EVOLUTION

Rapid and dynamic evolution of a giant Y chromosome in *Silene latifolia*

Takashi Akagi[1,2,3]†*, Naoko Fujita[1,4]†, Kenta Shirasawa[5], Hiroyuki Tanaka[6], Kiyotaka Nagaki[7], Kanae Masuda[1], Ayano Horiuchi[1], Eriko Kuwada[1], Kanta Kawai[1], Riko Kunou[1], Koki Nakamura[7], Yoko Ikeda[7], Atsushi Toyoda[8,9], Takehiko Itoh[6], Koichiro Ushijima[1], Deborah Charlesworth[10]

Some plants have massive sex-linked regions. To test hypotheses about their evolution, we sequenced the genome of *Silene latifolia*, in which giant heteromorphic sex chromosomes were first discovered in 1923. It has long been known that the Y chromosome consists mainly of a male-specific region that does not recombine with the X chromosome and carries the sex-determining genes and genes with other male functions. However, only with a whole Y chromosome assembly can candidate genes be validated experimentally and their locations determined and related to the suppression of recombination. We describe the genomic changes as the ancestral chromosome evolved into the current XY pair, testing ideas about the evolution of large nonrecombining regions and the mechanisms that created the present recombination pattern.

Chromosomal sex determination is common not only in animals, but also in a diversity of plant species. Most extant flowering plants are functional hermaphrodites, the sex chromosomes of which evolved much more recently than those of many animals and independently in different lineages [reviewed in (*1*)]. They are therefore interesting for studying sex chromosome establishment when separate sexes evolved from ancestral hermaphroditism. Close linkage between male and female-determining genes may be involved (*2*), and subsequent selection for sexual dimorphism involving traits that benefit only one sex (sexually antagonistic, or SA, traits) can lead to the establishment of polymorphisms closely linked to the sex-determining factor(s), favoring further recombination suppression (*3*). These ideas can explain the extensive male-specific regions of Y chromosomes (MSYs), that do not recombine with and become differentiated from the X chromosome. However, physically extensive MSYs may not require sexually antagonistic selection. For example, sex-determining factors may evolve within a genome region that was already recombinationally inactive (*4*). One alternative suggestion is that inversions carrying no SA factors may be more likely to persist on sex chromosomes than on autosomes (*5*, *6*). Moreover, sexual dimorphisms can be caused by the pleiotropic effects of the sex-determining factors themselves and do not require separate genetic factors with SA effects (*7*). The plant *Silene latifolia* is well suited for testing such ideas.

The first observation of sex chromosomes in a flowering plant was in *S. latifolia*, the karyotype of which is $2n = 22A + XY$ (*8*). Its "giant" Y chromosome (now known to be ~500 Mb; Fig. 1A) has been a model for studying heteromorphic sex chromosomes with large MSYs. After decades of cytogenetic and theoretical studies, empirical studies using limited numbers of molecular markers (*9*, *10*) revealed a pattern resembling the "evolutionary strata"

[1]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan. [2]Japan Science and Technology Agency (JST), PRESTO, Kawaguchi-shi, Saitama, Japan. [3]Kihara Institute for Biological Research, Yokohama City University, Yokohama, Kanagawa, Japan. [4]National Museum of Nature and Science, Tsukuba-shi, Ibaraki, Japan. [5]Kazusa DNA Research Institute, Kazusa-Kamatari, Kisarazu, Chiba, Japan. [6]School of Life Science and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo, Japan. [7]Institute of Plant Science and Resources, Okayama University, Kurashiki, Okayama, Japan. [8]Comparative Genomics Laboratory, National Institute of Genetics, Mishima, Shizuoka, Japan. [9]Advanced Genomics Center, National Institute of Genetics, Mishima, Shizuoka, Japan. [10]Institute of Ecology and Evolution, University of Edinburgh, Edinburgh, UK.
***Corresponding author. Email: takashia@okayama-u.ac.jp or claclatakashi@gmail.com**
†These authors contributed equally to this work.

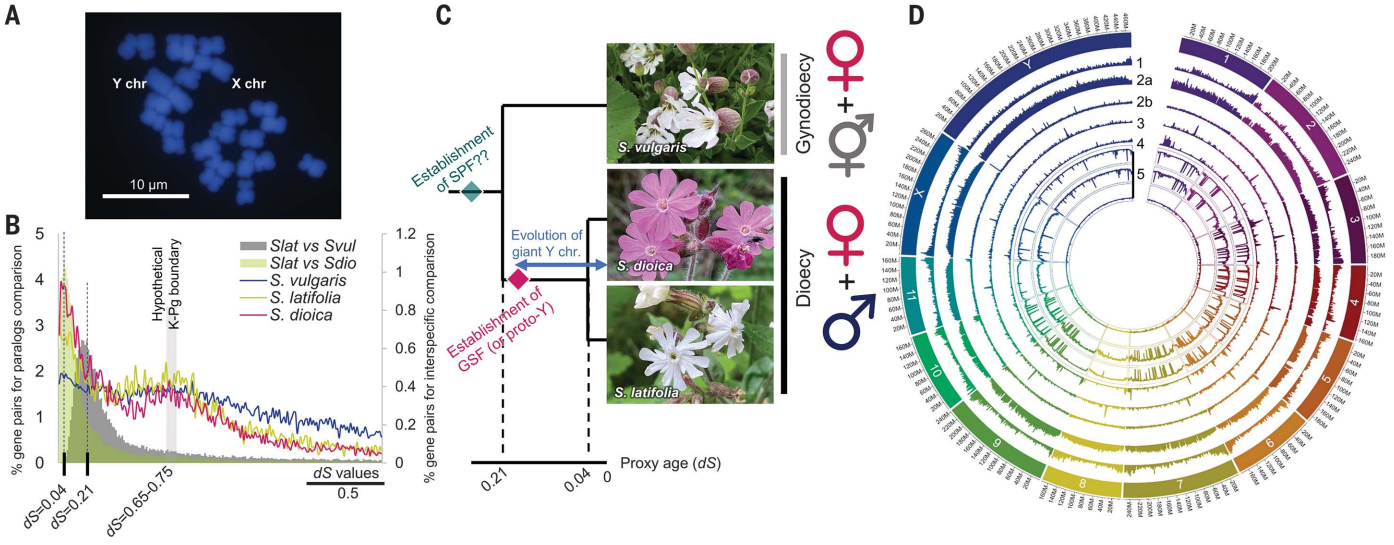

Fig. 1. Overview of the sex chromosomes and genome characteristics in *S. latifolia* and the two related species, *S. dioica* and *S. vulgaris*. (**A**) Cytogenetic observations of the giant Y and X chromosomes with 4′,6-diamidino-2-phenylindole (DAPI) staining. (**B**) Histogram of the pairwise dS values between paralogous genes within each genome and between orthologous genes in two species (*Slat* versus *Sdio* and *Slat* versus *Svul*). The interspecific dS values exhibited a peak at dS = 0.04 and 0.21, which corresponds the proxy age for divergence of *S. latifolia* and *S. dioica* and *S. latifolia* and *S. vulgaris*, respectively. (**C**) Estimated evolutionary context of the two sex-determining factors, SPF and GSF, in dioecious species (*S. latifolia* and *S. dioica*). (**D**) Bird's-eye views of the genomes of *S. latifolia*. Circle layer 1, gene density; layer 2, TE density (2a, LTR-type TEs; 2b, non-LTR); layer 3, standardized CENH3/H3 estimates; layer 4, standardized H3K9me2/H3 estimates; and layer 5, weighted DNA methylation rate (CG, CHG, and CHH, respectively, from the outer to the inner ring).

first discovered in humans (*11*), with Y-X sequence divergence lowest for genes closest to the pseudo-autosomal region that still recombines. Such divergence differences indicate that Y chromosome recombination suppression occurred at different times. However, important gaps in our understanding remain, including why the Y-linked region is so large and to what extent its lack of recombination has led to genetic degeneration, such as mammalian Y chromosomes (*12*). Recent advances in genome sequencing have yielded data suggesting that partial degeneration (*13*, *14*) and possibly dosage compensation (*15*–*18*) have evolved, but degeneration cannot be accurately quantified without an assembled Y chromosome.

Observations of Y chromosome deletions revealed that the *S. latifolia* Y chromosome carries two sex-determining factors, a stamen-promoting factor (SPF) and a gynoecium-suppressing factor (GSF) (*19*). This plant's Y therefore fits the "two-mutation model" (*20*), in which a loss-of-function mutation in an SPF in an ancestral hermaphrodite species creates females, and males arise by a dominant femaleness-suppressing mutation, GSF, in a closely linked gene. Consistent with this model, two genes have been found within physically small sex-determining genome regions in several dioecious plants (*21*–*24*). The very different situation in the *S. latifolia* Y chromosome suggests that, as the model proposes, selection favored closer linkage between initially recombining genes, with loss of Y-X recombination triggering repetitive sequence accumulation (*25*, *26*) and creating the giant Y chromosome.

Genome sequences of *Silene* species

The *S. latifolia* sex chromosomes, which are shared with its close dioecious relative, *S. dioica* in the subgenus Behenantha, evolved after divergence from *S. vulgaris*, a gynodioecious species (with female and hermaphrodite individuals, and $2n = 24$A chromosomes) (Fig. 1, B and C). We assembled chromosome-scale sequences of *S. latifolia* and *S. dioica* males and a *S. vulgaris* hermaphrodite using PacBio HiFi reads and Bionano optical mapping or Hi-C contact mapping, cytogenetic anchoring with molecular markers, and genetic maps (Fig. 1D and fig. S1). BUSCO (*27*) analysis found 93.1%, 99.7%, and 97.7% of universal single-copy orthologous eudicotyledonous plant genes in our *S. latifolia*, *S. dioica*, and *S. vulgaris*, genome sequences, respectively (table S1). Synonymous site divergence (dS) between *S. latifolia* and *S. dioica* and *S. latifolia* and

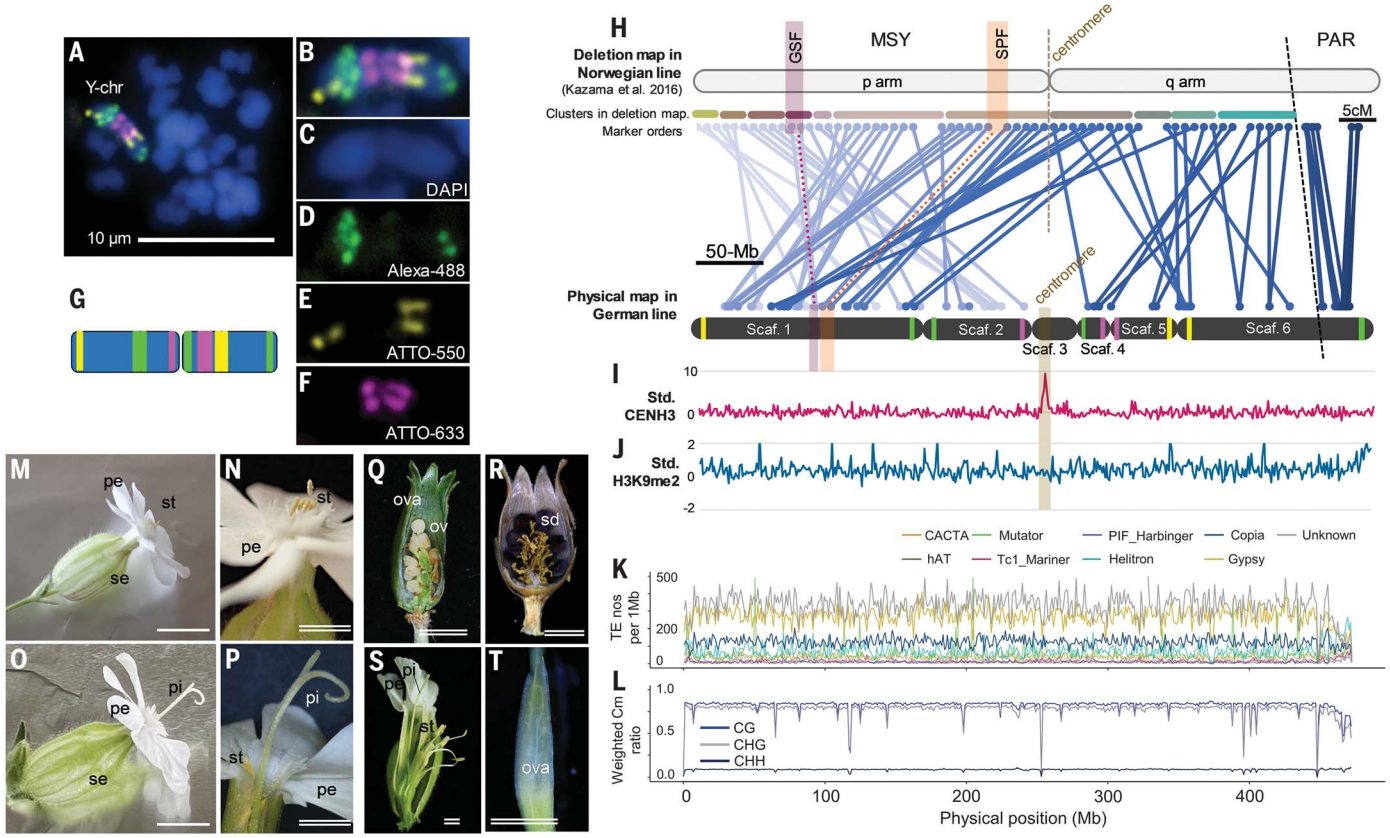

Fig. 2. Genomic context of the *S. latifolia* Y chromosome. (**A** to **G**) FISH anchoring of the genomic contigs. (A) Detection of the Y chromosome by myTag probes. (B) Enlarged image of the Y chromosome in (A) merging the signals with DAPI. (C to F) Staining for three probes, Alexa Fluor-488 (D), ATTO-550 (E), and ATTO-633 (F). (G) Ideogram of the Y chromosome based on the signals. (**H** to **L**) Genomic context of the Y chromosome. (H) Comparative analysis of the deletion map in a Norwegian line (*30*) (with only marker orders given because no assembly was available) and our genome sequence of a German line. The regions including the two sex-determining genes are highlighted. In the recombinationally active PAR, the genetic map positions from Bergero *et al*. (*10*) correlate with the physical positions in our genome sequences. (I and J) CENH3 content (standardized by H3 amounts) (I) exhibits a clear single peak, whereas the standardized H3K9me2 levels (J) fluctuate but show no major pattern along the chromosome. (K) TE distribution in 1-Mb windows. LTR-TEs, especially *Gypsy* and *unknown* classes, show high densities in the MSY compared with the PAR. (L) CG and CHG DNA methylation levels show a similar pattern. (**M** and **N**) VIGS of *Y-SlCLV3* resulted in the production of hermaphrodite flowers in a genetically male *S. latifolia*. Control male flowers [(M) and (N)] exhibited functional stamens (st) but no elongated pistils. (**O** to **R**) ALSV-induced gene silencing resulted in elongation of functional pistils (pi) with normal stamens in male plants [(O) and (P)], producing ovules (ov) (Q) to be viable seeds (sd) (R). (**S** and **T**) ALSV-infected male lines often induced imperfect sterile pistils. pe, petal; se, sepal; ova, ovary. Scale bars, 1 cm (single line) and 5 mm (double line).

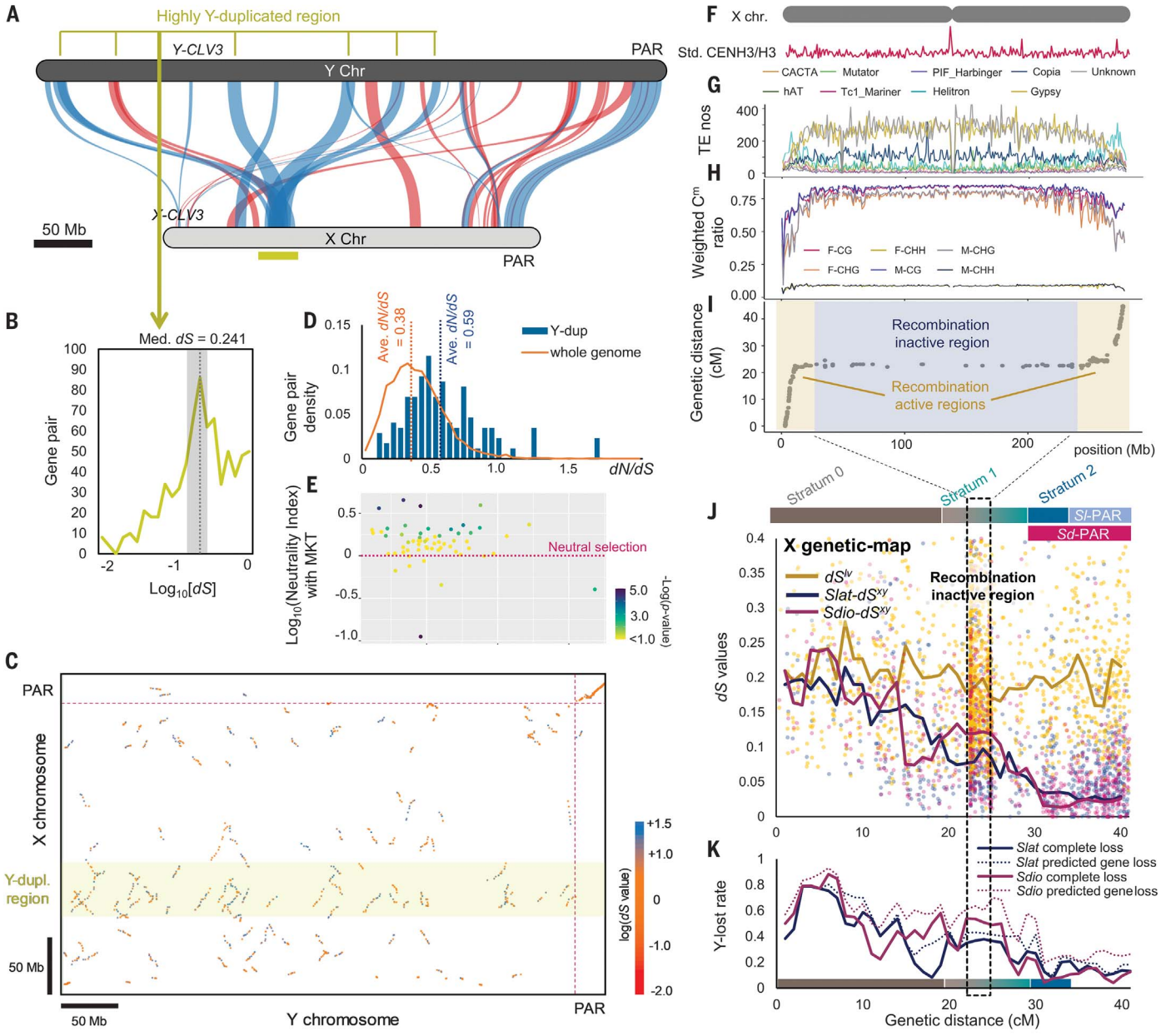

Fig. 3. Comparative analysis of Y and X chromosomes in *S. latifolia*. (**A**) Gene order–based synteny analysis between the X and Y chromosomes. Blue and red bands indicate synteny with forwarded and inverted directions, respectively. The X-linked region marked with a gold horizontal band includes sequences that are highly duplicated in the Y chromosome. (**B**) Pairwise dS values of genes in the Y-duplicated regions. The peak with dS values of 0.158 to 0.316 is highlighted in gray. (**C**) Locations of syntenic blocks in the X and Y chromosomes, with colors indicating the mean dS values between the X- and Y-linked sequences (see the key). (**D**) Values of dN/dS in the paralogous gene pairs with dS in the interval between 0.158 and 0.316; pairs are shown separately for the Y-duplicated region and the whole genome. (**E**) Distribution of the neutrality index in McDonald-Kreitman tests [the *x*-axis values are the dN/dS values given in (D)]. (**F** to **I**) Characteristics of the X chromosome. (F) Standardized CENH3 amount (corrected by H3 amount). (G) TE distribution with 1-Mb windows. (H) Weighted DNA methylation levels. (I) Genetic recombination map estimated using an F1 population of 228 individuals, showing that the regions with highly elevated LTR-TEs, especially the *Gypsy* and *unknown* classes, and CG and CHG DNA methylation levels corresponding with the recombination arrested region. (**J** and **K**) Evolutionary strata in terms of Y-X dS and of genetic degeneration in *S. latifolia* and *S. dioica*. For *S. dioica*, the genetic distances were defined with the orthologous genes in *S. latifolia*. The *x* axis shows positions in the X chromosome genetic map (fig. S15 shows the same data with positions in the X assembly). (J) dS values between X- and Y-linked alleles (*Slat-dS*^{XY} and *Sdio-dS*^{XY} in *S. latifolia* and *S. dioica*, respectively) and between *S. latifolia* X-linked alleles and their orthologs in *S. vulgaris* (*dS*^{lv}). The lines indicate median dS values in 1-cM bins. Based on Pettitt's change-point test using the estimated *Slat-dS*^{XY} values, we defined three strata, 0 (*dS*^{XY} > ~0.10), 1 (0.05 to 0.10), and 2 (0.01 to 0.05). (K) Proportion of genes ancestrally present (defined in the text) that have been lost from the *S. latifolia* and *S. dioica* Y chromosome. Complete loss means that the genome region corresponding to the gene is completely absent from the Y chromosome. Predicted gene loss means that the putative gene is not predicted in the Y sequence by our annotation or it is disrupted.

S. vulgaris average 0.04 and 0.21, respectively. Most sequences in these species are diploid despite an ancient genome duplication event at the K-Pg boundary (with dS = 0.65 to 0.75; Fig. 1B and fig. S1), which is consistent with other plant polyploidization events (*28*, *29*).

The *S. latifolia* Y chromosome sequence forms six super-contigs totaling ~480 Mb. Genetic mapping of markers in male meiosis revealed a large nonrecombining male-specific region of the Y chromosome (MSY) and a pseudoautosomal region (PAR) at one chromosome end

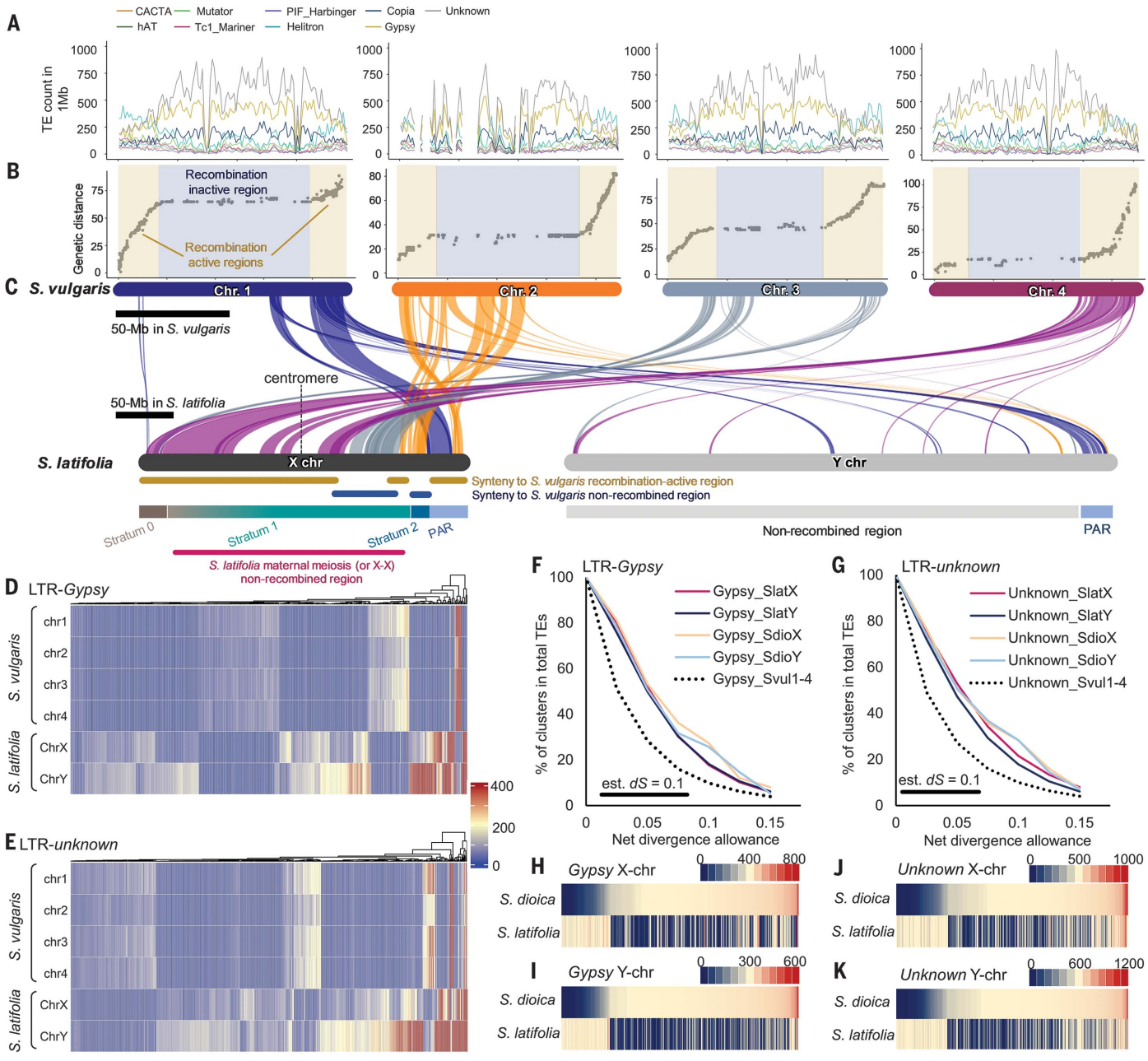

Fig. 4. Comparative genomic analysis between *S. latifolia* and *S. vulgaris*. (**A** to **C**) Genomic contexts of chromosomes 1 to 4 of *S. vulgaris* and their detailed relationship with the X and Y chromosomes of *S. latifolia*. (A) TE distribution. (B) Genetic recombination map showing wide, recombinationally inactive regions with accumulation of *Gypsy* and *unknown* LTR-TE classes (fig. S17). (C) Half of the X chromosome (and potentially also the ancestral Y chromosome) shows synteny with recombinationally active regions of *S. vulgaris* chromosome 4. The other half may be derived from fusions involving *S. vulgaris* chromosomes 1 to 3 regions, some of which are currently nonrecombining. (**D** and **E**) Sequence homology–based clustering of the *Gypsy* (D) and *unknown* LTR-TE classes (E) extracted from X/Y chromosomes in *S. latifolia* and chromosomes 1 to 4 in *S. vulgaris*. Many are specific to the X or Y chromosome or are lineage specific. (**F** and **G**) Numbers of *Gypsy* (F) and *unknown* type (G) TE clusters (defined as sequences with the divergence values indicated on the *x* axis) as proportions of the total numbers of each TE type in *S. latifolia* and *S. dioica*. The *x* axis shows clusters defined using threshold values of up to 15%, which approximately quantify the divergence of extant insertions from their ancestral sequences and therefore reflect the insertions' ages. The figure shows that the X and Y chromosomes of both species are rich in TE bursts that are older than bursts on the *S. vulgaris* chromosome 4. (**H** to **K**) Sequence homology–based clustering of the recent (>95% identity) *Gypsy* [(H) for X and (I) for Y] and *unknown* classes of LTR [(J) for X and (K) for Y] TEs extracted in *S. latifolia* and *S. dioica*. Most clusters indicate lineage-specific bursts.

(*10*, *15*). We anchored MSY contigs cytogenetically by fluorescence in situ hybridization (FISH) using dense probes specific to each contig end (see Fig. 2, A to H; fig. S2; and the materials and methods for details). All chromosomes are metacentric, but the Y arrangements differ within *S. latifolia*, as the arrangement in our family (using parent plants from Germany) differs from that in a Norwegian family (named K-line) used to determine marker orders by deletion mapping (*30*) (Fig. 2H and fig. S3). CENH3 antibody peaks and dense tandem repeats indicate the autosome centromere locations (figs. S4 and S5). No mature repeats were found in the Y centromeric region (Fig. 2I and fig. S5B), supporting rearrangements changing the relative Y chromosome centromere location. Consistent with this, the densities of H3K9me2 (a common heterochromatin epimark) and DNA methylation in young leaves (mainly in the CG and CHG contexts) are highly distributed throughout the MSY (Fig. 2, J to L). Long terminal repeats (LTRs) of the *Gypsy* and *Unknown* classes are enriched across the MSY and in autosomal pericentromeric low-recombination regions (Fig. 2K, figs. S6 and S7, and table S2), as discussed below.

Identification of the sex-determining gene candidates

Deletion mapping confirmed Westergaard's finding of MSY regions affecting male and female flower organs (*31*, *32*). We identified regions in our assembly syntenic with two previously identified; the GSF and SPF regions (figs. S8 and S9), both in the current Y*p* arm (*30*), which include 31 and 62 candidates, respectively (tables S3 and S4). Some candidate genes are absent from the *S. dioica* Y chromosome and were excluded from further consideration. Expression analysis in early flower primordia supported the previous GSF candidate (*33*), a *CLAVATA3*-like gene (*SlCLV3*) that may affect flower meristem size and gynoecium development. Its X and Y copies are highly diverged (dS^{XY} = 0.179), and in male flower bud primordia, we detected only *Y-SlCLV3* expression, consistent with the previous study. Experimental virus–induced gene silencing (VIGS) of Y-*SlCLV3* in male *S. latifolia* with apple-latent spherical virus (ALSV) (*34*) resulted in the formation of functional hermaphrodite flowers (Fig. 2, M to S; fig. S10; and table S6), validating this conclusion for the first time. One candidate within the SPF region is an ortholog of the immunophilin-like *FKBP42/TWISTED DWARF1* (*TWD1*) gene, which is required for androecium development in angiosperms (*35*, *36*). This gene is Y specific in *S. latifolia* and *S. dioica* and exhibits high expression in male flower primordia (table S4); the effects of silencing *TWD1* have not yet been tested.

Differentiation of X and Y chromosomes in *Silene*

The *S. latifolia* and *S. dioica* MSYs show little homology with their X counterparts (fig. S11). Gene-based syntenic block analyses detected many Y-specific duplications (Fig. 3, A to C, and

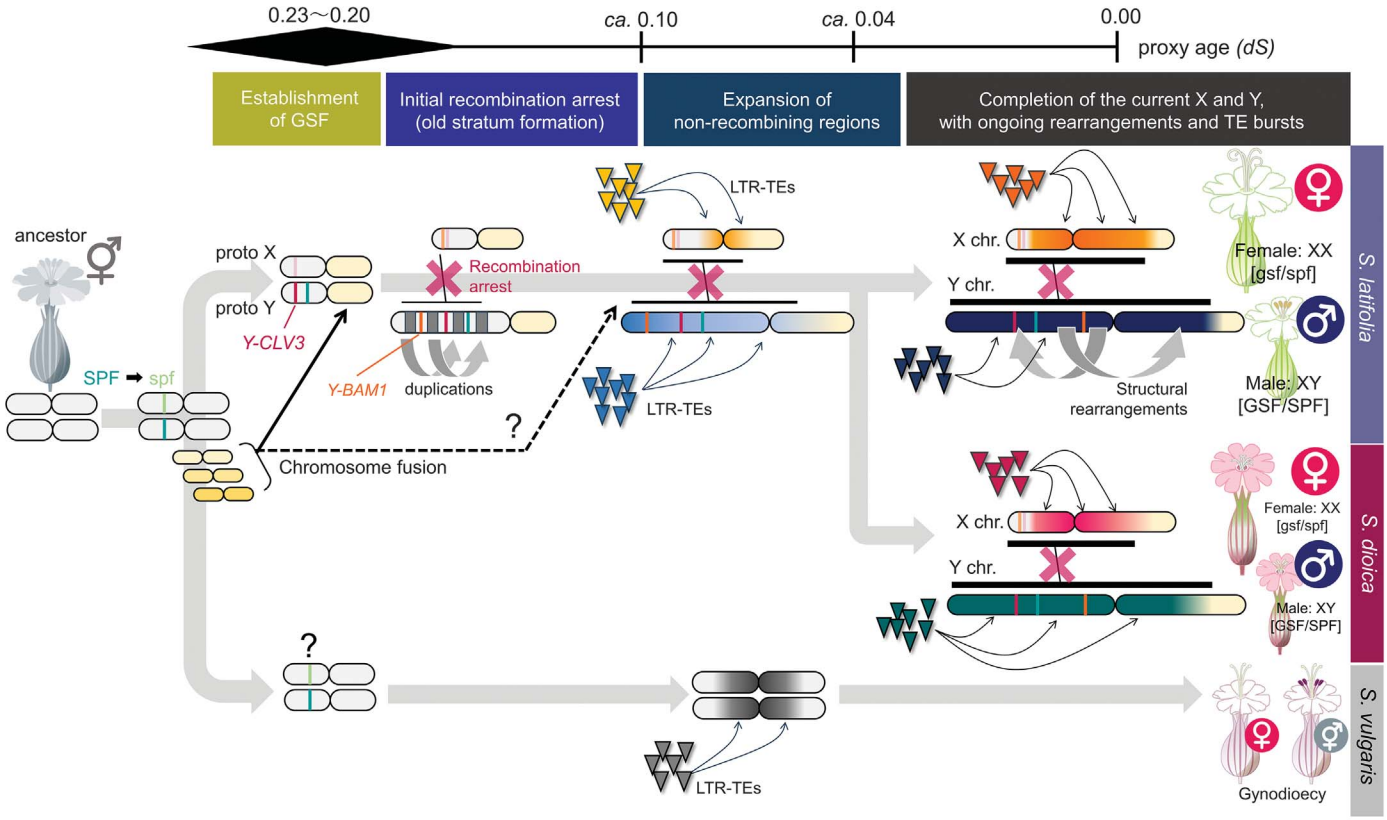

Fig. 5. Model for the rapid and dynamic sex chromosome evolution in *S. latifolia* and *S. dioica*. Under the two-mutation model for the evolution of separate males and females from a hermaphroditic plant, the SPF loss-of-function mutation occurred first, before the oldest nonrecombining stratum evolved. The SPF gene has not yet been identified. The second mutation (gain of a dominant gynoecium-suppressing function on the Y chromosome, the *Y-SlCLV3* allele), produces males. This gene has an X-linked allele, allowing its location to be identified as being within the oldest stratum, 0, consistent with its high Y-X divergence. Assuming that the SPF, *Y-SlCLV3*, and a possible sexually antagonistic gene, *SlBAM1*, were within the left ancestrally recombining region, selection would have favored suppressed recombination, explaining the evolution of stratum 0. After recombination stopped, the duplicated regions evolved on the Y chromosome left arm, also before the split of *S. latifolia* and *S. dioica*. Recombination subsequently stopped across the pericentromeric region in which it was previously infrequent. Each time recombination became suppressed, TEs and rearrangements accumulated, including accumulating to even higher than previous levels in the pericentromeric regions.

figs. S12 and S13), especially in *S. dioica*. One highly duplicated region (highlighted in gold in Fig. 3C and fig. S13) is shared by both species. In *S. latifolia*, these paralogs show high dS, peaking at 0.241 ± 0.022 (Fig. 3B), similar to the value of 0.213 for *S. latifolia*–*S. vulgaris* divergence and for the oldest recombination suppression event (Fig. 3J). Duplications therefore started accumulating immediately after XY recombination stopped, which was almost contemporaneous with the split of *S. latifolia* and *S. dioica* from the *S. vulgaris* lineage. These duplicates have much higher nonsynonymous site divergence (dN) than for genome-wide paralog pairs with similar dS values (0.158 to 0.316); the respective mean dN/dS values are 0.59 versus 0.38 ($P = 4.6 \times 10^{-26}$, Wilcoxon's test) (Fig. 2D). McDonald-Kreitman tests (*37*) indicated that a few of the *S. latifolia* duplicates may be positively selected ($P < 0.1$; Fig. 3E), possibly increasing their expression from the degenerating Y chromosome, but most are selected against or evolving nearly neutrally (neutrality index = 0.8~1.5, $P > 0.1$, G tests), consistent with the expected weaker purifying selection after duplication (*38*).

Recombination in *S. latifolia* female meiosis is also infrequent across much of the X chromosome (Fig. 3I) (*39*, *40*). Recombinationally inactive regions occupy most of the middle region of all chromosomes, including in *S. vulgaris*, supporting previous results suggesting the presence of physically large pericentromeric regions (*10*, *15*). Consistent with rare recombination, LTR retrotransposons (LTR-TEs) and DNA methylation densities in these regions are almost as high as in the MSY (Fig. 3, F to I, and figs. S6 and S7).

To test whether recombination suppression has occurred, and when, we estimated dS between complete X chromosome coding sequences and their MSY alleles. In both *S. latifolia* and *S. dioica*, genes distant from the PAR have the highest values (*Slat-dS*^{XY} and *Sdio-dS*^{XY}, respectively, in Fig. 3J). Within this oldest stratum, 0, median *Slat-dS*^{XY} values in 1-cM bins range from 0.10 to 0.23, close to the *S. latifolia* and *S. vulgaris* divergence (*dS*^{lv}) estimates. Using either X genetic or physical map positions, *dS*^{XY} declines toward the PAR, which recombines in both sexes (Fig. 3J), confirming results using X genetic map positions without an X assembly (*15*, *40*, *41*). Change-point tests (*42*) detected two changes in *Slat-dS*^{XY} values (with $P = 2.6 \times 10^{-11}$ and $P = 2.7 \times 10^{-3}$), indicating two recombination suppression events (Fig. 3J).

The availability of Y chromosome assemblies allows us to relate Y chromosome degeneration to strata ages. Following the approach developed for humans (*12*), we first defined ancestral genes that are present on the *S. latifolia* or *S. dioica* X chromosome and found in similar proportions across all homologous *S. vulgaris* regions (fig. S14). We used two degeneration criteria: (i) loss of the gene's complete sequence from the Y chromosome and (ii) genes not predicted on the Y chromosome (including disrupted genes). In many stratum 0 windows (corresponding to part of the X chromosome–recombining region to the left of the central rarely recombining region), >70% of *S. latifolia* and *S. dioica* genes are X specific (Fig. 3K), much greater degeneration than was previously estimated without a Y assembly (*43*). Degeneration is less in stratum 1, which includes parts of both recombining X regions plus the >200-Mb pericentromeric rarely recombining region (Fig. 3, I and J; the dotted box includes genes within this small genetic map interval). Ancestral gene losses are even rarer in the small stratum 2 near the PAR (Fig. 3K), which includes partially sex-linked genes in *S. dioica* that stopped recombining in *S. latifolia* since the species split (*41*, *44*, *45*) and shows slightly, but significantly, higher Y-X divergence (fig. S15).

Processes to establish recombinationally inactive regions in sex chromosomes

Chromosome-wide synteny analysis with *S. vulgaris* found that most of the *S. latifolia* and *S. dioica* X chromosomes correspond to a single *S. vulgaris* chromosome, chromosome 4, but portions correspond to three other *S. vulgaris* chromosomes (Fig. 4, A to C, and fig S16), indicating either fusions with other *S. vulgaris* chromosomes before the *S. latifolia* and *S. dioica* split or rearrangements in the *S. vulgaris* lineage. Stratum 0 and part of stratum 1 overlapping the X pericentromeric region correspond with a recombinationally active part of *S. vulgaris* chromosome 4 (Fig. 4C), confirming that the *S. latifolia* and *S. dioica* MSYs newly evolved suppressed recombination. Their homologous strata 0 and 1 regions are arranged differently, and therefore rearrangements must have independently evolved in their MSYs after recombination became suppressed.

In both *S. latifolia* and *S. vulgaris*, recombinationally inactive regions are highly enriched with LTR-*Gypsy* and LTR-*unknown* class TEs (Fig. 4, A and B, $r = -0.68$ to -0.6; figs. S6 and S7 and S17 and S18). To examine the potential involvement of TEs in sex chromosome expansion, we extracted *Gypsy* and *unknown*-class TEs from the *S. latifolia* X- and Y-chromosome sequences and the four syntenic *S. vulgaris* chromosome regions. Using a 20% net divergence threshold, most sequences are specific to one species or to the *S. latifolia* X or Y chromosome (Fig. 4, D and E, and table S6). In clusters shared by the *S. latifolia* X and Y chromosomes, or by both species, recently generated subclusters (bursts defined by dS < 0.1) tended to be more chromosome specific than older ones, as expected (fig. S19 and table S6). Divergence estimates of LTR coding sequences (Fig. 4, F and G) suggest that most insertions occurred recently, consistent with generally being deleterious (*25*). Insertions in the *S. latifolia* X and Y chromosomes tend to be older than those in *S. vulgaris*. Nevertheless, most bursts occurred after the oldest stratum 0 formed (dS > 0.1, see above), and were therefore not the cause of recombination arrest, but rather probably contributed to MSY expansion after recombination stopped. Furthermore, LTR-TE accumulation in the X and Y chromosomes occurred independently in *S. latifolia* and *S. dioica* (based on clusters having net divergence < 0.05; Fig. 4, H to K). Synteny analysis also detected rearrangements between the large *S. latifolia* and *S. dioica* X chromosome pericentromeric regions (fig. S20A), and even within *S. latifolia* (*39*) (figs. S20B and S21), suggesting ongoing rearrangements of these rarely recombining regions, perhaps involving TE activities.

What might have caused suppressed MSY recombination in regions that recombine between the X chromosomes in females? Under the two-gene hypothesis described above, both primary sex-determining genes should be in the region that first evolved suppressed recombination. The GSF, *Y-SlCLV3,* is indeed within stratum 0 (the terminal location of its X-linked copy probably reflects its ancestral position; it is between 84 and 86 Mb of our Y assembly, reflecting MSY rearrangements; Fig. 3A and fig. S8). However, because of the SPF candidate's Y specificity in *S. latifolia* and *S. dioica* and Y rearrangements, its X position cannot be assigned. We found a possible sexually antagonistic gene, *SlBAM1*, near the left end of stratum 0 (fig. S22, A to C) that potentially contributes to *S. latifolia*'s secondary sexual trait dimorphism(s). *BARELY ANY MERISTEM 1/2* (*BAM1/2*)–like receptor-like kinases often function in reproductive processes and flower morphological development through *CLV1/3*-related pathways (*46*–*48*). *SlBAM1* has male-specific *Y-BAM1* sequences in our *S. latifolia* and *S. dioica* individuals, and the gene is also present in the hermaphrodite *S. conica* (fig. S22B). Ectopic expression of *Y-BAM1* in *Nicotiana tabacum* under the regulation of its native promoter substantially increased flower numbers per inflorescence ($P = 0.39 \times 10^{-4}$) and reduced flower size ($P = 0.00015$), as in *S. latifolia* males (fig. S23 and table S7) (*49*). Thus, selection for closer linkage between the primary sex-determining factors and *SlBAM1* might have contributed to stratum 0 evolution.

It is unclear how or why the other strata evolved. Assuming that the ancestral state resembled the present X chromosome, stratum 1 formation involved cessation of recombination between stratum 0 and regions corresponding to parts of *S. vulgaris* chromosomes 2 and 3, including regions that are pericentromeric and recombinationally inactive in females and in

S. vulgaris (Fig. 4D). Sexually antagonistic polymorphisms in this region that were undetected in Y deletion experiments could have favored recombination suppression. However, alternatives are possible. It was recently proposed (*5*, *6*) that successive "lucky" inversions (or other changes) might become fixed among Y chromosomes by chance, extending the nonrecombining region without being favored through involvement of SA polymorphisms. In one model, re-evolution of recombination is prevented by evolution of low expression of Y-linked alleles, accompanied by higher expression of their X counterparts, a form of dosage compensation (*5*). We tested this by quantifying gene expression in young leaves of male and female *S. latifolia*. For genes in both stratum 0 and the probably still degenerating stratum 1, mean male values (normalized to equalize autosomal values in both sexes) are above half those in females (fig. S24 and table S8), confirming the previously reported partial compensation (*16*–*18*). However, in the younger, much less degenerated stratum 2, expression is similar in both sexes (fig. S24), whereas the Lenormand and Roze (*5*) model requires early repression of Y-linked alleles and up-regulation of their X alleles. The expression estimates also do not reflect DNA methylation patterns (fig. S25), as proposed previously (*50*).

A model for evolution of the *Silene* sex chromosomes

Figure 5 summarizes a model for the evolution of the *Silene* Y and X chromosomes based on our results supporting suppressed recombination between factors in X chromosome regions that formerly recombined. Under the two-mutation evolutionary scenario (*20*), the establishment of the (as yet unidentified) SPF mutation creating females was probably followed by the GSF (*SlCLV3*) mutation in a closely linked genome region. Selection favored suppressed XY recombination, creating stratum 0. This was possibly followed by further selection due to the Y-linked *SlBAM1* gene that we found to control a major sexual dimorphism in the species, strongly suggesting that its alleles are sexually antagonistic. Once recombination became rare, TE insertions and rearrangements accumulated in the MSY region, including Y-specific duplications (Fig. 3A and fig. S12A), further hindering recombination within stratum 0 and perhaps also with the pericentromeric region, enlarging the Y-linked region.

It is unknown how or why stratum 1 evolved. Fusions or Y duplications could have expanded the rarely recombining pericentromeric region by shifting crossovers toward the PAR, creating stratum 1. Y deletion maps suggest a pericentric inversion, which could have linked the initial stratum 0 MSY with segregating SA polymorphisms in the pericentromeric region or nearer the current PAR on the Y*q* arm. Our genome sequences thus help to explain the evolution of a giant Y chromoson and why the X chromosome is also large. The fewer X than Y structural rearrangements suggest that its large pericentromeric region does not completely lack recombination but occasionally recombines. TE insertions, including recent female-specific proliferation of some active retrotransposon types (*51*), can nevertheless enlarge this rarely recombining region, and our results show that the *S. latifolia* and *S. dioica* X chromosome pericentromeric regions are independently expanding through TE activities. However, rare recombination will prevent gene loss, and degeneration therefore mainly affects the Y chromosome. The Y chromosome is still actively rearranging, because we detected Y structural differences. It is also still degenerating, although stable heterochromatic epimarks, including DNA methylation, are present throughout the MSY. Many X-linked genes are already hemizygous in males, and dosage compensation has started to evolve in strata 0 and 1, but Y degeneration and X-linked allele up-regulation do not account for the youngest stratum, stratum 2.

REFERENCES AND NOTES

1. R. Ming, A. Bendahmane, S. S. Renner, *Annu. Rev. Plant Biol.* **62**, 485–514 (2011).
2. B. Charlesworth, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5618–5622 (1978).
3. D. Charlesworth, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160456 (2017).
4. D. Charlesworth, *New Phytol.* **224**, 1095–1107 (2019).
5. T. Lenormand, D. Roze, *Science* **375**, 663–666 (2022).
6. P. Jay, E. Tezenas, A. Véber, T. Giraud, *PLOS Biol.* **20**, e3001698 (2022).
7. T. Akagi *et al*., *Nat. Plants* **9**, 393–402 (2023).
8. K. B. Blackburn, *Nature* **112**, 687–688 (1923).
9. D. A. Filatov, F. Monéger, I. Negrutiu, D. Charlesworth, *Nature* **404**, 388–390 (2000).
10. R. Bergero, S. Qiu, A. Forrest, H. Borthwick, D. Charlesworth, *Genetics* **194**, 673–686 (2013).
11. B. T. Lahn, D. C. Page, *Science* **286**, 964–967 (1999).
12. M. A. Wilson Sayres, K. D. Makova, *Mol. Biol. Evol.* **30**, 781–787 (2013).
13. M. V. Chibalina, D. A. Filatov, *Curr. Biol.* **21**, 1475–1479 (2011).
14. R. Bergero, D. Charlesworth, *Curr. Biol.* **21**, 1470–1474 (2011).
15. A. S. Papadopulos, M. Chester, K. Ridout, D. A. Filatov, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13021–13026 (2015).
16. A. Muyle *et al*., *PLOS Biol.* **10**, e1001308 (2012).
17. A. Muyle *et al*., *Nat. Plants* **4**, 677–680 (2018).
18. M. Krasovec, Y. Kazama, K. Ishii, T. Abe, D. A. Filatov, *Curr. Biol.* **29**, 2214–2221.e4 (2019).
19. M. Westergaard, *Naturwissenschaften* **40**, 253–260 (1953).
20. B. Charlesworth, D. Charlesworth, *Am. Nat.* **112**, 975–997 (1978).
21. A. Harkess *et al*., *Nat. Commun.* **8**, 1279 (2017).
22. T. Akagi *et al*., *Plant Cell* **30**, 780–795 (2018).
23. T. Akagi *et al*., *Nat. Plants* **5**, 801–809 (2019).
24. M. Massonnet *et al*., *Nat. Commun.* **11**, 2902 (2020).
25. B. Charlesworth, P. Sniegowski, W. Stephan, *Nature* **371**, 215–220 (1994).
26. T. V. Kent, J. Uzunović, S. I. Wright, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160458 (2017).
27. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, *Bioinformatics* **31**, 3210–3212 (2015).
28. Tomato Genome Consortium, *Nature* **485**, 635–641 (2012).
29. Y. Van de Peer, E. Mizrachi, K. Marchal, *Nat. Rev. Genet.* **18**, 411–424 (2017).
30. Y. Kazama *et al*., *Sci. Rep.* **6**, 18917 (2016).
31. I. S. Donnison, J. Siroky, B. Vyskot, H. Saedler, S. R. Grant, *Genetics* **144**, 1893–1901 (1996).
32. R. Bergero, D. Charlesworth, D. A. Filatov, R. C. Moore, *Genetics* **178**, 2045–2053 (2008).
33. Y. Kazama *et al*., *Mol. Biol. Evol.* **39**, msac195 (2022).
34. N. Fujita *et al*., *Int. J. Mol. Sci.* **20**, 1031 (2019).
35. I. Kurek *et al*., *Plant Mol. Biol.* **48**, 369–381 (2002).
36. J. Liu, R. Ghelli, M. Cardarelli, M. Geisler, *J. Exp. Bot.* **73**, 4818–4831 (2022).
37. J. H. McDonald, M. Kreitman, *Nature* **351**, 652–654 (1991).
38. J. B. Walsh, *Genetics* **139**, 421–428 (1995).
39. J. Yue *et al*., *Curr. Biol.* **33**, 2504–2514.e3 (2023).
40. R. Bergero, A. Forrest, E. Kamau, D. Charlesworth, *Genetics* **175**, 1945–1954 (2007).
41. D. A. Filatov, *J. Evol. Biol.* **35**, 1696–1708 (2022).
42. A. N. Pettitt, *Appl. Stat.* **28**, 126–135 (1979).
43. G. A. Marais *et al*., *Curr. Biol.* **18**, 545–549 (2008).
44. J. L. Campos, S. Qiu, S. Guirao-Rico, R. Bergero, D. Charlesworth, *Heredity* **118**, 395–403 (2017).
45. S. Qiu *et al*., *Mol. Ecol.* **25**, 414–430 (2016).
46. B. J. DeYoung *et al*., *Plant J.* **45**, 1–16 (2006).
47. C. L. H. Hord, C. Chen, B. J. Deyoung, S. E. Clark, H. Ma, *Plant Cell* **18**, 1667–1680 (2006).
48. Z. L. Nimchuk, Y. Zhou, P. T. Tarr, B. A. Peterson, E. M. Meyerowitz, *Development* **142**, 1043–1049 (2015).
49. L. F. Delph, C. R. Herlihy, *Evolution* **66**, 1154–1166 (2012).
50. V. Bačovský, A. Houben, K. Kumke, R. Hobza, *Planta* **250**, 487–494 (2019).
51. J. Puterova *et al*., *BMC Genomics* **19**, 153 (2018).

ACKNOWLEDGMENTS

We thank I. M. Henry and L. Comai (Department of Plant Biology and Genome Center, University of California, Davis) for discussion and comments on this study. **Funding:** This work was supported by PRESTO from the Japan Science and Technology Agency (JST grant JPMJPR20D1) and the Japan Society for the Promotion of Science Grants-in-Aid for Transformative Research Areas (A) (grants 22H05172 and 22H05173 to T.A. and grant 22H05181 to K.S.; Grant-in-Aid for JSPS Fellow JSPS grant JP23KJ1615 to N.F., JSPS KAKENHI grant JP23K05226 to N.F.; grant 23H04747 to Y.I.; and grant 22H02598 to T.I.). **Author contributions:** Conceptualization: T.A., D.C.; Funding acquisition: T.A., N.F., K.S., Y.I.; Investigation: T.A., N.F., K.S., K.N., K.M., A.H., E.K., K.K., R.K., K.N., K.U.; Methodology: T.A., N.F., K.S., H.T., Y.I., A.T., T.I., K.U., D.C.; Project administration: T.A., D.C.; Supervision: T.A., K.U., D.C.; Visualization: T.A., N.F., H.T., K.N., K.M., A.H., K.U.; Writing – original draft: T.A., D.C.; Writing – review & editing: T.A., N.F., K.U., D.C. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All of the genome sequences and the annotated data were deposited to Plant GARDEN (https://plantgarden.jp/en/index) (*S. latifolia*: https://plantgarden.jp/en/list/t37657, *S. vulgaris*: https://plantgarden.jp/en/list/t42043, *S. dioica*: https://plantgarden.jp/en/list/t39879). The raw sequencing data have been deposited in the DDBJ database: Sequence Read Archives database (BioProject ID PRJDB16402, Run ID DRR496111-DRR496477 and DRR540087-DRR540088 for genome sequencing, Hi-C sequencing, DNA methylome and ChIP-seq data, and BioProject ID PRJDB16382, Run ID SAMD00635100-00635106 and SAMD00636369-00636376 for RNA-seq data in flower buds.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adk9074
Materials and Methods
Figs. S1 to S25
Tables S1 to S11
References (*52*–*98*)
MDAR Reproducibility Checklist

Submitted 18 September 2023; resubmitted 8 April 2024
Accepted 13 August 2024
10.1126/science.adk9074