Evolution of sex-biased gene expression during transitions to separate sexes in the *Silene* genus

Authors

Djivan Prentout[1,‡*], Aline Muyle[2,‡], Niklaus Zemp[3], Adil el Filali[4], Bastien Boussau[4], Pascal Touzet[5], Alex Widmer[6], Jos Käfer[7], Gabriel A.B. Marais[*4,8,9,10]

Affiliations

1 - Department of Biological Sciences, Columbia University, New York, USA

2 - Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), University of Montpellier, CNRS, 34 EPHE, IRD, Montpellier, France.

3 - Genetic Diversity Centre (GDC), ETH, Zürich, Switzerland

4 - Université de Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR 5558, 69622 Villeurbanne, France

5 - CNRS, UMR 8198 – Evo-Eco-Paleo, Univ. Lille, F-59000 Lille, France

6 - Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

7 - ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

8. CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO, Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal.

9. Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002, Porto, Portugal.

10. BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal.

* corresponding authors (djp2179@columbia.edu; gmarais@cibio.up.pt)

‡ These authors contributed equally to this work

Abstract

Sexual dimorphism is widespread among species with separate sexes and its extent is thought to be governed by the differential expression of thousands of genes between males and females (known as Sex-Biased Genes, hereafter SBGs). SBGs have been studied in numerous species, but rarely in a comparative way, which curtails our understanding of their evolution, especially during multiple independent transitions to separate sexes. We sequenced the transcriptomes of nine dioecious species, two gynodioecious species (separate females and hermaphrodites) and two hermaphrodite species from the *Silene* genus. Our dataset provides access to three independent transitions to dioecy (dating from less than 1 Myo to about 11 Myo). We demonstrated that male-biased expression emerges first during a transition to separate sexes, later followed by female-biased genes. Furthermore, we showed that, despite a mixture of selective regimes, positive selection significantly affects the evolution of some SBGs. Overall, this study provides new insights on the causes of SBG evolution during transitions to separate sexes.

Teaser

This study describes the evolution of sex-biased gene expression during a transition to separate sexes in plants.

MAIN TEXT

Introduction

Separate sexes (i.e. gonochorism in animals and dioecy in plants) is the sexual system of 95% of animals and 5% of flowering plants (*1–3*). The differences in the phenotypes of males and females (called sexual dimorphism) can affect the physiology, morphology, and other life history traits (*4–7*). The strength of sexual dimorphism varies widely between species and can be more important than phenotypic differences between individuals of the same sex (*8*). In several species with genetic sex determination, only one or two genes are sufficient to determine the sex of individuals (*9*, *10*). The sex determining genes then lead to the activation of a regulatory cascade where both transcription factors and hormones determine the differential expression of up to thousands of genes between males and females. The genes that are differentially expressed between males and females, the so-called Sex-Biased Genes (SBGs), are common in dioecious plants (from 2 to 17% of all expressed genes) and are distributed along the entire genome (*11–14*). Sex-Biased Gene Expression (SBGE) has been extensively described in several animals and to a lesser extent in some plant species (discussed in (*8*, *15*)). Previous studies have shown that the proportion of SBGs could vary significantly among tissues and developmental stages (*7*, *12*, *16*).

Despite numerous analyses of SBGE conducted to date, very few have been done in a comparative way. Therefore, the evolutionary forces at play remain an open question. For example, while a study in birds suggested that sexual selection (approximated by the intensity of sexual dimorphism) had driven the evolution of SBGE (17), converse results were found in cichlid fish (18), in the plant genus Leucadendron (19), and in brown algae (20), where genetic drift was likely to be the strongest evolutionary force driving SBGE. These studies question the common belief that the extent of sexual dimorphism is correlated to the number of SBGs (*8*).

In flowering plants, dioecy has evolved between 871 and 5,000 times independently (*1*), thus providing an exceptional opportunity for comparative analyses. Transitions from hermaphroditism to dioecy are thought to require an intermediate step, often through monoecy (female and male flowers on the same plant) or through gynodioecy (separate female and hermaphrodite individuals) (*2, 10, 17–19*). The latter assumes the invasion of the hermaphrodite population by a male-sterile (female) mutant, leading to gynodioecy (reviewed in (*2*)). Theoretical work suggests that hermaphrodites in gynodioecious populations gain most of their reproductive success through their male function (*17, 20, 21*), the loss of the female function in these individuals can be selected when it increases male fitness, which can lead to the evolution of dioecy. The steps to dioecy through the monoecious pathway have received less attention from modellers so the precise events and the associated selective pressures are less well formalised (*22*). To our knowledge, no comparative study has explored the evolution of SBGE in the monoecy nor the gynodioecy pathway in plants.

The *Silene* genus is a model for studying the evolution of plant sexual systems (*23*, *24*). At least three independent transitions to dioecy have been reported in *Silene* (*25–27*). It is likely that these transitions occurred through the gynodioecy pathway, as the genus contains many gynodioecious species. Dioecy evolved ~11 My ago in the *Melandrium* section, consisting of five dioecious species (*S. latifolia*, *S. dioica*, *S. heuffelii*, *S. marizii and S. diclinis*, (*25*, *28–30*)). Dioecy is probably younger in section *Otites* (~2.3 My; (*27*, *31*, *32*)) and likely of very recent origin in *S. acaulis* ssp *exscapa* (less than 1 My, (*26*)). XY sex chromosomes share a common origin in the *Melandrium* section (*33*). In the *Otites* section, *S. otites* has ZW sex chromosomes, while *S. pseudotites* and *S. colpophylla* have XY sex chromosomes (*27*, *31*). The ZW and XY systems evolved from different autosomes, although the exact evolutionary history (possibly involving introgression or turnover) is not known (*27*). So far, no sex chromosomes have been identified in *S. acaulis.* If a non-recombining region exists in *S. acaulis*, it is likely to be very small and carry only a few genes (*34*).

The repeated independent evolution of separate sexes with different ages of dioecy in *Silene* makes it an interesting model to study SBGE evolution in a comparative framework. So far SBGs have only been studied in *S. latifolia (12)*.

In this study, we characterise SBGE in the nine *Silene* dioecious species listed above (from the three independent transitions to dioecy) and two gynodioecious *Silene* species (*S. vulgaris* and *S. nutans)*. We use two hermaphrodite outgroup species (*S. paradoxa* and *S. viscosa*) to compare SBGE in dioecious and gynodioecious species to homologous hermaphrodite expression. With this dataset, we aimed to address the following questions: (1) Are there differences in the timing of the evolution of female- and male-biased genes? (2) Are gene expression changes occurring mostly in one sex, as previously suggested in *S. latifolia* (*12*, *35*)? (3) Do the same genes repeatedly become sex-biased in the independent transitions towards dioecy? (4) What evolutionary forces shape SBGE evolution, drift or selection?

Results

Transcriptome assemblies and mapping results

The *S. nutans* and *S. vulgaris* assemblies were composed of 23,836 genes and 31,526 genes, respectively (see Supplementary Table S1). A BUSCO (version 3.1.0; (*64*)) analysis showed that the *S. vulgaris* transcriptome assembly is more complete than that of *S. nutans* (Supplementary Table S1).

We mapped the RNA-seq data of the 9 dioecious species on both transcriptome assemblies and found that the mapping rate on *S. vulgaris* assembly was more similar among all the species (50%) than the one on *S. nutans* assembly (Supplementary Table S2). Coupled with the BUSCO results, we decided to keep the mapping on *S. vulgaris* for the rest of the analyses.

Numbers and proportions of SBGs in dioecious and gynodioecious species

We used three different tools to identify sex-biased genes: (1) DESeq2, (2) edgeR, (3) Limma-Voom. These three methods differ slightly in the way they assess sex-biased gene expression due to differences in the family distribution used for read counts. The two firsts rely on a negative binomial distribution and the latter on log-normal distribution. To limit biases, the genes with a log(foldchange)>2 and a p-value<10^{-4} in at least 2 of the three methods were considered as sex-biased. The number of sex-biased genes and the number of sex-biased genes identified as

autosomal or sex-linked by SEX-DETector are presented in Supplementary Table S3. Figure 1 represents the proportions of sex-biased genes among expressed genes for the eleven species.
For the rest of the analysis, we will report the genes over-expressed in hermaphrodites of the gynodioecious species as male-biased, as hermaphrodites of gynodioecious species reproduce mainly through the male function (*21*). Also, we checked for methodological biases by repeating the analyses with different *p*-values and fold-change thresholds for sex-biased genes inferences, and to test for an effect of sample size we repeated all analyses with 4 males and 4 females only for all species (Supplementary Figure S1 to S5 and Supplementary Tables S4 and S5). We found qualitatively similar results on these controls.

As shown in Figure 1**,** species in section Melandrium have the highest proportion of female- and male-biased genes. This difference is stronger for female-biased genes, with species in section *Melandrium* having about 4 times more female-biased genes than those in section *Otites*, and ten times more than gynodioecious species. The proportion of male-biased genes is much lower in gynodioecious species than in sections *Melandrium* and *Otites*, with *S. acaulis* being somewhat intermediate. Autosomal and sex-linked genes were identified for the eight dioecious species with a known pair of sex chromosomes. The pattern of SBG proportions is similar between autosomal genes and all genes, as expected as most SBGs are autosomal, with male-biased genes being more numerous than female-biased genes (Figure 1). Strikingly, however, female-biased genes are overrepresented in sex-linked genes compared to male-biased genes in section *Melandrium*, but not in section *Otites*.

SBGs accumulate over time after the evolution of separate sexes

We used a generalised linear model to investigate the relationship between the number of SBGs and the age of dioecy (using a negative binomial family, and by accounting for the different number of expressed genes among species through adding an offset, see Materials and Methods for details). A significant positive correlation was found between the age of dioecy and the number of female-biased, male-biased and the total number of SBGs ($p < 10^{-7}$, $R^2 > 0.9$, Supplementary Table S6, Fig. 2). The number of male-biased genes seemed to reach a plateau as the age of dioecy increased (Figure 2). We therefore used a polynomial regression for the relationship between male-biased genes and the age of dioecy (Materials and Methods equation 1, Supplementary Table S6), which explained more variance compared to a linear model. The same was true for the relationship between the number of total SBGs and the age of dioecy (Materials and Methods equation 1, Figure 2, Supplementary Table S6). On the other hand, the number of

female-biased genes significantly increased with the age of dioecy without reaching a plateau (Materials and Methods equation 2, Figure 2, Supplementary Table S6). The positive correlations between the number of SBGs and the age of dioecy remained significant when accounting for species phylogeny using generalised least squares (equation 4, Supplementary Table S6).

The positive correlations observed between the number of SBGs and the age of dioecy suggest it takes time for SBGE to evolve, as species with older dioecy have more SBGE. Male-biased genes evolve early, as they are already present in gynodioecious species, but, after a few My, their numbers reach a plateau in dioecious species. Female-biased genes evolve later, after the transition to dioecy. We did not detect a plateau for female-biased gene numbers over time, but older dioecious species should be studied to address this question.

High turnover of SBGs in *Silene*

The proportion of species-specific SBGs is larger than that of SBGs shared among several species (Figure 3). Indeed, there is no species for which the number of specific SBGs is smaller than the number of SBGs shared with at least another species. This result shows that there is a high turnover in SBGE evolution in *Silene*. This is consistent with studies in other organisms (*65*, *66*).

Selective pressures on SBGE evolution

In order to test whether SBGE evolved under positive selection, we calculated the Δ_X statistic for each species and each gene. This statistic summarises the change in expression of a gene in a focal species with respect to an outgroup species, normalised by the standard deviation in the focal species (see Materials and Methods). It is based on the double expectation that positive selection leads to larger than average changes in expression levels between species, as well as more similar expression levels between individuals within a species (*58*, *67*, *68*). We considered genes with Δ_X values higher than the 75 quantile as evolving under positive selection (as in (*12*); see below for a more stringent threshold). In order to test for an enrichment of selection in SBGE, the proportion of SBGs and unbiased genes evolving under selection were compared using a Chi-square test for each species, sex and type of sex-bias (Figure 4, Supplementary Tables S7 and S8).

As shown in Figure 4, female-biased expression is due both to increased female expression and decreased male expression compared to the hermaphrodite outgroup. Similarly, male-biased gene expression is due both to decreased expression in females and increased expression in males compared to the hermaphrodite outgroup. In order to assess whether these expression changes in

SBGs were driven by selection, we tested for an enrichment in selection in SBGs compared to unbiased genes (Supplementary Tables S7 and S8). First, in both sections *Melandrium* and *Otites*, female-biased gene expression that evolved through decreased expression in males compared to the hermaphrodite outgroup was enriched in positive selection compared to unbiased genes. Second, for the *Melandrium* section only, increased expression in females for the female-biased genes is also enriched in positive selection compared to unbiased genes. For male-biased gene expression, significant enrichment in positive selection was also found among genes for which expression decreased in females compared to the hermaphrodite outgroup. However, increased male expression compared to hermaphrodite outgroups was significantly depleted in positive selection for male-biased genes in eight out of eleven species, suggesting that male expression evolves mostly under drift or purifying selection in male-biased genes. To differentiate between drift and purifying selection, we considered expression variation within species (standard deviation). Male expression was the most variable in male-biased genes, especially when male expression increased compared to the outgroup (Supplementary Figure S6), suggesting drift was the driving force for the evolution of expression in this category.

Results were qualitatively similar when considering contigs with $\Delta_X > 10$ as under positive selection (as done in Scharmann *et al.* 2021 (*69*)), instead of using the quantile 75 (Supplementary Table S9). Using a threshold of $\Delta_X > 10$ is more stringent to infer positive selection because the species quantile 75 of the Δ_X ranged from 2.5 to 4.6. Results were also unaffected when we repeated the analyses on autosomal genes only (for species without sex chromosomes all genes were kept, for species with sex chromosomes, only genes inferred as autosomal by SEX-DETector were kept, Supplementary Table S10). As another control, we ran the same analysis on leaf tissues (instead of flower buds) in *S. latifolia* only, using *S. viscosa* leaf data as a hermaphrodite outgroup to compute the Δ_X. We detected a significant enrichment in selection for leaf male-biased genes in *S. latifolia* when male expression increased compared to the outgroup, as well as when female expression decreased compared to the outgroup (Supplementary Table S11).

Selected feminization of the X chromosome

We investigated whether the increase in the proportion of female-biased genes among sex-linked genes in the *Melandrium* section can be attributed to the feminization of the X chromosome. Feminization of the X is an enrichment of female-beneficial genes on the X, expected to occur because the X chromosome spends two thirds of its time in females and only one third of its time

in males (*70*, *71*). To test for X feminization, we compared the proportion of female-biased genes on the X that have an expression evolving under positive selection (1) to the proportion of male-biased genes on the X with an expression evolving under positive selection, and (2) to the proportion of female-biased genes on the autosomes with an expression evolving under positive selection. We found that the proportion of female-biased genes evolving under positive selection was significantly higher on the X compared to the other two categories in the *Melandrium* section (Chi-square test; see Supplementary Tables S12 and S13). This shows that the X is significantly enriched in female-biased genes under positive selection in the Melandrium section, suggesting active X feminization. Interestingly, we observed more positive selection signatures for a reduction of expression in males than for an increase of expression in females. These results tend to confirm a feminization of the X chromosomes in the *Melandrium* section. The numbers of female-biased genes in the *Otites* section were too small to conduct such an analysis.

Evolution of SBGs under drift

We hypothesised that if drift is a strong driver of sex-biased gene expression, species with lower effective population size may have more SBG. On the contrary, if selection is mostly driving the evolution of SBGs, species with higher effective population size (N_e) should have more SBGs. In order to test for an effect of the intensity of selection on the number of sex-biased genes, we ran a generalised linear model between the number of sex-biased genes and the synonymous nucleotide diversity π_S (a proxy for N_e, equations 3 and 5). No significant correlation was found in either direction (Supplementary Figure S7, Supplementary Table S14), suggesting that drift and selection do not explain differences in SBG numbers among species, or cancel each other.

As done by Scharmann *et al.* (2021), we tested whether sex-biased genes had an increased expression evolutionary rate compared to unbiased genes. To avoid rates being affected by SBGE evolution, we computed rates of SBGE evolution after removing the species for which the gene is sex-biased, so that the rates correspond to expression changes before the gene became sex-biased. We therefore compared expression evolution rates between genes that are always unbiased in all species and genes that are sex-biased in at least one species, after removing species with sex-biased expression from the analysis. We found, as Scharmann *et al.* (2021), that genes that are sex-biased in at least one species have a higher rate of expression change, measured as the mean of absolute PICs (91.91 on average) compared to genes that are always unbiased (51.93 on average, one-sided permutation test p-value $< 2.2 \times 10^{-16}$, Supplementary Figure S8). Scharmann *et al.* (2021) interpreted this result as an indication that genes that become sex-biased have faster-

evolving expression levels, even before becoming sex-biased (*69*). To further test this interpretation, we splitted the analysis of PICs for genes that were detected as evolving under positive selection in at least one species and genes that never evolve under positive selection (using the previous Δ_X analysis). We found that SBGs that evolve under positive selection in at least one species have the highest rate of expression change, even before becoming sex-biased (94.96 on average, permutation tests p-values < 10^{-3}, Supplementary Figure S8). The rate of expression change was lowest for unbiased genes that never evolve under positive selection (mean 10.94, permutation tests p-values < 10^{-13}, Supplementary Figure S9). Therefore both positive selection and drift seem to accelerate expression evolutionary rates.

Functional analysis of SBGs

We explored if the SBGs are enriched for functions or pathways linked to sexual reproduction. We tested the enrichment for four sets of genes (1) the whole set of male- or female-biased genes (2) the male- or female-biased genes in gynodioecious species (3) the male- or female-biased genes in dioecious species (4) the male- or female-biased genes under positive selection (using the previous Δ_X analysis, see Supplementary Table S15 for more information). We then looked for enrichment in functions explicitly linked to sexual selection in each of these sets of genes. We also produced wordcloud figures to highlight the main functions or pathways (Supplementary Figure S10 & S11). Only two GO terms are explicitly linked to a reproductive function (*GO:0090567: reproductive shoot system development* and *GO:0048437: floral organ development*). Those two GO terms are found while testing the whole set of female-biased genes against the whole annotation.

Discussion

We studied Sex-Biased Gene Expression (SBGE) in the flower buds of eleven sexually dimorphic *Silene* species, including nine dioecious species and two gynodioecious species. The nine dioecious species originated from three independent transitions to dioecy. The youngest transition occurred less than one million years ago and the oldest approximately eleven million years ago (*26*, *28*, *29*).

Overall, the 11 species displayed more male-biased genes than female-biased genes (Figure 1), which has already been observed in various plant species before (*15*). Two possible explanations for this observation are (i) that drift is stronger in males due to a more variable reproductive success and a smaller effective population size, or (ii) stronger sexual selection in males due to strong competition among males. The age of the dioecy is positively correlated with both the number of male-biased genes and the number of female-biased genes. The dynamics, however, seem to differ between sexes: male-biased gene numbers increase early and seem to reach a plateau, whereas female-biased genes evolve more gradually and continuously. A possible explanation for the plateau reached by the number of male-biased genes is that an equilibrium is attained among mutations, selection and drift. Mutations create new SBGs in dioecious species and selection and drift filter them through time.

In gynodioecious species, we consider genes with higher expression in the hermaphrodite individuals as male-biased, since the hermaphrodites' main reproductive output is through the male function (*20*). A possible explanation as to why male-biased genes are already present in gynodioecious species but female-biased genes are nearly absent is that a gene could become male-biased simply through the reduction of expression in female flowers, as a result of the loss of the male function. Therefore, although female flowers evolve first in the gynodioecy pathway, female-biased genes mainly evolve later, when female functions are suppressed in male flowers (i.e., when full dioecy evolves).

The patterns of male- and female-biased genes enrichment are reversed in the non-recombining region of the sex chromosomes in the Melandrium section (species from the oldest transition to dioecy, Figure 1). We tested whether the female-biased genes on the X chromosome bear footprints of positive selection, using the Δ_X statistic (see (*58*)). In brief, if selection drove the evolution of a SBG, we should observe a strong change in the expression level and a small variance between individuals of the same sex in a species. The Δ_X analysis showed that female-biased genes in the non-recombining region of the X chromosome are significantly enriched in positive selection compared to male-biased genes on the non-recombining region of the X chromosome or female-biased genes on autosomes. This result supports an active feminization of the X chromosome in the Melandrium section.

We used the Δ_X statistic to study the selective regime of all the SGBs in the different species. With the exception of female-biased genes in *S. acaulis* (the species from the most recent transition to dioecy and very few female-biased genes), the reduction of expression in males for female-biased genes, or reduction of expression in females for male-biased genes is enriched in genes with an expression evolving under positive selection for all the dioecious species. In contrast, the increase of expression is enriched in genes evolving under positive selection only in

females of the Melandrium section for the female-biased genes. This suggests that, even if SBGs result from changes in both sexes, increases in gene expression do generally not occur through positive selection, while the reduction in expression generally does. For gynodioecious species and *S. acaulis* (which evolved dioecy recently), the decrease in female expression of male-biased genes was significantly enriched in selection (Figure 4). These footprints of positive selection differ from a previous analysis in the *Leucadendron* genus (*69*), where sex-biased genes were not enriched in positive selection on expression levels compared to unbiased genes, suggesting that they mostly evolved under relaxed selection. In order to test whether this difference between the two studies was due to the sampled tissue (flower buds in *Silene* versus leaves in *Leucadendron*), we used leaf tissue available in *S. latifolia* and *S. viscosa* to compute the Δ_X in *S. latifolia* leaves (Supplementary Table S11). Leaf male-biased genes are significantly enriched in adaptive evolution compared to unbiased genes in *S. latifolia*. Therefore, the differences between *Silene* and *Leucadendron* are not attributable to the sampled tissue.

Some SBG categories are significantly enriched in positive selection (Figure 4), but other categories seem mostly driven by drift as they exhibit a high within-species variation in expression levels (Supplementary Figure S6). We therefore clearly observe the effects of both selection and drift on SBGE evolution in *Silene*. This is also visible when studying expression evolutionary rates (Supplementary Figure S9). In our *Silene* dataset, sex-biased genes that are never under selection in any species (according to the Δ_X analysis) have faster expression evolutionary rates than unbiased genes that are never under selection, suggesting that SBGE leads to faster expression evolutionary rates because of drift. Expression evolutionary rates are even further accelerated by selection, because SBGs that evolve under selection (according to the Δ_X analysis) in at least one species have higher expression evolutionary rates than SBGs that never evolve under selection (Supplementary Figure S9). We therefore confirm the theoretical prediction by Dapper and Wade (*72*) that SBGs mainly evolve under relaxed selection, with some exceptions. For some SBGs, positive selection is especially strong and contributes to accelerated rates of expression evolution in *Silene*.

To our knowledge, this analysis is the first comparative analysis of SBGE between gynodioecious species and dioecious species from independent transitions to dioecy. Indeed, despite numerous analyses of SBGE conducted to date, very few have been done in a comparative way (but see (*65*, *66*, *69*, *73*)). This limits our understanding of its evolution and the evolutionary forces that shape it, especially in a transition from hermaphroditism to dioecy or gonochorism. Here, we show that the proportion of sex-biased genes correlates with the age of dioecy and that, through the gynodioecious pathway, male-biased genes emerge first. Also, our results support a combined action of positive selection and genetic drift in the evolution of SBGE. This nuances

the theoretical prediction that most sex-biased genes should evolve under relaxed selection, simply because sex-specific expression reduces selection coefficient of a gene (*72*). Our study therefore calls for more investigation in other groups to enlighten why in some species SBGE seems to evolve mostly under drift (like *Leucadendron*), while other species show positive selection driving SBGE (such as *Silene*).

Materials and Methods

Data

- Crossing and sequencing for the *Melandrium* section

For *S. latifolia* and *S. vulgaris*, data have been reused from Zemp *et al*, (2016) and Zemp *et al.* (2018)(*12*),37)). For *S. dioica, S. heuffelii, S. marizii and S. diclinis*, we have crossed a female and a male from two different populations. Seeds from the crosses were sown to produce F1 individuals. Three flower buds (~2-3 mm) without the calyx were sampled for multiple females and males and their parents (see Supplementary Table S16 for details on sample sizes). High quality RNA (RIN > 9) has been extracted using the Plant Total RNA Mini Kit from Geneaid. RNAseq libraries were prepared using Ultra II RNA Library Prep Kit for Illumina from NEBNext. The libraries were then sequenced on a Hiseq4000 instrument using the paired end 150bp protocol at the Functional Genomic Center, Zurich.

- Sampling and sequencing of *S. acaulis*

Flower buds from males and females of *Silene acaulis* have been sampled from a natural population, in 2018, at “Rocher Blanc” (“massif d’Allevard”, France, between 2800m and 2900m above sea level).
The RNA preparation was done by the AGAP institute (Montpellier, France). For all individuals, fresh tissues were sent, then flash-frozen (Freshfreeze method). RNA was extracted following the SIGMA protocol and sequenced using the HiSeq3000-HWI-J00115 technology, producing 150bp paired reads.

- Published dataset for Otites section

For *S. otites* and *S. pseudotites*, we used publicly available data from Martin *et al.* (2019) (*31*). For *S. colpophylla*, we used data from Balounova *et al.* (2019)(*27*).

- Published dataset for hermaphrodite species, gynodioecious species, and *Dianthus chinensis*

For *S. nutans*, *S. paradoxa* and *S. viscosa* (the two hermaphrodite outgroups), as well as *Dianthus chinensis* we used publicly available RNAseq data from Muyle *et al.* (2021)(*36*) .

Transcriptome assemblies

We assembled a *de novo* transcriptome for both gynodioecious species: *S. vulgaris* and *S. nutans*. To this end, we used the tool DRAP (version 1.92,(*37*)), which allows merging and compacting several transcriptome assemblies into a single "meta-assembly".
For *S. vulgaris,* we first independently assembled two transcriptomes from two hermaphrodite individuals and one transcriptome from a female (tool RunDrap from Drap – default parameters). Then, we merged these three transcriptomes into a single one with RunMeta (from DRAP, default parameters).
For *S. nutans,* we independently assembled two hermaphrodite transcriptomes and two female transcriptomes that we merged in a single one (using the default parameters for RunDrap and RunMeta, as we did for *S. vulgaris*).
DRAP provides several assemblies. We used the one filtered with TransDecoder (https://github.com/TransDecoder/TransDecoder) in order to obtain the Open Reading Frame (ORF) of every gene.

Trimming, filtering and mapping of reads

We filtered PCR duplicates with Condetri filterPCRdupl(*38*). The reads with a Phred quality score lower than 64 were filtered with trimmomatic (option PE -phred64)(version 0.39; Bolger, Lohse, and Usadel 2014) and the adapters were removed using prinseq-lite (options -trim_tail_left 5 -trim_tail_left 5)(version 0.20.4; (*39*)). Once filtered, the reads were mapped on both transcriptome assemblies using GSNAP (option m 0.1)(version 2019-09-12 ;(*40*)). In order to increase the amount of reads mapped for distant species, we re-ran GSNAP with the SNP-tolerant option and SNPs identified in the first mapping (method previously employed in Prentout *et al.* (2020)(*14*)). This iteration has been done once for each species, except when mapping *S. vulgaris* and *S. nutans* on their respective transcriptome. Samtools version 1.9 (*41*) was used to remove unmapped reads and sort the mapping outputs to bam files (view -F 4 | sort).

Transcriptome annotation

We annotated the transcriptome of *S. vulgaris* as it is the one that we used for the rest of the analysis (see results section). We used diamond to blast the nucleotide CDS, provided by DRAP, against the nr database (nr version of April 2022)(diamond v2.0.4.142, blastx -p 4 --max-hsps 3 -e 1e-5 -f 5)(*42*). The diamond output (xml format) was provided to Blast2GO for the gene ontology mapping (Blast2GO Basic v6.0)(*43*, *44*).

To statistically test if a set of genes is enriched for specific functional pathways, we used two tools: GOstats (version 2.62.0, (*45*)) and clusterprofiler (version 3.15.0, (*46*)), and kept the GO term with an adjusted *p*-value < 0.05.

Sex-linked genes identification

Sex-linked genes are genes located in the non-recombining region of the Y (or W) chromosome or the homologous region on the X (or the Z) chromosome. In order to identify sex-linked genes in dioecious species, we first inferred the genotype of each individual with Reads2snp version 2.0 (*47*). We accounted for allelic expression bias, without filtering for paralogous SNPs and retained positions with a minimum coverage of 3 (-aeb -par 0 -min 3).

SEX-DETector (*48*), a tool based on the analyses of allele segregation within a cross, was used to identify sex-linked transcripts in every dioecious species of the Otites and Melandrium sections. This tool is based on an SEM algorithm and computes, for each gene, a posterior probability of being autosomal (P_A), XY (P_{XY}) or X-hemizygous ($P_{X\text{-}hemi}$). We considered a gene as sex-linked (or autosomal) if $P_{XY} + P_{X\text{-}hemi} > 0.8$ (respectively $P_A > 0.8$), and if at least one SNP was classified as sex-linked (or autosomal) without genotyping error.

Three models have been implemented in SEX-DETector: (1) XY sex chromosome system, (2) ZW sex chromosome system, (3) a system without sex chromosomes.

In all species, the type of sex chromosome has already been identify, so we run SEX-DETector with the corresponding model (i.e. ZW for *S. otites* and XY for the seven other species; (*27*, *28*, *31*)).

As mentioned in the introduction, no sex chromosomes have been identified in *S. acaulis* so far (*34*). Because their detection is beyond the scope of this analysis, we didn't detect sex-linked genes in this species.

Gene expression level and sex-biased gene identification

The expression level of each gene was computed with samtools idxstats (Version 1.9; (*41*)). Then, three tools were used to classify a gene as sex-biased: EdgeR (Version 3.36.0; (*49*)), DESeq2 (Version 1.34.0; (*50*)) and Limma-Voom (Version 3.50.3; (*51*)). We classified a gene as sex-biased if at least two of the three methods classified it as sex-biased with *p*-value lower than 10^{-4} and greater than 2 ($\log_2 FC > 1$).

To infer in which sex the expression changed compared to the ancestral expression level before separate sexes evolution, we computed the mean fragment per kilobase of transcript per million mapped reads (FPKM) for each gene in every individual, and then, the mean FPKM for each gene in each sex (males, females or hermaphrodites). The difference in mean expression between the focal dioecious (or gynodioecious) species and the closest fully hermaphrodite species (either *S. paradoxa* or *S. viscosa*) was used to infer an increase or a decrease in gene expression compared to ancestral expression before separate sexes evolution. For *S. diclinis*, *S. dioica*, *S. heuffelii*, *S. latifolia*, *S. marizii* and *S. vulgaris*, we used the hermaphrodite species *S. viscosa* as an outgroup. For *S. acaulis*, *S. colpophylla*, *S. nutans*, *S. otites* and *S. pseudotites*, we used the hermaphrodite species *S. paradoxa* as an outgroup.

Phylogenetic reconstruction

Dianthus chinensis was used as the outgroup species for the phylogenetic reconstruction (we used 2 hermaphrodite individuals). In order to build the species tree, we kept transcripts of autosomal genes for which the sequence was available in all individuals of all species ($N_{individuals} = 143$) with a maximum of 30% of Ns in the sequence (740 genes in total). In each individual, only one of the two alleles of each gene was kept for phylogenetic reconstruction. We concatenated the sequences of the 740 genes in each of the 143 individuals, which led to a sequence of 873,747 bp per individual. The phylogeny was inferred with IQ-TREE2 (version 2.1.3; (*52*)) using the GTR+G4 model and 100 replicates for the bootstraps.

Linear regressions between the number of sex-biased genes and the age of dioecy

The number of sex biased genes was correlated to the age of dioecy using a generalised linear model with negative binomial 1 family (appropriate for large count data) implemented in the glmmTMB R package (*53*). In order to account for the different number of sampled genes among species, the total number of expressed genes was log transformed and used as an offset. For the total number of sex-biased genes and the number of male-biased genes, the model that explained

the most variance was a raw polynomial of degree 2 (equation 1). For female-biased genes, a simpler model without polynomial effects was used (equation 2). Fit of the model to the data was checked using the DHARMa R package by looking at residual plots (*54*).

Number of sex-biased genes ~ poly(Age dioecy, 2, raw=T) + offset(log(Total_expressed+1)) (equation 1)

Number of sex-biased genes ~ Age dioecy + offset(log(Total_expressed+1)) (equation 2)

The regression output was plotted using the R package ggeffects to compute predicted values of the model with function ggpredict (*55*). The R^2 was computed using the function r.squaredGLMM of the R package MuMIn (*56*).

A similar model was run to test for an effect of the effective population size N_e (the synonymous nucleotide diversity π_S was used as a proxy for N_e). The negative binomial distribution was used and an offset was included:

Number of sex-biased genes ~ scale(π_S) + offset(log(Total_expressed+1)) (equation 3)

We also ran models that corrected for phylogenetic relationships among species using generalized least squares implemented in the nlme R package (*57*), using function gls and Martin's correlation structure. Since an offset cannot be implemented in such models, we did not include it. As the negative binomial family is not available is gls, we log transformed the number of sex-biased genes:

log(Number of sex-biased genes + 1) ~ scale(Age dioecy) + corMartins(phylogeny) (equation 4)

log(Number of sex-biased genes + 1) ~ scale(π_S) + corMartins(phylogeny) (equation 5)

<u>Δ_X analysis</u>

The Δ_X has become a widely used statistic in transcriptomics to evaluate selection pressures on gene expression (*58*). In order to compute the Δ_X for each sex, gene and species, the expression level of each gene and individual was first determined. The number of mapped reads was computed for each gene and individual using samtools version 1.10 idxstats and was normalised

to reads per kilobase per million mapped reads (RPKM) for each contig and individual separately as follows:

gene RPKM = raw number of mapped reads on gene / (contig length x individual mapped library size) x 10^9 (equation 6)

Then, the Δ_X was computed for each species and sex separately (i.e. for males and females separately) as follows:

Δ_X = abs(mean(focal species expression) – mean(outgroup)) / standard deviation(focal species expression) (equation 7)

Where abs stands for absolute value. The hermaphrodite species used as the outgroup are the same as in the "Gene expression level and sex-biased gene identification" section.

For each species, contigs with Δ_X values higher than the 75 quantile of Δ_X values were considered as under positive selection for expression evolution (as in (*12*)). We compared the proportion of SBGs and unbiased genes under selection using a Chi-square test and corrected for multiple tests using Benjamini & Hochberg correction (*59*).

Genes with sex-specific expression were excluded from Δ_X analyses, to avoid the confounding effect of including genes that are encoding tissues or functions sex-specific (*i.e.* androecium, gynoecium, etc), without the need to involve selection.

We also run this analysis on leaf tissues from *S. latifolia* and *S. viscosa*, to test if different tissues show different results.

Phylogenetically independent contrasts (PICs)

Phylogenetic independent contrasts (PICs) correspond to the amount of change (here change in expression level) between two taxa divided by the branch length separating them. The mean of the absolute standardised PICs (*60*) per gene was employed as a measure of the rate of expression evolution (*61*). For each gene, PICs were calculated based on the species tree and the mean RPKM expression values per species, using the pic function in APE (*62*). Species exhibiting sex-biased expressions were excluded when calculating PICs, so that the PICs only measure gene expression variation without sex bias (i.e. before the gene became sex-biased). Genes that are sex-biased in at least one species and genes that are always unbiased were then compared for their

mean absolute PICs on the basis of 10,000 permutations, using the function permTS in the R package perm v1.0 (*63*).

References

1. S. S. Renner, The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* **101**, 1588–1596 (2014).

2. C. Fruchard, G. A. B. Marais, "The Evolution of Sex Determination in Plants" in *Evolutionary Developmental Biology: A Reference Guide*, L. Nuño de la Rosa, G. B. Müller, Eds. (Springer International Publishing, Cham, 2021), pp. 683–696.

3. P. Jarne, J. R. Auld, ANIMALS MIX IT UP TOO: THE DISTRIBUTION OF SELF-FERTILIZATION AMONG HERMAPHRODITIC ANIMALS. *Evolution* . **60**, 1816–1824 (2006).

4. T. E. Dawson, M. A. Geber, "Sexual Dimorphism in Physiology and Morphology" in *Gender and Sexual Dimorphism in Flowering Plants*, M. A. Geber, T. E. Dawson, L. F. Delph, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999), pp. 175–215.

5. S. C. H. Barrett, J. Hough, Sexual dimorphism in flowering plants. *J. Exp. Bot.* **64**, 67–82 (2013).

6. E. Mori, G. Mazza, S. Lovari, "Sexual Dimorphism" in *Encyclopedia of Animal Cognition and Behavior*, J. Vonk, T. Shackelford, Eds. (Springer International Publishing, Cham, 2017), pp. 1–7.

7. S. Naqvi, A. K. Godfrey, J. F. Hughes, M. L. Goodheart, R. N. Mitchell, D. C. Page, Conservation, acquisition, and functional impact of sex-biased gene expression in mammals. *Science*. **365** (2019), doi:10.1126/science.aaw7317.

8. S. Grath, J. Parsch, Sex-Biased Gene Expression. *Annu. Rev. Genet.* **50**, 29–44 (2016).

9. D. Bachtrog, J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, R. Ming, N. Perrin, L. Ross, N. Valenzuela, J. C. Vamosi, Tree of Sex Consortium, Sex determination: why so many ways of doing it? *PLoS Biol.* **12**, e1001899 (2014).

10. S. S. Renner, N. A. Müller, Plant sex chromosomes defy evolutionary models of expanding recombination suppression and genetic degeneration. *Nat Plants*. **7**, 392–402 (2021).

11. A. Harkess, F. Mercati, H.-Y. Shan, F. Sunseri, A. Falavigna, J. Leebens-Mack, Sex-biased gene expression in dioecious garden asparagus (Asparagus officinalis). *New Phytol.* **207**, 883–892 (2015).

12. N. Zemp, R. Tavares, A. Muyle, D. Charlesworth, G. A. B. Marais, A. Widmer, Evolution of sex-biased gene expression in a dioecious plant. *Nat Plants*. **2**, 16168 (2016).

13. P. Veltsos, Y. Fang, A. R. Cossins, R. R. Snook, M. G. Ritchie, Mating system manipulation and the evolution of sex-biased gene expression in Drosophila. *Nat. Commun.* **8**, 2072 (2017).

14. D. Prentout, O. Razumova, B. Rhoné, H. Badouin, H. Henri, C. Feng, J. Käfer, G. Karlov, G. A. B. Marais, An efficient RNA-seq-based segregation analysis identifies the sex chromosomes of Cannabis sativa. *Genome Res.* **30**, 164–172 (2020).

15. A. Muyle, How different is the evolution of sex-biased gene expression between plants and animals? A commentary on: “Sexual dimorphism and rapid turnover in gene expression in pre-reproductive seedlings of a dioecious herb.” *Ann. Bot.* **123** (2019), pp. iv–v.

16. B. T. Mayne, T. Bianco-Miotto, S. Buckberry, J. Breen, V. Clifton, C. Shoubridge, C. T. Roberts, Large Scale Gene Expression Meta-Analysis Reveals Tissue-Specific, Sex-Biased Gene Expression in Humans. *Front. Genet.* **7**, 183 (2016).

17. B. Charlesworth, D. Charlesworth, A Model for the Evolution of Dioecy and Gynodioecy. *Am. Nat.* **112**, 975–997 (1978).

18. E. M. Golenberg, N. W. West, Hormonal interactions and gene regulation can link monoecy and environmental plasticity to the evolution of dioecy in plants. *Am. J. Bot.* **100**, 1022–1037 (2013).

19. S. C. H. Barrett, The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3**, 274–284 (2002).

20. D. G. Lloyd, Parental strategies of angiosperms. *N. Z. J. Bot.* **17**, 595–606 (1979).

21. D. G. Lloyd, The transmission of genes via pollen and ovules in gynodioecious angiosperms. *Theor. Popul. Biol.* **9**, 299–316 (1976).

22. J. Käfer, G. A. B. Marais, J. R. Pannell, On the rarity of dioecy in flowering plants. *Mol. Ecol.* **26**, 1225–1241 (2017).

23. G. Bernasconi, J. Antonovics, A. Biere, D. Charlesworth, L. F. Delph, D. Filatov, T. Giraud, M. E. Hood, G. A. B. Marais, D. McCauley, J. R. Pannell, J. A. Shykoff, B. Vyskot, L. M. Wolfe, A. Widmer, Silene as a model system in ecology and evolution. *Heredity* . **103**, 5–14 (2009).

24. I. Casimiro-Soriguer, M. L. Buide, E. Narbona, Diversity of sexual systems within different lineages of the genus Silene. *AoB Plants*. **7** (2015), doi:10.1093/aobpla/plv037.

25. C. Desfeux, B. Lejeune, Systematics of Euromediterranean Silene (Caryophyllaceae): evidence from a phylogenetic analysis using ITS sequences. *C. R. Acad. Sci. III*. **319**, 351–358 (1996).

26. G. Gussarova, G. A. Allen, Y. Mikhaylova, L. J. McCormick, V. Mirré, K. L. Marr, R. J. Hebda, C. Brochmann, Vicariance, long-distance dispersal, and regional extinction-recolonization dynamics explain the disjunct circumpolar distribution of the arctic-alpine plant Silene acaulis. *Am. J. Bot.* **102**, 1703–1720 (2015).

27. V. Balounova, R. Gogela, R. Cegan, P. Cangren, J. Zluvova, J. Safar, V. Kovacova, R. Bergero, R. Hobza, B. Vyskot, B. Oxelman, D. Charlesworth, B. Janousek, Evolution of sex determination and heterogamety changes in section Otites of the genus Silene. *Sci. Rep.* **9**, 1045 (2019).

28. G. A. B. Marais, A. Forrest, E. Kamau, J. Käfer, V. Daubin, D. Charlesworth, Multiple nuclear gene phylogenetic analysis of the evolution of dioecy and sex chromosomes in the genus Silene. *PLoS One*. **6**, e21915 (2011).

29. M. Krasovec, M. Chester, K. Ridout, D. A. Filatov, The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. *Curr. Biol.* **28**, 1832–1838.e4 (2018).

30. A. Rautenberg, L. Hathaway, B. Oxelman, H. C. Prentice, Geographic and phylogenetic

patterns in Silene section Melandrium (Caryophyllaceae) as inferred from chloroplast and nuclear DNA sequences. *Mol. Phylogenet. Evol.* **57**, 978–991 (2010).

31. H. Martin, F. Carpentier, S. Gallina, C. Godé, E. Schmitt, A. Muyle, G. A. B. Marais, P. Touzet, Evolution of Young Sex Chromosomes in Two Dioecious Sister Plant Species with Distinct Sex Determination Systems. *Genome Biol. Evol.* **11**, 350–361 (2019).

32. V. Slancarova, J. Zdanska, B. Janousek, M. Talianova, C. Zschach, J. Zluvova, J. Siroky, V. Kovacova, H. Blavet, J. Danihelka, B. Oxelman, A. Widmer, B. Vyskot, Evolution of sex determination systems with heterogametic males and females in silene. *Evolution*. **67**, 3669–3677 (2013).

33. R. Ming, A. Bendahmane, S. S. Renner, Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* **62**, 485–514 (2011).

34. D. Prentout, thesis, Université de Lyon (2021).

35. N. Zemp, A. Widmer, D. Charlesworth, Has adaptation occurred in males and females since separate sexes evolved in the plant Silene latifolia? *Proc. Biol. Sci.* **285** (2018), doi:10.1098/rspb.2017.2824.

36. A. Muyle, H. Martin, N. Zemp, M. Mollion, S. Gallina, R. Tavares, A. Silva, T. Bataillon, A. Widmer, S. Glémin, P. Touzet, G. A. B. Marais, Dioecy Is Associated with High Genetic Diversity and Adaptation Rates in the Plant Genus Silene. *Mol. Biol. Evol.* **38**, 805–818 (2021).

37. C. Cabau, F. Escudié, A. Djari, Y. Guiguen, J. Bobe, C. Klopp, Compacting and correcting Trinity and Oases RNA-Seq de novo assemblies. *PeerJ*. **5**, e2988 (2017).

38. L. Smeds, A. Künstner, ConDeTri--a content dependent read trimmer for Illumina data. *PLoS One*. **6**, e26314 (2011).

39. R. Schmieder, R. Edwards, Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. **27**, 863–864 (2011).

40. T. D. Wu, J. Reeder, M. Lawrence, G. Becker, M. J. Brauer, "GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality" in *Statistical Genomics: Methods and Protocols*, E. Mathé, S. Davis, Eds. (Springer New York, New York, NY, 2016), pp. 283–334.

41. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).

42. B. Buchfink, K. Reuter, H.-G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods*. **18**, 366–368 (2021).

43. A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. **21**, 3674–3676 (2005).

44. S. Götz, J. M. García-Gómez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talón, J. Dopazo, A. Conesa, High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **36**, 3420–3435 (2008).

45. S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics*. **23**, 257–258 (2007).

46. T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S.

Liu, X. Bo, G. Yu, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. **2**, 100141 (2021).

47. P. Gayral, J. Melo-Ferreira, S. Glémin, N. Bierne, M. Carneiro, B. Nabholz, J. M. Lourenco, P. C. Alves, M. Ballenghien, N. Faivre, K. Belkhir, V. Cahais, E. Loire, A. Bernard, N. Galtier, Reference-free population genomics from next-generation transcriptome data and the vertebrate-invertebrate gap. *PLoS Genet.* **9**, e1003457 (2013).

48. A. Muyle, J. Käfer, N. Zemp, S. Mousset, F. Picard, G. A. Marais, SEX-DETector: A Probabilistic Approach to Study Sex Chromosomes in Non-Model Organisms. *Genome Biol. Evol.* **8**, 2530–2543 (2016).

49. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **26**, 139–140 (2010).

50. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

51. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

52. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

53. M. E. Brooks, K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Machler, B. M. Bolker, glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400 (2017).

54. F. Hartig, DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.3. 3.0 (2020).

55. D. Lüdecke, Ggeffects: Tidy data frames of marginal effects from regression models. *J. Open Source Softw.* **3**, 772 (2018).

56. K. Barton, MuMIn : multi-model inference. R package version 1. 0. 0. *http://r-forge.r-project.org/projects/mumin/* (2009) (available at https://cir.nii.ac.jp/crid/1573387450573265536).

57. J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R. C. Team, Others, Nlme: linear and nonlinear mixed effects models. R package version 3.1-162 (2023).

58. P. D. Price, D. H. Palmer Droguett, J. A. Taylor, D. W. Kim, E. S. Place, T. F. Rogers, J. E. Mank, C. R. Cooney, A. E. Wright, Detecting signatures of selection on gene expression. *Nat Ecol Evol*. **6**, 1035–1045 (2022).

59. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).

60. J. Felsenstein, Phylogenies and the Comparative Method. *Am. Nat.* **125**, 1–15 (1985).

61. T. Garland Jr, Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Am. Nat.* **140**, 509–519 (1992).

62. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. **35**, 526–528 (2019).

63. M. P. Fay, P. A. Shaw, Exact and Asymptotic Weighted Logrank Tests for Interval Censored Data: The interval R package. *J. Stat. Softw.* **36** (2010), doi:10.18637/jss.v036.i02.

64. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. **31**, 3210–3212 (2015).

65. G. G. Cossard, O. Godfroy, Z. Nehr, C. Cruaud, J. M. Cock, A. P. Lipinska, S. M. Coelho, Selection drives convergent gene expression changes during transitions to co-sexuality in haploid sexual systems. *Nat Ecol Evol*. **6**, 579–589 (2022).

66. N. Lichilín, A. El Taher, A. Böhne, Sex-biased gene expression and recent sex chromosome turnover. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200107 (2021).

67. S. A. Rifkin, J. Kim, K. P. White, Evolution of gene expression in the Drosophila melanogaster subgroup. *Nat. Genet.* **33**, 138–144 (2003).

68. P. Khaitovich, G. Weiss, M. Lachmann, I. Hellmann, W. Enard, B. Muetzel, U. Wirkner, W. Ansorge, S. Pääbo, A neutral model of transcriptome evolution. *PLoS Biol.* **2**, E132 (2004).

69. M. Scharmann, A. G. Rebelo, J. R. Pannell, High rates of evolution preceded shifts to sex-biased gene expression in Leucadendron, the most sexually dimorphic angiosperms. *Elife*. **10** (2021), doi:10.7554/eLife.67485.

70. W. R. Rice, SEX CHROMOSOMES AND THE EVOLUTION OF SEXUAL DIMORPHISM. *Evolution*. **38**, 735–742 (1984).

71. D. Bachtrog, A dynamic view of sex chromosome evolution. *Curr. Opin. Genet. Dev.* **16**, 578–585 (2006).

72. A. L. Dapper, M. J. Wade, Relaxed Selection and the Rapid Evolution of Reproductive Genes. *Trends Genet.* **36**, 640–649 (2020).

73. P. W. Harrison, A. E. Wright, F. Zimmer, R. Dean, S. H. Montgomery, M. A. Pointer, J. E. Mank, Sexual selection drives evolution and rapid turnover of male gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4393–4398 (2015).

Acknowledgments

We would like to thank Theo Tricou for his advice with the phylogenetic reconstruction. This work was performed using the computing facilities of the CC LBBE/PRABI and of the Institut Français de Bioinformatique.

Funding: This project was supported through ANR grant ANR-14-CE19-0021-01 to G.A.B.M.

Author contributions:

Conceptualization: GABM, AW

Funding acquisition: GABM, AW

Project administration: GABM, AW

Software: DP, AM, BB

Formal analysis: DP, AM, AeF, BB

Investigation: DP, AM, NZ, BB, PT, AW, JK, GABM

Resources: NZ, AW, PT

Visualisation: DP, AM

Supervision: GABM, JK, AW

Writing—original draft: DP, AM, JK, GABM

Writing—review & editing: All authors

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: Raw reads are available in the NCBI under the project accession number SUB13729399.

Figures and Tables

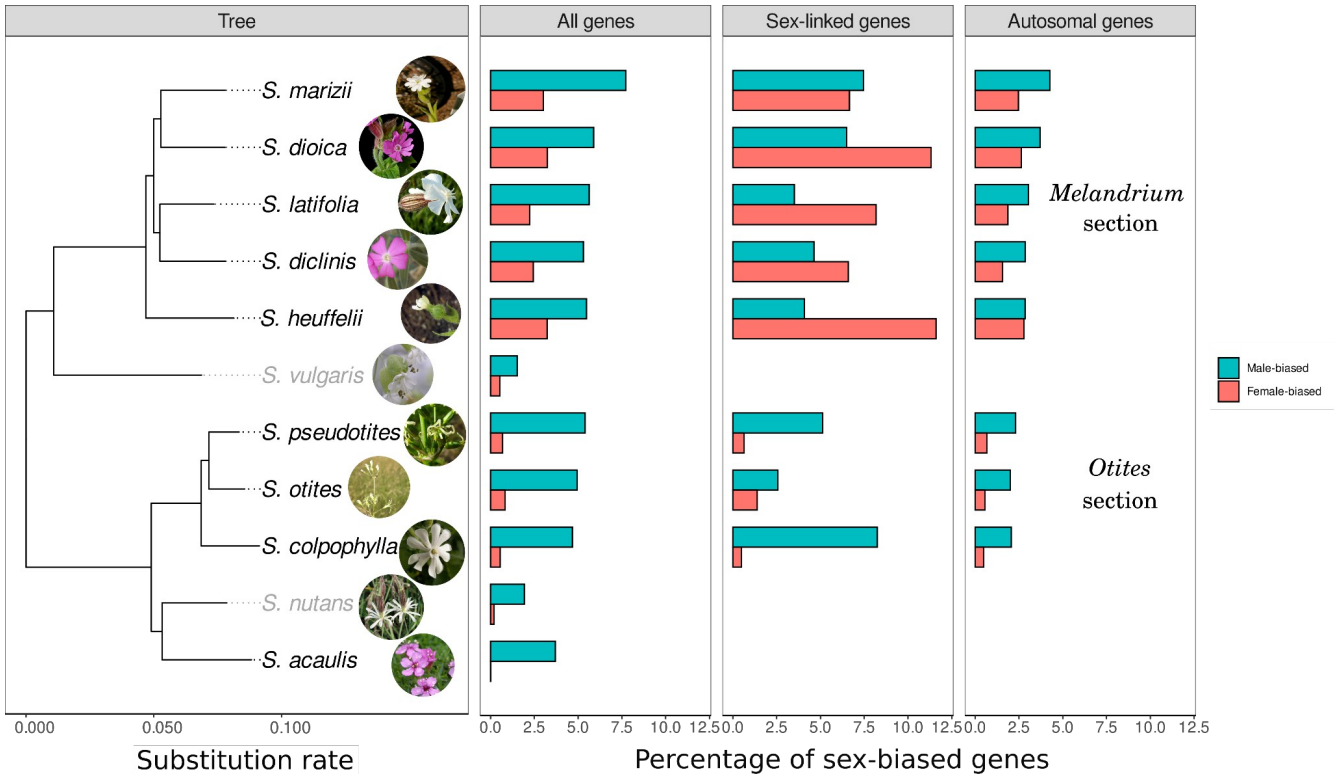

Figure 1: Phylogenetic reconstruction (left panel) of the eleven *Silene* species used in this study and sex-biased gene proportions in each species (the outgroup *Dianthus chinensis* was removed for this figure, see Supplementary Figure S12 for a complete phylogeny). Gynodioecious and dioecious species names are written in grey and black, respectively. The proportion of expressed genes which are female-biased (red) or male-biased (blue) is shown for different gene subsets: all expressed genes (middle left panel), sex-linked genes (middle right panel, only for Otites and Melandrium sections) autosomal genes (right panel, only for Otites and Melandrium sections). See Supplementary Table S3 for detailed gene numbers. (source of the *S. colpophylla* picture: www.earth.com).

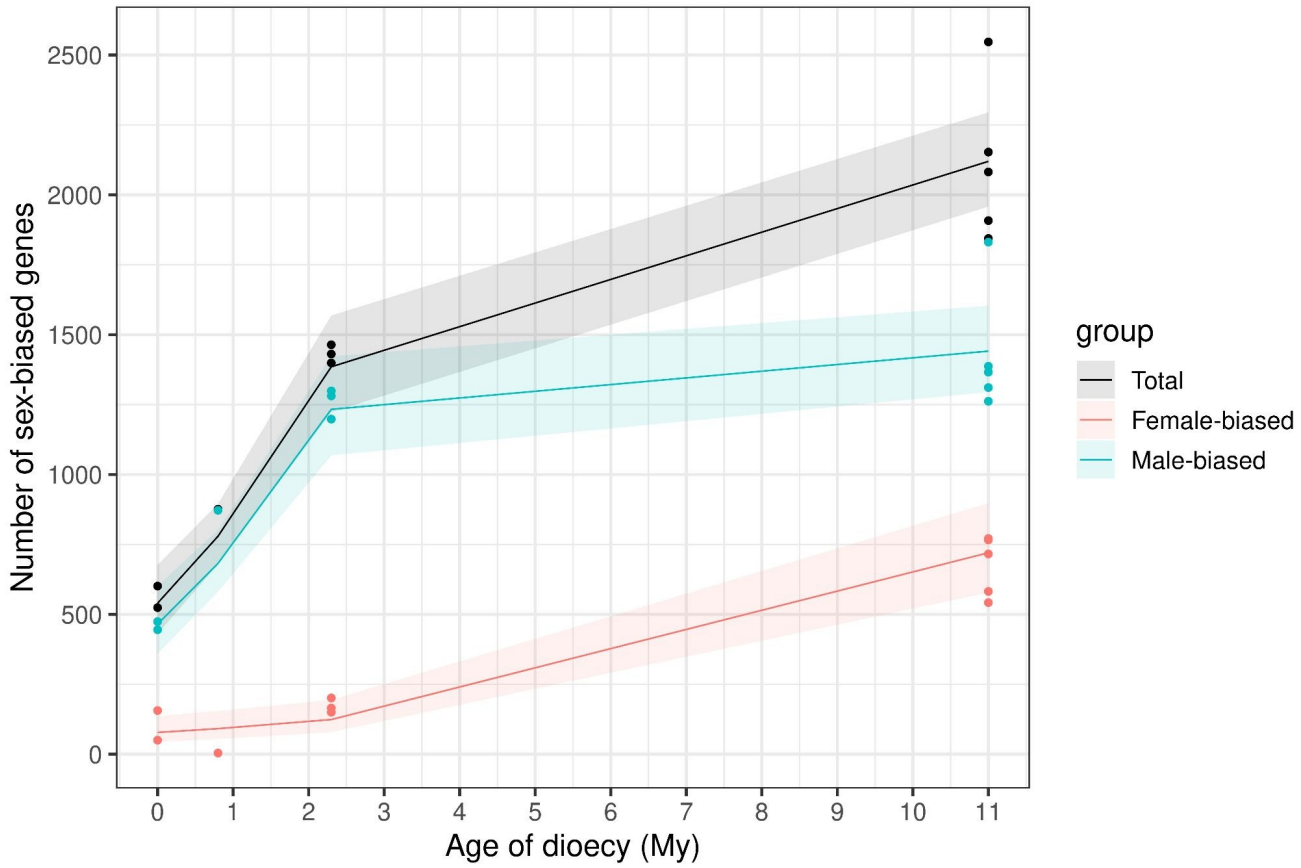

Figure 2: Number of sex biased genes as a function of the age of dioecy (separate sexes) in Million years (My). Gynodioecious species were plotted with age zero of dioecy, *S. acaulis* with age of dioecy 0.8 My, the Otites section 2.3 My and the Melandrium section 11 My. Total sex-biased gene numbers are shown in black, female-biased genes in red and male-biased genes in blue. Dots show the observed data, lines illustrate the predicted values by the generalised linear model detailed in Materials and Methods (equations 1 and 2), ribbons stand for the 95% confidence interval of predicted values. All regressions were significant with $p < 1x10^{-7}$ and $R^2>0.9$ (p-values and R^2 of the models can be found in Supplementary Table S4). For total sex-biased and male-biased genes, the best model included a plateau, specified by a polynomial regression (see Supplementary Table S4 for details). The number of female-biased genes kept increasing with the age of dioecy without reaching a plateau.

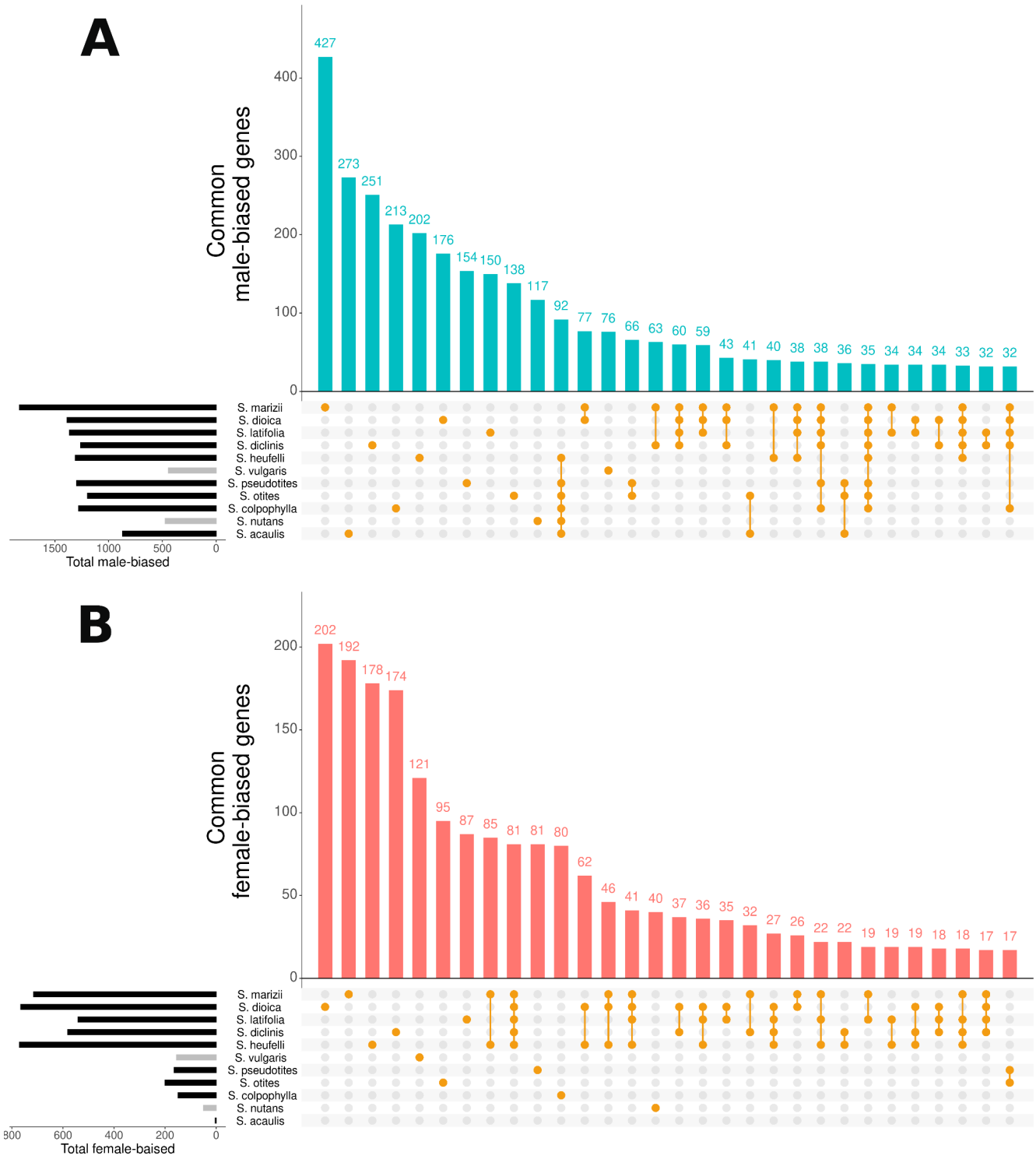

Figure 3: Histogram of male-biased genes (A) and female-biased genes (B) indicating the number of genes that are unique or shared between several species. The yellow dots under each bar indicate the species in which the genes are sex-biased (for example, 192 female-biased genes are unique to *S. marizii* and 85 are shared among *S. marizii* and *S.heuffelii*). For both graphs, only the 30 first bars have been represented (see Supplementary Figure S13 & S14 for additional data).

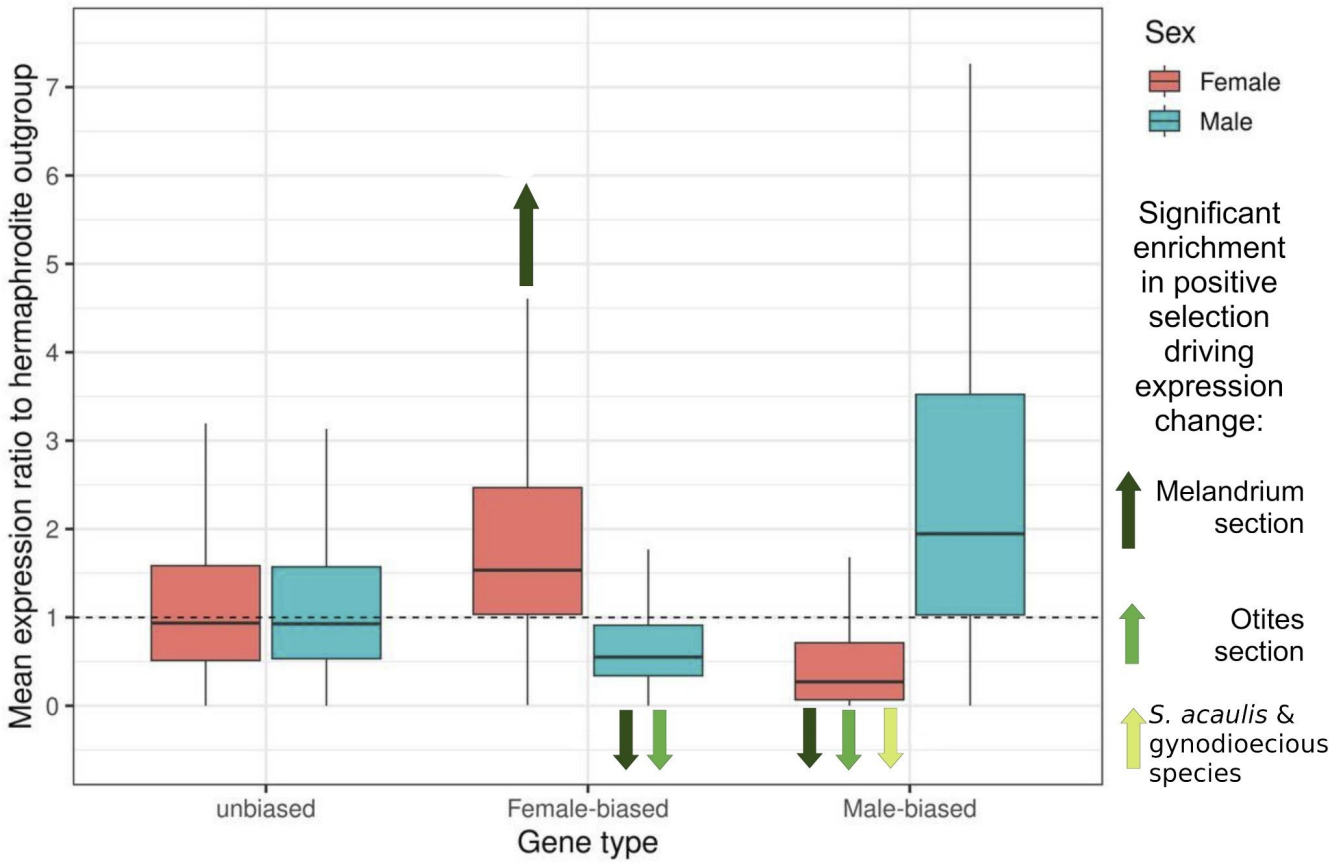

Figure 4: Boxplot of the expression ratio between focal species and their hermaphrodite outgroup. All species were plotted together (see Supplementary Figure S15 for a plot by groups of species). Values higher than one indicate higher expression in the focal dioecious species, and values below one lower expression in the focal species. The arrows summarise the results of the Δ_X analysis (Supplementary Tables S7 and S8). Dark-green arrows indicate that sex-biased genes were significantly enriched in selection for increased or decreased expression in four species of the *Melandrium* section. Medium-shaded green arrows indicate that sex-biased genes were significantly enriched in selection for decreased expression in three species of the *Otites* section. The light-green arrow indicates that male-biased genes were significantly enriched in selection for decreased female expression in *S. acaulis* and in gynodioecious species.