Effects of crossover interference on genomic recombination landscape

Mikko Kivikoski[*,1,2], **Pasi Rastas**[†,1], **Ari Löytynoja**[†] **and Juha Merilä**[*,§]

[*]Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, FI-00014 University of Helsinki, Finland, [†]Institute of Biotechnology, HiLIFE, FI-00014 University of Helsinki, Finland, [§]Division of Ecology and Biodiversity. Kadoorie Building, The University of Hong Kong, Pokfulam Road, Hong Kong, SAR

ABSTRACT Crossovers in meiosis increase genetic variability in gamete pool by breaking existing and creating new haplotypes, resulting in novel combinations of alleles. Frequencies of crossovers are known to differ between sexes in many species, but possible sex differences in localization of crossovers have seldom been studied in detail. Here, we address several long-standing questions regarding frequency and distribution of crossovers, with aid of in-depth analyses of ultra-dense sex-specific linkage maps of nine-spined (*Pungitius pungitius*) and three-spined (*Gasterosteus aculeatus*) sticklebacks. We observe that males have typically one crossover per chromosome, independently of the chromosome length, whereas in females the number of crossovers increases with the chromosome length, being typically more than two. We show that these differences largely explain more uniform crossover distribution of females and the fact that females have crossovers closer to the centromere. Furthermore, highly constrained crossover localization reveals that the relationship between linkage and genetic distance varies inside the chromosome. Our analyses show that recombination rate is highly heritable in females ($h^2 = 0.56 - 0.65$), and thus, able to evolve in response to selection. Similar results for the two species that diverged 26 million years ago suggest that our findings are general, at least in the context of stickleback fishes.

KEYWORDS Recombination; Heterochiasmy; Crossover interference; Mapping function; Heritability; *Pungitius pungitius*; Stickleback

Introduction

Increase of genetic variability through recombination is considered one of the key reasons for the ubiquity of sexual reproduction (Otto and Lenormand 2002). Crossovers increase recombination by breaking physical linkage between loci and generating novel combinations of alleles. Depending on the phenotypic effects of the loci and the underlying epistatic interactions, novel haplotypes may result in novel combinations of traits. Moreover, by breaking linkage between deleterious and beneficial loci, recombination facilitates removal of deleterious alleles from populations (Hill and Robertson 1966; Felsenstein 1974). On the other hand, recombination may disrupt beneficial allele combinations, making the presence or absence of crossovers fundamentally important for organisms' fitness and adaptation.

Localization of crossovers in chromosome is not random. The non-random localization of crossovers can be caused by crossover hotspots (Kauppi *et al.* 2004; Baudat *et al.* 2010), chromosomal inversions (Thompson and Jiggins 2014), sex-determining regions (Charlesworth 2017), repelling effects of centromere and telomere (Talbert and Henikoff 2010; Haenel *et al.* 2018; Sardell *et al.* 2018) or crossover interference. Crossover interference, sometimes called positive interference, refers to the deviation of the inter-crossover distances from a neutral expectation (Colombo and Jones 1997; Broman and Weber 2000; Otto and Payseur 2019). In other words: crossovers in the same bivalent occur farther from each other than would be expected by chance. Crossover interference was originally described by Sturtevant (1915) (see also Muller 1916), and elegantly demonstrated by Charles (1938). Although crossover interference was described more than a hundred years ago, its mechanistic underpinnings have remained elusive (Zickler and Kleckner 2016).

Crossovers are known to be suppressed near centromeres

Manuscript compiled: Monday 14th December, 2020

[1]These authors contributed equally to this work.

[2]Corresponding author: mikko.kivikoski@helsinki.fi, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, FI-00014 University of Helsinki, Finland

(Talbert and Henikoff 2010; Haenel *et al.* 2018), but centromere's effect on crossover localization and its relative importance in comparison to crossover interference has not been studied in detail. One explanation for this may be coincidental: the early work on crossover interference was done with X chromosome of *Drosophila melanogaster* (Muller 1916; Weinstein 1936; Charles 1938) in which the centromere is close to the chromosome end and does not produce evident break in crossover distribution (Fiston-Lavier *et al.* 2010). The same is seen in a more recent model organism, mouse, whose autosomes are all telocentric (Chmátal *et al.* 2014). Sequence-based studies are further complicated by the repetitive nature of centromeres and difficulties in determining their genetic locations in genome assemblies (Sardell *et al.* 2018; Varadharajan *et al.* 2019; Kivikoski *et al.* 2020).

In many species, sexes differ in crossover frequency (i.e. the number of crossovers in meiosis), a phenomenon referred to as heterochiasmy. These differences vary from achiasmy (no crossovers) in one sex to only slight or no difference in crossover frequency between sexes (Burt *et al.* 1991; see also Lenormand *et al.* 2016 for review). Several attempts have been made to identify characteristics that would predict the direction or magnitude of heterochiasmy. According to Haldane-Huxley rule, in species with one achiasmatic sex, it should always be the heterogametic sex (Burt *et al.* 1991). Based on an extensive comparative study, Lenormand and Dutheil (2005) concluded that haploid selection is likely the best explanation for heterochiasmy in plants. Mank (2009) reported that, across eutherian mammals, heterochiasmy correlates negatively with the degree of sexual dimorphism. In addition to sex differences in crossover frequencies, sexes are known to differ in spatial distribution of crossovers. In many species, paternal crossovers have been reported to occur more distally on chromosomes than maternal ones (e.g. Johnston *et al.* 2016, Sardell *et al.* 2018, Lloyd and Jenczewski 2019). However, it is currently unknown whether this or the reasons for heterochiasmy are generalizable over a broader range of taxa.

Crossover rates and quantification of crossover interference in different taxa and sexes have been in focus of active research over the past decades (reviewed by Otto and Payseur 2019) and more recently, the regional variation in crossover rate within chromosome has been studied in different species (Haenel *et al.* 2018). As many aspects of crossovers remain poorly described and general conclusions have been hard to come by, more studies performed on wild populations have been called forth (Otto and Payseur 2019). For instance, most estimates of crossover rate heritability come from laboratory and domesticated populations (e.g. Hadad *et al.* 1996 [*Zea mays*], Dumont *et al.* 2009 [mouse], Baudat *et al.* 2010 [human], Hunter *et al.* 2016 [*D. melanogaster*], Kadri *et al.* 2016 [cattle], but see: Johnston *et al.* 2016 [soay sheep] and Johnston *et al.* 2018 [red deer]) which are thought to be subject to recent directional selection for higher crossover rate, although evidence for this is controversial (Muñoz-Fuentes *et al.* 2015).

Crossovers have been studied by counting occurrences of novel phenotypes among offspring (Sturtevant 1915), by visual inspection of chiasmata (Hultén 1974) and by studying novel allele combinations with linkage maps (see Cann 1992 for a summary). Although high-throughput sequencing produces data that can be used to reconstruct linkage maps of unprecedented resolution, computational methods are still being developed to deal with high marker densities and noise in the resequencing data. As demonstrated by Rastas (2017), failure to deal with the high marker density leads to gross over estimation of genetic map lengths and crossover frequencies. In addition to development in linkage map reconstruction, linked read sequencing of gametes has been successfully used to detect the crossovers (Dréau *et al.* 2019).

Particularly important property of linkage maps is that they only contain information of half of the crossovers that occurred in meiosis. This stems from the fact that every chiasmata (and crossover, therefore) occurs between two chromatids, and that the homolog in a gamete is sampled from the four homologs (Weinstein 1936). Based on this random process, the number of observed crossovers in a randomly sampled meiotic product can be expressed with binomial coefficients (Goldcar and Fain 1988). This means that the crossover data derived from the linkage map is an outcome of a two-step random process, where the first step is the number of crossovers in the meiosis and the second step is the sampling of the meiotic product. Therefore, meiotic products with no traces of crossovers do not imply that there had not been any. Likewise, the fact that there are, on average, 0.5 crossovers per meiotic product which corresponds to map length of 50cM does not imply that there would be an obligate crossover as the average of 0.5 can be achieved with numerous combinations of crossover numbers. For instance, when estimating heritability of crossover rate, it is important to notice that the actual data (observed crossovers) may have different variance than the studied trait (crossovers in meiosis).

Crossover rate (i.e. frequency of crossovers in different parts of chromosome) is typically estimated by dividing genetic distance (centimorgan, cM) between two positions (typically SNP markers) by their genomic distance (base pair, bp) (see Kong *et al.* 2010 or Haenel *et al.* 2018 for example). However, due to additive nature of genetic distance, the relationship between the recombination and genetic distance is not constant (Liberman and Karlin 1984), which may bias these estimates. The genetic distance of the positions is obtained from a linkage map that is usually based on the entire data of all maternal and paternal homologs. Therefore, the resulting crossover distribution (i.e. variation of crossover rate along the chromosome) is derived from meiotic products which both vary in number of crossovers and originated in meioses with varying number of crossovers (e.g. 1, 2, 3...). When crossover rate is inferred from the entire data, the effect of crossover frequency (i.e. the number of crossovers in a homolog) on the crossover distribution cannot be assessed. Moreover, due to crossover interference, the distribution of crossovers depends on the crossover count in the entire bivalent, not only the number of crossovers in the particular meiotic product. Consequently, distribution of crossovers in meiotic products with given number of crossovers is a mixture of crossover distributions characteristic to meiosis with the given number or more crossovers in the bivalent.

The aim of this study was to investigate genomic distribution (i.e. localization) of crossovers in detail and in particular, how sex, centromere and crossover frequency (i.e. number of crossovers in meiosis) are related to it. Furthermore, long-standing questions regarding crossover interference, obligate crossover and heritability of crossover frequency were addressed. In order to generalize the findings, the studies were carried out with two species of wild origin stickleback fishes, the nine-spined (*Pungitius pungitius*) and the three-spined (*Gasterosteus aculeatus*) stickleback. For this, we used ultra-high-density sex-specific linkage maps comprised of 938 F_1-generation individuals from 87 nine-spined stickleback families and a comparable data set of 517 F_1-generation individuals from 60 three-

spined stickleback families (Kivikoski *et al.* 2020). With aid of the detailed linkage maps and improved version of the nine-spined stickleback reference genome (Kivikoski *et al.* 2020), we were able to study crossover rate throughout the genome by counting the crossovers in different parts of chromosomes instead of estimating the rate in units of centimorgan per base pairs from the entire data set (cf. Johnston *et al.* 2017, Haenel *et al.* 2018) and shed light on the role of centromeres in shaping the crossover distribution. With the large and precise data set, we were able to infer the frequency of crossovers that occurred in meiosis, but were not seen in the data. Finally, this approach allowed us to disentangle the effects of centromere and crossover frequency on the crossover distribution. Overall, our results highlight the potential of high-quality reference genomes and linkage maps for advancing our understanding of the biology of recombination in the wild.

Materials and Methods

Linkage maps

The crossover frequencies and localizations were estimated from the linkage maps described by Kivikoski *et al.* (2020). For the nine-spined stickleback, ultra-dense (22,468 informative markers) linkage maps were reconstructed with Lep-Map3 software (Rastas 2017) from a data set of 938 F_1-offspring from 87 families (41 females, each crossed with two males, and 5 females, each crossed with one male) as described in Kivikoski *et al.* (2020). The parental fish were wild caught individuals from Helsinki (60°13′N, 25°11′E), Finland and artificially crossed in laboratory facilities (Kivikoski *et al.* 2020, see also Rastas *et al.* 2016) . For the three-spined stickleback, linkage maps were based on 517 F_1-offspring from 60 families (30 males, each crossed with two females) and they were reconstructed with Lep-Map3 as well (see Kivikoski *et al.* 2020). The parents were wild caught from the Baltic Sea and the crossing scheme and sequencing data for them are described in detail in Pritchard *et al.* (2017).

Inference of crossover frequency

We used an expectation-maximization algorithm described by Yu and Feingold (2001) to estimate the probabilities for different number of crossovers in bivalents. The motivation of the algorithm is based on Weinstein's (1936) studies that, on average, half of the crossovers are not observed in a randomly sampled gamete. The algorithm was applied separately for both sexes and for each chromosome. We tested the obligate crossover hypothesis ($H_0 : p_0 = 0$, where p_0 is probability of no crossovers in bivalent) with likelihood ratio test described by Yu and Feingold (2001). We obtained the p-value for the test following their bootstrapping procedure with 1000 replicates and calculated the 95% confidence intervals for the maximum-likelihood estimate of p_0. Expectation-maximization algorithm, as well as the hypothesis testing procedure, were implemented and executed in R (ver. 3.4.4 R Core Team 2018).

Assessment of crossover interference

Crossover interference refers to a non-random localization of crossovers in respect to each other when there are multiple crossovers in the bivalent. Crossover interference can be studied with either mechanistic or statistical models (Otto and Payseur 2019). To quantify crossover interference in our data, we applied the commonly used gamma model (Broman and Weber 2000) and its derivative gamma-sprinkling model (Housworth and Stahl 2003; Otto and Payseur 2019). Of these two models, the gamma-sprinkling model has been suggested to provide a better fit to actual data (Otto and Payseur 2019).

We fitted the models separately for both sexes and for each chromosome using simStahl function of R-package xoi (ver. 0.68-3, Broman 2019). In our analyses, we assumed an obligate chiasmata (i.e. one crossover per bivalent) in every chromosome and maximum number of crossovers (max.conv parameter) of 2-9 depending on the highest possible crossover count according to maximum-likelihood estimates (see Inference of crossover frequency above; see also Fig. 1 and Supplementary Tables S1–S4 for numerical values). Model fit was assessed with a similar approach as in Broman and Weber (2000): 10 000 meiotic products were generated with simStahl function by setting the interference parameter (nu) and proportion of crossovers with no interference (p) to the maximum-likelihood estimates obtained from the fitted model. A chi-square test was used to test goodness of fit in terms of crossover counts in meiotic products (Broman and Weber 2000; Supplementary Tables S5–S9).

We assessed the effects of crossover interference across the centromeres by comparing the distance to centromere from the two crossovers closest to the given centromere in different chromosome arms (Colombo and Jones 1997; Broman and Weber 2000). Negative correlation of the distances would suggest that the crossover interference acts across the centromere. In testing this correlation, we followed the procedure described in Broman and Weber (2000).

To better understand the probability distribution underlying the observed crossover counts, we tested three simple models in comparison to the one inferred based on the method by Yu and Feingold (2001). The models were based on Poisson distribution, zero-truncated Poisson distribution and geometric distribution. In all cases the number of crossovers in meiosis was sampled from the given probability distribution with a parameter estimated from the empirical data. In the case of Poisson and geometric distributions, we added one to the sampled crossover number to fulfill the obligate crossover criterion. Then, the number of observed crossovers was sampled from the crossover count according to the binomial factors. The model fit was tested again with the goodness of fit test.

Characterization of crossover distributions

We defined the sex-specific crossover rate within certain interval (i.e. region of a chromosome) as the proportion of crossovers that occur in the given interval. This approach allows us to assess crossover rate in different subsets of data (e.g. offspring with two maternal crossovers or one paternal crossover) and study the effects of crossover interference on crossover distribution. To account for the fact that all crossovers in bivalent, not only those observed in the meiotic product, affect the crossover distribution, we defined and implemented an algorithm to decompose the mixtures of crossover distributions and to determine the characteristic distributions for each crossover class. The algorithm was implemented in R and is provided in the Supplementary material (Supplementary methods 1).

Decomposed crossover distributions were used to assess differences in male and female crossover distributions. Our hypothesis was that the between-sex differences in the crossover distributions are caused by the differences in crossover frequencies (i.e. number of crossovers in meiosis). To test this, we generated synthetic crossover data with crossover frequency of one sex and crossover distributions decomposed from the other sex. Thus,

to generate synthetic maternal crossover data, we assumed the crossover counts of paternal crossovers and crossover distributions decomposed from maternal data. This data set mimics the crossover distribution in a case where females had the same crossover frequencies as males. Synthetic paternal crossover data were generated similarly assuming maternal crossover frequencies. We evaluated these synthetic distributions against the original distributions by comparing the change in the expected value of the distribution and variance. To be meaningful, this analysis requires that the sexes have enough similarity in crossover frequencies in the chromosome in focus. Essentially, the distribution of crossovers in meioses with different crossover count should be interpreted in both sexes. If, for instance females, have always two or more crossovers in the chromosome and males have only one, it is not possible to interpret how two crossovers would be located in males or, vice versa, how one crossover would be located in females. Due to very different crossover counts for females and males (see Results), the approach was applicable to only two chromosomes of both species: chromosomes 7 and 10, in the nine-spined stickleback and chromosomes 8 and 12 in the three-spined stickleback.

In addition to simulation of crossovers, we also tested the hypothesis of unequal crossover frequencies explaining the observed differences in crossover distributions with a beta regression mixed model. The model was fitted with glmmTMB R-package (ver. 0.2.3, Brooks *et al.* 2017). We restricted the data set to the crossovers in the long chromosome arms and used the relative distance of the centromere from the nearest crossover as a response variable. Crossover count in the long arm was used as a fixed effect and offspring identity and chromosome were used as random effects to account for non-independence of the observations.

Heritability of crossover frequency

In order to study individual variation in crossover frequency (i.e. the number of crossovers in meiosis), we estimated the heritability of crossover frequency by using generalized linear mixed models (GLMMs). More specifically, we used an animal model inferred with Markov Chain Monte Carlo (MCMC) implemented in R-package MCMCglmm (ver. 2.29, Hadfield 2010) following Wilson *et al.* (2010). For every offspring, we used the total number of observed crossovers across the whole genome as a response variable. We estimated the heritability separately for paternal and maternal crossovers and in a similar manner for the nine and the three-spined stickleback (Table 1). For the nine-spined stickleback we excluded four single-offspring families from the analyses, because the crossovers could not be studied for in single-offspring families. This was done to ensure that families with no variation and with very deviating crossover count would not bias the heritability estimates upwards (see Supplementary methods 2 for discussion).

Crossover data are count data and could be assumed to be Poisson distributed. However, estimation, and especially interpretation, of heritabilities of Poisson distributed traits (or residuals) with generalized linear mixed models is challenging (de Villemereuil *et al.* 2016). Therefore, we estimated the heritabilities for maternal and paternal crossovers in two different ways. First, assuming the residuals to be Poisson distributed, and second, assuming that the residuals are normally distributed. The estimates derived under the assumption of Poisson distribution are reported both in latent and observed scales. We obtained the estimates in observed scale using function QGparams of the QGglmm R-package (ver. 0.7.2; de Villemereuil *et al.* 2016).

In order to assess the reliability of the estimates, we generated simulated data sets with only stochastic variation between individuals which should yield heritability estimate close to zero. We used weakly informative prior in all analyses following the example in the revised Supplementary File 5: Tutorial for MCMCglmm version of Wilson *et al.* (2010).

Data Availability

The linkage maps and computer code for replication of the analyses of this study are available in Github https://github.com/mikkokivikoski/sticklebackRecombination. In order to facilitate reproducibility of the analyses, all steps of the analyses following the linkage map reconstruction step are integrated into an Anduril2 workflow (Cervera *et al.* 2019).

Results

Females and males have different number of crossovers in meiosis, yet one obligate crossover

Like in many other taxa, in sticklebacks sexes differ in crossover frequencies: in both species females have more crossovers than males. Males typically have one crossover per bivalent in meiosis (corresponding to zero or one observed crossover) whereas females typically have two or more crossovers. In general, the number of observed crossovers are similar in the two species (Fig. 1, see supplementary Tables S1–S4 for numerical values). In both species, the number of crossovers increases as a function of chromosome length in females, whereas in males that increase is weak albeit statistically significant (Fig. 2). Number of crossovers per base pair decreases as a function of chromosome length in males, but not, or very little, in females (Fig. 2). Together, these results indicate that in sticklebacks the absolute number of crossovers per bivalent is highly constrained in males being typically one (Fig. 1), whereas in females the number of crossovers per base pair appears to be fairly constant in all 21 chromosomes (Fig. 2). We found strong evidence for the obligate crossover. The maximum-likelihood estimates for no crossovers in meiosis are zero or close to zero for both sexes and based on the likelihood ratio test there is no evidence for rejecting the null-hypothesis of at least one crossover in every meiosis (Supplementary Tables S1–S4). The only exception to this is chromosome 15 in male three-spined sticklebacks ($p<0.001$).

Assessment of crossover interference

Visual inspection of crossover distributions shows that crossover interference is evident in both sexes and in both species (Fig. 3A, see also Supplementary Fig. S1). Observation was confirmed in quantitative analyses, where gamma-sprinkling model provided a better fit to the data than the gamma model in all chromosomes, in both sexes and in both species (Supplementary Fig. S2, Supplementary Tables S5–S8). In the nine-spined sticklebacks the log-likelihood ratios between gamma and gamma-sprinkling model varied from 30 to 150 in females, chromosome 5 (LG5) being an exception with the ratio of 1.8, and from 70 to 190 in males. In the three-spined stickleback the log-likelihood ratios between gamma and gamma-sprinkling models varied from 4.8 to 85 in females, and from 31 to 94 in males (chromosome 19, chrXIX, being an exception with log-likelihood of 9.7).

The maximum-likelihood estimates of v (strength of crossover interference) were higher for paternal crossovers than for maternal crossovers (Supplementary Fig. S3). However, in many

nine and most three-spined stickleback chromosomes, the paternal estimates are not meaningful either because the estimates increase infinitely (linkage groups 3, 11 and 18 in the nine and 2, 3, 5, 6, 9, 10, 11, 13, 16, 17, 18, 20, 21 in the three-spined stickleback) or because the standard error is large indicating poor fit of the model (Supplementary Fig. S3). In general, the gamma-sprinkling model may not be applicable for our paternal data, where only a small proportion of meiotic products have multiple crossovers. However, this does not mean that crossover interference would not be detectable in males as well (cf. Fig. 3A). We assessed the model fit with a goodness of fit test. For both gamma and gamma-sprinkling models, the generated data showed poor fit with the empirical data across multiple chromosomes ($p < 0.05$, see Supplementary Fig. S2 and Supplementary Tables S5–S8). In contrast, a random process based on the probability mass functions estimated according to Yu and Feingold (2001) is in line with the empirical data ($p \geq 0.05$, Supplementary Fig. S2).

Interference causes crossovers to occur distantly to one another – in different chromosome arms for instance (Fig. 3A). However, the location of crossovers, once distant from each other, seems to be independent. For instance, we did not find correlation between distances from the centromere to pairs of crossovers in different chromosome arms, indicating that crossover interference does not act across the centromere in either species (Supplementary Fig. S4).

Sex-dependent crossover frequency leads to differences in crossover localization

We studied the previously reported pattern of lower crossover rate in the chromosome centers (Haenel *et al.* 2018). For instance, chromosome 1 of both species shows that when there are three crossovers in the meiotic product, the one in the middle occurs at the chromosome centre (Fig. 3A). In general, females and males had virtually no crossovers in the chromosome centers in metacentric chromosomes (Fig. 3B). In submetacentric chromosomes, the crossover rate was lower in the chromosome centers in males, but in females this trend was not as clear (Fig. 3B). In subtelocentric, acrocentric and telocentric chromosomes, paternal crossover distribution was skewed towards the end of the long arm of the chromosome, whereas in females the distribution was bimodal with slight decrease in the chromosome centers (Fig. 3B). In general, maternal crossover rate was slightly lower in chromosome centers but this trend varied across chromosomes and depended on crossover frequency (Fig. 3B). In maternal homologs with two crossovers in the long arm, no decrease in crossover rate was seen towards the center in submetacentric chromosomes (Fig. 3C). This can be explained by counterbalancing forces of centromere, telomere and crossover interference. These factors result in two crossovers being distantly located to each other while also maximizing their distances to the closest end of the arm, that being either the centromere or the telomere (Fig. 3C). The same trend is also seen in subtelocentric, acrocentric and telocentric chromosomes, where the outcome is low crossover rate in the centers of the chromosome arm that, for these chromosomes, are close the chromosome center as well (Fig. 3A-C). A by-product of multiple crossovers within a chromosome arm repelling each other is placement of crossovers in the vicinity of the centromere. As the crossover rates differ between sexes, crossovers occur closer to the centromere in females than in males (Supplementary Fig. S5). These patterns were similar in both stickleback species.

Effects of crossover count on crossover distribution raises the question whether a higher number of crossovers can explain observed differences in crossover distributions. Using the decomposed maternal and paternal crossover distributions (see Materials and Methods) for chromosomes 7 and 10 (nine-spined stickleback) and chromosomes 8 and 12 (three-spined stickleback), we tested if the sex differences in our crossover distributions can be explained by the difference in crossover frequencies. For this, we generated maternal and paternal crossover data assuming crossover frequency of the opposing sex. In all cases, the crossover distribution of the generated data moved to the direction of the empirical distribution of the opposing sex (Fig. 4). This was supported by similarity of the variance and the expected value of generated and empirical distributions (Fig. 4). Large differences in crossover frequencies induce uncertainty in decomposed crossover distributions and we cannot resolve whether the crossover frequency alone could explain the observed differences. However, our results strongly suggest that the spatial crossover distributions of the two sexes would be more similar with equal crossover frequencies.

The distribution of crossovers in a single meiotic product is expected to be shaped by all crossovers that occur in the bivalent (Mather 1938). Therefore, the localization of crossovers in two homologs with the same number of crossovers may differ greatly from each other. One example of this is in the nine-spined stickleback maternal homologs with one crossover in chromosome 1 (LG1) where the distribution of the single crossovers is fairly uniform (Fig. 3A). However, according to the maximum likelihood estimate of maternal crossover frequency in meiosis, there is always more than one crossover in the chromosome 1 (Fig. 1) and therefore, all the cases of one observed crossover descend from multi-crossover bivalents.

The results of the beta regression model show that in both sexes and in both species, the distance of crossovers from the centromere decreases as a function of crossover count ($p < 0.001$; Fig. 5). In the nine-spined stickleback the slope differs between sexes ($p < 0.001$), whereas in the three-spined stickleback, the difference is not significant ($p < 0.07$). Taken together, the results suggest that part of the sex differences in crossover distribution, here measured as crossover distance from centromere, are explained by differences in crossover frequencies, which in turn influences the prevalence of crossover interference. However, part of the differences remain to be explained by other factors.

Sex-specific heritability of crossover frequency

We obtained very different heritability estimates when assuming gaussian or Poisson distribution (see Discussion). Assuming normal distribution, the number of observed crossovers was found to be heritable in nine-spined stickleback females ($h^2 = 0.649$) and three-spined stickleback females and males ($h^2 = 0.558$ and $h^2 = 0.267$, respectively; Table 1). The heritability estimates indicate that in all cases, except in the nine-spined stickleback males ($h^2 = 0.086$), there is additive genetic variation in crossover frequency. To get further evidence for this and to take into account the stochasticity in the observed crossover data, we conducted a simulation assay. We generated data sets of maternal and paternal crossovers by assuming that all parents (of the same sex and species) have the same inferred crossover frequency (see Fig. 1) and similar family sizes as in our empirical data. We estimated the between-parent variance of the mean crossover count per offspring, and compared it to the empirical variance components. The empirical distribution of average

crossover count had greater variance than the 100 simulated data sets had on average, except for the nine-spined stickleback males (Supplementary Fig. S6). This confirms that there is additive genetic variance in crossover frequencies among female nine-spined sticklebacks and three-spined stickleback females and males, whereas little little or not at all among the nine-spined stickleback males.

Discussion

Crossovers are an important part of genetic recombination. Although crossovers have been studied for more than a century, several questions regarding their frequency and distribution in genomes have remained unsolved. As patterns generalizable across taxa have been proven difficult to formulate, in-depth studies of crossovers in different taxa are needed. Here we fill gaps in knowledge by reporting in-depth analyzes of crossover frequencies and distributions in sticklebacks. We base our findings on exceptionally dense linkage maps and propose a simple, yet elegant mechanistic explanation for the observed sex differences in crossover localization. Our findings are consistent in two distant genera of sticklebacks, proposing evolutionary conservation of crossover process at least at the level of taxonomic family.

Sexes differ in crossover frequency and localization

While our analyses re-confirm that female sticklebacks have more crossovers than males (Rastas *et al.* 2016, Sardell *et al.* 2018), with ultra dense linkage map we could study the consequences of this in great detail. We demonstrate that the higher crossover frequency in females leads to more uniform crossover distributions across chromosome, which in turn makes the recombination rate over large genomic regions to be highly sex specific (Fig. 3A-B). Interestingly, we found the crossover frequencies in males to be highly constrained, varying little between different chromosomes and minimally among individuals. In line with this, the crossover frequency was not found to be heritable in the nine-spined stickleback males, but highly so in females of both species. It appears that crossover frequency is under strict constraints in males and that males mostly have only the one obligate crossover in every bivalent. This is in stark contrast to females in which the number of crossovers increases as a function of chromosome length and appears to be constant per base pair.

We also found strong evidence for positive crossover interference (Fig. 3A), and that neither of the widely used gamma or gamma-sprinkling models provided good fit to our data (Supplementary Tables S5–S8). In general, gamma-sprinkling model fitted the data better than gamma model, and in females, the parameter estimates were similar to estimates reported earlier (Broman and Weber 2000 [human], Limborg *et al.* 2015 [sockeye salmon], Wang *et al.* 2016 [cattle]). However, it is unclear, whether in the case of the gamma-sprinkling model, the proportion of non-interfering crossovers (parameter p) are biologically realistic as little is known about different crossover pathways in sticklebacks. Interestingly, we found that simple probabilistic models can produce crossover counts that fit the paternal crossover data better or equally well than widely used mechanistic models of crossover formation (Supplementary Tables S6 and S8). This finding differs from the results presented by Broman and Weber (2000). For maternal crossovers, these models do not fit the data, likely because of the high among individual variation in the crossover frequency in stickleback females (Supplementary Tables S5 and S7). However, none of these models fit the data particularly well and given that these probabilistic models do not take the crossover positions into account, they also lack in their explanatory power. Because of the conflicting results of the different studies and models, there is a clear need for further studies, both empirical and methodological, on crossover interference and its relation to the crossover frequency (Otto and Payseur 2019).

The question whether crossover interference can act across the centromere dates back to 1940's as discussed by Hultén (1974). In an empirical study of grasshopper Colombo and Jones (1997) inferred that chiasma interference acts across the centromere. Studying humans, Broman and Weber (2000) came up with a similar conclusion, as did also Drouaud *et al.* (2007). Using our ultra dense data, we did not find evidence for the interference acting across the centromere (Supplementary Fig. S4). However, we found that two crossovers typically occur in different arms – likely due to interference (Fig. 3A) – but their locations seem to be determined by other crossovers in the same arm, centromere and telomere, rather than crossover interference across the centromere. More research is needed to investigate whether this discrepancy demonstrates diversity in crossover process or the effect of the data precision on the conclusions.

By decomposing the crossover distributions to different crossover classes, we discovered that the previously reported decrease of recombination rate towards the chromosome centers is likely caused by mixed effects of centromere and crossover interference (Fig. 3). Therefore, the distribution of crossovers along the chromosomes is shaped by mechanistic processes and studies of the ultimate reasons behind the distribution should take those into account. The results give strong support for our hypothesis that sex differences in crossover distributions are, at least partially, explained by differences in crossover frequencies and in the subsequent effects of crossover interference (Fig. 5).

Implications of sex-specific recombination

We found large differences in crossover frequencies between sexes. The crossover process has been investigated earlier in stickleback fishes but the analyses failed to detect and study heterochiasmy in detail. Roesti *et al.* 2013 studied three-spined sticklebacks but could not detect heterochiasmy due to unsequenced parents (see Corrigendum by Berner 2013). Sardell *et al.* 2018 reconstructed linkage maps from an hybrid cross with a Japan sea stickleback (*Gasterosteus nipponicus*) male and three-spined stickleback female which caused the sex-specific crossover rate to be possibly confounded with the inter-species differences and possibly biased because of the hybrid cross design.

Ultimate reasons for heterochiasmy are still elusive but recent studies have provided some new insights into mechanistic explanations (Sardell and Kirkpatrick 2020). One possibility is that the differences in crossover frequencies reflect the developmental and physiological differences between the sexes such as duration of the entire meiosis or chromosome packaging (Zickler and Kleckner 2016). However, males can have more crossovers than females, as in the case of sheep for instance (Johnston *et al.* 2016), which means that the crossover frequencies cannot only be defined by general differences in meiosis.

Regardless of the ultimate and proximate reasons for the sex difference in crossover frequencies, the outcome is that most of the variation and novel haplotypes are generated in the sex having more crossovers, females in the case of sticklebacks. The fact

that sexes differ also in localization of crossovers implies that linkage between certain loci is broken only in females. Concise example of this is that paternal crossovers occur in q-arm more often than would be expected by chance, whereas the opposite holds for maternal crossovers due to multiple crossovers and interference. However, higher number of maternal crossovers balance the difference of recombination rates between the chromosome arms making the ratio of their genetic lengths (cM) to match the ratio of genomic lengths (bp) (Supplementary Fig. S7).

Predicting recombination requires understanding of crossover process

In our linkage maps, genetic distance between adjacent markers is the observed frequency of recombinant meiotic products. In longer intervals, genetic distance between the markers is a sum of genetic distances between every adjacent pair of markers within that interval and, for instance, the genetic distance between the chromosome ends, i.e. the genetic length of a chromosome, is the sum of genetic distances between all adjacent markers. Therefore, in longer intervals where the distance is additive, the genetic distance does not tell the frequency of recombinant molecules as it is not measured as such units. In order to estimate recombinant frequency in context of a certain marker pair, the expected frequency of meiotic products with odd number of crossovers in between them is needed. For any non-zero number of crossovers in the bivalent, half of the meiotic products carry an odd number of crossovers and are therefore recombinants. Hence, to predict recombinant frequency, it is sufficient to know the probability for zero crossover between the markers *in the bivalent*. From this rationale follows that recombinant fraction r can be estimated as $r = \frac{1}{2}(1 - p_0)$, where p_0 is the likelihood of no crossovers between the markers in the bivalent (see Supplementary Fig. S8). If $p_0 = 0$, i.e. there is always at least one crossover between the markers in the bivalent, r becomes $\frac{1}{2}$, which in general terms means random segregation of the markers and no linkage between them. Notably, the recombinant fraction depends on p_0 only and remains the same for any non-zero number of crossovers between the markers, meaning that the relationship between (additive) genetic distance and recombination depends on the context of the markers. Mapping functions, such as Haldane's (1919) and Kosambi's (1943) functions, do not consider this relationship to vary between markers. Our data has poor fit with these often used mapping functions (Fig. 6; see also Supplementary Fig. S9) indicating that the recombination processes underlying those functions do not reflect the recombination process in the sticklebacks. In fact, it was suggested already in the pre-genomic era that those functions are not generally suitable for multilocus data (see Liberman and Karlin 1984).

Here, we have shown that sticklebacks have always at least one crossover per bivalent (Fig. 1) and the chromosome ends are not linked in any of the chromosomes. However, to understand linkage between two intervening loci, it is essential to know the likelihood for absence of crossovers, which requires understanding of the global crossover localization in the chromosome. For this, the probability for certain number of crossovers in the entire bivalent (see Fig. 1) and the distribution of crossovers in meiosis with different number of crossovers, i.e. the decomposed distributions, are needed. Our analyses show that in sticklebacks, and in the nine-spined stickleback in particular, the localization of the crossovers is constrained and shaped by crossover interference and centromere (Fig. 4). One consequence of the highly determined crossover positions is the higher variance in sibling relatedness than expected under uniform crossover localization. In the nine-spined sticklebacks, empirical variance of autosomal genome-wide sibling relatedness is $2.95 \cdot 10^{-3}$. That is statistically significantly higher ($p < 0.001$, $F_{7562} = 1.29$) than $2.23 \cdot 10^{-3}$ the variance perdicted by the estimator of Risch and Lange (1979), which assumes a uniform distribution of crossovers.

Methodological considerations

There is a possibility for misinterpretation of individual crossovers from linkage maps. However, the high marker density reduces the chances of crossover events being masked by multiple crossovers between two adjoining markers and phasing of the markers by Lep-Map3 (Rastas 2017) allows placing the crossovers between certain markers. Although errors in the reference genome could cause false positives in crossover detection, those should be unlikely in our data due to the high-quality of the reference genomes, especially that of the nine-spined stickleback (Kivikoski *et al.* 2020). Slight differences in crossover frequency between the species (Fig. 1) are likely due to differences in the number of families and offspring and reference genome quality which all affect precision of the linkage map. Even more detailed study of the differences between the sticklebacks would require similarly assembled reference genomes and similar crossing designs with equal sample sizes.

Heritability of the crossover frequency is an important quantity as it determines the evolvability of the trait in response to selection. Few studies have estimated heritability of crossover frequency in wild organisms (Johnston *et al.* 2016, 2018; Kawakami *et al.* 2019), and the accuracy of the available estimates is unclear due to the nature of the crossover data and the inherent difficulties it causes for estimation of the heritability. Since crossover frequency is measured as counts it could be assumed to follow a Poisson distribution. However, interpreting the results of an animal model (i.e. generalized linear mixed model with pedigree) applied to a Poisson distributed response variable is challenging because of the differences between the observed and expected scales, as well as due to the noise caused by the link function and the Poisson process itself (de Villemereuil *et al.* 2016). Few of the studies listed above include any discussion or consideration of the scale and distribution issues, which calls for caution in reading too much on their heritability estimates.

We obtained very different heritability estimates when assuming gaussian or Poisson distribution (Supplementary Tables S9–S10). Assessment with simulated data (where heritability is not expected) showed that the latent scale estimates under Poisson distribution were greater than the expected value (ca. 0.5 instead of zero), indicating that the estimates were unreliable (Supplementary Tables S11–S12). With the simulation-based assessment of the estimates, we feel confident that our heritability estimates, assuming normal distribution, should be in the right ballpark. All this said, we note that the estimates may contain some bias because of the study design: in our data, the nine-spined stickleback half-sib families were generated by mating females to two different males (rather than vice versa as usually; Falconer and Mackay 1996), which means that our estimates of additive genetic variance for nine-spined sticklebacks could include also maternal and dominance effects, if such were present. However, to the best of our knowledge, no one has as yet studied the possible maternal and dominance effects on crossover frequencies. Even if such were present, they would not

affect our inference on heritability of crossover counts in three-spined sticklebacks since contrary to the breeding design used in nine-spined sticklebacks, the design used for three-spined sticklebacks is expected to yield additive genetic variance estimates free of maternal and dominance effects.

Conclusions

Taken together, by studying explicit number and location of observed crossovers in nine- and three-spined sticklebacks, the results provide one of the most detailed analysis of individual variation in crossover frequencies to date. By inferring the total number of crossovers in meiosis, we show that the sexes differ in crossover frequency and that the number of crossovers is highly constrained in males. We further show that the location of crossovers in chromosome is affected by interplay of centromere, telomeres and crossover interference, which in turn is affected by the number of crossovers. As a result, the results demonstrate that commonly used mapping functions and relatedness (IBD) estimators do not describe the crossover process accurately. While we cannot answer the question regarding the ultimate or proximate reason for the difference of crossover frequency in sexes, we do show that it has clear mechanistic consequences in terms of crossover distribution and local recombination rate. The results further illustrate how different approaches used for estimation of heritability of crossover counts can yield dramatically different results, and consequently, great care should be exercised in choice of estimation method. High heritability of female crossover rate suggests that should recombination be under selection, it appears to be readily evolvable in both stickleback species.

Acknowledgements

Our research was supported by Academy of Finland (grants: 129662, 134728 and 218343 to JM; 322681 to AL), the Doctoral Programme in Wildlife Biology Research (LUOVA; funding to MK) as well support from the Helsinki Institute for Life Sciences (HiLife; grant to JM). We wish to acknowledge CSC – IT Center for Science, Finland, for access to computational resources.

Literature Cited

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober, *et al.*, 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science **327**: 836–840.

Berner, D., 2013 Corrigendum to the article 'Recombination in the threespine stickleback genome - patterns and consequences' by Roesti et al. 2013, in press for Molecular Ecology. Molecular Ecology **22**: 3652–3652.

Broman, K. W., 2019 *xoi: Tools for Analyzing Crossover Interference*. R package version 0.68-3.

Broman, K. W. and J. L. Weber, 2000 Characterization of human crossover interference. The American Journal of Human Genetics **66**: 1911–1926.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, *et al.*, 2017 glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal **9**: 378–400.

Burt, A., G. Bell, and P. H. Harvey, 1991 Sex differences in recombination. Journal of Evolutionary Biology **4**: 259–277.

Cann, H. M., 1992 CEPH maps. Current Opinion in Genetics & Development **2**: 393–399.

Cervera, A., V. Rantanen, K. Ovaska, M. Laakso, J. Nuñez-Fontarnau, *et al.*, 2019 Anduril 2: upgraded large-scale data integration framework. Bioinformatics **35**: 3815–3817.

Charles, D. R., 1938 The spatial distribution of cross-overs in x-chromosome tetrads of *Drosophila melanogaster*. Journal of Genetics **36**: 103–126.

Charlesworth, D., 2017 Evolution of recombination rates between sex chromosomes. Philosophical Transactions of the Royal Society B: Biological Sciences **372**: 20160456.

Chmátal, L., S. I. Gabriel, G. P. Mitsainas, J. Martínez-Vargas, J. Ventura, *et al.*, 2014 Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. Current Biology **24**: 2295–2300.

Colombo, P. C. and G. H. Jones, 1997 Chiasma interference is blind to centromeres. Heredity **79**: 214–227.

de Villemereuil, P., H. Schielzeth, S. Nakagawa, and M. Morrissey, 2016 General methods for evolutionary quantitative genetic inference from generalized mixed models. Genetics **204**: 1281–1294.

Dréau, A., V. Venu, E. Avdievich, L. Gaspar, and F. C. Jones, 2019 Genome-wide recombination map construction from single individuals using linked-read sequencing. Nature Communications **10**: 4309.

Drouaud, J., R. Mercier, L. Chelysheva, A. Bérard, M. Falque, *et al.*, 2007 Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. PLoS Genetics **3**: e106.

Dumont, B. L., K. W. Broman, and B. A. Payseur, 2009 Variation in genomic recombination rates among heterogeneous stock mice. Genetics **182**: 1345–1349.

Falconer, D. S. and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*. Longmans Green, Harlow, Essex, UK.

Felsenstein, J., 1974 The evolutionary advantage of recombination. Genetics **78**: 737–756.

Fiston-Lavier, A.-S., N. D. Singh, M. Lipatov, and D. A. Petrov, 2010 *Drosophila melanogaster* recombination rate calculator. Gene **463**: 18–20.

Goldcar, D. E. and P. R. Fain, 1988 Models of multilocus recombination: Nonrandomness in chiasma number and crossover positions. The American Journal of Human Genetics **43**: 38–45.

Hadad, R. G., T. W. Pfeiffer, and C. G. Poneleit, 1996 Repeatability and heritability of divergent recombination frequencies in the iowa stiff stalk synthetic (*Zea mays* l.). Theoretical and Applied Genetics **93-93**: 990–996.

Hadfield, J. D., 2010 Mcmc methods for multi-response generalized linear mixed models: The MCMCglmm R package. Journal of Statistical Software **33**: 1–22.

Haenel, Q., T. G. Laurentino, M. Roesti, and D. Berner, 2018 Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. Molecular Ecology **27**: 2477–2497.

Haldane, J. B. S., 1919 The combination of linkage values and the calculation of distances between the loci of linked factors. Journal of Genetics **8**: 299–309.

Hill, W. G. and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genetical Research **8**: 269–294.

Housworth, E. and F. Stahl, 2003 Crossover interference in humans. The American Journal of Human Genetics **73**: 188–197.

Hultén, M., 1974 Chiasma distribution at diakinesis in the normal human male. Hereditas **76**: 55–78.

Hunter, C. M., W. Huang, T. F. C. Mackay, and N. D. Singh, 2016 The genetic architecture of natural variation in recombination

rate in *Drosophila melanogaster*. PLoS Genetics **12**: e1005951.
Johnston, S. E., C. Bérénos, J. Slate, and J. M. Pemberton, 2016 Conserved genetic architecture underlying individual recombination rate variation in a wild population of soay sheep (*Ovis aries*). Genetics **203**: 583–598.
Johnston, S. E., J. Huisman, P. A. Ellis, and J. M. Pemberton, 2017 A high density linkage map reveals sexual dimorphism in recombination landscapes in red deer (*Cervus elaphus*). G3: Genes, Genomes, Genetics **7**: 2859–2870.
Johnston, S. E., J. Huisman, and J. M. Pemberton, 2018 A genomic region containing REC8 and RNF212B is associated with individual recombination rate variation in a wild population of red deer (*Cervus elaphus*). G3: Genes, Genomes, Genetics **8**: 2265–2276.
Kadri, N. K., C. Harland, P. Faux, N. Cambisano, L. Karim, *et al.*, 2016 Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212b affect recombination rate in cattle. Genome Research **26**: 1323–1332.
Kauppi, L., A. J. Jeffreys, and S. Keeney, 2004 Where the crossovers are: recombination distributions in mammals. Nature Reviews Genetics **5**: 413–424.
Kawakami, T., A. Wallberg, A. Olsson, D. Wintermantel, J. R. de Miranda, *et al.*, 2019 Substantial heritable variation in recombination rate on multiple scales in honeybees and bumblebees. Genetics **212**: 1101–1119.
Kivikoski, M., P. Rastas, A. Löytynoja, and J. Merilä, 2020 New gains with a slimmer genome – an automated approach to improve reference genome assemblies. (preprint posted august 18, 2020). bioRxiv. doi:10.1101/2020.08.18.255596. .
Kong, A., G. Thorleifsson, D. F. Gudbjartsson, G. Masson, A. Sigurdsson, *et al.*, 2010 Fine-scale recombination rate differences between sexes, populations and individuals. Nature **467**: 1099–1103.
Kosambi, D. D., 1943 The estimation of map distances from recombination values. Annals of Eugenics **12**: 172–175.
Lenormand, T. and J. Dutheil, 2005 Recombination difference between sexes: A role for haploid selection. PLoS Biology **3**: e63.
Lenormand, T., J. Engelstädter, S. E. Johnston, E. Wijnker, and C. R. Haag, 2016 Evolutionary mysteries in meiosis. Philosophical Transactions of the Royal Society B: Biological Sciences **371**: 20160001.
Liberman, U. and S. Karlin, 1984 Theoretical models of genetic map functions. Theoretical Population Biology **25**: 331–346.
Limborg, M. T., R. K. Waples, F. W. Allendorf, and J. E. Seeb, 2015 Linkage mapping reveals strong chiasma interference in sockeye salmon: Implications for interpreting genomic data. G3: Genes, Genomes, Genetics **5**: 2463–2473.
Lloyd, A. and E. Jenczewski, 2019 Modelling sex-specific crossover patterning in *Arabidopsis*. Genetics **211**: 847–859.
Mank, J. E., 2009 The evolution of heterochiasmy: the role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals. Genetics Research **91**: 355–363.
Mather, K., 1938 Crossing-over. Biological Reviews **13**: 252–292.
Muller, H. J., 1916 The mechanism of crossing-over. American Naturalist **50**: 193–221.
Muñoz-Fuentes, V., M. Marcet-Ortega, G. Alkorta-Aranburu, C. L. Forsberg, J. M. Morrell, *et al.*, 2015 Strong artificial selection in domestic mammals did not result in an increased recombination rate. Molecular Biology and Evolution **32**: 510–523.
Otto, S. P. and T. Lenormand, 2002 Resolving the paradox of sex and recombination. Nature Reviews Genetics **3**: 252–261.
Otto, S. P. and B. A. Payseur, 2019 Crossover interference: Shedding light on the evolution of recombination. Annual Review of Genetics **53**: 19–44.
Pritchard, V. L., H. M. Viitaniemi, R. J. S. McCairns, J. Merilä, M. Nikinmaa, *et al.*, 2017 Regulatory architecture of gene expression variation in the threespine stickleback *Gasterosteus aculeatus*. G3: Genes, Genomes, Genetics **7**: 165–178.
R Core Team, 2018 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Rastas, P., 2017 Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. Bioinformatics **33**: 3726–3732.
Rastas, P., F. C. F. Calboli, B. Guo, T. Shikano, and J. Merilä, 2016 Construction of ultradense linkage maps with lep-MAP2: Stickleback F_2 recombinant crosses as an example. Genome Biology and Evolution **8**: 78–93.
Risch, N. and K. Lange, 1979 Application of a recombination model in calculating the variance of sib pair genetic identity. Annals of Human Genetics **43**: 177–186.
Roesti, M., D. Moser, and D. Berner, 2013 Recombination in the threespine stickleback genome-patterns and consequences. Molecular Ecology **22**: 3014–3027.
Sardell, J. M., C. Cheng, A. J. Dagilis, A. Ishikawa, J. Kitano, *et al.*, 2018 Sex differences in recombination in sticklebacks. G3: Genes, Genomes, Genetics **8**: 1971–1983.
Sardell, J. M. and M. Kirkpatrick, 2020 Sex differences in the recombination landscape. American Naturalist **195**: 361–379.
Sturtevant, A. H., 1915 The behavior of chromosomes as studied through linkage. Z. Abstam. Vererbung. **13**: 234–287.
Talbert, P. B. and S. Henikoff, 2010 Centromeres convert but don't cross. PLoS Biology **8**: e1000326.
Thompson, M. J. and C. D. Jiggins, 2014 Supergenes and their role in evolution. Heredity **113**: 1–8.
Varadharajan, S., P. Rastas, A. Löytynoja, M. Matschiner, F. C. F. Calboli, *et al.*, 2019 A high-quality assembly of the nine-spined stickleback (*Pungitius pungitius*) Genome. Genome Biology and Evolution **11**: 3291–3308.
Wang, Z., B. Shen, J. Jiang, J. Li, and L. Ma, 2016 Effect of sex, age and genetics on crossover interference in cattle. Scientific Reports **6**.
Weinstein, A., 1936 The theory of multiple-strand crossing over. Genetics **21**: 155–199.
Wilson, A. J., D. Réale, M. N. Clements, M. M. Morrissey, E. Postma, *et al.*, 2010 An ecologist's guide to the animal model. Journal of Animal Ecology **79**: 13–26.
Yu, K. and E. Feingold, 2001 Estimating the frequency distribution of crossovers during meiosis from recombination data. Biometrics **57**: 427–434.
Zickler, D. and N. Kleckner, 2016 A few of our favorite things: Pairing, the bouquet, crossover interference and evolution of meiosis. Seminars in Cell & Developmental Biology **54**: 135–148.

Figures and Tables

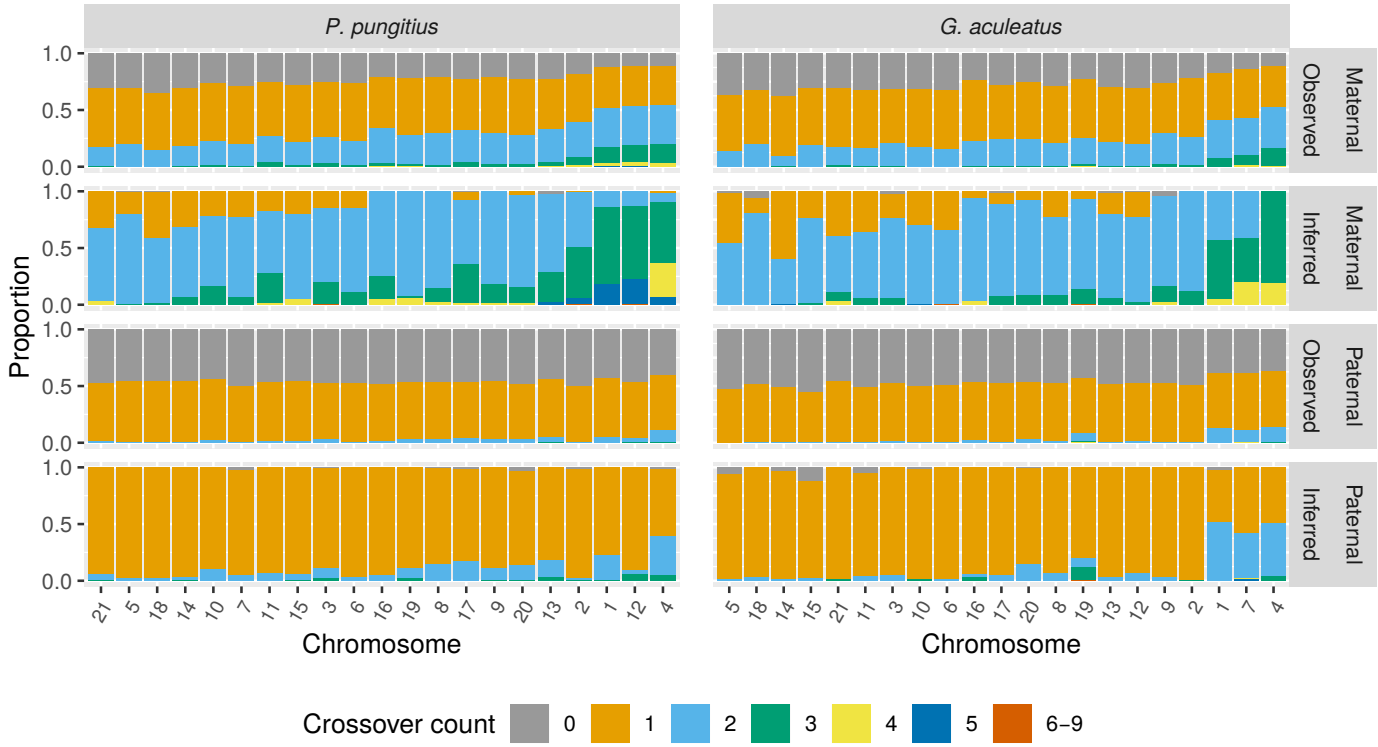

Figure 1 Observed and inferred crossover frequency distributions in the nine- (*P. pungitius*) and three-spined (*G. aculeatus*) sticklebacks. Each bar shows the proportion of offsprings with certain number of crossovers (Observed) and inferred proportions of meiosis with certain number of crossovers (Inferred) in maternal and paternal meiosis. Chromosomes are ordered from shortest to longest (in base pairs) and nine-spined chromosomes refer to linkage groups LG1-LG21 and three-spined chromosomes refer to linkage groups chrI-chrXXI. Crossover counts from 6 to 9 are grouped for readability.

Table 1 Heritability of crossover frequency in sticklebacks.

Species	Sex	h^2 (95% CI)[a]	V_A (95% CI)[b]	V_R (95% CI)[c]	V_P (95% CI)[d]
P. pungitius	maternal	0.649 (0.561,0.808)	13.4 (9.79,18.10)	6.72 (4.3,8.56)	18.7 (17.9,23)
P. pungitius	paternal	0.0862 (0.0261,0.131)	0.473 (0.163,0.849)	6.12 (5.22,6.66)	6.65 (5.96,7.32)
G. aculeatus	maternal	0.558 (0.296,0.826)	7.67 (4.31,17)	8.15 (3.61,11.7)	17.4 (14.8,21.6)
G. aculeatus	paternal	0.267 (0.118, 0.438)	1.97 (0.778, 3.33)	5 (4,6.23)	6.98 (6.2, 8.12)

[a] narrow-sense heritability (95% confidence interval)
[b] additive genetic variance
[c] residual variance
[d] phenotypic variance

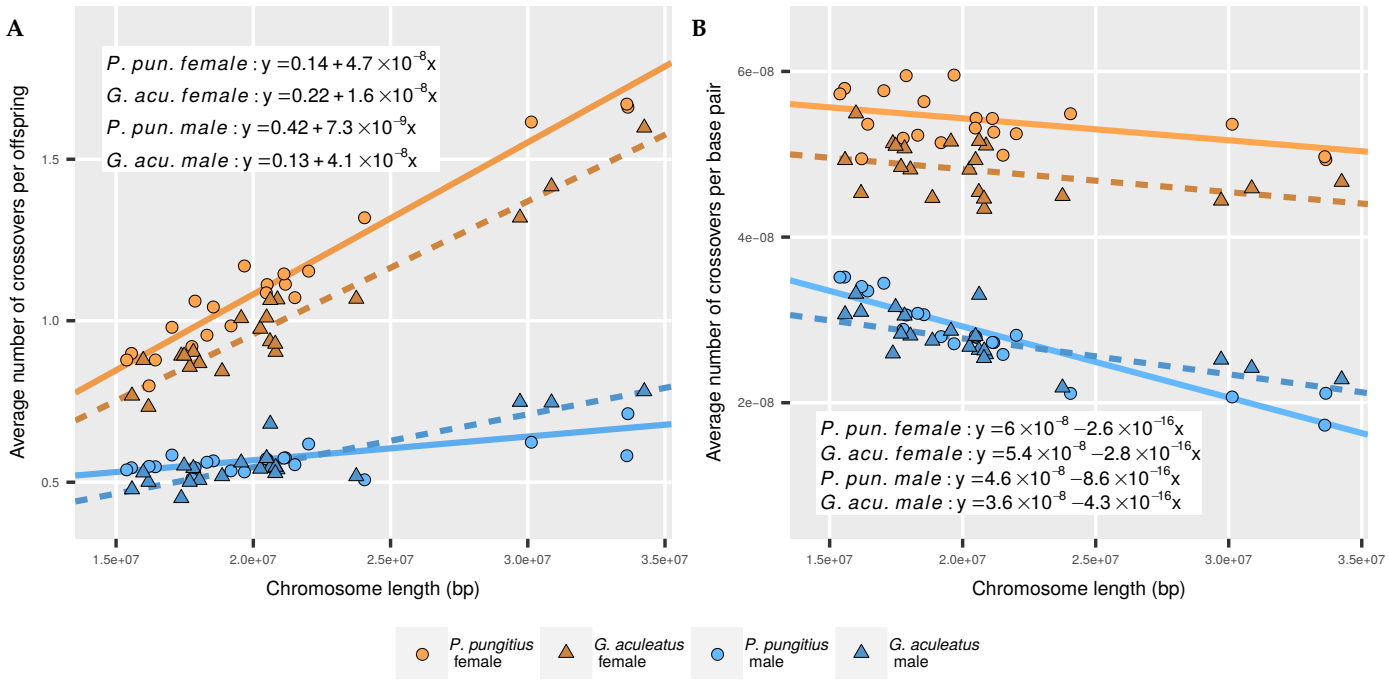

Figure 2 Crossover frequency as a function of chromosome length in nine- (*P. pungitius*, dot/solid line) and three-spined (*G. aculeatus*, triangle/dashed line) sticklebacks. (A) Number of crossovers per offspring increases as a function of chromosome length. In both species and sexes, the slopes differ from zero (*P. pungitius*: $p<0.005$; *G. aculeatus*: $p < 0.001$) and between the sexes (*P. pungitius*: $p<0.001$; *G. aculeatus*: $p<0.001$). (B) Number of crossovers per base pair decreases as a function of chromosome length in both species (*P. pungitius*: $p<0.04$; *G. aculeatus*: $p<0.05$). In the nine-spined stickleback this decrease is stronger in males ($p<0.005$), whereas in the three-spined stickleback the sexes do not differ ($p<0.3$). Sex chromosomes (LG12 and chrXIX) are shown in the plot but not included in the model fit.

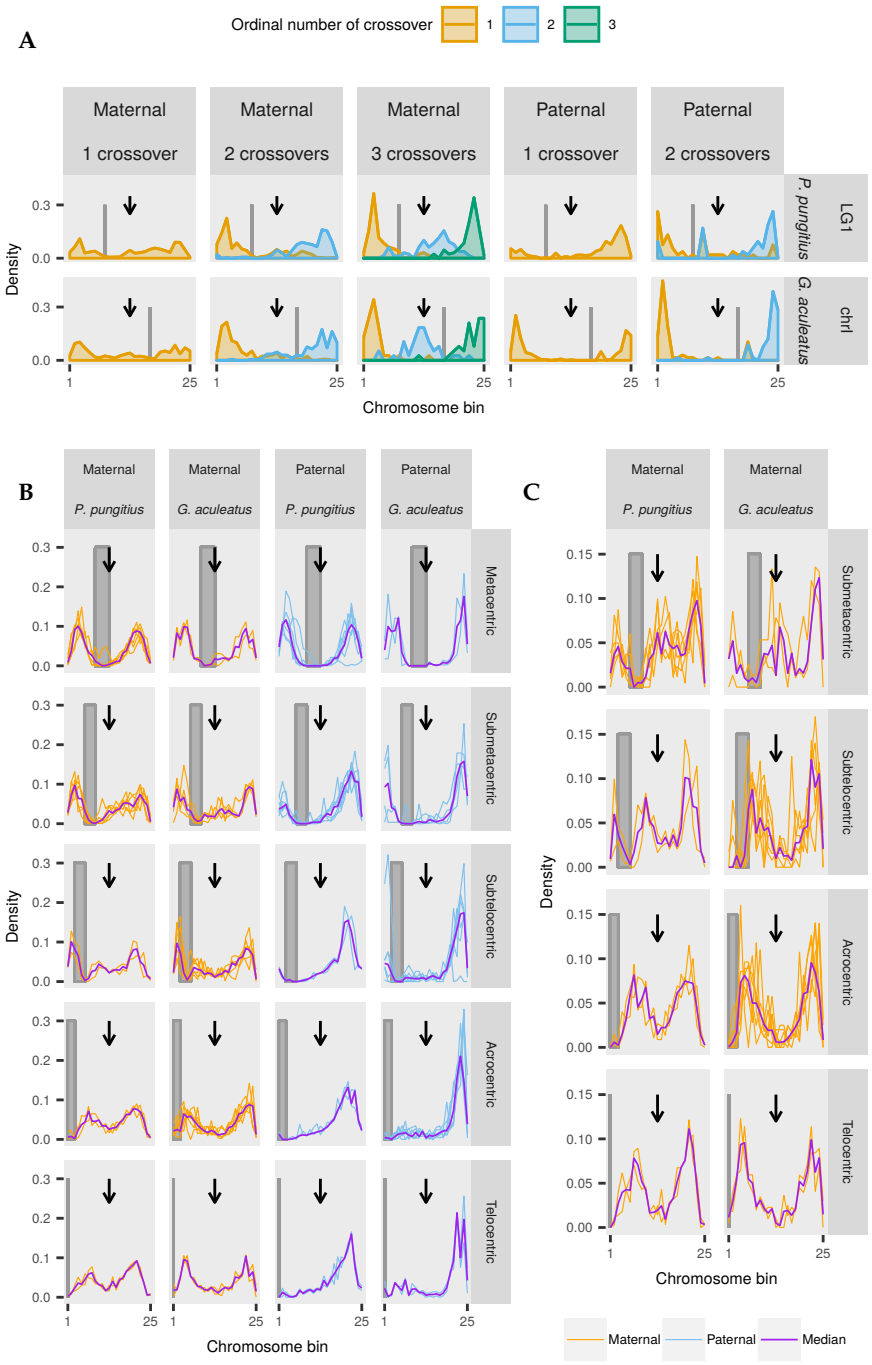

Figure 3 Chromosomal distribution of crossovers in nine- (*P. pungitius*) and three-spined (*G. aculeatus*) sticklebacks. Chromosomes are divided in 25 bins, grey bars and arrows indicate centromeric regions and chromosome midpoints, respectively. (A) Distribution of crossovers in the longest chromosomes for meiotic products with 1-3 observed maternal and 1-2 paternal crossovers. Colors correspond to the ordinal number of crossover starting from left. Spacing between adjacent crossovers indicates crossover interference, the third crossovers typically occurring at chromosome center. (B) Crossover distribution in chromosomes with different morphology. Orange and blue lines show maternal and paternal crossover distribution for each chromosome, the purple lines indicating the median. The crossover rate decreases towards the centromere and often towards the chromosome center. (C) Crossover distribution in meiotic products with two maternal crossovers in the long arm. Crossover interference pushes crossovers closer to the centromere and, especially in submetacentric chromosomes, closer to the chromosome center.

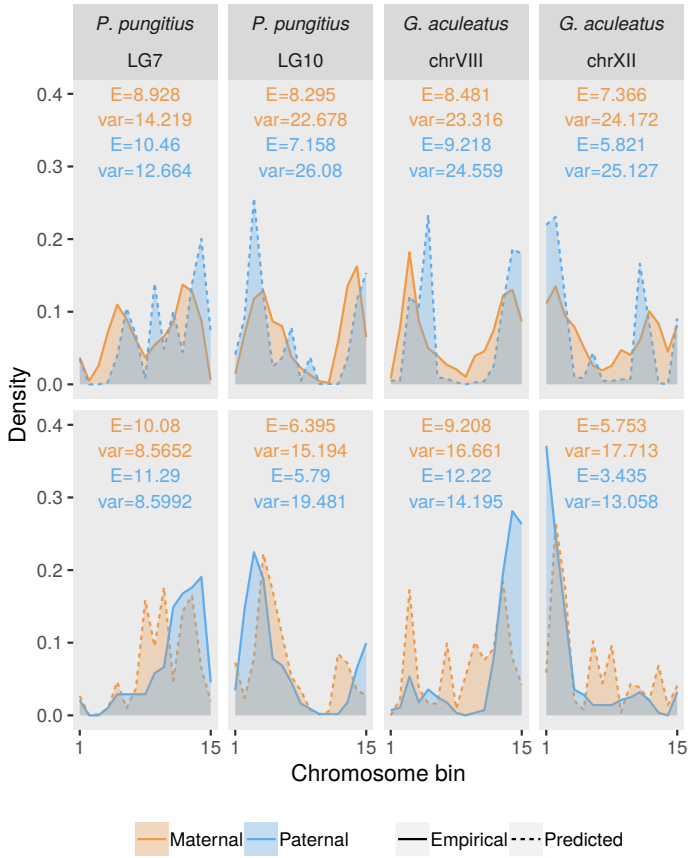

Figure 4 Distributions of observed crossovers of one sex (solid line) and crossovers simulated for the other sex (dashed line) assuming the crossover frequency of the oppose sex for two chromosomes in nine- (*P. pungitius*) and three-spined (*G. aculeatus*) sticklebacks. E and var refer to the expected value and the variance of the distribution.

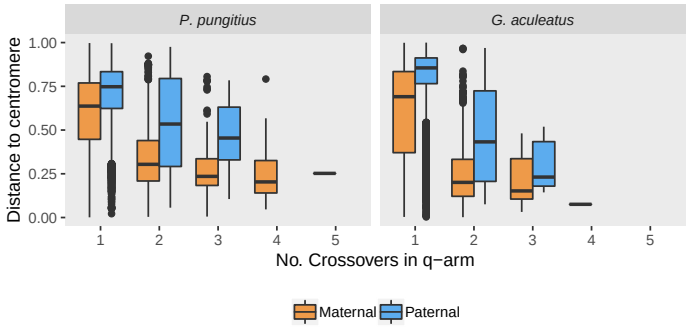

Figure 5 Distance to the centromere from the closest long-arm (q-arm) crossover in nine- (*P. pungitius*) and three-spined (*G. aculeatus*) sticklebacks. Maternal and paternal homologs are shown with orange and blue, respectively. Sex chromosomes are excluded from the analysis. Fit of the beta regression model:
Distance to centromere = $\alpha + \beta \cdot$ Crossover count $+ \sigma^2_{OFFSPRING} + \sigma^2_{CHR} + \epsilon$
P. pungitius maternal $= 1.31 - 0.93x + 0.0064 + 0.018; p < 0.001$
P. pungitius paternal $= 1.58 - 0.69x + 0.0064 + 0.018; p < 0.001$
G. aculeatus maternal $= 1.65 - 1.27x + 0.0060 + 0.023; p < 0.001$
G. aculeatus paternal $= 2.19 - 1.09x + 0.0060 + 0.023; p < 0.001$

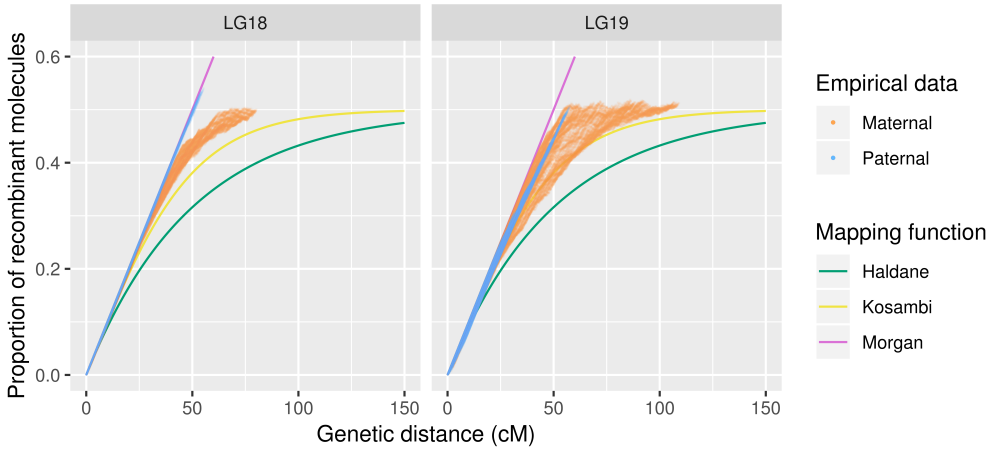

Figure 6 Relationship between genetic distance (cM) and proportion of recombinant molecules in two chromosomes of the nine-spined stickleback. Empirical data for each marker pair are shown with dots for maternal (orange) and paternal (blue) meioses. Three commonly used mapping functions are shown as reference. Paternal data matches well with Morgan's mapping function as males have mostly one crossover per chromosome. However, none of the mapping functions fits maternal data.