THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Slcyt, a Newly Identified Sex-Linked Gene, Has Recently Moved onto the X Chromosome in Silene latifolia (Caryophyllaceae)

OPEN ACCESS

Slcyt, a Newly Identified Sex-Linked Gene, Has Recently Moved onto the X Chromosome in *Silene latifolia* (Caryophyllaceae)

Vera B. Kaiser, Roberta Bergero, and Deborah Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences, The University of Edinburgh, Edinburgh, United Kingdom

The sex chromosomes of the plant species *Silene latifolia* (white campion) are very young (only 5–10 My old), and all 11 X-linked genes so far described have Y-linked homologues. Theory predicts that X chromosomes should accumulate a nonrandom set of genes. However, little is known about the importance of gene movements between the X and the autosomes in plants, or in any very young sex chromosome system. Here, we isolate from cDNA a new gene, *Slcyt*, on the *S. latifolia* X, which encodes a cytochrome B protein. We genetically mapped *SlCyt* and found that it is located ~1 cM from the pseudoautosomal region. Genes in this region of the X chromosome have low divergence values from their homologous Y-linked genes, indicating that the X only recently stopped recombining with the Y. Genetic mapping in *Silene vulgaris* suggests that *Slcyt* originally belonged to a different linkage group from that of the other *S. latifolia* X-linked genes. *Silene latifolia* has no Y-linked homologue of *Slcyt*, and also no autosomal paralogues seem to exist. *Slcyt* moved from an autosome to the X very recently, as the *Cyt* gene is also X linked in *Silene dioica*, the sister species to *S. latifolia*, but is probably autosomal in *Silene diclinis*, implying that a translocation to the X probably occurred after the split between *S. diclinis* and *S. latifolia*/*S. dioica*. Diversity at *Slcyt* is extremely low (π_{syn} = 0.16%), and we find an excess of high frequency–derived variants and a negative Tajima's *D*, suggesting that the translocation was driven by selection.

Introduction

In the evolution of animal sex chromosomes, gene movements occur both from and to the X chromosome. There are grounds to believe that this is not random but that selection acts to enrich the X (or Z in ZW systems) for genes with sex-specific functions, including sexually antagonistic genes, that is, genes that increase fitness in one sex but are deleterious in the other sex (Rice 1984). In mammals and Drosophila, the X chromosome has a nonrandom gene content, having an overrepresentation of male-specific genes in mammals (Wang et al. 2001; Lercher et al. 2003; Khil et al. 2004; Mueller et al. 2008) and an underrepresentation in Drosophila (Parisi et al. 2003; Oliver and Parisi 2004; Vicoso and Charlesworth 2006); in the chicken, female-specific genes are underrepresented on the Z chromosome (Kaiser and Ellegren 2006). Nonrandom distributions of sex-biased genes could evolve through the evolution of biased expression of initially nonbiased genes that were already located on the X. Alternatively, genes could be recruited onto the X from the autosomes. For instance, the *Drosophila melanogaster* Y carries mainly male function genes, none of which has an X-linked homologue (Brosseau 1960; Carvalho 2002). It could be advantageous for a female-beneficial antagonistic gene to move to the X if the amount of gene product is directly related to its copy number, because such a movement would lead to lower expression in males. However, this will depend on several factors, including details of the dosage compensation system. If dosage compensation occurs, and it acts on the whole chromosome or large X regions, such a change could be either selectively neutral or deleterious. If, however, expression in the two sexes is equalized by upregulating the X chromosome in males, as in Drosophila (Gupta et al. 2006), genes moved onto the X will not gain increased relative female expression. Finally, if dosage compensation occurs in females, for example, by inactivating one X, and increasing expression from the other X, as in mammals (Nguyen and Disteche 2006), translocating female-beneficial genes to compensated regions of the X might be disfavored in females, because expression relative to autosomal genes will often be reduced.

Both mammals and Drosophila have genetically degenerated Y chromosomes, and gene movements to and from their X chromosomes have been documented (Betrán et al. 2002; Emerson et al. 2004). These movements occurred over long evolutionary times, during most of which dosage compensation operated for at least some X-linked genes. The mammalian sex chromosomes evolved about 170 MYA, after the marsupial and Eutherian mammal lineages split from the ancestor of the platypus (Veyrunes et al. 2008), consistent with very high X–Y sequence divergence of a few genes (Lahn and Page 1999; Ross et al. 2005). In Drosophila, the XY system is at least 39 My old, and in *D. melanogaster*, the X is Muller's chromosome element A, the ancestral Drosophila X chromosome (Carvalho 2002; Ashburner et al. 2005). The bird sex chromosomes also stopped recombining very long ago (reviewed in Nam and Ellegren 2008).

Studying the evolution of the X gene content in younger sex chromosome systems should shed light on how this content evolves, and plants are of interest for such studies. The highest silent site divergence values between homologous X- and Y-linked genes of *Silene latifolia* (Caryophyllaceae) are just over 20%, suggesting that the sex chromosomes evolved only about 5–10 MYA (Filatov 2005; Nicolas et al. 2005; Bergero et al. 2007). Recombination in the regions nearest to the pseudoautosomal region (PAR) ceased much more recently than in other regions, and several genes near the *S. latifolia* PAR have X–Y silent site divergence below 5% (Nicolas et al. 2005; Bergero et al. 2007). Similarly, in mammals, genes on the X and Y fall into four or five "strata" of diminishing evolutionary ages as their locations get closer to the PAR, suggesting stepwise or gradual recombination suppression, though on a timescale much longer than in *S. latifolia* (Lahn and Page 1999; Nam and Ellegren 2008). These findings

are consistent with the hypothesis that, in both these X chromosomes, sexually antagonistic genes may have accumulated over time, leading to selective pressure for suppressed recombination in regions where recombination formerly occurred, to maintain associations between such genes and the sex-determining genes.

However, this is indirect evidence, as is all other currently available evidence for sexually antagonistic genes on evolving sex chromosomes (Ellegren and Parsch 2007; Mank et al. 2008; Mank and Ellegren 2009). One reason for studying plant sex chromosomes is the hope that such genes may be discovered. Dioecious plants have few secondary sexual differences, but flowers and inflorescences of males and females often differ (Lloyd and Webb 1977). In S. latifolia, the two sexes differ in their optimum numbers and sizes of flowers, leading to a genetic conflict, and females with fewer, larger flowers produce sons that have fewer, larger flowers than the average male (Delph et al. 2002, 2004; Prasad and Bedhomme 2006). In other dioecious plants, traits not associated with flower morphology also differ between the sexes, for example, leaf resin content, leaf and stem morphology, senescence patterns, and plant–herbivore interactions (Krischik and Denno 1990; Merzouki et al. 1996; Cornelissen and Stiling 2005), suggesting that genes expressed differentially in males and females may be common in dioecious plants. Currently, however, sex-specific expression in plants is known only for genes expressed exclusively in reproductive tissue (Sather et al. 2005; Yu et al. 2008).

The gene content of plant sex chromosomes is also currently largely unknown, except for the Y chromosomes of the moss Marchantia polymorpha (Ishizaki et al. 2002) and S. latifolia. In S. latifolia, 11 Y-linked genes have now been described, all apparently functional, except for the incomplete MROS3-Y gene (Guttman and Charlesworth 1998) and possibly SlssY (Filatov 2008). Here, we describe the first case of a gene movement to the S. latifolia X. The gene, Slcyt, was recently translocated from an autosome in S. latifolia and inserted close to the PAR of the X chromosome. This rearrangement could have led to suppressed recombination in the region, and, as we discuss below, the movement might have been driven by sexual antagonism. The Slcyt gene seems to have been affected by a recent selective sweep. The only other gene movement known in a plant sex chromosome system is the duplicative transposition of a gene to the Y, and the Y copy has increased expression in stamens, compared with the autosomal one, suggesting that this was probably also an advantageous gene movement (Matsunaga et al. 2003). Recently, it has also been discovered by cytogenetic studies that Silene diclinis has an X-autosome translocation not present in S. latifolia (Howell et al. 2009). No genes in the added region have yet been studied. Our present study suggests that the X-autosome translocation involving the Slcyt gene is a separate event from that in S. diclinis. It has so far been generally accepted that no additions to this sex chromosome pair had occurred during its evolution, but that it evolved from a single ancestral autosome that can be identified by its gene content (Filatov 2005), but these new results show that at least part of the X near the PAR has recently been added to one or both of the XY chromosome pair.

Materials and Methods
Silene DNA Samples

The study used S. latifolia, Silene dioica, S. diclinis, and Silene vulgaris plants from natural populations, which are described below. Silene dioica and S. diclinis have the same XY system as S. latifolia, that is, genes in the "older" regions of the S. latifolia Y chromosome stopped recombining with the X before this group of dioecious species split; the sequences of these genes form distinct X and Y clusters, rather than clustering by their species of origin (Nicolas et al. 2005). Silene dioica is the sister species to S. latifolia, and these species often hybridize in nature (Baker 1948; Filatov et al. 2001; Laporte et al. 2005). Silene vulgaris, the outgroup species used in this study, is gynodioecious and has no sex chromosomes (Desfeux et al. 1996).

Individuals were grown in the greenhouse at the University of Edinburgh and DNA was extracted from fresh leaves using the Fast DNA kit (Q-biogene), following the manufacturer's instructions.

Sex-linked and autosomal genes were identified in the mapping families F2005-4 and H2005-1, which are F2 families descended from crosses between geographically distant populations (Bergero et al. 2007). For putatively X-linked genes, we genotyped the two F1 individuals (parents of the F2) of H2005-1 and scored 92 F2 offspring for variants found in the maternal and/or paternal plant. For each gene, the inheritance patterns of the two variants within the family were compared with the pattern obtained for previously published X-linked genes.

For mapping genes in S. vulgaris, two families (named SV1 and SV2) were used, with 51 and 64 offspring, respectively. These families were generated by crossing a female plant (E2000 5/9, from Dijon, France) with two unrelated hermaphroditic plants (H2000-4 and 99K-10-4, from Flamanville, France, and Sussex, United Kingdom, respectively). A linkage map of the markers was inferred using the software JoinMap (Stam 1993).

Identifying Sex-Linked Genes

A set of S. latifolia gene sequences were isolated from a cDNA library constructed by a simplified version of the template-switching (TS) procedure of Matz et al. (1999). First-strand cDNA was synthesized from total RNA extracted from male leaf primordia, using 50 mM of oligo(dT)$_{21}$ (5′-GATCGATTTTTTTTTTTTTTTTTTTT-TVN-3′), 30 mM MgCl$_2$, and 200-U reverse transcriptase Superscript II (Invitrogen, Paisley, United Kingdom), following the manufacturer's recommendation. The TS adapter (5′-GGTTTTGGTAGTTCTGTGTGTTGGG-3′) was ligated to the 5′ ends of cDNAs in a 50-μl reaction containing 5 units of Klenow fragment $3′ → 5′$ exo$^−$ (New England Bioloabs, Ipswich, MA), 1× buffer 2 (New England Biolabs), 1 mM deoxynucleotide triphosphates, and 50 pmol of TS adapter. The reaction was carried out at 16 °C, overnight. After incorporation of the TS adapter, the cDNA was purified on a Qiagen spin column, polymerase chain reaction (PCR) amplified using the primer pair for the TS adapter and an oligo(dT)$_{21}$ primer,

Table 1
Lengths of the Region That Includes the Microsatellite SlX_STR1 in Intron 1 of *Slcyt* in Female *Silene latifolia* Individuals from Various Natural Populations

Individual	Location	Length of Amplicon (bp)
K2005-2B	France	171, 176
K2005-1B2	Germany	161, 176
G2005-5/3	Greece	148, 161
G2005-4	Greece	135, 145
J2006-1/3	Denmark	159, 177
E2004-17/3	Netherlands	149, 177
K2005-3A/3	Poland	155, 171
K2005-4/3	Austria	174, 214
K2005-9/3	Estonia	162, 171
K2005-2A/3	France	165, 222
K2005-7/3	Norway	165

and finally cloned in a T-tailing pBS vector (Stratagene, La Jolla, CA). We refer to this as the TS library.

Candidate sex-linked genes in *S. latifolia* were identified using a combination of segregation analysis of intron size variants (ISVS), using a universal primer (5′-GGT-TGGAGCTAGTGTTGTG-3′) labeled with 6-FAM or VIC (Applied Biosystems, Foster City, CA), and direct sequencing, as described in Bergero et al. (2007). Briefly, we first identified putative intron positions by comparing the *S. latifolia* cDNA sequences with the translated *Arabidopsis thaliana* and *Oryza sativa* (rice) genome sequences, using BlastX at www.ncbi.nlm.nih.gov. PCR primers were then designed from the *S. latifolia* cDNA sequences flanking putative introns, using the Oligonucleotide Properties Calculator available at http://www.basic.northwestern.edu/biotools/oligocalc.html. The PCR conditions using the labeled universal forward primer were generally as follows: 10 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 1 min, followed by 25 cycles of 94 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min; final extension at 72 °C for 30 min. PCR conditions without the universal labeled forward primer: 10 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 1 min, followed by 25 cycles of 94 °C for 30 s, 52 °C for 30 s, 72 °C for 1 min. To detect size differences that could be used as genetic markers, the PCR products were run on 1.5% agarose gels and inspected visually.

For sequences that did not yield suitable size variants, capillary electrophoresis was performed on an ABI 3730 capillary sequencer (Applied Biosystems). The labeled forward primer allowed length variants among the PCR products to be scored using the Genemapper software package 3.7 (Applied Biosystems). PCR amplicons that showed no length variants were directly sequenced, and sequences were examined in Sequencer 4.5 to detect variants suitable for segregation analysis after digestion with restriction enzymes. Finally, if no suitable restriction sites were found, genotyping for segregation analysis was performed by direct sequencing to detect polymorphic variants.

Obtaining the Complete *Slcyt* Sequence

Because the cDNA sequence that allowed us to determine whether *Slcyt* is sex linked (see Results) did not con-

tain the whole *Slcyt* coding sequence, nested PCR was used to obtain the 5′ end of the gene. The TS cDNA library was used as template in a first round of PCR using a primer for the TS adapter and the reverse gene-specific primer *Slcyt*_b_R (supplementary table 1, Supplementary Material online). One microliter of the first-round PCR mixture was used in a second round PCR with primer for the TS adapter and the gene-specific primer *Slcyt*_R, which binds internally to the PCR product from *Slcyt*_b_R (supplementary table 1, Supplementary Material online). The PCR products were cloned, sequenced in an ABI 3730 capillary sequencing machine (Applied Biosystems), and visualized using Sequencher 4.5 software.

Intron 1 proved difficult to amplify but was amplified with Phusion enzyme (Finnzymes, Espoo, Finland) and a PIKO 24 thermal cycler (Finnzymes), following the manufacturer′s instructions (PCR conditions: 40 s at 98 °C, followed by 10 cycles of 5 s at 98 °C, 5 s at 62 °C, 5 min at 72 °C; 25 cycles of 5 s at 98 °C, 5 s at 55 °C, 5 min at 72 °C; final extension at 72 °C for 5 min).

Analyses of DNA Sequence Diversity

To study DNA sequence diversity of *Slcyt* within *S. latifolia*, parts of the gene (starting within exon 2, and including the whole of intron 2 and exon3, and parts of the 3′ untranslated region (UTR), see fig. 2) were amplified in the same 48 European male plants, from 24 different European populations, covering the entire native range of the species (the samples are listed in supplementary table 2, Supplementary Material online) that were used to confirm absence of a Y-linked copy (see Results), using the primers *Slcyt*_ex2_cons_F and *Slcyt*_3_R (3′ UTR) (supplementary table 1, Supplementary Material online). The PCR products were directly sequenced and edited in Sequencher. Sequence diversity was analyzed using DnaSP software, which was also used for several tests of neutrality, including Tajima's D, Fu and Li's $D*$ and $F*$, Fay and Wu's H, and Fu's F statistics (Tajima 1989; Fu and Li 1993; Fu 1997; Fay and Wu 2000). The significance levels for these tests were calculated using coalescent simulations implemented in DnaSp, assuming no recombination (Filatov 2008), which is a conservative approach (Tajima 1989; Wall 1999). The Hudson-Kreitman-Aguadé (HKA) test (Hudson et al. 1987) was used to compare *Slcyt* diversity levels with those of *X7*, *Cyp-X*, *X4*, and *SlX9*, another new X-linked gene, which will be described elsewhere (Kaiser VB, Bergero R, Charlesworth D, in preparation).

In the course of studying sequence diversity of *Slcyt*, we identified a polymorphic $(TTA)_n$ microsatellite in intron 1 (referred to below as SlX_STR1). This was scored using the primers *Slcyt*-I1-F-univ.-1 and *Slcyt*-I1-R-2 (supplementary table 1, Supplementary Material online), and the sizes of the amplicons were determined using GeneMapper 3.7. Ten of 11 females from different natural populations were heterozygous for SlX_STR1 (table 1). To further test for complete sex linkage (and exclude a pseudoautosomal location for *Slcyt*), SlX_STR1 was amplified in the 48 males listed in supplementary table 2, Supplementary Material online. Finding heterozygous males would indicate the presence of a Y-linked copy, or a pseudoautosomal location.
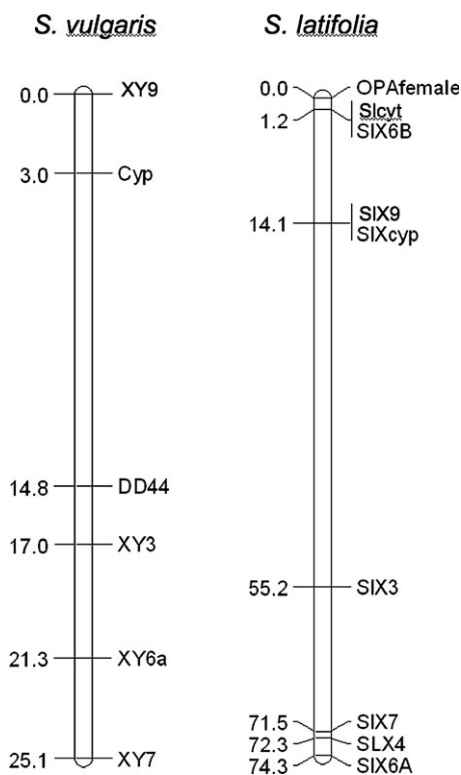
Fig. 1.—Genetic map of X-linked genes in *Silene latifolia* and their homologues in *Silene vulgaris*. The *cyt* gene falls into a different linkage group in *S. vulgaris* and is therefore not shown.

Results

Identification of Two New X-Linked Genes in *S. latifolia*

A total of 16 *S. latifolia* cDNA sequences were screened for sex linkage using the ISVS method, direct sequencing, and/or restriction digestion (see Materials and Methods). The PCR primers used for mapping are listed in supplementary table 1, Supplementary Material online. Two new genes were found to be sex linked, giving a total of 13 genes so far identified on the X and/or Y chromosome in this species. They were provisionally named *Slcyt* and *SlX9/SlY9*.

Sex linkage of the first gene, *Slcyt*, was established by direct sequencing and segregation analysis of an single nucleotide polymorphism (SNP) variant (G/C) in intron 2. The male parent of the F2005-4 mapping family previously studied (Bergero et al. 2007) carried a G, whereas the female parent was homozygous for C. Seven male offspring were scored, and all inherited the mother's variant, whereas all four female offspring scored were heterozygous for both variants, strongly suggesting that the gene is X linked. Further evidence for X-linkage is given below.

The other new sex-linked gene, *SlX9*, was mapped using the mapping family H2005-1 (Bergero et al. 2007), by scoring a size variant in intron 2 directly on an agarose gel. Variants in the *SlX9/SlY9* gene found exclusively in the father and male offspring (Y-linked sequences) were identified using segregation analysis of a marker that was heterozygous in both the maternal and paternal plants. This gene will be described in more detail elsewhere (Kaiser VB, Bergero R, Charlesworth D, in preparation).

Genetic Mapping of *Slcyt* and *SlX9*

Slcyt and *SlX9* were genetically mapped using segregation analysis among the 92 F2 offspring of the mapping family H2005-1. The maternal plant was heterozygous for both genes. For *Slcyt*, the F2 offspring were scored for SlX_STR1. Again, there were no heterozygous males among the offspring, and no recombinants were found between *Slcyt* and *SlX6b* among the 92 offspring scored, supporting the conclusion that *Slcyt* is X linked, and is located in the part of the X that does not recombine with the Y in males. *SlX6b* maps at a distance of ~1 cM from the pseudoautosomal marker, OPA (Bergero et al. 2007). One of the mother's 2 SlX_STR1 alleles appeared in only 2 of 43 males and 4 of 49 females. This aberrant ratio ($\chi^2 = 93.56$, $P < 0.001$, 1df) is similar to results for *SlX6b* (Bergero R, unpublished data) and suggests either a bias in transmission via female gametophytes or higher mortality of offspring carrying this maternal X chromosome. No such bias was detected in five other mapping families tested (results not shown), but these families confirmed X-linkage of *Slcyt*. No Y-linked copy was detected (see below).

The same intron size variant that was used to establish sex linkage was also used to map *SlX9* in relation to the other known X-linked genes. *SlX9* maps to the same position as *SlCypX*, a previously described sex-linked gene located about 14 cM from the pseudoautosomal marker (Bergero et al. 2007).

Mapping Orthologues of X-Linked Genes in *S. vulgaris*

Because no Y-linked copy was detected, we tested linkage of the *Slcyt* orthologue in *S. vulgaris* (which we denote by *Svcyt*), to test whether the gene has changed its chromosomal location. Segregating indels in introns were therefore used as markers to map as many *S. vulgaris* orthologues of *S. latifolia* X-linked genes as possible. Segregating ISVS and SNP markers were obtained for six *S. vulgaris* genes, the homologues of the X-linked genes *SlX3*, *CypX*, *SlX7*, *SlX6a*, *DD44X*, and *SlX9* and scored in the progeny of families SV1 and SV2 (see Materials and Methods). Neither mapping family had segregating indels in the introns of *Svcyt*, so the segregation of two SNPs in intron 1 was scored in family SV1 by direct sequencing; informative variants were present in both the maternal and paternal parents. Figure 1 shows the consensus map obtained for the two *S. vulgaris* families. The mapped genes fall into a single *S. vulgaris* linkage group, spanning 25 cM, except for *Svcyt*, suggesting that this gene moved to the X of *S. latifolia* from another location. We next describe the analysis of this gene in more detail.

Slcyt Gene Structure and Function

The complete coding sequence and most of the intron sequences were obtained for *Slcyt*. The gene structure was inferred by comparing the *S. latifolia* cDNA sequence with the genomic sequence (fig. 2). BlastX searches were performed at www.ncbi.nlm.nih.gov to identify homologous genes in other organisms and their functions. The best hit in the *A. thaliana* genome sequence was to a member
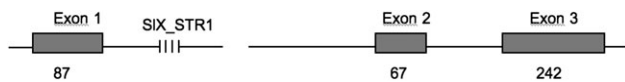
Fig. 2.—Gene structure of *Slcyt*. Sizes in bp of exons and the position of the microsatellite SlX_STR1 are given. There is a gap in the sequence of about 70 bp in intron 1, which has a total length of about 730 bp.

of a family of cytochrome B5 proteins (Blast e-value $2e^{-41}$) that has 61% amino acid identity to *Slcyt*. *Slcyt* also shows significant similarity to cytochrome B proteins in the dicotyledonous plant *Vernicia fordii* (Euphorbiaceae), and in the monocotyledons *Sorghum bicolor* and *Triticum monococcum* (Blast e-values $1e^{-46}$ to $4e^{-34}$). The homologue of the *Slcyt* gene was sequenced in *S. vulgaris*, *S. dioica*, and *S. diclinis*. Divergence values between *S. latifolia* and the other species are shown in table 2. The *Slcyt* gene has probably remained functional in the species studied. We isolated it from cDNA, showing that the gene is transcribed, and we found no premature stop codons or frameshifts in any of the sequences, and, in comparisons between *S. latifolia* and *S. vulgaris*, K_a is smaller than K_s, suggesting that purifying selection has been acting on the gene.

Searches for a Y-Linked Copy of *Slcyt* in *S. latifolia* and Its Close Relatives *S. dioica* and *S. diclinis*

No heterozygotes were found among 48 males screened for the microsatellite SlX_STR1 in the *Slcyt* gene, even though we found 19 length variants among these males (supplementary table 2, Supplementary Material online), and most females scored were heterozygous (table 1). It therefore seems likely that there is no Y copy. Given that *Slcyt* maps close to the PAR, divergence between the X-linked gene and any Y-linked copy is expected to be low, and conserved primers should amplify a Y-linked copy of *Slcyt*, if present.

To more rigorously test for the possibility that a Y-linked copy of *Slcyt* exists, several approaches were tried. First, primers designed to match regions of conserved amino acid identity between *Slcyt* in *S. latifolia* and the *A. thaliana* or *S. vulgaris* homologues were used to amplify genomic DNA and cDNA from male individuals. No heterozygous SNPs were detected in any of six males tested, four of which were F1 individuals from crosses between different populations. Second, 48 males from 24 different European populations (see above) were sequenced in our diversity analysis, and none carried any SNPs in the region.

Table 2
Synonymous (K_s) and Nonsynonymous (K_a) Divergence Values between *Slcyt* and the Sequences of the Orthologous *Cyt* Gene in Three Related Species

Species Compared	K_s (%)	K_a (%)	Length of Sequence (bp)
Silene latifolia–Silene Vulgaris	10	6	273
S. latifolia–Silene Dioica	1.6	1.3	288
S. latifolia–Silene diclinis	1.3	0.4	351

The lengths of the alignments are given in base pairs (bp).

Table 3
Lengths of the Region That Includes the Microsatellite SlX_STR1 in Intron 1 of the *Cyt* Gene, in 20 Offspring of a *Silene dioica* Female Sampled from a Wild Population (Family A2009-1, See Text)

Lengths Observed (bp)			Number of Females	Number of Males
146	168	180		
−	+	+	3	0
+	−	+	6	0
+	−	−	5	3
−	−	+	0	3
Total			14	6

We did not obtain any *Slcyt* sequence (not even the X sequence) using an approach in which amplification requires only one primer (either forward or reverse) to match the Y-linked copy (Guttman and Charlesworth 1998). The primers tested were *Slcyt*_ex3_cons_R, *Slcyt*–exon 2-R, *Slcyt*_ex1_cons_F, *Slcyt*_F (exon 3), Slcyt_b_R (exon 3), and *Slcyt*_ex2_cons_F.

To test whether there is a Y-linked copy in *S. dioica*, we amplified the microsatellite SlX_STR1 in 20 offspring of an *S. dioica* female from a population in Finland, whose genotype at this locus is unknown, using the same primer combination as for *S. latifolia* and scored the length variants using Genemapper. Among the progeny, there were three alleles, 146, 168, and 180 bp. Overall, nine females were heterozygous (three 168/180 and six 146/180) and five were 146-bp homozygotes (table 3). In all six males, only one allele was detected (either 146 or 180 bp, see table 3). This suggests X linkage and no Y copy (assuming a heterozygous 146/180 maternal parent, which mated with at least three males, carrying 146, 180, and 168 bp variants on the X). Autosomal inheritance is not excluded by these results alone (a 146/180 female mated with at least two males). However, it is very unlikely that all six male offspring would be homozygous by chance (Fisher's exact test, df = 1, $P < 0.05$).

We also genotyped 11 male and 6 female *S. dioica*, sampled from the wild. All male individuals again each had only one length variant, whereas four females were heterozygotes (table 4).

In *S. diclinis*, however, all three males and three females scored carried two length variants for *Cyt*. These plants were derived from seeds collected from one female from Spain and were therefore full or half siblings (family A2000-13). Direct sequencing of one male and one female individual from this family (individuals A2000-13-4 and A2000-13-10) confirmed that the PCR products (which span a region of 1,162 sites from exon 1 to within exon 3) were indeed two copies of *Cyt*, although they differ by about 70 bp in length due to an insertion in intron 1. There were, however, no SNP variants, which suggests that these are two alleles of the same gene, rather than paralogues. Without a mapping family or population samples from *S. diclinis*, we cannot currently test for sex linkage in this species, so the two copies could be autosomal or in the PAR of the X and Y. A pseudoautosomal location would result in incomplete sex linkage of two paternal variants in the F1 of a cross and in different variants being

Table 4
Lengths of the Region That Includes the Microsatellite SlX_STR1 in Intron 1 of *Cyt* in Male and Female *Silene dioica* Individuals from Natural Populations

Individual	Location	Sex	Length of Amplicon (bp)
99M24.2	Unknown	Male	129
9-1	Unknown	Male	146
FS01	Scotland	Male	139
FS02	Scotland	Male	136
FS03	Scotland	Male	133
FS04	Scotland	Male	136
FS05	Scotland	Male	146
FS06	Scotland	Male	139
FS07	Scotland	Male	136
FS08	Scotland	Male	136
FS09	Scotland	Male	133
99K24.1	France	Female	129,136
99K22.7	France	Female	129,136
FS10	Scotland	Female	136,146
FS11	Scotland	Female	136,146
FS12	Scotland	Female	139
FS13	Scotland	Female	139

incompletely associated with the sexes in population samples, so this is potentially testable in the future. It would be interesting to test this, because the *S. diclinis* X is now known to have undergone a translocation of an autosome region.

Sequence Diversity Analyses

To test whether the movement of the *Slcyt* gene to the *S. latifolia* X chromosome was a selected event, we analyzed this sequence further. *Silene latifolia* genomic DNA sequences were obtained for a region starting within exon 2, and including the whole of intron 2 and exon 3, and parts of the 3′ UTR, and diversity was analyzed using the European sample of *S. latifolia* males described above. We obtained direct sequences of 45 X-linked copies. The diversity of *Slcyt* was very low: Synonymous diversity, π_S, was 0.16% and nonsynonymous diversity, π_A, was 0.07%, based on 237 coding sites. The π_A/π_S ratio is 0.45. The aligned noncoding positions (173 sites) also yielded a low diversity estimate (0.16%).

The HKA test (Hudson et al. 1987) was used to compare *Slcyt* diversity levels with those at four other X-linked genes, *SlX9, X7, X4,* and *Cyp-X*, by testing the within-species polymorphism of *Slcyt*, after correcting for divergence from their *S. vulgaris* orthologues. The test was significant for the comparisons of *Slcyt* versus *SlX9, X4* and *Cyp-X* ($P < 0.05$, $P < 0.05$, and $P < 0.01$, respectively); for the comparison of *X7* versus *Slcyt*, $P < 0.06$, but the

Table 5
Tests of Neutrality Statistics *Slcyt*

Tajima's D	Fu and Li's D^*	Fu and Li's F^*	Fu's F	Fay and Wu's H
−1.6	−2.66	−2.72	−2.6	−1.82
$P < 0.05$	$P < 0.01$	$P < 0.01$	$P < 0.05$	$P < 0.01$

number of *X7* sites available was much smaller than for the other three genes (191 bp vs. 367, 863, 804 bp, for *SlX9, X4,* and *Cyp-X*, respectively), and the power was low. We therefore conclude that the low diversity at *Slcyt* is unlikely to be explained by neutral processes alone, such as a very low mutation rate at the locus.

The low *Slcyt* diversity suggests the possibility of a selective sweep. We therefore performed Tajima's *D* test, which compares the two estimates of diversity, the nucleotide diversity, π, based on pairwise comparisons of allele sequences, and θ (which is based on the number of segregating sites). At equilibrium under neutrality, the two estimates are expected to be the same, and Tajima's *D* will be close to zero (Tajima 1989), whereas recent directional selection (positive or purifying selection) leads to an excess of low frequency variants, and a negative Tajima's *D*. The other tests shown in table 5 also detect variants at frequencies differing from the neutral distribution. All the tests performed on *Slcyt* were significant; a skewed frequency spectrum at segregating sites, together with an excess of high-frequency variants, which represent the derived state, suggest the action of positive selection, rather than selection against deleterious variants.

Discussion

The sex chromosomes in *S. latifolia* are believed to have evolved from a single pair of ordinary autosomes, consistent with the fact that all 11 sex-linked genes previously described have Y-linked counterparts (Delichere et al. 1999; Atanassov et al. 2001; Moore et al. 2003; Filatov 2005; Nicolas et al. 2005; Bergero et al. 2007), and that all four genes previously mapped in *S. vulgaris* (Filatov 2005) are on a single autosome, and so is the newly discovered X-linked gene *SlX9* mapped here. *Slcyt* is the first gene discovered on the *S. latifolia* X that has been acquired from a different genomic location in the course of X-chromosome evolution in this genus, and is now located close to the *SlX6b* gene. In female meiosis, both genes map only about 1 cM from the pseudoautosomal marker, OPA, discovered by Di Stilio et al. (1998). Divergence between *SlX6b* and its Y-linked homologue is low ($K_s = 4.5\%$), showing that the X stopped recombining with the Y chromosome only recently at this locus (Bergero et al. 2007). All 46 males from the different European populations contained only a single, hemizygous copy of *Slcyt*. Thus, a translocation was probably involved, which is now fixed within the species. Because no autosomal paralogous copy was detected, the chromosomal segment probably moved onto the X chromosome only (and not the Y also) and was lost from another genomic region at the same time.

It is unknown whether the event affected *Slcyt* alone or whether neighboring genes were translocated in the same event. *Slcyt* contains introns, showing that it is not a retrogene, so a larger, segmental event is possible, forming a neo-X chromosome. If the event indeed translocated a substantial region into the *S. latifolia* PAR, it might have directly prevented recombination between the X and Y chromosomes in this region (and, once fixed, would not affect recombination between X chromosomes, though it

might initially have been deleterious for female fertility in heterozygotes, as is true for many translocations). This event could be similar to that in the medaka fish, where a duplication created a new copy of the autosomal *dmrt1* gene on another autosome, immediately isolating a small region close to the new copy from its homologue (Kondo et al. 2006). In medaka, unlike the case of *Slcyt*, the original copy is still present. More information on *S. vulgaris* or *S. latifolia* sequences, including sequencing bacterial artificial chromosome clones containing the *S. latifolia Slcyt* gene, should allow genes adjacent to *Slcyt* to be found, which would help determine the size of the insertion, and whether any of these genes could have been under sexually antagonistic selection, and could thus have selected for the translocation. Another interesting consequence of movement of a large genome region, containing several essential genes, from an autosome to the X, is that this could explain why YY plants in *S. latifolia* are inviable, because these plants would entirely lack these genes. This alternative to classical genetic degeneration of the Y chromosome has not previously been considered. Such a genome rearrangement would have to be sufficiently advantageous in females to drive the change, however.

Our evidence gives some support for the hypothesis that the translocation was driven by selection. *Cyt* appears to be X linked in *S. dioica* as well as *S. latifolia*, but not in *S. diclinis*, in which we found two copies in both females and males, so that it is probably still autosomal. Thus *Slcyt* probably moved to the X after the split of *S. latifolia*/*S. dioica* from *S. diclinis*. This is consistent with a skew in the frequency spectrum still being observable at segregating sites in the *Slcyt* gene. The combination of strongly reduced diversity and a highly skewed frequency spectrum of segregating sites, reflecting a recently reduced effective population size, suggests positive selection when the translocation became fixed (though we cannot exclude the possibility of a selective sweep after its fixation). A recent rapid nonselective fixation could also lead to a uniform haplotype at the locus (Tajima 1990). It seems unlikely, however, that such a nonselective event would produce the strongly significant results we find for Tajima's tests and the other tests in table 5, but this possibility should be tested in the future using neutral models.

If selection was involved, it need not have involved the *Slcyt* locus itself. The protein encoded is probably used in the mitochondrial electron transport system. *Slcyt* may therefore be a housekeeping gene, in which case sexual antagonism involving *Slcyt* itself might not be the main driver behind the gene's movement, and one would have to assume selection on another gene translocated in the same event (Jiang et al. 2001; Jin et al. 2001). However, a connection is well established between mitochondrial electron transport and cytoplasmic male sterility in plants, possibly due to the high metabolic needs of anther development (Warmke and Lee 1978; Chase 2007), and so it is possible that *Slcyt* has some important function in males, and that its loss would benefit females due to sexual antagonism. It will thus be very interesting in the future to compare gene expression of *Slcyt* in males and females and in sex-specific reproductive tissues.

Supplementary Material

Supplementary tables 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

Literature Cited

Ashburner M, Golic KG, Hawley RS. 2005. Drosophila: a laboratory handbook. Cold Spring Harbour (NY): Cold Spring Harbour Laboratory Press.

Atanassov I, Delichere C, Filatov DA, Charlesworth D, Negrutiu I, Moneger F. 2001. Analysis and evolution of two functional Y-linked loci in a plant sex chromosome system. Mol Biol Evol. 18:2162–2168.

Baker HG. 1948. Stages in invasion and replacement demonstrated by species of *Melandrium*. J Ecol. 36:96–119.

Bergero R, Forrest A, Kamau E, Charlesworth D. 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. Genetics. 175:1945–1954.

Betrán E, Thornton K, Long M. 2002. Retroposed new genes out of the X in Drosophila. Genome Res. 12:1854–1859.

Brosseau GE. 1960. Genetic analysis of the male fertility factors on the Y-chromosome of *Drosophila-Melanogaster*. Genetics. 45:257–274.

Carvalho AB. 2002. Origin and evolution of the Drosophila Y chromosome. CurrOpin Genet Dev. 12:664–668.

Chase CD. 2007. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet. 23:81–90.

Cornelissen T, Stiling P. 2005. Sex-biased herbivory: a meta-analysis of the effects of gender on plant–herbivore interactions. Oikos. 111:488–500.

Delichere C, Veuskens J, Hernould M, Barbacar N, Mouras A, Negrutiu I, Moneger F. 1999. SIY1, the first active gene cloned from a plant Y chromosome, encodes a WD-repeat protein. EMBO J. 18:4169–4179.

Delph LF, Gehring JL, Frey FM, Arntz AM, Levri M. 2004. Genetic constraints on floral evolution in a sexually dimorphic plant revealed by artificial selection. Evolution. 58:1936–1946.

Delph LF, Knapczyk FN, Taylor DR. 2002. Among-population variation and correlations in sexually dimorphic traits of *Silene latifolia*. J Evol Biol. 15:1011–1020.

Desfeux C, Maurice S, Henry JP, Lejeune B, Gouyon PH. 1996. Evolution of reproductive systems in the genus Silene. Proc R Soc Lond, Ser B Biol Sci. 263:409–414.

Di Stilio VS, Kesseli RV, Mulcahy DL. 1998. A pseudoautosomal random amplified polymorphic DNA marker for the sex chromosomes of *Silene dioica*. Genetics. 149:2057–2062.

Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nat Rev Genet. 8:689–698.

Emerson JJ, Kaessmann H, Betran E, Long MY. 2004. Extensive gene traffic on the mammalian X chromosome. Science. 303:537–540.

Fay JC, Wu CI. 2000. Hitchhiking under positive Darwinian selection. Genetics. 155:1405–1413.

Filatov DA. 2005. Evolutionary history of *Silene latifola* sex chromosomes revealed by genetic mapping of four genes. Genetics. 170:975–979.

Filatov DA. 2008. A selective sweep in or near the *Silene latifolia* X-linked gene SlssX. Genet Res. 90:85–95.

Filatov DA, Laporte V, Vitte C, Chartesworth D. 2001. DNA diversity in sex-linked and autosomal genes of the plant species *Silene latifolia* and *Silene dioica*. Mol Biol Evol. 18:1442–1454.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147:915–925.

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. Genetics. 133:693–709.

Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD, Eastman PS, Oliver B. 2006. Global analysis of X-chromosome dosage compensation. J Biol. 5:3.

Guttman DS, Charlesworth D. 1998. An X-linked gene with a degenerate Y-linked homologue in a dioecious plant. Nature. 393:263–266.

Howell EC, Armstrong SJ, Filatov DA. Forthcoming 2009. Evolution of Neo-sex Chromosomes in *Silene diclinis*. Genetics.

Hudson RR, Kreitman M, Aguadé M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics. 116:153–159.

Ishizaki K, Shimizu-Ueda Y, Okada S, Yamamoto M, Fujisawa M, Yamato KT, Fukuzawa H, Ohyama K. 2002. Multicopy genes uniquely amplified in the Y chromosome-specific repeats of the liverwort *Marchantia polymorpha*. Nucleic Acids Res. 30:4675–4681.

Jiang M, Ryu J, Kiraly M, Duke K, Reinke V, Kim SK. 2001. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. Proc Natl Acad Sci USA. 98:218–223.

Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nat Genet. 29:389–395.

Kaiser VB, Ellegren H. 2006. Nonrandom distribution of genes with sex-biased expression in the chicken genome. Evolution. 60:1945–1951.

Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat Genet. 36:642–646.

Kondo M, Hornung U, Nanda I, et al. (11 co-authors). 2006. Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. Genome Res. 16:815–826.

Krischik VA, Denno RF. 1990. Patterns of growth, reproduction, defense, and herbivory in the Dioecious shrub *Baccharis–Halimifolia* (Compositae). Oecologia. 83:182–190.

Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. Science. 286:964–967.

Laporte V, Filatov DA, Kamau E, Charlesworth D. 2005. Indirect evidence from DNA sequence diversity for genetic degeneration of the Y-chromosome in dioecious species of the plant Silene: the SlY4/SlX4 and DD44-X/DD44-Y gene pairs. J Evol Biol. 18:337–347.

Lercher MJ, Urrutia AO, Hurst LD. 2003. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. Mol Biol Evol. 20:1113–1116.

Lloyd DG, Webb CJ. 1977. Secondary sex characters in plants. Bot Rev. 43:177–216.

Mank JE, Ellegren H. 2009. Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. Evolution. 63:1464–1472.

Mank JE, Hultin-Rosenberg L, Webster MT, Ellegren H. 2008. The unique genomic properties of sex-biased genes: insights from avian microarray data. BMC Genom. 9:148.

Matsunaga S, Isono E, Kejnovsky E, Vyskot B, Dolezel J, Kawano S, Charlesworth D. 2003. Duplicative transfer of a MADS box gene to a plant Y chromosome. Mol Biol Evol. 20:1062–1069.

Matz M, Shagin D, Bogdanova E, Britanova O, Lukyanov S, Diatchenko L, Chenchik A. 1999. Amplification of cDNA ends based on template-switching effect and step-out PCR. Nucleic Acids Res. 27:1558–1560.

Merzouki A, M'Rabet RT, Mesa JM. 1996. Sex differentiation of dioecious hemp (*Cannabis sativa* L.) with a precocious morphocinetic character. Archiv Geobot. 2:165–169.

Moore RC, Kozyreva O, Lebel-Hardenack S, Siroky J, Hobza R, Vyskot B, Grant SR. 2003. Genetic and functional analysis of DD44, a sex-linked gene from the dioecious plant *Silene latifolia*, provides clues to early events in sex chromosome evolution. Genetics. 163:321–334.

Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA. 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet. 40:794–799.

Nam K, Ellegren H. 2008. The Chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. Genetics. 180:1131–1136.

Nguyen DK, Disteche CM. 2006. Dosage compensation of the active X chromosome in mammals. Nat Genet. 38:47–53.

Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, Vyskot B, Mouchiroud D, Negrutiu I, Charlesworth D, Moneger F. 2005. A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. Plos Biol. 3:47–56.

Oliver B, Parisi M. 2004. Battle of the Xs. Bioessays. 26:543–548.

Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B. 2003. Paucity of genes on the Drosophila X chromosome showing male-biased expression. Science. 299:697–700.

Prasad NG, Bedhomme S. 2006. Sexual conflict in plants. J Genet. 85:161–164.

Rice WR. 1984. Sex-chromosomes and the evolution of sexual dimorphism. Evolution. 38:735–742.

Ross MT, Grafham DV, Coffey AJ, et al. (279 co-authors). 2005. The DNA sequence of the human X chromosome. Nature. 434:325–337.

Sather DN, York A, Pobursky KJ, Golenberg EM. 2005. Sequence evolution and sex-specific expression patterns of the C class floral identity gene, SpAGAMOUS, in dioecious *Spinacia oleracea* L. Planta. 222:284–292.

Stam P. 1993. Construction of integrated genetic-linkage maps by means of a new computer package—Joinmap. Plant J. 3:739–744.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–595.

Tajima F. 1990. Relationship between DNA polymorphism and fixation time. Genetics. 125:447–454.

Veyrunes F, Waters PD, Miethke P, et al. (14 co-authors). 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res. 18:965–973.

Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet. 7:645–653.

Wall JD. 1999. Recombination and the power of statistical tests of neutrality. Genet Res. 74:65–79.

Wang PJ, McCarrey JR, Yang F, Page DC. 2001. An abundance of X-linked genes expressed in spermatogonia. Nat Genet. 27:422–426.

Warmke HE, Lee SLJ. 1978. Pollen abortion in T-cytoplasmic male-sterile corn (*Zea mays*)—suggested mechanism. Science. 200:561–563.

Yu Q, Steiger D, Kramer EM, Moore PH, Ming R. 2008. Floral MADS-box genes in trioecious papaya: characterization of AG and AP1 subfamily genes revealed a sex-type-specific gene. Trop Plant Biol. 1:97–107.

Hideki Innan, Associate Editor

Accepted July 2, 2009