

## When and how do sex-linked regions become sex chromosomes?

### Key words

Pseudo-autosomal region (PAR), partially sex-linked region, hemizyosity, genetic degeneration, sexual antagonism

### Word counts

**Abstract:** 160 words

**Introduction:** 704

**Testing for fully sex-linked regions in species with genetic sex-determination, and estimating their sizes, when they are present** 998

**The ages of sex-linked regions, including non-recombining regions, when present** 1313

**Genetic degeneration** 526

**Why are sex-determining regions often non-recombining?** 1449

**Conclusions: why do some sex chromosome pairs remain homomorphic?** 551

**TOTAL:** 5,541

**Abstract:** The attention given to heteromorphism and genetic degeneration of “classical sex chromosomes” (Y chromosomes in XY systems, and the W in ZW systems that were studied first and are best described) has perhaps created the impression that the absence of recombination between sex chromosomes is inevitable. I here argue that continued recombination is often to be expected,

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/evo.14196](https://doi.org/10.1111/evo.14196).

This article is protected by copyright. All rights reserved.

that absence of recombination is surprising and demands further study, and that the involvement of selection in reduced recombination is not yet well understood. Despite a long history of investigations of sex chromosome pairs, there is a need for more quantitative approaches to studying sex-linked regions. I describe a scheme to help understand the relationships between different properties of sex-linked regions. Specifically, I focus on their sizes (differentiating between small regions and extensive fully sex-linked ones), the times when they evolved, and their differentiation, and review studies using DNA sequencing in non-model organisms that are providing information about the processes causing these properties.

## Introduction

Despite great progress in understanding the evolution of non-recombining sex-linked genome regions, surprisingly many interesting questions remain, including how often sex-determining regions have evolved suppressed recombination, or why they did so when this has occurred. The gaps in knowledge can create an apparently confusing picture. I here argue that much of the diversity reflects well-understood biological processes acting in a diversity of organisms that evolved separate sexes independently, at different times in the past, with sex-linked regions on different chromosomes, rather than “many exceptions to the rules” (Furman et al. 2020). Some striking similarities between sex-linked regions in different organisms stem from their lack of recombination (and subsequent genetic degeneration of Y and W chromosomes). Such sex-linked regions have, however, been the focus of so much attention that these features have sometimes been viewed as ubiquitous (Ponnikas et al. 2018) and regarded as inevitable. I focus on understanding that can be gained from younger systems and sex-determining regions that are in the process of evolving.

Figure 1 shows changes if a sex-linked region appears in a genome region, and Figure 2 summarizes different situations that generate systems with different ages and properties, from newly evolved sex-determining genes to XY or ZW sex chromosomes. This simplified framework focuses attention on the following well-established concepts:

- 1) An important general “rule” is that absence of recombination is crucial for evolution of distinct sex chromosomes (or extensive sex-linked regions). Even low crossover rates prevent differentiation (Pamilo et al. 1987; Blaser et al. 2014).

- 2) Absence of recombination can arise in different ways (Figure 2). Newly evolved (young) sex-determining regions will often be small, though they can arise within physically large non-recombining regions. Sex-linked regions could potentially remain as small as the initial sex-determining gene, or could subsequently evolve into large recombinationally suppressed regions. As explained below, the role of selection in reducing recombination is not yet fully understood.
- 3) If a sex-determining region does not recombine, Y-X differentiation will become detectable, over time, contrasting with adjacent recombining, or “pseudo-autosomal”, regions (PAR in Figure 1). Importantly, increased sequence divergence, accumulation of repetitive sequences (expanding the region’s physical size), and genetic degeneration (potentially deleting genes, shrinking the Y-linked region) occur on different time-scales. Sex chromosome heteromorphism mainly reflects the net effects of accumulation of repeats and genetic degeneration (other changes that can contribute will be mentioned below).

Figures 1 and 2 about here

Many advances have occurred since molecular markers, and especially genome sequencing, became available. Many studies hope to discover sex-determining genes, and therefore start with locating the species’ fully sex-linked genome region. These studies also provide valuable information about differentiation. We can now discover whether a species with genetic sex-determination has a non-recombining region, and, if so, how large it is, and when recombination between sex chromosomes became suppressed. Figure 1 shows the situation in very general terms, based on studies of animal and plant sex chromosomes over many years, and has no quantitative x axis scale for either the time or the level of genetic degeneration.

To describe the observed diversity, Figure 1 distinguishes between “sex-linked regions”, which span a wide range of ages, sizes and levels of differentiation, and differentiated “sex chromosomes” (potentially with cytologically detectable heteromorphism), and to categorize species with genetic sex-determination into the types in Figure 2. As has long been suspected, some organisms are in the early stages of sex chromosome evolution. For example, the plant *Mercurialis annua* shows minimal sequence divergence between Y- and X-linked sequences, and little sign of genetic degeneration (Veltsos et al. 2019). In some animals, new sex-determining regions evolved recently through “turnover events” (reviewed in Vicoso 2019). In contrast, other well-studied animals, including mammals, *Drosophila* and many birds, have very old-established sex chromosomes where most of

the XY or ZW pair is non-recombining, and the Y or W is highly degenerated, having lost most of the genes carried on the X or Z counterpart. In plants, separate sexes evolved more recently (see below), but the sex chromosomes of several distantly related dioecious plants share these properties, including *Silene latifolia* (Papadopoulos et al. 2015), *Rumex* species (Grabowska-Joachimik et al. 2015; Crowson et al. 2017), *Cannabis sativa* (Prentout et al. 2020) and *Coccinia grandis* (Fruchard et al. 2020).

### **Detecting completely sex-linked regions, and estimating their sizes, when they are present**

#### *Cytogenetic studies and genome sequencing*

Cytogenetic studies were important in the discovery of many of the old-established, highly degenerated, sex chromosomes examples just mentioned (Swanson 1957; Westergaard 1958; Zrzava et al. 2018). Loss of genes' functions allowed deletions of large parts of the Y or W chromosomes, creating detectable YX heteromorphism, or even X0 systems, as in many grasshoppers and nematodes, where the sex-determining genes have been lost (and maleness or femaleness no longer involves an active male- or female-determining factor on the Y or W chromosome; sex is then determined by an X-autosome balance system, as in *Drosophila*, see Swanson 1957).

Heteromorphism can also occur by accumulation of repetitive sequences. In some birds, accumulation appears to have restored homomorphism after ZW heteromorphism had evolved, and the W chromosome had lost most genes (Rutkowska et al. 2012; Furo et al. 2017).

Several different approaches have been developed for detecting such regions in genome sequences, and discovering their sizes. Detailed information is now available about large completely sex-linked regions in species where cytogenetic studies had already established their existence: mammals (Skaletsky et al. 2003; Cortez et al. 2014) birds (Pigozzi 2011; Zhou et al. 2014; Schmid et al. 2015), reptiles (Schield et al. 2019), fish, including sticklebacks (Varadharajan et al. 2019; Peichel et al. 2020), Dipteran and Lepidopteran insects (Vicoso and Bachtrog 2015; Fraisse et al. 2017), and plants, including bryophytes (Allen 1917, 1932; Okada et al. 2001; Ishizaki et al. 2002; Marks et al. 2019). Physically smaller completely sex-linked regions that also carry multiple genes have been found in a moss (McDaniel et al. 2007; Carey et al. 2020), brown algae (Ahmed et al. 2014), the flowering plant papaya (Wang et al. 2012), and some cichlids (Gammerdinger and Kocher 2018).

With just a single individual of each sex, fully sex-linked regions that have undergone genetic degeneration and lost genes can be ascertained since the Y- or W-linked region is "hemizygous" and

shows sex-specific haploid depth of coverage compared with the other sex, or with genome regions that are not completely sex-linked. This approach can be used even in non-model species such as Paleognathous birds (Zhou et al. 2014), snakes (Schield et al. 2019) and Schistosomes (Vicoso and Bachtrog 2011). However, it will only detect Y- and W-linked regions that have been non-recombining for long enough to have lost many genes, or for sequences to have diverged sufficiently that they do not map to a reference genome of the other sex (“old” systems in Figure 1). This approach may therefore fail to detect non-recombining regions in species with younger sex chromosomes, or whose sex-linked regions include only a few genes.

### *Genetic mapping*

Genetic mapping within families can detect variants that co-segregate with the sex of the progeny. Combined with genome sequencing, very large numbers of variants in DNA (or RNA transcripts) can be mapped to positions in chromosomes’ physical maps in a species’ reference genome assembly. With dense markers, even small regions may be detectable. Genetic studies can also reveal that sex-determining genes are on different chromosomes in related species (suggesting creation of new sex-determining regions by turnover events or repeated evolution of new sex-determining genes, as mentioned in Figure 2). Turnovers have been detected in many animals (Vicoso 2019), and are starting to be discovered in plants (Tennessen et al. 2018; Xue et al. 2020; Yang et al. 2020). Finally, genetic maps are increasingly being used to separate the two haplotypes of the sex-linked region, by genotyping progeny individuals and a parent of the heterozygous sex (Zhou et al. 2020). Such “phasing” of variants in a fully sex-linked region is necessary for describing Y-X differences and estimating Y-linked regions’ ages (see below). However, the haplotypes in families refer only to variants in the phased individual, which may not be consistently sex-specific.

Genetic mapping has established that the plant *Mercurialis annua* has a genetic sex-determination system, with co-segregating Y-linked markers across a physically large region, at least in families (Veltsos et al. 2019). Family analysis will, however, over-estimate the physical size of fully sex-linked regions, because, in a small family, some partially sex-linked markers will co-segregate with fully sex-linked ones and be classified incorrectly. The true sizes are therefore often not accurately known, but only upper size estimates.

### *Genome-wide association approaches (GWAS) using sex as the phenotype*

To ascertain sex-specific variants, samples of multiple unrelated individuals from natural populations are needed, to ensure many past generations during which recombination can occur. Such genome-wide analyses often employ  $F_{ST}$  estimates between the sexes for sequences in non-degenerated fully sex-linked regions (Natri et al. 2013; Schultheiß et al. 2015; Gammerdinger and Kocher 2018), similarly to analyses aimed at discovering genes affecting other phenotypes. The sex-determining region can potentially be narrowed down to a physically small part of a chromosome, as in many fish species (Kamiya et al. 2012; Myosho et al. 2015; Gammerdinger and Kocher 2018; Conte et al. 2019). Figure 3 shows the wide range of recent size estimates in flowering plants.

#### **Figure 3 about here**

When the Y- and X-linked regions can be distinguished, this can reveal heteromorphism in species whose chromosomes that are too small for cytological detection. This approach can also detect “micro-heteromorphism”, when the two haplotypes of a physically small sex-linked region differ by rearrangements, such as inversions, as in papaya (Wang et al. 2012), and/or show deletion/duplication differences (Figure 5 below shows one example).

Studies of the physical sizes of fully sex-linked regions also provide information about the numbers of “ancestral” (X- or Z-linked) genes. Genetic degeneration can then be quantified as the proportion of genes that remaining on the Y or W chromosome, and the proportion that still appear to be functional. These important data are currently available only in humans (Sayres and Makova 2013) and *Drosophila* (e.g. Zhou and Bachtrog 2015), and a few plants (Papadopoulos et al. 2015; Wu and Moore 2015; Fruchard et al. 2020; Prentout et al. 2020).

#### **The ages of sex-linked regions, including non-recombining regions, when present**

As already explained, the different ways in which sex-determining genes can appear, and different evolutionary times since they originated (Figure 2), can, in principle, explain much of the otherwise puzzling differences in differentiation between Y- and X-linked regions. It is therefore important to estimate ages of such regions, when they are present, and to test whether some sex-determining regions have simply not had enough time for suppressed recombination to evolve.

The appearance of a sex-determining locus on a chromosome defines its oldest fully sex-linked region. A single sex-determining gene is necessarily sex-linked, and, if the alleles controlling the two sexes differ by a single mutation, this mutation is completely sex-linked. Other sequence variants in the gene might, however, recombine (albeit with a low recombination rate, as they are physically close). Such variants will show incomplete associations with the sexes, as in *fugu*; in this fish, recombination must occur often enough to separate variants in the 17.5 kb region identified (Kamiya et al. 2012).

Many fully sex-linked regions include multiple genes. This could be because the sex-determining locus evolved in a non-recombining genome region carrying other genes (Figure 2). In species where the region was ancestrally non-recombining, the time when the sex-determining locus appeared defines the age. However, recombination may have become suppressed subsequently (see the next section).

#### *Age estimates using phylogenetic data*

The ages of a sex-linked region in a set of related dioecious species can be estimated from a phylogeny based on divergence of sequences from those of their closest non-dioecious relatives. However, this relies on assuming that the dioecious species share the same oldest fully sex-linked region. When turnovers have occurred, replacing one sex-determining region with a new one, this approach could over-estimate ages. This creates problems for estimating the ages of single gene sex-determining systems, where approaches using a non-recombining genome region (see the next section), are unavailable. To exclude turnover events, evidence is needed that all the species share the same sex-determining gene, in the same genomic location. In fish taxa, small sex-determining regions are often found in different physical locations, or different chromosomes (e.g. Kamiya et al. 2012; Myosho et al. 2015; Ieda et al. 2018), and in *Diptera* (Post 1985; Mahajan and Bachtrog 2017; Meisel et al. 2020). Among flowering plants, it is currently unclear which species have old-established yet small sex-determining regions, and consequently it is not known how often small regions fail to evolve into large recombinationally suppressed regions.

*Y-X divergence estimates*

Once Y-linkage is established, by whatever process, Y sequences will start diverging from their X counterparts, providing direct information about the time since Y-X recombination stopped (Figure 1). If sex-linked genes have been ascertained in a dioecious species, one can use Y-X sequence divergence to test whether the Y became isolated before the split from the most closely related non-dioecious species (Lawson-Handley et al. 2004; Dixon et al. 2018). Even within a single dioecious species, Y-X and W-Z divergence estimates reflect the times when recombination stopped. However, some points should be noted. First, times should ideally be estimated in terms of synonymous site divergence ( $K_s$ ) values, as such sites are often weakly selected (though their divergence is not completely neutral, since selection opposes substitutions in some synonymous sites, e.g. Chamary et al. 2006; Parmley et al. 2006; Walsh et al. 2020).  $K_s$  is nevertheless roughly proportional to the natural evolutionary time unit, the number of generations (at least until many sites may have undergone multiple substitutions, an effect termed “saturation”, illustrated in Figure 1). Second, the common practice of translating synonymous site divergence values into estimated numbers of years should be avoided, as it precludes comparisons between species with different generation times (which are often poorly known).

There are also difficulties. In functional genes, selective constraints acting on many non-synonymous sites make divergence initially slower than for synonymous sites. Although this obviates the need to correct for saturation, this advantage is outweighed because these estimates are confounded with degeneration, as non-synonymous substitutions increase as genes lose functions (Figure 1). Estimating divergence using all sites, without distinguishing between synonymous and non-synonymous sites), is therefore problematic. It is preferable to separate genome sequences into coding regions, and to estimate synonymous site divergence.

Furthermore, highly degenerated sex-linked regions often contain few genes, making Y-X or W-Z divergence estimates unreliable or impossible, for example in *Drosophila* species (except when fusions with autosomes in species without recombination in males have created neo-Y regions carrying many newly Y-linked genes, see Bachtrog et al. 2008). Old sex-linked regions with low gene density and high repeat density may even remain undetected unless the complete genome can be assembled and sex differences in coverage assessed (see above). A further problem arises when an XY male genome sequence is assembled using a female reference assembly (XX, avoiding assembly



problems with diverged Y-linked sequences). If the Y-linked region includes genes not present on the X chromosome, such reference-based assembly will either leave these sequences unmapped, or incorrectly map them to the most similar autosomal sequences. Reference-quality genome sequences of both sexes are therefore needed (Wei and Bachtrog 2019; Xue et al. 2020).

#### *Evolutionary strata*

Y-X divergence estimates have revealed that, even in species with long-established sex-linked regions, parts of these chromosomes sometimes also subsequently stopped recombining, forming so called “evolutionary strata”, first detected in humans (Lahn and Page 1999; Skaletsky et al. 2003). This observation is important, because it shows that recombination has undergone an evolutionary change, rather than being the ancestral state for the region. The oldest human strata (Y-X synonymous site divergence,  $K_s$ , much higher than 20%) are shared with other mammals (Sandstedt and Tucker 2004; Cortez et al. 2014) include large numbers of ancestrally X-linked genes (369, 84 and 128, see Sayres and Makova 2013). Smaller non-recombining strata adjacent to the PAR boundary are younger, with  $K_s$  around 20%, and some are specific to only some mammal lineages. Strata with multiple genes are also found in birds (Zhou et al. 2014), snakes (Schield et al. 2019) and plants (Bergero et al. 2007; Wang et al. 2012; Zhang et al. 2020). However, Y-X and W-Z divergence estimates remain scarce.

#### *Other changes after recombination between the sex chromosome pair becomes suppressed*

Regions of suppressed recombination carrying new sex-determining genes also start accumulating Y-linked mutations causing genetic degeneration, repetitive sequences (reviewed by Bachtrog 2008), and adaptive changes, including the evolution of dosage compensation in response to degeneration (e.g. Mank et al. 2011; Disteche 2016; Gu and Walters 2017). Gene movements into the sex-linked region may also occur, or from the sex-linked region to an autosome (Bellott et al. 2010). These changes also accumulate over evolutionary times (Figure 1), but probably do not greatly affect the sizes of non-recombining regions (though more detailed future studies are needed). Their rates depend on many factors other than evolutionary times, and do not estimate such times. These changes can, however, be used to infer changes in recombination in non-model species where sex-

linkage cannot be tested genetically, or in old, highly degenerated systems, where divergence time estimates are not possible (see Figure 1). For example, similarly to the phylogenetic approach described above, analysis of sex differences in coverage revealed degeneration in Schistosoma lineages, initially in an old fully sex-linked region, and later independently in adjacent, but different, regions in two derived lineages (Picard et al. 2018),. Clearly, recombination became suppressed in several distinct events.

Similarly, accumulation of repetitive sequences compared with a suitable ‘outgroup’ suggests that a species has evolved a new non-recombining region. Such changes may be very fast once recombination stops (Charlesworth et al. 1994). Prominent accumulation of repetitive sequences is detected in recently evolved *Drosophila* neo-Y chromosomes (Bachtrog 2003), grasshopper neo-Y-linked regions (Palacios-Gimenez et al. 2020), and before loss of genes in plants including papaya (Wang et al. 2012), and in lizards (Matsubara et al. 2014).

### Genetic degeneration

Although the extent of genetic degeneration increases with the time a region has been evolving under full sex linkage, theoretical modelling has identified other important factors (reviewed by Bachtrog 2008). Degeneration rates may therefore differ greatly between different organisms. Together with the scarcity of quantitative degeneration data and divergence time estimates, this contributes to the seemingly confusing picture mentioned above. Many studies describe depth of coverage ratios in the two sexes, which merely detects regions with degenerated sequences. Few indicate the proportion that are hemizygous in males, and the number of XY gene pairs whose Y copy is a pseudogene, and species with partially degenerated sex-linked regions or strata have been little studied.

Testable predictions are nevertheless available. First, most models predict that degeneration will be faster in sex-linked regions with many genes (though a recent model predicts degeneration of regions with few genes Lenormand et al. 2020). Estimates of numbers of “ancestral” genes should allow tests of these ideas. Single-gene systems, and small chromosomes that acquire a sex-determining gene, such as micro-chromosomes of lizards (Matsubara et al. 2014), might be expected to degenerate slowly, and data from such non-model species should become available.

Second, degeneration has a non-linear time-course. Genes are predicted to initially lose functions rapidly by major effect mutations, followed by slower changes, and eventually deletion of sets of genes (Figure 1). Data from sex-linked regions at all degeneration stages are therefore needed. Plants, which include many species with small or young sex-linked regions, may be less suitable than animals, because selection in the haploid phase, including the pollen of flowering plants, may oppose degeneration (Bergero and Charlesworth 2011; Chibalina and Filatov 2011; Hough et al. 2014). However, considerable degeneration has been documented in several plants (see above), so data from plants are still needed.

At all stages of degeneration of a non-recombining region, the rates also depend on the specific properties of the genes present (e.g. Kramer et al. 2016; Rifkin et al. 2020; Bellott and Page 2021). A striking example is the neo-Y of *D. busckii*, which is more degenerated (with 58% non-functional genes) than the larger and older one in *D. miranda* (only 34% non-functional genes), probably because the latter evolved from a “dot” chromosome, whose genes show low selective constraints (Zhou and Bachtrog 2015).

The data currently available suggest that most animal strata with Y-X or W-Z  $K_s$  values above 20% show essentially complete degeneration of most ancestral geness. With  $K_s$  below this value, 50% or more of the ancestral genes present on the X are generally also present as likely functional copies on the Y, consistent with theoretical predictions (Bachtrog 2008). However, the  $K_s$  level and evolutionary time needed for strata to reach the stage of major gene loss, and for *de novo* evolution of dosage compensation, remain unclear. Deletions within fully sex-linked regions, contributing to heteromorphism, probably occur only in the late stages of degeneration, as large deletions are generally highly deleterious (Bull 1983; Manna et al. 2012; Bazrgar et al. 2013) unless the genes are all under weak selection, or the region has already degenerated and become a “gene desert” (Nóbrega et al. 2004).

### **Why are sex-determining regions often non-recombining?**

I next outline different situations that can account for lack of recombination in sex-determining regions. First, I describe two situations that involve selection for reduced recombination in systems

with polymorphisms for a sex-determining gene and a second gene. Collectively they can be termed “the sexually antagonistic polymorphism hypothesis”.

#### *Appearance of new sex-determining regions: two-gene systems*

A major hypothesis to explain why XY and ZW chromosome pairs do not recombine involves selection for closer linkage between two polymorphic loci. One such situation (Figure 4A) arises when separate sexes evolve de novo in an ancestrally cosexual species (for instance dioecious flowering plants that evolved from ancestors with hermaphrodite flowers, or in monoecious species, with each individual producing both male and female flowers). The evolution of females and males requires two mutations. In one model (Charlesworth and Charlesworth 1978), the two mutations are in separate genes, and each mutation causes sterility of one sex, and therefore acts antagonistically in the other sex. As reviewed by Westergaard (1958), dioecy in plants often involves a recessive loss-of-function male-sterility mutation that produces the females, and a femaleness-suppressing mutation that produces males (Figure 2A); hypothetically both mutations could occur in a single gene, though no example has yet been reported. Sex-determination then resembles supergenes controlling other polymorphisms that probably involve separate polymorphic mutations affecting different traits (Schwander et al. 2014; Charlesworth 2016). Unless the second mutation acts specifically only in one genotype of the first gene, loose linkage prevents establishment of a polymorphism. However, if a closely linked mutation does establish a polymorphism, tighter linkage may evolve.

#### **Figure 4 about here**

#### *Sexually antagonistic (SA) polymorphism near a sex-determining gene*

Figure 2B illustrates the other plausible biological situation involving a two gene polymorphism: when a sexually antagonistic (SA) polymorphism establishes in a gene closely linked to a sex-determining gene (Kirkpatrick and Guerrero 2014), or a turnover event involves a sex-determining gene appearing near a SA polymorphism (Bull 1983; van Doorn and Kirkpatrick 2007). As mentioned above, turnovers are well-documented in animals and plants. The first of these situations can also arise after de novo evolution of separate sexes has generated a single sex-determining gene. Figure 5 illustrates a naturally evolved case in the persimmon, *Diospyros lotus*; in this plant, an active

(genetically dominant) maleness factor has evolved; females are the “default” sex, developing only when the maleness gene is absent (Akagi et al. 2014; Akagi and Charlesworth 2019).

#### Figure 5 about here

When the heterogametic sex is achiasmate (or genome regions that undergo crossovers are strongly localized in one sex), a new sex-determining gene can define a new fully sex-linked region. Such male-specific crossover patterns are observed in distantly related fish, including the guppy (Bergero et al. 2019) and the stickleback (Sardell and Kirkpatrick 2019), and in frogs (Rodrigues et al. 2018). As reviewed recently (Charlesworth 2019), they may be commoner than is currently realised, as genetic maps are rarely estimated separately for the two sexes, and physical maps, to locate the genome regions where crossovers occur, are available from only a few species.

#### *Caveats concerning the SA polymorphism hypothesis*

Although both situations in Figure 4 generate selection for closer linkage with the sex-determining gene, a response to this selection will not happen unless heritable variation exists for recombination rates in the region. Second, the selection and dominance coefficients under which SA polymorphisms can be maintained are restrictive (Fry 2010). In theoretical studies with given selection and dominance coefficients, maintenance is most likely in regions very closely linked to sex-determining loci (Jordan and Charlesworth 2012; Kirkpatrick and Guerrero 2014). Therefore, even though mutations with SA effects can probably occur in many genes across the genomes of dioecious animals and plants (Connallon and Clark 2014), the numbers of genes in which SA polymorphisms could potentially be established are probably small (unless a sex-determining locus evolves within a large ancestrally non-recombining region). The waiting time until a suitable SA polymorphism becomes established will therefore often be long. The same applies to the model in Figure 4A (Charlesworth and Charlesworth 1978). However, if linkage is already close, closer linkage will evolve slowly (though an inversion preventing crossovers could nevertheless establish an extensive new Y-linked region, see Figure 2C).

On the other hand, SA conflicts can be resolved by evolving sex differences in gene expression, and need not lead to suppressed recombination. Vicoso *et al.* (2013) suggested that this might account for the persistence of large partially sex-linked regions in the Paleognathous birds (whereas

recombination has become suppressed in the homologous regions of the ZW pair of Neognathous birds). Sex biased expression has not been confirmed for emu partially sex-linked genes (Xu et al. 2019). Nevertheless, the idea is plausible, as conditions for expression level changes are less stringent than for changes in linkage, because there is no requirement for a SA polymorphism to be maintained. Indeed, after fixation of an allele that benefits one sex, but reduces fitness in the other, sex-specific control of its expression is favored. Overall, therefore, chromosome regions carrying sex-determining genes are not always expected to evolve suppressed recombination.

Surprisingly few studies have examined the evolution of sex-limitation of expression, though expression differences are detected at many loci. Genes encoding testis or egg proteins, or anther-specific proteins in plants, inevitably have sex-specific expression, but many genes expressed in other tissues might change from being expressed in both sexes, to having sex biased expression if conflicts have evolved. Changes producing over-or under-representation of genes with sex-biased expression (termed “masculinization” and “demasculinization” of sex chromosomes) have been documented, and may reflect the different balance of selection on mutations with sexually antagonistic effects (as reviewed, for example, by Vicoso and Charlesworth 2006; Bachtrog et al. 2010; Mank and Wright 2012).

#### *Alternatives to the SA polymorphism hypothesis*

As Figure 2 shows, the SA polymorphism hypothesis is only one possible explanation for suppressed recombination in sex-determining regions. I have already mentioned that a large fully sex-linked region could evolve without any selection favouring close linkage if a sex-determining gene arises within a region that already had a low recombination rate. Large “recombination deserts” are known in many plant genomes (Charlesworth 2019), and the sex-determining locus is within such a region in a species of the plant *Rumex* (Rifkin et al. 2020). Such a situation would also facilitate the establishment of SA polymorphisms because many genes within such a region will be closely linked to the sex-determining locus.

### *Small non-recombining regions*

At least two other possibilities can create small non-recombining regions, again without selection for loss of recombination. An insertion causing “micro-heteromorphism” (such as the duplication creating the non-paired maleness factor in Figure 5) might directly prevent pairing in the regions flanking the insertion, causing a few neighbouring genes to also become fully Y- or X-linked (pink). Single gene systems in poplar species appear to involve a male-determining gene that, like that in the persimmon, arose as a duplication that silenced a female-promoting gene (Müller et al. 2020; Xue et al. 2020). Small genome regions carrying multiple male-specific genes have been detected in two other plants, *Asparagus officinalis* (Harkess et al. 2020) and the date palm (Torres et al. 2018). A duplication also created a new sex-determining gene in the fish, medaka, *Oryzias latipes* (Kondo et al. 2006). These male-specific regions have no homologous counterpart with which to pair (formally they are hemizygous fully Y-linked regions, but distinct from the hemizygosity caused by genetic degeneration, where genes were lost from a Y-linked region).

Finally, the observation of slight changes in the location of the boundary between the completely and partially sex-linked (pseudo-autosomal, or PAR) regions suggests that small new sex-linked regions may arise in another manner. Two sets of genes (totalling only 17 X-linked genes (Sayres and Makova 2013), have become completely sex-linked genes in humans, but are PAR genes in other mammals, and the PAR boundaries in different mammals also differ slightly (Van Laere et al. 2008; Skinner et al. 2013), including among mouse species (White et al. 2012; Morgan et al. 2019). In plants, similar variation has been found between strains in the plant *Carica papaya* (Lappin et al. 2015), and in *Silene latifolia*, the boundary is not sharply defined, also suggesting a recent change (Krasovec et al. 2020). The reasons for such shifts are unknown, but heterochromatic regions enriched for repetitive sequences that evolved after a recombination-suppressed region evolved, might create micro-heteromorphism or actively inhibit crossing over in large fully sex-linked regions (Phillips and Ihssen 1985; Charlesworth et al. 1994). I am not aware of any explicit model or empirical evidence supporting such an idea, but detailed studies of PAR boundary changes may help understand such changes.

**Conclusions: why do some sex chromosome pairs remain homomorphic?**

I have argued that sex-determining regions need not necessarily evolve into large, multi-gene non-recombining regions, although young sex-linked regions with small Y-X divergence values can sometimes be extensive. Large suppressed recombination regions that still retain many Y- or W-linked genes, particularly those with more than a single stratum of divergence time, support the long-accepted view that the appearance of sex-determining genes has repeatedly triggered such changes in a remarkable diversity of different types of organisms, leading to differentiation. However, very recently evolved systems will initially be homomorphic (unless recombination is suppressed and there has been enough time for repetitive sequence accumulation), and some such chromosome pairs may continue to undergo crossovers.

The evolution of heteromorphism is not expected to be clock-like, or even a monotonic change. Systems old enough to have undergone genetic degeneration may sometimes remain homomorphic, as in the recently discovered case of skinks (Kostmann et al. 2021). The frequency of such cases is not yet clear, nor why heteromorphism is lacking. Some such situations may reflect technical difficulties in detecting differences in physically small chromosomes, and others could be taxa that (for unknown reasons) rarely undergo chromosome rearrangements, or where major repetitive sequence accumulation has not happened, or secondary loss of heteromorphism (see above).

In contrast, genetic degeneration is an inevitable decline, starting when recombination stopped (though different degeneration levels cannot yet be explained in precise quantitative terms).

Although neither the proportion of species where extensive reduced recombination regions evolved in response to the presence of a sex-determining locus, nor the role of selection in such changes (the SA polymorphism hypothesis) is yet fully understood, it is already clear that the times when sex-determining regions stopped recombining explains much of the variation in degeneration levels.

Chromosome rearrangements such as inversions may often be involved in recombination suppression. However, their presence does not demonstrate that suppressed recombination has evolved, as this need not involve rearrangements, and, rearrangements readily spread after recombination is suppressed. Moreover, a Y- or W-linked region's lack of recombination might reflect rearrangements that spread by genetic drift in a small population (reviewed by Ironside 2010; Ponnikas et al. 2018). This cannot explain an over-representation on sex chromosomes unless some further mechanism creates a higher rearrangement input rate on these chromosomes (such as the



remarkable apparent difference in the Dipteran blackflies, see Adler et al. 2016). Unfortunately, comparative tests of whether chromosomes carrying sex-determining loci have a special tendency to subsequently evolve non-recombining regions are hampered by a reporting bias: sex chromosome differences are readily detectable cytologically, without laborious surveys of all chromosomes in multiple individuals of a species. Even work to test whether sex chromosome rearrangement polymorphisms are commoner than autosomal ones is only just beginning (Anderson et al. 2020). Genome sequencing can potentially give unbiased information for such tests. However, such data are not yet easily obtainable in non-model species, as multiple individuals of both sexes are still required to discover sex-specific variants. Cytogenetic information can help choose suitable species for genome sequencing targeted at testing ideas such as the neutrality hypothesis, and classifying sex-determining regions into the types suggested here. In turn, results from the different types should lead to a better understanding of evolutionary changes in sex-determining loci and regions, and their timings.

**Acknowledgements:** I thank funding agencies for support for my research group's work on sex chromosomes in plants and animals, including current support from the European Research Council (grant number 695225 GUPPYSEX) and many colleagues for helpful discussions over many years.

## References

- Adler, P., A. Yildirim, Z. Onder, T. Tasci, O. Duzlu, O. Arslan, A. Ciloglu, B. Sari, N. Parmaksizoglu, and A. Inci. 2016. Rearrangement hotspots in the sex chromosome of the Palearctic black fly *Simulium bergi* (Diptera, Simuliidae). *Comparative Cytogenetics* 10:295–310
- Ahmed, S., J. M. Cock, E. Pessia, R. Luthringer, A. Cormier, M. Robuchon, L. Sterck, A. F. Peters, S. M. Dittami, E. Corre, M. Valero, J. Aury, D. Roze, Y. V. d. Peer, J. Bothwell, G. A. B. Marais, and S. M. Coelho. 2014. A haploid system of sex determination in the brown alga *Ectocarpus* sp. *Curr. Biol.* 24:1945–1957.
- Akagi, T. and D. Charlesworth. 2019. Pleiotropic effects of sex-determining genes in the evolution of dioecy in two plant species. *Proc. Royal Soc. B* 286:20191805.
- Akagi, T., I. M. Henry, R. Tao, and L. Comai. 2014. A Y-chromosome–encoded small RNA acts as a sex determinant in persimmons. *Science* 346:646–650.
- Allen, C. E. 1917. A chromosome difference correlated with sex differences in *Sphaerocarpos*. *Science*

46:466-467.

- Allen, C. E. 1932. Sex-inheritance and sex determination. *Am. Nat.* 66:97-107.
- Anderson, N. W., C. E. Hjelman, and H. Blackmon. 2020. The probability of fusions joining sex chromosomes and autosomes. *Biol. Lett.* 16.
- Bachtrog, D. 2003. Accumulation of Spock and Worf, two novel non-LTR retrotransposons, on the neo-Y chromosome of *Drosophila miranda*. *Mol. Bio. Evol.* 20:173-181.
- Bachtrog, D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* 179:1513-1525.
- Bachtrog, D., E. Hom, K. Wong, X. Maside, and P. d. Jong. 2008. Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biology* 9:R30.
- Bachtrog, D., N. Toda, and S. Lockton. 2010. Dosage compensation and demasculinization of X chromosomes in *Drosophila*. *Curr. Biol.* 20:1476-1481.
- Bazrgar, M., H. Gourabi, M. Valojerdi, P. Yazdi, and H. Baharvand. 2013. Self-correction of chromosomal abnormalities in human preimplantation embryos and embryonic stem cells. *Stem Cells and Development* 22:2449-2456.
- Bellott, D. and D. Page. 2021. Dosage-sensitive functions in embryonic development drove the survival of genes on sex-specific chromosomes in snakes, birds, and mammals. *Genome Res.*
- Bellott, D., H. Skaletsky, T. Pyntikova, E. Mardis, T. Graves, C. Kremitzki, L. Brown, S. Rozen, W. C. Warren, R. K. Wilson, and D. Page. 2010. Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. *Nature* 466:612-616.
- Bergero, R. and D. Charlesworth. 2011. Preservation of the Y transcriptome in a 10MY old plant sex chromosome system. *Curr. Biol.* 21:1470-1474.
- Bergero, R., A. Forrest, E. Kamau, and D. Charlesworth. 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* 175:1945-1954.
- Bergero, R., J. Gardner, B. Bader, L. Yong, and D. Charlesworth. 2019. Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proc. Natl. Acad. Sci. U. S. A.* 116:6924-6931.
- Blaser, O., S. Neuenschwander, and N. Perrin. 2014. Sex-chromosome turnovers: the hot-potato model. *Am. Nat.* 183:140-146.
- Bull, J. J. 1983. *Evolution of Sex Determining Mechanisms*. Benjamin/Cummings, Menlo Park, CA.
- Carey, S. B., A. C. P. J. Jenkins, Shenqiang Shu, John T. Lovell, A. S. F. Maumus, George P. Tiley, Noe Fernandez-Pozo, Kerrie Barry, Cindy Chen, Mei Wang, Anna Lipzen, Chris Daum, Christopher A. Saski, Jordan C. McBreen, Roth E. Conrad, Leslie M. Kollar, Sanna Olsson, Sanna Huttunen, Jacob B. Landis, J. Gordon Burleigh, Norman J. Wickett, Matthew G. Johnson, Stefan A. Rensing, Jane Grimwood, J. Schmutz, and S. F. McDaniel. 2020. Chromosome fusions shape an ancient UV sex chromosome system. *bioRxiv:2020.2007.2003.163634*.

- Chamary, J. V., J. L. Parmley, and L. D. Hurst. 2006. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature Rev. Genet.* 7:98-108.
- Charlesworth, B. and D. Charlesworth. 1978. A model for the evolution of dioecy and gynodioecy. *Am. Nat.* 112:975-997.
- Charlesworth, B., P. Sniegowski, and W. Stephan. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215-220.
- Charlesworth, D. 2016. The status of supergenes in the 21st century: Recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. *Journal of Evolutionary Applications* 9:74-90.
- Charlesworth, D. 2019. Young sex chromosomes in plants and animals. *New Phytol.* 224:1095-1107.
- Chibalina, M. and D. Filatov. 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr. Biol.* 21:1475-1479.
- Connallon, T. and A. G. Clark. 2014. Evolutionary inevitability of sexual antagonism. *Proceedings of the Royal Society B: Biological Sciences* 281:20132123.
- Conte, M., R. Joshi, E. C. Moore, S. Nandamuri, W. Gammerdinger, R. Roberts, K. Carleto, S. Lien, and T. Kocher. 2019. Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *GigaScience* 4:giz030.
- Cortez, D., R. Marin, D. Toledo-Flores, L. Froidevaux, A. Liechti, P. D. Waters, F. Grützner, and H. Kaessmann. 2014. Origins and functional evolution of Y chromosomes across mammals. *Nature* 508:488-493.
- Crowson, D., S. C. H. Barrett, and S. I. Wright. 2017. Purifying and positive selection influence patterns of gene loss and gene expression in the evolution of a plant sex chromosome system. *Mol. Bio. Evol.* 34:1140-1154.
- Disteche, C. 2016. Dosage compensation of the sex chromosomes and autosomes. *Seminars in Cell and Developmental Biology* 56:9-18.
- Dixon, G., J. Kitano, and M. Kirkpatrick. 2018. The origin of a new sex chromosome by introgression between two stickleback fishes. *Molecular Biology and Evolution*, 36:28-38.
- Fraisse, C., M. Picard, and B. Vicoso. 2017. The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. *Nature Communications* 8:1486.
- Fruchard, C., H. Badouin, D. Latrasse, R. S. Devani, A. Muyle, B. Rhoné, S. S. Renner, A. K. Banerjee, A. Bendahmane, and G. A. B. Marais. 2020. Evidence for dosage compensation in *Coccinia grandis*, a plant with a highly heteromorphic XY system. *Genes* 11:787.
- Fry, J. D. 2010. The genomic location of sexually antagonistic variation: some cautionary comments. *Evolution* 64:1510-1516.
- Furman, B., D. Metzger, I. Darolti, A. E. Wright, B. Sandkam, P. Almeida, J. Shu, and J. Mank. 2020. Sex chromosome evolution: so many exceptions to the rules. *Genome Bio. Evol.* 12:750-763.

- Furo, I. d. O., R. Kretschmer, M. D. Santos, C. d. L. Carvalho, R. Gunski, P. O'Brien, M. Ferguson-Smith, M. Cioffi, and E. C. d. Oliveira. 2017. Chromosomal mapping of repetitive DNAs in *Myiopsitta monachus* and *Amazona aestiva* (Psittaciformes, Psittacidae) with emphasis on the sex chromosomes. *Cytogenet. Genome Res.* 15:151-160.
- Gammerdinger, W. and T. Kocher. 2018. Unusual diversity of sex chromosomes in African Cichlid fishes. *Genes* 9:480.
- Grabowska-Joachimciak, A., A. Kula, T. Książczyk, J. Chojnicka, E. Sliwinska, and A. Joachimciak. 2015. Chromosome landmarks and autosome-sex chromosome translocations in *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosome Res.* 23:187-197.
- Gu, L. and J. R. Walters. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and bedeviled by details. *Genome Bio. Evol.* 9: 2461–2476.
- Harkess, A., K. Huang, R. van der Hulst, B. Tissen, J. L. Caplan, A. Koppula, M. Batish, B. C. Meyers, and J. H. Leebens-Mack. 2020. Sex determination by two Y-linked genes in garden asparagus. *Plant Cell* 32: 1790–1796.
- Ieda, R., S. Hosoya, S. Tajima, K. Atsumi, T. Kamiya, A. Nozawa, Y. Aoki, S. Tasumi, T. Koyama, O. Nakamura, S. Suzuki, and K. Kikuchi. 2018. Identification of the sex-determining locus in grass puffer (*Takifugu niphobles*) provides evidence for sex-chromosome turnover in a subset of *Takifugu* species. *PLoS ONE* 13:e0190635.
- Ironside, J. 2010. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *Bioessays* 32:718-726.
- Ishizaki, K., Y. Shimizu-Ueda, S. Okada, M. Yamamoto, M. Fujisawa, K. T. Yamato, H. Fukuzawa, and K. Ohyama. 2002. Multicopy genes uniquely amplified in the Y chromosome-specific repeats of the liverwort *Marchantia polymorpha*. *Nucleic Acids Res.* 30:4675-4681.
- Jordan, C. and D. Charlesworth. 2012. The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* 66:505–516.
- Kamiya, T., W. Kai, S. Tasumi, A. Oka, T. Matsunaga, M. Mizuno, M. Fujita, H. Suetake, S. Suzuki, S. Hosoya et al. 2012. A trans-species missense SNP in *amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genetics* 8:e1002798.
- Kirkpatrick, M. and R. Guerrero. 2014. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197:531-541.
- Kondo, M., U. Hornung, I. Nanda, S. Imai, T. Sasaki, A. Shimizu, S. Asakawa, H. Hori, M. Schmid, N. Shimizu, and M. Schartl. 2006. Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. *Genome Res.* 16:815-826.
- Kostmann, A., L. Kratochvíl, and M. Rovatsos. 2021. Poorly differentiated XX/XY sex chromosomes are widely shared across skink radiation. *Proc. Royal Soc. B* 288:20202139.
- Kramer, M., P. Rao, and S. Ercan. 2016. Untangling the contributions of sex-specific gene regulation and X chromosome dosage to sex-biased gene expression in *C. elegans*. *Genetics* 204:355-.

- Krasovec, M., Y. Zhang, and D. Filatov. 2020. The location of the pseudoautosomal boundary in *Silene latifolia*. *Genes* 11.
- Lahn, B. T. and D. C. Page. 1999. Four evolutionary strata on the human X chromosome. *Science* 286:964-967.
- Lappin, F., C. Medert, K. Hawkins, S. Mardonovich, M. Wu, and R. Moore. 2015. A polymorphic pseudoautosomal boundary in the *Carica papaya* sex chromosomes. *Molecular and Genetical Genomics* 290:1511-1522.
- Lawson-Handley, L. J., H. Ceplitis, and H. Ellegren. 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* 167:367-376.
- Lenormand, T., F. Fyon, E. Sun, and D. Roze. 2020. Sex chromosome degeneration by regulatory evolution. *Curr. Biol.* in press.
- Mahajan, S. and D. Bachtrog. 2017. Convergent evolution of Y chromosome gene content in flies. *Nature Communications* 18:785.
- Mank, J. E., D. J. Hosken, and N. Wedell. 2011. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. *Evolution* 65:2133-2144.
- Mank, J. E. and A. E. Wright. 2012. The scope and strength of sex-specific selection in genome evolution. *J. Evol. Biol.* 26:1841–1853.
- Manna, F., R. Gallet, G. Martin, and T. Lenormand. 2012. The high-throughput yeast deletion fitness data and the theories of dominance. *J. Evol. Biol.* 25:892-903.
- Marks, R., J. Smith, Q. Cronk, C. Grassa, and D. McLetchie. 2019. Genome of the tropical plant *Marchantia inflexa*: implications for sex chromosome evolution and dehydration tolerance. *Scientific Reports* 9:8722.
- Matsubara, K., S. Sarre, A. Georges, J. Graves, Y. Matsuda, and T. Ezaz. 2014. Highly differentiated ZW sex microchromosomes in the Australian *Varanus* species evolved through rapid amplification of repetitive sequences. *PLOS ONE* 9:e95226.
- McDaniel, S. F., J. H. Willis, and A. J. Shaw. 2007. A linkage map reveals a complex basis for segregation distortion in an interpopulation cross in the moss *Ceratodon purpureus*. *Genetics* 176:2489-2500.
- Meisel, R. P., P. U. Olafson, K. Adhikari, F. D. Guerrero, K. Konganti, and J. B. Benoit. 2020. Sex chromosome evolution in Muscid flies. *G3: Genes, Genomes, Genetics* 10:1341–1352.
- Morgan, A. P., T.A. Bell, J. J. Crowley, and F. P.-M. d. Villena. 2019. Instability of the pseudoautosomal boundary in house mice. *Genetics* 212:469-487.
- Müller, N., B. Kersten, A. Montalvão, N. Mähler, C. Bernhardsson, K. Bräutigam, Z. Lorenzo, H. Hoenicka, V. Kumar, M. Mader, B. Pakull, K. M. Robinson, M. Sabatti, C. Vettori, P. Ingvarsson, Q. Cronk, N. Street, and M. Fladung. 2020. A single gene underlies the dynamic evolution of poplar sex determination.
- Myosho, T., Y. Takehana, S. Hamaguchi, and M. Sakaizumi. 2015. Turnover of sex chromosomes in

- celebensis group medaka fishes. *G3-Genes Genomes Genetics* 5:2685-2691.
- Natri, H. M., T. Shikano, and J. Merilä. 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Mol. Bio. Evol.* 30:1131-1144.
- Nóbrega, M. A., Y. Zhu, I. Plajzer-Frick, V. Afzal, and E. M. Rubin. 2004. Megabase deletions of gene deserts result in viable mice. *Nature* 431:988-993.
- Okada, S., T. Sone, M. Fujisawa, S. Nakayama, M. Takenaka, K. Ishizaki, K. Kono, Y. Shimizu-Ueda, T. Hanajiri, K. T. Yamato, H. Fukuzawa, A. Brennicke, and K. Ohyama. 2001. The Y chromosome in the liverwort *Marchantia polymorpha* has accumulated unique repeat sequences harboring a male-specific gene. *Proc. Natl. Acad. Sci. U. S. A.* 98:9454-9459.
- Palacios-Gimenez, O. M., J. Koelman, M. Palmada-Flores, T. M. Bradford, K.K. Jones, S. J. B. Cooper, T. Kawakami, and A. Suh. 2020. Comparative analysis of morabine grasshopper genomes reveals highly abundant transposable elements and rapidly proliferating satellite DNA repeats. *BMC Biology* 18:199.
- Pamilo, P., M. Nei, and W.-H. Li. 1987. Accumulation of mutations in sexual and asexual populations. *Genet. Res. Camb.* 43:135-146.
- Papadopoulos, A. S. T., M. Chester, K. Ridout, and D. A. Filatov. 2015. Rapid Y degeneration and dosage compensation in plant sex chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 112:13021–13026.
- Parmley, J. L., J. V. Chamary, and L. D. Hurst. 2006. Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. *Mol. Bio. Evol.* 23:301-309.
- Peichel, C., S. McCann, J. Ross, A. Naftaly, J. Urton, J. Cech, J. Grimwood, J. Schmutz, R. Myers, D. Kingsley, and D. White. 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biology* 21:177.
- Phillips, R. and P. Ihssen. 1985. Identification of sex chromosomes in lake trout (*Salvelinus namaycush*). *Cytogenet. Cell Genet.* 39:14–18.
- Picard, M., C. Cosseau, S. Ferré, T. Quack, C. Grevelding, Y. Couté, and B. Vicoso. 2018. Evolution of gene dosage on the Z-chromosome of Schistosome parasites. *elife*:e35684.
- Pigozzi, M. 2011. Diverse stages of sex-chromosome differentiation in tinamid birds: evidence from crossover analysis in *Eudromia elegans* and *Crypturellus tataupa*. *Genetica* 139:771-777.
- Ponnikas, S., H. Sigeman, J. Abbott, and B. Hansson. 2018. Why do sex chromosomes stop recombining? *Trends Genet.* 34:492-503.
- Post, R. 1985. Sex-chromosome evolution in *Simulium erythrocephalum* (Diptera, Simuliidae). *Heredity* 54:149-158.
- Prentout, D., O. Razumova, B. Rhoné, H. Badouin, H. Henri, C. Feng, J. Käfer, G. Karlov, and G. A. B. Marais. 2020. An efficient RNA-seq-based segregation analysis identifies the sex chromosomes of *Cannabis sativa*. *Genome Res.* 164-172.
- Rifkin, J., F. Beaudry, Z. Humphries, B. Choudhury, S. Wright, and S. Barrett. 2020. Widespread recombination suppression facilitates plant sex chromosome evolution. *Mol. Bio. Evol.* in

press.

- Rodrigues, N., T. Studer, C. Dufresnes, and N. Perrin. 2018. Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Mol. Bio. Evol.*
- Rutkowska, J., M. Lagisz, and S. Nakagawa. 2012. The long and the short of W chromosomes: no evidence for avian W degeneration. *Biol. Lett.* 8:636-638.
- Sandstedt, S. A. and P. K. Tucker. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Res.* 14:267-272.
- Sardell, J. and M. Kirkpatrick. 2019. Sex differences in the recombination landscape. *Am. Nat.* 195:361–379.
- Sayres, M. and K. Makova. 2013. Gene survival and death on the human Y chromosome. *Mol. Bio. Evol.* 30:781-787.
- Schild, D. R., D. C. Card, N. R. Hales, B. W. Perry, G. M. Pasquesi, H. Blackmon, R. H. Adams, A. B. Corbin, C. F. Smith, B. Ramesh, J. P. Demuth, E. Betrán, M. Tollis, J. M. Meik, S. P. Mackessy, and T. A. Castoe. 2019. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. *Genome Res.* 29:590-601.
- Schmid, M., J. Smith, D. Burt, B. Aken, P. Antin, A. Archibald, C. Ashwell, P. Blackshear, C. Boschiero, C. Brown et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825 - 837.
- Skinner, B., K. Lachani, CA Sargent, and N. Affara. 2013. Regions of XY homology in the pig X chromosome and the boundary of the pseudoautosomal region. *BMC Genetics* 14:1-7.
- Swanson, C. P. 1957. *Cytology and Cytogenetics*. Prentice Hall, New York.
- Tennessen, J. A., N. Wei, S. C. K. Straub, R. Govindarajulu, A. Liston, and T. Ashman. 2018. Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLoS Biol.*
- Torres, M., L. S. Mathew, I. Ahmed, I. K. Al-Azwan, R. Krueger, D. Rivera-Núñez, Y. A. Mohamoud, A. G. Clark, K. Suhre, and J. A. Malek. 2018. Genus-wide sequencing supports a two-locus model for sex-determination in Phoenix. *Nature Communications* 9:3969.
- van Doorn, G. and M. Kirkpatrick. 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature* 449:909-912.
- Van Laere, A.-S., W. Coppieters, and M. Georges. 2008. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. *Genome Res.* 18:1884-1895.
- Varadharajan, S., P. Rastas, A. Löytynoja, M. Matschiner, F. C. Calboli, B. Guo, A. J. Nederbragt, K. Jakobsen, and J. Merilä. 2019. A high-quality assembly of the nine-spined stickleback (*Pungitius pungitius*) genome. *Genome Bio. Evol.* 11:3291-3308.
- Veltsos, P., K. Ridout, M. Touns, S. González-Martínez, A. Muyle, O. Emery, P. Rastas, V. Hudzieczek, B. Vyskot, G. Marais, D. Filatov, and J. Pannell. 2019. Early sex-chromosome evolution in the diploid dioecious plant *Mercurialis annua*. *Genetics* 212:815-835.

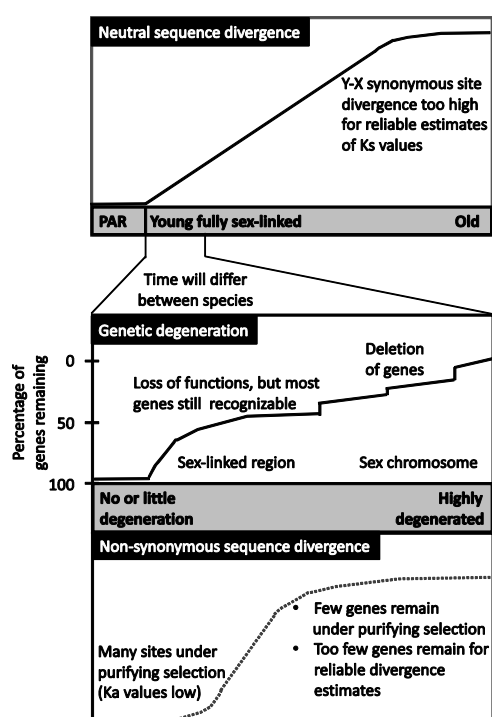
- Vicoso, B. 2019. Molecular and evolutionary dynamics of animal sex-chromosome turnover. *Nature Ecology and Evolution* 3:632–1641.
- Vicoso, B. and D. Bachtrog. 2011. Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite. *Genome Bio. Evol.* 3:230-235.
- Vicoso, B. and D. Bachtrog. 2015. Numerous transitions of sex chromosomes in Diptera. *PLoS Biol.* 13:e1002078.
- Vicoso, B. and B. Charlesworth. 2006. Evolution on the X chromosome: unusual patterns and processes. *Nature Rev. Genet.* 7:645-653.
- Vicoso, B., V. Kaiser, and D. Bachtrog. 2013. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proc. Natl. Acad. Sci. U. S. A.* 110:6453-6458.
- Walsh, I. M., M. A. Bowman, I. F. S. Santarriag, A. Rodriguez, and P. L. Clark. 2020. Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness. *Proc. Soc. Exp. Biol. Med.* 117:3528-3534.
- Wang, J., J. Na, Q. Yu, A. R. Gschwend, J. Han, F. Zeng, R. Aryal, R. VanBuren, J. E. Murray, W. Zhang et al. 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc. Natl. Acad. Sci. U. S. A.* 109:13710-13715.
- Wei, K. H.-C. and D. Bachtrog. 2019. Ancestral male recombination in *Drosophila albomicans* produced geographically restricted neo-Y chromosome haplotypes varying in age and onset of decay. *PLoS Genetics* 15:e1008502.
- Westergaard, M. 1958. The mechanism of sex determination in dioecious plants. *Adv. Genet.* 9:217-281.
- White, M. A., A. Ikeda, and B. A. Payseur. 2012. A pronounced evolutionary shift of the pseudoautosomal region boundary in house mice. *Mammalian Genome* 23:454-466.
- Wu, M. and P. H. Moore. 2015. The evolutionary tempo of sex chromosome degradation in *Carica papaya*. *J. Mol. Evol.* 80:265-277.
- Xu, L., S. Sin, P. Grayson, S. Edwards, and T. Sackton. 2019. Evolutionary dynamics of sex chromosomes of Paleognathous birds. *Genome Bio. Evol.* 11:2376–2390.
- Xue, L., H. Wu, Y. Chen, X. Li, J. Hou, J. Lu, S. Wei, X. Dai, M. S. Olson, J. Liu, M. Wan, D. Charlesworth, and T. Yin. 2020. Evidences for a role of two Y-specific genes in sex determination in *Populus deltoides*. *Nature Communications* 11:5893.
- Yang, W., D. Wang, Y. Li, Z. Zhang, S. Tong, M. Li, X. Zhang, L. Zhang, L. Ren, X. Ma, R. Zhou, B. J. Sanderson, K. Keefover-Rin, T. Yin, L. B. Smart, J. Liu, S. P. DiFazio, M. Olson, and T. Ma. 2020. A general model to explain repeated turnovers of sex determination in the Salicaceae. *Mol. Bio. Evol.*
- Zhou, Q. and D. Bachtrog. 2015. Ancestral chromatin configuration constrains chromatin evolution on differentiating sex chromosomes in *Drosophila*. *Plos Genetics* 11:e1005331.



- Zhou, Q., J. Zhang, D. Bachtrog, N. An, Q. Huang, E. D. Jarvis, M. T. P. Gilbert, and G. Zhang. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* 346:1246338.
- Zhou, R., D. Macaya-Sanz, C. Carlson, J. Schmutz, J. W. Jenkins, D. Kudrna, A. Sharma, L. Sandor, S. Shu, K. Barry, G. A. Tuskan, T. Ma, J. Liu, M. Olson, L. B. Smart, and S. P. DiFazio. 2020. A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Genome Biology* 21:38.
- Zrzava, M., I. Irena, M. Dalikova, J. Sichova, E. Ounap, S. Kubickova, and F. Marec. 2018. Sex chromosomes of the iconic moth *Abraxas grossulariata* (Lepidoptera, Geometridae) and its congener *A. sylvata*. *Genes* 9:UNSP 279.

## Figure legends

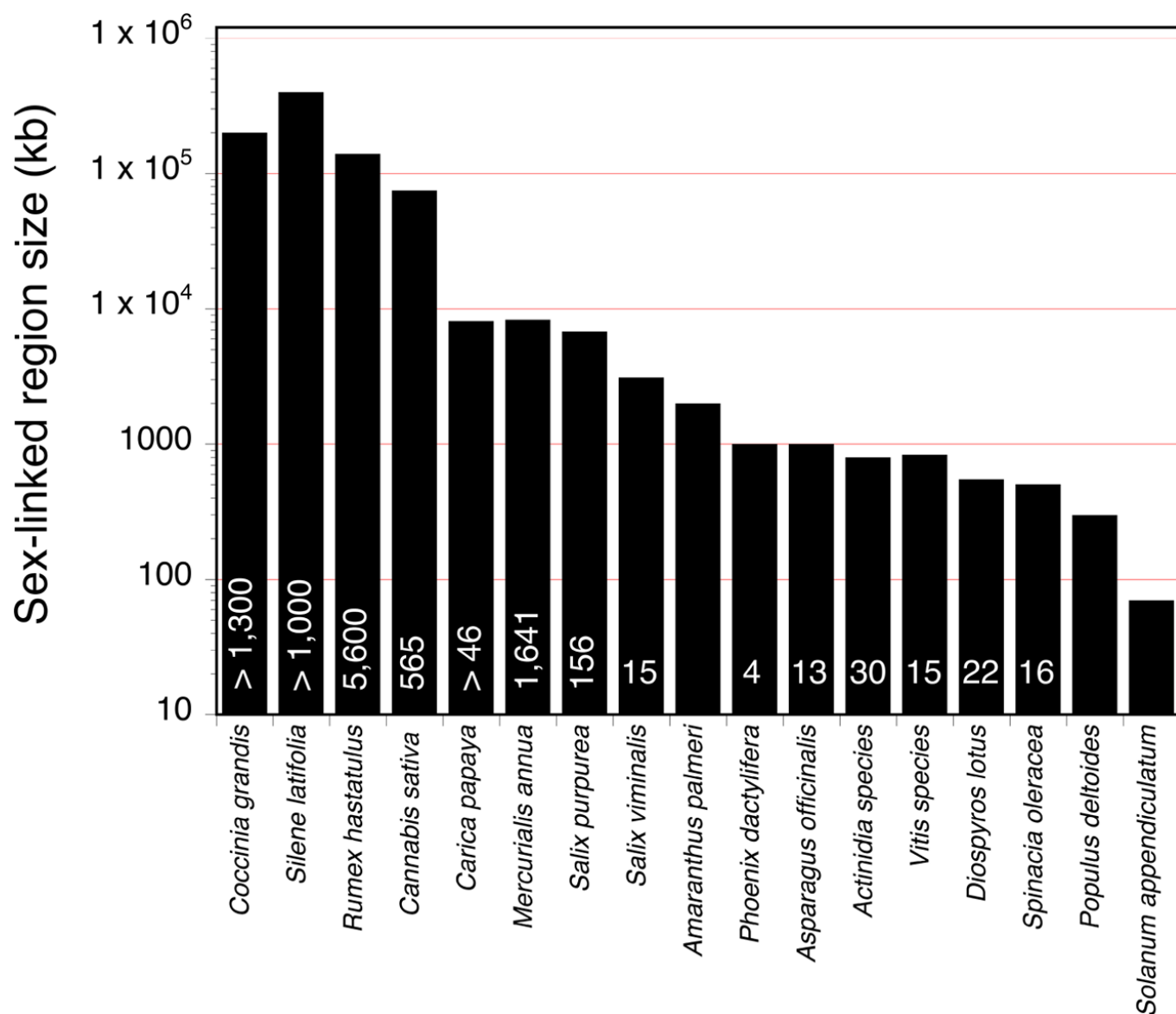
**Figure 1.** Schematic diagram to illustrate the stages of evolution of a completely sex-linked region after recombination stops, in cases where it does stop. The left-hand end is a pseudo-autosomal region (PAR that continues to recombine). The absence of numbers on the x axis, and the y axis for divergence, is deliberate, as the diagrams are intended to show general concepts. Specific examples, and measures of divergence and degeneration, are described in the text. Importantly, the time-scales for different processes differ (as indicated by the lines connecting the short segment of time corresponding to neutral divergence and the degeneration process); specifically, as drawn, loss of selective constraints is shown occurring in an evolutionary time corresponding to small, rather than large, neutral divergence. and complete degeneration might follow shortly after this stage, or might take much longer.



**Figure 2:** Scheme to show the different ways in which a small region carrying the sex determining gene or genes can arise, and the different situations where these regions could expand in size to form extensive fully sex-linked regions that, given enough time, would be predicted to evolve the complete set of sex chromosome properties (divergence of Y- or W-linked sequences, accumulation of repetitive sequences, chromosome rearrangements and genetic degeneration). In the two situations involving polymorphisms in two genes that create selection for closer linkage (right-hand columns) an inversion could suppress recombination across a physically large region, or “evolutionary stratum”, that includes many genes without sex-related functions.

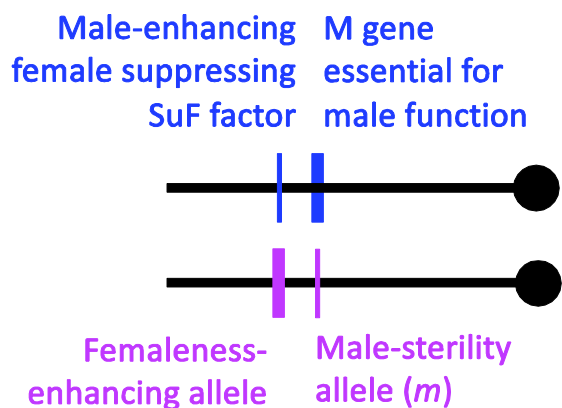
Situations producing new sex determining regions		Direct effect on recombination	Creation of large non-recombining regions		
			Without selection	Possible changes due to selection	
De novo evolution of separate sexes	2 genes (Figure 4A shows a situation that arises in plants)	None	Evolution of sex-determining genes within an already non-recombining region, potentially large (e.g. a pericentromeric region)	Recombinants between male-sterility and female-suppressing genes are disfavored	Sexually antagonistic polymorphisms may establish in the region, selecting for suppressed recombination
	Single gene system evolves (Figure 5, shows a plant example)				
Turnover events (see Figure 2B)	New sex-determining gene arises			Creates a small hemizygous non-recombining region (see Figure 5)	
	Duplication of pre-existing gene into new genomic location				

**Figure 3.** Size estimates of sex-determining regions from flowering plants. The estimates from different species are rough, and are based on different approaches (see main text), and the smallest estimates are shown, where possible (see the text for an explanation of the problem of size over-estimates); in species with large, heteromorphic XY pairs, and in *C. papaya*, which shows micro-heteromorphism, this estimate refers to the X-linked haplotype. The numbers on the bars indicate rough estimated numbers of genes in each species' sex-determining region, when the number has been estimated.

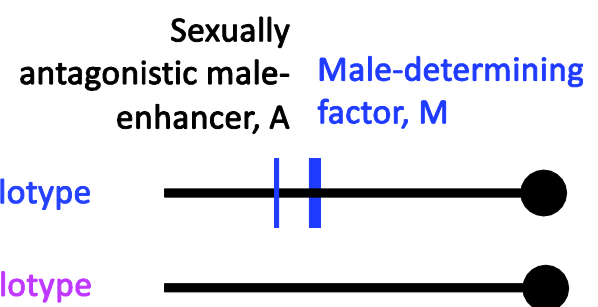


**Figure 4:** Situations where selection favours recombination suppression. **A.** When the ancestral population is cosexual, females can arise by mutations in a male-function gene ( $M \rightarrow m$ ), and males by a femaleness-suppressing mutation ( $F \rightarrow SuF$ ) in the same genome region; the Y haplotype will then carry  $M$  and  $SuF$ . This is a “supergene” model for sex-determination. **B.** In a situation after a turnover event that creates a new maleness ( $M$ ) factor in an ancestral population of males and females, the  $M$  factor defines a Y haplotype, and a sexually antagonistic maleness-enhancing factor,  $A$ , establishes a polymorphism in the same genome region. **C.** In either case, recombinants between the two haplotypes will have disfavoured combinations of alleles at the two loci, and a non-recombining region (possibly extensive) could evolve.

**A** Male in a species evolved from a hermaphrodite or monoecious ancestor



**B** Male in a species with a single male-determining gene



**C** Completely sex-linked region evolves



**Figure 5:** Scenario in which mutations in two genes produce a single gene sex-determining system, based on the situation in the persimmon (Akagi et al. 2014). In the ancestral population a gene (shown as pink outlined symbols) promotes maleness by permitting expression of another gene necessary for some male function, such as an anther function. A first mutation in the first gene (pink filled symbols) causes femaleness by inhibiting this gene's expression. A duplication of the first gene then occurs (blue arrow), causing suppression of expression of either allele of the first gene (symbolized by lines connected to the femaleness gene and the standard symbol for inhibition of gene expression). The male function is therefore not inhibited and these carriers develop as males. Selection for a 1:1 sex ratio will lead to the pink allele replacing the ancestral allele, as shown at the bottom right of the figure, leaving a population in which sex is controlled solely by the presence or absence of the duplicate gene, which can thus be termed Y-linked. The duplication could be on a different chromosome from the first gene (as in the persimmon), or at a different location on the same chromosome (as shown in the figure). In either case, it could prevent recombination in the region (as symbolized by the blue unpaired Y-linked region in the diagram).

