Tansley review

Plant contributions to our understanding of sex chromosome evolution

Deborah Charlesworth
Institute of Evolutionary Biology, University of Edinburgh, Ashworth Lab, King’s Buildings, W. Mains Road, Edinburgh, EH9 3FL, UK

Summary

A minority of angiosperms have male and female flowers separated in distinct individuals (dioecy), and most dioecious plants do not have cytologically different (heteromorphic) sex chromosomes. Plants nevertheless have several advantages for the study of sex chromosome evolution, as genetic sex determination has evolved repeatedly and is often absent in close relatives. I review sex-determining regions in non-model plant species, which may help us to understand when and how (and, potentially, test hypotheses about why) recombination suppression evolves within young sex chromosomes. I emphasize high-throughput sequencing approaches that are increasingly being applied to plants to test for non-recombining regions. These data are particularly illuminating when combined with sequence data that allow phylogenetic analyses, and estimates of when these regions evolved. Together with comparative genetic mapping, this has revealed that sex-determining loci and sex-linked regions evolved independently in many plant lineages, sometimes in closely related dioecious species, and often within the past few million years. In reviewing recent progress, I suggest areas for future work, such as the use of phylogenies to allow the informed choice of outgroup species suitable for inferring the directions of changes, including testing whether Y chromosome-like regions are undergoing genetic degeneration, a predicted consequence of losing recombination.
I. Introduction: advantages of plants for the study of sex chromosome evolution

Most flowering plants have hermaphroditic flowers, and only a minority have separate male and female flowers (monoecy or dioecy). Among dioecious plants, with male and female flowers separated in distinct individuals, some species have environmental, not genetic, control of sex determination (Policansky, 1981; Zimmerman, 1991; Pannell, 1997), and those with genetic sex determination often do not have cytologically differentiated sex chromosomes (Westergaard, 1958; Ming et al., 2011; Renner, 2014). By contrast, separate sexes and heteromorphic sex chromosomes are common in many familiar animal groups (Bachtrog, 2012). Nevertheless, plants have several advantages for research on sex chromosomes, because genetic sex determination has evolved repeatedly among angiosperms, and independently in different families (Charlesworth, 1985; Ming et al., 2011; Renner, 2014). Compared with the best-studied animal systems (Bellott et al., 2014; Cortez et al., 2014; Zhou et al., 2014), many flowering plant sex chromosomes probably evolved very recently (Marais et al., 2011; Renner, 2014); yet, as will be illustrated below, similarities with animal systems are striking. Table 1 summarizes the main advantages of using dioecious plants to study sex chromosome evolution, and to test hypotheses about sex chromosome evolution derived from theoretical modelling.

There is now too much published work on plant sex chromosomes and their evolution to include in a single article. I therefore focus on recombination suppression, the defining characteristic of sex chromosomes, which leads to the evolution of the other unusual characteristics of sex chromosomes, genetic degeneration and accumulation of repetitive sequences on the sex chromosomes, which I mention only briefly. I review progress that has come through genetic and molecular evolutionary studies, illustrating how this has involved the combination of approaches, including DNA sequencing, and resequencing of multiple individuals of the same sex and species for genetic and population genetic tests of sex linkage, together with sequencing for phylogenetic studies, and place events in well-established time frames.

II. Sex chromosomes and estimation of their ages from sequence divergence

I define plant sex chromosomes as genome regions of these species that carry the ‘SEX’ locus that controls the sexes of individuals, and that do not recombine. Rather than using the term sex chromosomes, I shall often use ‘fully sex-linked regions’ (and full sex linkage), because of the diversity among plants with genetically controlled dioecy — some have extensive non-pairing regions that show heteromorphism between the sexes, like many animal sex chromosomes, but many have no detectable cytological differences (recently reviewed by Renner, 2014). Silene latifolia is an example of sex chromosome heteromorphism. It has an XY system, and males are the heterozygous sex, as in mammals (Bellott et al., 2014; Cortez et al., 2014). The Y is largely non-recombining, with XY pairing only in a small pseudo-autosomal region (PAR) at one tip (Westergaard, 1958; Ming et al., 2011; Renner, 2014).

### Table 1 Some characteristics favourable for the study of the genetics and evolution of plant sex chromosomes (the main text provides examples from plant studies)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Advantages</th>
<th>Specific evolutionary questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A range of sex chromosome ages exists, including recently evolved and old established systems</td>
<td>The time at which recombination stopped can be estimated using a molecular clock, as it is often not long, and sequence differences will not be saturated, but will reflect times at which recombination was suppressed</td>
<td>1. Which species without cytologically visible heteromorphism have sex-linked regions that include genes other than the sex-determining genes? 2. Does recombination suppression always evolve, even in old established systems, or does it sometimes fail to evolve (and, if so, why)?</td>
</tr>
<tr>
<td></td>
<td>The earliest stages of sex chromosome evolution can be studied</td>
<td>1. Is there a tendency for chromosomal heteromorphism, heterochromatinization, ZW systems and X-autosome balance systems to be associated with older established systems? 2. How did recombination suppression evolve (gradually, or in distinct recombination suppression events affecting genome regions with many genes), and how often do such events happen? 3. Did repetitive sequences accumulate before genes started to lose functions, or does their accumulation contribute to loss of functions?</td>
</tr>
<tr>
<td>Closely related non-dioecious outgroup species often exist</td>
<td>The directions of changes during sex chromosome evolution can be studied</td>
<td>1. Have plant X and/or Y chromosomes adapted to the new dioecious state? 2. Have plant Y chromosomes degenerated genetically? If so, what is the time course?</td>
</tr>
<tr>
<td>Dioecy evolved repeatedly</td>
<td>The phylogenetic context is often available, so that the directions of changes during sex chromosome evolution can be studied</td>
<td>1. Have plant X and/or Y chromosomes adapted to the new dioecious state? 2. Have plant Y chromosomes degenerated genetically? If so, what is the time course?</td>
</tr>
</tbody>
</table>

© 2015 The Authors
New Phytologist © 2015 New Phytologist Trust

www.newphytologist.com
Filatov *et al.*, 2008); mapping of genic markers suggests a single PAR (Bergero *et al.*, 2013), although an amplified fragment length polymorphism (AFLP) map suggests two (Scotti & Delph, 2006). However, unlike many animal Y chromosomes, the fully sex-linked region still carries hundreds of genes (Bergero & Charlesworth, 2011; Chibalina & Filatov, 2011; Muyle *et al.*, 2012). By contrast, the fully Y-linked region in papaya is only about 10% of chromosome 1 (Liu *et al.*, 2004; Wang *et al.*, 2012). Some diploid plants have ZW systems, in which females are ZW heterozygotes and males are ZZ homozygotes (Westergaard, 1958), as in birds (Zhou *et al.*, 2014) and Lepidoptera (Suetsugu *et al.*, 2013); these include *Fragaria* (strawberry) species (Spigler *et al.*, 2008; Goldberg *et al.*, 2010) and *Salix* (Alstrom-Rapaport *et al.*, 1998). Other systems, including those in haploid plants, will be described below.

The time at which recombination stopped can be estimated using DNA sequence divergence between genes present on the Y as well as the X chromosome, together with a ‘molecular clock’ for synonymous or silent site divergence per year. Higher X–Y divergence values correspond to greater times since recombination suppression. In both humans (Lahn & Page, 1999) and the plant *S. latifolia* (Bergero *et al.*, 2007), divergence increases with the distance from the PAR (in X chromosome genetic or physical maps; as these Y chromosomes are extensively rearranged, distances on the Y chromosome are not informative; Skaletsky *et al.*, 2003; Bergero *et al.*, 2008). Therefore, suppressed recombination must have spread from an early non-recombining region, the oldest ‘evolutionary stratum’ (Lahn & Page, 1999), towards younger ‘strata’ closer to the current PAR. X–Y divergence in the older *S. latifolia* stratum is similar to that in the youngest of the five strata in humans (Skaletsky *et al.*, 2003), and the *Silene* XY pair probably evolved c. 5–10 million yr ago (MYA; Nicolas *et al.*, 2005).

A sex chromosome system may be older than its oldest stratum, because recombination suppression in a sex-determining region usually takes time to evolve (Section VIII). However, recombination suppression may pre-date the evolution of separate sexes. In several well-studied plants, results from the combination of genetic and physical mapping reveal large genome regions with infrequent crossing over the surrounding centromeres, with crossovers restricted to the ends of chromosomes, for example, in maize (Rodgers-Melnicka *et al.*, 2015). These regions may include substantial proportions of genes; in barley, for example, c. 20% of genes are estimated to be located in such regions (Baker *et al.*, 2014). If sex-determining loci evolve in such a region, the oldest stratum will be contemporaneous with the sex-determination system (Fig. 1).

In what follows, I stress the importance of estimating the ages of non-recombining regions in order to understand several important aspects of sex chromosome evolution. Young sex chromosome systems are well suited for the study of the early stages of evolution of recombination suppression and the evolution of these characteristics; in older animal systems, these processes can only be studied over a coarse timescale that cannot reveal much detail. Young evolutionary strata in plant sex chromosomes are also of interest for the study of the time course of genetic degeneration, including gene losses from Y chromosomes.

### III. Which plants have sex chromosomes?

Genetic maps can detect the presence of sex-linked regions in dioecious species. In papaya, for example, a large set of AFLP molecular variants was first mapped in a full-sib family (Liu *et al.*, 2004). The completely sex-linked region is small, making bacterial artificial chromosome (BAC) sequencing of the region possible, which showed that the X-linked region includes only 3.5 Mb (it is flanked by much larger PARs), and carries c. 50 genes with apparently functional Y-linked copies (Wang *et al.*, 2012). Assembly of the physical map of the homologous Y-like region suggests that part of the sex-linked region is probably in a pericentromeric region (Yu *et al.*, 2007; Zhang *et al.*, 2008). Silent site divergence values for XY pairs suggest that inversions occurred, suppressing recombination, c. 7 and 1.7 MYA, implying that a new
recombination-suppressed region has formed since the older stratum evolved (Fig. 1).

Ideally, sex linkage in a family should be confirmed by showing that Y-linked variants are found only in males in a wider sample of genotypes (from natural population samples, or from multiple cultivars of crop species) to exclude partially sex-linked genes in the PAR that did not yield recombinants in the particular cross studied. This is unnecessary for papaya because, although this is clearly a young system, the sequence divergence across part of the X–Y region is c. 7% for silent sites, much higher than between alleles in recombining regions of the genome, including the collinear regions adjoining the fully sex-linked sequences (Wang et al., 2012). This strongly suggests complete sex linkage.

Now that large numbers of genetic markers can be developed in non-model organisms, using high-throughput approaches, it will be possible to discover how many other plant cases like papaya exist, without major cytologically detectable sex chromosome heteromorphism, but with fully sex-linked regions carrying multiple genes, and to assess how many plants evolved dioecy so recently that their sex-determining loci have not yet evolved non-recombining sex chromosome-like regions.

I next outline other approaches that have demonstrated that sex-linked regions have evolved in dioecious plants.

When divergence data are not available, genetic mapping in related species can help test whether recombination suppression has evolved in dioecious plants. If the SEX locus of a dioecious species is in a genome region of suppressed recombination, but the homologous region of a non-dioecious relative recombines, this would suggest that recombination suppression evolved following the evolution of genetically controlled dioecy, rather than being the ancestral state (unless the SEX locus is in a pericentromeric region whose extent or location has changed between the species). A dioecious close relative of papaya, Vasconcellea parviflora, has been shown to have a homologous SEX locus, based on cytogenetic detection of heterochromatin in the centromere-proximal regions of the homologous chromosomes of the two species. This result also shows that the papaya and V. parviflora sex chromosomes are not truly homomorphic (Iovene et al., 2015) – their heteromorphism is minor, but detectable with refined modern cytological methods, consistent with the sequencing results showing that the papaya Y-linked region is larger than the X region (Wang et al., 2012).

These examples illustrate the value of genetic mapping within families. However, studies of wider population samples can also be used to discover sex-linked regions and to establish which is the SEX region, nor the age of this system, has yet been estimated. It is a heteromorphic XY system (Siljak-Yakovlev et al., 1996) that may be ancient, as many other species in the palm family are also dioecious (Renner, 2014).

IV. Haploid plants

Many haploid plants have sex chromosomes, as determined either from morphology differences in the karyotypes of male and female gametophytes (Bull, 1983; Ming et al., 2011), or from the existence of one or more sex-linked genetic markers (Immler & Otto, 2015). The male- and female-determining chromosomes of haploids are now often called V and U, respectively (Bachtrog et al., 2011), to emphasize that the SEX region is never homozygous and can therefore never recombine; until physical maps are produced, it is not possible to relate sequence divergence to the genetic map position in the sex-linked region. The older literature, including studies in Marchantia polymorpha, called them X/Y systems (Okada et al., 2001; Yamato et al., 2007). Marchantia polymorpha has highly heteromorphic sex chromosomes, and divergence between alleles of the few sex-linked gene pairs studied is extremely high, indicating an ancient system.

By contrast, genetic mapping in the moss Ceratodon purpureus that, as in papaya, only markers in the middle of the linkage group with the SEX locus show full sex linkage (McDaniel et al., 2007). Sequence data for all site types in coding plus (predominantly) non-coding regions of U- and V-linked allele sequences suggest that evolutionary strata may exist in C. purpureus. Divergence between four of eight U–V gene pairs studied is only c. 1–3%, but two genes have divergence of almost 7% (McDaniel et al., 2013); one is long enough to reliably suggest high divergence, either indicating a longer time since recombination stopped between the U and V regions, or a higher mutation rate and/or lesser selective constraint (divergence from the related species is also high for the high U–V divergence gene, consistent with the last two possibilities). Even the highest silent site divergence currently found suggests, however, that this is not an ancient system (although genes with higher divergence may be discovered when more genes are analysed).

V. Plants with very small, or no, non-recombining regions

Genetic mapping (or related methods that can detect such regions even if they are small, such as bulk segregant analysis) has yet to be applied in many plants with genetic sex determination, and could reveal non-recombining regions in many plants not currently classified as having sex chromosomes. Indeed, a major currently unanswered question is whether the number of plants with sex-linked regions is currently under-estimated. Such studies are, however, limited by marker density, and very small non-recombining regions may be missed as a result of insufficient marker density. Indeed, in several plants, genetic sex determination has been established, and a SEX locus controlling gender has been mapped, but no fully sex-linked marker has been found. In kiwifruit (Actinidia chinensis), for example, mapping with 644 microsatellite markers still failed to detect any fully sex-linked markers. Other species in which the recombination status of the SEX locus is currently uncertain include spinach (Khattak et al., 2006), asparagus (Telgmann-Rauber et al., 2007) and Populus species (Yin et al., 2008; Pakull et al., 2011). Such species may, of
course, truly lack non-recombining SEX regions. They may either not yet have evolved fully sex-linked regions, or may be single-gene systems, which can evolve when a new gene takes over control of flower sex determination after dioecy has become established, replacing an existing sex-determining gene (Bull, 1983; van Doorn & Kirkpatrick, 2007; Vuilleumier et al., 2007; Blaser et al., 2013); Fig. 1(c) illustrates such an event. Takeovers are known in several animal taxa, including insects (Wilkins, 1995; Beye et al., 2003) and fish (Ross et al., 2009; Myoshio et al., 2012).

To map SEX loci, high-throughput methods, including RAD-Seq (Baird et al., 2008) or RNA-Seq transcriptome sequencing (Bergero & Charlesworth, 2011; Chibalina & Filatov, 2011; Muyle et al., 2012; Hough et al., 2014), can now generate large numbers of markers, overcoming the problem of low marker density. Alternatively, given the large resources sometimes available in crop plants, high-density linkage maps can be obtained by ‘target-sequence capture’ (Mamanova et al., 2010). In Fragaria vesca, for example, a map was made by first obtaining a low coverage genome sequence, then identifying polymorphisms, and then using an enrichment approach to obtain short sequences (200 bp) surrounding each polymorphism. This allowed the genotyping of 5417 genes in a mapping family (Tennesen et al., 2013). Another recently developed approach that may be helpful in plants ascertainment fully sex-linked sequences by searching short read genome sequence data from multiple individuals for k-mers (short sequences of length k) that appear only in one sex (Carvalho & Clark, 2013). In the section on sex-determining genes below, I outline how this approach has ascertained Y-linked sequences in persimmon (Diospyros lotus) in the Ebenaceae (Akagi et al., 2014).

A major difference between ancient animal sex chromosomes and the Y-linked regions of the plants just discussed (with the possible exception of M. polymorpha) is the minor extent of gene loss in plants (see section on genetic degeneration below). In plants, sex-linked genes therefore cannot be ascertained by genome sequencing, using their lower coverage in the heterogametic sex. Moreover, the assembly of short-read sequences will be difficult, because of sequence divergence and the accumulation of repetitive sequences, which occur in non-recombining genome regions (Charlesworth et al., 1994), including plant sex chromosomes (Kubat et al., 2014). Assemblies of the human and papaya Y chromosomes involved deep sequencing of single males, avoiding variants that might confuse assembly, and are restricted to non-heterochromatic regions (Hughes et al., 2010; Wang et al., 2012). Finally, multiple individuals of each sex are needed to determine sex-specific sequences and to distinguish fully and partially sex-linked regions. Even with large family sizes, lack of recombinants does not exclude rare recombination. Population genetic approaches can, however, detect recombination in past generations, even if it occurs very rarely. In papaya, for example, the SEX region adjoins a ‘collinear region’ in which the sequenced X and Y chromosomes appear to have the same genes in the same order, unlike the older Y-linked strata, whose assembly includes many rearrangements (Wang et al., 2012). Divergence between the single X and Y sequences so far available is low in the collinear region, indicating that recombination must have continued after it had stopped in the two strata defined by the Y region inversions described above. Sequence differences between the single X- and Y-linked regions so far sequenced may merely be variants that happen to be carried on these particular chromosomes, and not sex linked in the species as a whole, and so this region may prove to be partially sex linked.

An approach that does include multiple individuals is bulk segregant analysis. This has recently been successful in the grape vine (Vitis vinifera), an XY system. The fully sex-linked genome region is only c. 150 kb, and includes only a few fully sex-linked genes, and so recombination suppression has clearly not yet extended across any substantial genome region (Picq et al., 2014), consistent with the lack of chromosomal heteromorphism.

The observation that much of the chromosome carrying the SEX locus recombinates does not necessarily mean that it is a young system. Instead, it may be an ancient system whose non-recombining region has remained limited to just a part of the chromosome. This is the case in the Ratite birds (Pigozzi, 2011; Vicoso et al., 2013; Zhou et al., 2014). In Vitis, most wild species are dioecious, suggesting that dioecy is ancestral, and so this could be a plant example of an old established system that has not evolved recombination suppression. However, divergence between Vitis sex-linked genes has not yet been estimated, and so the time at which the sex-determining region evolved is not yet known.

VI. Comparative genetic mapping

Genetic mapping is also important for the detection of differences in the chromosome carrying the SEX region. This can occur when genetic sex determination evolves independently (or re-evolves after loss of dioecy). For example, genes that are sex linked in S. latifolia do not show sex linkage in Silene oitze and S. colpohylla (Mrackova et al., 2008). Phylogenetic analyses of sex systems in Silene (Mrackova et al., 2008; Marais et al., 2011) suggest that these species do not have a dioecious common ancestor, and so dioecy probably evolved de novo in the two lineages, and involved different genes.

Events in which a new gene takes over control of gender can also cause the SEX loci of related species to be on non-homologous chromosomes (Fig. 1c), or on a new location on the same chromosome, as in some animal cases of takeovers (Uno et al., 2008). In Populus, as in papaya, the SEX loci appear to be within pericentromeric regions (Gerald et al., 2015) but, although these regions map to the same linkage group, their locations differ greatly in the physical maps of different Populus species (Yin et al., 2008; Pakull et al., 2011). If confirmed, this suggests a takeover event by a new sex-determining gene on the same chromosome. Independent evolution of separate sexes in different Populus lineages has not yet been excluded, however, even though almost all Salicaceae are dioecious. Data on the ages of the systems, phylogenetic analysis and genome sequencing should help to distinguish between the possibilities. Takeovers or independent evolution both predict that different species should have different sets of genes at their SEX loci, unlike a chromosome rearrangement. Independent evolution predicts that the times since recombination stopped should differ (although this might not be detectable if all species have young systems), whereas takeover events generating single-locus systems
may not have been followed by recombination suppression in the surrounding genome region.

VII. Why does suppressed recombination evolve?

The repeated evolution of regions without crossing over between sex chromosomes strongly suggests a causal connection with the evolution of sex-determining regions (only the centromeric and pericentromeric regions of autosomes generally have suppressed crossing over). The evolutionary strata of sex chromosomes discussed above prove that suppressed recombination often evolves after a sex-determining system is established. Some disadvantage to recombinant genotypes must clearly be involved. Such situations probably occur both during the initial evolution of dioecy, and also later, as males and females evolve in the absence of constraints imposed by the other sex functions. Briefly, as illustrated in Fig. 2, separate sexes in plants probably often evolved from hermaphroditic or monoecious ancestors, often called ‘cosexual’ species (Lloyd, 1982).

The change from cosexuality to dioecy probably involves a mutation creating females (a male-sterility mutation in an initially hermaphroditic species, or a mutation suppressing some or all female flowers in an initially monoecious species, or replacing them with male flowers), and then one or more female-suppressing mutations, creating males or male-biased plants (Westergaard, 1958; Charlesworth & Charlesworth, 1978). (I can find no cases in which dioecy in plants evolved from environmental sex determination, although this seems possible in principle).

In this scenario, male-promoting mutations (suppressing femaleness) clearly reduce female fitness, and are therefore most likely to spread if linked to the gene causing femaleness, which minimizes the conflict. If a two-gene polymorphism results, selection against recombinants will generate linkage disequilibrium (with the X associated with male sterility and the Y with female-suppressor alleles, see Fig. 2). Suppressed recombination is therefore favoured, and may evolve (Charlesworth & Charlesworth, 1980; Bull, 1983), creating a male-determining Y chromosome.

**Fig. 2** Evolution of sex-determining and sex-linked genome regions. (a) Evolution of sex-determining genes in a genome region starting from a hermaphrodite ancestor, and of suppressed recombination in the region, forming a sex chromosome-like region, showing disadvantageous recombinants between the proto-Y and the proto-X chromosomes (the reciprocal recombinant would be hermaphroditic, and is not shown). If the female suppressor has male-specific expression (or evolves expression restricted to males), it can spread through the entire population, and create a single-gene sex-determining system (bottom left). PAR, the pseudo-autosomal region (partially sex-linked region); MSY, the male-specific, or fully Y-linked, region. (b) Evolution of sex-determining genes in a monoecious ancestor.
If, however, a male-specific female-suppressing mutation occurs, no harm is caused to females; if sufficiently advantageous in males, such a mutation can spread, even if unlinked to the femalelessness gene, yielding a single gene sex-determining system (Fig. 2a), and no selection for closer linkage with the gene causing femalelessness (Muller, 1932; Charlesworth & Charlesworth, 1978).

Much genetic evidence supports the two-gene model for plant sex determination, rather than one with sex-specific gene actions. First, in several species (or intercrops of dioecious plants with close non-dioecious relatives), three allelic types at the sex-determining locus control whether individuals are (1) females, (2) males, or (3) hermaphrodites or monocious functional hermaphrodites (Westergaard, 1958). Second, in papaya and grape vine, humans have selected individuals that have Y-linked regions that do not suppress female functions (Wang et al., 2012; Picq et al., 2014). Similarly, in S. latifolia, deletions detectable through the loss of Y-linked markers can create hermaphrodites and neuter plants (Fujita et al., 2011 and references therein). The Y-linked regions of these plants must therefore carry suppressors of female functions whose loss does not affect male functions, and distinct maleness factor(s) elsewhere on the chromosome. Third, in the strawberry species Fragaria virginiana, two closely, but not completely, linked genes with the expected phenotypes have been found (Spigler et al., 2008), whereas a related Fragaria species has suppressed recombination (Goldberg et al., 2010). The sex-determining regions of these species probably evolved independently (Goldberg et al., 2010), but the results nevertheless suggest recombination suppression evolving in response to a two-locus polymorphism.

The evolution of dioecy probably often involves further sexually antagonistic mutations, lending to further selection to suppress recombination, and potentially generating younger strata. For example, dioecy has often evolved from monocacy (Renner, 2014), and full maleness may involve successive increases in the proportion of investment in male flowers (Fig. 2b), each involving sexually antagonistic ‘trade-offs’, because each must decrease the proportion invested in female flowers. Variable degrees of maleness are indeed seen in the monocotyledon Sagittaria latifolia, in the Alismataceae (Dorken & Barrett, 2004), Spinacia oleracea in the Chenopodiaceae (Onoder et al., 2011) and Urtica dioica (Glawe & de Jong, 2009). Similarly, when the ancestral state is hermaphroditism, evolution of dioecy often involves ‘inconstant males’ with partial female function (e.g. producing some fruits in favourable conditions). Species in which genetic variation in male functions seems likely include Antennaria dioica in the Asteraceae (Ubsch, 1936) and Euponymus europaeus in the Celastraceae (Webb, 1979), but these have not been investigated using genetic markers to map the factors. Even after complete unisexuality has evolved, male and female functioning may be suboptimal, and improvements to each sex may often reduce functions of the other. In S. latifolia, for example, female fecundity is enhanced by making large flowers, but fertility is highest for males with many small flowers (Delph & Herlihy, 2012). Just as outlined above for sterility mutations, a mutation benefitting one sex at the expense of the other is most likely to invade but not spread throughout the population; if such a polymorphism is established, it creates selection for reduced recombination with the sex-determining locus, if it arises at a locus closely linked to the SEX locus (Rice, 1987; Jordan & Charlesworth, 2012).

Testing for the trade-offs and conflicts assumed in these scenarios, and for the involvement of sexually antagonistic polymorphisms in the PAR regions of sex chromosomes, is clearly a major task for future work. An approach that can potentially detect sexually antagonistic variation is quantitative trait locus (QTL) analysis within the two sexes separately, as proposed and implemented in S. latifolia (Scotti & Delph, 2006; Delph et al., 2010). This detected several autosomal and PAR QTLs, and, interestingly, the latter appeared only in the analysis of males, implying that their phenotypic effects are not expressed in females. Such male-specific expression is consistent with a past conflict between the sexes that has been resolved in later evolution, as seems to have occurred for some sexually selected male coloration genes in the PAR of a fish, the guppy, Poecilia reticulata (Lindholm & Breden, 2002). Male benefit alleles with male-specific expression no longer harm females, and will spread throughout the population; some other selection is therefore required to maintain the QTL variation, perhaps environmental differences (Scotti & Delph, 2006). In S. latifolia, for example, thin leaves appear to be disadvantageous to males only in dry years (Delph et al., 2011). The S. latifolia QTL analysis used dominant AFLP markers, but codominant markers now available in this plant’s PAR, and obtainable in other plants, will permit future analyses of variation in natural populations. This may detect factors whose conflict has not been resolved, corresponding to the situation that creates selection for reduced recombination in the theoretical models of sexually antagonistic PAR genes.

VIII. Recombination suppression: mechanisms

Non-recombining regions may eventually evolve to encompass a large region of the chromosome carrying the sex-determining loci or locus. Studies of young plant sex chromosomes may be valuable for the investigation of the mechanistic basis of recombination suppression, and whether it generally involves infrequent, large-scale events, such as inversions, or smaller shifts in the position of the PAR boundary.

If chromosome inversions cause recombination suppression in SEX regions (Lahn & Page, 1999), the region will often include many non-sex-determining genes. In papaya, two Y chromosome inversions indeed seem to be involved (including 10 genes with both X and Y copies in the older stratum, and 16 in the newer; Wang et al., 2012). In closely related dioecious Vasconcellea species, alleles of several papaya fully sex-linked genes are not associated with gender (Gschwend et al., 2011). Unlike the Silene situation described above, this probably does not reflect independent evolution of dioecy in Carica and Vasconcellea, as bacterial artificial chromosome-fluorescence in situ hybridization (BAC-FISH) experiments found sex-linked regions including several homologous sequences in similar locations on the largest chromosome of both species (Iovene et al., 2015). Recombination suppression has therefore probably remained restricted in V. paraflora to a genome region near the SEX locus, whereas it has spread across a wider region in C. papaya. This is testable by sequencing to ask whether...
V. parviflora genes homologous to C. papaya genes in the older sex-linked stratum have distinct X and Y haplotypes, like those of papaya. The alternative, that the V. parviflora long arm has become a new recombination-suppressed stratum, seems unlikely, because the chromosomal positions of all nine relevant BACs in an outgroup, Jacaratia spinosa, were found to be similar to those in V. parviflora, and so the inversions probably occurred in the C. papaya lineage.

Recombination suppression mechanisms other than inversions may, however, exist, including modifiers controlling the number of crossover events, restricting them to certain genome regions or restricting crossing over to only one sex. In one of the two human PARs, for example, crossovers are localized very differently in male and female meiosis (Hinch et al., 2014). Some young sex chromosome systems may still be in the process of undergoing recombination suppression. If recombination varies between individuals, or between closely related species that can be interbred, genetic studies can potentially identify the factors involved. In some populations of frog species, male specificity of microsatellite alleles differs between populations, implying that the XY pair shows suppressed recombination only in some populations (Dufresnes et al., 2014). Apparently similar variation was inferred for an anonymous sequence marker within the plant species Bryonia dioica (Oyama et al., 2009), which should be studied further. In S. latifolia, recombination suppression appears to vary between families for several genes (Bergero et al., 2013).

IX. Old established sex chromosome systems

Old sex chromosomes also exist among plants, for example among liverworts (Okada et al., 2001), but have been less studied than young plant systems. As molecular approaches and phylogenetic analyses are extended to the study of further plant sex chromosome systems, it will be interesting to include taxa with high proportions of dioecious species, such as the palm, Vitaceae and Ebenaceae (including Diospyros lotus, see later) families, to test whether dioecy is ancestral and to estimate the time at which recombination stopped, or has evolved several times (as may be the case in the Salicaceae already discussed).

Old systems are particularly interesting for the investigation of genetic degeneration and repetitive sequence accumulation, which occur over large evolutionary timescales. The potentially large range of ages of dioecious plant sex chromosome systems will allow the time course of sex chromosome evolution to be studied. Old plant systems may also help us to understand why recombination suppression sometimes fails to evolve.

The evidence for old established systems is currently incomplete, and age estimates based on sequence divergence are lacking. There are currently no dense genetic maps for plants that seem likely to have old XY systems and, so far, genetic mapping in these systems has largely used non-genic markers, such as AFLPs and microsatellites. These are excellent for testing for a non-recombining (sex chromosome-like) region, for determining which is the heterozygous sex and for estimating the proportion of the chromosome that are fully sex linked. However, as explained above, estimating the age of a sex chromosome system, and the time at which recombination stopped, requires X–Y sequence divergence estimates, based on the ascertainment of sex-linked genes and their sequencing.

In the absence of divergence data, the observation that a sex chromosome system is heterochromatic and heteromorphic might be thought to suggest that it is old established, especially in plant families that include distantly related dioecious species, such as date palms (Al-Mahmoud et al., 2012). For example, Rumex acetosa belongs to a clade that may have been dioecious for 15–16 MYA (Navajas-Perez et al., 2005), but X–Y divergence has not been estimated. Its Y chromosomes are heterochromatic (Shibata et al., 2000; Mariotti et al., 2008), unlike those of other cytologically well-studied plants, such as S. latifolia and S. dioica (Grabowska-Joachimiak & Joachimiak, 2002; Kubat et al., 2014), which are estimated to be younger (Section III). However, heterochromatin can evolve rapidly, as in papaya. Another example is Coccinia grandis, within a wholly dioecious genus of 27 species (in the Cucurbitaceae, another family with many dioecious species, often with XY heteromorphism). Its male genome C-value is 10% larger than that of females, indicating that the Y is much larger than the X chromosome, and the entire Y is heterochromatic (Sousa et al., 2012); yet, phylogenetic analysis suggests that these characteristics evolved recently (Holstein & Renner, 2011).

Sex chromosome heteromorphism can also arise in young systems, for example through fusions with autosomes, as in Rumex hastatus (Smith, 1964) and, possibly, spinach (Araratjan, 1939). The systems in Cannabis sativa (Peil et al., 2003; Sakamoto et al., 2005) and Humulus lupulus (hops) in the Cannabaceae, whose Y chromosome is heterochromatic (Westergaard, 1958), are probably much older.

Studies of old systems are also needed to test the prediction that other sex-determining systems are derived from XY systems (Charlesworth & Charlesworth, 1978). Again plants may be very helpful, as systems with male-determining Y chromosomes probably evolve first, as outlined above, butZW systems also exist, and it can be tested whether the frequencies of such systems increase over time. X-autosome balance systems are also probably derived from XY systems (and potentially allow the loss of the Y chromosome, and evolution of an X0 male genotype). However, it has again not yet been demonstrated that such species tend to be older than other plant sex-determining systems. The absence of carpel development in males or stamen development in females, as in hops, may also indicate an ancient system (but might simply be caused by a long history of unisexual flowers, for example because dioecy has evolved from monoecy); so far, only one fully sex-linked genetic marker locus has been found in hops (Jakse et al., 2008).

X. Genetic degeneration: the need for empirical data in a phylogenetic setting

Ancient systems are also of great interest for the study of genetic degeneration (gene loss or loss of function). In diploid organisms, only the Y chromosomes are predicted to degenerate, because X chromosomes recombine in the XX females, whereas Y-linked regions do not, and are subject to several processes that allow detrimental mutations to increase in frequency in the population of Y-linked alleles, or even to become fixed in this population, as
recently reviewed by Bachtrog (2008). In haploid plants, however, the complete lack of recombination across the entire sex-linked region predicts similar degeneration of both U and V chromosomes (Bull, 1983). Genes affecting non-sex functions should not degenerate or become lost, and so the female-determining U region should lose only male function genes, and the male-determining V region only female function genes (Fig. 3c).

Haploid plants with separate sexes of gametophytes are ideal for the study of this prediction. In *M. polymorpha*, a species whose sex chromosomes carry highly diverged sequences, the V chromosome has been studied in detail, but analysis of the U chromosome is currently incomplete (Okada et al., 2001). In the brown alga, *Ectocarpus siliculosus*, however, *c*. 24 genes were found in the fully sex-linked regions (either the U or V regions, or both), seven of which were not detected in the V and nine in the U chromosome (Ahmed et al., 2014). This is in apparent agreement with Bull’s prediction; however, without an outgroup, gene movements onto one sex chromosome, but not the other, cannot be excluded. To determine whether suitably close non-dioecious relatives exist (and avoid species that might have reverted from dioecy to a non-dioecious state), phylogenetic relationships of the species must be known. This is often difficult for closely related species, a frequent situation relevant to the evolution of sex chromosomes. Nevertheless, among plants, sets of species should exist with good phylogenies well suited for future work estimating ancestral character states, and changes in states (Maddison & Leduc-Robert, 2013).

Diploid dioecious plants also have extended haploid life cycle stages, which may also cause genetic degeneration of non-recombining sex chromosomal regions to be minor (Fig. 3b). Around two-thirds of plant genes are expressed in male gametophytes of angiosperms (Tanksley et al., 1981; Gorla et al., 1986; Honys & Twell, 2003). Therefore, only genes with no important pollen functions should be lost from plant SEX regions, or lose their functions; the limited evidence so far about the loss of genes from the *S. latifolia* Y chromosome is consistent with this expectation (Guttman & Charlesworth, 1998; Chibalina & Filatov, 2011). Degeneration might thus be restricted to *c*. one-third of genes (or possibly somewhat higher if the expression of some pollen-expressed genes is not important, and purifying selection maintaining their functions is consequently weak). The few current estimates, from the unrelated plants *S. latifolia* and *R. hastatulus*, suggest that fewer than 30% of Y-linked genes have lost expression (Bergero & Charlesworth, 2011; Chibalina & Filatov, 2011; Hough et al., 2014). By contrast, such regions are almost completely degenerated in the best-studied animals, such as species of *Drosophila* (Muller, 1950), mammals (Skaletsky et al., 2003) and those birds that have extensive fully W-linked regions (Zhou et al., 2014), and possibly in part of the much younger Y chromosome of the threespine stickleback (Ross & Peichel, 2008; Yoshida et al., 2014). Large genome regions that stopped recombining recently and carry many genes driving the degeneration processes, such as the neo-Y chromosome of *Drosophila miranda*, have quickly lost functions of large fractions of genes (Bachtrog et al., 2008).

However, the regions of the two plants so far studied that recently became fully sex linked probably include many fewer genes than the *D. miranda* region, so that the small extent of gene losses in these young systems is not surprising. It will be interesting to study older plant systems.

Genetic degeneration in young plant sex chromosomes, and in young evolutionary strata in older systems, is also of interest. The first step after recombination is suppressed between Y- and X-linked regions may be the accumulation of repetitive sequences, including transposable elements. Such insertions may decrease the expression of Y-linked alleles, even before mutations in the coding regions, or in non-coding regions that control the gene’s expression. This appears to be the case in *Drosophila albomicans* (Zhou & Bachtrog, 2012).

However, plants with sex-determining loci within rarely recombining pericentromeric regions, such as papaya and *Populus* species, are not well suited for the study of genetic degeneration, because the accumulation of maladaptive sequence changes and of repetitive sequences is also expected in pericentromeric genome regions (Charlesworth et al., 1986). It will therefore be difficult to detect extra effects of the evolution of sex-determining genes in the region. For example, in papaya, gene density is low in the sex-linked region, but this is not wholly caused by the loss of genes; accumulation of repeated sequences has also reduced gene density (Wang et al., 2012).

1. Dosage compensation

In sex chromosome systems in which Y-linked gene expression is reduced, or Y-linked genes have been lost, dosage compensation has sometimes evolved, and it is therefore interesting to test whether X-linked alleles of plant genes whose Y-linked copies have lost function are expressed at higher levels in males than females. There is currently no clear evidence that this occurs in *S. latifolia* or *R. hastatulus*, but partial compensation cannot yet be excluded (Chibalina & Filatov, 2011; Muyle et al., 2012; Hough et al., 2014; Bergero et al., 2015).
XI. Plant sex-determining loci

To identify sex-linked regions and to determine whether males or females are the heterozygous sex, it is not necessary to find the gene(s) controlling male or female development. As explained above, it suffices to find genetic markers, even anonymous ones, such as AFLPs, that co-segregate with sex. However, plant sex-determining loci are interesting in several ways, including for the identification of the hypothesized two or more genes causing male and female sterility during the evolution of dioecy. If sex-determining genes can be discovered, sequence divergence between their sex-linked alleles may also help to estimate the time at which recombination first stopped. With the possibility of dense marker development and genome sequences, renewed efforts are being made to identify plant sex-determining genes, and progress can be expected in the next few years.

The approach of testing known flower development genes has been largely superseded by high-throughput sequencing methods. Searches have found MADS-box and ABC(DE) genes involved in flower whorl development on the sex chromosomes of *S. latifolia* and *Asparagus officinalis* (Matsumaga *et al.*, 2003; Park *et al.*, 2003; Cegan *et al.*, 2010; Nishiyama *et al.*, 2010; Penny *et al.*, 2011). However, genes that control floral organ identity are not generally promising candidates. They might be involved in species with complete absence of one sex organ (Type I of Mitchell & Diggle, 2005; Ramos *et al.*, 2014). In many dioecious plants, however, both male and female floral organs are initiated in flowers of both sexes, and the development of opposite sex organs is later interrupted.

Alternative approaches also encounter difficulties because of the numerous candidates whose loss of function can produce male or female sterility. For example, as mentioned above, deletion mapping of the *S. latifolia* Y chromosome has established that separate loci exist whose deletion causes abortion or incomplete development of stamens, or removes the suppression of pistils that occurs in wild-type males, creating hermaphrodite flowers (Farbos *et al.*, 1999; Lardon *et al.*, 1999; Zluvova *et al.*, 2005; Bergero *et al.*, 2008; Fujita *et al.*, 2011). However, these deletions probably involve the loss of many fully sex-linked genes other than those causing these phenotypes, and this is supported by the observation that pollen carrying deleted Y chromosome regions often has low ability to fertilize ovules (Lardon *et al.*, 1999). When the sex-linked region is large, it will be difficult to identify the genes responsible for the evolution of dioecy unless small deletions can be generated and identified using dense mapping of sequences lost from deleted genotypes.

This problem also hinders attempts to identify genes involved in gender determination using mutations, including mutations induced by ethyl methanesulfonate (EMS) or irradiation (Ohnishi, 1985; Christensen *et al.*, 1998; Hony & Twell, 2004; Wellmer *et al.*, 2006; Chang *et al.*, 2011), or by studying genes with different expression in flower buds of the two sexes. Moreover, many genes have stamen- or pistil-specific expression, and will be non-expressed in buds of one sex purely because the relevant structures are absent. Distinguishing such downstream acting genes from the sex determiners themselves requires the establishment of sex linkage. If, however, the fully sex-linked region includes many genes, the problem of having too many candidates with suitable function is not eliminated. In addition, expression differences may not be involved (for example, male sterility can involve mutations in coding sequences, and the mutant alleles may be present in mRNA).

Small sex-linked regions may offer the best prospects for the identification of the sex-determining genes, because fewer candidates need to be considered. A candidate Y-linked gene has been proposed in persimmon (Akagi *et al.*, 2014). This study started by identifying sex-linked genes, using pools of males and females from a full-sib family, and their sex linkage was confirmed in samples of unrelated males and females. Efforts were made to ensure that most Y-linked genes present in transcripts were detected by employing RNA-Seq, and 22 expressed sequences were identified. The total length of sex-specific sequences was only 1 Mb, suggesting a small fully sex-linked region. One candidate for involvement in sex determination was found. This gene (named *OGI*) is expressed only in male flower buds. *OGI* is a duplication onto the Y-linked region of an autosomal gene called *MeGI* that expresses a male-suppressing regulatory RNA in females. Because no X copy exists, X–Y divergence cannot be estimated, but divergence from the presumed autosomal progenitor is high, and Y-linked *OGI* sequences were detected in other species of Ebenaceae, suggesting an old established Y-linked duplication. Low divergence was found between the X- and Y-linked alleles of other sex-linked genes (silent site divergence of 12 XY allele pairs was below 2%), suggesting that a younger stratum evolved recently.

The proposed scenario for sex determination in persimmon is that the Y-linked *OGI* gene opposes *MeGIs* male-suppressing action. This form of gene action that could act in the heterozygous state, and should increase male functions, and the processes in the two sexes may indeed conflict, as proposed for the female suppressor in the two-gene model outlined in Fig. 2(a). It is currently unclear how females evolved. The *MeGI* male-suppressing factor is autosomal, and is therefore unlikely to represent the male-sterility gene in the two-gene model. *OGI* could therefore be an example of a single-locus sex-determining gene that evolved by a take-over event, if searches fail to find a femaleness factor.

Although, as mentioned already, reversals and re-evolution of dioecy can complicate comparative studies and hinder inferences of the ages of the origins of dioecy, they may also be very helpful in revealing the genetic basis of dioecy (Westergaard, 1958), and molecular studies of such hermaphrodites could help to identify plant sex-determining genes. In papaya and *Vitis*, hermaphrodites are commercially successful crop plants. The Y chromosomes in these hermaphrodites do not suppress female functions, but their sequences are very similar to those of males (Picq *et al.*, 2014; Van Buren *et al.*, 2015), and they probably have no large deletions, making them ideal for the identification of the gene whose loss causes reversion to hermaphroditism, a good candidate for the female suppressor involved in the evolution of dioecy.

The hypothesized X-linked genes responsible for the male sterility of females in dioecious plant species are likely to be even
harder to identify, but this may be possible in systems in which suppressed recombination has not yet evolved. If two incompletely linked sex-determining genes exist, hermaphrodite recombinants and recombinants with the male sterility allele of females and the female suppressor of males should arise. With the modern ability to identify the region and genotype closely linked markers, as in *Fragaria* species (Tennesen et al., 2013), it should be possible to check that these phenotypes are indeed associated with recombination, and to pinpoint both genes.

Once the genes are identified in some plant species, this will open up the way for testing whether the same genes are sex linked in other dioecious plants. Given that large numbers of genes affect flower and inflorescence development, different genes may be involved in different angiosperm lineages, rather than the same genes being repeatedly involved. If so, plants will differ from major animal groups, such as insects, which share sex-determination pathways across major taxa (Saccone et al., 2002; Beye et al., 2003; Pomiankowski et al., 2004; Pane et al., 2005). In plants, sterility factors may have to be identified, and their actions investigated, in individual genera and species. Moreover, it should not be assumed that the sex-determining genes necessarily function during flower development, or cause sterility. In monoecious plants, a state that is ancestral to many dioecious species (Renner & Ricklefs, 1995; Renner & Won, 2001), they might instead control the proportions of male and female flowers, perhaps at developmental stages before flower parts are initiated (Fig. 2b). Unisexuality may be much more ancient than dioecy, and early, complete abortion of male or female parts may be ancestral.

XII. Conclusions

It is now technically feasible to use young sex chromosomes in nonmodel plants to test hypotheses about the initial evolution of suppressed recombination, and to study the time course of later evolution of sex chromosomes in older systems, as has been initiated in some animal systems (Bachtrog et al., 2009). Young plant sex-linked systems should also be suitable for the investigation of the earliest adaptations to dioecy, which have so far been little studied. The change from cosexuality to unisexuality may be accompanied by considerable expression changes, if unisexuals are released from conflicts between the two sex functions, so that changes can occur to optimize expression in each sex. For example, the non-dioecious *S. vulgaris* appears to be suitable as an outgroup for the study of the evolution of changes in expression in the dioecious species *S. latifolia* (Marais et al., 2011). Because hermaphrodite *S. vulgaris* individuals have both stamens and pistils, differences in unisexual individuals of the dioecious species that are caused directly by the loss of these structures should be distinguishable from changes in expression of genes that are expressed in non-sex-specific structures. Such studies can potentially discover genes that can be expressed in both sexes, but that change when dioecy evolves, and evaluate whether, as has been predicted, the sex chromosomes, including the PAR, carry unexpectedly large numbers of such genes (Vicoso & Charlesworth, 2006; Vicoso et al., 2013).

Acknowledgements

I thank S. C. H. Barrett, S. I. Wright and J. Leebens-Mack for comments on an earlier version of the manuscript.

References


[References for the document are listed here, with specific bibliographic information for each source.]


New Phytologist

provide genetic evidence of an XY chromosome system and recombination arrest and allow to trace paternal lineages in date palm. New Phytother 197: 409–415.


Ming R, Bendahmane A, Renner S. 2011. 


