbance and biological invasion. *Diversity and Distributions* 20: 1345–1354.
- Guo, Y. P., J. Saukel, R. Mittermayr, and F. Ehrendorfer. 2005. AFLP analyses demonstrate genetic divergence, hybridization, and multiple polyploidization in the evolution of *Achillea* (Asteraceae-Anthemideae). *New Phytologist* 166: 273–289.

- Gussarova, G., M. Popp, E. Vitek, and C. Brochmann. 2008. Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): Recent radiations in an old genus. *Molecular Phylogenetics and Evolution* 48: 444–460.
- Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *Plos Genetics* 5: 11.
- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software* 33: 1–22.
- Hall, M. C., D. B. Lowry, and J. H. Willis. 2010. Is local adaptation in *Mimulus guttatus* caused by trade-offs at individual loci? *Molecular Ecology* 19: 2739–2753.
- Han, T. S., Q. J. Zheng, R. E. Onstein, B. M. Rojas-Andres, F. Hauenschild, A. N. Muellner-Riehl, and Y. W. Xing. 2020. Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist* 225: 571–583.
- Hedren, M. 2003. Plastid DNA variation in the *Dactylorhiza incarnata/maculata* polyploid complex and the origin of allotetraploid *D. sphagnicola* (Orchidaceae). *Molecular Ecology* 12: 2669–2680.
- Hegarty, M. J., and S. J. Hiscock. 2005. Hybrid speciation in plants: New insights from molecular studies. *New Phytologist* 165: 411–423.
- Hegde, S. G., J. D. Nason, J. M. Clegg, and N. C. Ellstrand. 2006. The evolution of California's wild radish has resulted in the extinction of its progenitors. *Evolution* 60: 1187–1197.
- Heinze, B. 2011. Towards a quantitative description of landscape, demography and flowering phenology effects on realized hybridization potential. *Molecular Ecology* 20: 2233–2235.
- Hellstrom, K., P. Rautio, A. P. Huhta, and J. Tuomi. 2004. Tolerance of an annual hemiparasite, *Euphrasia stricta* agg., to simulated grazing in relation to the host environment. *Flora* 199: 247–255.
- Hersh, E., J. Grimm, and J. Whitton. 2016. Attack of the clones: Reproductive interference between sexuals and asexuals in the *Crepis agamic* complex. *Ecology and Evolution* 6: 6473–6483.
- Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.
- Hoffman, A., J. Plötner, N. B. Pruvost, D. G. Christiansen, S. Röthlisberger, L. Choleva, P. Mikulicek, et

- al. 2015. Genetic diversity and distribution patterns of diploid and polyploid hybrid water frog populations (*Pelophylax esculentus* complex) across europe. *Molecular Ecology* 24: 4371–4391.
- Hollingsworth, M. L., J. P. Bailey, P. M. Hollingsworth, and C. Ferris. 1999. Chloroplast DNA variation and hybridization between invasive populations of Japanese knotweed and giant knotweed (*Fallopia*, Polygonaceae). *Botanical Journal of the Linnean Society* 129: 139–154.
- Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, S. Ratnasingham, M. van der Bank, M. W. Chase, et al. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 12794–12797.
- Hollingsworth, P. M., L. E. Neaves, and A. D. Twyford. 2017. Using DNA sequence data to enhance understanding and conservation of plant diversity at the species level. In S. Blackmore, and S. Oldfield [eds.], *Plant conservation science and practice : The role of botanic gardens*, Cambridge University Press, Cambridge.
- Hollingsworth, P. M., J. Squirrell, M. L. Hollingsworth, A. Richards, and R. M. Bateman. 2006. Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis*. *Current taxonomic research on the British & European flora, BSBI*: 27–44.
- Hollister, J. D., B. J. Arnold, E. Svedin, K. S. Xue, B. P. Dilkes, and K. Bomblies. 2012. Genetic adaptation associated with genome-doubling in autotetraploid *Arabidopsis arenosa*. *PLOS Genetics* 8.
- Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50: 346–363.
- Huang, K., R. Whitlock, M. C. Press, and J. D. Scholes. 2012. Variation for host range within and among populations of the parasitic plant *Striga hermonthica*. *Heredity* 108: 96–104.
- Hunter, J. D. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9: 90–95.
- Husband, B. C., S. J. Baldwin, and J. Suda. 2013. The incidence of polyploidy in natural plant populations: Major patterns and evolutionary processes. In G. J. D. J, and W. J [eds.], *Plant genome diversity*, Springer, Vienna.
- Husband, B. C., and H. A. Sabara. 2004. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist* 161: 703–713.

- Irwin, J. A., and R. J. Abbott. 1992. Morphometric and isozyme evidence for the hybrid origin of a new tetraploid radiate groundsel in York, England. *Heredity* 69: 431–439.
- Jay, P., A. Whibley, L. Frezal, M. A. R. de Cara, R. W. Nowell, J. Mallet, K. K. Dasmahapatra, and M. Joron. 2018. Supergene evolution triggered by the introgression of a chromosomal inversion. *Current Biology* 28: 1839–1845.
- Joel, J. M., J. Gressel, and L. J. Musselman. 2013. Parasitic Orobanchaceae. Springer, Berlin, New York.
- Johnson, S. D. 2018. Natural hybridization in the orchid flora of South Africa: Comparisons among genera and floristic regions. *South African Journal of Botany* 118: 290–298.
- Johnston, S. A., T. P. M. Dennijs, S. J. Peloquin, and R. E. Hanneman. 1980. The significance of genic balance to endosperm development in interspecific crosses. *Theoretical and Applied Genetics* 57: 5–9.
- Jombart, T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Jones, D. A., and R. Turkington. 1986. Biological flora of the British Isles - *Lotus corniculatus* L. *Journal of Ecology* 74: 1185–1212.
- Jones, L., A. D. Twyford, C. R. Ford, T. C. G. Rich, H. Davies, L. L. Forrest, M. L. Hart, et al. Barcode UK: a DNA barcoding resource for the flowering plants and conifers of the United Kingdom. *Molecular Ecology Resources*.
- Jonstrup, A., M. Hedren, and S. Andersson. 2016. Host environment and local genetic adaptation determine phenotype in parasitic *Rhinanthus angustifolius*. *Botanical Journal of the Linnean Society* 180: 89–103.
- Kamvar, Z. N., J. F. Tabima, and N. J. Grunwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: 14.
- Karron, J. D., R. T. Jackson, N. N. Thumser, and S. L. Schlicht. 1997. Outcrossing rates of individual *Mimulus ringens* genets are correlated with anther-stigma separation. *Heredity* 79: 365–370.
- Katoh, K., and D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30: 772–780.
- Kay, K. M. 2006. Reproductive isolation between two closely related hummingbird-pollinated neotropical gingers. *Evolution* 60: 538–552.

- Keith, A. M., D. D. Cameron, and W. E. Seel. 2004. Spatial interactions between the hemiparasitic angiosperm *Rhinanthus minor* and its host are species-specific. *Functional Ecology* 18: 435–442.
- Kellner, A., C. M. Ritz, and V. Wissemann. 2012. Hybridization with invasive *Rosa rugosa* threatens the genetic integrity of native *Rosa mollis*. *Botanical Journal of the Linnean Society* 170: 472–484.
- Kelly, C. K., D. L. Venable, and K. Zimmerer. 1988. Host specialization in *Cuscuta costariensis* - an assessment of host use relative to host availability. *Oikos* 53: 315–320.
- Kim, M., M. L. Cui, P. Cubas, A. Gillies, K. Lee, M. A. Chapman, R. J. Abbott, and E. Coen. 2008. Regulatory genes control a key morphological and ecological trait transferred between species. *Science* 322: 1116–1119.
- Knaus, B. J., and N. J. Grunwald. 2017. VCFR: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* 17: 44–53.
- Koch, L. 1959. Zur Entwicklungsgeschichte der Rhinanthaceen (ii *Euphrasia officinalis* L.). *Jahrbücher für Wissenschaftliche Botanik* 22: 1–34.
- Kohler, C., O. M. Scheid, and A. Erilova. 2010. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics* 26: 142–148.
- Kolar, F., M. Certner, J. Suda, P. Schonswetter, and B. C. Husband. 2017. Mixed-ploidy species: Progress and opportunities in polyploid research. *Trends in Plant Science* 22: 1041–1055.
- Kolar, F., S. Pisova, E. Zaveska, T. Fer, M. Weiser, F. Ehrendorfer, and J. Suda. 2015. The origin of unique diversity in deglaciated areas: Traces of pleistocene processes in north-european endemics from the *Galium pusillum* polyploid complex (Rubiaceae). *Molecular Ecology* 24: 1311–1334.
- Koutecky, P., T. Badurova, M. Stech, J. Kosnar, and J. Karasek. 2011. Hybridization between diploid *Centaurea pseudophrygia* and tetraploid *C. jacea* (Asteraceae): The role of mixed pollination, unreduced gametes, and mentor effects. *Biological Journal of the Linnean Society* 104: 93–106.
- Kowal, R. R., E. J. Judziewicz, and J. Edwards. 2011. *Packera insulae-regalis* (Asteraceae, Senecioneae), a new species endemic to Isle Royale, Michigan, USA. *Brittonia* 63: 343–354.
- Kramina, T. E., I. G. Meschersky, G. V. Degtjareva, T. H. Samigullin, Y. S. Belokon, and I. A. Schanzer. 2018. Genetic variation in the *Lotus corniculatus* complex (Fabaceae) in northern Eurasia as inferred from nuclear microsatellites and plastid *trnL-trnF* sequences. *Botanical Journal of the Linnean Society* 188: 87–116.

- Krasnov, B. R., S. Morand, D. Mouillot, G. I. Shenbrot, I. S. Khokhlova, and R. Poulin. 2006. Resource predictability and host specificity in fleas: The effect of host body mass. *Parasitology* 133: 81–88.
- Kreiner, J. M., P. Kron, and B. C. Husband. 2017a. Evolutionary dynamics of unreduced gametes. *Trends in Genetics* 33: 583–593.
- Kreiner, J. M., P. Kron, and B. C. Husband. 2017b. Frequency and maintenance of unreduced gametes in natural plant populations: Associations with reproductive mode, life history and genome size. *New Phytologist* 214: 879–889.
- Kuzmina, M. L., T. W. A. Braukmann, A. J. Fazekas, S. W. Graham, S. L. Dewaard, A. Rodrigues, B. A. Bennett, et al. 2017. Using herbarium-derived DNAs to assemble a large-scale DNA barcode library for the vascular plants of Canada. *Applications in Plant Sciences* 5: 1700079.
- Lack, A. J. 1995. Relationships and hybridisation between British species of *Polygala* - evidence from isozymes. *New Phytologist* 130: 217–223.
- Lafon-Placette, C., and C. Kohler. 2016. Endosperm-based postzygotic hybridization barriers: Developmental mechanisms and evolutionary drivers. *Molecular Ecology* 25: 2620–2629.
- Lammi, A., P. Siikamaki, and V. Salonen. 1999. The role of local adaptation in the relationship between an endangered root hemiparasite *Euphrasia rostkoviana*, and its host, *Agrostis capillaris*. *Ecography* 22: 145–152.
- Lamont, B. B., T. He, N. J. Enright, S. L. Krauss, and B. P. Miller. 2003. Anthropogenic disturbance promotes hybridization between *Banksia* species by altering their biology. *Journal of Evolutionary Biology* 16: 551–557.
- Laport, R. G., R. L. Minckley, and J. Ramsey. 2016. Ecological distributions, phenological isolation, and genetic structure in sympatric and parapatric populations of the *Larrea tridentata* polyploid complex. *American Journal of Botany* 103: 1358–1374.
- Larson, E. L., R. M. Tinghitella, and S. A. Taylor. 2019. Insect hybridization and climate change. *Frontiers in Ecology and Evolution* 7: 11.
- Lawrence, W. J. C. 1936. The origin of new forms in *Delphinium*. *Genetica* 18: 109–115.
- Leavitt, S. D., L. A. Johnson, T. Goward, and L. L. St Clair. 2011. Species delimitation in taxonomically difficult lichen-forming fungi: An example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. *Molecular Phylogenetics and Evolution* 60: 317–332.

- Leggett, H. C., A. Buckling, G. H. Long, and M. Boots. 2013. Generalism and the evolution of parasite virulence. *Trends in Ecology & Evolution* 28: 592–596.
- Lenth, R. 2020. Emmeans: Estimated marginal means, aka least-squares means.
- Levin, D. 1975. Minority cytotype exclusion in local plant populations. *Taxon* 24: 35–43.
- Lexer, C., M. E. Welch, J. L. Durphy, and L. H. Rieseberg. 2003. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Molecular Ecology* 12: 1225–1235.
- Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.
- Liebst, B. 2008. Do they really hybridize? A field study in artificially established mixed populations of *Euphrasia minima* and *E. salisburgensis* (Orobanchaceae) in the Swiss Alps. *Plant Systematics and Evolution* 273: 179–189.
- Liebst, B., and J. Schneller. 2008. Seed dormancy and germination behaviour in two *Euphrasia* species (Orobanchaceae) occurring in the Swiss Alps. *Botanical Journal of the Linnean Society* 156: 649–656.
- Lievens, E. J. P., J. Perreau, P. Agnew, Y. Michalakis, and T. Lenormand. 2018. Decomposing parasite fitness reveals the basis of specialization in a two-host, two-parasite system. *Evolution Letters* 2: 390–405.
- Lihova, J., J. F. Aguilar, K. Marhold, and G. N. Feliner. 2004. Origin of the disjunct tetraploid *Cardamine amporitana* (Brassicaceae) assessed with nuclear and chloroplast DNA sequence data. *American Journal of Botany* 91: 1231–1242.
- Lim, J., M. J. Crawley, N. De Vere, T. Rich, and V. Savolainen. 2014. A phylogenetic analysis of the British flora sheds light on the evolutionary and ecological factors driving plant invasions. *Ecology and Evolution* 4: 4258–4269.
- Lim, K. Y., G. Werlemark, R. Matyasek, J. B. Bringloe, V. Sieber, H. El Mokadem, J. Meynet, et al. 2005. Evolutionary implications of permanent odd polyploidy in the stable sexual, pentaploid of *Rosa canina* L. *Heredity* 94: 501–506.
- Lischer, H. E. L., and L. Excoffier. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.

- Love, A., and D. Love. 1943. The significance of differences in the distribution of diploids and polyploids. *Hereditas* 29: 145–163.
- Lowe, A., S. Harris, and P. Ashton. 2004. Ecological genetics : Design, analysis, and application. Blackwell Pub., Malden, Mass. ; Oxford.
- Lowe, A. J., and R. J. Abbott. 2004. Reproductive isolation of a new hybrid species, *Senecio eboracensis* Abbott & Lowe (Asteraceae). *Heredity* 92: 386–395.
- Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363: 3009–3021.
- Ludwig, S., A. Robertson, T. C. G. Rich, M. Djordjevic, R. Cerovic, L. Houston, S. A. Harris, and S. J. Hiscock. 2013. Breeding systems, hybridization and continuing evolution in avon gorge *Sorbus*. *Annals of Botany* 111: 563–575.
- Lumaret, R., and E. Barrientos. 1990. Phylogenetic relationships and gene flow between sympatric diploid and tetraploid plants of *Dactylis glomerata* (Gramineae). *Plant Systematics and Evolution* 169: 81–96.
- Luo, A., C. Ling, S. Y. W. Ho, and C. D. Zhu. 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology* 67: 830–846.
- Luo, Y. H., and A. Widmer. 2013. Herkogamy and its effects on mating patterns in *Arabidopsis thaliana*. *Plos One* 8.
- Mackinnon, M. J., and K. Marsh. 2010. The selection landscape of malaria parasites. *Science* 328: 866–871.
- Mallet, J. 2007. Hybrid speciation. *Nature* 446: 279–283.
- Mandakova, T., A. Kovarik, J. Zozomova-Lihova, R. Shimizu-Inatsugi, K. K. Shimizu, K. Mummenhoff, K. Marhold, and M. A. Lysak. 2013. The more the merrier: Recent hybridization and polyploidy in *Cardamine*. *Plant Cell* 25: 3280–3295.
- Manton, I. 1950. Problems of cytology and evolution in the Pteridophyta. Cambridge University Press.
- Marques, I., J. Loureiro, D. Draper, M. Castro, and S. Castro. 2018. How much do we know about the frequency of hybridisation and polyploidy in the Mediterranean region? *Plant Biology* 20: 21–37.

- Martin, M. D., and T. C. Mendelson. 2018. Hybrid sterility increases with genetic distance in snubnose darters (Percidae: Etheostoma). *Environmental Biology of Fishes* 101: 215–221.
- Martin, N. H., and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61: 68–82.
- Mason, A. S., and J. C. Pires. 2015. Unreduced gametes: Meiotic mishap or evolutionary mechanism? *Trends in Genetics* 31: 5–10.
- Matthies, D. 1998. Influence of the host on growth and biomass allocation in the two facultative root hemiparasites *Odontites vulgaris* and *Euphrasia minima*. *Flora* 193: 187–193.
- Matthies, D. 2017. Interactions between a root hemiparasite and 27 different hosts: Growth, biomass allocation and plant architecture. *Perspectives in Plant Ecology Evolution and Systematics* 24: 118–137.
- Mayr, E. 1963. Animal species and evolution. Belknap Press of Harvard University Press, Cambridge,
- Metherell, C., and F. Rumsey. 2018. Eyebrights (*Euphrasia*) of the UK and Ireland. Botanical Society of Britain; Ireland.
- Meyer-Lucht, Y., and S. Sommer. 2005. MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). *Molecular Ecology* 14: 2233–2243.
- Migdalek, G., J. Nowak, M. Saluga, E. Cieslak, M. Szczepaniak, M. Ronikier, T. Marcussen, et al. 2017. No evidence of contemporary interploidy gene flow between the closely related European woodland violets *Viola reichenbachiana* and *V. riviniana* (sect. *Viola*, Violaceae). *Plant Biology* 19: 542–551.
- Milne, R. I., S. Terzioglu, and R. J. Abbott. 2003. A hybrid zone dominated by fertile F(1)s: maintenance of species barriers in *Rhododendron*. *Molecular Ecology* 12: 2719–2729.
- Mitchell, N., L. G. Campbell, J. R. Ahern, K. C. Paine, A. B. Giroldo, and K. D. Whitney. 2019. Correlates of hybridization in plants. *Evolution Letters*: 16.
- Mitchell, R. J., J. D. Karron, K. G. Holmquist, and J. M. Bell. 2004. The influence of *Mimulus ringens* floral display size on pollinator visitation patterns. *Functional Ecology* 18: 116–124.
- Moghe, G. D., and S. H. Shiu. 2014. The causes and molecular consequences of polyploidy in flowering plants. In C. W. Fox, and T. A. Mousseau [eds.], Year in evolutionary biology, Annals of the New York Academy of Sciences, 16–34. Blackwell Science Publ, Oxford.

Mojica, J. P., Y. W. Lee, J. H. Willis, and J. K. Kelly. 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. *Molecular Ecology* 21: 3718–3728.

Monnahan, P., F. Kolar, P. Baduel, C. Sailer, J. Koch, R. Horvath, B. Laenen, et al. 2019. Pervasive population genomic consequences of genome duplication in *Arabidopsis arenosa*. *Nature Ecology & Evolution* 3: 457–+.

Moore, T. E., G. A. Verboom, and F. Forest. 2010. Phylogenetics and biogeography of the parasitic genus *Thesium* L. (Santalaceae), with an emphasis on the Cape of South Africa. *Botanical Journal of the Linnean Society* 162: 435–452.

Morgan, C., H. K. Zhang, C. E. Henry, F. C. H. Franklin, and K. Bomblies. 2020. Derived alleles of two axis proteins affect meiotic traits in autotetraploid *Arabidopsis arenosa*. *Proceedings of the National Academy of Sciences of the United States of America* 117: 8980–8988.

Motten, A. F., and J. L. Stone. 2000. Heritability of stigma position and the effect of stigma-anther separation on outcrossing in a predominantly self-fertilizing weed, *Datura stramonium* (Solanaceae). *American Journal of Botany* 87: 339–347.

Moyle, L. C., M. S. Olson, and P. Tiffin. 2004. Patterns of reproductive isolation in three angiosperm genera. *Evolution* 58: 2110–2110.

Mudrak, O., and J. Leps. 2010. Interactions of the hemiparasitic species *Rhinanthus minor* with its host plant community at two nutrient levels. *Folia Geobotanica* 45: 407–424.

Musselman, L. J. 1980. The biology of *Striga*, *Orobanche*, and other root-parasite weeds. *Annual Review of Phytopathology* 18: 463–489.

Nason, J. D., N. C. Ellstrand, and M. L. Arnold. 1992. Patterns of hybridisation and introgression in populations of oaks, manzanitas, and irises. *American Journal of Botany* 79: 101–111.

Nelson-Jones, E., D. Briggs, and A. Smith. 2002. The origin of intermediate species of the genus *Sorbus*. *Theoretical and Applied Genetics* 105: 953–963.

Neuffer, B., H. Auge, H. Mesch, U. Amarell, and R. Brandl. 1999. Spread of violets in polluted pine forests: Morphological and molecular evidence for the ecological importance of interspecific hybridization. *Molecular Ecology* 8: 365–377.

Nguyen, L. T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.

Nickrent, D. L. 2020. Parasitic angiosperms: How often and how many? *Taxon* 69: 5–27.

Nickrent, D. L., and L. J. Musselman. 2017. Parasitic plants. In B. H. Ownley, and R. N. Trigiano [eds.], *Plant pathology: Concepts and laboratory exercises*, CRC Press.

Novak, S. J., D. E. Soltis, and P. S. Soltis. 1991. Ownby's *Tragopogons* - 40 years later. *American Journal of Botany* 78: 1586–1600.

Novikova, P. Y., I. G. Brennan, W. Booker, M. Mahony, P. Doughty, A. R. Lemmon, E. M. Lemmon, et al. 2020. Polyploidy breaks speciation barriers in Australian burrowing frogs *Neobatrachus*. *Plos Genetics* 16.

Oberle, B., R. A. Montgomery, J. B. Beck, and E. J. Esselman. 2012. A morphologically intergrading population facilitates plastid introgression from diploid to tetraploid *Dodecatheon* (Primulaceae). *Botanical Journal of the Linnean Society* 168: 91–100.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437.

Palme, A. E., Q. Su, S. Palsson, and M. Lascoux. 2004. Extensive sharing of chloroplast haplotypes among European birches indicates hybridization among *Betula pendula*, *B. pubescens* and *B. nana*. *Molecular Ecology* 13: 167–178.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.

Paradis, E., and K. Schliep. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Parisod, C., R. Holderegger, and C. Brochmann. 2010. Evolutionary consequences of autopolyploidy. *New Phytologist* 186: 5–17.

Paule, J., N. D. Wagner, K. Weising, and G. Zizka. 2017. Ecological range shift in the polyploid members of the South American genus *Fosterella* (Bromeliaceae). *Annals of Botany* 120: 233–243.

Paun, O., F. Forest, M. F. Fay, and M. W. Chase. 2009. Hybrid speciation in angiosperms: Parental divergence drives ploidy. *New Phytologist* 182: 507–518.

- Pedersen, A. B., and A. Fenton. 2007. Emphasizing the ecology in parasite community ecology. *Trends in Ecology & Evolution* 22: 133–139.
- Pellicer, J., S. Clermont, L. Houston, T. C. G. Rich, and M. F. Fay. 2012. Cytotype diversity in the *Sorbus* complex (Rosaceae) in Britain: Sorting out the puzzle. *Annals of Botany* 110: 1185–1193.
- Petit, C., F. Bretagnolle, and F. Felber. 1999. Evolutionary consequences of diploid-polyplloid hybrid zones in wild species. *Trends in Ecology & Evolution* 14: 306–311.
- Pignone, D., and K. Hammer. 2016. Parasitic angiosperms as cultivated plants? *Genetic Resources and Crop Evolution* 63: 1273–1284.
- Pinheiro, F., F. de Barros, C. Palma-Silva, D. Meyer, M. F. Fay, R. M. Suzuki, C. Lexer, and S. Cozzolino. 2010. Hybridization and introgression across different ploidy levels in the Neotropical orchids *Epidendrum fulgens* and *E. puniceoluteum* (Orchidaceae). *Molecular Ecology* 19: 3981–3994.
- Popelka, O., B. Travnick, P. Sikova, M. Jandova, and M. Duchoslav. 2019. Natural hybridization between diploid *Ficaria calthifolia* and tetraploid *Ficaria verna* subsp. *verna* in central Europe: Evidence from morphology, ecology and life-history traits. *Preslia* 91: 179–212.
- Poulin, R., and M. R. Forbes. 2012. Meta-analysis and research on host-parasite interactions: Past and future. *Evolutionary Ecology* 26: 1169–1185.
- Prancl, J., Z. Kaplan, P. Travnick, and V. Jarolimova. 2014. Genome size as a key to evolutionary complex aquatic plants: Polyploidy and hybridization in *Callitrichie* (Plantaginaceae). *Plos One* 9: 15.
- Prentis, P. J., E. M. White, I. J. Radford, A. J. Lowe, and A. R. Clarke. 2007. Can hybridization cause local extinction: A case for demographic swamping of the Australian native *Senecio pinnatifolius* by the invasive *Senecio madagascariensis*? *New Phytologist* 176: 902–912.
- Preston, C. D., and D. A. Pearman. 2015. Plant hybrids in the wild: Evidence from biological recording. *Biological Journal of the Linnean Society* 115: 555–572.
- Price, P. W. 1980. Evolutionary biology of parasites. Princeton University Press, Princeton.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, et al. 2007.

PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559–575.

Pywell, R. F., J. M. Bullock, K. J. Walker, S. J. Coulson, S. J. Gregory, and M. J. Stevenson. 2004. Facilitating grassland diversification using the hemiparasitic plant *Rhinanthus minor*. *Journal of Applied Ecology* 41: 880–887.

Qiu, T., Z. Y. Liu, and B. Liu. 2020. The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports* 47: 5549–5558.

Raggi, L., E. Bitocchi, L. Russi, G. Marconi, T. F. Sharbel, F. Veronesi, and E. Albertini. 2015. Understanding genetic diversity and population structure of a *Poa pratensis* worldwide collection through morphological, nuclear and chloroplast diversity analysis. *Plos One* 10.

Ramsey, J., H. D. Bradshaw, and D. W. Schemske. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* 57.

Ramsey, J., and T. S. Ramsey. 2014. Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society B-Biological Sciences* 369.

Ramsey, J., and D. W. Schemske. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589–639.

Ramsey, J., and D. W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.

Ranallo-Benavidez, T. R., K. S. Jaron, and M. C. Schatz. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications* 11: 10.

Ravi, M., M. P. A. Marimuthu, and I. Siddiqi. 2008. Gamete formation without meiosis in *Arabidopsis*. *Nature* 451: 1121–1124.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ren, R., H. F. Wang, C. C. Guo, N. Zhang, L. P. Zeng, Y. M. Chen, H. Ma, and J. Qi. 2018. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Molecular Plant* 11: 414–428.

- Renny-Byfield, S., and J. F. Wendel. 2014. Doubling down on genomes: Polyploidy and crop plants. *American Journal of Botany* 101: 1711–1725.
- Rhymer, J. M., and D. Simberloff. 1996. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* 27: 83–109.
- Rice, A., L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel, et al. 2015. The Chromosome Counts Database (CCDB) - a community resource of plant chromosome numbers. *New Phytologist* 206: 19–26.
- Rice, A., P. Smarda, M. Novosolov, M. Drori, L. Glick, N. Sabath, S. Meiri, et al. 2019. The global biogeography of polyploid plants. *Nature Ecology & Evolution* 3: 265–273.
- Rieseberg, L. H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28: 359–389.
- Rieseberg, L. H., M. A. Archer, and R. K. Wayne. 1999. Transgressive segregation, adaptation and speciation. *Heredity* 83: 363–372.
- Rieseberg, L. H., and N. C. Ellstrand. 1993. What can molecular and morphological markers tell us about plant hybridisation? *Critical Reviews in Plant Sciences* 12: 213–241.
- Rieseberg, L. H., C. Vanfossen, and A. M. Desrochers. 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375: 313–316.
- Rieseberg, L. H., and J. H. Willis. 2007. Plant speciation. *Science* 317: 910–914.
- Rieseberg, L., T. Wood, and E. Baack. 2006. The nature of plant species. *Nature* 440: 524–527.
- Riihimäki, M., and O. Savolainen. 2004. Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). *American Journal of Botany* 91: 1036–1045.
- Roberts, H. F. 1929. Plant hybridization before Mendel. Princeton University Press.
- Robertson, A., T. C. G. Rich, A. M. Allen, L. Houston, C. Roberts, J. R. Bridle, S. A. Harris, and S. J. Hiscock. 2010. Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Molecular Ecology* 19: 1675–1690.
- Roccaforte, K., S. E. Russo, and D. Pilson. 2015. Hybridization and reproductive isolation between diploid *Erythronium mesochoreum* and its tetraploid congener *E. albidum* (Liliaceae). *Evolution* 69: 1375–1389.

- Roff, D. A. 2002. Life history evolution. Sinauer Associates, Sunderland, Mass.
- Roodt, R., and J. J. Spies. 2003. Chromosome studies in the grass subfamily Chloridoideae. II. An analysis of polyploidy. *Taxon* 52: 736–746.
- Roumet, C., C. Urcelay, and S. Diaz. 2006. Suites of root traits differ between annual and perennial species growing in the field. *New Phytologist* 170: 357–368.
- Rowley, G. 1967. Chromosome studies and evolution in *Rosa*. *Bulletin du Jardin botanique National de Belgique* 37: 45–52.
- Rowntree, J. K., D. F. Barham, A. J. A. Stewart, and S. E. Hartley. 2014. The effect of multiple host species on a keystone parasitic plant and its aphid herbivores. *Functional Ecology* 28: 829–836.
- Ruhsam, M., P. M. Hollingsworth, and R. A. Ennos. 2011. Early evolution in a hybrid swarm between outcrossing and selfing lineages in *Geum*. *Heredity* 107: 246–255.
- Ruhsam, M., P. M. Hollingsworth, and R. A. Ennos. 2013. Patterns of mating, generation of diversity, and fitness of offspring in a *Geum* hybrid swarm. *Evolution* 67: 2728–2740.
- Ruhsam, M., T. Jacobs, K. Watson, and P. M. Hollingsworth. 2015. Is hybridisation a threat to *Rumex aquaticus* in Britain? *Plant Ecology & Diversity* 8: 465–474.
- Rümer, S., D. D. Cameron, R. Wacker, W. Hartung, and F. Jiang. 2007. An anatomical study of the haustoria of *Rhinanthus minor* attached to roots of different hosts. *Flora-Morphology, Distribution, Functional Ecology of Plants* 202: 194–200.
- Salonen, V., H. Setala, and S. Puustinen. 2000. The interplay between *Pinus sylvestris*, its root hemiparasite, *Melampyrum pratense*, and ectomycorrhizal fungi: Influences on plant growth and reproduction. *Ecoscience* 7: 195–200.
- Sarkar, D. 2008. Lattice: Multivariate data visualization with R. Springer, New York.
- Schneeweiss, G. M. 2007. Correlated evolution of life history and host range in the nonphotosynthetic parasitic flowering plants *Orobanche* and *Phelipanche* (Orobanchaceae). *Journal of Evolutionary Biology* 20: 471–478.
- Scott, R., and A. Bolbol. 2013. Seed development in interploid hybrids. In J. Chen, and J. Birchler [eds.], *Polypliod and hybrid genomics*, 271–290. Wiley.
- Seel, W. E., R. E. Cooper, and M. C. Press. 1993. Growth, gas-exchange and water-use efficiency of the

facultative hemiparasite *Rhinanthus minor* associated with hosts differing in foliar nitrogen concentration. *Physiologia Plantarum* 89: 64–70.

Seel, W. E., and M. C. Press. 1994. Influence of the host on 3 sub-arctic annual facultative root hemiparasites 2. Gas-exchange characteristics and resource use-efficiency. *New Phytologist* 127: 37–44.

Segarra-Moragues, J. G., L. Villar, J. Lopez, E. Perez-Collazos, and P. Catalan. 2007. A new Pyrenean hybrid *Cirsium* (Asteraceae) as revealed by morphological and molecular analyses. *Botanical Journal of the Linnean Society* 154: 421–434.

Shimizu-Inatsugi, R., J. Lihova, H. Iwanaga, H. Kudoh, K. Marhold, O. Savolainen, K. Watanabe, et al. 2009. The allotetraploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*. *Molecular Ecology* 18: 4024–4048.

Shipunov, A. B., M. F. Fay, Y. Pillon, R. M. Bateman, and M. W. Chase. 2004. *Dactylorhiza* (Orchidaceae) in European Russia: Combined molecular and morphological analysis. *American Journal of Botany* 91: 1419–1426.

Silverside, A. 1990. A guide to eyebrights (*Euphrasia*) (ii) the *E. rostkoviana* group . *Wild Flower Magazine* 418: 31–34.

Simpson, G. G. 1951. The species concept. *Evolution* 5: 285–298.

Sochor, M., R. J. Vasut, T. F. Sharbel, and B. Travnicek. 2015. How just a few makes a lot: Speciation via reticulation and apomixis on example of European brambles (*Rubus* subgen. *Rubus*, Rosaceae). *Molecular Phylogenetics and Evolution* 89: 13–27.

Soltis, D. E., R. J. A. Buggs, J. J. Doyle, and P. S. Soltis. 2010. What we still don't know about polyploidy. *Taxon* 59: 1387–1403.

Soltis, D. E., P. S. Soltis, and J. A. Tate. 2004. Advances in the study of polyploidy since Plant Speciation. *New Phytologist* 161: 173–191.

Soltis, P. S., and D. E. Soltis. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588.

Spallek, T., M. Mutuku, and K. Shirasu. 2013. The genus *Striga*: A witch profile. *Molecular Plant Pathology* 14: 861–869.

- Spaniel, S., K. Marhold, B. Filova, and J. Zozomova-Lihova. 2011. Genetic and morphological variation in the diploid-polyplid *Alyssum montanum* in Central Europe: Taxonomic and evolutionary considerations. *Plant Systematics and Evolution* 294: 1–25.
- Spoelhof, J. P., R. Keeffe, and S. F. McDaniel. 2020. Does reproductive assurance explain the incidence of polyploidy in plants and animals? *New Phytologist* 227: 14–21.
- Spoelhof, J. P., P. S. Soltis, and D. E. Soltis. 2017. Pure polyploidy: Closing the gaps in autoploid research. *Journal of Systematics and Evolution* 55: 340–352.
- Stace, C. 1975. Hybridization and the flora of the British Isles. Academic Press.
- Stace, C. 2019. New flora of the British Isles. 4th ed. C & M Floristics.
- Stace, C. A., and M. J. Crawley. 2015. Alien plants. William Collins.
- Stace, C. A., C. D. Preston, and D. Pearman. 2015. Hybrid flora of the British Isles. Botanical Society of Britain & Ireland, Bristol.
- Stearns, S. C. 1992. The evolution of life histories. Oxford University Press.
- Stebbins, G. L. 1971. Chromosomal evolution in higher plants. Edward Arnold, London, UK.
- Stebbins, G. L. 1956. Cytogenetics and evolution of the grass family. *American Journal of Botany* 43: 890–905.
- Stebbins, G. L. 1959. The role of hybridization in evolution. *Proceedings of the American Philosophical Society* 103: 231–251.
- Stelkens, R., and O. Seehausen. 2009. Genetic distance between species predicts novel trait expression in their hybrids. *Evolution* 63: 884–897.
- Stenberg, P., M. Lundmark, S. Knutelski, and A. Saura. 2003. Evolution of clonality and polyploidy in a weevil system. *Molecular Biology and Evolution* 20: 1626–1632.
- Stephenson, A. G. 1981. Flower and fruit abortion: Proximate causes and ultimate functions. *Annual Review of Ecology and Systematics* 12: 253–279.
- Stone, H. 2013. Evolution and conservation of tetraploid *Euphrasia* L. in Britain. PhD thesis. University of Edinburgh.

- Stöck, M., J. Ustinova, D. K. Lamatsch, M. Schartl, N. Perrin, and C. Moritz. 2010. A vertebrate reproductive system involving three ploidy levels: Hybrid origin of triploids in a contact zone of diploid and tetraploid pale arctic green toads (*Bufo viridis* subgroup). *Evolution* 64: 944–959.
- Suarez-Gonzalez, A., C. Lexer, and Q. C. B. Cronk. 2018. Adaptive introgression: A plant perspective. *Biology Letters* 14: 8.
- Sutherland, B. L., and L. F. Galloway. 2017. Postzygotic isolation varies by ploidy level within a polyploid complex. *New Phytologist* 213: 404–412.
- Sutkowska, A., P. Boron, T. Warzecha, J. Debowski, and J. Mitka. 2017. Hybridization and introgression among three *Aconitum* (Ranunculaceae) species of different ploidy levels in the Tatra Mountains (Western Carpathians). *Plant Species Biology* 32: 292–303.
- Svensson, B. M., and B. A. Carlsson. 2004. Significance of time of attachment, host type, and neighbouring hemiparasites in determining fitness in two endangered grassland hemiparasites. *Annales Botanici Fennici* 41: 63–75.
- Tate, J., D. Soltis, and P. Soltis. 2005. Polyploidy in plants. In R. Gregory [ed.], *The evolution of the genome*, 371–426. Academic Press.
- Tesitel, J., P. Riha, S. Svobodova, T. Malinova, and M. Stech. 2010. Phylogeny, life history evolution and biogeography of the Rhinanthonid Orobanchaceae. *Folia Geobotanica* 45: 347–367.
- Těšitel, J., L. Plavcova, and D. D. Cameron. 2010. Heterotrophic carbon gain by the root hemiparasites, *Rhinanthus minor* and *Euphrasia rostkoviana* (Orobanchaceae). *Planta* 231: 1137–1144.
- Thompson, J. N., and K. F. Merg. 2008. Evolution of polyploidy and the diversification of plant-pollinator interactions. *Ecology* 89: 2197–2206.
- Thompson, W. P. 1930. Causes of difference in success of reciprocal interspecific crosses. *American Naturalist* 64: 407–421.
- Thorogood, C. J., F. J. Rumsey, S. A. Harris, and S. J. Hiscock. 2009. Gene flow between alien and native races of the holoparasitic angiosperm *Orobanche minor* (Orobanchaceae). *Plant Systematics and Evolution* 282: 31–42.
- Thorogood, C. J., F. J. Rumsey, and S. J. Hiscock. 2009. Host-specific races in the holoparasitic angiosperm *Orobanche minor*: Implications for speciation in parasitic plants. *Annals of Botany* 103: 1005–1014.

- Thorsson, A., S. P. Palsson, A. Sigurgeirsson, and K. Anamthawat-Jonsson. 2007. Morphological variation among *Betula nana* (diploid), *B. pubescens* (tetraploid) and their triploid hybrids in Iceland. *Annals of Botany* 99: 1183–1193.
- Todesco, M., M. A. Pascual, G. L. Owens, K. L. Ostevik, B. T. Moyers, S. Hubner, S. M. Heredia, et al. 2016. Hybridization and extinction. *Evolutionary Applications* 9: 892–908.
- Tsai, Y. H. E., and P. S. Manos. 2010. Host density drives the postglacial migration of the tree parasite, *Epifagus virginiana*. *Proceedings of the National Academy of Sciences of the United States of America* 107: 17035–17040.
- Tschirren, B., L. L. Bischoff, V. Saladin, and H. Richner. 2007. Host condition and host immunity affect parasite fitness in a bird-ectoparasite system. *Functional Ecology* 21: 372–378.
- Twyford, A. D. 2017. New insights into the population biology of endoparasitic Rafflesiaceae. *American Journal of Botany* 104: 1433–1436.
- Twyford, A. D. 2018. Parasitic plants. *Current Biology* 28: 857–859.
- Twyford, A. D., and R. A. Ennos. 2012. Next-generation hybridization and introgression. *Heredity* 108: 179–89.
- Twyford, A. D., and J. Friedman. 2015. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* 69: 1476–1486.
- Twyford, A. D., C. A. Kidner, and R. A. Ennos. 2015. Maintenance of species boundaries in a neotropical radiation of *Begonia*. *Molecular Ecology* 24: 4982–4993.
- Vale, P. F., and T. J. Little. 2009. Measuring parasite fitness under genetic and thermal variation. *Heredity* 103: 102–109.
- Valentine, D. H., and S. R. J. Woodell. 1960. Seed incompatibility in *Primula*. *Nature* 185: 778–779.
- Vallejo-Marin, M. 2012. *Mimulus peregrinus* (Phrymaceae): A new British allopolyploid species. *Phytokeys* 14: 1–14.
- Vallejo-Marin, M., and S. J. Hiscock. 2016. Hybridization and hybrid speciation under global change. *New Phytologist* 211: 1170–1187.

- Vandijk, P., M. Hartog, and W. Vandelen. 1992. Single cytotype areas in autopolyploid *Plantago media* L. *Biological Journal of the Linnean Society* 46: 315–331.
- W., E. S. 2016. Using DECIPHER v2.0 to analyze big biological sequence data in R. *The R Journal* 8: 352–359.
- Wang, J., S. Dong, L. Yang, A. Harris, H. Schneider, and M. Kang. 2020. Allopolyploid speciation accompanied by gene flow in a tree fern. *Molecular Biology and Evolution*.
- Wang, N., J. S. Borrell, W. J. A. Bodles, A. Kuttapitiya, R. A. Nichols, and R. J. A. Buggs. 2014. Molecular footprints of the Holocene retreat of dwarf birch in Britain. *Molecular Ecology* 23: 2771–2782.
- Wang, X. M., G. Gussarova, M. Ruhsam, N. de Vere, C. Metherell, P. M. Hollingsworth, and A. D. Twyford. 2018. DNA barcoding a taxonomically complex hemiparasitic genus reveals deep divergence between ploidy levels but lack of species-level resolution. *Aob Plants* 10: 13.
- Watkins, A. E. 1932. Hybrid sterility and incompatibility. *Journal of Genetics* 25: 125–162.
- Webb, C. O., and M. J. Donoghue. 2005. Phyloomatic: Tree assembly for applied phylogenetics. *Molecular Ecology Notes* 5: 181–183.
- Weinkoop, S., S. Baginsky, and W. Weckwerth. 2010. *Arabidopsis thaliana* as a model organism for plant proteome research. *Journal of Proteomics* 73: 2239–2248.
- Wendel, J. F. 2015. The wondrous cycles of polyploidy in plants. *American Journal of Botany* 102: 1753–1756.
- Wendel, J. F., S. A. Jackson, B. C. Meyers, and R. A. Wing. 2016. Evolution of plant genome architecture. *Genome Biology* 17: 14.
- Westbury, D. B., and N. P. Dunnett. 2007. The impact of *Rhinanthus minor* in newly established meadows on a productive site. *Applied Vegetation Science* 10: 121–129.
- Westwood, J. H., J. I. Yoder, M. P. Timko, and C. W. dePamphilis. 2010. The evolution of parasitism in plants. *Trends in Plant Science* 15: 227–235.
- Wettstein, R. 1895. Der Saison-Dimorphismus als Ausgangspunkt fir die Bildung neuer Arten im Pflanzenreiche. *Berichte der Deutschen Botanischen Gesellschaft*. 13: 303–313.
- Wettstein, R. von. 1896. Zur Systematik der Europäischen Euphrasia-Arten. *Österreichische Botanische Zeitschrift* 46: 381–386.

- Whitney, K. D., J. R. Ahern, L. G. Campbell, L. P. Albert, and M. S. King. 2010. Patterns of hybridization in plants. *Perspectives in Plant Ecology Evolution and Systematics* 12: 175–182.
- Whitton, J. 2004. One down and thousands to go - dissecting polyploid speciation. *New Phytologist* 161: 610–612.
- Wicaksono, A., S. Mursidawati, L. A. Sukamto, and J. A. T. da Silva. 2016. *Rafflesia* spp.: propagation and conservation. *Planta* 244: 289–296.
- Wickham, H. 2016. Ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.
- Wickham, H., R. François, L. Henry, and K. Müller. 2020. Dplyr: A grammar of data manipulation.
- Widmer, A., C. Lexer, and S. Cozzolino. 2009. Evolution of reproductive isolation in plants. *Heredity* 102: 31–38.
- Wilkins, D. A. 1963. Plasticity and establishment in *Euphrasia*. *Annals of Botany* 27: 533 –&.
- Wilson, J., A. Perry, J. R. Shepherd, M. Duran-Castillo, C. E. Jeffree, and S. Cavers. 2020. Invasion, isolation and evolution shape population genetic structure in *Campanula rotundifolia*. *Aob Plants* 12.
- Windsor, D. A. 1998. Most of the species on earth are parasites. *International Journal for Parasitology* 28: 1939–1941.
- Wittzell, H. 1999. Chloroplast DNA variation and reticulate evolution in sexual and apomictic sections of dandelions. *Molecular Ecology* 8: 2023–2035.
- Wood, T. E., N. Takebayashi, M. S. Barker, I. Mayrose, P. B. Greenspoon, and L. H. Rieseberg. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13875–13879.
- Yang, Z., E. K. Wafula, L. A. Honaas, H. Zhang, M. Das, M. Fernandez-Aparicio, K. Huang, et al. 2015. Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. *Molecular biology and evolution* 32: 767–790.
- Yeo, P. F. 1978. A taxonomic revision of *Euphrasia* in Europe. *Botanical Journal of the Linnean Society* 77: 223–334.
- Yeo, P. F. 1961. Germination, seedlings, and the formation of haustoria in *Euphrasia*. *Watsonia* 5: 11–22.

- Yeo, P. F. 1956. Hybridisation between diploid and tetraploid species of *Euphrasia*. *Watsonia* 3: 253–269.
- Yeo, P. F. 1973. The Azorean species of *Euphrasia*. *Boletim do Museu de História Natural do Funchal* 17: 74–83.
- Yeo, P. F. 1954. The cytology of British species of *Euphrasia*. *Watsonia* 3: 101–108.
- Yeo, P. F. 1968. The evolutionary significance of the speciation of *Euphrasia* in Europe. *Evolution* 22: 736–747.
- Yeo, P. F. 1964. The growth of *Euphrasia* in cultivation. *Watsonia* 6: 1–24.
- Yoder, J. I. 1999. Parasitic plant responses to host plant signals: A model for subterranean plant–plant interactions. *Current Opinion in Plant Biology* 2: 65–70.
- Yoneyama, K., A. A. Awad, X. Xie, K. Yoneyama, and Y. Takeuchi. 2010. Strigolactones as Germination Stimulants for Root Parasitic Plants. *Plant and Cell Physiology* 51: 1095–1103.
- Young, M. R., H. C. Proctor, J. R. deWaard, and P. D. N. Hebert. 2019. DNA barcodes expose unexpected diversity in canadian mites. *Molecular Ecology* 28: 5347–5359.
- Yu, G. C., D. K. Smith, H. C. Zhu, Y. Guan, and T. T. Y. Lam. 2017. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8: 28–36.
- Zhang, J. H., and M. J. Lechowicz. 1994. Correlation between time of flowering and phenotypic plasticity in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany* 81: 1336–1342.
- Zhang, X. T., S. C. Zhang, Q. Zhao, R. Ming, and H. B. Tang. 2019. Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data. *Nature Plants* 5: 833–845.
- Zhou, M. M., G. Q. Yang, G. L. Sun, Z. H. Guo, X. Gong, and Y. Z. Pan. 2020. Resolving complicated relationships of the *Panax bipinnatifidus* complex in southwestern China by RAD-seq data. *Molecular Phylogenetics and Evolution* 149.
- Zlonis, K. J., and B. L. Gross. 2018. Genetic structure, diversity, and hybridization in populations of the rare arctic relict *Euphrasia hudsoniana* (Orobanchaceae) and its invasive congener *Euphrasia stricta*. *Conservation Genetics* 19: 43–55.

- Zohary, D., and U. Nur. 1959. Natural triploids in the orchard grass, *Dactylis glomerata* L. polyploid complex and their significance for gene flow from diploid to tetraploid levels. *Evolution* 13: 311–317.
- Zohren, J., N. A. Wang, I. Kardailsky, J. S. Borrell, A. Joecker, R. A. Nichols, and R. J. A. Buggs. 2016. Unidirectional diploid-tetraploid introgression among British birch trees with shifting ranges shown by restriction site-associated markers. *Molecular Ecology* 25: 2413–2426.
- Zopfi, H. J. 1993. Ecotypic variation in *Rhinanthus alectorolophus* (Scopoli) Pollich (Scrophulariaceae) in relation to grassland management 2. The genotypic basis of seasonal ecotypes. *Flora* 188: 153–173.
- Zopfi, H. J. 1995. Life history variation and infraspecific heterochrony in *Rhinanthus glacialis* (Scrophulariaceae). *Plant Systematics and Evolution* 198: 209–233.
- Zopfi, H. J. 1998. The genetic basis of ecotypic variants of *Euphrasia rostkoviana* Hayne (Scrophulariaceae) in relation to grassland management. *Flora* 193: 41–58.