

---

# MATING SYSTEM AND THE EVOLUTION OF RECOMBINATION RATES IN SEED PLANTS

---

July 17, 2024

## Abstract

Meiotic recombination is a central mechanism underlying sexual reproduction among eukaryotes. In many species, the recombination rate is strongly constrained by chromosome size, as the number of crossovers per chromosome generally ranges between one and no more than a few (around three to five). Yet, recombination rates are variable and can evolve between species, in particular when they differ in their reproductive system. According to theory, indirect selection towards higher recombination rates is expected to be stronger in populations with a non-random mating system, such as selfing species compared to randomly mating species. To test for the impact of the mating system on the evolution of recombination rates, we leveraged a dataset with genetic maps, genome sizes, chromosome numbers and life history traits in 200 seed plant species. After controlling for the chromosome size effect, the phylogeny and map quality, we found a joint positive effect of the mating system and longevity on recombination rates, with higher recombination rates in mixed-mating and selfing species. We also found that mixed-mating and selfing species had a significantly higher number of crossovers in larger chromosomes than outcrossing species, suggesting selection for relaxed crossover interference in these former species. Our results point to the mating system as an important factor potentially shaping the evolution of recombination despite the strong constraints imposed on the number of crossovers per chromosome.

**Keywords:** Plants, Recombination, Life history traits, Selfing, Longevity, Genetic map length

## 20 Introduction

21 Meiotic crossovers occur in all sexually reproducing eukaryotes, and lead to the reciprocal exchange of genetic  
 22 material between homologous chromosomes. The average number of crossovers per chromosome is typically  
 23 comprised between one and three in most species (Stapley et al. 2017; Fernandes et al. 2018; Brazier and  
 24 Glémén 2022), which is usually thought to reflect mechanical constraints acting on chromosomal segregation  
 25 during the first meiotic division: at least one crossover per bivalent seems required to ensure the proper  
 26 disjunction of homologs, while too many crossovers may possibly impair correct segregation (Koehler et al.  
 27 1996 – but see Fernandes et al. 2018). However, genetic variation for the number and position of crossovers  
 28 along chromosomes can be observed at different scales: between broad taxa, between closely related species,  
 29 between populations of the same species and between individuals from the same population (e.g. Stapley  
 30 et al. 2017; Dumont, Broman, and Bret A Payseur 2009; Johnston et al. 2016; Brazier and Glémén 2022;  
 31 Peñalba et al. 2020; Samuk et al. 2020). Besides their effect on chromosomal segregation, crossovers also  
 32 lead to genetic recombination, eroding linkage disequilibria (LD) among loci. While an important amount  
 33 of theoretical work has explored the conditions under which breaking LD may provide a selective advantage  
 34 (Lenormand and Sarah P Otto n.d.; Agrawal 2006; Sarah P. Otto 2009), assessing whether such indirect  
 35 benefits may explain patterns of recombination variation between populations and species remains difficult  
 36 (Dapper and Bret A. Payseur 2017; Ritz, Noor, and Singh 2017).

37 Theoretical models have identified different mechanisms generating indirect selection on recombination  
 38 rates, corresponding to different possible sources of LD among selected loci. Selection may generate LD in the  
 39 presence of epistatic interactions among loci, since fitter combinations of alleles tend to increase in frequency.  
 40 When selection remains constant in space and time, breaking these fitter allelic combinations decreases the  
 41 mean fitness of offspring ("recombination load"). Nevertheless, increased recombination may be favored  
 42 when LD impedes adaptation, which is the case when LD is negative (meaning that beneficial alleles tend  
 43 to be associated with deleterious alleles at other loci). Epistatic interactions generate negative LD when the  
 44 combined effect of beneficial alleles at different loci is decreased compared to the multiplicative expectation,  
 45 a situation described as negative epistasis. However, negative epistasis can favor high recombination rates  
 46 only when it is sufficiently weak, so that the recombination load is not too strong (N. H. Barton 1995).  
 47 In finite populations, an additional source of negative LD is selective interference (also known as the Hill-  
 48 Robertson effect; Hill and Robertson 1966; Felsenstein 1974): indeed, while the stochasticity of mutation  
 49 and reproduction in finite populations may generate either positive or negative LD, situations where LD  
 50 is negative tend to persist longer, favoring increased recombination rates (Felsenstein and Yokoyama 1976;  
 51 Nicholas H Barton and Sarah P Otto 2005; Keightley and Sarah P Otto 2006; Hartfield, Sarah P Otto, and  
 52 Keightley 2010; Roze 2021; Bergman, Sarah P Otto, and Feldman 1995).

53 A possible approach to assess whether recombination rate variation may be explained by such indirect  
 54 selective forces could consist in testing to what extent populations or species that differ in some biological or  
 55 ecological characteristics present systematic differences in recombination rates, in a way that would corre-  
 56 spond to theoretical predictions. For example, the models cited above predict that recombination rates may  
 57 be higher in populations that have undergone recent episodes of directional selection, either due to nega-  
 58 tive epistasis or to selective interference between linked beneficial mutations (favoring genetic modifiers that  
 59 increase recombination). While this prediction is corroborated by the fact that increased recombination is  
 60 observed more often than decreased recombination after artificial selection experiments (Table 1 in Sarah P.  
 61 Otto and Nick H. Barton 2001), assessing whether a natural population has recently undergone directional  
 62 selection is generally difficult.

63 Interestingly, the mating system of organisms can vary greatly among species (sometimes among popula-  
 64 tions from the same species), and may have strong effects on indirect selection on recombination. In partic-  
 65 ular, self-fertilization is a common reproductive mode in hermaphroditic species, about 20% of Angiosperm  
 66 species being predominantly selfing (Barrett 2002). The effect of selfing on the evolution of recombination  
 67 stems both from the increase in homozygosity of individuals caused by inbreeding, and from correlations  
 68 in homozygosity among loci that are generated as long as some level of outcrossing is maintained in the  
 69 population. As shown by the recent analysis of Stetsenko and Roze (2022) focusing on the effect of deleteri-  
 70 ous mutations, the decreased efficiency of recombination caused by homozygosity has two contrasted effects  
 71 on the evolution of a recombination modifier: (1) it decreases indirect selection by reducing the effect of  
 72 the modifier on effective recombination rates (indirect selection on recombination vanishing under complete  
 73 selfing); (2) it increases the magnitude of LD, enhancing indirect selection on recombination. As a result,  
 74 selection for recombination (either due to selective interference or to negative epistasis) is often predicted  
 75 to be maximal for selfing rates slightly below one. Additionally, and as shown previously by Roze and  
 76 Lenormand (2005), correlations in homozygosity among loci generated by partial selfing may strongly favor  
 77 recombination in the presence of a particular form of negative epistasis (negative dominance-by-dominance  
 78 epistasis). In that case, selection for recombination may be maximized for moderate selfing rates (see Figure  
 79 5c in Stetsenko and Roze 2022).

80 These theoretical results thus show that self-fertilization should generally enhance indirect selection for  
 81 recombination, except under complete selfing. Do partially selfing species exhibit higher recombination  
 82 rates than outcrossing ones? Cytological studies from different genera of Angiosperms suggest that it may  
 83 indeed be the case, as self-pollinating species tend to show higher numbers of chiasmata per bivalent than  
 84 outcrossing species from the same genus (Roze and Lenormand 2005; Ross-Ibarra 2007). This positive  
 85 correlation was found to be significant across all species, as well as within each genus (Roze and Lenormand  
 86 2005). Furthermore, comparisons of the genetic maps of the highly selfing *Arabidopsis thaliana* and of its  
 87 outcrossing relative *Arabidopsis lyrata* also show higher recombination rates in *A. thaliana* (Kuittinen et al.

88 2004; Hansson et al. 2006; Kawabe et al. 2006). Since then, the number of species for which genetic maps  
 89 are available keeps increasing rapidly, opening new possibilities to test whether recombination rates tend to  
 90 correlate with certain biological traits of organisms.

91 In this article, we present an analysis of a dataset consisting of about 200 plant species for which genome  
 92 size, number of chromosomes and genetic map length are available. Species were categorized based on their  
 93 mating system and other life-history traits. The results show that species classified as mixed mating or selfing  
 94 tend to have higher chromosome map lengths than outcrossing species, thus confirming the positive effect  
 95 of selfing on recombination found in previous studies. Furthermore, mixed-mating and selfing species with  
 96 larger chromosomes tend to have higher numbers of crossovers per chromosome than species with smaller  
 97 chromosomes, while this trend is not observed in outcrossing species. Among the other traits tested, only  
 98 longevity showed a significant association with recombination, with medium-lived species presenting lower  
 99 recombination rates than short-lived or long-lived species.

## 100 Materials and Methods

### 101 Recombination dataset

102 The original dataset with total genetic map length (cM), number of chromosomes and genome physical  
 103 size (Mb) was retrieved from Stapley et al. (2017). We added genetic map length and genome sizes for  
 104 24 new species to the 184 species already in Stapley et al. (2017). This final dataset contained a list of  
 105 208 species. We returned to original publications to retrieve the exact ploidy level, the number of markers  
 106 on the final map and the number of progeny. In order to properly control for the chromosome size effect  
 107 already assessed in this data (Stapley et al. 2017), we also defined new genomic variables, such as the average  
 108 chromosome genetic map length (cM/chromosome) and the residuals of the linear regression of recombination  
 109 rates (cM/Mb) as a function of average chromosome size (Mb).

### 110 Life History Traits

111 Based on a literature search, we gathered Life History Traits data for every species (Table S1). We first  
 112 retrieved mating system information, using the coarse classification outcrossing / mixed / selfing, when  
 113 directly given in the literature. To complete the characterization of the mating system, we collected other  
 114 reproductive traits: sexual system (andromonoecy, dioecy, gynodioecy, hermaphroditism, monoecy), self  
 115 incompatibility (SI) status (self-incompatible and self-incompatible) and outcrossing rate, which we used to  
 116 characterize the mating system: dioecious and SI species were classified as outcrossing; based on outcrossing  
 117 rate ( $t$ ) we classified species as outcrossing for  $t > 0.80$ , mixed for  $0.2 < t < 0.8$  and selfing for  $t < 0.2$ . We  
 118 added the life form (herb, liana, shrub, tree, vine) and the life span (annual, biannual, perennial), which are  
 119 associated with the mating system. In addition it has been proposed that long life-span should select more

recombination (Burt and Bell 1987), although the underlying mechanism is unclear. As some categories were represented by only a few species and some traits were correlated, we defined three categories reflecting longevity: annual species, perennial non-woody species (herbs and vines), and woody perennials (trees, shrubs and lianas). We also included the cultivation status (cultivated, wild), as domestication has been also proposed to select higher recombination rate because of recent and prolonged episodes of directional selection (Ross-Ibarra 2007; Burt and Bell 1987). Finally, we categorized species as diploid (ploidy level = 2) or polyploid (ploidy level > 2). We added more phylogenetic levels (class, family) and we grouped species in large phylogenetic families (Superasterids, Lilioids + Alismatids Superrosids, Commelinids, Magnoliids, Basal eudicots, Conifers). The full dataset is available as Table S1.

### 129 Phylogenetic signal

We used the 'phytools' R package to manipulate phylogenetic tree data and to statistically test for a phylogenetic signal (Revell 2024). The phylogenetic time-calibrated supertree used for the comparative phylogenetic dataset was retrieved from Smith and Brown (Smith and Brown 2018). Tip labels were curated to remove subspecies and cultivar annotations and one tip per species was retained. Species missing in the tree were assigned randomly to a sister species of the same genus. The tree was forced to be ultrametric. We computed two phylogenetic signal metrics, Blomberg's K and Pagel's lambda. We also fitted three evolution models (Brownian Motion, Ornstein-Ulhenbeck and Early Burst) to the distribution of recombination rates, chromosome genetic map length and the residuals of the regression presented above.

### 138 Statistical analyses

All statistical analyses were performed on R version 4.3.3 (R Core Team 2022). Linear regressions were performed with the R base function 'lm'. We used the 'caper' R package to fit Phylogenetic Generalized Linear models with the function 'pgls' (Orme et al. 2018). We performed a forward stepwise model selection based on Anova and AIC/BIC criterion. We tested the significance of predictors with the 'anova' R function. The validity of the models was evaluated by visually checking the normality of the residuals (Q-Q plot and histogram), the homoscedasticity of the residuals, and the observed vs. fitted values and the residuals vs. fitted values to detect correlated errors.

## 146 Results

### 147 Dataset

After removing species with no information on the mating system, we kept 200 plant species, including 190 angiosperms and 10 gymnosperms. The map quality was globally good with 20 species with more than 200 markers (18 maps without information on the number of markers). The minimal number was 82 markers and

the highest number was 64,263 markers, with a median number of 998 markers. The number of progenies ranged from 43 to 3,480, with a median number of 149 progenies. We had a sampling covering the whole angiosperm phylogeny plus a few gymnosperms (Figure S1). A few families had a lot of species (e.g. Poaceae, Fabaceae, Rosaceae and Brassicaceae had more than 10 species), but a majority of species was the single representative of their family.

Among the 200 species, the number of chromosomes ranged from 5 to 90, with 170 diploid and 30 polyploids. We had 45 selfing species, 37 mixed mating species and 118 outcrossing species (including 21 dioecious species). Seventy-four species were long-lived (woody perennial), 47 were medium-lived (non-woody perennial) and 79 short-lived species (annual). We had 49 self-compatible species (SC) and 56 self-incompatible (SI). We had 109 cultivated and 91 wild species, though some species can be both cultivated and wild.

## Control for the strong chromosome size effect

Based on previous studies (Brazier and Glémén 2022; Haenel et al. 2018) we postulated that chromosome size was the main determinant of recombination rates. Indeed we confirmed the strong negative correlation (log-log relationship) between chromosome size and recombination rate in our dataset (Figure 1A). Commelinids and conifers had amongst the largest chromosome size, while other plant species were on a lower range of chromosome size. The average genetic map length was significantly yet only weakly correlated with chromosome size (Spearman's  $\rho = 0.21$ ,  $p = 0.002$ , Figure 1B) and not with the number of chromosomes (Spearman's  $\rho = -0.05$ ,  $p = 0.49$ , Figure S2). The distribution of chromosome genetic map length was fat-tailed with a few chromosomes larger than 200 cM (range 44-347 cM/chromosome, Figure 1C). Finally, average chromosome map lengths varied homogeneously among phylogenetic families, except for Lilioids + Alismatids and basal Eudicots counting only a few species (Figure 1D, compare with recombination rates (cM/Mb) in Figure 2 in Stapley et al. 2017). Overall our results suggest that the absolute number of COs per chromosome (i.e. average chromosome genetic map length) are susceptible to evolve in a reasonable range independently of chromosome size.

## Weak phylogenetic signal

After controlling for the chromosome size effect, we also investigated how recombination rates evolved amongst the plant phylogeny. In order to test the significance and strength of the phylogenetic signal, we estimated two phylogenetic signal metrics (Blomberg's K and Pagel's lambda) and fitted three competing phylogenetic models (Brownian Motion, Ornstein-Uhlenbeck and Early-burst) to the continuous distribution of chromosome average genetic length and the residuals of the regression of recombination rates as a function of average chromosome size (Table 1). Blomberg's K were not significantly different from zero while Pagel's Lambda was around 0.5, both indicating a weak phylogenetic signal. Based on the maximum

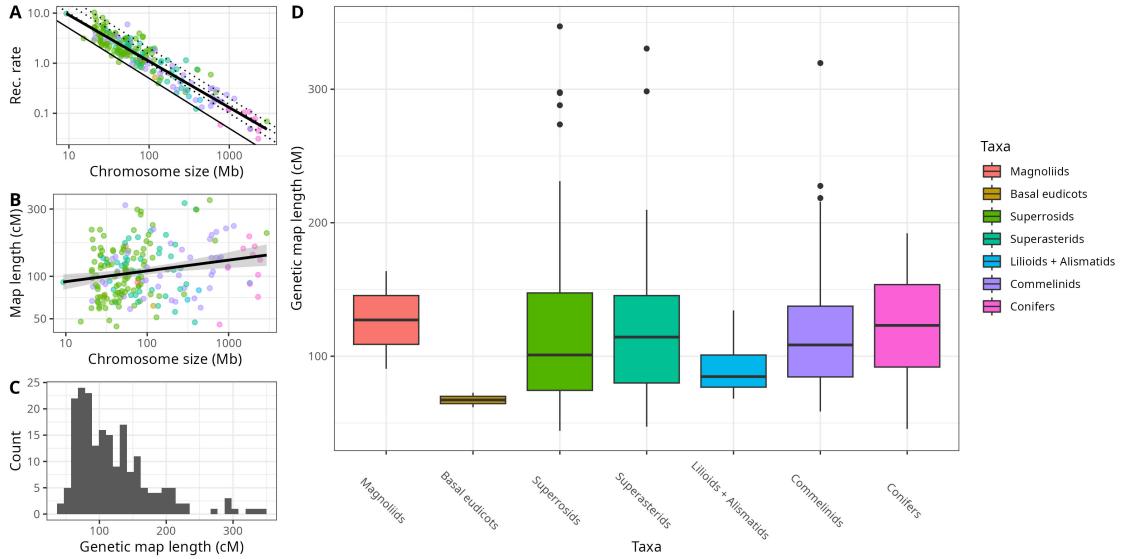


Figure 1: Distribution of recombination rates and genetic map length ( $n = 200$  species). (A) Average chromosome size (Mb, log scale) is negatively correlated with recombination rates (cM/Mb, log scale). (B) Average genetic map length (cM) is weakly correlated to average chromosome size (Mb) (C) Density of chromosome genetic map length. (D) Boxplots of chromosome genetic map length per phylogenetic family. Compare with recombination rates (cM/Mb) in Figure 2 in Stapley et al. (2017).

184 Log-Likelihood criterion, the Ornstein-Uhlenbeck phylogenetic model was always preferred, suggesting that  
 185 the recombination rate evolved under a constrained range.

Table 1: Phylogenetic signal for the two proxies of recombination rates. Blomberg's K and Pagel's Lambda (and their respective p-value) and the Log-Likelihood of the three phylogenetic models (Brownian Motion, Ornstein-Uhlenbeck and Early-burst, respectively). The best model according to the Log-Likelihood criterion is in bold.

Trait	K	K p-value	$\lambda$	$\lambda$ p-value	Log-lik BM	Log-lik OU	Log-lik EB
Genetic map length	0.09	0.117	0.46	0.01	-1155.09	<b>-1091.53</b>	-1155.09
Residuals	0.1	0.032	0.47	0.01	-173.16	<b>-110.77</b>	-173.16

### 186 Joint effect of the mating system and longevity

187 Based on ANOVAs and AIC/BIC criterion, we run a step forward model selection for Linear Regression  
 188 and Phylogenetic Generalized Least Squares (Table S2). For both average chromosome map length and the  
 189 residuals, the two significant predictors were always the mating system first, then longevity based on ANOVA  
 190 ( $p < 0.05$ ) and AIC (lower AIC value). However these effects were weak (Figure S3, S4), as the null model  
 191 was always preferred by the BIC criterion. The effect of other life history traits was weak or null. The more  
 192 complex models including both mating system and longevity were significantly better than the two models  
 193 with a single predictor, except for the Linear Model with the average chromosome map length as response  
 194 variable ( $df = 2$ ,  $F = 1.9095$ ,  $p = 0.151$ ). The effect was stronger when the phylogeny was accounted for  
 195 ( $p = 0.035$  for 'lm' vs.  $p = 0.008$  for 'pgls'). Overall, we observed a joint effect of the mating system and  
 196 longevity on proxies of recombination rates (Figure 2, S5).

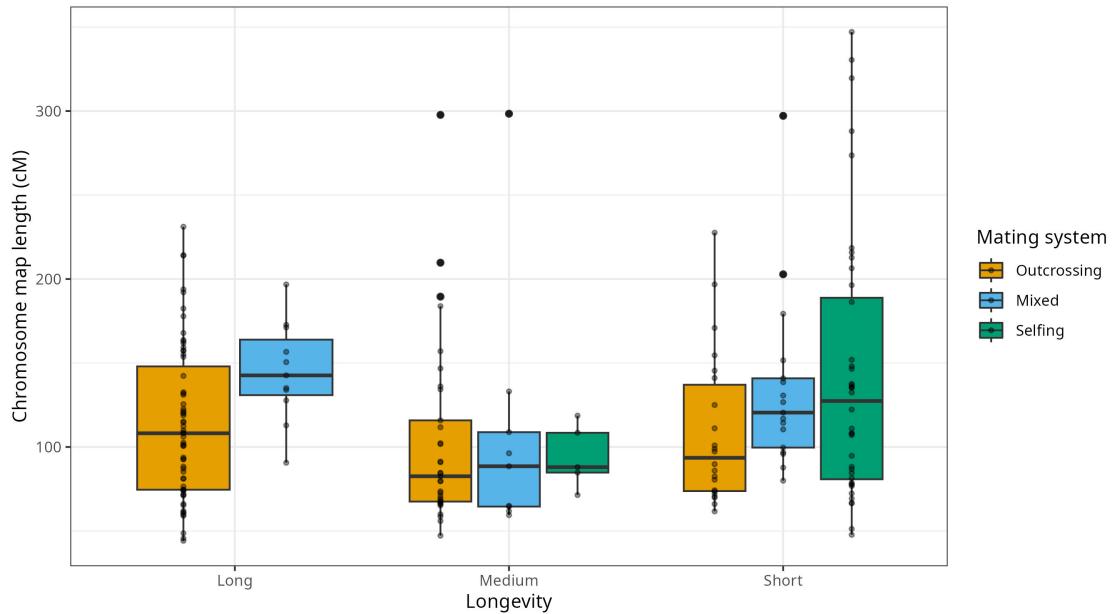


Figure 2: Recombination rates depend on the mating system and longevity. (A) The combined effect of the mating system and longevity on the average chromosome genetic map length. Each point is a species ( $n = 200$  species).

Table 2: Model fit and parameter estimates of the best model (PGLS residuals = mating system + longevity,  $\lambda = 0.480$ , F-statistic = 4.069, df = 195, p = 0.0034, Adjusted R-squared = 0.0581). Significance codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 ' 1.

Response	Df	p	Parameter	Estimate	Std. Error	t	p
Mating system	2	0.0078 **	(Intercept)	-0.0789	0.2145	-0.368	0.7133
			Mixed-mating	0.1805	0.0824	2.1898	0.0297 *
			Selfing	0.1857	0.0891	2.0839	0.0385 *
Longevity	2	0.0443 *	Medium longevity	-0.2201	0.0941	-2.3395	0.0203 *
			Short longevity	-0.0741	0.0994	-0.7461	0.4565

197 As a final model we selected the Phylogenetic Generalized Least Squares with the residuals as a response  
 198 variable ( $\lambda = 0.480$ , F-statistic = 4.069, df = 195, p = 0.003433, Adjusted R-squared = 0.0581, Table  
 199 2) where selfing and mixed mating species had significantly higher recombination than outcrossing species  
 200 whereas medium-lived species had significantly lower recombination (Table 2). Both the mating system and  
 201 longevity were significant in the ANOVA (df = 2, Sum of Squares = 0.0076, Mean Squares = 0.0038, F =  
 202 4.9710, p = 0.0078, and df = 2, Sum of Squares = 0.0048, Mean Squares = 0.0024, F = 3.1666, p = 0.0443,  
 203 respectively). The validity of the model was successfully assessed with diagnostic plots. The model with the  
 204 average genetic map length yielded similar results.

205 We observed the same results when we plotted the joint effect of the mating system and longevity on  
 206 recombination rates (Figure 2, S5). Outcrossing species had on average lower recombination rates for the  
 207 three longevity categories though the difference was less clear in medium-lived species. On average, selfing  
 208 and mixed mating system species had higher recombination rates. There was no selfing species in long-lived

209 species. Medium-lived species had lower recombination rates on average, with globally a maximum of two  
 210 COs per chromosome (i.e. 100 cM) while long and short-lived generally exceeded two COs.

211 **Selection for higher recombination varies among phylogenetic families**

212 We investigated if there could be a family-specific effect of the mating system or longevity. Our phylogenetic  
 213 sampling was sparse across the plant phylogeny but we managed to subset three independent families with  
 214 at least ten species each and two species for each mating system. The family-specific selfing effect was not  
 215 clear among the three families (Figure 3A). The increased recombination rate in selfing species was clear in  
 216 Poaceae but weak or even reversed in Brassicaceae and Fabaceae. We observed only a few mixed-mating  
 217 species. For Fabaceae and Poaceae we observed a clear short-lived effect with increased recombination rates  
 218 (Figure 3B). We did not have enough medium-lived species to conclude for Brassicaceae. We did not have  
 219 enough samples to investigate the joint effect or to perform a proper statistical analysis.

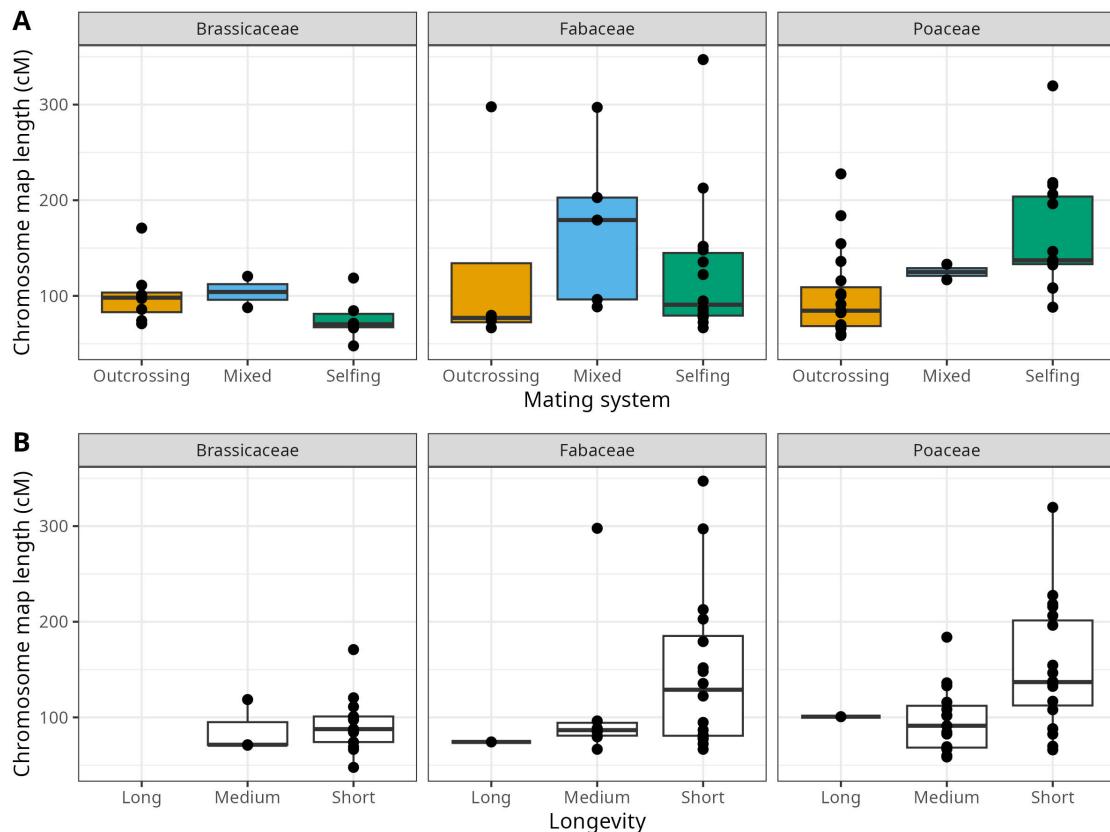


Figure 3: Chromosome genetic map length as a function of the mating system (same colors as in Figure 2)  
 or longevity in three different phylogenetic families ( $n = 91$  species). Each point is a species.

220 **Selection towards extra crossovers in selfing and mixed mating species**

221 We analyzed the dataset slightly differently by considering that there is one mandatory CO per chromosome  
 222 and then the additional number of CO can increase or not with chromosome size, depending on the strength  
 223 of CO interference. If CO interference is limited, larger chromosomes should have on average a higher  
 224 number of COs. On the contrary, if interference is the limiting factor, evolution towards more COs on larger  
 225 chromosomes should be prevented. We tested this idea by analyzing separately for each mating system the  
 226 slope of the average chromosome genetic map length as a function of chromosome physical size (Figure 4).  
 227 Species with larger chromosomes seemed to have more COs per chromosome in selfing and mixed-mating  
 228 species but not in outcrossing species. The intercept was similar among the three mating systems (Figure 4).  
 229 We tested the significance of this effect with a PGLS model with an interaction term between the chromosome  
 230 size and the mating system. We also used the number of chromosomes as a covariate, since the Genome  
 231 Wide Recombination Rate can be increased by increasing the total number of chromosomes (chromosome  
 232 genetic map length = chromosome size\*mating system + number of chromosomes,  $\lambda = 0.468$ , F-statistic =  
 233 3.732, df = 193, p = 0.0016, Adjusted R-squared = 0.0761). There was a trend for the interaction between  
 234 chromosome size and the mating system (p = 0.056), as well as a trend for the number of chromosomes (p  
 235 = 0.071; see Table S3 for model fit and parameters). While there was a slight trend for the chromosomal  
 236 genetic map length to increase with chromosome size in selfing species (interaction chromosome size:selfing,  
 237 coefficient = 0.069, std. error = 0.0359, t = 1.9229, p = 0.056), the number of chromosomes had a negative  
 238 effect on the chromosomal genetic map length (coefficient = -0.946, std. error = 0.5213, t = -1.8147, p =  
 239 0.0711).

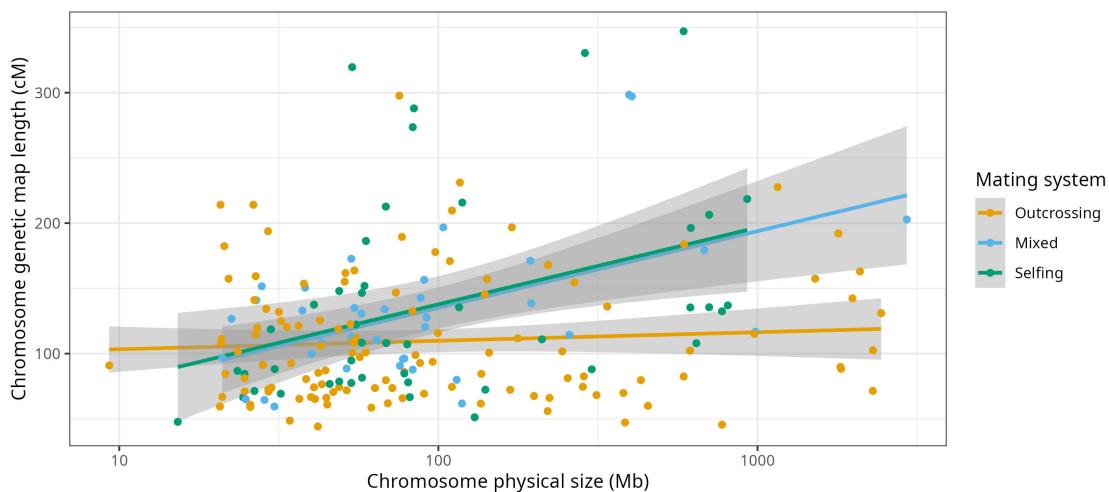


Figure 4: Selection towards higher CO rates in larger chromosomes for selfing and mixed-mating species (n = 118, 37 and 45 for outcrossing, mixed-mating and selfing species, respectively). The linear regression line and its 95% confidence interval for each mating system were estimated with the ggplot2 'geom\_smooth' function (Wickham 2016).

**240 The effect is robust to map quality**

241 In order to control if the mating system and longevity effect were robust to differences in map quality among  
 242 species, we tested the influence of marker density and number of progenies on the significance of the results.  
 243 We added either marker density or number of progenies to the Phylogenetic Generalized Least Squares model  
 244 with the residuals as a response variable. Marker density had a significant positive effect on recombination  
 245 rates ( $p = 0.028035$ , see parameter estimates in Table S4) but it did not change the significance of selfing  
 246 and medium-lived species ( $p = 0.0469$  and  $0.0318$ , respectively) while mixed-mating species remained as a  
 247 trend ( $p = 0.0681$ ). Progeny number was not significant at all ( $p = 0.5988$ ).

**248 Discussion**

249 In order to test for an effect of the mating system on the evolution of recombination rates, we compared  
 250 genetic maps in 200 plant species differing by their mating system and other life-history traits. We found  
 251 a joint positive effect of the mating system and longevity on recombination rates with variation across  
 252 phylogenetic families. We also found that mixed-mating and selfing species had a significantly higher number  
 253 of crossovers in larger chromosomes compared to outcrossing species. These results have implications for the  
 254 evolution of recombination given the constraints imposed on the number of crossovers per chromosome.

**255 Higher CO rates in selfing and mixed-mating species**

256 Among life history traits, the mating system has the main significant effect on recombination rates after  
 257 controlling for the strong chromosome size effect. Indeed, chromosome size is a major determinant of re-  
 258 combination rates in Eukaryotes (Brazier and Glémin 2022; Stapley et al. 2017; Haenel et al. 2018) but  
 259 our results show that the average chromosome genetic map length (i.e. the average number of COs) is only  
 260 weakly proportional to chromosome size. The average number of COs per chromosome varies between one  
 261 and four for most species, with some species exceeding six COs per chromosome (300 cM). Despite con-  
 262 straints on the number of COs per chromosome, as supported by the choice of the OU model (Table 1),  
 263 there is still possible variation upon which selection can act.

264 Selfing and mixed-mating species have higher recombination rates on average than outcrossing species,  
 265 thus supporting the theoretical predictions of a positive effect of selfing on recombination(Roze and Lenor-  
 266 mand 2005; Stetsenko and Roze 2022). Under realistic parameter values, Stetsenko and Roze (2022) found  
 267 that selfing increases selection for recombination to compensate for the decreased efficacy of recombination.  
 268 However, they predict that selection for recombination vanishes for selfing rates approaching one as, in this  
 269 case, homozygosity is too high for recombination to have an effect. This results in a non-monotonic increase  
 270 in recombination with selfing, with the selfing rate maximizing selection for recombination depending on pa-  
 271 rameter values, but typically being quite high. Our data are not precise enough to directly assess the nature

272 of the relationship between selfing rate and recombination rate. However, we observed the main difference  
 273 between outcrossing and mixed-mating + selfing, with no significant difference between mixed mating and  
 274 selfing, which is in agreement with the non-linear relationship predicted by theory assuming particular forms  
 275 of epistasis between deleterious mutations.

276 Our study also matches well with previous empirical results. Cytological measures of the number of chiasmata per bivalent per meiosis in several plant genera tend to confirm this overall increase of recombination  
 277 with selfing (reviewed in Roze and Lenormand 2005 and Ross-Ibarra 2007). This positive correlation was  
 278 found to be significant across all species as well as within genera (Roze and Lenormand 2005). However,  
 279 chiasma frequency only represents an indirect measure of the genome-wide recombination rate and is only  
 280 available in a dozen genera. Similar results were found from genetic maps, pointing at higher recombina-  
 281 tion rates in the highly selfing *Arabidopsis thaliana* compared to its outcrossing relative *Arabidopsis lyrata*  
 282 (Hansson et al. 2006; Kawabe et al. 2006; Kuittinen et al. 2004), but it was only in a single pair of species.  
 283

284 We observed a significant joint effect of the mating system and longevity. Recombination rates are similar  
 285 at the extreme life-spans, in short and long-lived species (annual and woody perennial, respectively), while  
 286 medium-lived species (perennial non-woody species such as herbs and vines) experience lower recombination  
 287 rates on average. Taking longevity into account reveals a stronger effect of the mating system in short-  
 288 and long-lived species than in medium-lived species (note that long-lived selfing species are exceedingly rare  
 289 and absent from our dataset). In mammals, Burt and Bell (Burt and Bell 1987) proposed that a long  
 290 life-span should select more recombination but the mechanism remains unclear. In plants, a meta-analysis  
 291 on chiasma frequencies found lower recombination in perennials than in annual species, but they did not  
 292 distinguish medium- and long-lived perennials and did not control for the mating system as we did (Koella  
 293 1993).

#### 294 Indirect evidence of reduced CO interference to increase the CO number

295 The CO number per chromosome per meiosis is strongly constrained by CO assurance and interference  
 296 (Wang et al. 2015). While CO assurance guarantees at least one CO, the maximum number of COs is  
 297 strongly limited by CO interference, up to a maximum of four COs in most species (Brazier and Glémén 2022;  
 298 Stapley et al. 2017). This limited evolvability probably explains why the mating system and longevity have  
 299 a marginal effect that is significant only after controlling for the strong chromosome size effect. The proxies  
 300 of recombination rates have a weak phylogenetic signal and the Ornstein-Uhlenbeck model of evolution is  
 301 always preferred, suggesting stabilizing selection on the number of COs.

302 However indirect selection on the strength of CO interference might be a possible mechanism to allow  
 303 evolution towards a higher CO number, especially in non-random mating populations. The strength of CO  
 304 interference itself varies dramatically among species (Sarah P. Otto and Bret A. Payseur 2019) and CO  
 305 interference is thought to evolve in finite populations as a way to reduce selective interference among loci

306 (Nicholas H Barton and Sarah P Otto 2005; Keightley and Sarah P Otto 2006; Roze and Nick H Barton 2006).  
 307 Indeed we observe a higher number of COs per chromosome in mixed-mating and selfing species, particularly  
 308 in larger chromosomes, while this trend is not observed in outcrossing species. This suggests that the higher  
 309 number of COs in mixed-mating and selfing species may evolve through relaxed CO interference allowing  
 310 more COs to occur on the same chromosome. On the contrary, CO interference may have remained strong  
 311 in outcrossing species, limiting the CO number to its minimum independently of chromosome size.

312 One interesting question is whether one or two extra COs are sufficient to efficiently compensate for  
 313 the reduced efficiency of selection due to selfing. Considering that most of the genetic shuffling is done via  
 314 the independent assortment of chromosomes (inter-chromosomal shuffling), a more efficient way to increase  
 315 the recombination rate would be increasing the chromosome number to produce more COs on the whole  
 316 (Veller, Kleckner, and Nowak 2019). Selection for increased number of COs per chromosome, increased by  
 317 the effect of selfing, should thus be stronger in species with a small number of chromosomes. The number  
 318 of COs per chromosome can evolve more rapidly than the number of chromosomes, possibly on the same  
 319 timescale as selfing rates (Henderson and Bomblies 2021; Whitehead et al. 2018). It is thus more likely that  
 320 the number of chromosomes constrains the evolution of CO number rather than being the direct target of  
 321 selection for increased genetic shuffling. We found that at a chromosome level, the number of chromosomes  
 322 was negatively associated with the chromosome genetic map length, though not significantly. At a genome  
 323 wide level, Stapley et al. (2017) found a weak positive effect of chromosome number on the genome wide  
 324 recombination rate (cM/Mb) in plants, but not in animals and fungi.

### 325 Strength and limitations of our dataset

326 The effect of the mating system and longevity is moderate, most likely because it is biologically much  
 327 weaker than the dominant chromosome size effect. We were able to detect it by leveraging a large curated  
 328 dataset from Stapley et al. (2017) plus 24 new species with a sampling across seed plants (angiosperms and  
 329 gymnosperms). We also found that the results were robust to differences in map quality (approximated by  
 330 marker density and number of progeny).

331 However, our dataset remains limited to detect complex effects. The interaction between chromosome  
 332 size and the mating system is only a trend. We did not detect significant effects for other life-history traits,  
 333 despite strong expectations for some of them (e.g. cultivation status). There is multiple evidence of an effect  
 334 of domestication on recombination rates in plants (Ross-Ibarra 2007; Dreissig, Mascher, and Heckmann  
 335 2019; Fuentes et al. 2021; Schreiber et al. 2022). However, we were not able to distinguish cultivated and  
 336 domesticated species in our dataset and many species can be either strongly domesticated or wild relative  
 337 populations. For such traits, a pairwise sampling of closely related species/populations with contrasted life  
 338 history traits (wild vs. domesticated) should be more powerful than our large unstructured sampling. In  
 339 particular, comparing species with the same number of chromosomes and similar genome size would avoid

340 the strong confounding effect of chromosome size. It is likely that such an approach would also strengthen  
 341 the detection of the effect of the mating system.

342 It is difficult to empirically test theories on the evolution of recombination and comparative analyses are  
 343 a valuable approach as a first step to identify relevant factors. Despite the limitations discussed above, our  
 344 results point to the mating system as an important factor potentially shaping the evolution of recombination.  
 345 Further detailed studies of recombination patterns in species with contrasting selfing rates is a promising  
 346 approach to better understand the forces acting on the evolution of recombination and, in turn, the evolution  
 347 of selfing rates.

### 348 Competing interests

349 The authors declare no conflicts of interest.

### 350 Data availability

351 All scripts and data necessary to reproduce this study are available at the OSF repository  
 352 <https://osf.io/XXX>.

### 353 References

- 354 Agrawal, Aneil F. (2006). “Evolution of Sex: Why Do Organisms Shuffle Their Genotypes?” In: *Current  
 355 Biology* 16.17, R696–R704. doi: [10.1016/j.cub.2006.07.063](https://doi.org/10.1016/j.cub.2006.07.063).
- 356 Barrett, Spencer C. H. (2002). “The Evolution of Plant Sexual Diversity”. In: *Nature Reviews Genetics* 3.4,  
 357 pp. 274–284. doi: [10.1038/nrg776](https://doi.org/10.1038/nrg776).
- 358 Barton, N. H. (1995). “A General Model for the Evolution of Recombination”. In: *Genetical Research* 65.2,  
 359 pp. 123–144. doi: [10.1017/S0016672300033140](https://doi.org/10.1017/S0016672300033140).
- 360 Barton, Nicholas H and Sarah P Otto (2005). “Evolution of Recombination Due to Random Drift”. In:  
 361 *Genetics* 169.4, pp. 2353–2370.
- 362 Bergman, Aviv, Sarah P Otto, and Marcus W Feldman (1995). “On the Evolution of Recombination in  
 363 Haploids and Diploids: I. Deterministic Models”. In: *Complexity* 1.1, pp. 57–67.
- 364 Brazier, Thomas and Sylvain Glémin (2022). “Diversity and Determinants of Recombination Landscapes in  
 365 Flowering Plants”. In: *PLOS Genetics* 18.8. Ed. by Ian R. Henderson, e1010141. doi: [10.1371/journal.pgen.1010141](https://doi.org/10.1371/journal.pgen.1010141).
- 367 Burt, Austin and Graham Bell (1987). “Mammalian Chiasma Frequencies as a Test of Two Theories of  
 368 Recombination”. In: *Nature* 326.6115, pp. 803–805. doi: [10.1038/326803a0](https://doi.org/10.1038/326803a0).

- 369 Dapper, Amy L. and Bret A. Payseur (2017). "Connecting Theory and Data to Understand Recombination  
 370 Rate Evolution". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1736,  
 371 p. 20160469. DOI: [10.1098/rstb.2016.0469](https://doi.org/10.1098/rstb.2016.0469).
- 372 Dreissig, Steven, Martin Mascher, and Stefan Heckmann (2019). "Variation in Recombination Rate Is Shaped  
 373 by Domestication and Environmental Conditions in Barley". In: *Molecular Biology and Evolution* 36.9.  
 374 Ed. by Michael Purugganan, pp. 2029–2039. DOI: [10.1093/molbev/msz141](https://doi.org/10.1093/molbev/msz141).
- 375 Dumont, Beth L, Karl W Broman, and Bret A Payseur (2009). "Variation in Genomic Recombination Rates  
 376 Among Heterogeneous Stock Mice". In: *Genetics* 182.4, pp. 1345–1349. DOI: [10.1534/genetics.109.105114](https://doi.org/10.1534/genetics.109.105114).
- 377 Felsenstein, Joseph (1974). "The Evolutionary Advantage of Recombination". In: *Genetics* 78.2, pp. 737–756.  
 378 DOI: [10.1093/genetics/78.2.737](https://doi.org/10.1093/genetics/78.2.737).
- 380 Felsenstein, Joseph and Shozo Yokoyama (1976). "The Evolutionary Advantage of Recombination. II. Indi-  
 381 vidual Selection for Recombination". In: *Genetics* 83.4, pp. 845–859.
- 382 Fernandes, Joiselle Blanche et al. (2018). "Unleashing Meiotic Crossovers in Hybrid Plants". In: *Proceedings  
 383 of the National Academy of Sciences* 115.10, pp. 2431–2436. DOI: [10.1073/pnas.1713078114](https://doi.org/10.1073/pnas.1713078114).
- 384 Fuentes, Roven Rommel et al. (2021). "Domestication Shapes Recombination Patterns in Tomato". In:  
 385 *Molecular Biology and Evolution*. Ed. by Michael Purugganan, msab287. DOI: [10.1093/molbev/msab287](https://doi.org/10.1093/molbev/msab287).
- 386 Haenel, Quiterie et al. (2018). "Meta-Analysis of Chromosome-Scale Crossover Rate Variation in Eukaryotes  
 387 and Its Significance to Evolutionary Genomics". In: *Molecular Ecology* 27.11, pp. 2477–2497. DOI: [10.1111/mec.14699](https://doi.org/10.1111/mec.14699).
- 389 Hansson, Bengt et al. (2006). "Comparative Gene Mapping in *Arabidopsis Lyrata* Chromosomes 1 and 2 and  
 390 the Corresponding *A. Thaliana* Chromosome 1: Recombination Rates, Rearrangements and Centromere  
 391 Location". In: *Genetical Research* 87.2, pp. 75–85. DOI: [10.1017/S0016672306008020](https://doi.org/10.1017/S0016672306008020).
- 392 Hartfield, Matthew, Sarah P Otto, and Peter D Keightley (2010). "The Role of Advantageous Mutations  
 393 in Enhancing the Evolution of a Recombination Modifier". In: *Genetics* 184.4, pp. 1153–1164. DOI: [10.1534/genetics.109.112920](https://doi.org/10.1534/genetics.109.112920).
- 395 Henderson, Ian R. and Kirsten Bomblies (2021). "Evolution and Plasticity of Genome-Wide Meiotic Re-  
 396 combination Rates". In: *Annual Review of Genetics* 55.1, annurev-genet-021721-033821. DOI: [10.1146/annurev-genet-021721-033821](https://doi.org/10.1146/annurev-genet-021721-033821).
- 398 Hill, W. G. and Alan Robertson (1966). "The Effect of Linkage on Limits to Artificial Selection". In: *Genetical  
 399 Research* 8.3, pp. 269–294. DOI: [10.1017/S0016672300010156](https://doi.org/10.1017/S0016672300010156).
- 400 Johnston, Susan E et al. (2016). "Conserved Genetic Architecture Underlying Individual Recombination  
 401 Rate Variation in a Wild Population of Soay Sheep (*Ovis Aries*)". In: *Genetics* 203.1, pp. 583–598. DOI:  
 402 [10.1534/genetics.115.185553](https://doi.org/10.1534/genetics.115.185553).

- 403 Kawabe, Akira et al. (2006). "Comparative Gene Mapping in *Arabidopsis Lyrata* Chromosomes 6 and 7 and  
404 *A. Thaliana* Chromosome IV: Evolutionary History, Rearrangements and Local Recombination Rates".  
405 In: *Genetical Research* 88.1, pp. 45–56. DOI: [10.1017/S0016672306008287](https://doi.org/10.1017/S0016672306008287).
- 406 Keightley, Peter D and Sarah P Otto (2006). "Interference among Deleterious Mutations Favours Sex and  
407 Recombination in Finite Populations". In: *Nature* 443.7107, pp. 89–92.
- 408 Koehler, Kara E. et al. (1996). "Recombination and Nondisjunction in Humans and Flies". In: *Human  
409 Molecular Genetics* 5.Supplement\_1, pp. 1495–1504. DOI: [10.1093/hmg/5.Supplement\\_1.1495](https://doi.org/10.1093/hmg/5.Supplement_1.1495).
- 410 Koella, Jacob C. (1993). "Ecological Correlates of Chiasma Frequency and Recombination Index of Plants".  
411 In: *Biological Journal of the Linnean Society* 48.3, pp. 227–238. DOI: [10.1111/j.1095-8312.1993.tb00889.x](https://doi.org/10.1111/j.1095-8312.1993.tb00889.x).
- 412 Kuittinen, Helmi et al. (2004). "Comparing the Linkage Maps of the Close Relatives *Arabidopsis Lyrata* and  
413 *A. Thaliana*". In: *Genetics* 168.3, pp. 1575–1584. DOI: [10.1534/genetics.103.022343](https://doi.org/10.1534/genetics.103.022343).
- 414 Lenormand, Thomas and Sarah P Otto (n.d.). "The Evolution of Recombination in a Heterogeneous Envi-  
415 ronment". In: (), p. 16.
- 416 Orme, David et al. (2018). *Caper: Comparative Analyses of Phylogenetics and Evolution in r*. Manual.
- 417 Otto, Sarah P. (2009). "The Evolutionary Enigma of Sex". In: *The American Naturalist* 174.S1, S1–S14. DOI:  
418 [10.1086/599084](https://doi.org/10.1086/599084).
- 419 Otto, Sarah P. and Nick H. Barton (2001). "Selection for Recombination in Small Populations". In: *Evolution*  
420 55.10, pp. 1921–1931. DOI: [10.1111/j.0014-3820.2001.tb01310.x](https://doi.org/10.1111/j.0014-3820.2001.tb01310.x).
- 421 Otto, Sarah P. and Bret A. Payseur (2019). "Crossover Interference: Shedding Light on the Evolution of  
422 Recombination". In: *Annual Review of Genetics* 53.1, pp. 19–44. DOI: [10.1146/annurev-genet-040119-093957](https://doi.org/10.1146/annurev-genet-040119-093957).
- 423 Peñalba, Joshua V. et al. (2020). "Genome of an Iconic Australian Bird: High-quality Assembly and Linkage  
424 Map of the Superb Fairy-wren (*Malurus Cyaneus*)". In: *Molecular Ecology Resources* 20.2, pp. 560–578.  
425 DOI: [10.1111/1755-0998.13124](https://doi.org/10.1111/1755-0998.13124).
- 426 R Core Team (2022). *R: A Language and Environment for Statistical Computing*. Manual. R Foundation for  
427 Statistical Computing. Vienna, Austria.
- 428 Revell, Liam J. (2024). "phytools 2.0: An Updated R Ecosystem for Phylogenetic Comparative Methods  
429 (and Other Things)." In: *PeerJ* 12, e16505. DOI: [10.7717/peerj.16505](https://doi.org/10.7717/peerj.16505).
- 430 Ritz, Kathryn R., Mohamed A.F. Noor, and Nadia D. Singh (2017). "Variation in Recombination Rate:  
431 Adaptive or Not?" In: *Trends in Genetics* 33.5, pp. 364–374. DOI: [10.1016/j.tig.2017.03.003](https://doi.org/10.1016/j.tig.2017.03.003).
- 432 Ross-Ibarra, J. (2007). "Genome Size and Recombination in Angiosperms: A Second Look". In: *Journal of  
433 Evolutionary Biology* 20.2, pp. 800–806. DOI: [10.1111/j.1420-9101.2006.01275.x](https://doi.org/10.1111/j.1420-9101.2006.01275.x).
- 434 Roze, Denis (2021). "A Simple Expression for the Strength of Selection on Recombination Generated by  
435 Interference among Mutations". In: *Proceedings of the National Academy of Sciences* 118.19, e2022805118.  
436 DOI: [10.1073/pnas.2022805118](https://doi.org/10.1073/pnas.2022805118).

- 439 Roze, Denis and Nick H Barton (2006). “The Hill–Robertson Effect and the Evolution of Recombination”.  
 440 In: *Genetics* 173.3, pp. 1793–1811.
- 441 Roze, Denis and Thomas Lenormand (2005). “Self-Fertilization and the Evolution of Recombination”. In:  
 442 *Genetics* 170.2, pp. 841–857. DOI: [10.1534/genetics.104.036384](https://doi.org/10.1534/genetics.104.036384).
- 443 Samuk, Kieran et al. (2020). “Natural Selection Shapes Variation in Genome-wide Recombination Rate in  
 444 *Drosophila Pseudoobscura*”. In: *Current Biology* 30.8, 1517–1528.e6. DOI: [10.1016/j.cub.2020.03.053](https://doi.org/10.1016/j.cub.2020.03.053).
- 445 Schreiber, Mona et al. (2022). “Recombination Landscape Divergence between Populations Is Marked by  
 446 Larger Low-Recombining Regions in Domesticated Rye”. In: *Molecular Biology and Evolution*. Ed. by  
 447 Michael Purugganan, msac131. DOI: [10.1093/molbev/msac131](https://doi.org/10.1093/molbev/msac131).
- 448 Smith, Stephen A. and Joseph W. Brown (2018). “Constructing a Broadly Inclusive Seed Plant Phylogeny”.  
 449 In: *American Journal of Botany* 105.3, pp. 302–314. DOI: [10.1002/ajb2.1019](https://doi.org/10.1002/ajb2.1019).
- 450 Stapley, Jessica et al. (2017). “Variation in Recombination Frequency and Distribution across Eukary-  
 451 otes: Patterns and Processes”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences*  
 452 372.1736, p. 20160455. DOI: [10.1098/rstb.2016.0455](https://doi.org/10.1098/rstb.2016.0455).
- 453 Stetsenko, Roman and Denis Roze (2022). “The Evolution of Recombination in Self-Fertilizing Organisms”.  
 454 In: *Genetics* 222.1. Ed. by A Agrawal, iyac114. DOI: [10.1093/genetics/iyac114](https://doi.org/10.1093/genetics/iyac114).
- 455 Veller, Carl, Nancy Kleckner, and Martin A. Nowak (2019). “A Rigorous Measure of Genome-Wide Genetic  
 456 Shuffling That Takes into Account Crossover Positions and Mendel’s Second Law”. In: *Proceedings of the*  
 457 *National Academy of Sciences* 116.5, pp. 1659–1668. DOI: [10.1073/pnas.1817482116](https://doi.org/10.1073/pnas.1817482116).
- 458 Wang, Shunxin et al. (2015). “Meiotic Crossover Patterns: Obligatory Crossover, Interference and Home-  
 459 ostasis in a Single Process”. In: *Cell Cycle* 14.3, pp. 305–314. DOI: [10.4161/15384101.2014.991185](https://doi.org/10.4161/15384101.2014.991185).
- 460 Whitehead, Michael R. et al. (2018). “Plant Mating Systems Often Vary Widely Among Populations”. In:  
 461 *Frontiers in Ecology and Evolution* 6, p. 38. DOI: [10.3389/fevo.2018.00038](https://doi.org/10.3389/fevo.2018.00038).
- 462 Wickham, Hadley (2016). *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

463 **Supplementary figures**

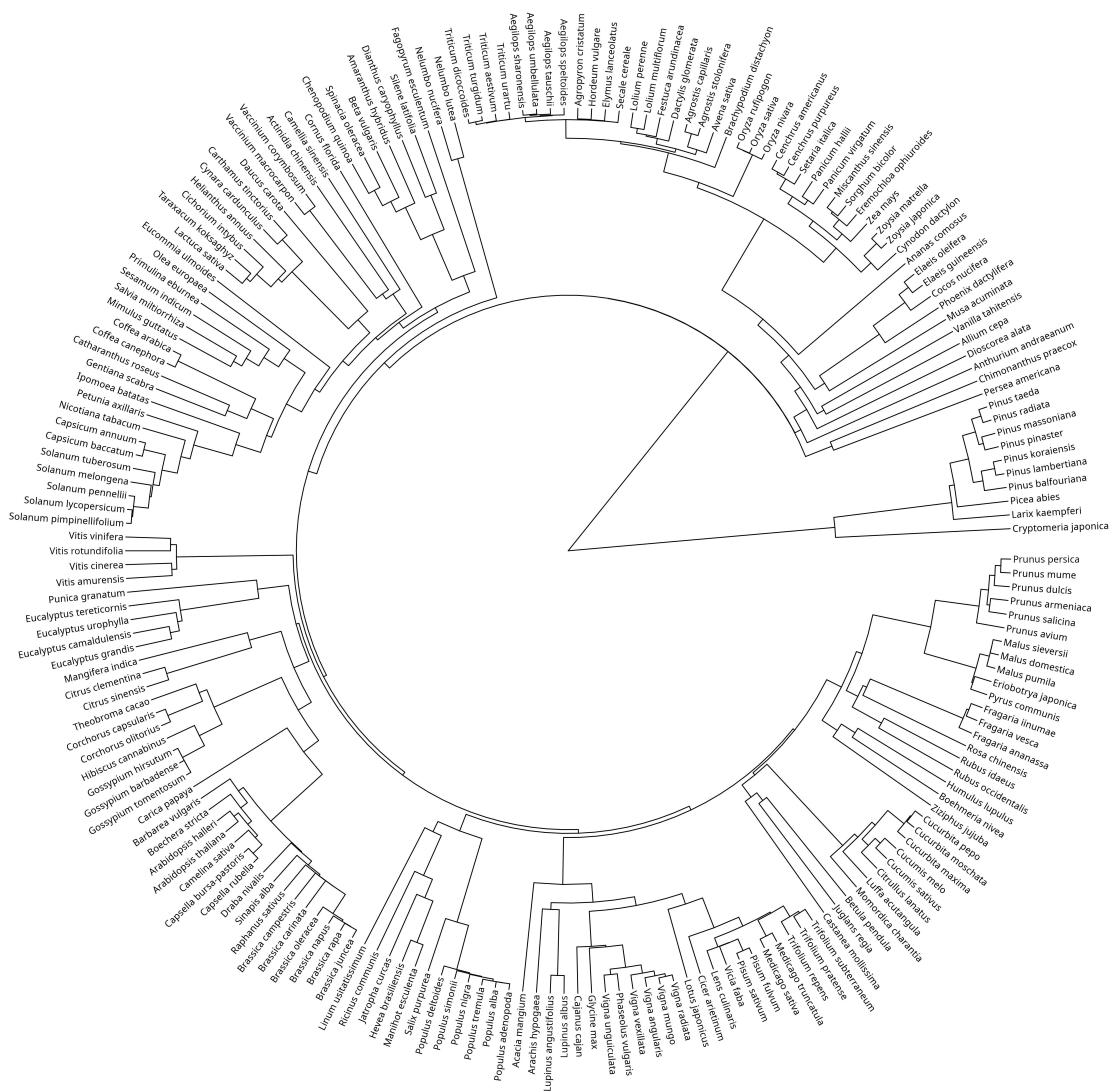


Figure S1: Phylogeny of the complete dataset based on the ultra-metric tree from Smith and Brown (2018).

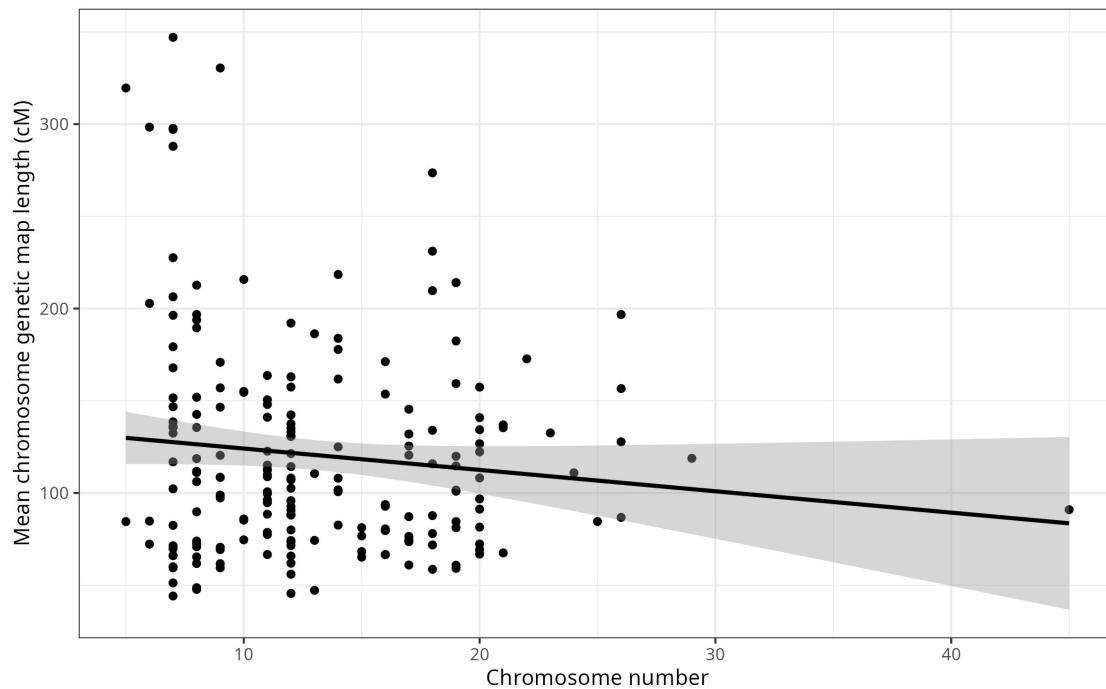


Figure S2: Average chromosome map length is not correlated with the haploid chromosome number (Spearman's  $\rho = -0.08$ ,  $p = 0.27$ ).

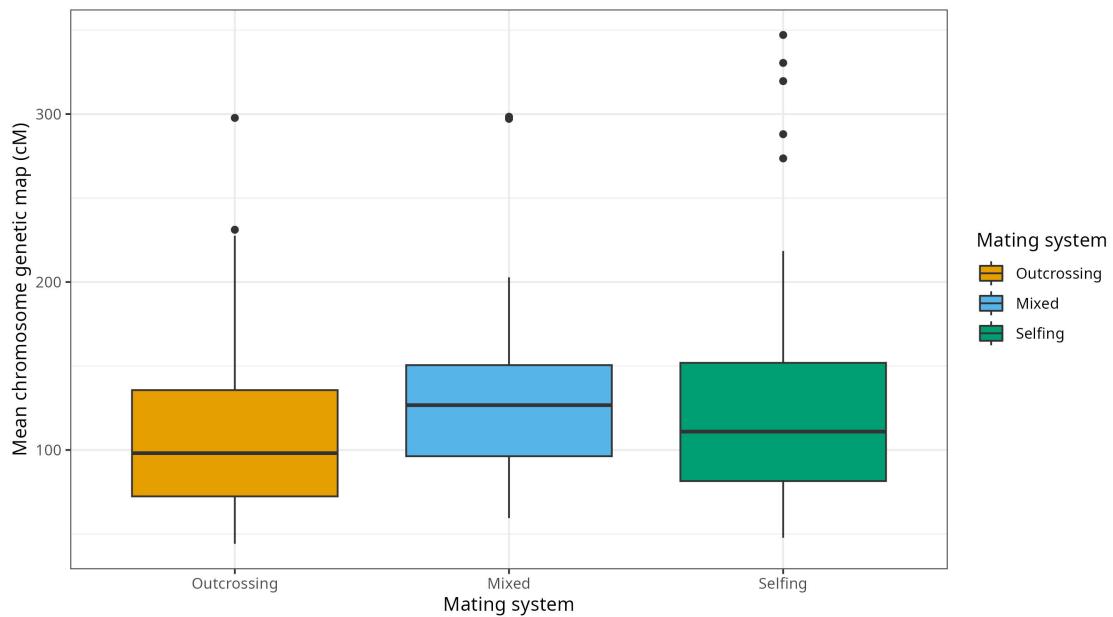


Figure S3: Chromosome map length as a function of the mating system.

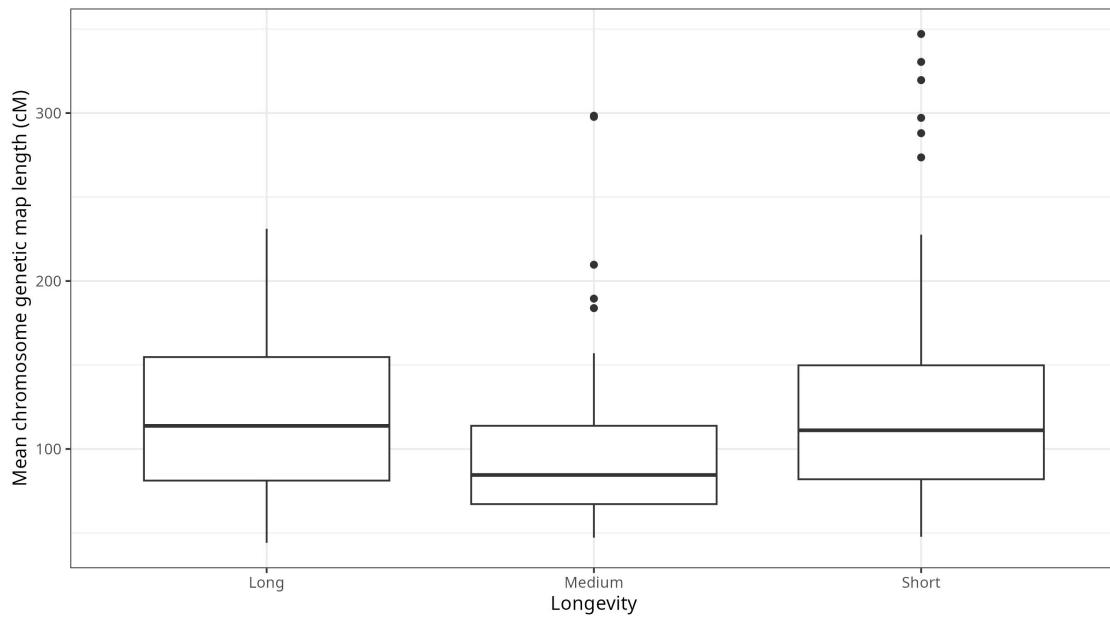


Figure S4: Chromosome map length as a function of longevity.

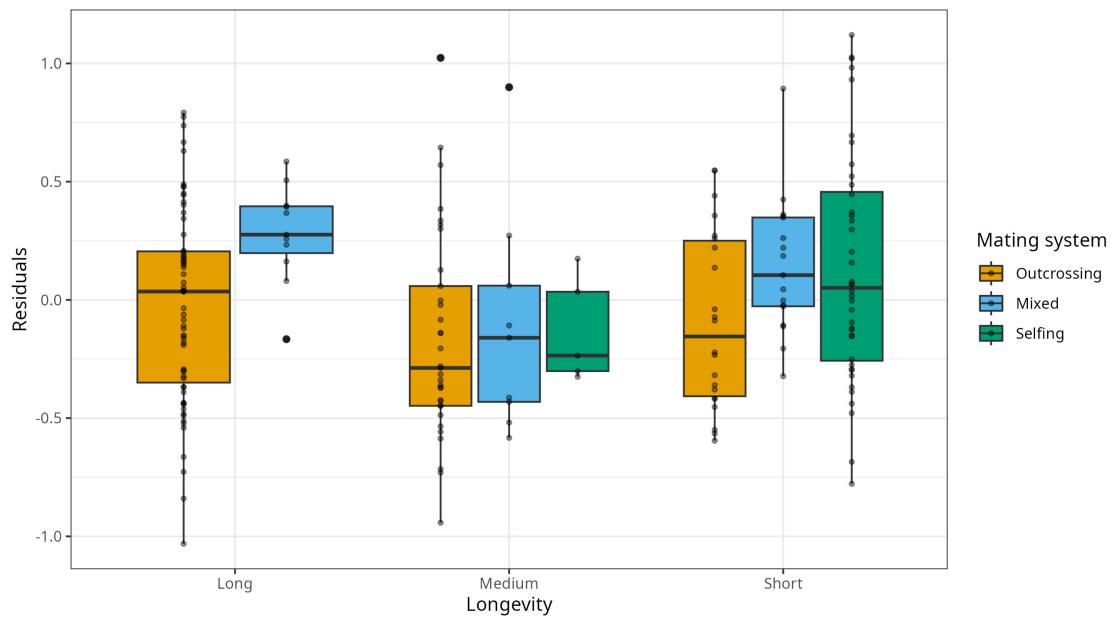


Figure S5: The recombination rates depend on the mating system and longevity. The combined effect of the mating system and longevity on the residuals of the regression recombination rate (cM/Mb) as a function of chromosome size(Mb).

<sup>464</sup> **Supplementary tables**

Table S1: Complete dataset with genetic map length, genome size, genomic characteristics (e.g. ploidy, number of chromosomes) and life history traits, associated with references.

Table S2: Forward model selection steps for the chromosome genetic map length and the residuals of the regression recombination rate as a function of chromosome size. The Anova p-value and AIC/BIC values are provided for each 'lm' and 'pgls' model.

Table S3: Model fit and parameters estimates for the 'pgls' model testing the effect of the mating system, chromosome size, the interaction of the mating system and chromosome size, and the number of chromosomes. Chromosome genetic map length = mating system\*chromosome size + chromosome number.

Table S4: Model fit and parameters estimates for the 'pgls' model testing the effect of marker density and number of progeny. Chromosome genetic map length = mating system + longevity + marker density + number of progeny.