

# A new biological species in the *Mercurialis annua* polyploid complex: functional divergence in inflorescence morphology and hybrid sterility

Wen-Juan Ma<sup>1,†,\*</sup>, Luis Santos del Blanco<sup>1,‡,§</sup> and John R. Pannell<sup>1</sup>

<sup>1</sup>Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

\*For correspondence. E-mail [wenjuanma84@gmail.com](mailto:wenjuanma84@gmail.com)

†Present address: Department of Biology, Amherst College, Amherst, MA 01002, USA.

‡Present address: Sustainable Forest Management Research Institute, University of Valladolid – INIA, Avda. Madrid, 34071 Palencia, Spain.

§Present address: Biotecnología Forestal Aplicada, Calle Curtidores 17, 34004 Palencia, Spain

Received: 7 August 2018 Returned for revision: 18 March 2019 Editorial decision: 25 March 2019 Accepted: 27 March 2019

- **Background and Aims** Polyploidy has played a major role in the origin of new plant species, probably because of the expansion of polyploid populations in the species' ecological niche, and because reproductive isolation can be established between a new polyploid population and its diploid progenitor species. It is well established that most polyploid species are polyphyletic, with multiple independent origins, and that polyploid genomes may undergo rapid change after their duplication and hybridization associated with their origin. We considered whether multiple independent origins and rapid genomic change might lead to reproductive isolation between polyploid populations of the same ploidy but with potentially different evolutionary histories.
- **Methods** We tested our hypothesis by assessing differences in DNA content and morphology, the evolution of reproductive isolation, and the phylogenetic placement of two broadly sympatric hexaploid lineages of the wind-pollinated annual plant *Mercurialis annua* hitherto regarded as populations of the same species.
- **Key Results** The two hexaploid lineages of *M. annua* have slightly divergent DNA content, and distinct inflorescence morphology. They also fall into largely different clades of a chloroplast phylogeny and are reproductively isolated from one another.
- **Conclusions** The distinct evolutionary histories of the two hexaploid lineages of *M. annua* have contributed to the remarkable reproductive diversity of the species complex. It seems likely that reproductive interference between them will eventually lead to the displacement of one lineage by the other via pollen swamping. Thus, whereas polyploidization can contribute to speciation, diversification might also be compromised by reproductive interference.

**Keywords:** Male-like inflorescence, polyploidy, functional divergence, female sterility, reproductive isolation, phylogeny, evolutionary origin.

## INTRODUCTION

Whole genome duplication and polyploidy have played major roles in the diversification of land plants (Stebbins, 1950; Levin, 1983; Soltis *et al.*, 2014a, b). Recent polyploids account for about 35 % of flowering plant species (Wood *et al.*, 2009), and probably all angiosperm lineages have experienced polyploidy in their distant past (Soltis *et al.*, 2009). The success of polyploids, particularly allopolyploids, is thought to be due to the genomic flexibility and increased heterozygosity that results from the combination of two or more genomes, conferring on individuals an ability to colonize a broader range of environments than their diploid progenitors (Soltis *et al.*, 2004, 2014; Hegarty and Hiscock, 2008; but see Glennon *et al.*, 2014). Polyploidization may also contribute to diversification through the reciprocal loss of duplicated genes, as shown in yeast and *Arabidopsis* (Scannell *et al.*, 2006, 2007a, b; Bikard *et al.*, 2009). Yet if polyploid lineages were the result of a unique event of hybridization and genome duplication, as they were thought to be (Otto, 2007), the adaptive flexibility of having multiple genomes should be limited by the narrow genetic bottleneck of

their origin. However, it has been clear for more than two decades that most polyploid species are polyphyletic, i.e. the result of multiple, independent episodes of hybridization and genome duplication (Soltis and Soltis, 1999). The polyphyletic origin of polyploid species probably contributes substantially to their commonly high genetic diversity, ecological breadth and evolutionary potential (Lewis, 1980; Pannell *et al.*, 2004; Soltis *et al.*, 2004, 2014a; Beest *et al.*, 2012).

The high genetic diversity observed in polyploid species may also be due to an accelerated rate of genome reorganization and other mutational processes that follow hybridization and genome duplication. It is not straightforward to characterize such processes in long-established polyploid species, but analysis of the immediate descendants of synthetic polyploids has revealed that these processes can be dramatic. For instance, Song *et al.* (1995) observed changes in restriction fragment patterns in each of the first few generations after synthesizing polyploids from several *Brassica* species, attributing these changes to chromosomal rearrangements, point mutations, gene conversion and DNA methylation. Subsequent research indicates that

accelerated evolution of new polyploid lineages is probably frequent (Chen and Ni, 2006; Chen *et al.*, 2007), but not ubiquitous (e.g. Liu *et al.*, 2001). A revealing example is provided by the high cytological and morphological variation observed in the arctic polyploid species *Saxifraga cernua*, this variation probably having arisen multiple times (Brochmann *et al.*, 1998).

The polyphyletic origin and rapid genome evolution of polyploid lineages should not only contribute to their ecological breadth, but also to their potential for subsequent diversification. Substantial attention has been paid to establishing the extent and implications of reproductive isolation between polyploids and their diploid relatives, through both experimental crosses (Maceira *et al.*, 1993; Husband *et al.*, 2002; Buggs and Pannell, 2006) and genome analysis (Song *et al.*, 1995; Soltis and Soltis, 2000; Chen and Ni, 2006). Although reproductive isolation between polyploids and their diploid progenitors as a result of their differing chromosome numbers may be an important engine of diversification (Thompson and Lumaret, 1992; Ramsey and Schemske, 1998; Soltis and Soltis, 2009; but see Baduel *et al.*, 2018 for a discussion of evidence for cross-ploidy gene flow), different independent origins of the same polyploid taxon might also contribute to diversification if they are reproductively isolated from one another, particularly if they have been subject to rapid evolutionary change following hybridization and/or genome duplication. Yet there appears to be an almost complete lack of evidence for or against the hypothesis that polyploid populations of the same ploidy levels with different origins might be reproductively isolated. Combined with the potentially accelerated evolution of new polyploid lineages due to genome reorganization (Song *et al.*, 1995; Soltis and Soltis, 2000) and changes in gene expression (Chen and Ni, 2006), such reproductive isolation would evidently have important taxonomic and ecological implications.

In the present study, we reveal that broadly sympatric hexaploid monoecious populations of the European wind-pollinated annual plant *Mercurialis annua* (Euphorbiaceae) belong to at least two different lineages that are strongly reproductively isolated from one another, and that probably had an independent origin. Our study was stimulated by the observation of morphological differences in inflorescence architecture among populations and an interest in the functional significance of that variation. The *M. annua* species complex displays remarkable variation in its sexual systems, including dioecy, monoecy and the rare sexual system androdioecy, where males co-occur with monoecious hermaphrodites (Pannell, 1997a, 2008; Obbard *et al.*, 2006a,b; Russell and Pannell, 2015). Against this background, variation in inflorescence architecture among monoecious populations allows valuable comparisons among divergent inflorescence strategies for outcross siring success (Santos del Blanco *et al.*, 2019), but it also prompts questions concerning its origins.

In the *M. annua* species complex, diploid populations are dioecious, but hexaploid populations are either monoecious or androdioecious (Obbard *et al.*, 2006a,b). Diploid and polyploid males hold their flowers in erect ‘pedunculate’ inflorescences above the plant (Fig. 1A), whereas diploid females and hexaploid monoecious individuals (which are effectively modified pollen-producing females) typically hold both their male and their female flowers in sub-sessile inflorescences in the leaf axils (Hesse and Pannell, 2011a, b; Fig. 1B). However, we recently

documented the existence in some populations of apparently hexaploid monoecious individuals of *M. annua* that, like males, hold virtually all their male flowers on erect peduncles rather than in the sub-sessile axillary inflorescences that are more typical of hexaploid *M. annua*, whereas female flowers are found on short pedicels in the leaf axils or, occasionally, on the erect peduncles along with male flowers (Fig. 1C; Santos del Blanco *et al.*, 2019). The non-pedunculate form of hexaploid monoecious *M. annua* (which we here label ‘P–’) occurs widely around the Iberian Peninsula in disturbed ruderal habitats, often together with males in androdioecious populations (Pannell, 1997a; Buggs and Pannell, 2006; Obbard *et al.*, 2006a). The pedunculate phenotype (labelled ‘P+’), although broadly sympatric with the P– phenotype along the Spanish Mediterranean coastal areas and occupying very similar habitats, occurs rarely in mixed populations with the P– phenotype. A previous study found that the P+ phenotype is substantially better at dispersing pollen than the P– phenotype and enjoys much greater siring success in experimental populations, without any obvious phenotypic advantage to the P– phenotype (Santos del Blanco *et al.*, 2019). One might expect such a superior strategy for pollen dispersal to spread quickly throughout the species range, particularly as it does not appear very costly in terms of other fitness components (Buggs and Pannell, 2006). However, the results we present here indicate that the P– and P+ phenotypes, which are both hexaploid and broadly sympatric, nonetheless belong to two different biological species.

We characterized the variation between the P– and P+ phenotypes of hexaploid *M. annua* in terms of inflorescence and vegetative morphology, ploidy and DNA quantity, their position in a phylogeny of the genus based on both nuclear and chloroplast markers (Obbard *et al.*, 2006a), and experimental crosses. Although the two phenotypes are morphologically and ecologically very similar, all our data are consistent with the hypothesis that they belong to two different evolutionary lineages. Our study shows that the *M. annua* species complex is even more diverse than hitherto thought. The finding of reproductive isolation between two polyploid lineages with the same ploidy level but apparently different hybridization pathways points to a potentially important but underexplored mode of diversification in flowering plants. It also raises interesting questions concerning the long-term evolutionary maintenance of distinct lineages that occupy similar habitats and are susceptible to strongly detrimental effects of reproductive interference through the generation of sterile hybrid progeny.

## MATERIALS AND METHODS

### Source populations

We sampled plants of hexaploid *M. annua* from a total of 26 source populations (Supplementary Data Table S1) in 2013 in eastern Spain, in the region where populations of P– and P+ individuals broadly co-occur (Supplementary Data Fig. S1). Seeds were pooled from ~30 individuals from each of 13 populations with P– and 13 populations with P+ individuals (Fig. 1). Plants were raised from these seeds under uniform conditions in a glasshouse at the University of Lausanne, Switzerland, and their seed progeny were harvested for use in various experiments as detailed below.

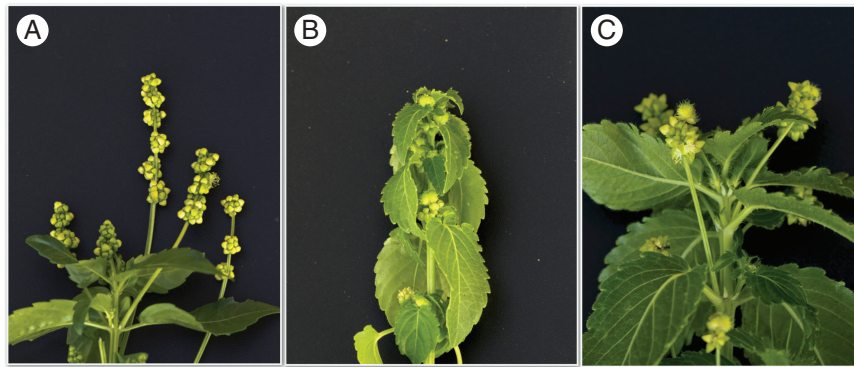


FIG. 1. Images of three different sexual phenotypes of hexaploid *Mercurialis annua*: (A) a male individual, showing flowers on peduncles; (B) a P– monoecious individual, with male and female flowers held in leaf axils; and (C) a P+ monoecious individual, with flowers held on erect inflorescence stalks, or ‘peduncles’. Images courtesy of Xinji Li.

#### Common garden experiment and morphological trait measurements

We recorded morphological trait values for individuals from 21 populations of *M. annua* in our common garden experiment. Seeds were first sown in bulk in seedling trays, and seedlings were then transplanted individually into 15-cm pots ~2 weeks after emergence. Pots were arranged in a complete randomized block design, with 11 blocks and 21 pots per block, i.e. one per population. Plants were watered as needed and grown under uniform glasshouse conditions (23–25 °C, 50 % humidity, 12:12 light–dark cycle).

Plants were harvested 7 weeks after transplanting, and the following morphological traits were measured and recorded: plant height; length of one of each of the pairs of branches emanating from the first three basal nodes (branches 0, 1 and 2); length of third, fourth and fifth internodes; total number of male peduncles on the plant; length and number of male flowers on five randomly chosen male peduncles; number of peduncles with fruits instead of male flowers at their apex; number of peduncles protruding above the plant’s uppermost leaves; and the biomass of all male flowers on the plant (following their removal and drying at 40 °C for 1 week). For one leaf emanating from each of the third, fourth and fifth internodes on the main shoot, we also estimated the leaf blade area, leaf blade circularity, leaf blade length, leaf blade perimeter, leaf blade width, petiole length and leaf perimeter area, based on scanning and image analysis using the software LeafJ (Maloof *et al.*, 2013). We also recorded whether the leaves had curly margins or not, as well as the presence or absence of trichomes on stems. Finally, we recorded total male flower biomass, total seed biomass and total above-ground plant biomass for each plant. Seed number and the size distribution of seeds were obtained with an Elmor C3 seed counter (SA. Ltd, Schwyz, Switzerland). Male and female reproductive effort was estimated as male-flower and seed biomass, respectively, divided by above-ground plant biomass. An index of sex allocation was estimated as male-flower biomass divided by the total biomass of pollen and seeds together.

We used principal component analysis (PCA) based on all traits to characterize the overall morphological differences between the two phenotypes. We used generalized linear mixed models (GLMMs) using the lme4 package in R version 3.4.3 (R Core Team, 2017) to assess differences in morphology

between the two phenotypes for specific traits, with phenotype defined as a fixed effect, and population and block defined as random effects. Gaussian models were used for all variables, except for the presence or absence of trichomes or curly leaves, which were analysed with binomial models.

#### Genome size estimation

We used flow cytometry to estimate the genome size and ploidy level of three to five individuals from each population of P– and P+ phenotypes, modifying the protocols described by Obbard *et al.* (2006a) and Ma *et al.* (2013). Briefly, we chopped a mix of 10 mg of leaf material from *M. annua* with 5 mg of a fresh mature leaf of tomato (*Solanum lycopersicum*) as an internal control in 1000 µL ice-cold Galbraith buffer (Galbraith *et al.*, 1983). We strained the mix through a 35-µm membrane filter of Falcon tubes (12 × 75 mm), and stained the strained DNA samples with 60 µL DAPI (0.02 µL mL<sup>-1</sup>, 1.2 µL). The stained DNA samples were mixed in a gentle vortex, and then kept at 4 °C for ~30 min before analysis with a flow cytometer (BD LSRFortessa, BD Biosciences, Franklin Lakes, NJ, USA). Data were analysed using the software BDFACSDiva, with the combination of populations from the same and other ploidy levels of *M. annua*, as described and analysed by Obbard *et al.* (2006a). Finally, the target genome size (pg) was calculated as (mean fluorescence target/mean fluorescence control) × genome control (Hare and Johnston, 2012). We used a GLM with the multcomp function in R version 3.4.3 (R Core Team, 2017) to compare the genome size among the sampled populations.

#### Crossing array for the production of hybrid and non-hybrid progeny

We conducted reciprocal crosses between sets of P– and P+ monoecious individuals to assess evidence for post-zygotic reproductive isolation, choosing two populations of the P– phenotype (P-GAU, P-GET; Supplementary Data Table S1) and two populations of the P+ phenotype (P+\_ACA, P+\_BNS). Each of the four combinations of crosses was replicated three times (i.e. there were 12 crossing arrays in total). Arrays were established as squares of 49 plants, with 25 P– and 24 P+ individuals



in each array, arranged in spatial alternation. Plants were harvested after 5 weeks of growth and mating in their array, and 30 progeny from the set of P– or P+ mothers per array were raised on a glasshouse bench in a complete randomized block design. For each progeny individual, we recorded above-ground biomass, and female and male reproductive efforts. We determined the paternity for each progeny individual (i.e. whether it had been sired by a P+ or P– individual) on the basis of four microsatellite markers (see below).

To assess the male fertility of hybrids between P+ and P– parents, we first checked the morphology of anthers under a stereomicroscope (Wild M3, Heerbrugg, Switzerland), and counted the number of anthers, number of pollen grains and pollen size using a particle size analyser (Micromeritics Elzone II, Norcross, GA, USA). Because different lineages of *M. annua* are able to pollinate and fertilize one another, including between ploidy levels, we further assayed the siring potential of the  $F_1$  progeny by allowing them to pollinate the stigmas of diploid *M. annua* females and assaying their siring success; we performed this assay on diploid females rather than hexaploid monoecious individuals to avoid the complication of self-pollination. In total, we established three isolated mating ‘triplets’, with each triplet comprising one  $F_1$  (progeny of P– and P+) hybrid individual together with two diploid females; the triplet crosses were established individually in one of three separated pollen-proof incubators (~23 °C, 50 % humidity, 12: 12-h light–dark cycle). We also established three replicate crosses between diploid *M. annua* females and males as a control to ensure that culture conditions allowed effective crossing under similar conditions in a glasshouse.

#### Paternity assignment for progeny from crossing arrays

For DNA extraction, the fresh leaf material was first dried in an oven at 40 °C for 3 d. The dry leaves (10–15 mg) were then ground to a powder with beads using a TissueLyser II machine (Qiagen, Valencia, CA, USA) for subsequent DNA extraction, following the Biosprint 96 protocol provided by the manufacturer. All plant tissue samples were then processed using the BioSprint 96 workstation with the BS 96DNA Plant program, resulting in 200 µL DNA in elution buffer (Qiagen).

We determined the paternity class of each progeny on the basis of population-specific microsatellite markers developed previously in the lab (primer sequences and PCR conditions are described in Korbecka et al., 2011). We used these microsatellites to distinguish between siring by other P– or P+ individuals in each array, as well as to estimate average selfing rates. We genotyped 30 plants from progeny produced by each of the two types of monoecious mothers from each of the 12 replicate arrays (a total of 720 plants). Individuals were scored for each of four microsatellite markers (Mh14, Mh15, Mh19, Mh52, Supplementary Data Table S2) that are known to provide good separation of the hexaploid populations sampled (Korbecka et al., 2010). All four markers were amplified in a single multiplexed PCR, following the protocol described in Korbecka et al. (2010). We processed the samples in an ABI 3100 sequencer (Applied Biosystems, Carlsbad, CA, USA), and analysed the results with GeneMapper v.4.0 (Applied Biosystems). Individual genotypes were classified as having been sired by

a father of the same or the competing phenotype in the array. We estimated selfing rates using the software RMES (David et al., 2007). This was only possible for the populations of P– individuals, as populations with P+ individuals had almost no genetic variability for the markers used (Supplementary Data Table S2); for P+ individuals we also distinguished between P– and P+ sires, but could not estimate selfing rates. This limitation did not affect the interpretation of  $F_1$  progeny between P+ and P– individuals.

#### Sanger sequencing and phylogeny construction

To investigate the phylogenetic relationship between P– and P+ individuals in our samples and with other lineages in the genus, we constructed phylograms on the basis of the same loci as those used for phylogenetic reconstruction of the genus by Obbard et al. (2006a), i.e. two chloroplast markers, *trnL-trnF* and *matK-trnK*, as well as the nuclear internal transcribed spacer 2 (*ITS2*). For each individual sampled, we collected fresh leaves, dried them and performed DNA extraction as described above. DNA quality of leaves from two P– populations was insufficient for genotyping (due to poor plant health prior to sampling), and we excluded these populations from the following phylogenetic analysis described below. We used cloning to identify potentially different copies of *ITS2* in the hexaploid genome. Because cloning can separate individual sequence copies, sampling a number of colonies can potentially allow identification of all *ITS2* copies for a particular genome. Cloning was performed using the Promega pGEM-T Easy Vector System II, following the manufacturer’s protocol. The PCR and cloning products were sequenced by Microsynth (Balgach, Switzerland). Raw sequences were trimmed with Geneous v.2 (<http://www.geneious.com>; Kearse et al., 2012), and ambiguous nucleotides were checked by eye and corrected manually. After trimming, sequences of all individuals were the same length for a given locus: *ITS2*, *trnL-trnF* and *matK-trnK* sequences were 679, 415 and 560 bp in length, respectively. All primer sequences as well as the protocol for PCRs are described in Obbard et al. (2006a).

We used MrBayes (v.3.2.6) for phylogenetic reconstruction (Ronquist et al., 2012), combining our own sequences for nine P– and 11 P+ populations with those downloaded from the NCBI database (populations from the same and other ploidy levels of *M. annua*, as described and analysed in Obbard et al. (2006a)). MrBayes uses Bayesian inferences and Markov chain Monte Carlo (MCMC) methods to estimate the posterior distribution of model parameters. First, sequences for each marker were aligned using the multiple sequence alignment function in ClustalW (Larkin et al., 2007). For the aligned sequences of each marker, the best partition scheme and nucleotide substitution model were selected using PartitionFinder (Lanfear et al., 2012). Aligned sequences were analysed in MrBayes, seeking the best partition scheme and substitution model parameters with an MCMC chain of 10 million iterations. The trace files generated by Bayesian MCMC runs were verified using Tracer software until all parameter values sampled from the chain reached their optima (<http://tree.bio.ed.ac.uk/software/tracer/>). Finally, the generated tree file was visualized in FigTree software (<http://tree.bio.ed.ac.uk/software/figtree/>).

## RESULTS

*Morphological trait differences between inflorescence phenotypes*

Most vegetative traits differed significantly between the P– and P+ phenotypes, but there was also substantial variation among populations within each phenotype (Table 1). Overall, PCA separated the two phenotypes into two clusters, albeit with some overlap (Fig. 2). Although most leaf traits (i.e. leaf blade area, leaf blade perimeter, leaf blade width, leaf petiole length and leaf perimeter area) did not differ between the two phenotypes, the leaf margins of the P+ phenotype were significantly more undulating or ‘curly’ than those of the P– phenotype (Table 1). The first and second PCA axes accounted for 33.3 and 18.8 %, respectively, of the overall morphological variation (Fig. 2).

The P+ phenotype had ~25 times more peduncles than the P– phenotype ( $P < 0.0001$ , Fig. 3A), and peduncles on P+ plants were overall seven times longer than those on P– plants ( $P < 0.0001$ , Fig. 3B). Plants of the P+ phenotype also invested significantly more resources toward male function (in terms of male reproductive effort) than did those of the P– phenotype ( $P < 0.0001$ , Fig. 3C), and correspondingly fewer resources to female function (in terms of female reproductive effort;  $P < 0.0001$ , Fig. 3D). Accordingly, the sex allocation of the P+ phenotype was skewed substantially towards the male function relative to the P– phenotype (Fig. 3E), although the P+ phenotype had overall significantly lower above-ground biomass than the P– phenotype ( $P < 0.0001$ , Fig. 3F).

*Differences in DNA content between inflorescence phenotypes*

Flow cytometry confirmed that both P– and P+ phenotypes are probably hexaploid, i.e. DNA content corresponds roughly to that observed for hexaploid *M. annua* previously (Obbard *et al.*, 2006a), and was about three times larger than for diploids of the species complex. However, we nonetheless found significant variation in genome size among populations within each lineage assayed, particularly for the hexaploid populations (Fig. 4). Specifically, individuals of the P+ phenotype had somewhat higher DNA content ( $2.05 \pm 0.055$  pg) than those of the P– phenotype ( $1.81 \pm 0.025$  pg), and varied more among populations ( $P < 0.001$ , Fig. 4).

*Reciprocal crossing experiment, hybrid viability and hybrid sterility*

Progeny of crosses between P+ and P– individuals were fully viable, and indeed the above-ground biomass of  $F_1$  hybrids was two- to three-fold higher than that of progeny resulting from crosses within each phenotype ( $P < 0.001$ , Fig. 5A; Supplementary Data Table S2). However,  $F_1$  hybrids were largely female-sterile, regardless of the direction of the cross, with very few viable seeds produced (most seeds were aborted before they reached their mature size;  $P < 0.001$ , Fig. 5B). We were unable to compare the siring ability of the  $F_1$  individuals in a fully balanced experiment, but it would seem that the male fertility of  $F_1$  hybrids is probably also impaired:

TABLE 1. Measures of morphological traits for the two distinctive monoecious phenotypes

Trait	P– (mean $\pm$ s.d.)	P+ (mean $\pm$ s.d.)	<i>P</i> value	Bonferroni-corrected <i>P</i> value
Leaf blade area (mm <sup>2</sup> )	1410 $\pm$ 434	1460 $\pm$ 299	0.54	1
Leaf blade circularity	0.38 $\pm$ 0.04	0.35 $\pm$ 0.04	<b>0.03</b>	0.9
Leaf blade length (mm)	52.40 $\pm$ 8.90	52.60 $\pm$ 6.20	<b>0.02</b>	0.6
Leaf blade perimeter (mm)	212.00 $\pm$ 35.40	226.90 $\pm$ 22.30	0.84	1
Leaf blade width (mm)	33.20 $\pm$ 5.60	34.90 $\pm$ 4.10	0.09	1
Blade to petiole length ratio	3.47 $\pm$ 1.18	3.56 $\pm$ 1.28	0.90	1
Leaf petiole length (mm)	17.30 $\pm$ 5.00	16.70 $\pm$ 4.30	0.67	1
Leaf perimeter to area ratio	0.014 $\pm$ 0.003	0.014 $\pm$ 0.002	0.95	1
Proportion of plants with curly leaves	0.06 $\pm$ 0.05	0.71 $\pm$ 0.10	<b>&lt;0.001</b>	<b>0.003</b>
Above-ground biomass (g)	4.33 $\pm$ 0.98	3.8 $\pm$ 0.91	<b>0.02</b>	0.6
Length of branch 0 (cm)	17.50 $\pm$ 5.97	12.90 $\pm$ 4.82	0.24	1
Length of branch 1 (cm)	15.40 $\pm$ 5.63	12.00 $\pm$ 3.99	<b>0.03</b>	0.9
Length of branch 2 (cm)	15.70 $\pm$ 5.68	12.10 $\pm$ 4.71	<b>0.05</b>	1
Plant height (cm)	30.40 $\pm$ 7.77	23.10 $\pm$ 6.35	<b>0.01</b>	0.3
Length of 4th internode (mm)	44.40 $\pm$ 11.90	36.10 $\pm$ 10.70	<b>0.04</b>	1
Length of 5th internode (mm)	44.10 $\pm$ 13.80	35.20 $\pm$ 12.40	<b>0.05</b>	1
Length of 3rd internode (mm)	40.10 $\pm$ 9.43	31.90 $\pm$ 8.48	<b>0.01</b>	0.3
Proportion of plants with trichomes	0.11 $\pm$ 0.10	0.03 $\pm$ 0.04	0.19	1
Average number of clusters of male flowers	0.94 $\pm$ 0.75	1.92 $\pm$ 0.76	<b>&lt;0.001</b>	<b>0.003</b>
Length of peduncle (mm)	7.4 $\pm$ 8.89	54.80 $\pm$ 23.60	<b>&lt;0.001</b>	<b>0.003</b>
Number of peduncles with fruits	0.59 $\pm$ 1.85	11.37 $\pm$ 9.58	<b>&lt;0.001</b>	<b>0.003</b>
Number of peduncles	1.36 $\pm$ 4.03	32.90 $\pm$ 20.90	<b>&lt;0.001</b>	<b>0.003</b>
Number of apical peduncles	0.18 $\pm$ 0.8	7 $\pm$ 10.90	<b>&lt;0.001</b>	<b>0.003</b>
Male reproductive effort	26.60 $\pm$ 16.02	47.40 $\pm$ 33.70	<b>&lt;0.001</b>	<b>0.003</b>
Pollen biomass (mg)	117.0 $\pm$ 62.70	188.20 $\pm$ 91.70	<b>&lt;0.001</b>	<b>0.003</b>
Seed biomass (g)	2.31 $\pm$ 0.28	2.66 $\pm$ 0.48	<b>0.01</b>	0.3
Seed number	526.80 $\pm$ 222.60	285.20 $\pm$ 190.10	<b>&lt;0.001</b>	<b>0.003</b>
Female reproductive effort	120.20 $\pm$ 41.90	71.80 $\pm$ 38.40	<b>&lt;0.001</b>	<b>0.003</b>
Seed size ( $\mu$ m)	231.50 $\pm$ 102.10	113.10 $\pm$ 87.70	<b>&lt;0.001</b>	<b>0.003</b>
Sex allocation	0.19 $\pm$ 0.12	0.4 $\pm$ 0.25	<b>&lt;0.001</b>	<b>0.003</b>

*P* values are derived from generalized linear mixed models (GLMMs), and bonferroni adjusted *P* values were provided for correcting multiple test. *P* values < 0.05 are in bold.

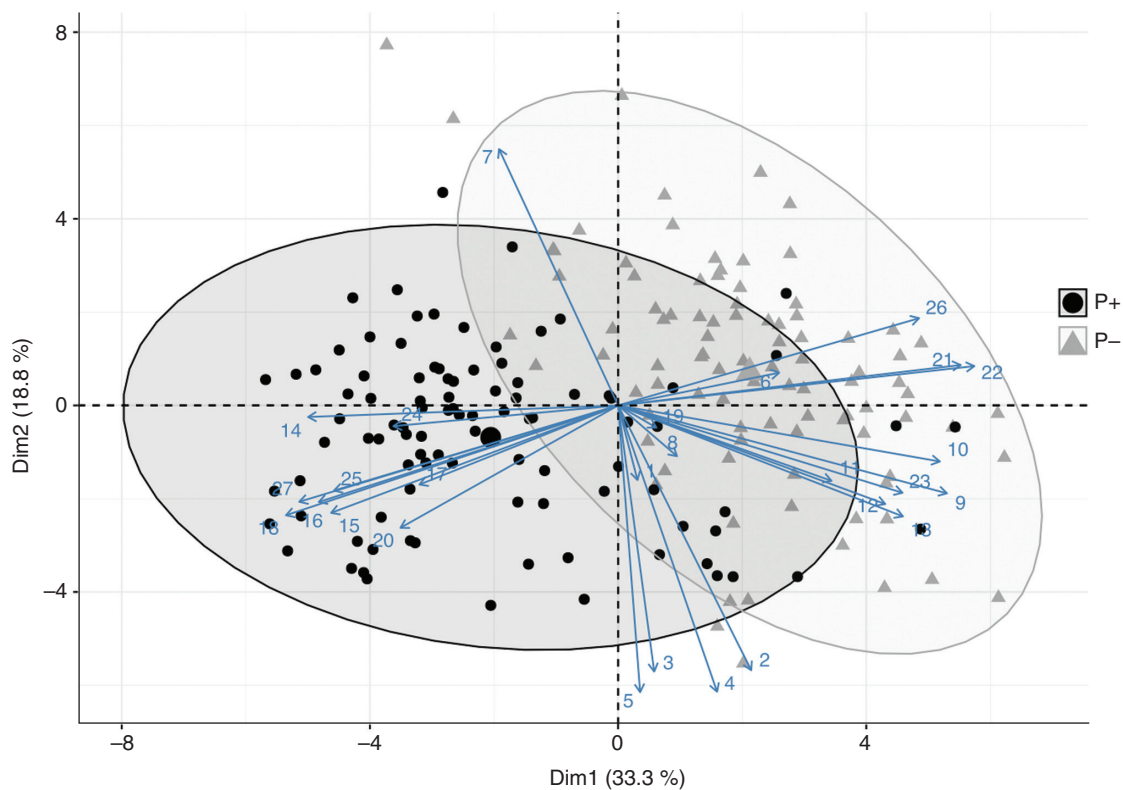


FIG. 2. Principal component plot of multiple morphological traits for P+ and P- monoecious hexaploid individuals of *Mercurialis annua*. Solid black circles and solid grey triangles denote P+ and P- populations, respectively. Numbers refer to traits as follows: 1, petiole length; 2, leaf blade length; 3, leaf blade width; 4, leaf blade area; 5, leaf blade perimeter; 6, leaf blade circularity; 7, leaf blade perimeter to area ratio; 8, leaf blade to petiole ratio; 9, plant height; 10, average branch length; 11, length of 3rd internode; 12, length of 4th internode; 13, length of 5th internode; 14, presence of curly leaves; 15, number of peduncles; 16, numbers of apical peduncles; 17, number of peduncles with fruits; 18, average peduncle length; 19, presence of trichomes on stems; 20, pollen biomass; 21, seed biomass; 22, seed number; 23, above-ground biomass; 24, individual seed weight; 25, pollen allocation; 26, seed allocation; 27, sex allocation.

although  $F_1$  individuals produced similar numbers of anthers and pollen grains to their parents (see [Supplementary Data Fig. S2A–F](#)), they sired only two seeds on diploid females in a growth chamber, whereas hexaploid hermaphrodites from either parental P- or P+ populations sired a large number of seeds under broadly similar conditions in a glasshouse.

#### Phylogenetic reconstruction

**Plastid sequences.** Phylogenetic analysis of a concatenation of *trnL-trnF* and *matK-trnK* chloroplast sequences strongly supports a clade that includes all but one of the P- monoecious populations. Within this clade, all individuals from populations characterized by the P- morphology were clustered with those from tetraploid populations of *M. annua*, as well as those from two hexaploid populations from Morocco and Catalonia ([Fig. 6](#)). All but one of the P+ monoecious individuals formed a separate clade, which included those of two other hexaploid *M. annua* populations from Seville and from Catalonia, as well as from diploid populations of *M. annua* and tetraploid *M. canariensis*. The exceptional individual from the female-like monoecious population P-<sub>HOR</sub> falls into the largely P+ clade, while the exceptional individual from the male-like monoecious population P+<sub>BAN</sub> falls into a clade with individuals of the P- phenotype.

**Nuclear sequences.** Given that all P+ and P- individuals sampled were probably hexaploid, they should each have more than one *ITS2* sequence copy. Indeed, phylogenetic reconstruction based on the nuclear marker *ITS2* points to an allopolyploid hybrid origin for the hexaploid lineages, with all P+ or P- individuals having *ITS2* copies from both diploid and tetraploid *M. annua*, on the one hand, and diploid *M. huetii*, on the other ([Fig. 7](#)). This result confirms the findings of [Obbard et al. \(2006a\)](#) for hexaploid *M. annua* generally. Furthermore, within the clade of sequences representing the diploid *M. annua* *ITS2* copy, phylogenetic analysis grouped together individuals of tetraploid *M. canariensis*, tetraploid *M. annua*, and all hexaploid *M. annua* populations, including those with both P+ and P- morphology. All hexaploid individuals of *M. annua* fell into a clade with *M. huetii*, which excluded individuals of *M. canariensis* and tetraploid *M. annua* ([Fig. 7](#)).

## DISCUSSION

#### Origin and taxonomy of a new biological species in the *M. annua* species complex

Three lines of evidence suggest that individuals of *M. annua* with the P+ phenotype belong to a different biological species from those with the more widely distributed P- phenotype,



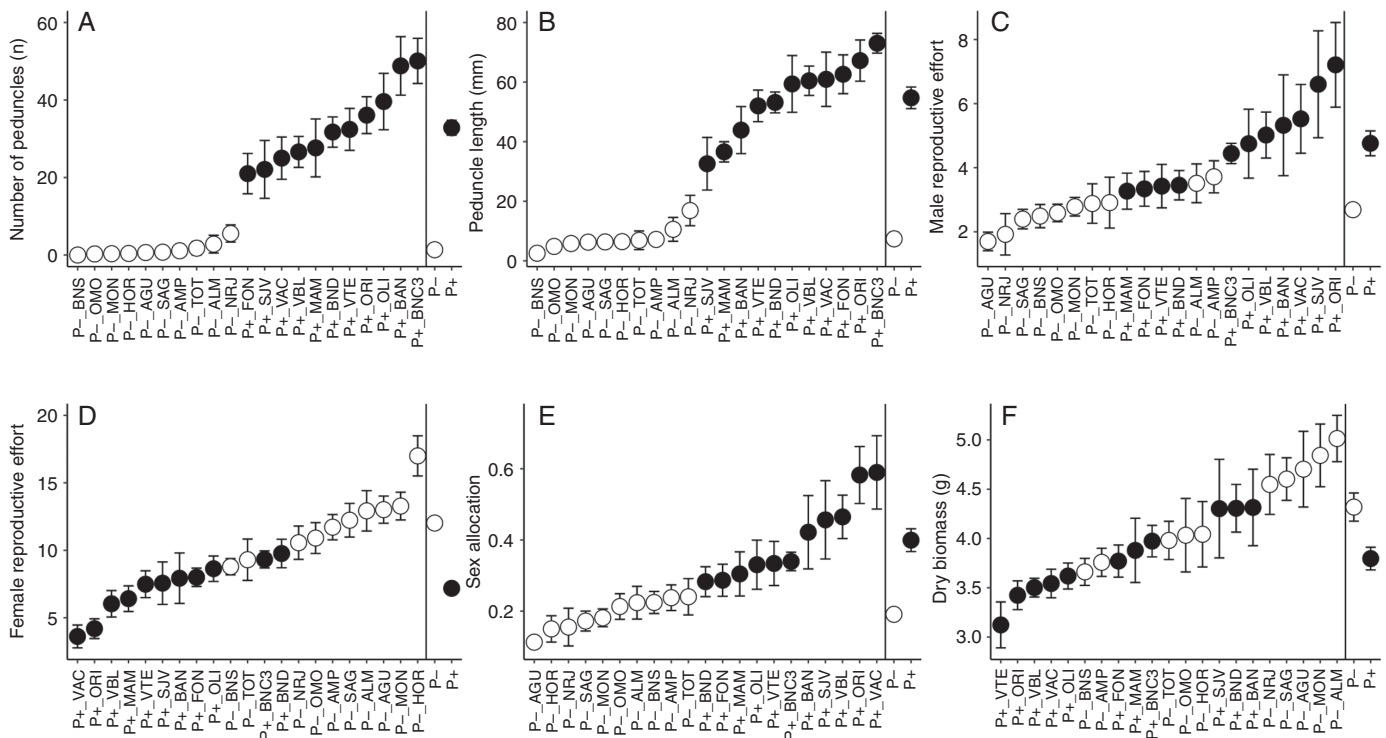


Fig. 3. Morphological trait measurements (mean and standard error) of P+ and P- monoecious hexaploid populations of *Mercurialis annua* from the common-garden experiment: (A) number of peduncles, (B) peduncle length, (C) male reproductive effort, (D) female reproductive effort, (E) sex allocation, and (F) above-ground biomass. Grey and black circles denote P- and P+ populations, respectively.

which occurs either in purely monoecious or, together with males, in androdioecious populations: (1) individuals of the P+ lineage are morphologically distinct from those of the P- lineage; (2)  $F_1$  progeny produced in crosses between the two phenotypes are largely (but not completely) female-sterile and show signs of male sterility, too; and (3) individuals of the two phenotypes fall largely into separate clades of a chloroplast-based gene tree. We also found that  $F_1$  progeny from crosses between the two phenotypes were substantially larger than the non-hybrid progeny. This result may point to heterosis or hybrid vigour, consistent with both parental lineages being inbred (Semel et al., 2006; Hochholdinger and Hoecker, 2007; Wang et al., 2015), and as suggested by the low genetic diversity observed in P+ populations (Santos del Blanco et al., 2019), but also in many P- populations (Obbard et al., 2006b; Pujol et al., 2009, 2010). Alternatively, it is possible that some of the vigour of hybrid individuals is attributable to trade-offs resulting from their lower reproductive effort.

Both biological species of *Mercurialis* studied here appear to be hexaploid hybrids, probably allohexaploids generated from a cross between *M. annua* and *M. huetii*. In contrast to the chloroplast gene tree, which separates the two hexaploid lineages, the nuclear gene tree fails to group them into different clades. This pattern might be the result of incomplete lineage sorting, or of recent or ongoing gene flow between the two lineages. In the absence of further sampling, it is not possible to distinguish between these two possibilities (Maddison and Knowles, 2006; Joly et al., 2009; Zhou et al., 2017). However, incomplete lineage sorting appears to be a less likely explanation if both lineages arose through a narrow bottleneck, as expected for

new polyploids (e.g. Willyard et al., 2009). The possibility of ongoing gene flow would be consistent with the current broad sympatry of the two lineages.

Polyploidization and hybridization have played an important role in the diversification of the annual clade of *Mercurialis* (Obbard et al., 2006a), as they have for flowering plants in general (Pannell et al., 2004; Adams and Wendel, 2005; Otto, 2007; Soltis et al., 2014b). Diploid *M. annua* and its diploid sister species *M. huetii* have previously been hypothesized as the original parents of two tetraploid lineages (monoecious *M. annua*, currently found in south-western Morocco; and dioecious *M. canariensis*, currently endemic to the Canary Islands), as well as one hexaploid lineage (androdioecious *M. annua*, i.e. the P- lineage, widespread in northern Morocco and around the Iberian Peninsula) (Obbard et al., 2006b). The two diploid *Mercurialis* species hybridize easily, but hybrid progeny suffer high sterility in both their male and their female functions (Russell and Pannell, 2015). The present study now suggests that the P+ hexaploid lineage, comprising the morphologically and functionally distinct male-like monoecious individuals, has also been generated by genome duplication and hybridization. The remarkable feature of this process of diversification in *Mercurialis* is that both pairs of independently derived polyploid lineages (the two tetraploid and the two hexaploid lineages) have contrasting sexual systems: tetraploids that are either dioecious or monoecious; and hexaploids that are either androdioecious (males coexisting with female-like P- monoecious individuals), monoecious P+ (individuals with their distinct male-like inflorescence) or monoecious P- (populations with female-like P- monoecious individuals without males).

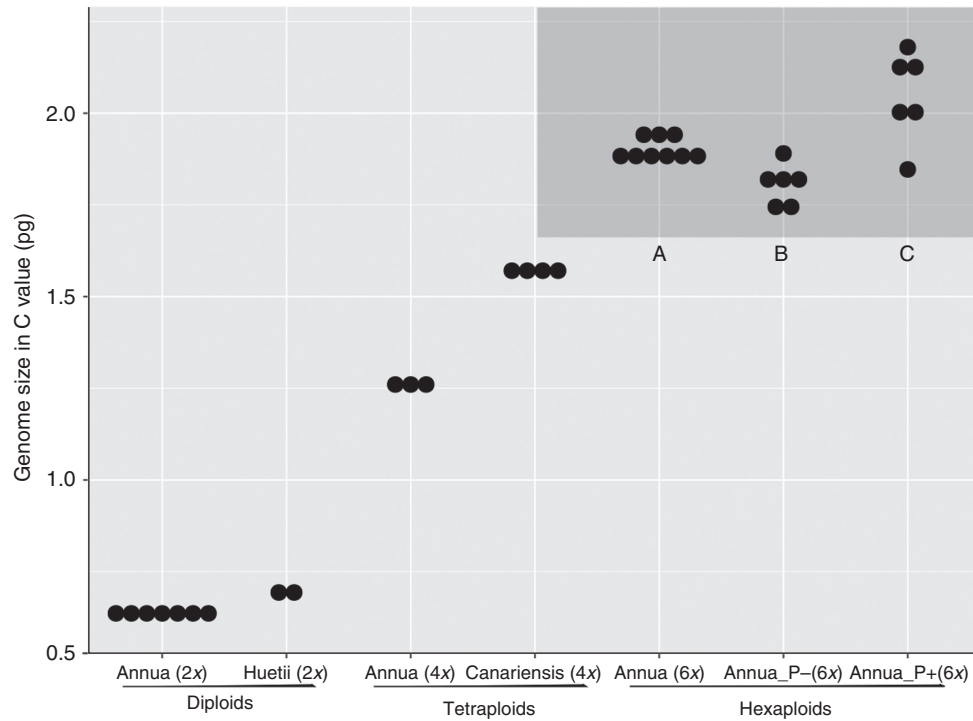


FIG. 4. Mean genome size (measured as C value of DNA content, in pg) of different *Mercurialis* lineages, with ploidy levels indicated below each species. Each solid dot is the mean value of several individuals from one population. Data from *Mercurialis* lineages are from [Obbard et al. \(2006a\)](#). The grey area indicates DNA content of hexaploid lineages, and A, B and C denote significance at  $P < 0.001$  for the corresponding pairwise comparisons.

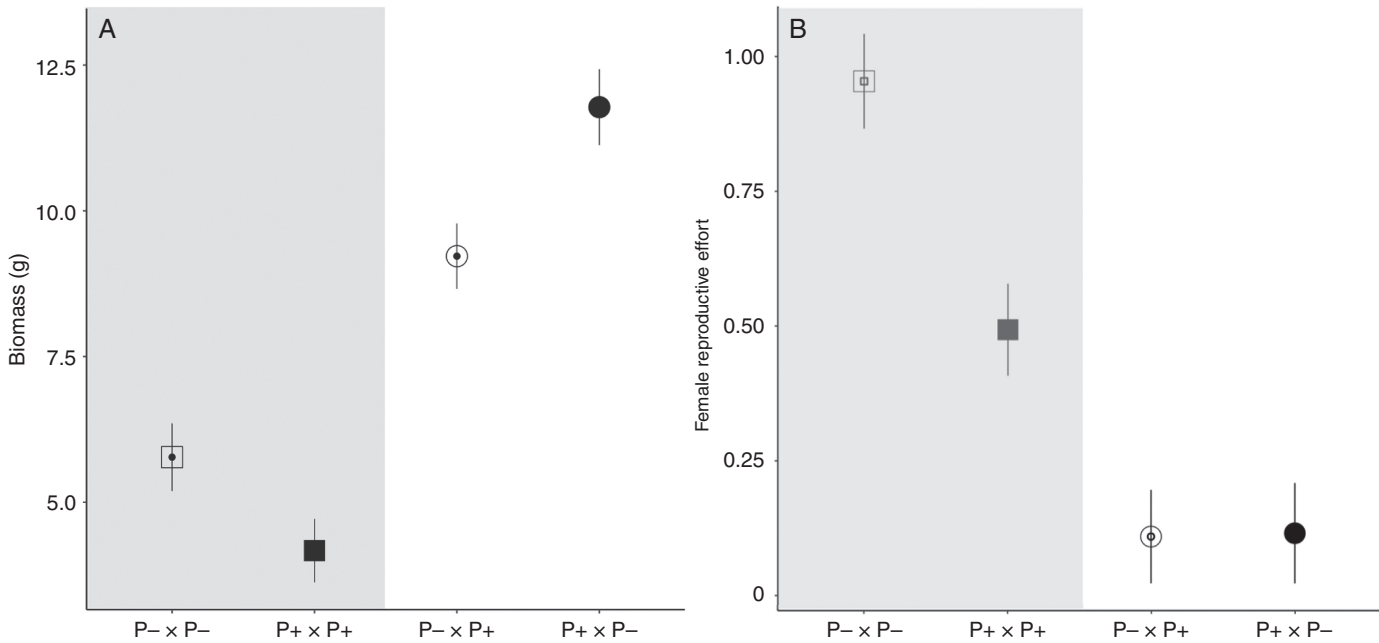


FIG. 5. (A) Above-ground biomass and (B) female reproductive effort for offspring from control crosses (in the grey area, square symbol) and reciprocal crosses (round symbol) between individuals with female-like inflorescence morphology (P-) and male-like inflorescence morphology (P+). The maternal parent is indicated first in each cross type, with symbols for P- and P+ being open and closed, respectively.

Further sampling, both of genotypes and of more nuclear loci, will be required to establish with more confidence the evolutionary paths that led to the origin of the two different hexaploid lineages. [Figure 8](#) recapitulates the scheme proposed by [Obbard et al., \(2006a\)](#) for the relationships between the diploid,

the tetraploid and the P- hexaploid lineages of the annual mercuries, modified with hypotheses for the origin of the P+ lineage. At least two scenarios seem plausible. One scenario is that the P- and P+ lineages are the result of two independent polyploid hybridization events involving ancestors of the current



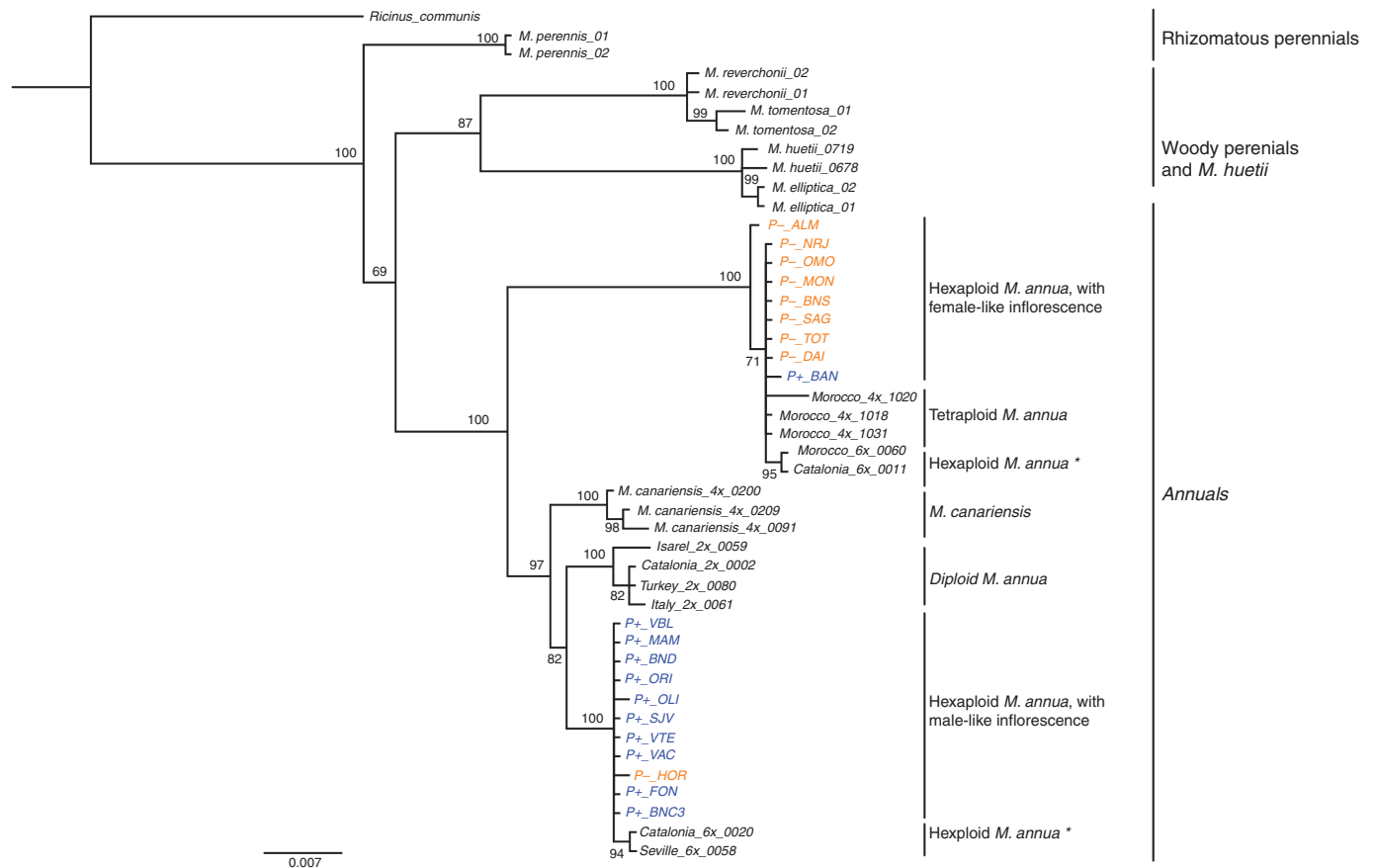


FIG. 6. MrBayes consensus tree based on the concatenated *trnL-trnF* and *matK-trnK* chloroplast datasets, using the best-fit partition model of nucleotide substitution with 10 million MCMC burn-in iterations. P- and P+ monoecious populations are highlighted in orange and blue, respectively. Asterisks denote hexaploid *Mercurialis annua* populations for which inflorescence structure was not measured. The scale bar at the bottom are in the units of number of base substitutions per site.

diploid dioecious lineages *M. huetii* and *M. annua* (Fig. 8A). The other scenario is that both lineages are the result of a single polyploid hybridization event that generated an ancestral hexaploid lineage, and that this population subsequently diversified into the two separate P- and P+ lineages (Fig. 8B). The current overlapping geographical distributions of these two lineages in south-eastern Spain would seem to argue more against the latter scenario, but it is plausible that the two lineages originated or diverged in allopatry and only subsequently came to occupy their current range, with occasional recent gene flow between them explaining the observed sharing of ITS variation.

Regarding the origin of the two hexaploid lineages, it does not seem possible to attribute much significance to their association with distinct chloroplast haplotypes; this pattern might simply be the result of differential chloroplast capture from a related lineage subsequent to divergence. Chloroplast capture, which results from hybridization between lineages with different chloroplast haplotypes and subsequent introgression by repeated asymmetrical backcrossing (Jackson *et al.*, 1999; Tsitrone *et al.*, 2003; Acosta and Premoli, 2010), has been common in angiosperm evolution and is a probable cause of incongruence between chloroplast gene trees and species trees in many phylogenies (Mort *et al.*, 2002; Fehrer *et al.*, 2007; Baldwin *et al.*, 2011; Dong *et al.*, 2012). Indeed, *M. huetii* itself

has a chloroplast that is divergent from that of its sister species *M. annua*, and that probably introgressed from perennial species of the genus (Obbard *et al.*, 2006a). Thus, although a more complete picture of the evolutionary origin of the hexaploid phenotypes of *M. annua* will require analysis of more markers across the genome, the current limited sampling of both chloroplast and nuclear markers already points to a complex history involving genome duplication, hybridization and potentially on-going gene flow and introgression among diverging lineages. It is of course also possible that the P+ population in the P- clade is the result of convergent evolution.

The taxonomy and systematics of the genus *Mercurialis* are clearly in need of revision. Several different names refer to the various annual lineages of *Mercurialis*. The two diploid dioecious lineages, *M. annua* and *M. huetii*, would appear to be species with acceptable names; although having the same ploidy, these species are morphologically and genetically distinct, and their hybrids are largely sterile (although viable). The nomenclature of the tetraploid and hexaploid lineages that probably descend from ancestors of *M. annua* and *M. huetii*, however, is inconsistent and problematic. These lineages have variously been grouped together under the same label as *M. annua* or *M. ambigua*. The epithet *ambigua* unifies tetraploid and hexaploid lineages that are morphologically

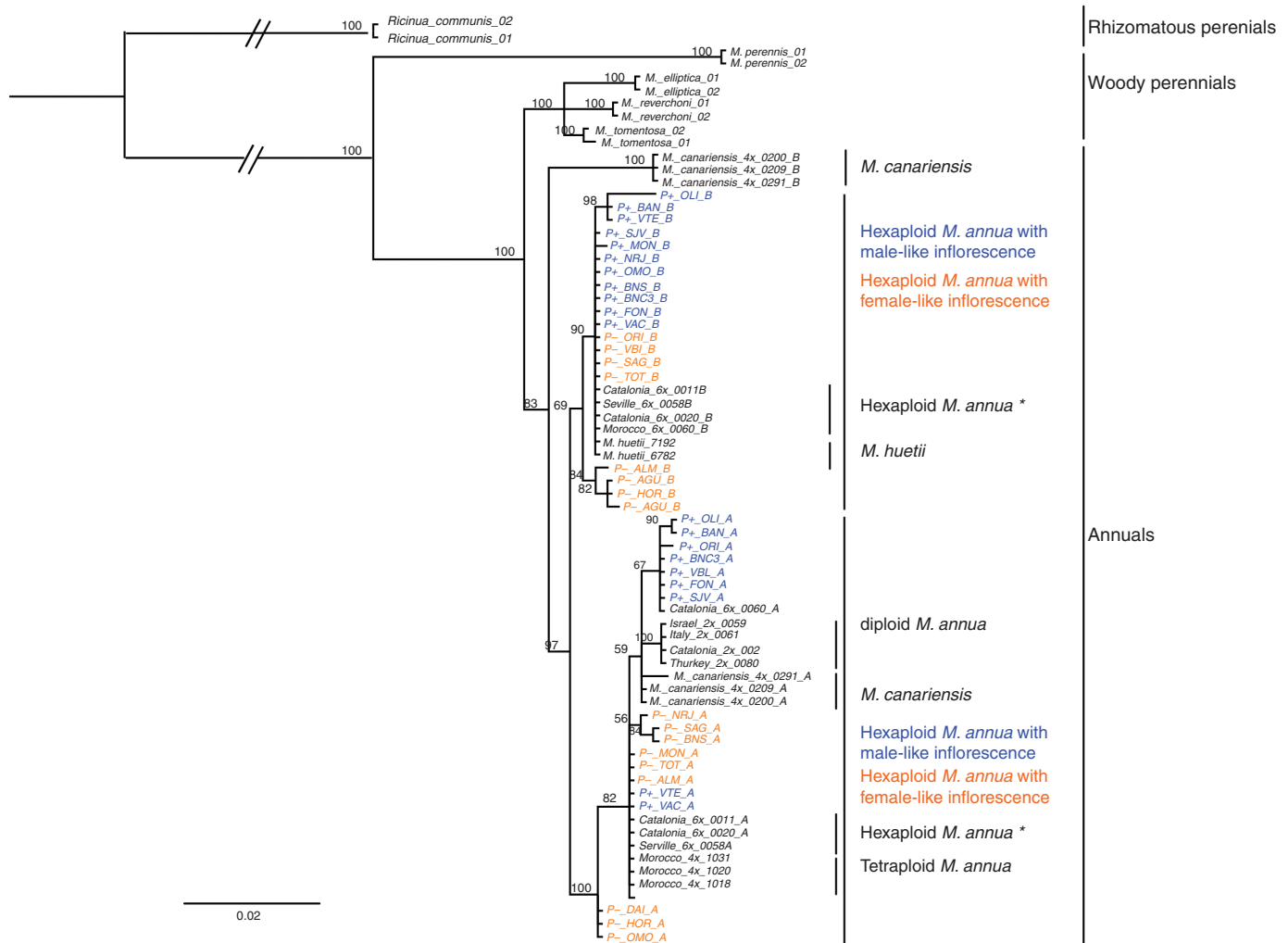


FIG. 7. MrBayes consensus tree based on the *ITS* datasets, using the best-fit partition model of nucleotide substitution with 10 million MCMC burn-in iterations. P- and P+ monoecious populations are highlighted in orange and blue, respectively. The scale bar at the bottom are in the units of number of base substitutions per site.

difficult to distinguish (Obbard *et al.*, 2006a), but that are reproductively isolated from one another (Russell and Pannell, 2015). By contrast, the epithet *annua* unifies dioecious diploid *M. annua* with largely monoecious tetraploid and hexaploid lineages, i.e. includes one of their putative diploid parents (*M. annua*) with the tetraploid and hexaploid lineages, but excludes the other putative progenitor (*M. huetii*) as well as tetraploid dioecious *M. canariensis*. Moreover, both *annua* and *ambigua* also exclude lineages with higher ploidy levels that occur in North Africa, Corsica and Sardinia, which have been sometimes referred to as *M. monoica* (Durand, 1963; Pannell, 1997b; Pannell *et al.*, 2004). The current study now reveals that the hexaploid populations of *M. annua* (*sensu lato*) comprise at least two different lineages that are morphologically more distinct from one another than is the common P- phenotype from tetraploid *M. annua* (*sensu lato*) (Obbard *et al.*, 2006a). Although it may have been convenient to refer loosely to the *M. annua* species complex with the exclusion of *M. huetii*, a comprehensive revision of the clade would now be useful, not least because the lineages referred to are both reproductively isolated from one another and morphologically and functionally distinct.

#### Maintenance of ploidy and reproductive diversity in the *M. annua* species complex

Although the different lineages of the *M. annua* species complex, including *M. huetii*, differ in their sexual systems, they occupy very similar ruderal and roadside habitats. How then are they being maintained? This question is particularly pressing for lineages on the Iberian Peninsula that are in broad sympatry with one another, not least because they are mobile colonizers of disturbed habitats. The problem of coexistence applies particularly to the newly characterized P+ phenotype, but it is general to the whole species complex. For example, in both north-eastern and north-western Spain, diploid dioecious *M. annua* is parapatric to hexaploid *M. annua*, whose populations are almost exclusively monoecious at the zones of contact. Here, diploid individuals appear to be physiologically superior to their hexaploid counterparts, but are also displacing them as a result of asymmetrical pollen swamping and hybridization that results in sterile progeny (Buggs and Pannell, 2006). Indeed, historical data indicate that this demographic displacement is occurring at a rate of several kilometres per year in both zones of contact and might be expected to lead to the ultimate

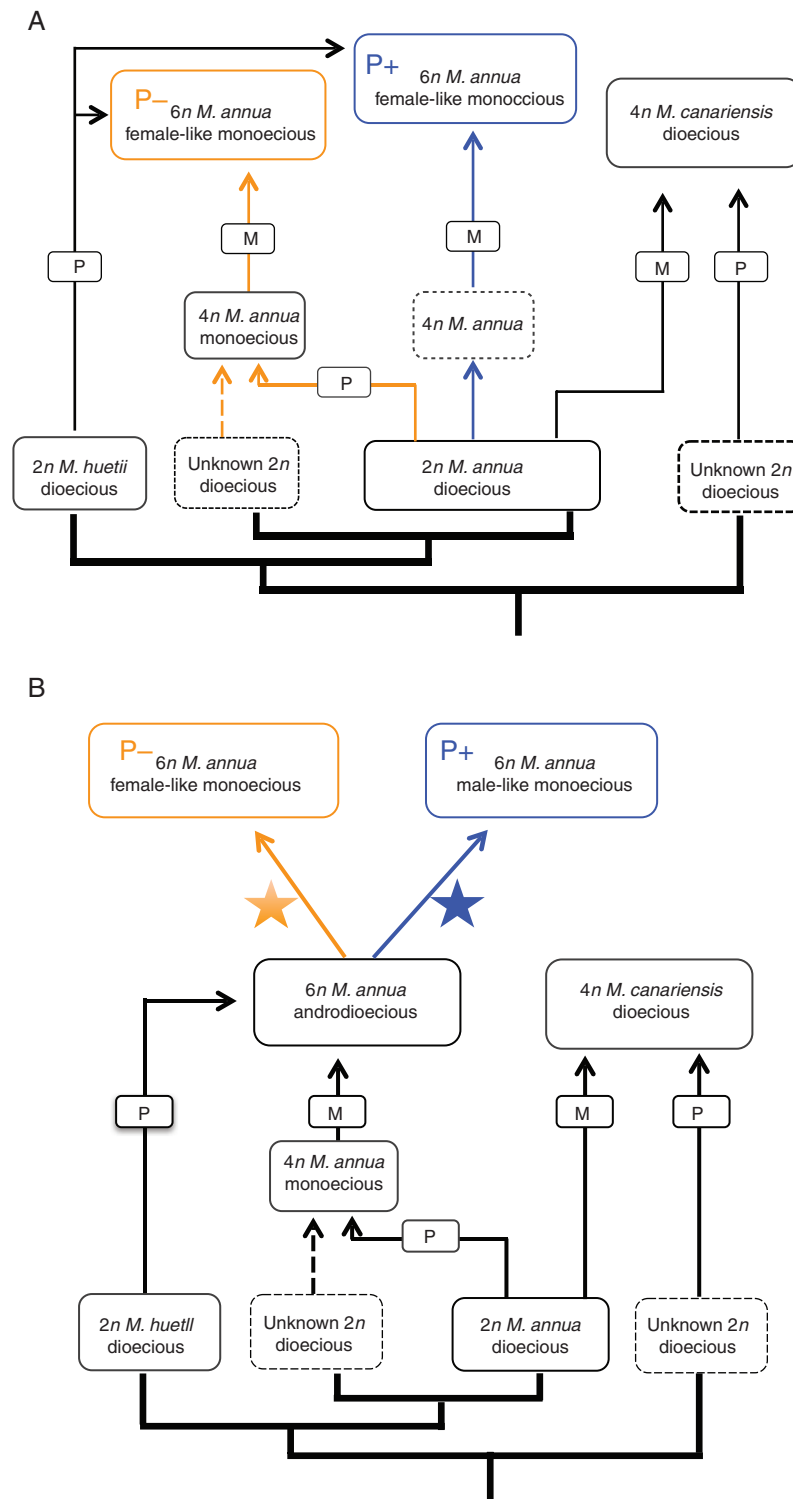


FIG. 8. Modified scheme of two possible scenarios for the evolutionary origin of the P+ hexaploid monoecious phenotype within the *Mercurialis* species complex. Scenario A describes two independent evolutionary origins of P+ and P- hexaploid monoecious phenotypes. Here, the P+ hexaploid phenotype is derived from a possible autopolyploidization of diploid dioecious *M. annua* to form an unknown tetraploid *Mercurialis* species (indicated with blue lines), followed by hybridization with a diploid *M. huetii* male. The P- hexaploid lineage is then possibly derived from hybridization between an unknown maternal diploid *Mercurialis* species and a diploid *M. annua* male, followed by autopolyploidization to form tetraploid *M. annua* (orange lines), finally followed by hybridization with a diploid *M. huetii* male. Scenario B depicts a single origin for both P+ and P- phenotypes. Here, the evolutionary origin of hexaploid androdioecious *M. annua* follows the hypothesis of *Obbard et al. (2006a)*. Both P- and P+ evolved from hexaploid androdioecious *M. annua*, during which reproductive isolation evolved, with chloroplast capture between different hexaploid lineages or between hexaploid and diploid *M. annua*. Figure modified from *Obbard et al. (2006a)*. A star associated with orange or blue denotes a different diversifying process.



of demise of P– hexaploid *M. annua* entirely, as a result of diploid expansion (Buggs and Pannell, 2006).

Given the superiority in both pollen production and dispersal ability of the P+ over the P– hexaploid hermaphrodites (Santos del Blanco *et al.*, 2019), we might expect that, over time, the P+ phenotype will eventually displace the P– phenotype as a result of pollen swamping similar to that experienced by the P– lineage when it encounters dioecious diploid *M. annua* (Buggs and Pannell, 2006). Indeed, in the common garden experiment in which P+ and P– individuals were grown together at equal frequencies, the great majority of seeds produced by both phenotypes were sired by P+ individuals (Supplementary Data Table S2). We speculate that displacement of the P– by the P+ lineage should be even more rapid than by diploids, because only half of all diploid individuals (i.e. the males) produce pollen. Determining whether the P+ lineage could resist displacement by diploids, or could even displace the diploids themselves, will require array experiments similar to those conducted by Buggs and Pannell (2006) and Santos del Blanco *et al.* (2019).

A plausible alternative scenario to the demographic displacement of P– by the P+ phenotype through pollen swamping is the selective introgression into the genomic background of the P– lineage of genes responsible for the pedunculate inflorescence. As discussed above, there is already some evidence for this possibility from the genotypes observed in our small sample in this study: one individual with a male-like inflorescence had a chloroplast haplotype more typically associated with the P– lineage (and vice versa). Given the much higher siring success, and probably greater fitness, of individuals with male-like inflorescences (Santos del Blanco *et al.*, 2019), the selective introgression of the genes responsible would seem highly likely, and it is surprising that the male-like inflorescence has not already swept through the range of the P– phenotype. Such selective introgression is well known for other plant traits. For instance,  $C_4$  photosynthesis that is advantageous under conditions of high insolation, high temperature and drought is known to have introgressed into formerly  $C_3$  plants that had presumably moved into environments in which  $C_4$  photosynthesis was favoured (Christin *et al.*, 2008; Gowik and Westhoff, 2011). Similarly, advantageous mutations altering that function of the key protein involved in photosynthesis, Rubisco, appear to have moved among species of the genus *Schiedea* by selective introgression (Kapralov *et al.*, 2011). Other similar examples of selective introgression have been recorded in animals (Albertson *et al.*, 2003; Quinn *et al.*, 2007).

#### *Independent origins of polyploidy as a further engine for diversification?*

The *M. annua* species complex is remarkable for its sexual-system variation, but it is wholly unremarkable in terms of its ploidy variation: within-species variation in ploidy is a hallmark of many plant species (Lewis, 1980; Soltis *et al.*, 2007; Beest *et al.*, 2012). Indeed, genome duplication has contributed substantially to plant diversification in general (Stebbins, 1950; Levin, 1983; Soltis *et al.*, 2014a), probably both by expanding or shifting the environmental niches that new polyploid populations can occupy (Beest *et al.*, 2012), but also by establishing instantaneous reproductive isolation between a newly derived polyploid population and its lower-ploidy progenitor species

(Ramsey and Schemske, 1998; Seehausen, 2004; Husband and Sabara, 2011). Our study indicates that populations with the same ploidy level but with different origins may also be reproductively isolated from one another, although the breakdown of homoploid reproductive barriers with rising ploidy is known for some lineages (Ehrendorfer, 1996; Lafon-Placette *et al.*, 2017).

To the extent that rapid genetic changes immediately following genome duplication and/or hybridization associated with polyploidy should occur independently in lineages with independent origins, our conclusion should not be so surprising – although the polyphyletic origins of one and the same ploidy level are perhaps an underappreciated engine of plant diversification. In the case of *M. annua* hexaploids, the occupancy of P– and P+ lineages of very similar habitats and their coincident flowering phenologies are likely to bring about the displacement of one lineage by the other due to pollen swamping (just as the P– lineage is currently being displaced by the diploid lineages in northern Spain; Buggs and Pannell, 2006), diminishing the likelihood of a net positive effect on diversification. However, a scenario of multiple origins of polyploidy might have a more positive effect on long-term diversification when divergent populations are maintained in allopatry by occupying different ecological niches, or when reproductive interference is avoided through the evolution of prezygotic isolation. Such a scenario would seem to be more likely in animal- than in wind-pollinated species (e.g. Moyle *et al.*, 2004; Rieseberg *et al.*, 2006; Rieseberg and Willis, 2007; Lowry *et al.*, 2008).

#### SUPPLEMENTARY DATA

Supplementary data are available online at <https://academic.oup.com/aob> and consist of the following. Fig. S1: Sampling locations for both P– and P+ monoecious hexaploid populations of *M. annua* along the east coast of the Iberian Peninsula. Fig. S2: Comparison of male flowers between  $F_1$  hybrids of P+ and P– lineages and hexaploid monoecious individuals and between these  $F_1$  hybrids and hexaploid males. Table S1: Sampled populations, their original location, abbreviation and the experiments carried out in this study. Table S2: Source populations for reciprocal crosses between P+ and P– phenotypes, and the genotypes of four microsatellite markers.

#### FUNDING

This work was supported by the Sinergia grant CRSII3\_147625 from the Swiss National Science Foundation to J.R.P., Nicolas Perrin and Mark Kirkpatrick.

#### ACKNOWLEDGEMENTS

Sarah Beuvier and Laure Olazuaga assisted in the wet-lab and with phenotyping plants, Nicolas Ruth helped care for plants in the glasshouse, Jörn Gerchen helped with phenotyping plants, Xinji Li provided some hexaploid pollen data and help with pollen counting, Martha Serrano helped with phylogenetic tree construction, and Lucía de Soto Suárez provided advice on analysis of leaf traits. The computations were performed at the Vital-IT Center for high-performance computing of the SIB Swiss Institute of Bioinformatics (<http://www.vital-it.ch>). We thank members of the Pannell lab for stimulating discussion and anonymous reviewers for helpful suggestions on the manuscript.

## LITERATURE CITED

- Acosta CM, Premoli AC. 2010.** Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). *Molecular Phylogenetics and Evolution* **54**: 235–242.
- Adams KL, Wendel JF. 2005.** Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8**: 135–141.
- Albertson RC, Streelman JT, Kocher TD. 2003.** Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proceedings of the National Academy of Sciences* **100**: 5252–5257.
- Baduel P, Bray S, Vallejo-Marin M, Kolář F, Yant L. 2018.** The “polyploid hop”: shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers in Ecology and Evolution* **6**: 117.
- Baldwin BG, Kalisz S, Scott Armbruster W. 2011.** Phylogenetic perspectives on diversification, biogeography, and floral evolution of *Collinsia* and *Tonella* (Plantaginaceae). *American Journal of Botany* **98**: 731–753.
- Beest M, Roux JJ Le, Richardson DM, Brysting AK, Suda J, Kubes M. 2012.** The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany* **109**: 19–45.
- Bikard D, Patel D, Le Mette C, et al. 2009.** Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323**: 623–626.
- Brochmann C, Xiang QY, Brunsfeld SJ, Soltis DE, Soltis PS. 1998.** Molecular evidence for polyploid origins in *Saxifraga* (Saxifragaceae): The narrow arctic endemic *S. svalbardensis* and its widespread allies. *American Journal of Botany* **85**: 135–143.
- Buggs RJA, Pannell JR. 2006.** Rapid displacement of a monoecious plant lineage is due to pollen swamping by a dioecious relative. *Current Biology* **16**: 996–1000.
- Chen ZJ, Ni Z. 2006.** Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *BioEssays* **28**: 240–252.
- Chen ZJ, Ha M, Soltis D. 2007.** Polyploidy: Genome obesity and its consequences. *New Phytologist* **174**: 717–720.
- Christin P, Salamin N, Muasya AM, Roalson EH, Russier F, Besnard G. 2008.** Evolutionary switch and genetic convergence on *rbcl* following the evolution of C 4 photosynthesis. *Molecular Biology and Evolution* **25**: 2361–2368.
- David P, Pujol B, Viard F, Castella V, Goudet J. 2007.** Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology* **16**: 2474–2487.
- Dong W, Liu J, Yu J, Wang L, Zhou S. 2012.** Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* **7**: 1–9.
- Durand B. 1963.** Le complexe *Mercurialis* annual L. s.l.: une étude biosystématique. *Annales des Sciences Naturelles, Botanique, Paris* **12**: 579–736.
- Ehrendorfer F. 1996.** Reproductive biology in island plants. *Philosophical Transactions of The Royal Society of London B: Biological Sciences* **351**: 725–733.
- Fehrer J, Gemeinholzer B, Chrtek J, Bräutigam S. 2007.** Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in *Pilosella* hawkweeds (Hieracium, Cichorieae, Asteraceae). *Molecular Phylogenetics and Evolution* **42**: 347–361.
- Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E. 1983.** Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**: 1049–1051.
- Glennon KL, Ritchie ME, Segraves KA. 2014.** Evidence for shared broad-scale climatic niches of diploid and polyploid plants. *Ecology Letters* **17**: 574–582.
- Govik U, Westhoff P. 2011.** The path from C3 to C4 photosynthesis. *Plant Physiology* **155**: 56–63.
- Hare EE, Johnston JS. 2012.** Genome size determination using flow cytometry of propidium iodide-stained nuclei. In: Orgogozo V, Rockman MV (eds.). *Molecular methods for evolutionary genetics*. Humana Press, New York, 3–12.
- Hegarty MJ, Hiscock SJ. 2008.** Genomic clues to the evolutionary success of polyploid plants. *Current Biology* **18**: R435–R444.
- Hesse E, Pannell JR. 2011a.** Sexual dimorphism in androdioecious *Mercurialis annua*, a wind-pollinated herb. *International Journal of Plant Sciences* **172**: 49–59.
- Hesse E, Pannell JR. 2011b.** Sexual dimorphism in a dioecious population of the wind-pollinated herb *Mercurialis annua*: the interactive effects of resource availability and competition. *Annals of Botany* **107**: 1039–1045.
- Hochholdinger F, Hoecker N. 2007.** Towards the molecular basis of heterosis. *Trends in Plant Science* **12**: 427–432.
- Husband BC, Sabara HA. 2011.** Reproductive isolation and their diploid autotetraploids *Chamerion* progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist* **161**: 703–713.
- Husband BC, Schemske DW, Burton TL, Goodwillie C. 2002.** Pollen competition as a unilateral reproductive barrier between sympatric diploid and tetraploid *Chamerion angustifolium*. *Proceedings of the Royal Society B: Biological Sciences* **269**: 2565–2571.
- Jackson HD, Steane DA, Potts BM, Vaillencourt RE. 1999.** Chloroplast DNA evidence for reticulate evolution in *Eucalyptus* (Myrtaceae). *Molecular Ecology* **8**: 739–751.
- Joly S, McLenachan PA, Lockhart PJ. 2009.** A statistical approach for distinguishing hybridization and incomplete lineage sorting. *The American Naturalist* **174**: E54–E70.
- Kapralov MV, Kubien DS, Andersson I, Filatov DA. 2011.** Changes in Rubisco kinetics during the evolution of C4 photosynthesis in *Flaveria* (Asteraceae) are associated with positive selection on genes encoding the enzyme. *Molecular Biology and Evolution* **28**: 1491–1503.
- Kearse M, Moir R, Wilson A, et al. 2012.** Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics Applications Note* **28**: 1647–1649.
- Korbecka G, Rymer PD, Harris SA, Pannell JR. 2010.** Solving the problem of ambiguous paralogy for marker loci: microsatellite markers with diploid inheritance in allohexaploid *Mercurialis annua* (Euphorbiaceae). *Journal of Heredity* **101**: 504–511.
- Korbecka G, Hamilton A, Pannell JR. 2011.** Mixed mating in androdioecious *Mercurialis annua* inferred using progeny arrays and diploid-acting microsatellite loci in a hexaploid background. *Annals of Botany* **107**: 1057–1061.
- Lafon-Placette C, Johannessen IM, Hornslien KS, et al. 2017.** Endosperm-based hybridization barriers explain the pattern of gene flow between *Arabidopsis lyrata* and *Arabidopsis arenosa* in Central Europe. *Proceedings of the National Academy of Sciences* **114**: E1027–E1035.
- Lanfear R, Calcott B, Ho SYW, Guindon S. 2012.** PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**: 1695–1701.
- Larkin MA, Blackshields G, Brown NP, et al. 2007.** Clustal W and Clustal X version 2.0. *Bioinformatics Applications Note* **23**: 2947–2948.
- Levin DA. 1983.** Polyploidy and novelty in flowering plants. *The American Naturalist* **122**: 1–25.
- Lewis WH. 1980.** Polyploidy in species populations. In: Lewis WH (ed.). *Polyploidy: biological relevance*. Plenum, New York, 103–144.
- Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF. 2001.** Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* **44**: 321–330.
- Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008.** The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 3009–3021.
- Ma W-J, Kuijper B, de Boer JG, et al. 2013.** Absence of complementary sex determination in the parasitoid wasp genus *Asobara* (Hymenoptera: Braconidae). *PLoS One* **8**: e60459.
- Maceira NO, Jacquard P, Lumaret R. 1993.** Competition derivative between diploid and glomerata for autotetraploid *Dactylis glomerata* L. from Galicia. Implications for the establishment of novel polyploid populations. *New Phytologist* **124**: 321–328.
- Maddison WP, Knowles LL. 2006.** Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* **55**: 21–30.
- Maloo JN, Nozue K, Mumbach MR, Palmer CM. 2013.** LeafJ: an ImageJ plugin for semi-automated leaf shape measurement. *Journal of Visualized Experiments* **71**: e50028.
- Mort M, Soltis DE, Soltis PS, Francisco-Ortega J. 2002.** Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. *Systematic Botany* **27**: 271–288.
- Moyle LC, Olson MS, Tiffin P. 2004.** Patterns of reproductive isolation in three angiosperm genera. *Evolution* **58**: 1195–1208.
- Obbard DJ, Harris SA, Buggs RJA, Pannell JR. 2006a.** Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). *Evolution* **60**: 1801–1815.
- Obbard DJ, Harris SA, Pannell JR. 2006b.** Sexual systems and population genetic structure in an annual plant: testing the metapopulation model. *The American Naturalist* **167**: 354–366.

- Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell* **131**: 452–462.
- Pannell J. 1997a. Variation in sex ratios and sex allocation in androdioecious *Mercurialis annua*. *Journal of Ecology* **85**: 57–69.
- Pannell J. 1997b. Widespread functional androdioecy in *Mercurialis annua* L. (Euphorbiaceae). *Biological Journal of the Linnean Society* **61**: 95–116.
- Pannell J. 2008. The maintenance of gynodioecy and androdioecy in a metapopulation. *Evolution* **51**: 10–20.
- Pannell JR, Obbard DJ, Buggs RJA. 2004. Polyploidy and the sexual system: What can we learn from *Mercurialis annua*? *Biological Journal of the Linnean Society* **82**: 547–560.
- Pujol B, Zhou S, Sanchez J, Pannell JR. 2009. Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences* **106**: 15379–15383.
- Pujol B, Obbard DJ, Pannell JR. 2010. Symptoms of population range expansion: lessons from phenotypic and genetic differentiation in hexaploid *Mercurialis annua*. *Plant Ecology and Diversity* **3**: 103–108.
- Quinn TP, Hodgson S, Flynn L, Hilborn R, Rogers DE. 2007. Direction selection by fisheries and the timing of sockeye salmon (*Oncorhynchus nerka*). *Ecological Applications* **17**: 731–739.
- R Core Team. 2017. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29**: 467–501.
- Rieseberg LH, Willis JH. 2007. Plant speciation. *Science* **317**: 910–914.
- Rieseberg LH, Wood TE, Baack EJ. 2006. The nature of plant species. *Nature* **440**: 524–527.
- Ronquist F, Teslenko M, Van Der Mark P, et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539–542.
- Russell JRW, Pannell JR. 2015. Sex determination in dioecious *Mercurialis annua* and its close diploid and polyploid relatives. *Heredity* **114**: 262–271.
- Santos del Blanco L, Tudor E, Pannell JR. 2019. Low siring success of females with an acquired male function illustrates the legacy of sexual dimorphism in constraining the breakdown of dioecy. *Ecology Letters* **22**: 486–497.
- Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440**: 341–345.
- Scannell DR, Butler G, Wolfe KH. 2007a. Yeast genome evolution - the origin of the species. *Yeast* **24**: 929–942.
- Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH. 2007b. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. *Proceedings of the National Academy of Sciences* **104**: 8397–8402.
- Seehausen O. 2004. Hybridization and adaptive radiation. *Trends in Ecology & Evolution* **19**: 198–207.
- Semel Y, Nissenbaum J, Menda N, et al. 2006. Overdominant quantitative trait loci for yield and fitness in tomato. *Proceedings of the National Academy of Sciences* **103**: 12981–12986.
- Soltis DE, Soltis PS. 1999. Polyploidy : recurrent formation and genome evolution. *Trends in Ecology & Evolution* **14**: 348–352.
- Soltis DE, Soltis PS, Tate JA. 2004. Advances in the study of polyploidy since plant speciation. *New Phytologist* **161**: 173–191.
- Soltis DE, Soltis PS, Schemske DW, et al. 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *International Association for Plant Taxonomy* **56**: 13–30.
- Soltis DE, Albert VA, Leebens-Mack J, et al. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* **96**: 336–348.
- Soltis DE, Visger CJ, Soltis PS. 2014a. The polyploidy revolution then ... and now: Stebbins revisited. *American Journal of Botany* **101**: 1057–1078.
- Soltis PS, Soltis DE. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences* **97**: 7051–7057.
- Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* **60**: 561–588.
- Soltis PS, Liu X, Marchant DB, Visger CJ, Soltis DE. 2014b. Polyploidy and novelty: Gottlieb's legacy. *Philosophical Transactions of The Royal Society of London B: Biological Sciences* **369**: 20130351.
- Song K, Lu P, Tang K, Osborn TC. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 7719–7723.
- Stebbins GL. 1950. *Variation and evolution in plants*. New York: Columbia University Press.
- Thompson JD, Lumaret R. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology and Evolution* **7**: 302–307.
- Tsitrona A, Kirkpatrick M, Levin DA. 2003. A model for chloroplast capture. *Evolution* **57**: 1776–1782.
- Wang H, Fang Y, Wang L, et al. 2015. Heterosis and differential gene expression in hybrids and parents in *Bombyx mori* by digital gene expression profiling. *Scientific Reports* **5**: 8750.
- Willyard A, Cronn R, Liston A. 2009. Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution* **52**: 498–511.
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences* **106**: 13875–13879.
- Zhou Y, Duvaux L, Ren G, Zhang L, Savolainen O, Liu J. 2017. Importance of incomplete lineage sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. *Heredity* **118**: 211–220.