

RESEARCH ARTICLE

Open Access

Comparative analysis of a plant pseudoautosomal region (PAR) in *Silene latifolia* with the corresponding *S. vulgaris* autosome

Nicolas Blavet^{1*†}, Hana Blavet^{2†}, Radim Čegan², Niklaus Zemp¹, Jana Zdanska², Bohuslav Janoušek², Roman Hobza^{2,3} and Alex Widmer¹

Abstract

Background: The sex chromosomes of *Silene latifolia* are heteromorphic as in mammals, with females being homogametic (XX) and males heterogametic (XY). While recombination occurs along the entire X chromosome in females, recombination between the X and Y chromosomes in males is restricted to the pseudoautosomal region (PAR). In the few mammals so far studied, PARs are often characterized by elevated recombination and mutation rates and high GC content compared with the rest of the genome. However, PARs have not been studied in plants until now. In this paper we report the construction of a BAC library for *S. latifolia* and the first analysis of a > 100 kb fragment of a *S. latifolia* PAR that we compare to the homologous autosomal region in the closely related gynodioecious species *S. vulgaris*.

Results: Six new sex-linked genes were identified in the *S. latifolia* PAR, together with numerous transposable elements. The same genes were found on the *S. vulgaris* autosomal segment, with no enlargement of the predicted coding sequences in *S. latifolia*. Intergenic regions were on average 1.6 times longer in *S. latifolia* than in *S. vulgaris*, mainly as a consequence of the insertion of transposable elements. The GC content did not differ significantly between the PAR region in *S. latifolia* and the corresponding autosomal region in *S. vulgaris*.

Conclusions: Our results demonstrate the usefulness of the BAC library developed here for the analysis of plant sex chromosomes and indicate that the PAR in the evolutionarily young *S. latifolia* sex chromosomes has diverged from the corresponding autosomal region in the gynodioecious *S. vulgaris* mainly with respect to the insertion of transposable elements. Gene order between the PAR and autosomal region investigated is conserved, and the PAR does not have the high GC content observed in evolutionarily much older mammalian sex chromosomes.

Keywords: BAC library, Pseudoautosomal region, PAR, *Silene latifolia*, Sex chromosome, Evolution

Background

Heteromorphic sex chromosomes (XY/ZW) can often be distinguished from autosomes by the absence of recombination in at least a part of their length and Y/W chromosome degeneration [1,2]. Plants with sex chromosomes have evolved rarely but repeatedly in many plant lineages, and sex chromosomes have reached various levels of differentiation [3]. In *Asparagus officinalis* and *Carica papaya*

for example, X and Y chromosomes have diverged little and recombine along most of their length [4-6] whereas in *Rumex acetosa* and *Silene latifolia*, the sex chromosomes in males are largely non-recombining [7,8]. In *S. latifolia*, X and Y chromosomes can recombine only in the regions known as pseudoautosomal regions (PARs). Westergaard [9] originally identified one PAR on each of the q-arms of *Silene latifolia* sex chromosomes. Later, Lengerova *et al.* (2003) [8] using fluorescent in situ hybridization (FISH) revealed that the X PAR is located on the p-arm, whereas the PAR on the Y chromosome is located on the q-arm. More recently, Scotti and Delph [10] proposed that PARs exist on both ends of the X and Y chromosomes, similar to

* Correspondence: nicolas.blavet@env.ethz.ch

†Equal contributors

¹Institute of Integrative Biology (IBZ), ETH Zürich, Universitätstrasse 16, Zürich, 8092, Switzerland

Full list of author information is available at the end of the article

the situation in humans [11]. A further similarity to mammalian sex chromosomes is that the *S. latifolia* sex chromosomes diverged gradually [12], which led to the formation of evolutionary strata. Comparisons between the evolutionarily young *S. latifolia* sex chromosomes (about 10 million years [12,13]) and those of eutherian mammals (about 110 million years [14]) have revealed that similar processes are involved in the evolution of sex chromosomes in both animals and plants.

The sex chromosomes of *S. latifolia* most likely evolved from a single pair of autosomes as previously shown [12,13], with one autosome of the gynodioecious relative *S. vulgaris*, a species lacking sex chromosomes, carrying homologues of *S. latifolia* sex-linked genes [12,13]. *Silene latifolia* and *S. vulgaris* have the same haploid chromosome number ($n = 12$), but differ substantially in genome size. The *Silene latifolia* haploid genome is 2646 Mbp in females [15], with the X chromosome being about 400 Mbp in length [16], whereas the haploid genome size of *S. vulgaris* is 1103 Mbp [17] and autosomes are about 100 Mbp long.

In this study we analyzed a part of the *S. latifolia* PAR located on the p-arm of the X chromosome and on the q-arm of the Y chromosome (henceforth referred to as PAR) and of the corresponding *S. vulgaris* autosome, in order to study collinearity and divergence between these chromosome parts and to assess whether the *S. latifolia* PAR has characteristics in common with animal PARs. Furthermore, we investigate whether the *S. latifolia* size increase relative to *S. vulgaris* reflects the increase in size of the entire X chromosome or more closely resembles the increase seen in *S. latifolia* autosomes.

In mammalian genomes, PARs have several interesting properties including increased GC content, higher mutation rates and a level of recombination higher than in the rest of the genome [18,19] due to the necessity for crossing over in this region [20]. PARs in mice and the human PAR1 appear to serve a critical function in spermatogenesis, as indicated by the fact that their absence prevents X and Y chromosome segregation during male meiosis, causing male sterility [21-23]. However, PARs differ widely in size among mammals (covering about 4 % of the Y chromosome in humans [24,25] and mice [26], about 8 % in cattle [27] and about 24 % in dogs [28,29]), with most eutherians sharing the same genes situated closest to the telomere but having the pseudoautosomal boundary (PAB), separating the PAR from the sex-specific part of the sex chromosomes, at variable positions [27]. In mice, the PAB is located in the gene *Fxy*. Exons 1-3 are located in the X specific part, while exons 4-10 are located in the PAR [30]. The segment of this gene located in the PAR has a higher GC content than its X-specific portion [30,31].

In order to analyze the *S. latifolia* PAR and the corresponding region on the *S. vulgaris* autosome, we first established and screened a bacterial artificial chromosome

(BAC) library of *S. latifolia* with the marker ScOPA09 that has previously been found to be located in the PAR of the closely related dioecious species *S. dioica* [32] and has successfully been identified and used for mapping *S. latifolia* sex chromosomes [12]. The marker ScOPA09 is located in the *S. latifolia* PAR which is known to recombine once per generation in males [33] and makes up about 10 % of the Y chromosome [34,35]. In *S. vulgaris*, the marker OPA is lacking. We therefore first sequenced a clone of the *S. latifolia* BAC library containing the marker ScOPA09. Sequencing was performed by Sanger and 454 pyrosequencing to explore the suitability of different sequencing strategies for BAC assembly. From these sequences we identified new markers and used them to screen the *S. vulgaris* BAC library for a homologous clone. Both BAC sequences were assembled into >100,000 bp-long scaffolds using GS De Novo Assembler (Roche).

Here we present the results of a genomic comparison between an area located in the PAR of the X chromosome p-arm and in the q-arm of the Y chromosome of the dioecious plant species *S. latifolia* and its homologous autosomal area in the closely related gynodioecious species *S. vulgaris*. Our results identify the first physically mapped genes located in the *Silene* PAR and reveal characteristics of a plant pseudoautosomal region.

Results and discussion

Our study reports the first comparative analysis of a BAC sequence from a plant pseudoautosomal region and the corresponding autosomal area in a related species that lacks sex chromosomes. Comparative mapping of a limited number of sex-linked genes in *S. latifolia* and autosomal genes in *S. vulgaris* has previously demonstrated large-scale synteny between the X chromosome of *S. latifolia* and one *S. vulgaris* autosome [12,13]. Our results provide the first evidence for small-scale synteny and strong collinearity at the gene level within a restricted region of the *S. latifolia* sex-chromosomes, the PAR located in the p-arm of the X chromosome and in the q-arm of the Y chromosome, and the corresponding *S. vulgaris* autosomes.

BAC sequencing, assembly and annotation

The 454 paired-end sequencing of both a *S. latifolia* BAC clones containing marker ScOPA09 and of a homologous BAC clone from *S. vulgaris* gave more than 150,000 reads for each BAC clone. These were assembled into 171,870 bp and 116,096 bp-long scaffolds for *S. latifolia* and *S. vulgaris*, respectively.

We found a total of twenty-eight homologous sequences (seventeen different accession numbers) with the *A. thaliana* proteome. Of these, nine were found in both *Silene* species, two were identified only in *S. vulgaris* (one of them twice), and six were found only in *S. latifolia*. A total of 16 out of 28 sequences are most likely transposable or

Table 1 Putative transposable elements identified in *Silene* BAC clones

Found in	TAIR accession	Function	ProDom annotation	<i>Silene</i> repeats coverage
<i>S. latifolia/S. vulgaris</i> ^a	ATMG00860	Mitochondrion/ hypothetical protein	Transposable element	0 %
<i>S. latifolia/S. vulgaris</i>	AT4G23160	Cysteine-rich receptor-like protein kinase 8/ polyprotein	Transposable element	80 %
<i>S. latifolia</i>	ATMG00710	Mitochondrion/ hypothetical protein	Transposable element	100 %
<i>S. vulgaris</i> ^a	AT3G01410	Putative RNase H	Transposable element	77 %
<i>S. latifolia</i>	ATMG00310	Mitochondrion/ hypothetical protein	Transposable element	49 %
<i>S. latifolia</i>	AT5G41980	Uncharacterized protein	Transposable element	0 %
<i>S. latifolia/S. vulgaris</i>	AT2G01050	Nucleic acid binding / zinc ion binding/ uncharacterized protein	Transposable element	25 %
<i>S. latifolia/S. vulgaris</i>	AT1G43760	Uncharacterized protein	Transposable element	20 %
<i>S. latifolia</i>	AT4G20520	RNA binding / RNA-directed DNA polymerase/ uncharacterized protein	Transposable element	0 %
<i>S. latifolia</i>	ATMG01250	Mitochondrion/ hypothetical protein	Transposable element	23 %

^aThe sequence has been found twice in the BAC clone.

Putative transposable elements were identified by BLASTX searches against the *Arabidopsis thaliana* proteome (TAIR10) with an E-value cut-off of 1E-4, ProDom annotations and coverage analysis with the *Silene* repeated sequences database [37].

repeated elements as indicated by their annotations extracted from the protein domain family database ProDom [36] and repeat coverage (Table 1). The repeat coverage is based on BLAST hits with a *Silene* repeated elements library [37].

Among the nine sequences shared by the two *Silene* species, the areas matching TAIR accessions AT4G23160 (*CRK8*), AT2G01050, AT1G43760 and ATMG00860 contain repeated elements. Moreover, in *S. vulgaris*, the part matching ATMG00860 is contained in a match with the transposon sequence Q3I6J4_SILLA, which also includes the sequence matching accession number AT3G01410 (Table 1). Using annotated *Silene* transposable elements [34] we found that the region matching *CRK8* is similar to a *Copia*-like retrotransposon [38], and that the region matching Q3I6J4_SILLA is similar to a *Retand*-like retrotransposon (see Kejnovský *et al.* (2006) for description [39]). Moreover, the sequences matching AT2G01050, AT1G43760 and *CRK8* are found in both scaffolds at different positions, which provide further evidence that these sequences are transposable elements.

The five remaining sequences correspond to new pseudoautosomal genes. They are homologues of the *A. thaliana* genes *ESPI* (AT4G22970), *BIP1* (AT5G28540), *ACBP1* (AT5G53470), and of genes AT5G53500 and AT5G41970. These latter two genes we named PAR1 and PAR2, respectively (Table 2).

Finally, two other gene sequences, AT3G15000 and AT4G27700, that we named PAR3 and SVA1, were found on the BACs of *S. latifolia* and *S. vulgaris*, respectively (Table 2). However, because the BAC sequences only partly overlap, we do not have the homologous copies of these genes in the other species. The PAR3 gene in *S. latifolia* corresponds to a putative pseudoautosomal gene. The *S.*

vulgaris SVA1 sequence is homologous to a gene coding for a rhodanese protein in *A. thaliana* (information collected from TAIR [<http://arabidopsis.org/>]).

Genes located in PARs close to the PAB (less than 50 cM) often present sex-specific expression [40]. Using RNA-seq data from Muyle *et al.* [41] we found no evidence for sex-biased expression of the genes located in the *S. latifolia* PAR (Additional file 1: Table S1).

GC content

In mammalian PARs, high recombination associated with biased gene conversion (BGC) [42-44] results in a high GC content [27,30,45,46]. Comparisons of GC and GC3 content between the PAR and non-PAR regions of the human X chromosome revealed a higher GC and GC3 content in the pseudoautosomal region [42]. Further studies of the human PAR revealed that the GC content decreases from 64 % close to the telomeric region to 55 % in the middle of the PAR and is only 38 % close to the pseudoautosomal boundary (PAB) [19]. Similar declines of the GC content were also found in other mammals, including cattle [27] and murine species [47].

A recent analysis of sequence polymorphisms in plants has revealed that the mating system affects GC content, with a higher content in outcrossing compared to selfing taxa being observed, but this effect is significant only in Poaceae that are known to have unusual GC contents [48]. Whether this effect is due to BGC and why it is observed only in Poaceae, however, is not clear. To date, evidence for BGC in plant sex chromosomes is lacking, but given the relatively small size of the investigated PAR (about 10 % of the Y chromosome [34,35]) in *S. latifolia*, which is comparable to PAR size in mammals, and the fact that recombination occurs during meiosis, a higher recombination rate is

Table 2 Putative genes identified in *Silene* BAC clones

Gene name	Found in	TAIR accession	Putative function	ProDom annotation	%identity
<i>ESP1</i>	<i>S. latifolia/S. vulgaris</i>	AT4G22970	Separase	Separase	34/36
<i>BIP1</i>	<i>S. latifolia/S. vulgaris</i>	AT5G28540	Luminal-binding protein 1	ATP-binding	80/74
<i>ACBP1</i>	<i>S. latifolia/S. vulgaris</i>	AT5G53470	Acyl-CoA-binding domain-containing protein 1	Lipid-binding	44/42
<i>PAR1</i>	<i>S. latifolia/S. vulgaris</i>	AT5G53500	WD-40 repeat family protein	Hydrolase phosphatase	41/42
<i>PAR2*</i>	<i>S. latifolia/S. vulgaris</i>	AT5G41970	Uncharacterized protein	Metal dependent	56/70
<i>PAR3</i>	<i>S. latifolia</i>	AT3G15000	Uncharacterized protein	Plastid developmental protein DAG	54/-
<i>SVA1</i>	<i>S. vulgaris</i>	AT4G27700	Rhodanese-like domain-containing protein	Rhodanese	-/71

% identity with *A. thaliana* sequence is given first for *S. latifolia* and then for *S. vulgaris*. * *PAR2* is truncated approximately by half in *S. latifolia*. Putative genes were identified using BLASTX searches against the *Arabidopsis thaliana* proteome (TAIR10) with an E-value cut-off of 1E-4, ProDom annotations were used to determine putative function or detect repeats and transposable elements.

expected in this region as compared with other regions of the genome.

In contrast to mammals, the *Silene latifolia* PAR can directly be compared with a homologous autosomal region in a closely related species. If the *S. latifolia* PAR has an increased GC content compared to autosomes, then this should be detectable in a comparative analysis. We then determined the GC and GC3 contents for each gene (Table 3). A comparison between *S. latifolia* and *S. vulgaris* revealed no significant difference (GC content: $t = 0.0638$, $p\text{-value} = 0.9507$; GC3 content: $t = 0.0521$, $p\text{-value} = 0.9597$). Moreover we determined the GC and the GC3 content of nine sex-linked genes that had previously been identified and are located in the sex-specific region of the X chromosome [49-57] (Additional file 2: Table S2). No difference was found in the GC and GC3 content between the newly identified PAR genes and the genes located in the sex-specific region of the X chromosome ($t = 0.601$, $p\text{-value} = 0.5656$ and $t = 0.6295$, $p\text{-value} = 0.5437$, respectively). These results indicate that the pattern typical for mammalian PARs is not present in the investigated part of the *S. latifolia*

PAR. This may indicate that the *S. latifolia* PAR maintains its "autosomal" features, as the sex chromosomes in this species are evolutionarily young. Alternatively, the studied region of the *S. latifolia* PAR might be close to the pseudoautosomal boundary where the GC content is lower than in more distal PAR areas, as has been found to be the case in most mammals [19,27,46]. Indeed, the ScOPA09 marker was estimated to be located 15 cM from the pseudoautosomal boundary (PAB) in *S. dioica* [32]. We obtained a very similar estimate of 11 cM for *S. latifolia* in this study (both calculations were done using Kosambi and Haldane mapping functions, for details see Additional file 3: Table S3). Even though we presently do not know the physical distance between the PAB and the BAC clone studied here, the results of our genetic analysis clearly show that all genes identified in this study are located in the PAR, as evidenced by their cosegregation with marker ScOPA09 (Additional file 4: Table S4). Furthermore, these genes are recombining with the same recombination frequency of 11 % with the PAB (Additional file 3: Table S3).

Structure comparison

Large-scale collinearity between the *S. latifolia* X chromosome and *S. vulgaris* autosome has repeatedly been reported in studies of *S. latifolia* sex chromosome evolution [13,52,58,59]. In addition to large-scale collinearity we here report the presence of small-scale collinearity spanning five genes whose linear arrangement is conserved. We also show that the length of these genes is identical between the studied *Silene* species. Indeed, while we assessed whether the investigated *S. latifolia* PAR presents signs of chromosome enlargement, because the X chromosome is about four times the size of a *S. vulgaris* autosome, we analyzed exon and intron lengths of the five genes previously reported (Table 4). We then compared with a Student's t-test whether the average size of both introns and exons of the different genes between both species were similar and we found no significant difference between *S. latifolia* and *S. vulgaris* ($t = -0.0817$, $p\text{-value} =$

Table 3 GC and GC3 content comparison

Gene name	<i>ESP1</i>	<i>BIP1</i>	<i>ACBP1</i>	<i>PAR1</i>	<i>PAR2</i>
Sl sequence length (bp)	5416	2013	1125	2049	507
Sl exon GC content (%)	42.4	47.3	48.3	42.3	48.3
Sl exon GC3 content (%)	41.2	53.5	42.9	38.1	49.7
Sv sequence length (bp)	5416	2013	1125	2049	507
Sv exon GC content (%)	42.4	47.3	47.7	42.4	48.3
Sv exon GC3 content (%)	41.1	53.7	42.9	37.3	49.1
Sl putative gene GC content*	38.2	42.3	38.8	40.4	39.9
Sv putative gene GC content*	37.5	42.8	39.2	40.2	40.6

Sl: *Silene latifolia*, Sv: *S. vulgaris*.

*Exon and intron sequences were taken into account.

GC and GC3 contents were computed on gene sequences of identical size in both *S. latifolia* and *S. vulgaris*. *PAR2* is truncated in the BAC sequence of *S. latifolia* and *ESP1* is truncated in *S. vulgaris*. In these cases, only partial sequences with coverage in both species were compared.

Table 4 Comparison of gene sizes

Gene name	ESP1	BIP1	ACBP1	PAR1	PAR2
Sl sequence length (bp)	15455	3401	4909	3107	1127
Sl exons	24	6	6	7	3
Sl exon length (bp)	5416	2019	1143	2049	510
Sl intron length (bp)	10039	1382	3766	1058	617
Sv sequence length (bp)	16677	3217	5042	3043	1138
Sv exons	24	6	6	7	3
Sv exon length (bp)	5422	2046	1143	2049	507
Sv intron length (bp)	11255	1171	3899	994	631
% identity: entire gene	80.7	85.6	73.6	92.6	92.8
% identity: exons	96.5	96.9	92.9	97.1	95.9
Sv/Sl introns	1.1	0.9	1.0	0.9	1.0

Sl: *Silene latifolia*, Sv: *S. vulgaris*.

Genes PAR2 and ESP1 are truncated in *S. latifolia* and *S. vulgaris* respectively, due to their positions at the end of the investigated BAC clones. For these genes, only the fragment covered in both species was reported below.

0.9369 and $t = -0.005$, $p\text{-value} = 0.9961$ for intron and exon comparisons respectively). The conserved intron size may indicate a functional role in gene regulation. Indeed, introns enlarged by repetitive elements were found to affect gene expression in rice [60]. However, substantial differences in length occur in intergenic regions due to transposable element insertions. Figure 1 presents the global alignment of both BAC sequences. Intergenic regions are highly diverged in size and consequently major gaps are visible in the alignment. We considered as intergenic all regions in-between the five genes described in this paper for which there are copies in both *Silene* species. The total length of the intergenic region in the *S. latifolia* BAC is 115,909 bp, and 71,866 bp in the *S. vulgaris* BAC. This difference of 44,043 bp is highly significant ($\chi^2 = 30823.1$, $df = 3$, $p\text{-value} < 2.2e-16$) and corresponds to a 61 % increase of the *S. latifolia* chromosome size as compared to *S.*

vulgaris. This increase is due to the insertion of transposable elements in the *S. latifolia* PAR region.

Microsatellite comparison

We found 577 and 377 microsatellite loci in *S. latifolia* and *S. vulgaris* respectively. A comparison of the average proportion of microsatellite loci (mono-, di-, tri- and tetra-nucleotide microsatellite) between both *Silene* species using a Student's t-test revealed no significant difference ($t = 0.5461$, $p\text{-value} = 0.5867$). A previous analysis of microsatellites in plants revealed a negative correlation between microsatellite frequency and genome size [61]. The *Silene latifolia* X chromosome is about four times larger than the *S. vulgaris* autosome. However, our results revealed that the *S. latifolia* PAR contains a similar density of microsatellite repeats as the *S. vulgaris* autosomes, suggesting that microsatellites play no or only a minor role in the size increase of the *Silene latifolia* PAR. Then, we searched for long-mer microsatellite accumulation in *S. latifolia* and *S. vulgaris* (see Methods), which are expected to be rare in the PAR [62]. We found one (ATC)₁₀ microsatellite locus in *S. latifolia* and one occurrence each of (T)₃₅, (ATA)₁₆, (AAG)₁₉ and (TTA)₂₅ in *S. vulgaris*. The low density of long-mer microsatellites observed here confirms the previously reported paucity of microsatellite repeats on the *S. latifolia* X chromosome PAR [62].

Transposable element insertion

We found three transposable elements containing long terminal repeats (LTR) in *S. latifolia* and two in *S. vulgaris*. The estimates of the invasion of these elements vary from about 17,200 years ago to 766,000 years ago (Additional file 5: Table S5). These elements were inserted after the divergence of the *Silene* sex chromosomes approximately 5 ~ 10 million years ago [12], which may be an indication of highly active transposable elements in the PAR. However, given that we have observed a smaller than expected size increase in the PAR (about 1.6x instead of about 4x), we hypothesize

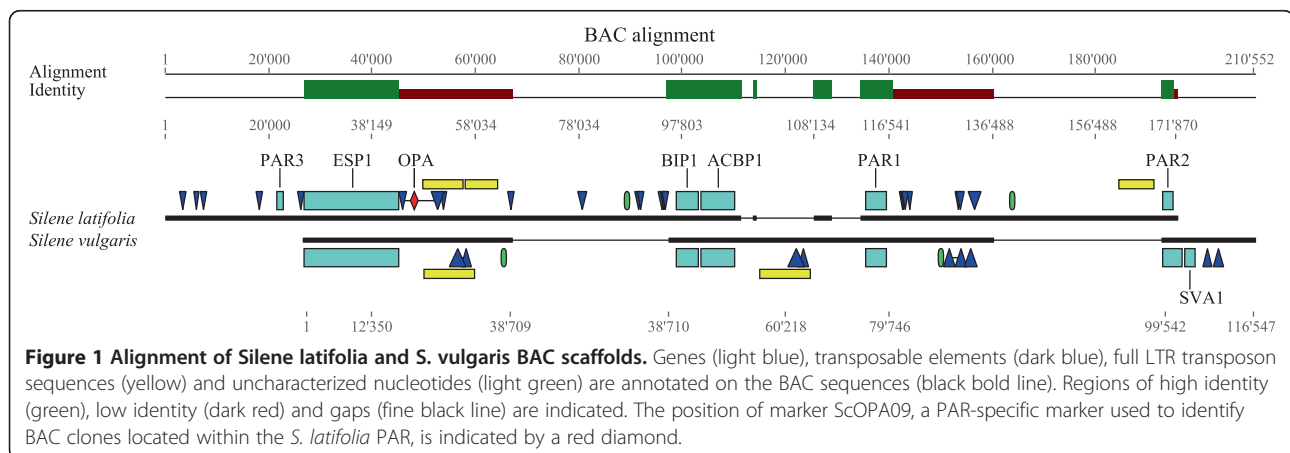


Figure 1 Alignment of *Silene latifolia* and *S. vulgaris* BAC scaffolds. Genes (light blue), transposable elements (dark blue), full LTR transposon sequences (yellow) and uncharacterized nucleotides (light green) are annotated on the BAC sequences (black bold line). Regions of high identity (green), low identity (dark red) and gaps (fine black line) are indicated. The position of marker ScOPA09, a PAR-specific marker used to identify BAC clones located within the *S. latifolia* PAR, is indicated by a red diamond.

that the enlargement of the X chromosome occurs mainly in the non-PAR areas of the X chromosome and is due to large-scale accumulation of different tandem repeats [63,64] and retrotransposons [65,66]. The observed larger size of the studied *S. latifolia* PAR segment in comparison to the *S. vulgaris* autosome is close to the approximate difference in size between autosomes in *S. latifolia* and *S. vulgaris* and may therefore not reflect the size increase seen in the sex chromosomes.

Conclusions

In this study we present the first analysis of a fragment belonging to the *S. latifolia* pseudoautosomal region located in the p-arm of the X chromosome and the q-arm of the Y chromosome. The analysis of BAC sequences revealed five new pseudoautosomal genes that are conserved in size and linear arrangement between *S. latifolia* and *S. vulgaris*, indicating small-scale gene collinearity between the X chromosome and the corresponding autosomal region. No increase in GC or GC3 content was found in the studied PAR area, indicating that either the evolutionarily young *S. latifolia* PAR is not GC rich or alternatively, that the studied region is close to the pseudoautosomal boundary, where no increase in GC content is expected. A structural comparison revealed that non-coding regions of the *S. latifolia* PAR contain multiple transposable and repeated elements and are overall about 61 % longer than in *S. vulgaris*. This size increase is similar to the size difference between *S. latifolia* and *S. vulgaris* autosomes and may therefore reflect a genome-wide, rather than a sex chromosome-specific trend. Our study reports the first comparative analysis of a partial pseudoautosomal region in a plant which we compare to a closely related species lacking sex chromosomes, thereby providing new insights into genome size and sex chromosome evolution in *Silene latifolia*.

Methods

BAC preparation

S. latifolia and *S. vulgaris* were grown from seeds in a climate chamber. Fresh leaf material was harvested after initiation of flowering. *Silene latifolia* leaves were snap frozen in liquid nitrogen, packaged in dry ice and shipped to Amplicon Express, Pullman, Washington, where the BAC library was constructed from high molecular weight (HMW) genomic DNA following the method described by Tao et al. (2002) [67]. The *S. vulgaris* BAC library was assembled in the Institute of Experimental Botany of the AS CR Laboratory of Molecular Cytogenetics and Cytometry, Olomouc. In summary, DNA was partly digested with *Hind*III and inserted into the pECBAC1 vector. Ligations were transformed into DH10B *E. coli* cells (Invitrogen) and plated on LB agar containing appropriate concentrations of chloramphenicol, X-gal and IPTG. Clones were robotically

selected with a Genomic Solution G3 and transferred into 384 well plates, grown for 18 h, replicated and frozen at -80° C. In order to identify positions and plate numbers of each clone, they were placed on a grid in duplicate on Hybond N + (Amersham, Biosciences) nitrocellulose membranes in a 4x4 pattern. The membranes were incubated and processed as described by Bouzidi et al. (2006) [68]. The *S. latifolia* BAC library was arrayed on six membranes of 18,432 colonies and one membrane containing 9,216 clones. The average insert-size of the library is 128 kb. The *S. vulgaris* BAC library was arrayed on three membranes with 18,432 colonies each. The average insert-size of the library was 110 kb [50]. Probes for radioactive hybridizations were labeled with α^{32} P using the Prime-It II Random Primer Labelling Kit (Stratagene) according to the manufacturer's protocol. The presence of the marker in positive BACs was verified by PCR. BAC DNA was isolated with the Large Construct Kit (Qiagen). The *S. latifolia* BAC clone containing the marker ScOPA09 was then sequenced. In order to test which method was most suitable for subsequent BAC assembly, we used three sequencing methods: shotgun Sanger sequencing, 3 kb and 8 kb 454 paired-end pyrosequencing on a GS-FLX machine (Roche). Sanger sequencing was performed by the GATC Biotech laboratory in Konstanz, Germany, and 454 sequencing by the Functional Genomics Center Zurich (FGCZ). As the ScOPA09 marker is not present in *S. vulgaris*, we developed suitable markers from neighboring loci using the *S. latifolia* BAC sequence in order to identify a homologous BAC clone in *S. vulgaris*. Primers used for screening the BAC library are presented in Table 5.

Sequencing, assemblies and annotations

Assembly of the initial shotgun sequences of the *S. latifolia* BAC lead to the identification of numerous relatively short contigs because of the presence of multiple repeat regions. We therefore tested two different 454 pair-end sequencing methods (3 kb and 8 kb) in order to overcome these problems. Both approaches provided similar results, but because the 8 kb paired-end sequencing method requires more DNA, we used 3 kb paired-end sequencing for the

Table 5 PCR primers used to amplify fragments of genes located in the *S. latifolia* PAR

Gene name	Forward-primer (5' – 3')	Reverse-primer (5' – 3')	Product size (bp)
<i>ESP1</i>	AAATACCCAGCCCGTAGCTT	TGCTCAATACATGCCTCCAG	414
<i>BIP1</i>	CGAAAGATGAAGCTCCCAAG	CCCTTCTGTGCCAAACCGTA	990
<i>ACBP1</i>	TTAGCCCTGGCAGTCATCTT	AGGAAGTGTGTCCGGTGGAG	252
<i>PAR1</i>	TTTTCTCAGGCCATAATGC	GGCTACCGAGAACACCATGT	245
<i>PAR2</i>	CTCCAAATTCTCGGGTTC	GCTCAAACACTCCACCAACA	176

The annealing temperature for all primer pairs is 60°C.

corresponding *S. vulgaris* BAC clone (Additional file 6: Table S6).

All 454 sequences were *de novo* assembled using Roche GS De Novo Assembler with 98 % minimum overlap identity and 60 bp as minimum overlap length. We set the expected depth parameter with reference to the estimated size of the BAC (determined by pulse field gel electrophoresis (PFGE)) and the number of reads sequenced. In addition to 454 sequencing and assembly, we used targeted Sanger sequencing to close gaps remaining in the *S. latifolia* scaffold after the assembly process. On *S. vulgaris* we successfully used the Epicentre transposon insertion Kit (EZ1982K) to fill a 987 bp-long gap remaining after assembly. Nevertheless, a few short stretches remain unsequenced after the assembly. Both scaffolds were submitted to GenBank under accession numbers JN574439-JN574440.

The scaffolds were then annotated by similarity using BLASTX [69] with UniProtKB (Swiss-Prot + TrEMBL, 13 July 2010), the *Arabidopsis thaliana* proteome (TAIR10_20100802) and transposable elements (TAIR9_TE) [www.arabidopsis.org] with an E-value cut-off of 1E-4. We used the exon prediction tool Genscan to identify coding sequences with *A. thaliana* as the training model [70] and used ProDom [36] and a *Silene* repeated element database [37] to detect *Silene*-specific repeats and transposable elements. Transposable element annotations were then completed using annotated *S. latifolia* repeats [34].

Genetic analysis

In order to verify the pseudoautosomal location of the BAC clone studied here, we performed segregation analyses using nucleotide polymorphisms in the genes *ESPI*, *PAR2*, *ACBPI* and *ScOPA09*. First, PCR products of these genes were amplified and sequenced to look for partially sex-linked restriction polymorphisms segregating in the pseudobackcross population RB1. This population was prepared by crossing a female plant from a Swiss population, with a male plant obtained from a cross between a female of an inbred line, U9 (from Utrecht, kindly provided by Prof. Sarah Grant), and male plant from a Swiss population. Putative restriction polymorphisms (CAPS) were verified by restriction analysis, and genotyping was performed in 76 DNA samples available. For the list of primers and restriction enzymes used, see Additional file 7: Table S7. The observed recombination frequencies were used for the calculation of map distances using the onemap package of R [71] with both the Kosambi and the Haldane mapping functions [72,73].

Structure comparison

Using BLAST [69] annotation results we searched for genes and transposable elements, also using Perfect Microsatellite Repeat Finder (<http://sgdp.iop.kcl.ac.uk/>

[nikammar/repeatfinder.html](http://sgdp.iop.kcl.ac.uk/nikammar/repeatfinder.html)) with default parameters (minimum number of repeats = 3, minimum repeat unit length = 2 and maximum repeat unit length = 100). Mononucleotide repeats were identified manually. Only mononucleotide microsatellites with a minimum repeat number greater than or equal to 12 were counted. We assessed the frequency of all mono-, di-, tri- and tetranucleotide microsatellites and tested whether a smaller density of these microsatellites in *S. latifolia* than in *S. vulgaris* was associated with the observed genome size difference between species [61]. Furthermore, we estimated the frequency of long-mer microsatellite stretches (mononucleotides ≥ 30 repeats, dinucleotides ≥ 15 repeats and trinucleotides ≥ 10 repeats) in both species. We used a combination of *A. thaliana* annotation and Genscan [70] exon prediction to compare the size of intergenic regions and introns, and we measured the percentage identity of exons and introns and the GC content at third codon positions (GC3). In order to take into account the possibility that a gene was truncated due to its position at the beginning or end of a BAC, the gene fragment for which we have coverage in both species was used in calculations, whereas gaps were excluded. Both the GC and GC3 contents of the PAR genes were then compared with the GC and GC3 contents of nine X-linked genes (*DD44X*, *SIAP3X*, *SIX1*, *SlCypX*, *SIX3*, *SIX4*, *SIX7*, *SIX9* and *SlssX*) located in the non-recombining part of the *S. latifolia* sex chromosomes [49-57]. In order to test whether average GC and GC3 contents differ between species we used Student's t-tests in R [71].

Transposable element analysis

We identified LTRs of transposable elements using LTR_Finder [74] set to the default parameters. We then aligned paired LTR sequences and determined the number of mutations (substitutions and insertions/deletions) between them. We estimated the age of the LTR invasion using the following equation: $N/(2*L*K)$, where N is the number of base substitutions between the two LTRs, L is the length of the LTR and K the base substitution per site per year [75]. We used a value of $K = 23E-9$ as the average substitution rate per site per year [76].

Additional files

Additional file 1: Table S1. PAR gene expression. Description: Expression data were extracted from Muyle et al. [41].

Additional file 2: Table S2. Comparison of GC and GC3 contents in *Silene latifolia* sex-linked genes. Description: GC and GC3 contents were calculated for known sex-linked genes located in the non-recombining part of the *S. latifolia* X chromosome. CDS sequences were extracted from GenBank.

Additional file 3: Table S3. Genetic analyses of recombination events between the PAB and markers SCOPA09, *ESP1*, *PAR2* and *ACBP1*. Description: Standard errors are indicated in brackets.

Additional file 4: Table S4. Genetic analysis of the linkage between the marker SCOPA09 and each gene *ESP1*, *PAR2* and *ACBP1*.

Additional file 5: Table S5. LTR insertion time estimates Description: Positions and characteristics of LTRs (Primer binding site (PBS), polypurine tract (PPT), both 5'- and 3'-LTR size) found using LTR_Finder. Time was computed as described by Liu et al. (2008): Nucleotide differences / (2 x LTR length x K). K: substitutions per site per year = 23E-9. SI = *S. latifolia*, Sv = *S. vulgaris*.

Additional file 6: Table S6. Results of the different BAC sequencing approaches and assemblies. Description: The same *S. latifolia* BAC clone was sequenced by shotgun Sanger sequencing and 454 pyrosequencing with different paired-end libraries.

Additional file 7: Table S7. Primers and restriction enzymes used for genetic analysis. Description: The annealing temperature for all primer pairs was 60°C.

Acknowledgements

We would like to thank L. Poveda, M. Kuenzli and W. Qi from the Functional Genomic Center Zurich (FGCZ) for assistance relating to sequencing, T. Torossi, C. Michel and the Genetic Diversity Center (GDC) at ETH Zürich for technical support, S. Zoller for bioinformatics support and M. Scarborough for help with English writing. We further acknowledge support by J. Macas and E. Kejnovský who provided sequences of repeated elements and J. Bartoš who participated in BAC library analysis. This study was supported by an ETH Zurich grant (TH-07 06-3) to AW, by Czech Science Foundation grants (522/09/0083 to RH and P501/12/G090 to BJ) and Centre of the Region Haná for Biotechnological and Agricultural Research grant (ED0007/01/01) to RH.

Author details

¹Institute of Integrative Biology (IBZ), ETH Zürich, Universitätstrasse 16, Zürich, 8092, Switzerland. ²Institute of Biophysics, Laboratory of Plant Developmental Genetics, Academy of Sciences of the Czech Republic, v.v.i. Kralovopolska 135, Brno, CZ-61200, Czech Republic. ³Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Sokolovská 6, Olomouc, CZ-77200, Czech Republic.

Authors' contributions

NB performed data analysis and drafted the manuscript. HB screened the *S. latifolia* BAC library and assisted with writing the manuscript. RC screened the *S. vulgaris* BAC library and helped with transposable element analysis. JZ and BJ performed genetic analyses. BJ also assisted with writing the manuscript. RH participated in the coordination of the study and helped with data analysis, interpretation and drafting of the manuscript. AW conceived the study, coordinated and supervised all stages and helped draft the manuscript. All authors read and approved the final manuscript.

Received: 8 September 2011 Accepted: 8 June 2012

Published: 8 June 2012

References

- Wilson MA, Makova KD: **Genomic Analyses of Sex Chromosome Evolution.** *Annu Rev Genomics Hum Genet* 2009, **10**:333–354.
- Bergero R, Charlesworth D: **The evolution of restricted recombination in sex chromosomes.** *Trends Ecol Evol* 2009, **24**(2):94–102.
- Jamilena M, Mariotti B, Manzano S: **Plant sex chromosomes: molecular structure and function.** *Cytogenet Genome Res* 2008, **120**(3–4):255–264.
- Reamon-Buttner SM, Schondelmaier J, Jung C: **AFLP markers tightly linked to the sex locus in *Asparagus officinalis* L.** *Mol Breeding* 1998, **4**(2):91–98.
- Reamon-Buttner SM, Jung C: **AFLP-derived STS markers for the identification of sex in *Asparagus officinalis* L.** *Theor Appl Genet* 2000, **100**(3–4):432–438.
- Liu ZY, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu QY, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH, Ming R: **A primitive Y chromosome in papaya marks incipient sex chromosome evolution.** *Nature* 2004, **427**(6972):348–352.
- Lengerova M, Vyskot B: **Sex chromatin and nucleolar analyses in *Rumex acetosa* L.** *Protoplasma* 2001, **217**(4):147–153.
- Lengerova M, Moore RC, Grant SR, Vyskot B: **The sex chromosomes of *Silene latifolia* revisited and revised.** *Genetics* 2003, **165**(2):935–938.
- Westergaard M: **Aberrant Y chromosomes and sex expression in *Melandrium album*.** *Hereditas* 1946, **32**(3–4):419–443.
- Scotti I, Delph LF: **Selective Trade-offs and Sex-Chromosome Evolution in *Silene latifolia*.** *Evolution* 2006, **60**(9):1793–1800.
- Freije D, Helms C, Watson M, Donis-Keller H: **Identification of a second pseudoautosomal region near the Xq and Yq telomeres.** *Science* 1992, **258**(5089):1784–1787.
- Nicolas M, Marais G, Hykelova V, Janoušek B, Laporte V, Vyskot B, Mouchiroud D, Negrutiu I, Charlesworth D, Moneger F: **A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants.** *PLoS Biol* 2004, **3**(1):47–56.
- Filatov DA: **Evolutionary history of *Silene latifolia* sex chromosomes revealed by genetic mapping of four genes.** *Genetics* 2005, **170**(2):975–979.
- Veyrunes F, Waters PD, Miethke P, Rens W, McMillan D, Alsop AE, Gruzner F, Deakin JE, Whittington CM, Schatzkammer K, Kremitzki CL, Graves T, Ferguson-Smith MA, Warren W, Graves JAM: **Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes.** *Genome Res* 2008, **18**(6):965–973.
- Costich D, Meagher T, Yurkew E: **A rapid means of sex identification in *Silene latifolia* by use of flow cytometry.** *Plant Molecular Biology Reporter* 1991, **9**(4):359–370.
- Ming R, Moore PH: **Genomics of sex chromosomes.** *Curr Opin Plant Biol* 2007, **10**(2):123–130.
- Šíroky J, Lysak MA, Doležel J, Kejnovský E, Vyskot B: **Heterogeneity of rDNA distribution and genome size in *Silene* spp.** *Chromosome Res* 2001, **9**(5):387–393.
- Jeffreys AJ, May CA: **Intense and highly localized gene conversion activity in human meiotic crossover hot spots.** *Nat Genet* 2004, **36**(2):151–156.
- Chen JF, Lu F, Chen SS, Tao SH: **Significant positive correlation between the recombination rate and GC content in the human pseudoautosomal region.** *Genome* 2006, **49**(5):413–419.
- Burgoyne PS: **Genetic homology and crossing over in the X and Y chromosomes of mammals.** *Hum Genet* 1982, **61**(2):85–90.
- Mohandas TK, Speed RM, Passage MB, Yen PH, Chandley AC, Shapiro LJ: **Role of the pseudoautosomal region in sex-chromosome pairing during male meiosis: meiotic studies in a man with a deletion of distal Xp.** *Am J Hum Genet* 1992, **51**(3):526–533.
- Burgoyne PS, Mahadevaiah SK, Sutcliffe MJ, Palmer SJ: **Fertility in Mice Requires X-Y Pairing and a Y-Chromosomal Spermiogenesis Gene-Mapping to the Long Arm.** *Cell* 1992, **71**(3):391–398.
- Matsuda Y, Moens PB, Chapman VM: **Deficiency of X-Chromosomal and Y-Chromosomal Pairing at Meiotic Prophase in Spermatocytes of Sterile Interspecific Hybrids between Laboratory Mice (*Mus-Domesticus*) and *Mus-Spretus*.** *Chromosoma* 1992, **101**(8):483–492.
- Bachtrog D, Charlesworth B: **Towards a complete sequence of the human Y chromosome.** *Genome Biol* 2001, **2**(5):reviews1016.1011–reviews1016.1015.
- Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, Frankish A, Lovell FL, Howe KL, Ashurst JL, Fulton RS, Sudbrak R, Wen G, Jones MC, Hurler ME, Andrews TD, Scott CE, Searle S, Ramser J, Whittaker A, Deadman R, Carter NP, Hunt SE, Chen R, Cree A, Gunaratne P, et al: **The DNA sequence of the human X chromosome.** *Nature* 2005, **434**(7031):325–337.
- Perry J, Palmer S, Gabriel A, Ashworth A: **A Short Pseudoautosomal Region in Laboratory Mice.** *Genome Res* 2001, **11**(11):1826–1832.
- Das PJ, Chowdhary BP, Raudsepp T: **Characterization of the Bovine Pseudoautosomal Region and Comparison with Sheep, Goat, and Other Mammalian Pseudoautosomal Regions.** *Cytogenet Genome Res* 2009, **126**(1–2):139–147.
- Langford C, Fischer P, Binns M, Holmes N, Carter N: **Chromosome-specific paints from a high-resolution flow karyotype of the dog.** *Chromosome Res* 1996, **4**(2):115–123.
- Young A, Kirkness E, Breen M: **Tackling the characterization of canine chromosomal breakpoints with an integrated in-situ/in-silico approach: The canine PAR and PAB.** *Chromosome Res* 2008, **16**(8):1193–1202.
- Perry J, Ashworth A: **Evolutionary rate of a gene affected by chromosomal position.** *Curr Biol* 1999, **9**(17):987–989.

31. Montoya-Burgos JI, Boursot P, Galtier N: **Recombination explains isochores in mammalian genomes.** *Trends in genetics: TIG* 2003, **19**(3):128–130.
32. Di Stilio VS, Kesseli RV, Mulcahy DL: **A pseudoautosomal random amplified polymorphic DNA marker for the sex chromosomes of *Silene dioica*.** *Genetics* 1998, **149**(4):2057–2062.
33. Westergaard M: **The mechanism of sex determination in dioecious flowering plants.** *Adv Genet* 1958, **9**:217–281.
34. Čermák T, Kubát Z, Hobza R, Koblížková A, Widmer A, Macas J, Vyskot B, Kejnovský E: **Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes.** *Chromosome Res* 2008, **16**(7):961–976.
35. Filatov DA, Howell EC, Groutides C, Armstrong SJ: **Recent Spread of a Retrotransposon in the *Silene latifolia* Genome, Apart From the Y Chromosome.** *Genetics* 2009, **181**(2):811–817.
36. Corpet F, Servant F, Gouzy J, Kahn D: **ProDom and ProDom-CG: tools for protein domain analysis and whole genome comparisons.** *Nucleic Acids Res* 2000, **28**(1):267–269.
37. Macas J, Kejnovský E, Neumann P, Novák P, Koblížková A, Vyskot B: **Next Generation Sequencing-Based Analysis of Repetitive DNA in the Model Dioecious Plant *Silene latifolia*.** *PLoS One* 2011, **6**(11):e27335.
38. Kalendar R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH: **Analysis of plant diversity with retrotransposon-based molecular markers.** *Heredity* 2011, **106**(4):520–530.
39. Kejnovský E, Kubát Z, Macas J, Hobza R, Mracek J, Vyskot B: **Retand: a novel family of gypsy-like retrotransposons harboring an amplified tandem repeat.** *Mol Genet Genomics* 2006, **276**(3):254–263.
40. Otto SP, Pannell JR, Peichel CL, Ashman T-L, Charlesworth D, Chippindale AK, Delph LF, Guerrero RF, Scarpino SV, McAllister BF: **About PAR: The distinct evolutionary dynamics of the pseudoautosomal region.** *Trends in genetics: TIG* 2011, **27**(9):358–367.
41. Muyle A, Zemp N, Deschamps C, Mousset S, Widmer A, Marais GAB: **Rapid De Novo Evolution of X Chromosome Dosage Compensation in *Silene latifolia*, a Plant with Young Sex Chromosomes.** *PLoS Biol* 2012, **10**(4):e1001308.
42. Galtier N, Piganeau G, Mouchiroud D, Duret L: **GC-content evolution in mammalian genomes: The biased gene conversion hypothesis.** *Genetics* 2001, **159**(2):907–911.
43. Gabriel M: **Biased gene conversion: implications for genome and sex evolution.** *Trends Genet* 2003, **19**(6):330–338.
44. Duret L, Galtier N: **Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes.** *Annu Rev Genomics Hum Genet* 2009, **10**(1):285–311.
45. Filatov DA: **A gradient of silent substitution rate in the human pseudoautosomal region.** *Mol Biol Evol* 2004, **21**(2):410–417.
46. Raudsepp T, Chowdhary BP: **The horse pseudoautosomal region (PAR): characterization and comparison with the human, chimp and mouse PARs.** *Cytogenet Genome Res* 2008, **121**(2):102–109.
47. Huang SW, Friedman R, Yu N, Yu A, Li WH: **How strong is the mutagenicity of recombination in mammals?** *Mol Biol Evol* 2005, **22**(3):426–431.
48. Glémin S, Bazin E, Charlesworth D: **Impact of mating systems on patterns of sequence polymorphism in flowering plants.** *Proceedings of the Royal Society B: Biological Sciences* 2006, **273**(1604):3011–3019.
49. Moore RC, Kozyreva O, Lebel-Hardenack S, Široký J, Hobza R, Vyskot B, Grant SR: **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* 2003, **163**(1):321–334.
50. Čegan R, Marais GAB, Kubeková H, Blavet N, Widmer A, Vyskot B, Doležel J, Šafář J, Hobza R: **Structure and evolution of *Apetala3*, a sex-linked gene in *Silene latifolia*.** *BMC Plant Biol* 2010, **10**:180.
51. Delichère C, Veuskens J, Hernould M, Barbacar N, Mouras A, Negrutiu I, Monéger F: **SIY1, the first active gene cloned from a plant Y chromosome, encodes a WD-repeat protein.** *EMBO J* 1999, **18**(15):4169–4179.
52. Qiu S, Bergero R, Forrest A, Kaiser VB, Charlesworth D: **Nucleotide diversity in *Silene latifolia* autosomal and sex-linked genes.** *P Roy Soc B-Biol Sci* 2010, **277**(1698):3283–3290.
53. Bergero R, Forrest A, Kamau E, Charlesworth D: **Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: Evidence from new sex-linked genes.** *Genetics* 2007, **175**(4):1945–1954.
54. Kaiser VB, Bergero R, Charlesworth D: **A new plant sex-linked gene with high sequence diversity and possible introgression of the X copy.** *Heredity* 2011, **106**(2):339–347.
55. Atanassov I, Delichère C, Filatov DA, Charlesworth D, Negrutiu I, Monéger F: **Analysis and Evolution of Two Functional Y-Linked Loci in a Plant Sex Chromosome System.** *Mol Biol Evol* 2001, **18**(12):2162–2168.
56. Marais GAB, Nicolas M, Bergero R, Chambrier P, Kejnovský E, Monéger F, Hobza R, Widmer A, Charlesworth D: **Evidence for degeneration of the Y chromosome in the dioecious plant *Silene latifolia*.** *Curr Biol* 2008, **18**(7):545–549.
57. Filatov DA: **Substitution rates in a new *Silene latifolia* sex-linked gene, *SlssXY*.** *Mol Biol Evol* 2005, **22**(3):402–408.
58. Matsunaga S: **Sex chromosome-linked genes in plants.** *Genes Genet Syst* 2006, **81**(4):219–226.
59. Kaiser VB, Bergero R, Charlesworth D: **Slcyl, a Newly Identified Sex-Linked Gene, Has Recently Moved onto the X Chromosome in *Silene latifolia* (Caryophyllaceae).** *Mol Biol Evol* 2009, **26**(10):2343–2351.
60. Guo X, Wang Y, Keightley P, Fan L: **Patterns of selective constraints in noncoding DNA of rice.** *BMC Evol Biol* 2007, **7**(1):208.
61. Morgante M, Hanafey M, Powell W: **Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes.** *Nat Genet* 2002, **30**(2):194–200.
62. Kubát Z, Hobza R, Vyskot B, Kejnovský E: **Microsatellite accumulation on the Y chromosome in *Silene latifolia*.** *Genome* 2008, **51**(5):350–356.
63. Hobza R, Lengerova M, Svoboda J, Kubeková H, Kejnovský E, Vyskot B: **An accumulation of tandem DNA repeats on the Y chromosome in *Silene latifolia* during early stages of sex chromosome evolution.** *Chromosoma* 2006, **115**(5):376–382.
64. Hobza R, Kejnovský E, Vyskot B, Widmer A: **The role of chromosomal rearrangements in the evolution of *Silene latifolia* sex chromosomes.** *Mol Genet Genomics* 2007, **278**(6):633–638.
65. Kejnovský E, Hobza R, Čermák T, Kubát Z, Vyskot B: **The role of repetitive DNA in structure and evolution of sex chromosomes in plants.** *Heredity* 2009, **102**(6):533–541.
66. Čegan R, Vyskot B, Kejnovský E, Kubát Z, Blavet H, Šafář J, Doležel J, Blavet N, Hobza R: **Genomic Diversity in Two Related Plant Species with and without Sex Chromosomes - *Silene latifolia* and *S. vulgaris*.** *PLoS One* 2012, **7**(2):e31898.
67. Tao Q, Wang A, Zhang H-B: **One large-insert plant-transformation-competent BIBAC library and three BAC libraries of Japonica rice for genome research in rice and other grasses.** *TAG Theoretical and Applied Genetics* 2002, **105**(6):1058–1066.
68. Bouzidi MF, Franchel J, Tao QZ, Stormo K, Mráz A, Nicolas P, Mouzezar S: **A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions.** *Theor Appl Genet* 2006, **113**(1):81–89.
69. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**(17):3389–3402.
70. Burge C, Karlin S: **Prediction of complete gene structures in human genomic DNA.** *J Mol Biol* 1997, **268**(1):78–94.
71. R Development Core Team: **R: A Language and Environment for Statistical Computing.** 2011 [http://www.R-project.org].
72. Kosambi DD: **The estimation of map distances from recombination values.** *Ann Hum Genet* 1943, **12**(1):172–175.
73. Haldane JBS: **The combination of linkage values and the calculation of distance between the loci of linked factors.** *J Genet* 1919, **8**:299–309.
74. Xu Z, Wang H: **LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons.** *Nucleic Acids Res* 2007, **35**:W265–W268. Web Server issue.
75. Liu Z, Yue W, Li DY, Wang RRC, Kong XY, Lu K, Wang GX, Dong YS, Jin WW, Zhang XY: **Structure and dynamics of retrotransposons at wheat centromeres and pericentromeres.** *Chromosoma* 2008, **117**(5):445–456.
76. Wolfe KH, Li WH, Sharp PM: **Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs.** *Proc Natl Acad Sci* 1987, **84**(24):9054–9058.

doi:10.1186/1471-2164-13-226

Cite this article as: Blavet et al.: Comparative analysis of a plant pseudoautosomal region (PAR) in *Silene latifolia* with the corresponding *S. vulgaris* autosome. *BMC Genomics* 2012 **13**:226.