MLH1 Results Outline

null

Table of Contents

[Genome-wide recombination rate estimates for both sexes 2](#_Toc39156110)

[Partitioning variation in recombination rate 8](#_Toc39156111)

[Within mouse variance in CO count per cell 9](#_Toc39156112)

[DSB. Evolution of genome-wide recombination rate is associated with evolution of double strand breaks 10](#_Toc39156113)

[Evolution of genome-wide recombination rate is reflected at the single chromosome level 13](#_Toc39156114)

[Q1. Sex Differences in the Length of the Synaptonemal Complex 14](#_Toc39156115)

[Q1. Sex Differences in the Positions of Single Crossovers 17](#_Toc39156116)

[Q1. Sex Differences in CO Interference (IFD) 18](#_Toc39156117)

[Additional Determinants of Genome-wide Recombination Rate Evolution in Males 19](#_Toc39156118)

[Q2 SC Length 20](#_Toc39156119)

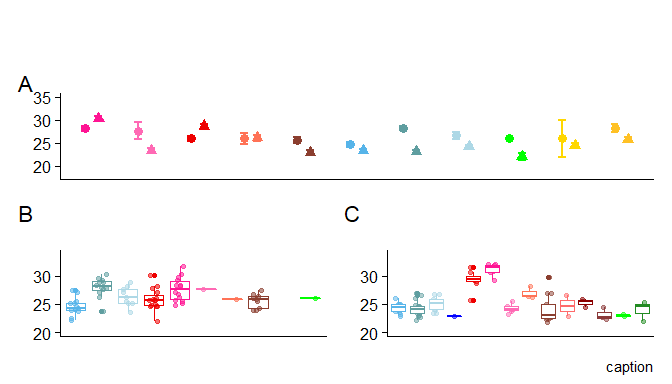
[Q2.1CO rec landscape evolution is decoupled from gwRR evolution 21](#_Toc39156120)

[Q2 Evolution of interference is associated with genome wide recombination rate evolution 22](#_Toc39156121)

[References 23](#_Toc39156122)

**sex differences in Celegan meiosis**

Kelly, William G., Sara M. Fielder, and Rieke Kempfer. “Multiple Sex-Specific Differences in the Regulation of Meiotic Progression in C. elegans.” bioRxiv (2020).



# Genome-wide recombination rate estimates for both sexes

We used counts of MLH1 foci per cell to estimate genome-wide recombination rates in 14 wild-derived inbred strains sampled from three subspecies of house mice: *M. m. domesticus*, *M. m. musculus* and *M. m. molossinus*. Mean MLH1 foci counts for 161 mice were quantified from an average of 21.99 spermatocytes per male (for a total of 1,867 spermatocytes) and 18.54 oocytes per female (for a total of 1,409 oocytes).

Graphical comparisons between the two sexes reveal several patterns (Figure 1A). First, recombination rate differs between females and males in most strains. Second, the direction and magnitude of heterochiasmy varies among strains. Although the majority of strains show higher recombination rates in females (following the pattern in laboratory mice), two musculus strains and one molossinus strain exhibit male-biased heterochiasmy. Hence, relative recombination rates in the two sexes are evolutionarily labile.

The general patterns of heterochiasmy in house mouse are displayed in Figure 1A. Taking note of the direction and the magnitude of the sex differences, our results confirm two general patterns: i) genome-wide recombination rates averages are greater in females compared to males (female biased heterochiasmy) and ii) the degree of heterochiasmy (Female:Male ratio) is generally low, ranging from 1.22 in *domesticusG* to 1.06 in *domesticusLEW*. Three notable exceptions of male biased heterochiasmy are the strains *musculusPWD*, *molossinusMSM* and *musculusSKIVE*, with heterochiasmy values of 0.91, 0.93 and 0.99 respectively.

Separately examining the mouse means of MLH1 foci per cell for each sex points to distinct patterns of variation (Figure 1B-C). Female recombination rates are evenly distributed around the sex-wide mean of approximately 25 MLH1 foci per cell (Figure 1B). In stark contrast, males separate more clearly into two groups of strains with high rates (near 30 MLH1 foci per cell) and low rates (near 23 MLH1 foci per cell) (Figure 1C).

# Partitioning variation in recombination rate

To situate variation in recombination rate within an evolutionary framework, we fit a series of models including subspecies, strain, and sex, to mean MLH1 foci counts from 187 mice. We began with a full mixed model (M1, see Methods), which showed that strain (random effect p < 10^{-6}), sex (p = 1.5510^{-8}), subspecies (p=1.7210^{-4}), and subspecies sex (p = 3.110^{-5}) each significantly affect recombination rate.

After a general linear model including all factors as fixed effects (M2) revealed only weak contributions of subspecies, we focused on additional models designed to illuminate the role of strain and sex. A general linear model with these two variables (M3) identified two strains with particularly strong effects on recombination rate: musculusMSM (p = 3.9910^{-6}) and domesticusG (p = 1.0410^{-6}). In addition, two strains exhibit strain-by-sex interactions: molossinusMSM (p = 1.2610^{-4}) and musculusPWD (p = 3.8610^{-4}).

We next fit general linear models separately for 192 males and 144 females (M4; see methods). In the male dataset, three strains significantly affect recombination rate: musculusPWD ((glm; p = 6.3110^{-8}; effect = 6.11 foci), and musculusSKIVE (glm; p = 0.01; 0; effect = 3.8), and *molossinusMSM* (glm; p=2.4210^{-12}; effect 6.99).

These three strains point to rapid evolution in recombination rate in spermatocytes; we subsequently refer to them collectively as “high-recombination” strains. Analysis of the female dataset points to four strains with significant effects on recombination rate: domesticusG (p = 2.510^{-6}), molossinusMSM (p = 6.2410^{-6}), domesticusLEW (p = 0.01), and musculusPWD (p= 0.02). Strain effect sizes in females are modest in magnitude (ranging from 1 to 4 foci) compared to those in males. Together, these results demonstrate heritable differences in the genome-wide recombination rate evolving in a highly sex-specific manner over short evolutionary timescales.

# Within mouse variance in CO count per cell

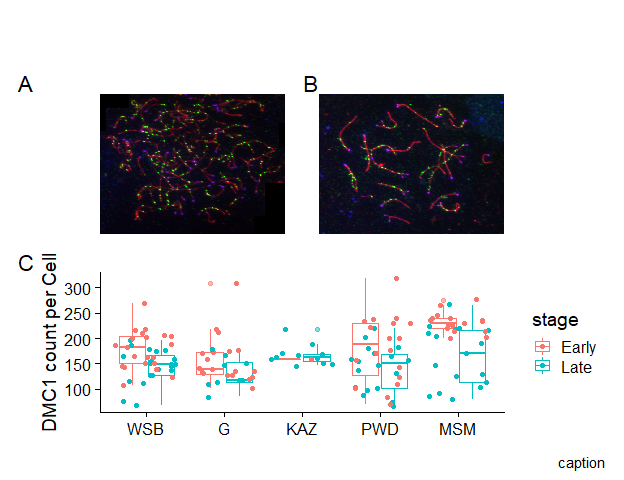
Counting MLH1 foci in multiple oocytes for each female and multiple spermatocytes for each male allowed us to examine determinants of the within-mouse variance in recombination rate. To do this, we considered the same models as above, but replaced mean MLH1 foci count with within-mouse variance in MLH1 foci count per cell as the dependent variable. Sex is the only variable that significantly affects recombination rate in both the mixed model (M1) (p < 10^{-6}) and general linear models (M2) (p = 0.03) and M3 (p = 0.03).

In general, females have almost twice as much variance in MLH1 foci per cell compared to males (Figure 1). Since estimates of within-mouse variance may be more susceptible to technical error from the staining protocol, we repeated the analyses using a subset of cells with higher quality scores (quality score 1 and 2, see Materials and Methods). The results are similar: sex is the strongest effect (M1 p < 10^{-6} <round(LmerMLH1\_M1\_Q12\_VAR\_results$sex.results, 5)>; M2 p = 2.310^{-4}; also M3 p = 2.2810^{-4}). When both quality-curated and full datasets are considered, strain does not significantly and consistently affect variance in MLH1 foci count per cell in either sex. These results suggest that within-mouse variance in recombination rate evolves independently of mean recombination rate.

**(sex specific results - no efficient strain effect for males. In females LEW - but this is due to cell quality difference, what about quality score and cell number within the models?**

**re-read Lenzi et al – heterogeneity in oogenesis**

# DSB. Evolution of genome-wide recombination rate is associated with evolution of double strand breaks



In an attempt to localize the male-specific evolution of crossover number to steps of the meiotic pathway, we counted foci from a marker for double strand breaks (DSBs), DMC1, in prophase spermatocytes. DMC1 foci were scored from a total of 76 early zygotene-stage and 76 late zygotene-stage spermatocytes from three low-recombination strains (musculusKAZ , domesticusWSB , and domesticusG) and two high-recombination strains (musculusPWD and molossinusMSM).

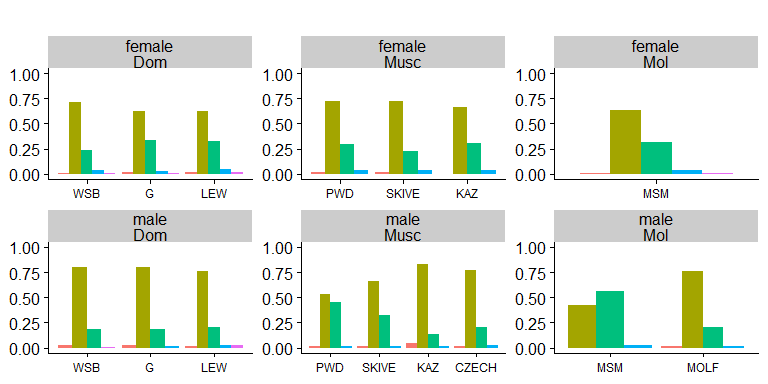
The high-recombination strains have significantly more DMC1 foci than the low-recombination strains in early zygotene cells (t-test, p < 10^{-6} <ttest.HighLow.L.pval>). In contrast, the two strain groups do not differ in DMC1 foci counted in late zygotene cells (t-test, p = 0.66).

After DSB formation, DSBs are repaired as either non-crossovers (NCO) or as crossovers (COs), with the vast majority being repaired as NCOs. Thus the ratio of CO:DSB is a partial indicator of the proportion of DSBs which are designated as COs. The ratios, calculated for DMC1 means from both stages, are not significantly different between the high and low strain groups (t-test, p = 0.94 and p = 0.11 for early zygotene and late zygotene ratios, respectively). (connection to barier)

This comparison raises the possibility that the evolution of crossover number is primarily due to processes that precede the crossover/non-crossover decision. Cole et al. (2012) suggest that early zygotene counts for DMC1 are most relevant for predicting the final CO number due to the crossover homeostasis process.

This result, combined with the large difference in MLH1 count between high and low strains, might indicate that the CO/NCO decision contributes to strain differences in recombination rate.

# Evolution of genome-wide recombination rate is reflected at the single chromosome level



To examine the connection between evolution of the genome-wide recombination rate and changes to the recombination landscape, we used an image analysis pipeline to measure properties of single bivalents ((Peterson, Miller, and Payseur 2019)).

This algorithm substantially speeds the accurate measurement of bivalents, but has the limitation that not all bivalents per cell can be isolated due to overlapping bivalents. In this dataset, isolation rates per cell range from 0.51 (molossinusMSM male) to 0.72 (musculusKAZ female).

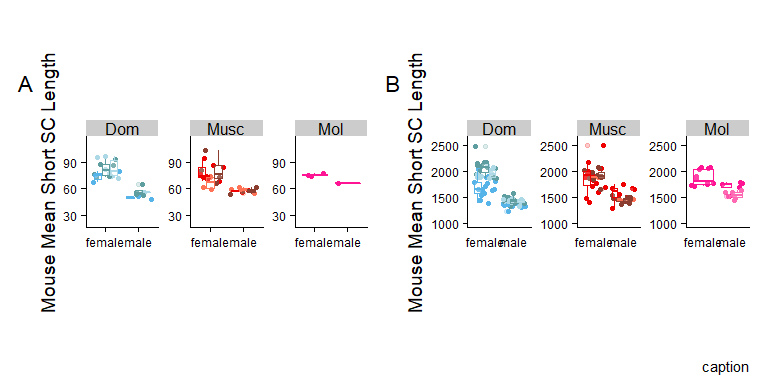
From the total set of cell images, 10,458 bivalent objects were isolated by the image analysis software. After a human curation step ((Peterson, Miller, and Payseur 2019)), 9,829 single-bivalent observations remained. We assume that the isolation of bivalents within cells is unbiased.

Ninety-six percent of single bivalents in our pooled dataset (n = 34,982) have either one or two crossovers (Figure X). The proportion of one-crossover (1CO) to two-crossover (2CO) bivalents distinguishes the high vs. low recombining strains (Figure X). High-recombination strains are enriched for 2CO bivalents at the expense of 1CO bivalents: proportions of 2CO bivalents are 0.33 ( *musculusSKIVE* ) in musculusSKIVE, 0.44 in musculusPWD, and 0.53 in molossinusMSM.

Following patterns in the genome-wide recombination rate, male musculusPWD and male molossinusMSM have 2CO proportions that are more similar to each other than to strains from their own subspecies (chi-square tests; musculusPWD vs. musculusKAZ p = 3.1510^{-33}; molossinusMSM vs. molossinusMOLF p = 4.7210^{-13}).

Next, we use this single-bivalent dataset to focus on aspects of the recombination landscape along chromosomes. We address two main questions. First, which traits are sexually dimorphic? Second, which traits differ between males from high-recombination vs. low-recombination strains?

# Q1. Sex Differences in the Length of the Synaptonemal Complex



In many mammalian species, the synaptonemal complex (SC) is longer in females (Cahoon and Libuda (2019), Sardell and Kirkpatrick (2020)). Yet, the majority of these observations come from species with female-biased heterochiasmy. Our wider survey of recombination in house mice provides an opportunity to determine whether sex differences in chromatin compaction (SC length) are reversed when heterochiasmy is male-biased. In addition, if SC length is a strong determinant of the genome-wide recombination rate, male musculusPWD and male molossinusMSM should have longer SC lengths than the other strains.

We use the reduce data set of short bivalents to remove the effects of the paired X-X bivalent in females. The mouse average of short bivalents are significantly longer in females than males in all strains (t-test; p < 0.05) except musculusSKIVE, which my be attributed to low sample sizes (p = 0.11). The ratios of the mouse mean SC across sexes range across strains from 1.15 (musculusMSM) to 1.49 (domesticusWSB).

Females have significantly longer total SC than males in each strain tested separately (t-test; p < 0.05). That females have longer SCs is also supported by mixed models and general linear models with covariates, which identify sex as the most consistently significant effect (p < 0.05). Additionally, there are some significant subspecies and strain effects (p < 0.05), indicating that SC length has evolved among strains and subspecies, however further investigation is beyond the scope of this paper.

In summary, two approaches for measuring and analyzing SC length indicate that females have longer SCs (chromosome axes), even in strains where males have more MLH1 foci per cell. Furthermore, joint consideration of MLH1 foci and total SC length suggests that males from high-recombination strains have less “space” in which to place their additional crossovers compared to females.

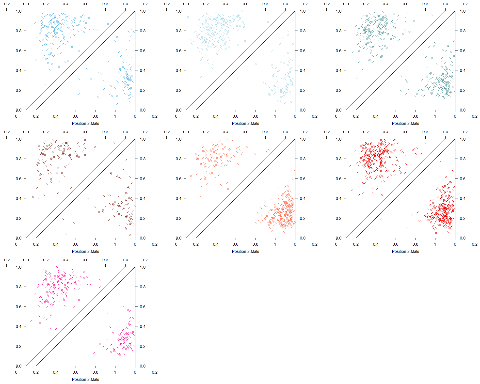
# Q1. Sex Differences in the Positions of Single Crossovers

We used the single crossover bivalent (1CO) normalized position to compared general location of crossovers while controlling for differences in total SC length. Given that, in most of the strains we surveyed, the majority of bivalents are observed to contain one crossover (focus) – thus justifying the use of single crossover bivalents – to discribe this metric / trait.

In all strains, the landscape across 1CO bivalents is significantly different in females and males. Normalized foci positions tend to be more central in females 0.56 and closer to the telomere in males 0.68 (t-test; p = 2.9210^{-22}).

Sex is also the most significant effect on focus position in a mixed model (M1: p = 1.2610^{-25}) and in general linear models (M2: p = 1.3310^{-7}; M3: p = 1.3310^{-7}). These sex differences in the placement of foci on 1CO bivalents follow a pattern observed across a variety of mammalian species (Sardell and Kirkpatrick 2020).

# Q1. Sex Differences in CO Interference (IFD)



Example of Sex differences in IFD distributions

There is no strong signal of sex differences in raw mean inter-focal distances (IFDraw) across the full set of strains. A marginally significant difference between the sexes (t-test; p = 0.07) is driven by one strain, *domesticusG* (t-test without domesticusG; p = 0.02). This result indicates that females and males exhibit a similar level of interference when it is measured in physical (SC) units.

In contrast, males have significantly longer normalized mean inter-focal distances (IFDnorm) than females in seven out of eight strains (t-tests; p < 1.4910^{-12}),

with only musculusKAZ showing no difference (t-test; p = 0.33). Examination of IFDnorm distributions indicates that female IFDnorm values are centered at approximately 50% and show a slight enrichment of low (<25%) values, whereas males are enriched for higher values. Mixed models and general linear models of IFDnorm support the inference of stronger interference in males: sex is the most significant variable (M1 - LRT **random**: p = 6.7410^{-14} glm *M2* 0.01 *M3* 0.01). When interference is measured in physical SC units (IFDraw), the differences between sexes is low and only slightly significant (data not shown). In summary, controlling for differences in SC length (chromatin compaction) using IFDnorm indicates that interference is stronger in males, whereas consideration of IFDraw shows that the sexes exhibit a similar level of interference on the physical (SC) scale.

**address the difference is pattern signal for IFD triangles and boxplots** (make boxplots?)

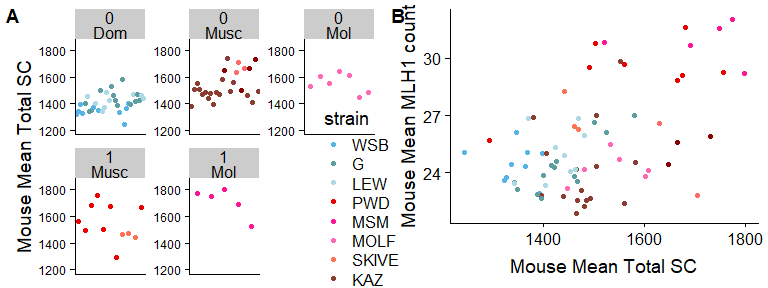
# Additional Determinants of Genome-wide Recombination Rate Evolution in Males

Next, we used the contrast between males in the high-recombination strains (musculusPWD , musculusSKIVE, and molossinusMSM) and low-recombination strains to identify features of the recombination landscape associated with evolutionary transitions in the genome-wide recombination rate. The main objective of this section is to test for significant correlations between features of the recombination landscape and the evolution of mean MLH1 foci per cell. In comparing the male specific single bivalent based metrics the first criteria for analysis is a significant differences between the high and low recombining groups and the second step is testing for significant subspecies and strain effects (M4 and M4.2).

Previous empirical work suggests basic predictions for the relationship between the genome wide recombination rate and two aspects of the recombination landscape. **SC length is expected to be positively associated with genome-wide recombination rate because of the loop-axis structure** **(zickler Kleckner 1999, Merier 2015?).** Crossover interference strength is expected to be negatively associated with genome-wide recombination rate as interference strength is reflected as the space between crossovers on the same chromosome (on chromosomes of a finite length, increasing interference strength can result in ‘pushing’ additional crossovers off of the chromosomes.

Following this logic we predict (1) musculusPWD will have greater SC length and weaker interference than musculusSKIVE, which in turn will have longer SC and weaker interference compared to the other *musculus* strains, (2) molossinusMSM will have longer SC and weaker interference compared to molossinusMOLF, and 3) *domesticus* strains will have similar SC length and crossover interference.

# Q2 SC Length



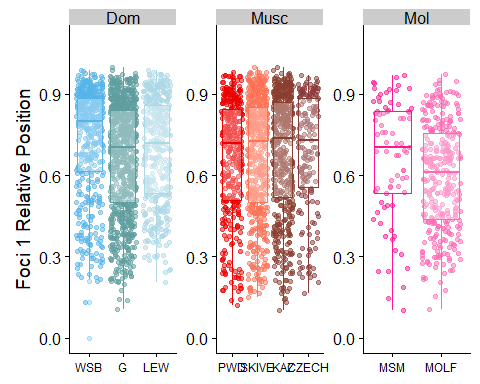
Confirming the basic predictions, there is a positive correlation between mean MLH1 foci per cell and total SC (Spearmans’ r = 0.48; p = 2.2410^{-10}). Nevertheless, mean total SC only partially differentiates high-recombination and low-recombination strains (Figure X). **Bret’s note on positive correlation, ’A correlation across mice, across strains, or both?**

While high-recombination strains have significantly greater mouse mean total SC area in the total dataset (t-test; p = 0.01), separate tests by strain show that only within *molossinus* are high- and low-recombination strains significantly different (t-test; molossinus: p = 0.03; musculus: p = 0.87). **the values above are diferent between editions**

Additionally, the mouse means for the reduced (short and long) bivalent datasets are not significantly different between high-recombination and low-recombination strains (t-test; short: p = 0.88; long: p = 0.18).

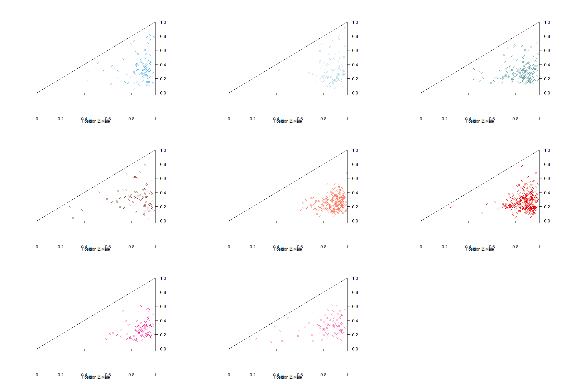
**the values between versions are slightly different** In a general linear model with total SC as the dependent variable, two subspecies effects are significant (p = musculus 1.2410^{-6}, molossinus p = 10^{-6}). In general linear models with reduced bivalent means as dependent variables, several subspecies and strain effects reach significance (p < 0.05) but they are not consistent across models, indicating to some extent the chromatin compaction evolution is decoupled from evolution in mean MLH1 foci per cell. **Bret’s comment; How can you differentiate between this biological conclusion and uncertainty due to the size of the datasets and statistical approaches?**

# Q2.1CO rec landscape evolution is decoupled from gwRR evolution



The normalized 1CO position is not significantly different between high-recombination and low-recombination strains for the total pooled data (t-test; p = 0.24) and also when examined within subspecies (t-test; p = 0.41 and p = 0.07 for *musculus* and *molossinus*, respectively). **While there are significant strain effects for domesticusWSB and molossinusMOLF in a general linear model with normalized position as the dependent variable (Figure X)**, *this evolution of the 1CO positioning is decoupled from the total genome-wide recombination rate.*

# Q2 Evolution of interference is associated with genome wide recombination rate evolution



Mouse averages for both IFDraw and IFDnorm are significantly longer in high-recombination strains (t-test; IFDnorm: p = 7.7410^{-7}; IFDraw: p = 8.7810^{-6}). This pattern is confirmed by separate comparisons within *musculus* (t-test; IFDnorm: p = 2.0410^{-5}; IFDraw: p = 1.9410^{-4}) and within *molossinus* (IFDnorm: p= 0.17; IFDraw: p = 0.08). Similar results are seen with general linear models for both IFDraw and IFDnorm: only effects associated with high-recombination strains are significant (p < 0.05).

That IFDraw and IFDnorm show similar patterns eliminates variation in SC lengths and bivalent sizes as primary explanations. We determined that the main difference in IFDnorm distributions between high-recombination and low-recombination strains is an enrichment of IFDnorm values under 30% in low-recombination strains. The frequency of IFDnorm values that fall below 30% ranges from 8.2% (domesticusG) to 16% (musculusKAZ) in low-recombination strains, whereas high-recombination strains all show such frequencies below 5% (0%, 1.3%, and 3.3% for musculusSKIVE, molossinusMSM, and musculusPWD, respectively).

In summary, the level of interference is a significant predictor of evolution in the genome-wide recombination rate, but SC length and crossover position on 1CO bivalents are not. However, the pattern is in the opposite direction to our prediction: high-recombination strains have stronger interference.

# References

Cahoon, Cori K, and Diana E Libuda. 2019. “Leagues of Their Own: Sexually Dimorphic Features of Meiotic Prophase I.” *Chromosoma*. Springer, 1–16.

Cole, Francesca, Liisa Kauppi, Julian Lange, Ignasi Roig, Raymond Wang, Scott Keeney, and Maria Jasin. 2012. “Homeostatic Control of Recombination Is Implemented Progressively in Mouse Meiosis.” *Nature Cell Biology* 14 (4). Nature Publishing Group: 424–30.

Peterson, April L, Nathan D Miller, and Bret A Payseur. 2019. “Conservation of the Genome-Wide Recombination Rate in White-Footed Mice.” *Heredity* 123 (4). Nature Publishing Group: 442–57.

Sardell, Jason M., and Mark Kirkpatrick. 2020. “Sex Differences in the Recombination Landscape.” *The American Naturalist* 195 (2): 361–79. doi:[10.1086/704943](https://doi.org/10.1086/704943).