

Research



Cite this article: Gerchen JF, Veltsos P, Pannell JR. 2022 Recurrent allopolyploidization, Y-chromosome introgression and the evolution of sexual systems in the plant genus *Mercurialis*. *Phil. Trans. R. Soc. B* **377**: 20210224.
<https://doi.org/10.1098/rstb.2021.0224>

Received: 2 November 2021
Accepted: 3 January 2022

One contribution of 15 to a theme issue ‘Sex determination and sex chromosome evolution in land plants’.

Subject Areas:
evolution

Keywords:
sex chromosome, phylogenetics, androdioecy

Author for correspondence:
J. F. Gerchen
e-mail: jorn.gerchen@unil.ch

[†]Present address: Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas, USA.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5893446>.

Recurrent allopolyploidization, Y-chromosome introgression and the evolution of sexual systems in the plant genus *Mercurialis*

J. F. Gerchen, P. Veltsos[†] and J. R. Pannell

Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

JFG, 0000-0002-0821-7912; PV, 0000-0002-8872-6281; JRP, 0000-0002-0098-7074

The plant genus *Mercurialis* includes dioecious, monoecious and androdioecious species (where males coexist with hermaphrodites). Its diversification involved reticulate evolution via hybridization and polyploidization. The Y chromosome of the diploid species *Mercurialis annua* shows only mild signs of degeneration. We used sequence variation at a Y-linked locus in several species and at multiple autosomal and pseudoautosomal loci to investigate the origin and evolution of the Y chromosome across the genus. Our study provides evidence for further cases of allopolyploid speciation. It also reveals that all lineages with separate sexes (with one possible exception) share the same ancestral Y chromosome. Surprisingly, males in androdioecious populations of hexaploid *M. annua* carry a Y chromosome that is not derived from either of its two putative progenitor lineages but from a more distantly related perennial dioecious lineage via introgression. These results throw new light on the evolution of sexual systems and polyploidy in *Mercurialis* and secure it as a promising model for further study of plant sex chromosomes.

This article is part of the theme issue ‘Sex determination and sex chromosome evolution in land plants’.

1. Introduction

Dioecious plants provide valuable material for studying the earliest stages of sex-chromosome evolution [1,2]. Because dioecy is usually derived from hermaphroditism, their sex chromosomes must postdate that transition and may thus often be relatively young. In contrast to animals, in which separate sexes are ancestral and conserved and the same genes are thus often involved in determining sex (reviewed in [3]), the genetic details of sex determination in plants tend to differ among species [4–8]. The convergent evolution of dioecy in flowering plants thus offers opportunities for understanding not only why and how separate sexes evolve but also how sex determination and sex chromosomes evolve in different independent lineages.

Populations or species with sexual system variation point to particularly recent evolutionary transitions that allow close comparisons of sex chromosomes in different contexts. Several such cases have been studied in some detail. The quantitative dimensions of sex allocation, including reversions from dioecy to hermaphroditism, were explored by Lloyd [9] in the genus *Lep tinella* (syn. *Cotula*), which shows variation among lineages between combined and separate sexes. Important insights have also been gained from other systems with similar sexual-system variation (e.g. *Wurmbea* [10,11], *Sagittaria* [12,13], *Ecballium* [14,15] and *Mercurialis* [16,17]). Among-species variation in sex determination has been studied in the poplar and willow family (Salicaceae) and in wild strawberries, *Fragaria*. In the Salicaceae, sex-chromosome evolution

has involved the independent adoption of homologues of a single gene with a role in sex determination in different species [7,18,19], as in some animal groups. In octoploid *Fragaria*, femaleness is determined by a cassette of genes whose chromosomal position differs between species, probably owing to translocations between homeologous chromosomes [20], and with the continued maintenance of separate sexes during diversification. Dioecy was also probably conserved during divergence of two closely related lineages of *Rumex hastatulus*, where one lineage retained an ancestral XY sex determination system and the other acquired a second Y chromosome through an X-autosome fusion [21–23].

Most work on sexual system transitions and the evolution of sex chromosomes has concerned the evolution of dioecy from hermaphroditism, which was long regarded as an evolutionary dead end [24]. However, it is now clear that reversions to hermaphroditism have been frequent [25,26], probably driven by selection of self-fertile hermaphrodites that confer reproductive assurance [27–29]. Something like this is likely to have occurred in *Vitis*, where hermaphrodites probably resulted from recombination between male- and female-sterility loci in a sex-determining region and domestication that favoured uniparental reproduction [30]. Similarly, cultivated *Carica papaya* is hermaphroditic, with a modified Y chromosome, whereas wild populations are dioecious [31]. In a process not unlike domestication, experimental evolution of dioecious populations of the plant *Mercurialis annua* demonstrated a rapid transition from dioecy to hermaphroditism in just a few generations via the selection of ‘leaky’ sex expression in females following the removal of males [32].

Another important question concerns the effect of polyploidy on sexual systems and sex determination [33], not least because so many angiosperms have a recent history of polyploidization [34,35]. Indeed, polyploidy is found in several plant clades in which sex chromosomes have been studied, e.g. *Salix* [36], *Silene* [37], *Rumex* [38] and *Fragaria* [20]. Genome duplication can lead to problems of meiosis and, in species in which sex-determination depends on the relative dosage of the sex-determining alleles, can cause failure of the sex-determining system, precipitating a transition from dioecy to hermaphroditism [33]. When polyploidization involves hybridization between two dioecious species, it is also pertinent to ask whether dioecy is maintained in the allopolyploid hybrid, and if so, which of the two progenitor species contributed the sex-determination locus (if indeed either of them did).

Here, we present a phylogenetic analysis of the evolution of sexual systems and sex chromosomes in the genus *Mercurialis* (Euphorbiaceae), in which sex chromosomes have been evolving in the context of transitions both between sexual systems and among ploidy levels [16,17,39]. *Mercurialis* is an almost exclusively European genus with several perennial and annual species, most of which are dioecious (table 1 and figure 1). Both perenniality and dioecy are ancestral in the genus, and monoecy has evolved in annual species that diversified in the context of both genome duplication and hybridization [17,40]. Diploid *M. annua* has an XY sex chromosome system, and crosses among several of the annual species suggest that they all have the same sex determination system and probably the same sex chromosomes [41]. Veltsos *et al.* [42,43] studied the genome and sex chromosomes of diploid *M. annua* and found evidence for a large but only mildly degenerate sex-determining region, and Li *et al.* [44] showed that YY males (that thus lacked an X chromosome) were fully viable,

though partially sterile, consistent with the low divergence between the X and Y. Our study builds on this work by asking how the *M. annua* sex chromosomes relate to those in the other *Mercurialis* species with separate sexes.

Our phylogenetic analysis represents an advance on those of Obbard *et al.* [17] and Ma *et al.* [40], which were based on sequence variation at nuclear internal transcribed spacer and plastid loci. By contrast, we considered the topologies of phylogenetic trees for multiple loci on both autosomes and the sex chromosomes, including a locus that amplifies only in males in most species of the genus and that thus allowed us to consider the topology of the sex-determining region itself. We were particularly interested in discovering the phylogenetic origin of androdioecy in hexaploid *M. annua* [45]. Androdioecy, a rare sexual system in both plants and animals in which males co-occur with hermaphrodites [46–48], has usually evolved from hermaphroditism via the breakdown of dioecy [28,49], but Obbard *et al.* [17] suggested that androdioecy in hexaploid *M. annua* may have evolved following hybridization between tetraploid monoecious *M. annua* and *Mercurialis huetii*, from which it was thought to have derived its Y chromosome. Our results clearly reject this hypothesis: the Y chromosome in hexaploid androdioecious *M. annua* appears to be the result of unusual sex-chromosome introgression from a more distantly related perennial species. Our analysis also points to previously unrecognized allopolyploidization in both annual and perennial *Mercurialis* species and indicates that dioecy was probably maintained through the process of polyploidization.

2. Material and methods

(a) Species sampled

Table 1 lists all species sampled in our study, and figure 1 shows the present distribution of the annual species, our prime focus. For lineages with males, we included one male and either one female or one monoecious individual. For diploid and hexaploid *M. annua*, we included samples from the eastern and western ranges of their distributions. We also included two recently described hexaploid monoecious individuals with male-like inflorescences [40]. Tetraploid *M. annua* has been previously described as monoecious [17,39], but we have since found a few males in tetraploid populations south of Casablanca, and we included one of these in our dataset (electronic supplementary material, table S2).

(b) Overview of the development of new phylogenetically informative loci

We developed 24 phylogenetically informative loci for our study. All but one of them amplified in all males, females and hermaphrodites, so that their polymerase chain reaction (PCR) products contained mixtures of multiple sequences representing alleles or diverged homeologous sequences in polyploids. The inferred chromosomal location of these loci was based on the linkage map developed by Veltsos *et al.* [42]; we refer them as either ‘autosomal’ (located on one of the autosomes), ‘pseudoautosomal’ (located on the sex chromosomes, but outside the region of suppressed recombination in diploid *M. annua* males) or ‘sex-linked’ (located on the sex chromosomes inside the region of suppressed recombination in diploid *M. annua* males). In addition, one locus amplified in males as a single copy, regardless of ploidy, indicating tight linkage to the sex-determining region on the Y chromosome across the genus. We refer to this locus as ‘male-specific’. All loci were based on genomic regions covered by

Table 1. *Mercurialis* lineages used in this study. (Two different chromosome counts were published for *Mercurialis elliptica*, the lower count is more likely to be correct but needs verification.)

lineage	sexual system	life history	chromosome count	ploidy
<i>Mercurialis annua</i>	dioecy	annual	16	diploid
<i>Mercurialis annua</i>	monoecy/androdioecy	annual	32	tetraploid
<i>Mercurialis annua</i>	monoecy/androdioecy	annual	48	hexaploid
<i>Mercurialis canariensis</i>	dioecy	annual	32	tetraploid
<i>Mercurialis huetii</i>	dioecy	annual	16	diploid
<i>Mercurialis reverchonii</i>	dioecy	perennial	26	tetraploid
<i>Mercurialis tomentosa</i>	dioecy	perennial	26	tetraploid
<i>Mercurialis elliptica</i>	dioecy	perennial	42 (220)	hexaploid

aligned exon-capture datasets independent of gene models, i.e. the amplified regions were mostly exonic (but could also include intronic or intergenic regions) and were not complete genes. Further details about the development of the loci are given in the electronic supplementary material.

(c) Development of phylogenetically informative loci on sex chromosomes and autosomes

Development and analysis of phylogenetically informative loci in allopolyploids requires accounting for multiple homeologous or allelic copies of the same sequence that must be distinguished to infer origins of allopolyploid subgenomes [50]. We thus used long-read sequencing to obtain reads that span the complete amplified region of PCR products without the need for cloning [51]. We then phased homeologous sequences from multiple loci, based on phylogenetic information and using an iterative approach, and generated multi-locus phylogenies with these loci. The pipeline used is described in the electronic supplementary material.

(d) Polymerase chain reaction amplification and sequencing of autosomal and sex-linked loci

We selected 24 primer pairs, 15 of which were located on the sex chromosome and one or two on each of the autosomes (electronic supplementary material, figure S3). We performed PCR with these (electronic supplementary material, table S1) for each of 24 samples of perennial and annual *Mercurialis* species (electronic supplementary material, table S2) using 1× Qiagen Hotstart PCR Buffer, 0.2 μm dNTPS, 0.2 μm forward and reverse primer and 0.001 U μl⁻¹ Hotstart Taq Polymerase (Qiagen). The PCR protocol was 15 min of initial denaturation at 95°C, followed by 35 cycles of amplification with initial denaturation for 30 s at 94°C, annealing for 30 s and elongation at 73°C for 1 min, and final elongation at 73°C for 5 min.

After amplification, we pooled PCR products for each sample in approximately equimolar amounts, based on the intensity of agarose gel bands, cleaned the pools using cleanNGS beads (Labgene) according to the manufacturer's instructions and quantified DNA using a Qubit fluorometer (Fisher Scientific). In a second step, we ligated individual barcodes to each of these pools and combined pools to obtain a single barcoded Oxford Nanopore sequencing library. We followed the recommendations for native barcoding of genomic DNA using the ligation sequencing kit and the native barcoding extension 1–12 and 13–24 from Oxford Nanopore. About 100 fmol of the final library were sequenced on an Oxford Nanopore 9.4.1 flowcell. Raw reads were base-called and demultiplexed using GUPPY 4.0.15 (Oxford Nanopore).

In polyploid samples, the resulting long reads represent a mixture of alleles from different homeologues. Owing to the high error rates of Oxford Nanopore reads, final alignments have to be based on a consensus of multiple reads representing unique template sequences, which requires clustering of raw reads. We developed a new approach, which clusters reads at each sample and locus using a custom Python script (02_cluster_pcr_reads; see the electronic supplementary material for details).

Our dataset could contain multiple sequences per locus for the polyploid genomes, representing fixed differences between, or allelic variation within, homeologous sequences. In a first step, we built single-locus phylogenies, which contained both types of sequences using RAxML 8.2.12 (parameters '-f d', '-d', '# 100', '-m GTRGAMMA', 1000 bootstrap replicates). These single-locus phylogenies may be limited in size and phylogenetic resolution. To overcome these limitations, we also built multi-locus phylogenies. These should represent differences between homeologues and should not include the within-locus allelic variation that was removed using a custom Python script (03_subset_alignment; see the electronic supplementary material for details).

In general, polyploid sequences cannot be simply concatenated for generating multi-locus phylogenies, because the phase of homeologous sequences from different loci of the same polyploid individual is typically unknown. We sequentially phased homeologous sequences for multiple sets of loci with a custom workflow implemented in Python and Biopython 1.79 (04_multilocus_phylogeny; see the electronic supplementary material for details). We ran this analysis for the full dataset including all loci and independently for three subsets of the autosomal, pseudoautosomal and sex-linked loci.

(e) Development of a male-specific polymerase chain reaction locus

We previously identified 17 exonic loci with Y-linked inheritance throughout the diploid *M. annua* species range, based on the published exon-capture dataset [52], and confirmed their Y linkage with PCR and Sanger sequencing [42]. We extended the PCR assay for one of these loci (g3639/gm56331, located on contig56631 of v. 1.3 of the *M. annua* genome assembly), using DNA from male and female/monoecious samples of annual and perennial lineages of the *M. annua* species complex. We then used the previously described newly developed exon-capture dataset to develop a longer sex-linked sequence at the same locus, which amplified in annual lineages as well as in perennial *Mercurialis reverchonii*, *Mercurialis tomentosa* and *Mercurialis elliptica*. Reads were aligned against v. 1.3 of the *M. annua* genome assembly using the previously described pipeline. We then generated consensus sequences based on aligned exon-capture reads at

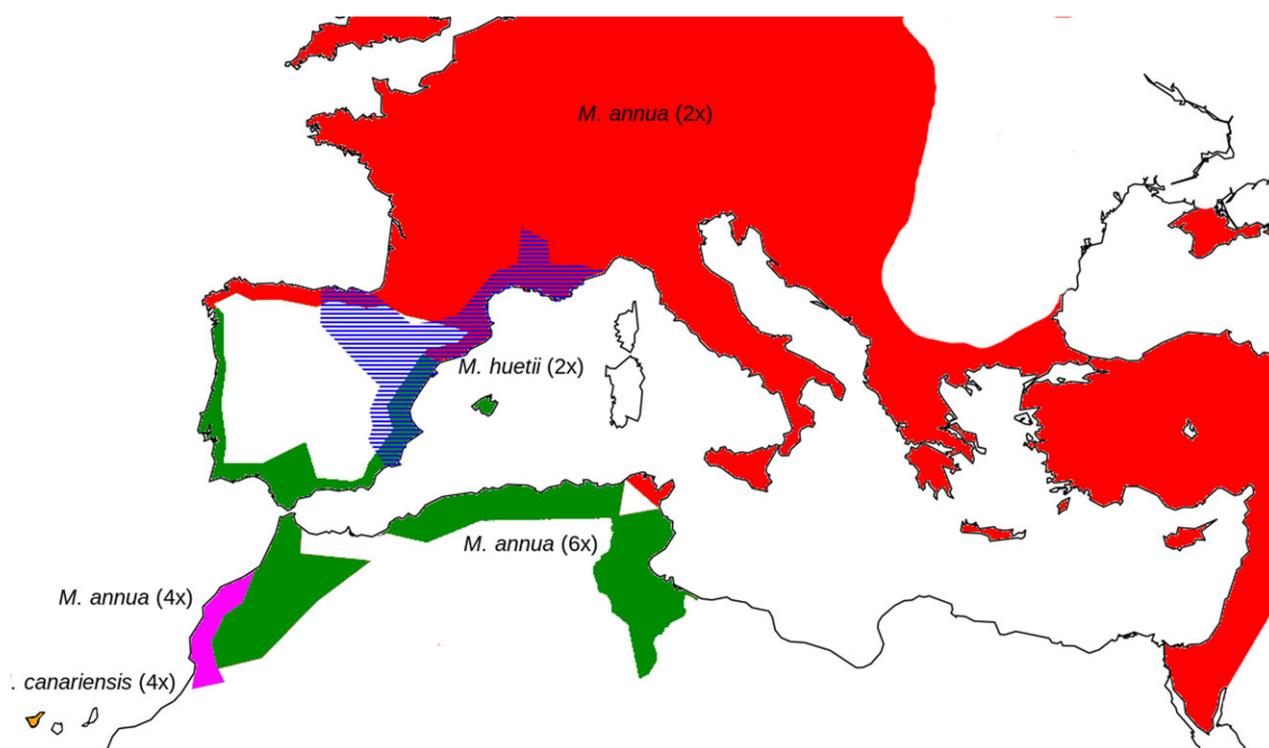


Figure 1. Distribution of lineages of annual *Mercurialis* species across Europe, North Africa and the Middle East. Red, diploid dioecious *Mercurialis annua*; purple, tetraploid monoecious *M. annua*; green, hexaploid androdioecious *M. annua*; orange, tetraploid dioecious *Mercurialis canariensis*; blue, diploid dioecious *Mercurialis huetii*.

the sex-linked contig for males from annual and perennial lineages of *Mercurialis*, using a pileup from the Python implementation of SAMTOOLS 1.9 [53] and marking variable sites with Ns. We used these sequences to search for PCR primers using primer3-plus [54] and developed four different primer pairs that target the same locus in different lineages (electronic supplementary material, table S1).

To test for sex-specificity of the loci, we applied them to males and females or monoecious (negative control) samples of annual and perennial *Mercurialis* species using 1× Hotstart PCR buffer (Qiagen), 0.2 µm dNTPS, 0.2 µm forward and reverse primer and 0.001 U µl⁻¹ Hotstart Taq Polymerase (Qiagen). The PCR protocol was 15 min of initial denaturation at 95°C followed by 35 cycles of amplification with initial denaturation for 30 s at 94°C, annealing at 60°C for 30 s and elongation at 73°C for 1 min. Final elongation was at 73°C for 5 min. Afterwards, we visualized bands using gel electrophoresis and sent a subset of PCR products of males that amplified successfully for Sanger sequencing to Microsynth (Balgach, Switzerland).

We aligned the *Mercurialis* sequences using MAFFT 7.475 [55] and aligned each of the diploid *M. annua* sequences against the NCBI nucleotide collection [56] using blastn [57] to identify the best unique outgroup sequence. We added the sequence of the resulting best BLAST hit to the alignment, realigned the sequences using MAFFT and generated a phylogenetic tree using RAxML 8.2.12 (parameters '-m GTR' and 100 bootstrap replicates) [58].

3. Results

(a) Phylogenetic inference based on the autosomal and sex-chromosome loci

We generated a total 2.2 Gb of Oxford nanopore long-reads from PCR amplicons of the newly developed

phylogenetically informative loci from autosomal and sex-linked loci (electronic supplementary material, table S1 and figure S3). Upon inspection of alignments of demultiplexed reads to the *M. annua* genome, we noted that a significant proportion of reads for primer pair 1_24 were much shorter than expected, based on the location of primers at the *M. annua* genome assembly. We interpreted this as a result of non-specific PCR amplification and excluded this locus from further analyses. For the remaining loci, we excluded a total of 13 (2.4%) low-coverage samples that had a sequencing depth smaller than 10 times the ploidy at the target locus. The full concatenated and phased alignment, which included all 23 remaining loci, was 22 928 bp long.

The phylogenetic tree resulting from analysis of the autosomal, pseudoautosomal and sex-linked sequences includes five major monophyletic clades with a bootstrap support of 100, each of which contained homeologous sequences from multiple polyploid species (figure 2). The most basal clade relative to the *M. perennis* outgroup contains sequences from males and females from the three perennial lineages *M. reverchonii*, *M. elliptica* and *M. tomentosa* (perennial clade 1). Annual clade 1 contains sequences from each of the tetraploid and hexaploid *M. annua* lineages as well as from *Mercurialis canariensis*. In addition, perennial clade 2 contains sequences from all samples from the perennial lineages. Annual clade 2 contains sequences from the male and female *M. huetii* samples and all hexaploid *M. annua* individuals as well as sequences from one male and female perennial *M. elliptica*. Annual clade 3 contains sequences from all *M. annua* individuals of either sex and ploidy level (including the newly described monoecious hexaploids with male-like inflorescences; [40]), as well as from both male and female *M. canariensis* samples (figure 2).

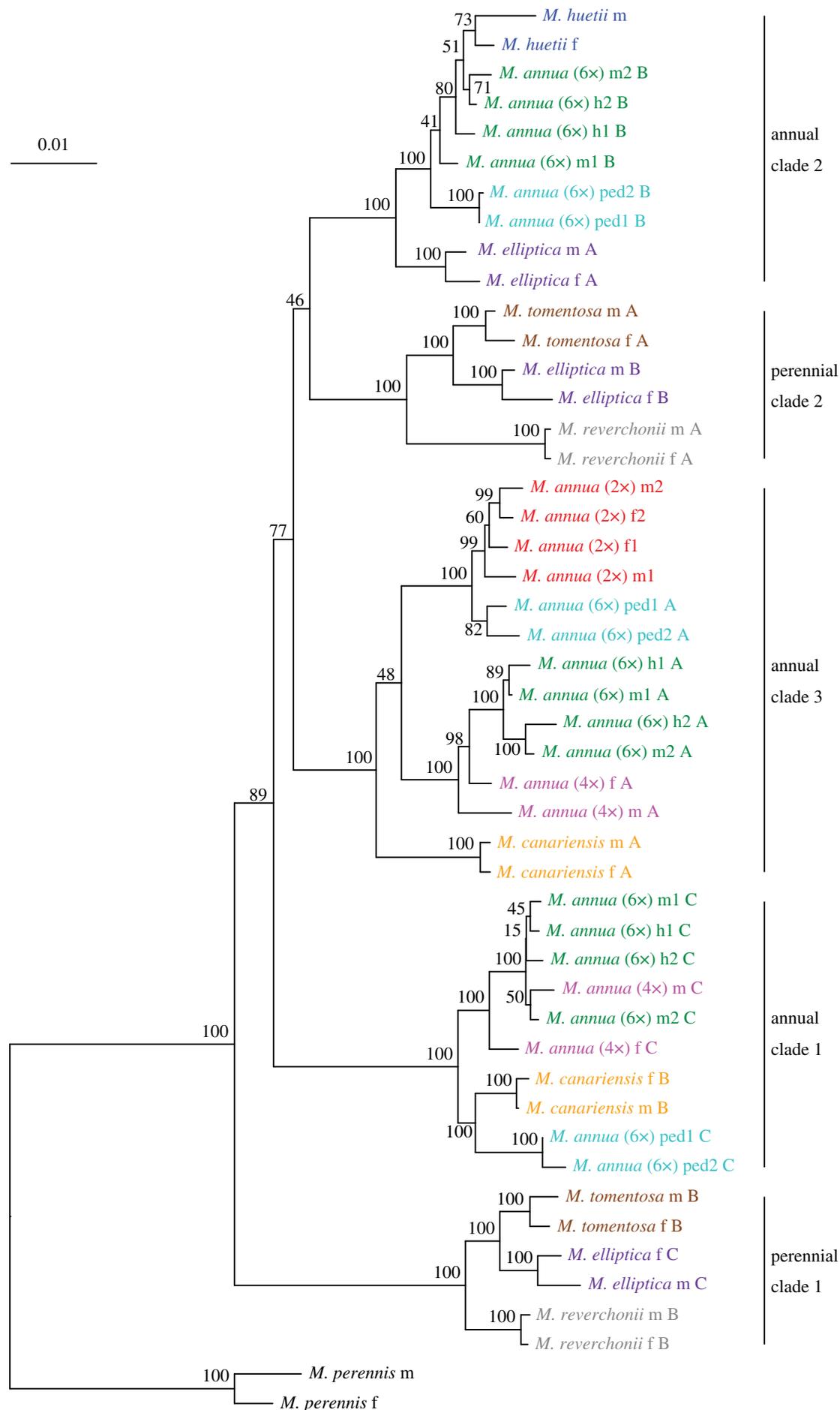


Figure 2. Multi-locus phylogeny based on 23 phylogenetically informative loci located on the sex chromosomes and autosomes. Values next to nodes indicate bootstrap support. The scale bar indicates the substitution rate. m, f, h and ped at the end of species names at the tips describe male, female, monoecious individuals and monoecious individuals with male-like inflorescences, respectively. A, B and C indicate different homeologous sequences from the same polyploid individual.

All five clades were well-supported in the phylogenetic trees, based on the three subsets of the full dataset, i.e. autosomal, pseudoautosomal and sex-linked loci (electronic supplementary material, figures S4–S29), with the exception of annual clade 3 in the autosomal set of loci, which did not contain sequences of male and female *M. canariensis*; these were placed as a sister group to annual clade 1. Our analysis is consistent with that of Obbard *et al.* [17] in a number of respects, not least in the inference that both dioecy and perenniality are ancestral.

Homeologous sequences from all polyploid species are found in multiple monophyletic clades, implying an allopolyploid origin for all of them. The result confirms inferences by Obbard *et al.* [17] and implies additional allopolyploidization events. Obbard *et al.* [17] inferred that tetraploid *M. annua* was an autopolyploid, but the presence of its sequences in both annual clade 3 and annual clade 1 suggest that it is allopolyploid. In addition, the two hexaploid lineages of *M. annua* are both allopolyploid, each with three progenitors, one from each of the three annual clades. This contrasts with Obbard *et al.*'s [17] inference that all hexaploid populations were the result of hybridization between an ancestor of the inferred autotetraploid *M. annua* and an ancestor of diploid *M. huetii*. Finally, sequences of all three perennial lineages are found in diverged clades and are thus allopolyploids too. A single sequence of *M. elliptica* was found in annual clade 2, also indicating allopolyploidy.

While the lineages contained in five major clades could be clearly defined, the phylogenies based on subsets of data showed some differences in the relationships among clades. Relationships among clades were the same for the pseudoautosomal and full datasets (figure 2; electronic supplementary material, figure S5). In the dataset based on autosomal loci, the placement of annual clade 2, perennial clade 2 and annual clade 1 differed from the full dataset (electronic supplementary material, figure S4). Here, annual clade 2 and annual clade 3 formed a sister clade, while the perennial clade 2 and annual clade 3 formed a sister clade in the full dataset. Analysis of the sex-linked loci pointed to two monophyletic clades (electronic supplementary material, figure S6), one with the perennial species and annual clade 2, and the other with annual clade 1 and annual clade 3. Uncertainties in the relative placement of these clades are reflected in the low bootstrap support at some of the deeper nodes in the phylogenetic trees.

(b) Phylogenetic inference based on the male-specific locus

The previously described male-specific locus for diploid *M. annua* was successfully amplified in a male-specific fashion in diploid dioecious *M. annua* and *M. huetii*, tetraploid dioecious *M. canariensis* and hexaploid androdioecious *M. annua* (electronic supplementary material, figure S1). This is, to our knowledge, the first evidence of a common Y chromosome in the *M. annua* species complex. Analysis of the longer sequence (707 bp) at the same locus also showed Y-specific amplification among most species with males (figure 3). We could not amplify the locus in the rare males in tetraploid *M. annua*, nor in monoecious individuals with pedunculate inflorescences of hexaploid *M. annua*, which Ma *et al.* [40] found was phylogenetically distinct from androdioecious hexaploid *M. annua*.

Alignment of the male-specific sequences revealed two deletions and one insertion that were specific to hexaploid *M. annua*, *M. reverchonii*, *M. elliptica* and *M. tomentosa* but absent in the other annual lineages. The best BLAST hit ($E = 1 \times 10^{-117}$, 73% identity) of the male-specific *Mercurialis* sequence was a messenger RNA (XM_002528314.3) from *Ricinus communis*, with its function predicted as serine-rich adhesin for platelets. Consistent with the pattern of indels shown in figure 3, the tree based on the male-specific locus placed hexaploid *M. annua* into a clade with the perennial *Mercurialis* species, with *M. elliptica* its closest inferred relative, while the other annual lineages formed a distinct clade (figure 3). This topology differs from that inferred on the basis of all other loci (figure 2), in which the hexaploid lineage is absent from both of the two main perennial clades (which include sequences from all three perennial lineages). Together, these results imply that the part of the Y chromosome of hexaploid *M. annua* must have been acquired through introgression from a lineage in the perennial clade, probably a close ancestor of *M. elliptica*. Significantly, the autosomal and pseudoautosomal loci, including the six sex-linked loci, shared a similar topology (electronic supplementary material, figures S5, S6 and S15–S20), indicating that only a small part of the Y chromosome including the sex-determining locus was introgressed.

4. Discussion

Our phylogenetic analyses of *Mercurialis*, based on sequence variation at 24 autosomal and sex-linked loci (one tightly linked to the sex-determination locus), suggest the evolutionary scenario summarized in figure 4. This scenario: (i) points to additional allopolyploidization events in the genus, permitting a more nuanced assessment of the relationship between sexual systems and ploidy; and (ii) allows both the identification of the source of the Y chromosome in two of the annual allopolyploid species as well as a phylogenetic explanation for the origin of androdioecy in hexaploid *M. annua*.

(a) The robustness of our approach for phylogenetic inference

To distinguish sequences at the same locus that represent informative differences between homeologues and those that represent differences between alleles deal with the substantial challenges of phylogeny reconstruction involving allopolyploids [50], we assumed that divergence between allelic sequences should be smaller than between homeologous sequences and that the number of expected homeologous sequences should be equal to half the sample ploidy. When comparing the divergence between sequence pairs for putative allelic sequences and homeologous sequences in our dataset, there was not always a complete separation between both sequence types (electronic supplementary material, figure S30). This could result from natural variation in the divergence between homeologous sequences across loci and individuals. However, if sequences representing one homeologue were missing from the dataset, allelic sequences could be mistaken for homeologues, which could have strong effects on the inference of phylogenies and the phasing of homeologous sequences based on them.

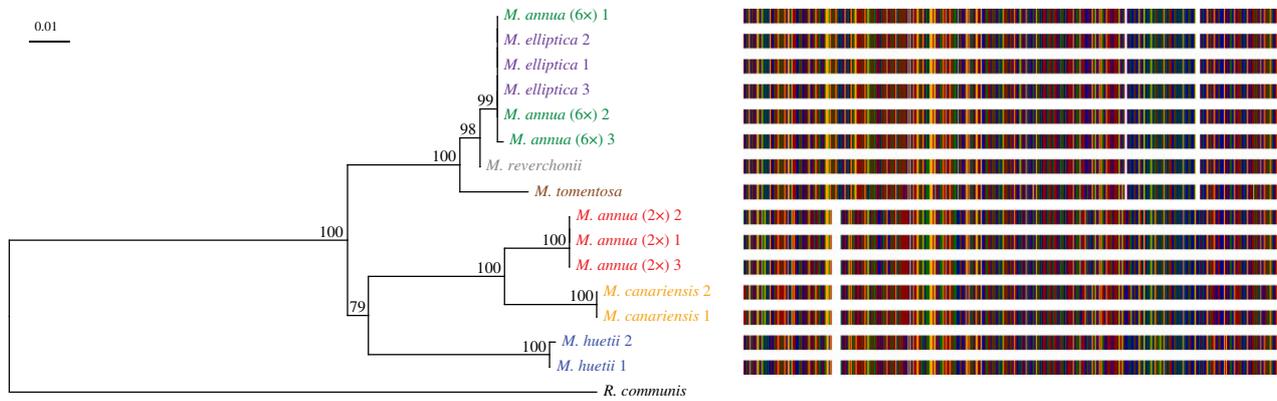


Figure 3. Phylogeny and alignment for the Y-specific PCR locus. Values next to nodes of the tree indicate bootstrap support, and the scale bar indicates the substitution rate. Bars on the right side of the tip labels represent the multiple sequence alignment for each sequence, with coloured bars representing nucleotides (red, A; blue, T; orange, C; green, G) and white areas indicating deletions.

Such dropout of homeologous sequences could occur if variation in the prime region at the target locus were in one of the homeologues, or in cases of gene conversion or structural variation between homeologues. Such a possibility seems unlikely to have affected our analysis, given the generally well-supported separation of homeologous sequences into discrete clades (electronic supplementary material, figure S4–S6). Only in the autosomal dataset did putative homeologous sequences for *M. canariensis* show a pattern incongruent with these well-defined clades (electronic supplementary material, figure S4).

(b) Allopolyploidy and sexual-system variation in *Mercurialis*

Our study confirms that a transition from a perennial to an annual life history preceded the transition among the annual species from dioecy to monoecy. It seems likely that this transition to monoecy coincided with the loss of the Y chromosome and the evolution in females of a male function in response to selection favouring enhanced 'leaky' sex expression [27,29], which remains common in dioecious *M. annua* [59,60]. Our analyses also confirm an allopolyploid origin for both hexaploid androdioecious *M. annua* involving the diploid *M. annua* and *M. huetii* lineages and tetraploid *M. canariensis* (involving the diploid *M. annua* and an unknown lineage) and indicate that tetraploid *M. annua* is an allopolyploid (and not an autopolyploid as suggested by Obbard *et al.* [17]), that tetraploid dioecious perennial *M. tomentosa* and *M. reverchonii* are allopolyploids with two divergent genomes and that *M. elliptica* is an allohexaploid with an additional genome sharing ancestry with both diploid *M. huetii* and hexaploid androdioecious *M. annua*.

Allopolyploidy has clearly played a major role in the diversification of *Mercurialis* but has not impacted sexual-system transitions in a straightforward way. Unlike the transition from diploidy to allotetraploidy in *M. annua*, which coincided with a transition to monoecy, none of the other polyploidization events resulted in the evolution of monoecy. Dioecy may have broken down with polyploidization in these lineages and then re-evolved, but this seems unlikely. For most of the allopolyploid lineages (apart from hexaploid *M. annua*, which we discuss below), the topology for the Y-linked locus is congruent with that for the autosomal and pseudoautosomal loci used, allowing the parsimonious

inference that the Y chromosome (and thus dioecy) was maintained through the ploidy transition.

Despite the loose association between ploidy and sexual system among annual species of *Mercurialis* (dioecious diploid species *M. annua* and *M. huetii*, and the largely monoecious tetraploid and hexaploid lineages of *M. annua*; [16,39]), there would appear to be no meaningful constraint linking these two traits at the genus level, and it is more reasonable to seek a functional explanation for the evolution of monoecy in polyploid *M. annua*. The fact that strong selection on sex allocation in experimental populations of diploid *M. annua* could bring about a transition from dioecy to monoecy in just a few generations [32] is consistent with this view: monoecy can just as well evolve in diploids. In angiosperms more generally, there is also no clear association between the sexual system and ploidy, with polyploidy associated with both transitions to separate sexes and combined sexes (reviewed in [33]).

Finally, it is clear that the X and Y chromosomes are not only shared among the annual species of *Mercurialis*, as inferred from simple between-species crosses by Russell & Pannell [41], but also that the same sex chromosomes are shared with perennial species *M. tomentosa*, *M. reverchonii* and *M. elliptica*. The ancestral XY system has thus evidently been conserved during the genus' diversification and, significantly, there is no evidence of sex-chromosome turnover in *Mercurialis*. This finding sets a foundation for a comparative analysis of the same Y chromosome that has been evolving independently among different species.

(c) The origin of the Y chromosome in allopolyploid species of *Mercurialis*

Our study allows us to infer which progenitor species contributed the Y chromosome to the allopolyploid species that have retained or regained separate sexes in *Mercurialis*. It is now clear that dioecious tetraploid *M. canariensis* derives its Y chromosome from the diploid *M. annua* lineage, the most widespread lineage of the genus with a distribution across much of Europe and around the Mediterranean Basin [39,61]. *Mercurialis canariensis* has only been sampled on the island of Tenerife several thousand kilometres from the nearest populations of its one extant progenitor, diploid *M. annua* in northern Spain [62]. Nevertheless, phylogeographical inference based on gradients in genetic diversity across

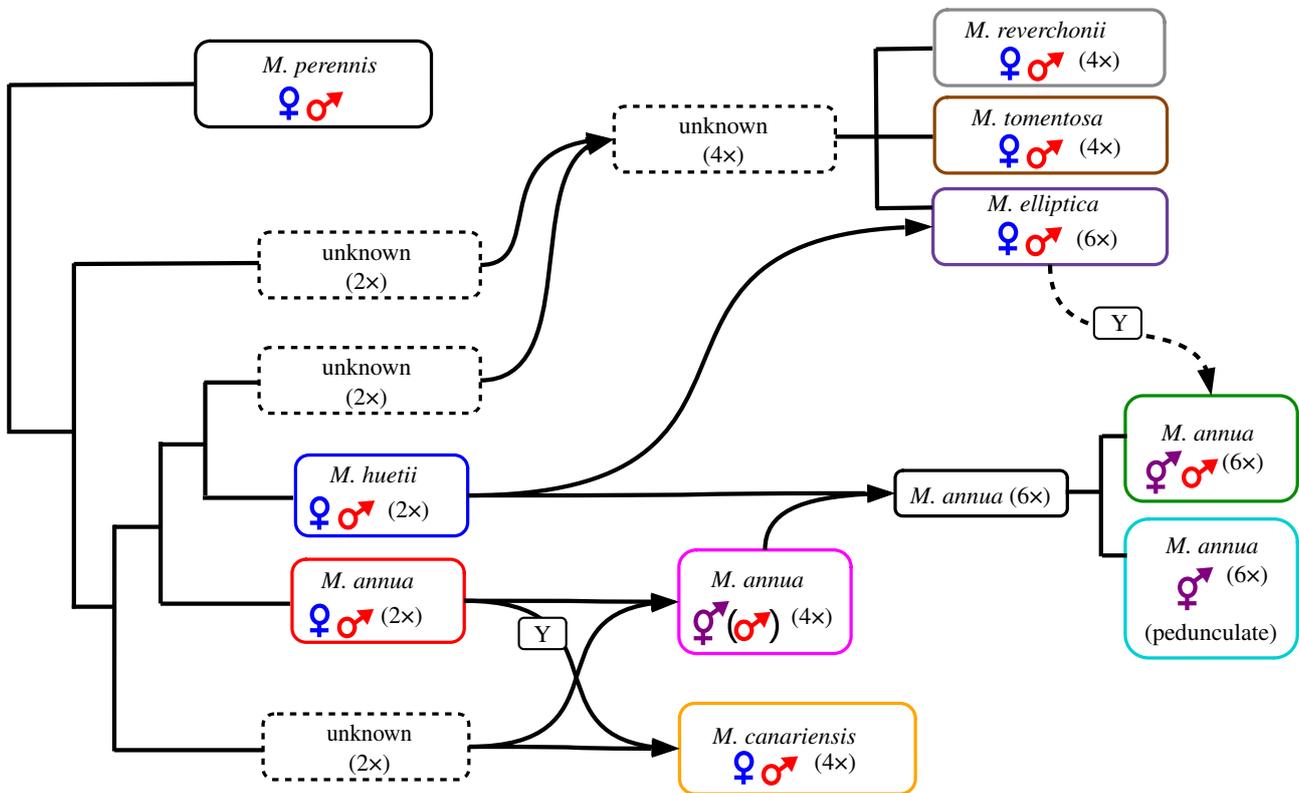


Figure 4. Model for the evolution of polyploid *Mercurialis* lineages. Arrows indicate polyploidization events, and dashed arrow indicates Y chromosome introgression. Lines indicate homoploid speciation events.

Europe point to a dynamic history of migrations, with *M. annua* having occupied much of Central and Western Europe from a refugium in the eastern Mediterranean, perhaps in present-day Turkey or the Middle East [63]. The range expansion into Western Europe was likely post-Pleistocene, so that *M. annua* probably occupied Northwest Africa much earlier, when *M. canariensis* likely originated. The same reasoning applies to the origin of tetraploid *M. annua*, which is found in central Western Morocco, far from the current distribution of its diploid progenitor *M. annua* [39,61]. Tetraploid *M. annua* is largely monoecious, but rare males have been observed and were sampled in our study here. Unfortunately, we were unable to amplify the male-linked locus in these males, and we thus remain ignorant of the origin of the Y chromosome in these populations.

Our most interesting finding is evidence for the origin of the Y chromosome in hexaploid androdioecious *M. annua* via introgression from one of the more distantly related perennial dioecious species, most probably an ancestor to *M. elliptica* (which is itself hexaploid). This interpretation is convincingly supported by the shared indels in the male-specific locus in both hexaploid *M. annua* and the basal perennial species. Importantly, other than the male-specific locus, we could find no other loci that were contributed by any of the perennial lineages to the *M. annua* hexaploid genome. It would thus seem that hybridization between hexaploid *M. annua* and an ancestor of *M. elliptica* was followed by backcrossing to the annual lineage, with retention of only the sex-determining region carrying the male-specific locus.

The introgression of the sex-determining region of the Y chromosome of hexaploid *M. annua* might have been owing to strong positive selection on the male-determining sequence. Hexaploid *M. annua* occurs as a metapopulation of both monoecious and androdioecious populations [64].

Monoecious individuals have a reproductive advantage over males at low density or during colonization, but males have a siring advantage at high density, so that the Y chromosome quickly rises to intermediate frequencies following immigration [64]. This metapopulation model [65] could also apply to hexaploid *M. annua* at a higher genealogical level. If hexaploid *M. annua* was originally monoecious (with combined sexes maintained by selection for a reproductive assurance because of the possibility of uniparental reproduction at low density), an introgressed male-determining Y chromosome might have been strongly favoured in high-density populations because of the high siring success that male phenotype confers [66,67].

It is also interesting that we could find no evidence for the introgression of other genomic regions into hexaploid *M. annua* from the contributor of its sex-linked region. Hexaploid *M. annua* is an annual species that occupies disturbed habitats such as ploughed fields, roadsides and other ruderal habitats. Its ecological preferences are thus different from those of the perennial species from which its Y chromosome is derived (they are small woody shrubs that occupy less disturbed, often wooded, habitats). If much of their genome had been influenced by selection for their perennial life history, little of it would have been able to persist in the face of negative selection in a frequently disturbed habitat. This possibility would be worth investigating further.

Data accessibility. The Python scripts developed for this study are available on a publicly available GitHub repository under https://github.com/jgerchen/mercurialis_phylogeny. Raw sequencing data are available on the National Center of Biotechnology Information databases (<https://www.ncbi.nlm.nih.gov/>) under Bioproject PRJNA794737. Additional data are provided in the electronic supplementary material.

Authors' contributions. J.F.G.: conceptualization, data curation, formal analysis, methodology, software, visualization, writing—original draft, writing—review and editing; P.V.: formal analysis, investigation, writing—original draft, writing—review and editing; J.R.P.: conceptualization, funding acquisition, project administration, supervision, validation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by funding from the Swiss National Science Foundation (grant nos. 31003A_163384 and 310030_185196) and the University of Lausanne.

Acknowledgements. We thank Ludovic Rüedi for help with testing the approach for generating phylogenetically informative loci in polyploids and Santiago C. González-Martínez for providing additional exon-capture data used for generating new primer pairs. The manuscript benefited from helpful comments by Susanne Renner and two anonymous reviewers.

References

- Charlesworth D. 2015 Plant contributions to our understanding of sex chromosome evolution. *New Phytol.* **208**, 52–65. (doi:10.1111/nph.13497)
- Charlesworth D. 2016 Plant sex chromosomes. *Annu. Rev. Plant Biol.* **67**, 397–420. (doi:10.1146/annurev-arplant-043015-111911)
- Beukeboom LW, Perrin N. 2014 *The evolution of sex determination*. Oxford, UK: Oxford University Press.
- Akagi T *et al.* 2019 Two Y-chromosome-encoded genes determine sex in kiwifruit. *Nat. Plants* **5**, 801–809. (doi:10.1038/s41477-019-0489-6)
- Harkess A *et al.* 2017 The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat. Comm.* **8**, 10. (doi:10.1038/s41467-017-01064-8)
- Harkess A, Huang K, van der Hulst R, Tissen B, Caplan JL, Koppula A, Batish M, Meyers BC, Leebens-Mack J. 2020 Sex determination by two Y-linked genes in garden asparagus. *Plant Cell* **32**, 1790–1796. (doi:10.1105/tpc.19.00859)
- Muller NA *et al.* 2020 A single gene underlies the dynamic evolution of poplar sex determination. *Nat. Plants* **6**, 630. (doi:10.1038/s41477-020-0672-9)
- Zhou R *et al.* 2020 A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Gen. Biol.* **21**, 19. (doi:10.1186/s13059-020-1952-4)
- Lloyd DG. 1975 Breeding systems in *Cotula* IV. Reversion from dioecy to monoecy. *New Phytol.* **74**, 125–145. (doi:10.1111/j.1469-8137.1975.tb01346.x)
- Case AL, Barrett SCH. 2004 Environmental stress and the evolution of dioecy: *Wurmbea dioica* (Colchicaceae) in western Australia. *Evol. Ecol.* **18**, 145–164. (doi:10.1023/B:EVEC.0000021152.34483.77)
- Case AL, Graham SW, Macfarlane TD, Barrett SCH. 2008 A phylogenetic study of transitions in sexual system in Australasian *Wurmbea* (Colchicaceae). *Int. J. Plant Sci.* **169**, 141–156. (doi:10.1086/523368)
- Dorken ME, Friedman J, Barrett SCH. 2002 The evolution and maintenance of monoecy and dioecy in *Sagittaria latifolia* (Alismataceae). *Evolution* **56**, 31–41. (doi:10.1111/j.0014-3820.2002.tb00847.x)
- Dorken ME, Barrett SCH. 2004 Sex determination and the evolution of dioecy from monoecy in *Sagittaria latifolia* (Alismataceae). *Proc. R. Soc. Lond. B* **271**, 213–219. (doi:10.1098/rspb.2003.2580)
- Mather K. 1949 Genetics of dioecy and monoecy in *Ecballium*. *Nature* **163**, 926. (doi:10.1038/163926a0)
- Costich DE, Meagher TR. 1992 Genetic variation in *Ecballium elaterium* (Curcubitaceae): breeding systems and geographic distribution. *J. Evol. Biol.* **5**, 589–601. (doi:10.1046/j.1420-9101.1992.5040589.x)
- Pannell JR, Dorken ME, Pujol B, Berjano R. 2008 Gender variation and transitions between sexual systems in *Mercurialis annua* (Euphorbiaceae). *Int. J. Plant Sci.* **169**, 129–139. (doi:10.1086/523360)
- Obbard DJ, Harris SA, Buggs RJA, Pannell JR. 2006 Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). *Evolution* **60**, 1801–1815. (doi:10.1111/j.0014-3820.2006.tb00524.x)
- Yang WL *et al.* 2021 A general model to explain repeated turnovers of sex determination in the Salicaceae. *Mol. Biol. Evol.* **38**, 968–980. (doi:10.1093/molbev/msaa261)
- Capel B. 2017 Vertebrate sex determination: evolutionary plasticity of a fundamental switch. *Nat. Rev. Genet.* **18**, 15. (doi:10.1038/nrg.2017.60)
- Tennessen JA, Wei N, Straub S, Govindarajulu R, Liston A, Ashman TL. 2018 Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLoS Biol.* **16**, 22. (doi:10.1371/journal.pbio.2006062)
- Beaudry FEG, Barrett SCH, Wright SI. 2020 Ancestral and neo-sex chromosomes contribute to population divergence in a dioecious plant. *Evolution* **74**, 256–269. (doi:10.1111/evo.13892)
- Rifkin JL, Beaudry FEG, Humphries Z, Choudhury BI, Barrett SCH, Wright SI. 2021 Widespread recombination suppression facilitates plant sex chromosome evolution. *Mol. Biol. Evol.* **38**, 1018–1030. (doi:10.1093/molbev/msaa271)
- Hough J, Hollister JD, Wang W, Barrett SC, H, Wright SI. 2014 Genetic degeneration of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc. Natl Acad. Sci. USA* **111**, 7713–7718. (doi:10.1073/pnas.1319227111)
- Heilbuth JC. 2000 Lower species richness in dioecious clades. *Am. Nat.* **156**, 221–241. (doi:10.1086/303389)
- Käfer J, de Boer HJ, Mousset S, Kool A, Dufay M, Marais GAB. 2014 Dioecy is associated with higher diversification rates in flowering plants. *J. Evol. Biol.* **27**, 1478–1490. (doi:10.1111/jeb.12385)
- Käfer J, Marais GAB, Pannell JR. 2017 On the rarity of dioecy in flowering plants. *Mol. Ecol.* **26**, 1225–1241. (doi:10.1111/mec.14020)
- Crossman A, Charlesworth D. 2014 Breakdown of dioecy: models where males acquire cosexual functions. *Evolution* **68**, 426–440. (doi:10.1111/evo.12283)
- Pannell JR. 2001 A hypothesis for the evolution of androdioecy: the joint influence of reproductive assurance and local mate competition in a metapopulation. *Evol. Ecol.* **14**, 195–211. (doi:10.1023/A:1011082827809)
- Ehlers BK, Bataillon T. 2007 'Inconstant males' and the maintenance of labile sex expression in subdioecious plants. *New Phytol.* **174**, 194–211. (doi:10.1111/j.1469-8137.2007.01975.x)
- Massonnet M *et al.* 2020 The genetic basis of sex determination in grapes. *Nat. Commun.* **11**, 12. (doi:10.1038/s41467-020-16700-z)
- Yu Q, Navajas-Pérez R, Tong E, Robertson J, Moore PH, Paterson AH, Ming R. 2008 Recent origin of dioecious and gynodioecious Y chromosomes in papaya. *Trop. Plant Biol.* **1**, 49–57. (doi:10.1007/s12042-007-9005-7)
- Cossard GG, Gerchen JF, Li XJ, Cuenot Y, Pannell JR. 2021 The rapid dissolution of dioecy by experimental evolution. *Curr. Biol.* **31**, 1277. (doi:10.1016/j.cub.2020.12.028)
- Ashman TL, Kwok A, Husband BC. 2013 Revisiting the dioecy-polyploidy association: alternate pathways and research opportunities. *Cytogenet. Genome Res.* **140**, 241–255. (doi:10.1159/000353306)
- Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009 The frequency of polyploid speciation in vascular plants. *Proc. Natl Acad. Sci. USA* **106**, 13 875–13 879. (doi:10.1073/pnas.0811575106)
- Otto SP, Whitton J. 2000 Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437. (doi:10.1146/annurev.genet.34.1.401)
- Wagner ND, He L, Horandl E. 2020 Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Front. Plant Sci.* **11**, 15. (doi:10.3389/fpls.2020.01077)
- Popp M, Oxelman B. 2007 Origin and evolution of North American polyploid *Silene* (Caryophyllaceae). *Am. J. Bot.* **94**, 330–349. (doi:10.3732/ajb.94.3.330)
- Singh RB, Smith BW. 1971 Mechanism of sex determination in *Rumex acetosella*. *Theor. Appl. Genet.* **41**, 360. (doi:10.1007/BF00277336)
- Durand B. 1963 Le complexe *Mercurialis annua* L. s.l.: une étude biosystématique. *Ann. Sci. Nat. Bot. Paris* **12**, 579–736.

40. Ma WJ, del Blanco LS, Pannell JR. 2019 A new biological species in the *Mercurialis annua* polyploid complex: functional divergence in inflorescence morphology and hybrid sterility. *Ann. Bot.* **124**, 165–178. (doi:10.1093/aob/mcz058)
41. Russell JRW, Pannell JR. 2015 Sex determination in dioecious *Mercurialis annua* and its close diploid and polyploid relatives. *Heredity* **114**, 262–271. (doi:10.1038/hdy.2014.95)
42. Veltsos P, Cossard G, Beaudoin E, Genséric B, Roux C, Gonzalez-Martinez SC, Pannell JR. 2018 The size and content of the sex-determining region of the Y chromosome in dioecious *Mercurialis annua*, a plant with homomorphic sex chromosomes. *Genes* **9**, 277. (doi:10.3390/genes9060277)
43. Veltsos P *et al.* 2019 Early sex-chromosome evolution in the diploid dioecious plant *Mercurialis annua*. *Genetics* **212**, 815–835. (doi:10.1534/genetics.119.302045)
44. Li XJ, Veltsos P, Cossard GG, Gerchen J, Pannell JR. 2019 YY males of the dioecious plant *Mercurialis annua* are fully viable but produce largely infertile pollen. *New Phytol.* **224**, 1394–1404. (doi:10.1111/nph.16016)
45. Pannell J. 1997 Widespread functional androdioecy in *Mercurialis annua* L. (Euphorbiaceae). *Biol. J. Linn. Soc.* **61**, 95–116. (doi:10.1111/j.1095-8312.1997.tb01779.x)
46. Pannell JR. 2002 What is functional androdioecy? *Func. Ecol.* **16**, 858–869. (doi:10.1046/j.1365-2435.2002.06893.x)
47. Charlesworth D. 1984 Androdioecy and the evolution of dioecy. *Biol. J. Linn. Soc.* **23**, 333–348. (doi:10.1111/j.1095-8312.1984.tb01683.x)
48. Weeks SC, Benvenuto C, Reed SK. 2006 When males and hermaphrodites coexist: a review of androdioecy in animals. *Integ. Comp. Biol.* **46**, 449–464. (doi:10.1093/icb/ijc048)
49. Wolf DE, Takebayashi N. 2004 Pollen limitation and the evolution of androdioecy from dioecy. *Am. Nat.* **163**, 122–137. (doi:10.1086/380493)
50. Rothfels CJ. 2021 Polyploid phylogenetics. *New Phytol.* **230**, 66–72. (doi:10.1111/nph.17105)
51. Rothfels CJ, Pryer KM, Li FW. 2017 Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytol.* **213**, 413–429. (doi:10.1111/nph.14111)
52. Gonzalez-Martinez SC, Ridout K, Pannell JR. 2017 Range expansion compromises adaptive evolution in an outcrossing plant. *Curr. Biol.* **27**, 2544–2549. (doi:10.1016/j.cub.2017.07.007)
53. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data P. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)
54. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012 Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115. (doi:10.1093/nar/gks596)
55. Katoh K, Misawa K, Kuma K, Miyata T. 2002 MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066. (doi:10.1093/nar/gkf436)
56. NCBI. 2018 Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **46**, D8–D13. (doi:10.1093/nar/gkx1095)
57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. (doi:10.1016/S0022-2836(05)80360-2)
58. Stamatakis A. 2014 RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
59. Cossard GG, Pannell JR. 2019 A functional decomposition of sex inconstancy in the dioecious, colonizing plant *Mercurialis annua*. *Am. J. Bot.* **106**, 722–732. (doi:10.1002/ajb2.1277)
60. Cossard GG, Pannell JR. 2021 Enhanced leaky sex expression in response to pollen limitation in the dioecious plant *Mercurialis annua*. *J. Evol. Biol.* **34**, 416–422. (doi:10.1111/jeb.13720)
61. Durand R, Durand B. 1992 Dioecy, monoecy, polyploidy and speciation in the annual mercuries. *Bull. Soc. Bot. France Lett. Bot.* **139**, 377–399. (doi:10.1080/01811797.1992.10824974)
62. Obbard DJ, Pannell JR, Harris SA. 2006 *Mercurialis canariensis* (Euphorbiaceae), a new endemic to the Canary Islands. *Kew Bull.* **61**, 99–106.
63. Obbard DJ, Harris SA, Pannell JR. 2006 Sexual systems and population genetic structure in an annual plant: testing the metapopulation model. *Am. Nat.* **167**, 354–366. (doi:10.1086/499546)
64. Dorken ME, Pannell JR. 2008 Density-dependent regulation of the sex ratio in an annual plant. *Am. Nat.* **171**, 824–830. (doi:10.1086/587524)
65. Pannell J. 1997 The maintenance of gynodioecy and androdioecy in a metapopulation. *Evolution* **51**, 10–20. (doi:10.1111/j.1558-5646.1997.tb02383.x)
66. Eppley SM, Pannell JR. 2007 Density-dependent self-fertilization and male versus hermaphrodite siring success in an androdioecious plant. *Evolution* **61**, 2349–2359. (doi:10.1111/j.1558-5646.2007.00195.x)
67. Santos del Blanco L, Tudor E, Pannell JR. 2018 Low siring success of females with an acquired male function illustrates the legacy of sexual dimorphism in constraining the breakdown of dioecy. *Ecol. Lett.* **22**, 486–497. (doi:10.1111/ele.13207)