Evolutionary potential for genomic islands of sexual divergence on recombining sex chromosomes

Sarah P. Otto

Department of Zoology & Biodiversity Research Centre, University of British Columbia, Vancouver, BC Canada, V6T 1Z4

Summary

- Differentiated sex chromosomes are thought to develop through the accumulation of polymorphisms at loci subject to opposing selection between males and females, and/or between haploids and diploids. As sex chromosomes differentiate, reduced recombination becomes favored between selected loci and the sex-determining region, strengthening genetic associations between alleles favored in a sex and the corresponding sex chromosome.
- Here a model is analyzed to explore whether polymorphism at one sexually or ploidally antagonistic locus facilitates the spread of rare alleles at other loci experiencing antagonistic selection, promoting further differentiation of the sex chromosomes.
- It is found that antagonistic polymorphisms can spread and capture other such loci, building 'genomic islands' of differentiation on sex chromosomes, but the conditions are very restrictive, requiring the loci to be strongly selected, tightly linked and distant from the sex-determining region. Epistatic interactions can facilitate the promotion of polymorphism among selected loci, but only if preferentially favoring heterozygotes.
- Although these results apply to any taxa, plants provide a fertile ground for testing these and related theories given the recurrent evolutionary transitions to dioecy, which provide multiple opportunities to track the early evolution of sex chromosomes.

Introduction

According to now classic evolutionary theory, heteromorphic sex chromosomes evolve from recombining autosome-like ancestors via the accumulation of alleles that are beneficial to males on the Y (or Z) and that are beneficial to females on the X (or W), followed by natural selection for modifiers that reduce recombination between these loci and the sex-determining region (SDR; Fisher, 1931; Nei, 1969; Charlesworth & Charlesworth, 1980; Bull, 1983; Rice, 1987; Lenormand, 2003), ultimately allowing degeneration of genes on the Y (or Z; Rice, 1987). The coupling of selected alleles with specific sex chromosomes is analogous to the evolution of local adaptation through the spread of alleles beneficial in different spatial locations, followed by natural selection for modifiers that reduce gene flow (Balkau & Feldman, 1973; Wiener & Feldman, 1993) or that reduce recombination among the loci contributing to local adaptation (Pyłkowski et al., 1998; Lenormand & Otto, 2000). In each case, genetic associations are expected to accumulate such that alleles that function better together in one sex or location are found more often together in that sex or location, generally favoring the reduction of processes such as migration and recombination that break apart these associations (with some exceptions, see Otto, 2014).

Although there is now a substantial body of theory supporting the view that sex chromosomes can, but need not, evolve in this way (reviewed by Otto, 2014), there is a gap when it comes to understanding whether selectively maintained polymorphisms facilitate the maintenance of polymorphism at neighboring loci on the sex chromosomes. Intuitively, one might argue that pre-existing polymorphisms further differentiate the sex chromosomes, building stronger genetic associations that tie together alleles that work well within a sex to the appropriate sex chromosome. To the extent that polymorphisms facilitate maintenance of polymorphism at neighboring loci, the differentiation of sex chromosomes in pseudo-autosomal regions (PAR) may be a self-reinforcing and expanding process, favoring modifiers that reduce recombination over broader spans of the sex chromosome.

Previous work on autosomal loci has shown that a sexually antagonistic polymorphism at one locus, A, increases the parameter space at which selection can maintain polymorphism at a second locus, B (Patten et al., 2010), a process that can act even in hermaphrodites (Olito, 2017). Furthermore, linkage disequilibria accumulate that associate alleles favored in females (say, AB) and those that are favored in males (say, ab), which increases the amount of genetic variation that can be ascribed to sexually antagonistic selection (Patten et al., 2010; Ubeda et al., 2011). In these studies, locus A facilitates the maintenance of...
polymorphism at locus B, even though selection acts independently on each locus within each sex (multiplicative epistasis). The facilitative effect is strongest when selection is strong and when the selected loci are completely linked \((R=0)\), with the conditions allowing maintenance of polymorphism shrinking as the rate of recombination rises, approaching the conditions in the absence of locus A as \(R\) approaches \(1/2\).

Likewise, it is known that the existence of a sex-determining region facilitates the maintenance of polymorphism by sexually antagonistic selection at linked sites within the PAR (Bull, 1983; Rice, 1987; Jordan & Charlesworth, 2012; Charlesworth et al., 2014). Again, the lower the recombination rate between the selected locus A and the sex-determining region (denoted by a lower case \(r\)) to distinguish it from recombination between selected loci, \(R\), the broader the conditions for the maintenance of polymorphism. Specifically, with XY sex-determination, a population of \(AA\) or \(aa\) residents (‘res’) can be invaded by the opposite allele whenever the largest solution for the eigenvalue, \(\lambda\), of

\[
\lambda^3 - \lambda^2 \left( \frac{W_{aa}^g}{2W_{res}^g} + \frac{W_{Aa}^g}{W_{res}^g} (1 - r) \right) + \frac{W_{AA}^g}{2W_{res}^g} \left( W_{Aa}^g (1 - 2r) = 0, \right.
\]

Eqn 1(a)

is \(> 1\), where \(W_{res}^g\) represents the fitness of genotype \(g\) in males (sex \(\beta\)) or females (sex \(\alpha\)) (see Bull, 1983, p. 266). Equation 1a can be used to derive the conditions under which there is a ‘protected polymorphism,’ where both alleles \(A\) and \(a\) spread when rare. With sexually antagonistic selection favoring allele \(a\) in males and allele \(A\) in females, a protected polymorphism occurs at a locus A completely linked to the SDR \((r=0)\) if

\[
\frac{1}{2} \frac{W_{Aa}^g}{W_{aa}^g} \left( 1 + \frac{W_{Aa}^g}{W_{aa}^g} \right) > 1 \quad \& \quad \frac{W_{Aa}^g}{W_{AA}^g} > 1, \quad \text{Eqn 1(b)}
\]

whereas an unlinked locus on the sex chromosome or an autosome \((r=1/2)\) requires:

\[
\frac{1}{2} \left( \frac{W_{AA}^g}{W_{Aa}^g} + \frac{W_{aA}^g}{W_{aa}^g} \right) > 1 \quad \& \quad \frac{1}{2} \left( W_{Aa}^g + W_{aa}^g \right) > 1, \quad \text{Eqn 1(c)}
\]

(see Jordan & Charlesworth, 2012 for further exploration of the role of recombination on the maintenance of polymorphism). Thus, polymorphism can be maintained at a locus subject to sexually antagonistic selection that is tightly linked to the SDR whenever heterozygous females have sufficiently high fitness (i.e. \(W_{Aa}^g\) is sufficiently greater than \(W_{aa}^g\) that conditions in Eqn 1(b) are satisfied). Without linkage, however, both the female-beneficial allele A and the male-beneficial allele a must confer sufficiently high fitness in heterozygotes that the average heterozygous fitness across males and females is higher than either homozygote (conditions in Eqn 1c).

The above results suggest that combining a polymorphism at locus A and sex linkage would further facilitate the maintenance of polymorphism at a second sex-linked locus B. In this case, one might expect ‘genomic islands of sexual divergence’ to evolve on undifferentiated or neo-sex chromosomes. That is, one might expect to see regions with high levels of linkage disequilibrium between selected loci with a predominance of haplotypes that bear alleles favored either in males or in females, akin to genomic islands that can arise in the presence of spatially divergent selection (Nosil et al., 2009; Yeaman et al., 2016), especially when selected sites are closely linked relative to the strength of selection (Bürger & Akerman, 2011; Akerman & Bürger, 2013; Charlesworth & Barton, 2018). Here, assuming that locus A is at equilibrium on the sex chromosome and introducing rare alleles at locus B, it is shown that the conditions for maintaining polymorphism at B are not substantially impacted by locus A unless selection is strong, loci A and B are tightly linked, and neither locus is tightly linked to the sex-determining region. Surprisingly, tighter linkage between a selected locus and the sex-determining region causes it to have a much milder effect on the conditions for polymorphism at another locus of interest. It is concluded herein that genomic islands of sexual divergence are unlikely to arise except under special circumstances (e.g. with certain forms of epistasis or with strong selection between tightly linked selected loci).

Plants provide an excellent testing ground for theories, such as this, about the evolution of sex chromosomes, as there are hundreds of independent transitions to separate male and female sexes (Charlesworth, 1985; Renner, 2014; Goldberg et al., 2017), with many species at different stages of sex chromosome evolution (Westergaard, 1958; Ming et al., 2011; Charlesworth, 2016). Although phenotypic selection is thought to be more similar between male and female plants compared to animals (driving less sexual dimorphism, primarily in floral and genetically correlated traits; e.g. Delph, 2007), the opportunity for opposing selection pressures between haploid and diploid phases is substantially greater, with more extensive gene expression and selection in the haploid phase of plants compared to animals (see recent review by Immel & Otto, 2018). The models explored in this paper thus allow selection to differ between males and females and between haploid and diploid phases. This contribution ends with a discussion of emerging data on sex chromosomes in dioecious plants that may be used to test this and related theories on the evolution of sex chromosomes.

**Description**

The analysis tracks the frequency of alleles at a selected locus, B (with alleles \(B^h\)), within a system where the A locus (with alleles \(A^d\)) is at a polymorphic equilibrium and both are linked to the sex-determining region (SDR). It is assumed, for concreteness, that the SDR is based on a Y-dominant male-determining allele, with XY individuals being male and XX being female, although the results apply to ZW systems by considering ‘male’ to refer the heterogametic sex and ‘female’ to refer
to the homogametic sex. Both A and B reside in the PAR and potentially recombine with each other and the SDR. Importantly, the degeneration of Y-linked genes that can occur within nonrecombining regions is prevented by recurrent recombination with the X, so that the Y-linked alleles are assumed functional in this model (not hemizygous). For gene order SDR-A-B, the recombination rate between B and A is $R^g$ in females and $R^s$ in males, between A and the SDR is $r$ (only relevant in XY males), and double recombination events occur at rate $\chi$. Given appropriate choices of $r$, $R^g$, and $\chi$, any order of the loci can be modeled (see Appendix A1).

Except where noted, multiplicative interactions are assumed within each sex ($W^{sex}_{AABB} = W^{sex}_{AA} W^{sex}_{BB}$). Selection in the haploid stage and meiotic drive are included in the supplementary analysis package (Supporting Information Notes S1), developed in Mathematica v.8.0 (Wolfram Research Inc., Champaign, IL, USA). With haploid and diploid selection, the conditions under which a polymorphism can be maintained are diverse and include overdominance, sexually antagonistic selection and ploidyly antagonistic selection (Immler et al., 2012). Nevertheless, the impacts of polymorphism at locus A on the ability of selection to maintain a polymorphism at locus B are similar across these forms of selection, so the focus in this text is on sexually antagonistic selection in diploids except where noted.

Because associations develop between the genotypes in each sex, the frequencies of $AB$, $aB$, $Ab$ and $ab$ gamete combinations must be tracked in X-bearing female gametes (denoted $X^e_\phi$), in X-bearing male gametes (denoted $X^m_\phi$), and in Y-bearing male gametes (denoted $Y^m_\phi$). Censusing is carried out within each of these categories separately, such that the sum of $X^e_\phi$ across genotypes is one, as is the sum of $X^m_\phi$ and the sum of $Y^m_\phi$. The life cycle proceeds from a census of haploid gametes in males and females, followed by gamete union (random mating between eggs and sperm), diploid selection, meiosis with recombination and, finally, haploid selection. The recursions for the case of diploid selection are given in Appendix A1 and with haploid and diploid selection, as well as meiotic drive, in the Supporting Information Mathematica file (Notes S1).

The first step is to determine the equilibria at locus A. Two cases are analytically tractable: weak selection or tight linkage between A and the SDR. For these two cases the A locus is assumed to be at a polymorphic equilibrium and analytical results are obtained for the inversion conditions for allele $b$ when rare and for allele $B$ when rare. Then the conditions for a protected polymorphism, where both alleles can invade when rare, are compared to the conditions that would apply in the absence of locus A (i.e. to Eqn 1, replacing $A/a$ with $B/b$). Finally, the conditions under which B is polymorphic are compared numerically, allowing a range of linkage relationships and selection strengths.

Weak selection

Considering weak selection, the fitness of an individual with genotype $g$ can be written as $W^{sex}_g = 1 + \delta^{sex}_g$, where $\delta^{sex}_g$ is small (i.e. of the order of a small quantity, $\varepsilon$). With locus B fixed, the change in frequency of $X^e_A$, $X^m_A$, and $Y^m_A$ is calculated across one generation. Setting this change to zero and solving to leading order in the selection terms, the equilibrium frequency of allele $A$ is:

$$p_A = \frac{s_{Aa} - s_{AA} + s_{Ab} - s_{Aa}}{2s_{AA} - s_{AA} + 2s_{Ab} - s_{Ab} + \chi},$$

Eqn 2(a)

where $X^e_A = X^m_A = Y^m_A = p_A$. This equilibrium is valid and stable if:

$$s_{Aa} - s_{AA} + s_{Ab} - s_{Aa} > 0 \text{ and } s_{Ab} - s_{Aa} + s_{Ab} - s_{Ab} > 0.$$

Eqn 2(b)

Next the conditions are determined under which a rare allele, say $b$, can spread at the B locus. If selection is weak at locus A (of order $\varepsilon$) but strong at locus B (of order one), then it is trivially true that locus A can have only a minor effect on the spread of alleles at locus B because selection at locus A drops out in a stability analysis when analyzed to leading order.

Assuming instead that selection is similarly weak at both loci, allele $b$ can be introduced into a population at equilibrium (Eqn 2a) and its rate of spread determined from the leading eigenvalue in a stability analysis. To leading order in $\varepsilon$, allele $b$ can spread when rare when:

$$\dddot{p}_A \dot{p}_A^2 + \ddot{p}_A^2 + 2\ddot{p}_A (1 - \dot{p}_A) \dddot{p}_A + \dot{p}_A^2 + (1 - \ddot{p}_A)^2 \dddot{p}_A + \dot{p}_A^2 > 0,$$

Eqn 3(a)

where $\dddot{p}_A$ is the difference in fitness between the mutant $Bb$ and resident $BB$ individuals carrying genotype $g$ at the A locus (e.g. $\dddot{p}_A = W^{sex}_{AABb} - W^{sex}_{AAABB}$). A similar equation applies when allele B is rare instead, with $bb$ replacing $BB$ as the reference genotype in $\dddot{p}_A$. Eqn 3 is simply the fitness difference between $Bb$ and $BB$, averaged over the sexes and the genetic backgrounds at the A locus. Thus, allele $b$ spreads when rare if $Bb$ individuals are fitter $BB$ individuals in this average sense. Importantly, linkage to locus A does not play a role (to this order), and the only role that A plays is through epistasis affecting the $\delta^{sex}_g$ terms. With multiplicative interactions within each sex (i.e. no epistasis: $W^{sex}_{AABB} = W^{sex}_{AA} W^{sex}_{BB}$), alleles $b$ and $B$ can both invade when rare if:

$$s_{Bb} - s_{BB} + s_{Bb} - s_{BB} > 0 \text{ and } s_{Bb} - s_{BB} + s_{Bb} - s_{BB} > 0.$$

Eqn 3(b)

Hence, with weak selection, the conditions for the maintenance of variation at locus B are unaffected by polymorphism at locus A in the absence of epistasis (Eqs 2b, 3b are equivalent). It is worth noting that this weak selection approximation assumes that linkage is loose relative to selection (i.e. the recombination rates are all larger than order $\varepsilon$); Eqn 3(b) thus coincides with the weak selection conditions for a protected
polymorphism at an autosomal locus (Eqn 2.6 of Parsons, 1961).

Next, weak selection is considered but allowing tight linkage between loci A and B \((R^{\text{sex}} \sim 0(\epsilon))\). In this case, allele \(b\) can spread when rare if either of the following conditions hold to leading order in \(\epsilon\):

\[
\frac{\hat{\delta}_{AA}}{2} + \frac{\hat{\delta}_{Aa}}{2} + (1 - \hat{\rho}_{A}) \frac{\delta_{AA} + \delta_{Aa}}{2} > 0, \quad \text{Eqn 4(a)}
\]

\[
\hat{\rho}_{A} \frac{\delta_{AA} + \delta_{Aa}}{2} + (1 - \hat{\rho}_{A}) \frac{\delta_{AA} + \delta_{Aa}}{2} > 0. \quad \text{Eqn 4(b)}
\]

That is, the average fitness of allele \(b\) must be greater than the resident \(B\) allele, but this condition need only be true on the fittest genetic background at locus \(A\), allowing the haplotype \(ba\) to spread when conditions in Eqn 4(a) is met or haplotype \(bb\) when Eqn 4(b) is met. Similar conditions hold for the spread of allele \(B\) into a population fixed for \(b\) (replacing \(BB\) with \(bb\) in \(R^{\text{sex}}\)). The maintenance of polymorphism at locus \(B\) can be substantially easier to satisfy if epistasis tends to contribute positively to the \(\hat{\delta}_{\text{sex}}\) terms in Eqs 3 and 4 (see section below on epistasis).

Without epistasis, however, both \(\hat{\rho}_{A}\) and the fitnesses at locus \(A\) (e.g. \(W_{AA}^{b}\)) factor out of Eqn 4, so once again the conditions for a protected polymorphism reduce to Eqn 3(b), as if locus \(A\) did not exist. This is true also when recombination is of the same order as selection \((R^{\text{sex}} \sim 0(\epsilon))\). This result is counterintuitive: if allele \(a\) benefits males and allele \(b\) benefits males, then surely coupling the two alleles with tight linkage should allow \(b\) to spread more easily (ditto for female beneficial alleles \(A\) and \(B\))? This surprising result can be understood by considering the implication of the weak selection assumption. With weak selection, the genotype frequencies change slowly, allowing each genotype to experience selection in males and females equally often. At equilibrium, alleles \(A\) and \(a\) reach a point (Eqn 2a) at which the average fitness of each allele across both sexes is the same. Consequently, even when allele \(b\) arises with allele \(a\) where both increase the fitness of males, the dynamics of \(b\) will be determined slowly by selection, over which time the rise in frequency of \(a\) in males will exactly counterbalance the decline in frequency of \(a\) in females. Allele \(b\) will thus experience no boost from associations with neighboring sexually antagonistic loci when selection is weak and epistasis absent, regardless of the degree of linkage.

**Tight linkage between \(A\) and the SDR**

When locus \(A\) is tightly linked to the sex-determining region \((r\) small\), the allele frequencies on the \(X\) in males and females and on the \(Y\) can be substantially different (Clark, 1987; Otto, 2014).

At equilibrium, the \(Y\) chromosome is restricted to males and becomes nearly fixed for the allele that is fittest in males, say allele \(a\), with \(A\) nearly absent \((\hat{Y}_{A}^{\text{sex}} \sim O(\epsilon))\). The \(X\) chromosome, however, experiences different selection in males and females and may either remain polymorphic or become fixed for the \(A\) allele. Allowing arbitrarily strong selection, the resulting equilibria to leading order in \(r\) are:

\[
\hat{X}_{A}^{\text{sex}} = \frac{W_{AA}^{b} W_{AA}^{b} + W_{Aa}^{a} W_{Aa}^{b} - 2 W_{Aa}^{a} W_{Aa}^{b}}{2 \left( W_{AA}^{b} W_{AA}^{b} + W_{Aa}^{a} W_{Aa}^{b} - W_{Aa}^{a} W_{Aa}^{b} - W_{AA}^{b} W_{AA}^{b} \right)},
\]

\[
\hat{X}_{A}^{\text{sex}} = \frac{W_{AA}^{b} W_{AA}^{b} + W_{Aa}^{a} W_{Aa}^{b} - 2 W_{Aa}^{b} W_{Aa}^{b}}{W_{AA}^{b} \left( W_{AA}^{b} + W_{Aa}^{a} \right)^{2} - 2 \left( W_{Aa}^{b} + W_{Aa}^{a} \right) W_{AA}^{b} W_{AA}^{b}},
\]

\[
\hat{Y}_{A}^{\text{sex}} = 0 [\text{Equilibrium (A)}],
\]

\[
\hat{X}_{A}^{\text{sex}} = \hat{X}_{A}^{\text{sex}} = 1 & \hat{Y}_{A}^{\text{sex}} = 0 [\text{Equilibrium (B)}].
\]

Two similar equilibria exist with allele \(A\) fixed in males \((\hat{Y}_{A}^{\text{sex}} \sim 1 - O(\epsilon))\), which are not stable when sexually antagonistic selection favors \(a\) in males (Clark, 1987; Otto, 2014). The stability conditions of the equilibria given in Eqn 5 are given in Otto (2014), who notes that equilibrium (A) and (B) cannot be simultaneously stable.

Next the spread is considered of a rare \(b\) allele introduced into a population where the \(A\) locus is at equilibrium (A) or (B). The invasion conditions for allele \(b\) when rare (and for allele \(B\) when rare) are presented in Notes S1. Again, epistasis between loci \(A\) and \(B\) affects the conditions for a protected polymorphism (see section on epistasis).

When epistasis is absent and the system is at equilibrium with \(X\) and \(Y\) chromosomes fixed for different alleles (equilibrium B), a cubic equation of the form of Eqn 1(a) factors out of the characteristic polynomial in a local stability analysis and determines stability. Consequently, locus \(A\) again has no impact on the invasion conditions for alleles at locus \(B\). Essentially, all daughters will be \(XA/AX\) and all sons \(Xa/Ya\), regardless of the selection coefficients at locus \(A\). As the reproductive values of males and females are equal, selection at locus \(A\) has no influence on the relative success of the various chromosomes, and the fate of rare alleles at locus \(B\) depends only on selection at locus \(B\) and its linkage to the SDR.

The same is not true at equilibrium (A). Even when epistasis is absent, genetic associations can develop that affect the spread of alleles at the \(B\) locus when the \(X\) chromosome is polymorphic at locus \(A\). Let us again define alleles such that allele \(b\) is favored when rare in males. According to a stability analysis, allele \(b\) then spreads on the \(Y\) chromosome as long as \(W_{AB}^{b} > W_{BB}^{b}\), which does not depend on the \(A\) locus. A protected polymorphism requires this condition and that \(B\) can spread when rare, given its associations with the \(X\) chromosome. When locus \(B\) also is tightly linked to the SDR \((r, R^{2}, \text{and} R^{2}\) all small\), allele \(B\) can spread if either the \(XAB\) or \(XaB\) chromosome imparts a higher fitness upon its carriers:
where $\bar{W}_{\text{sex}}^r$ is the average fitness of residents (bearing allele $b$) of a particular sex. In the absence of locus $A$ or if selection is weak at locus $A$, these conditions are equal and equivalent to the first condition in Eqn 1(b) but for locus $B$. Eqn 6 indicates that strong selection at locus $A$ can facilitate the spread of rare alleles at locus $B$ on the $X$ chromosome because the alleles at locus $B$ need only perform well when carrying either allele $A$ or $a$, which is always easier to satisfy than if the rare allele must perform well when averaging over the alleles at the $A$ locus. With loose linkage between locus $B$ and the SDR, the stability analysis is unwieldy (a fifth-degree characteristic polynomial in $\lambda$, which depends on the fitnesses at locus $A$). Even so, if selection is weak (regardless of the degree of linkage) or if $B$ is loosely linked ($R_{\text{sex}}^r$ near $1/2$, regardless of the strength of selection), a term factors from the characteristic polynomial that depends only on the fitness effects of locus $B$, with the form given by Eqn 1(a).

Thus, it is concluded that when locus $A$ is tightly linked to the SDR and epistasis is absent, the maintenance of polymorphism can be aided by the presence of polymorphism at $A$, but only under fairly restrictive conditions: (1) the $X$ chromosome is polymorphic at locus $A$ (i.e. at equilibrium (A)), (2) selection is strong and (3) locus $B$ also is sufficiently linked to the SDR.

Incorporating haploid selection

The results summarized above apply as well when haploid selection is included, as detailed in Notes S1. In particular, with weak nonepistatic selection, the fitness effects of locus $A$, both haploid and diploid, drop out of the stability analysis. Specifically, incorporating haploid selection, the conditions in Eqn 3(b) for the spread of both allele $B$ and $b$ when rare assuming weak selection relative to recombination become:

$$
\frac{W_{\text{sex}}^r \left( W_{\text{sex}}^r \bar{X}_A^2 + W_{\text{sex}}^r \left( 1 - \bar{X}_A^2 \right) \right)}{W_{\text{sex}}^r} \geq 1,
$$

Eqn 6(a)

or

$$
\frac{W_{\text{sex}}^r \left( W_{\text{sex}}^r \bar{X}_A^2 + W_{\text{sex}}^r \left( 1 - \bar{X}_A^2 \right) \right)}{W_{\text{sex}}^r} \geq 1,
$$

Eqn 6(b)

mentioned above, namely that locus $A$ is at equilibrium ($A$) under strong selection and with tight linkage to both locus $B$ and the SDR. In this case, Eqn 6 with haploid selection becomes:

$$
\frac{1}{2} \left( \frac{V_{B} \bar{W}_{Bb}^r V_{A} \left( W_{\text{sex}}^r \bar{X}_A^2 + W_{\text{sex}}^r \left( 1 - \bar{X}_A^2 \right) \right)}{W_{\text{sex}}^r} \right)
$$

$$
\geq 1,
$$

Eqn 8(a)

or

$$
\frac{1}{2} \left( \frac{V_{B} \bar{W}_{Bb}^r V_{A} \left( W_{\text{sex}}^r \bar{X}_A^2 + W_{\text{sex}}^r \left( 1 - \bar{X}_A^2 \right) \right)}{W_{\text{sex}}^r} \right)
$$

$$
\geq 1,
$$

Eqn 8(b)

where $V_{i}^r$ is the haploid fitness of allele $i$ in pollen (sex $\delta$) or ovules (sex $\gamma$) and where $W_{\text{sex}}^r$ is the average fitness of resident haplotypes following diploid then haploid selection in a particular sex. Meiotic drive also is explored in Notes S1, where locus $A$ is again found not to influence the maintenance of polymorphism at locus $B$ with weak selection. The tight linkage case differs slightly, however, if meiotic drive at locus $A$ impacts the drive experienced at locus $B$, because the gametes produced by an individual then depend on how all of its haplotypes are driven (see Notes S1).

Role of epistasis

Epistatic fitness interactions between loci subject to antagonistic selection may be commonplace, particularly when fitness depends on composite traits affecting behavior or morphology that experience different selection pressures in males and females (Arnqvist et al., 2014). Intuitively, epistasis should facilitate the maintenance of polymorphism if alleles that are favored in females (say $A$ and $B$) interact well to further raise female fitness and likewise for alleles favored in males (say $a$ and $b$), strengthening the degree
of sex-specific adaptation. An examination of Eqns 3 and 4 assuming weak selection indicates, however, that epistasis that predominantly raises the fitness of homozygotes (e.g. raising the fitness of AABB females and aabb males) reduces the conditions under which a rare \( b \) allele will spread within a resident population of BB individuals, because \( \delta_{AA}^{RE} = W_{AABB}^{RE} - W_{AABB} \) is lowered if epistasis in females simply raises \( W_{AABB} \) and reduces the conditions under which the rare \( B \) allele will spread in a population of bb individuals, because \( \delta_{bb}^{RE} = W_{aabb}^{RE} - W_{aabb} \) is lowered if epistasis in males simply raises \( W_{aabb}^{RE} \). Thus, epistatic interactions that only act to increase the fitness of double homozygotes hinder the maintenance of polymorphism, making it more likely that allele \( B \) or \( b \) will remain fixed.

Conversely, interactions that primarily raise the fitness of double heterozygotes \( W_{AaBb}^{RE} \) in each sex, perhaps by providing sufficient gene products of both \( A \) and \( B \) to enhance female fitness and of \( a \) and \( b \) to enhance male fitness would increase the conditions under which which polymorphism is favored, by raising \( \delta_{AA}^{RE} = W_{AABB}^{RE} - W_{AABB} \) for both sexes in a resident population of BB individuals and \( \delta_{bb}^{RE} = W_{aabb}^{RE} - W_{aabb} \) in a resident population of bb individuals.

More complex cases of epistasis, where gene interactions affect the fitness of both homozygotes and heterozygotes, can help or hinder the maintenance of polymorphism, depending on whether fitness is elevated predominantly in heterozygotes (helps) or in homozygotes (hinders). For example, Arnqvist et al. (2014) numerically explored a case of epistatic interactions that reduced the fitness of double heterozygotes \( W_{AaBb}^{RE} \) lowered by \( \xi_{B}^{RE} \) and \( W_{AaBb}^{RE} \) by \( \xi_{A}^{RE} \) but reduced the fitness of double homozygotes even more \( W_{aabb}^{RE} \) lowered by \( 4\xi_{b}^{RE} \) and \( W_{aabb}^{RE} \) by \( 4\xi_{a}^{RE} \), see Table 1 in Arnqvist et al. (2014) for complete details), finding that such epistasis could, but did not always, increase the parameter space in which polymorphism was maintained. Another case that yields straightforward predictions occurs with dominant epistatic interaction where any female carrying both \( A \) and \( B \) has an elevated fitness \( W_{AABB}^{RE}, W_{AABB}^{RE}, W_{AABB}^{RE}, W_{AABB}^{RE} \) raised by \( \xi_{A} \) and any male carrying both \( a \) and \( b \) has an elevated fitness \( W_{aabb}^{RE}, W_{aabb}^{RE}, W_{aabb}^{RE}, W_{aabb}^{RE} \) raised by \( \xi_{b} \). Working through the effects of this form of epistasis on the weak selection results (Eqns 3, 4) indicates that it always helps maintain polymorphism because epistasis only increases fitness of heterozygous males when \( b \) is rare or heterozygous females when \( B \) is rare. For example, when \( b \) is rare, epistasis in females has no effect on the \( \delta_{B}^{RE} \) (because both \( Bb \) and \( BB \) are raised equally), whereas both \( \delta_{bb} = W_{aabb}^{RE} - W_{aabb} \) and \( \delta_{BB} = W_{AABB}^{RE} - W_{AABB} \) are raised by \( \xi_{b}^{RE} \), expanding the conditions given by Eqns 3 and 4 for a protected polymorphism if \( \xi_{b} > 0 \) and \( \xi_{B} > 0 \).

Thus, epistatic interactions that cause female-beneficial alleles to work well together in females and male-beneficial alleles to work well together in males will not universally favor the maintenance of variation. Although such epistasis does improve sex-specific adaptation, these gene interactions can destabilize a polymorphism if one or the other homozygote at locus \( B \) becomes too fit, as also highlighted by Arnqvist et al. (2014). Only if epistasis primarily raises the fitness of \( Bb \) heterozygotes, at least while alleles \( b \) or \( B \) are rare, does it help maintain polymorphism.

The above discussion focused on weak selection, assuming that locus \( A \) is loosely linked to the SDR (the selected loci, \( A \) and \( B \), could be either loosely linked, yielding the conditions for a protected polymorphism given by Eqn 3, or tightly linked, yielding Eqn 4). A similar weak selection analysis when \( A \) is tightly linked to the SDR shows again that the conditions for the maintenance of polymorphism are positive functions of the \( \delta_{BB}^{RE} \) terms, expanding when epistasis raises the fitness of \( Bb \) heterozygotes and shrinking when epistasis raises the fitness of resident homozygotes, whether at equilibrium (A) or (B).

Numerical exploration

In order to complement the analysis, the conditions under which both a male-beneficial allele \( b \) and a female-beneficial allele \( B \) could invade when rare were numerically calculated, allowing for strong selection and a range of recombination rates (assuming no crossover interference \( \chi = r R^{2} \)).

Each curve in Fig. 1 illustrates the parameter combinations delineating when allele \( b \) can invade when rare (right of the curves nearer the \( y \)-axis) and when allele \( B \) can invade (above the curves nearer the \( x \)-axis). Thus, a protected polymorphism occurs in the region between two curves of the same type and color. The recombination between the \( B \) locus and the SDR was set to 0.01 (purple), 0.1 (blue), 0.3 (orange) and 0.5 (all loci unlinked, black), with locus \( A \) exactly midway (order SDR-A-B, with no sex effects
on recombination $R^2 = R^2$). Selection was assumed to be additive for both loci ($s_{bb} = (s_{AA} + s_{ab})/2$), with no epistasis between the loci. Selection against allele $b$ in females ($s_{bb}$) varies along the y-axis, with selection against allele $B$ in males ($s_{BB}$) varying along the x-axis, both measured relative to the fittest allele in each sex ($s_{BB} = s_{AA} = 0$). To explore the impact of selection at locus A on the conditions allowing polymorphism at locus B, selection was varied against allele $a$ in females and against allele $A$ in males from no selection ($s_{aa} = s_{AA} = 0$, solid curves), to moderately strong selection ($s_{aa} = s_{AA} = -0.1$, dashed curves), to very strong selection ($s_{aa} = s_{AA} = -0.5$, dotted curves), again measured relative to the fittest allele in each sex ($s_{AA} = s_{bb} = 0$).

Consistent with the analyses described above, when linkage between the adjacent loci is either very tight (purple curves) or very loose (black curves), the impact of locus A on the conditions maintaining polymorphism is negligible. Only when linkage is intermediate (e.g. 10 cM between locus A and the SDR, blue) and selection very strong (e.g. $s_{aa} = s_{AA} = -0.5$, dotted curves) does the parameter space for the maintenance of polymorphism open up appreciably. Allowing different strengths of selection in males and females at locus A (Fig. S1) or altering the dominance coefficient at locus A (Fig. S2) has a very minor effect on these curves. If the dominance coefficients are also altered at locus B, the curves shift, but the impact of locus A remains minor (Fig. S3). Moving the relative positions of the loci causes locus A to have less of an effect when it is located closer to the SDR than to locus B, as predicted analytically (Fig. S4a), but to have a greater impact when located closer to locus B than the SDR (Fig. S4b). Even then, effects are appreciable only if selection at the A locus is strong.

Finally, epistatic interactions were explored (Fig. 2), assuming that epistasis improves fitness of the double heterozygote by 5% in each sex, as might be the case if double heterozygotes benefit enough from having both alleles $A$ and $B$ to enhance female fitness and both $a$ and $b$ to enhance male fitness. This form of epistasis was a focus because it has no effect on fitnesses at the $A$ locus when the $B$ locus is fixed, making it easier to visualize the effects that epistasis alone is having, without having to consider whether this is coming from a shift in (or loss of) polymorphism at the $A$ locus. As can be seen, even a 5% fitness boost to double heterozygotes greatly increases the parameter space allowing a protected polymorphism when there is a preexisting polymorphism at locus A (dashed and dotted curves) compared to the case when A is fixed (solid curves), especially when selection is weak.

**Discussion**

This work explores whether polymorphism at sex-linked loci subject to opposing selection pressures – between males and females and/or between haploids and diploids – facilitates the maintenance of polymorphism at other loci. Maintaining a polymorphism generally requires a delicate balance of selective pressures, which is easier to achieve with linkage to the sex-determining region (SDR). As shown in previous studies (reviewed in Otto *et al.*, 2011; Charlesworth *et al.*, 2014), if selection is too strong in one sex (or ploidy level), or if dominance of the allele

---

**Fig. 2** Effect of epistasis favoring double heterozygotes for maintaining a polymorphism at locus B. Identical to Figure 1 except that the AaBb individuals have a 5% higher fitness in both males and females, a form of epistasis that facilitates the maintenance of polymorphism. The solid curves illustrate the conditions for polymorphism in the absence of locus A and are identical to Figure 1. With a polymorphism at locus A, however, epistasis expands the region in which a polymorphism is maintained (dashed curves for $s_{aa} = s_{AA} = -0.1$, dotted curves for $s_{aa} = s_{AA} = -0.5$), especially when selection at the B locus is weak to moderate. The purple dashed and dotted curves near the $y$-axis have moved to negative values and thus the male-beneficial allele $b$ can invade over the entire parameter space shown.

when favorable is too low, or if linkage to the SDR is too loose, then selection will tend to drive one allele or the other to fixation (see also Eqn 1). The goal of this study was to determine if the occurrence of one polymorphism facilitates the maintenance of variation at other loci subject to antagonistic selection. If so, then there could be a shift, like a phase transition, from sex chromosomes that harbor little variation to ones that harbor substantial variation in the recombining portion.

It was found, however, that polymorphism at one locus subject to opposing selective pressures does not strongly facilitate the maintenance of variation at other sex-linked loci, at least in the absence of epistasis. A second selected locus has no effect, to leading order, on the conditions maintaining polymorphism at a locus of interest when selection is weak, even if the two selected loci are tightly linked to each other or to the sex-determining region. This result is not too surprising when selection is weak and the selected loci recombine frequently: genetic associations would be dissipated by recombination faster than they would build up by selection. It is perhaps surprising that there is no effect even for two loci that are completely linked, but this result can be understood by noting that the genetic associations that could facilitate the maintenance of rare alleles depend on the product of selection at locus A and B and hence are negligible relative to selection on B itself (Notes S1).

There also is no effect if selection is strong and the pre-existing polymorphism (at locus A) is tightly linked to the sex-determining region, except under special conditions (requiring strong selection requiring strong selection...
at locus A, which must be at the equilibrium given by Eqn 5(a) but not Eqn 5(b), and requiring locus B to be sufficiently linked to the SDR as well. Essentially, the nonrecombining parts of the sex-determining region are already completely genetically isolated from one another and adding another locus to this (locus A) makes no difference. By analogy to the development of genomic islands of divergence in a spatial model, the SDR essentially experiences no migration; whereas linkage to the SDR has a dramatic effect on the maintenance of polymorphism of locus B, associating another locally (sexually) adapted polymorphism at locus A with the SDR has no additional effect, because it cannot make the chromosomes any more isolated than they already are with no migration.

With epistasis, there can be a phase transition, with one polymorphism facilitating another polymorphism, and so on, but only if epistasis, like dominance, tends to increase the fitness of heterozygotes (Fig. 2). Arnqvist et al. (2014) noted that this form of epistasis (negative) emerges naturally if the loci affect a composite trait that exhibits a concave fitness surface with different peaks in the two sexes. It is an open empirical question whether such negative epistasis is common. If instead the most fit homozygote in each sex (AB/AB in females, ab/ab in males, say) is particularly fit because it only bears alleles that work well in that sex (a form of positive epistasis), polymorphism would be hindered.

These conclusions appear at odds with those of Patten et al. (2010), who found that sexually antagonistic polymorphism at one autosomal locus, A, increases the parameter space at which selection can maintain polymorphism at a second autosomal locus, B, assuming additive selection of equal strength at both loci that acts in opposite directions in males and females. Even when A and B are fully linked (the best-case scenario), their condition (11) under which polymorphism can be maintained becomes vanishingly small under weak selection. Only when selection is strong and linkage between A and B sufficiently tight does the parameter space open up substantially. This is consistent with the results found herein (e.g. Fig. 1), with one additional restriction: for another selected locus to have an impact on the conditions maintaining polymorphism, neither locus can be too near the sex-determining region, or else the SDR will be the overriding factor setting those conditions.

One caveat to these results is that they focus on the long-term maintenance of polymorphism. Transient dynamics also will shape patterns of diversity on the sex chromosomes and may frequently involve sexually antagonistic selection, as fitness in males and females shift in response to a changing environment (Connallon & Clark, 2014). The same is likely true for ploidal antagonistic selection. For example, a shift in pollinator abundance could substantially alter the strength of pollen competition, causing different alleles to be favored at ploidal antagonistic loci. Such transient dynamics can drive both divergence between the sex chromosomes (Kirkpatrick & Guerrero, 2014) and transitions among sex chromosomes (e.g. Scott et al., 2018). Furthermore, Hill-Robertson effects can cause interference between selected loci, which was ignored in this paper. Selective interference in the pseudo-autosomal regions (PAR) of sex chromosomes has been shown to cause the accumulation of deleterious mutations alongside sexually antagonistic loci and could potentially favor the maintenance of some recombination on the sex chromosomes (Cavoto et al., 2018).

Antagonistic selection between haploid and diploid phases is an important mechanism by which genetic variation can be maintained, especially in plants (Immler et al., 2012; Immler & Otto, 2018). A majority of genes in plants are expressed in both sporophytic and gametophytic stages, with evidence that selection is, on average, more effective on such genes (e.g. Arunkumar et al., 2013 in Capsella, and Chettoor et al., 2014 in maize). Opposing selection pressures between haploids and diplpoids, as well as between males and females, can drive the evolution of sex chromosome features, including evolutionary strata, reduced recombination rates, and turnovers between different sex chromosomes (Scott & Otto, 2017; Scott et al., 2018).

Recent empirical work in Rumex indeed finds tell-tale signatures of haploid selection playing an important role in the evolution of sex chromosomes: a bias towards pollen expression on the autosomes that evolved into sex chromosomes and an enrichment of genes highly expressed in pollen on the Y chromosome (Sandler et al., 2018). In papaya, the extent of the nonrecombining region of the sex chromosome is polymorphic, with the longer variant capturing only four genes, one of which is homologous to an Arabidopsis gene FAB1D that is broadly expressed and affects pollen development and male fertility (Lappin et al., 2015). Further work is needed, however, to determine the fitness effects of these loci in male and female sporophytes and gametophytes.

Although conflicting selection pressures between haploid and diploid stages, as well as between males and females, can promote the maintenance of polymorphism at any one locus, it is found herein that the conditions for maintaining polymorphism are little affected by other such loci unless they interact epistatically or selection is strong with tight linkage between the selected loci (e.g. see Eqns 7, 8).

Thus, the theory developed here predicts that ‘genomic islands of sexual divergence’ should not be expected, with selectively maintained polymorphisms clustered together in the recombining portions of sex chromosomes, beyond the fact that the SDR itself will be a hotspot for the accumulation of nearby polymorphisms. The finding that antagonistic polymorphisms should not show a strong tendency to cluster, except when tightly linked to one another and under strong selection, may help identify genes experiencing antagonistic selection, because such genes are less affected by one another (e.g. exhibiting separated high divergence (Fst) peaks between males and females; Kirkpatrick & Guerrero, 2014). Furthermore, without a strong tendency to build associations between loci, models with one selected locus may continue to serve as a good proxy for evolution within the PAR. For example, Qiu et al. (2016) distinguish three regions when examining Fst between the sex chromosomes, based on the findings of Kirkpatrick & Guerrero (2014): (1) a very small region that is so tightly linked to the SDR that even neutral sites will diverge between the sex chromosomes (for 4 Ne r values < −5), (2) a broader region in which antagonistic selection can generate divergence (for r values less than the strength of selection) and (3) more distal regions in which selection is too weak relative to
recombination with the SDR to expect divergence. Except under special circumstances (e.g., strongly selected sites within a single gene or closely linked genes), this guide should serve well, even with multiple loci under selection. Using this framework, Qiu et al. (2016) searched for examples of X–Y sequence divergence in Silene latifolia, finding several potential loci but none that could be definitively placed in region (2) implicating selection.

Based on the results presented herein, an acceleration in the accumulation of polymorphisms within the recombining region of sex chromosomes also should not be expected. Rather, the view that is most consistent with the current work is that polymorphisms at sexually and ploidally antagonistic loci accumulate nearly independently of one another, followed by suppression of recombination. It is this expansion of the nonrecombining SDR, rather than polymorphisms begetting further polymorphism, that facilitates differentiation even further along the sex chromosome and drives the evolutionary transition to heteromorphic sex chromosomes.

Acknowledgements

This work was supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC RGPIN-2016-03711). I would like to thank Mark Kirkpatrick and Matt Pennell for helpful discussions and Stephen Wright and three anonymous reviewers for many helpful suggestions on the manuscript.

ORCID

Sarah P. Otto https://orcid.org/0000-0003-3042-0818

References


Conditions for maintaining a polymorphism at locus B with sex differences in selection.

Fig. S2 Conditions for maintaining a polymorphism at locus B with partial dominance at locus A.

Fig. S3 Conditions for maintaining a polymorphism at locus B with partial dominance at both loci.

Fig. S4 Conditions for maintaining a polymorphism at locus B with unequally spaced loci.

Notes S1 MATHEMATICA file with derivations in .nb and .pdf formats (see separate files).

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.

Appendix A1

Recursions

The recursion equations consider only diploid selection (see Notes S1 for the extension to haploid selection and meiotic drive). Random union of gametes generates diploid females at frequency $x_{ij} = X_i^j X_j^i$ and diploid males at frequency $y_{ij} = X_i^j Y_j^i$, where $i$ and $j$ refer to the gamete genotype numbered according to 1: $AB$, 2: $aB$, 3: $Ab$, 4: $ab$.

Selection alters the diploid frequencies to $x_{ij} = W_{ij}^x x_i y_j / W_{ij}^y$ in females and $y_{ij} = W_{ij}^x y_i x_j / W_{ij}^y$ in males, with mean female fitness $W_{ij}^x = \sum_{i=1}^{4} \sum_{j=1}^{4} W_{ij}^x y_{ij}$ and mean male fitness $W_{ij}^y = \sum_{i=1}^{4} \sum_{j=1}^{4} W_{ij}^y y_{ij}$.

Females then undergo meiosis with recombination between the A and B loci at rate $R^x$:

$$X_{1ij}^0 = \frac{\sum_{j=1}^{4} x_{ij}}{R^x (x_{i4} - x_{i3})}, \quad \text{Eqn A1(a)}$$

$$X_{2ij}^0 = \frac{\sum_{j=1}^{4} x_{ij}}{R^x (x_{i4} - x_{i3})}, \quad \text{Eqn A1(b)}$$

$$X_{2ij}^0 = \frac{\sum_{j=1}^{4} x_{ij}}{R^x (x_{i4} - x_{i3})}, \quad \text{Eqn A1(c)}$$

$$X_{1ij}^0 = \frac{\sum_{j=1}^{4} x_{ij}}{R^x (x_{i4} - x_{i3})}, \quad \text{Eqn A1(d)}$$

Males undergo meiosis with recombination between the A and B loci at rate $R^y$, between the A locus and the SDR at rate $r$, and between the B locus and the SDR at rate $R^y + r - 2x$. Assuming gene order SDR-A-B, $x$ represents the rate of double recombination, but different gene orders can be modeled by choosing appropriate values of $R^x$, $r$, and $x$ (e.g. for gene order SDR-B-A, the rate of double recombination can be shown to equal $R-x$). The resulting frequencies of the X-bearing sperm are:

$$X_{1ij}^0 = \frac{\sum_{j=1}^{4} y_{ij}}{r (y_{i2} - y_{i1}) - (R^y + r - 2x) (y_{i3} - y_{i1}) - (R^y + r - x) y_{i4} + (r - x) y_{i1} + x y_{i3} + (R^y - x) y_{i2}}, \quad \text{Eqn A2(a)}$$

$$X_{2ij}^0 = \frac{\sum_{j=1}^{4} y_{ij}}{r (y_{i2} - y_{i1}) - (R^y + r - 2x) (y_{i3} - y_{i2}) - (R^y + r - x) y_{i4} + (r - x) y_{i2} + x y_{i3} + (R^y - x) y_{i4}}, \quad \text{Eqn A2(b)}$$

$$X_{2ij}^0 = \frac{\sum_{j=1}^{4} y_{ij}}{r (y_{i4} - y_{i3}) - (R^y + r - 2x) (y_{i3} - y_{i1}) - (R^y + r - x) y_{i4} + (r - x) y_{i3} + x y_{i1} + (R^y - x) y_{i4}}, \quad \text{Eqn A2(c)}$$
\[
X_{ab}^{3'} = \left( \sum_{j=1}^{4} y_{ij} \right) - r(y_{i3} - y_{i4}) - (R^3 + r - 2\chi)(y_{i42} - y_{i24}) - (R^3 + r - \chi)y_{i41} + (r - \chi)y_{i14} + \chi y_{i32} + (R^3 - \chi)y_{i23},
\]
Eqn A2(d)

\[
Y_{AB}^{3'} = \left( \sum_{j=1}^{4} y_{ij} \right) - r(y_{i1} - y_{i2}) - (R^3 + r - 2\chi)(y_{i31} - y_{i13}) - (R^3 + r - \chi)y_{i41} + (r - \chi)y_{i14} + \chi y_{i32} + (R^3 - \chi)y_{i23},
\]
Eqn A3(a)

The frequencies among the Y-bearing sperm are the same, but with the genotypic subscripts reversed \(y_{ij}^{3'}\) becoming \(y_{ji}^{3'}\), for example:

The equilibria and stability properties of these recursions are analyzed in the Mathematica file in Notes S1.