npg

ORIGINAL ARTICLE

Single locus sex determination and female heterogamety in the basket willow (*Salix viminalis* L.)

P Pucholt[1], A-C Rönnberg-Wästljung[1] and S Berlin

Most eukaryotes reproduce sexually and a wealth of different sex determination mechanisms have evolved in this lineage. Dioecy or separate sexes are rare among flowering plants but have repeatedly evolved from hermaphroditic ancestors possibly involving male or female sterility mutations. Willows (*Salix* spp.) and poplars (*Populus* spp.) are predominantly dioecious and are members of the Salicaceae family. All studied poplars have sex determination loci on chromosome XIX, however, the position differs among species and both male and female heterogametic system exists. In contrast to the situation in poplars, knowledge of sex determination mechanisms in willows is sparse. In the present study, we have for the first time positioned the sex determination locus on chromosome XV in *S. viminalis* using quantitative trait locus mapping. All female offspring carried a maternally inherited haplotype, suggesting a system of female heterogamety or ZW. We used a comparative mapping approach and compared the positions of the markers between the *S. viminalis* linkage map and the physical maps of *S. purpurea*, *S. suchowensis* and *P. trichocarpa*. As we found no evidence for chromosomal rearrangements between chromosome XV and XIX between *S. viminalis* and *P. trichocarpa*, it shows that the sex determination loci in the willow and the poplar most likely do not share a common origin and has thus evolved separately. This demonstrates that sex determination mechanisms in the Salicaceae family have a high turnover rate and as such it is excellent for studies of evolutionary processes involved in sex chromosome turnover.
Heredity (2015) **114**, 575–583; doi:10.1038/hdy.2014.125; published online 4 February 2015

INTRODUCTION

Sexual reproduction is the most common mode of reproduction among eukaryotes and results in mixing and reshuffling of alleles of the maternal and paternal genomes, yielding novel allelic combinations in future generations. In most animals, sexual reproduction has led to the evolution of separate sexes with male and female sexual organs on different individuals. In contrast, in flowering plants, only about 5% are dioecious and have male and female flowers on separate individuals (Charlesworth, 2002; Renner, 2014). Instead, the great majority, or about 90% are hermaphrodites with male and female sexual organs on the same flower, while another 5% are monoecious (Charlesworth, 2002), having separate male and female flowers on one individual. The fact that dioecy is so rare and that it has a scattered taxonomic distribution suggests that it is a derived trait and that hermaphroditism is the ancestral angiosperm state (Charlesworth, 2002). When dioecy evolves from hermaphroditism, two new mutations are required, that is, one male and one female sterilizing (Charlesworth *et al.*, 2005). Selection will favour suppression of recombination between the two mutations and with time the non-recombining region can expand as sexually antagonistic mutations accumulate near the sex determination locus (=SD locus), benefitting from linkage disequilibrium. A side effect of suppressed recombination is that selection will be inefficient, so that deleterious mutations can accumulate, which with time can result in structural changes in these chromosomal regions, which has happened with mammalian male-specific Y chromosomes (Charlesworth and Charlesworth, 2000; Bachtrog, 2008) and avian female-specific W chromosomes (Berlin and Ellegren, 2004; Hillier *et al.*, 2004; Moghadam *et al.*, 2012). Among many plants (Charlesworth and Mank, 2010) and lower vertebrates such as fish, amphibians and reptiles, sex chromosomes often display little differentiation, which is assumed to be the result of high turnover rates of the sex chromosomes (Perrin, 2009). This can happen if new sex determination mutations appear, replacing the old, before degeneration of the old sex chromosomes had time to happen (Perrin, 2009). High turnover rates can also result in transitions between male and female heterogametic sexual systems. Frequent sex chromosome turnover has been documented in several fish species (Mank and Avise, 2009) and different populations of an amphibian species even display alternate heterogametic systems (Miura, 2007). In plants, sex chromosome turnover has not been frequently encountered, although in the *Silene* plant genus, transition from male to female heterogamety did take place in the subsection Otites (Slancarova *et al.*, 2013). Also in the *Populus* genus, transitions between male and female heterogamety seem to have happened, as is described below.

The *Populus* and the *Salix* genera belong to the Salicaceae plant family where most species are dioecious, thus suggesting that dioecy is old and evolved before the two genera diverged some 45 Mya (Boucher *et al.*, 2003; Manchester *et al.*, 2006). Despite the fact that dioecy is old, heteromorphic sex chromosomes have not been found

Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala, Sweden
[1]These authors contributed equally to this work.
Correspondence: Dr S Berlin, Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, SLU, Box 7080, 750 07 Uppsala, Sweden.
E-mail: sofia.berlin@slu.se

in any poplar or willow species studied so far. Instead, poplars have their SD loci on chromosome XIX, with a peritelomeric localization in *P. nigra* (Gaudet *et al.*, 2008) and *P. trichocarpa* (Yin *et al.*, 2008) and a pericentromeric localization in *P. tremula*, *P. tremuloides* (Pakull *et al.*, 2009; Kersten *et al.*, 2012) and *P. alba* (Paolucci *et al.*, 2010). In *P. trichocarpa*, a 1-Mbp long, female-specific fragment with suppressed recombination is located in the sex determining region, supporting female heterogamety (Yin *et al.*, 2008). In *P. alba*, sex was mapped on the female map but not on the male map, also supporting female heterogamety (Paolucci *et al.*, 2010). In *P. tremuloides*, instead a 2-Mbp long male-specific region was found, supporting male heterogamety (Kersten *et al.*, 2014). In *P. nigra*, sex was mapped to the male map and not to the female map, suggesting male heterogamety also in this species (Gaudet *et al.*, 2008). Based on these observations it seems likely that different sex determination mechanisms exist in the *Populus* genus (Gaudet *et al.*, 2008; Paolucci *et al.*, 2010; Tuskan *et al.*, 2012) with both male and female heterogamety present among closely related species, observations that support high rates of sex chromosome turnover in this plant genus.

Sex determination in willows has not been as extensively studied as in poplars, however, by investigating biased sex ratios in *Salix viminalis* (Subgenus: *Vetrix*), a polygenic model for sex determination was early on proposed (Alström-Rapaport *et al.*, 1997) and the discovery of a RAPD marker (UBC354$_{560}$) predominantly linked to females led to the hypothesis that females may possess some chromosomal fragment that is not present in males (Alström-Rapaport *et al.*, 1998). Furthermore, Semerikov *et al.* (2003) mapped a SD locus to one linkage group, however, which chromosome that linkage group belong to could not be determined at that time. Two SCAR markers (SCAR 354 and SCAR AE08) were developed from the original RAPD marker that were sequenced and analysed for sex linkage in a wider material of *S. viminalis*, where female linkage was confirmed for both markers (Gunter *et al.*, 2003a). In *S. eriocephala*, SCAR AE08 was also significantly associated with females, whereas SCAR 354 was only marginally associated, however, the authors concluded that different willow species may share the same sex determination mechanism (Gunter *et al.*, 2003b). In *S. caprea* on the other hand, no complete sex linkage was demonstrated albeit using a large number of markers, which led to the conclusion that different willow species do not share the same sex determination mechanism (Semerikov *et al.*, 2003). Later on, SCAR 354 was sequenced in 12 different willow species and successful amplification was demonstrated only in species in the *Vetrix* and the *Chamaetia* subgenera, whereas those species in which amplification failed was in the *Salix* subgenus, possibly indicating that the locus is not present in this subgenus (Temmel *et al.*, 2007). Interestingly, the SCAR 354 sequence mapped to chromosome XV in the *P. trichocarpa* genome (Temmel *et al.*, 2007).

Given the dynamic nature of sex determination among willows and poplars, they together represent an attractive system for studies of processes involved in sex chromosome turnover. In the present study, the first objective was to identify and position loci associated with sex in *S. viminalis* by performing quantitative trait locus (QTL) mapping and to determine if the sexual system is male or female heterogametic. Sex was mapped in one mapping population, and then the interval for the major QTL was finemapped by increasing the number of genotyped individuals as well as the number of markers in the interval. The *S. purpurea* genome assembly was used to analyse the genomic content of the SD locus. The second objective was to investigate if *S. viminalis* and *P. trichocarpa* have homologous SD loci, possibly possessing the same sex determination mechanism.

Chromosome XIX was therefore finemapped and marker positions between *S. viminalis* and *P. trichocarpa* were compared to find evidence for translocations of chromosomal segments.

MATERIALS AND METHODS

Plant material and phenotyping
The S_3 mapping population (Berlin *et al.*, 2010) is composed of full siblings from a cross between the male *S. viminalis* clone 81084 and the female *S. viminalis* clone 78195. This cross has previously been demonstrated to produce individuals with variation in their resistance to the gall midge *Dasineura marginemtorquens* (Bremi) (Höglund *et al.*, 2005). Crosses were made in 2000 and 2004 and the plants were reared in a greenhouse before planting in the field. The plants are maintained in a plant orchard in Pustnäs, south of Uppsala, Sweden (59.80°N, 17.67°E, 25m AOD). Gender determination was done twice in 2010 and in 2014.

QTL mapping
QTL mapping was performed using phenotypic data from 282 individuals and a genome-wide linkage map presented in Berlin *et al.* (2010) where the S_3 linkage map for chromosome XV consists of two linkage groups (XV-1 and XV-2; here called the old map). The analysis was made with the programme MapQTL 6.0 (Van Ooijen, 2009) using the non-parametric Kruskal–Wallis method and interval mapping with a regression model. The genome was scanned at 1 cM intervals to determine putative QTLs associated with sex. A significance threshold value, estimated as the logarithm of odds ratio (LOD), to determine significant QTLs was estimated with a permutation test of 1000 repetitions. A genome-wide threshold for a significant QTL was set at $P = 0.05$. The two LOD confidence interval for each QTL were estimated using the LOD drop-off method (Lander and Botstein, 1989) based on the LOD-value at the peak position of the QTL. The markers closest to the 2-LOD region of the QTL were used to compare the lengths in cM of the QTL regions between the new and old linkage map. The proportion of the phenotypic variation explained by each significant QTL was estimated (% Expl.). Maternal allelic effects were estimated as $A_m = ((\mu_{ac} + \mu_{ad}) - (\mu_{bc} + \mu_{bd}))/2$, paternal allelic effects were estimated as $A_p = ((\mu_{ac} + \mu_{bc}) - (\mu_{ad} + \mu_{bd}))/2$ and the allelic interaction effect were estimated as $A_i = ((\mu_{ac} + \mu_{bd}) - (\mu_{ad} + \mu_{bc}))/2$ where μ_{ac}, μ_{bc}, μ_{ad} and μ_{bd} are the estimated phenotypic means of the four genotypic classes ac, bc, ad and bd obtained from an ab × cd cross (Van Ooijen, 2009).

Finemapping
With the aim of achieving a more precise localization of the sex determination QTL, new single-nucleotide polymorphism (SNP) markers with putative positions on the *S. viminalis* chromosome XV were genotyped in more individuals. Two hundred primer pairs were placed in exons spanning introns of approximately 500–1000 base pairs on chromosome XV in the *P. trichocarpa* genome v. 1.1 (Tuskan *et al.*, 2006). Genomic DNA was extracted from frozen young leaves with the FastDNA Kit (MP Biomedicals, Stockholm, Sweden). The gene fragments were amplified by PCR with DNA from each of the male and the female parent in a total volume of 15 µl. Each reaction contained 10 ng genomic DNA, 1 × PCR buffer II (Applied Biosystems, Stockholm, Sweden), 2.5 mM MgCl$_2$ (Applied Biosystems), 0.2 mM dNTP mix (Fermentas, Stockholm, Sweden), 0.5 µM of each primer (Invitrogen, Stockholm, Sweden) and 0.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems). The PCRs were run on a MyCycler thermal cycler (Bio-Rad Laboratories, Stockholm, Sweden) with a PCR profile consisting of 10 min denaturation at 95 °C followed by 35 cycles of 30 s denaturation at 95 °C, 30 s annealing at 58 °C and 1 min extension at 72 °C with a final 10 min 72 °C step. Amplification success was determined by agarose gel electrophoresis. The PCR products were cleaned with 1 µl of a mixture of Exonuclease I (New England BioLabs, Stockholm, Sweden) and Shrimp Alkaline Phosphatase (Fermentas) for every 5 µl of PCR product and sequenced at Macrogen Inc. (Seoul, South Korea) on both strands using the forward and reverse PCR primers as sequencing primers. The chromatograms were visually inspected and edited using Lasergene SeqMan vs 9.1.0 (DNASTAR) and SNPs were identified. The new markers were named XV_*fx*, where *x* is an arbitrary number and *f* refer to 'finemapping'. In addition to these markers, two previously developed markers were genotyped: XV-5_sa and DT_6_sa

(Berlin *et al.*, 2010). SNPs were genotyped using the Illumina GoldenGate Assay at the SNP&SEQ Technology Platform, Science for Life Laboratory at Uppsala University, Sweden, in 531 individuals, including the parental genotypes and the previously used 282 individuals (thus 247 new individuals were genotyped).

Markers were also developed evenly spaced on chromosome XIX in the *P. trichocarpa* genome. Primer pairs were designed and sequenced in both the male and the female parent in a similar manner as for the primers on chromosome XV and SNPs were genotyped together with the markers on chromosome XV. The new markers were named XIX_*fx*, adopting a similar nomenclature as for the markers on chromosome XV. In addition to the newly developed markers, eight previously developed markers were genotyped (Supplementary Table 1).

A new linkage map was constructed combining genotypes obtained with the old and the new markers using JoinMap 4.0 software (Van Ooijen, 2006) with a LOD of 4.0 for marker grouping. In addition to the genotypes, sex as a morphological marker was included in the linkage analyses. The mapping population was assigned as a cross pollinating population where markers could be polymorphic in one or both of the parents. The Kosambi mapping function and the Joinmap default settings was used for map construction. In case of more than one round of map building, we used the second round without allowing the programme to force any problematic markers into the map. Mendelian segregation of the different markers was tested with χ^2 analysis. Maps were drawn using MapChart vs 2.2 (Voorrips, 2002). QTL mapping was conducted as described above using the new linkage map and phenotype data on sex from all 529 individuals in the mapping population.

Inferring the haplotypes in the sex determination QTL

In order to further narrow down the sex determination QTL, each individual's two haplotypes in the major QTL interval were inferred using the statistical software package PHASE v. 2.1 (Stephens *et al.*, 2001; Stephens and Scheet, 2005). The analysis was restricted to the 10 markers surrounding the peak position of the QTL, that is, XV_f2, XV-5_sa, R_33_sa_pIII, R_51_sa, XV_f6, XV_f7, XV_f11, XV_f5, XV_f12 and XV_f21. PHASE adopts an iterative scheme to perform the inferences. Default values were used, which means 100 iterations, a thinning interval of 1 and a burn-in of 100. In order to obtain robust results the algorithm was applied five times with different seeds for the random number generator and consistency among runs was evaluated by examination of the haplotype frequency estimates. The inferred haplotypes were subsequently associated with the phenotypic data on sex.

Comparative mapping and estimates of recombination rates

SNP marker positions were determined in the *Salix purpurea* (Subgenus: Vetrix) genome assembly v 1.0, (DOE-JGI, http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Spurpurea), in the *Salix suchowensis* genome assembly (Dai *et al.*, 2014) and in the *P. trichocarpa* genome assembly v. 3 (http://www.phytozome.net/poplar.php; Tuskan *et al.*, 2006) using BLAST searches with the consensus DNA sequences from the male and the female parent (the marker sequences) containing the SNP markers on the *S. viminalis* chromosome XV. Similarly, marker positions were determined in the *P. trichocarpa* genome assembly for SNPs on the *S. viminalis* chromosome XIX. The positional information from the best BLAST hits was used to draw physical maps of chromosome XV in *P. trichocarpa* and *S. purpurea* and

scaffolds of *S. suchowensis* that are part of chromosome XV and of chromosome XIX in *P. trichocarpa*. Also the positions of previously developed markers SCAR 354 (Genbank accession no. AY192565) and SCAR AE08 (Genbank accession no. AY192566) were examined by BLAST searches in the different genomes. In some instances, the best BLAST hit of the marker sequence was not on the homologous chromosome of *P. trichocarpa* and *S. purpurea*. In these cases, the position of the second best BLAST hit was used for the map construction.

The *S. viminalis* linkage map was aligned to the physical maps and the positions of homologous markers in the different species were shown by lines. The *S. viminalis* SD locus, that is, the interval with the markers XV-5_sa, R_33_sa_pIII and R_51_sa was analysed using the *S. purpurea* genome. Gene models located within the genomic interval were retrieved from the genome browser at http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Spurpurea.

Recombination rates in base pairs per centimorgan were estimated between pairs of markers (a marker and its closest neighbour) across the *S. viminalis* genome. Only pairs of markers with a minimum of 0.1 cM distance in the genetic map and 10 base pair distance in the physical map were used in the analysis. Physical positions of 309 SNPs distributed over the whole genome (Berlin *et al.*, 2010) were obtained by BLAST searches against the *S. purpurea* genome as previously described. Genetic distances were taken from the linkage maps. The results were visualized in a histogram.

RESULTS

QTL mapping with old linkage map

Sex could be determined for 272 individuals out of the 282 that were genotyped, of which 137 were females and 135 were males. The analysis identified one highly significant QTL on chromosome XV associated with sex that explained 84.7% of the variation (Table 1). A large difference in the effect of female alleles was identified while this was not seen for the male alleles (Table 1). In the Kruskal–Wallis test, all markers on the two linkage groups on chromosome XV, except for the marker XV-5_sa, which was polymorphic only in the male, were significantly associated with sex. The markers R_33_sa_pIII and R_51_sa had particularly high significances ($K > 200$; Table 2).

Finemapping and QTL mapping with new linkage map

Two hundred primer pairs were positioned along the *P. trichocarpa* chromosome XV and were tested for amplification in both the female and the male parent. In all, 82 primer pairs resulted in successful amplification (41.0%) and were directly sequenced. After inspection of the chromatograms, 40 SNPs were identified and selected for genotyping. Of the 39 SNPs that resulted in informative genotypes, 35 mapped to chromosome XV, 1 to chromosome VI and 3 to chromosome XII (Supplementary Table 1). When performing BLAST searches with the marker sequences containing the SNP markers that mapped to VI and XII to the *P. trichocarpa* genome, the best BLAST hits were also on VI and XII (Supplementary Table 1), suggesting that the intended gene fragments had not been amplified. The gene fragments that mapped to chromosome XII are likely the paralogous

Table 1 Results from the QTL analysis using the old and new maps

Chr.	N	Marker[a]	Pos. (cM)	LOD	LOD thresh.[b]	Var. expl. (%)[c]	Female alleles	Male alleles	Interaction	QTL interval (cM)[d]	QTL-marker interval (cM)[e]
XV (old)	282	R_51_sa	16.7	111.1	4.1	84.7	−0.988	−0.007	−0.016	6.2	45.6
IX_2 (new)	529	IX_16_sa	44.4	4.6	4.1	4.0	−0.091	−0.257	0.138	20.8	23.8
XV (new)	529	sex	39.3	439.2	4.1	97.9	1.000	0.004	0.004	0.2	1.8

Abbreviations: Chr., chromosome; Pos., position; thresh., threshold.
[a]Marker closest to the peak position of the QTL.
[b]Genome-wide LOD threshold at 0.05 level.
[c]Phenotypic variation explained by the QTL.
[d]Two-LOD QTL interval.
[e]Region between markers closest to the position on each side of the LOD peak position representing a drop of two LODs.

Table 2 Results from the Kruskal–Wallis analysis showing all markers on the old map of chromosome XV (LG 1 and 2)

Position, cM (LG)	Marker	K	P value	N	Mean of genotype			Mean of genotype		Mean of genotype	
					hh	*hk*	*kk*	*ll*	*lm*	*nn*	*np*
0.000 (1)	XV-5_sa	1.679	NS	261						1.460	1.540
4.894 (1)	WPMS5	18.862	1×10^{-7}	75	1.210	1.454	1.869				
0.000 (2)	DT_6_sa	122.301	1×10^{-7}	261				1.160	1.846		
14.744 (2)	R_33_sa_pIII	229.151	1×10^{-7}	261				1.023	1.962		
14.744 (2)	R_51_sa	229.151	1×10^{-7}	261				1.962	1.023		
45.567 (2)	R_59_sa	27.157	1×10^{-7}	258	1.208	1.510	1.701				
53.582 (2)	R_32_sa_pl	13.383	1×10^{-6}	261				1.604	1.376		

Abbreviations: K, the Kruskal–Wallis test statistic based on ranks; LG, linkage groups; NS, not significant.
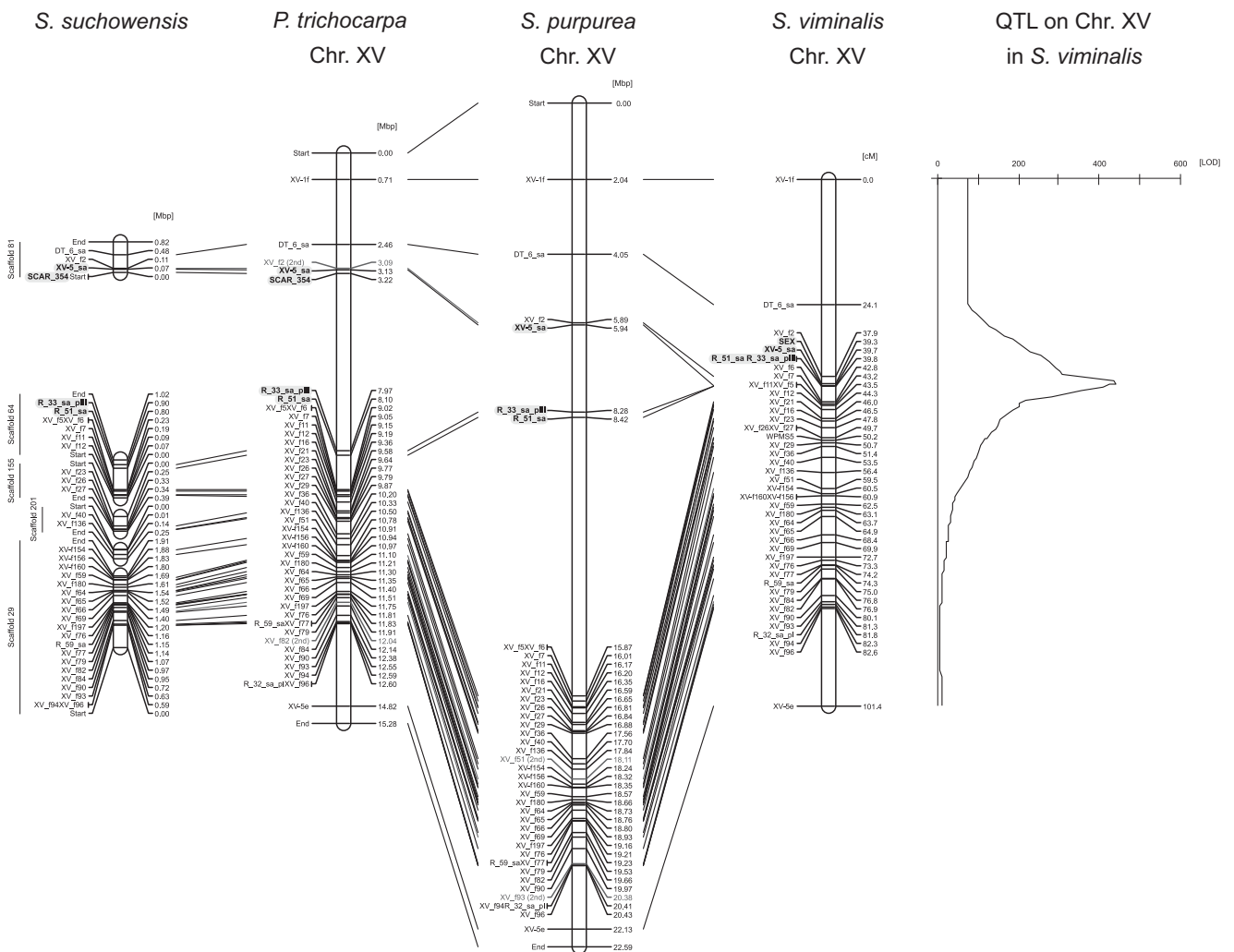
Figure 1 Comparison of maps of chromosome XV in *S. viminalis, S. purpurea, S. suchowensis* and *P. trichocarpa* and the LOD curve showing the position of the major QTL. The physical maps of *S. purpurea, S. suchowensis* and *P. trichocarpa* show the positions of the markers that were determined by BLAST searches with the *S. viminalis* marker sequences to the respective genomes. For *S. suchowensis*, scaffolds containing at least two markers were aligned according to the marker order in *S. viminalis* the distance between scaffolds is unknown. The *S. viminalis* map represents the new linkage map of chromosome XV after integration of the finemapping markers. Homologous markers between the maps were connected by lines. For some markers, the best BLAST hits in *P. trichocarpa* or *S. purpurea* were not on chromosome XV. In these cases, the second best hits were used, which is labelled also in the figure. Markers in the sex determination locus (XV-5_sa, SCAR_354, R_33_sa_pIII and R_51_sa) are highlighted. The marker SCAR_354 was not genotyped in the *S. viminalis* mapping population and is therefore not present on the genetic map.
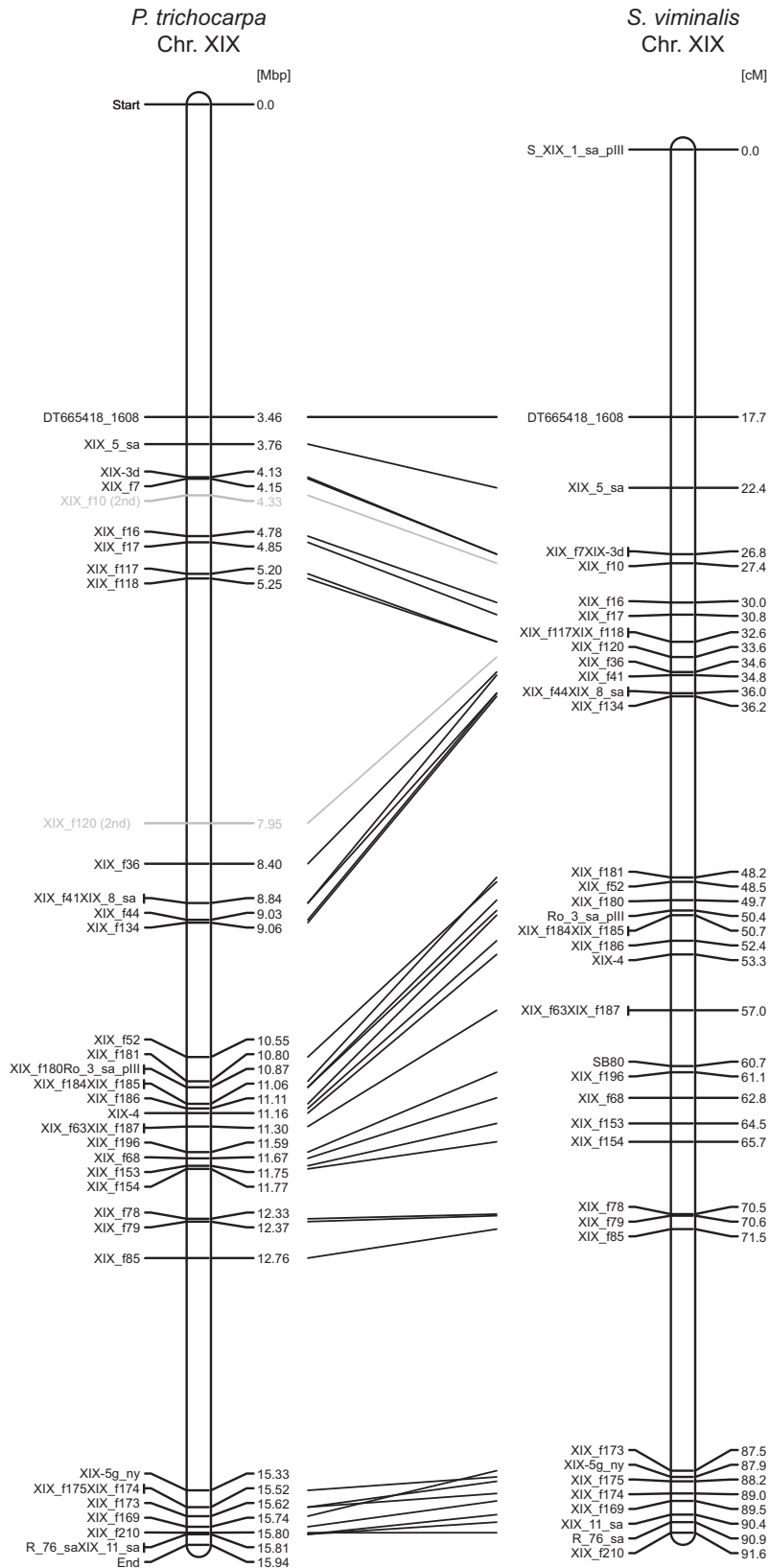
Figure 2 *S. viminalis* and *P. trichocarpa* chromosome XIX. The *P. trichocarpa* map shows the positions of the markers that were determined by BLAST searches of the *S. viminalis* marker sequences to the *P. trichocarpa* genome v. 3. The *S. viminalis* map represents the new linkage map of chromosome XIX. Homologous markers between the *P. trichocarpa* and the *S. viminalis* maps were connected by lines. For some markers, the best BLAST hits in the *P. trichocarpa* genome were not on chromosome XIX. In these cases, the second best hits were used, which is labelled also in the figure.

gene fragments to those on chromosome XV, as large parts of chromosome XII and XV contain paralogous regions that originate from the whole-genome duplication (that is, the Salicoid event) that took place before the divergence of poplars and willows (Tuskan *et al.*, 2006). Fourteen of the new markers that mapped were polymorphic in both parental genotypes (2:1:1 segregation), nine in the maternal parent and twelve in the paternal parent only (3:1 segregation; Supplementary Table 1). The morphological marker 'sex' mapped close to XV-5_sa on chromosome XV (Figure 1).

A total of 210 primer pairs were developed on *P. trichocarpa* chromosome XIX, of which 96 resulted in successful amplification (45.7%) that were directly sequenced. Forty-two SNPs were identified and genotyped. Of the 36 SNPs that resulted in informative genotypes, 31 mapped to chromosome XIX, 1 to chromosome VII, 2 to chromosome XIII, 1 to chromosome XVIII and 1 did not map. BLAST searches with the marker sequences containing the SNP markers that mapped to chromosome XIII and XVIII to the *P. trichocarpa* genome, showed that the best BLAST hits were also on XIII and XVIII (Supplementary Table 1), suggesting that the intended gene fragments were not amplified. In the case of the marker sequences on chromosome XIII, the paralogous gene fragments could have been amplified as XIII and XIX contain paralogous regions that originate from the Salicoid event (Tuskan *et al.*, 2006). However, the best BLAST hit of the marker sequence containing the marker that mapped to chromosome VII was on chromosome XIII, suggesting that again the paralogue had been amplified. However, the fact that it mapped to chromosome VII in *S. viminalis* could indicate that some rearrangement has happened between chromosome VII and XIII since the split between *S. viminalis* and *P. trichocarpa*. Eight of the new markers that mapped were polymorphic in both parental genotypes, thirteen in the maternal parent and fourteen in the paternal parent only (Supplementary Table 1).

Sex could be determined for 523 individuals out of the 529 that were genotyped. Of these, 271 carried catkins with female flowers and were thus scored as females and 252 carried catkins with male flowers and were scored males. This resulted in a slightly female biased sex ratio (1:0.93).

QTL mapping revealed, as in the previous analysis, a highly significant QTL for sex on chromosome XV, which explained as much as 97.9% of the variation (Table 1). The peak position was at the morphological marker 'sex' very close to the marker XV-5_sa and the tightly linked markers R_33_sa_pIII and R_51_sa. The difference in

the effects of the female alleles was 1, which means that one of the alleles was completely associated with sex as the male sex was scored as 1 and the female sex as 2. The 2-LOD region connected to the closest markers was significantly reduced compared with the analysis with the old map, because of the much tighter linkage map (Table 1). The non-parametric Kruskal–Wallis analysis also supported a strong signal in this region as nine markers had exceptionally high K values ($K > 200$; Table 3). The markers XV-5_sa and XV_f5 were, however, non-significant as they were only polymorphic in the males (Table 3). In addition to the QTL on chromosome XV, there was also a significant QTL on chromosome IX that explained 4% of the variation of the sex phenotype. For this QTL, the difference between the male alleles was higher compared with between female alleles.

Haplotype inference and analyses of phenotypic effects

By applying a statistical threshold of $P < 0.1$, haplotypes were inferred in 502 individuals (259 females and 243 males) for 10 markers in the QTL on chromosome XV (XV_f2, XV-5_sa, R_33_sa_pIII, R_51_sa, XV_f6, XV_f7, XV_f11, XV_f5, XV_f12 and XV_f21). Four common haplotypes were found, ranging between 23 and 25% in frequency, representing the non-recombinant parental haplotypes. Recombination was marginally more common in the maternal parent compared with the paternal parent as 32 or 6.4% of the haplotypes with a maternal origin and 26 or 5.2% of the haplotypes with a paternal origin were recombinant haplotypes. However, a larger number of different recombinant haplotypes were formed in the maternal parent compared with the paternal parent as among the recombinant haplotypes with a maternal origin there were 10 different haplotypes, and among the recombinant haplotypes with a paternal origin there were only 5 different haplotypes (Table 4). When associating these haplotypes to sex, an interesting pattern emerged. The haplotypes with a paternal origin were to the same degree found in male and female offspring (Table 4). On the other hand, the haplotypes with a maternal origin were found either in male or female offspring (Table 4). The Maternal-1 parental haplotype was female specific and as a consequence the Maternal-2 parental haplotype was male specific (Table 4). A similar pattern was also found for the recombinant haplotypes, where Maternal-3 to Maternal-7 were female specific and Maternal-8 to Maternal-12 were male specific (Table 4). This agrees with results from the QTL mapping that showed large differences between maternal alleles and a complete association of one allele with sex (Table 1). Common to all female-specific haplotypes were the

Table 3 Results from the Kruskal–Wallis analysis showing the most significant markers on the new map of chromosome XV

Position, cM	Marker	K	P value	N	Mean of genotype			Mean of genotype		Mean of genotype	
					hh	hk	kk	ll	lm	nn	np
24.1	DT_6_sa	249.6	1×10^{-7}	515				1.159	1.856		
37.8	XV_f2	483.5	1×10^{-7}	512				1.008	1.981		
39.3	Sex	511.0	1×10^{-7}	516				1.004	2.000		
39.7	XV-5_sa	0.168	NS	516						1.524	1.506
39.8	R_51_sa	262.0	1×10^{-7}	267				2.000	1.008		
39.8	R_33_sa_pIII	262.0	1×10^{-7}	267				1.008	2.000		
42.8	XV_f6	221.0	1×10^{-7}	515	1.983	1.541	1.058				
43.2	XV_f7	215.0	1×10^{-7}	514	1.059	1.539	1.975				
43.5	XV_f11	431.6	1×10^{-7}	513				1.966	1.048		
43.5	XV_f5	0.006	NS	515						1.516	1.513
44.3	XV_f12	204.2	1×10^{-7}	515	1.967	1.537	1.074				

Abbreviations: K, the Kruskal–Wallis test statistic based on ranks; NS, not significant.

Table 4 Haplotype inferences in the major QTL and associations between haplotypes and sex

Origin	Haplotype	Number	Frequency of females	Frequency of males
Paternal-1, parental	AGGACAACGC	250	0.53	0.47
Paternal-2, parental	AAGAAGAGAC	226	0.50	0.50
Paternal-3, recombinant	AGGACAACAC	4	0.25	0.75
Paternal-4, recombinant	AAGAAGAGGC	2	0.50	0.50
Paternal-5, recombinant	AGGAAGAGAC	4	0.75	0.25
Paternal-6, recombinant	AAGACAACGC	15	0.53	0.47
Paternal-7, recombinant	AGGAAGAGGC	1	1	0
Maternal-1, parental	GGAAAGACAA	242	1	0
Maternal-2, parental	AGGGCATCGC	228	0	1
Maternal-3, recombinant	GGAACATCGC	9	1	0
Maternal-4, recombinant	GGAAAATCGC	1	1	0
Maternal-5, recombinant	GGAAAGACGC	3	1	0
Maternal-6, recombinant	GGAAAGACAC	2	1	0
Maternal-7, recombinant	AGAAAGACAA	2	1	0
Maternal-8, recombinant	AGGGCATCGA	3	0	1
Maternal-9, recombinant	AGGGCATCAA	1	0	1
Maternal-10, recombinant	GGGGCATCGC	4	0	1
Maternal-11, recombinant	AGGGCGACAA	2	0	1
Maternal-12, recombinant	AGGGAGACAA	5	0	1

Sex was recorded in 523 plants where 271 carried catkins with female flowers and 252 carried catkins with male flowers. By applying a statistical threshold of $P<0.10$, haplotypes were inferred in 502 individuals (259 females and 243 males). The female-specific haplotype at the SD locus was GAA at the markers XV-5_sa, R_33_sa_pIII and R_51_sa (underlined).
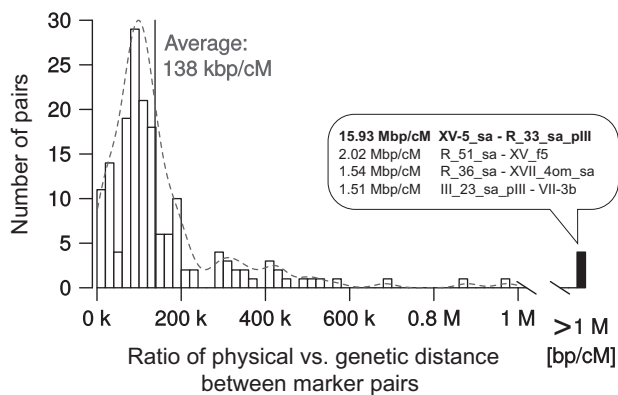
Figure 3 Distribution of pairwise recombination rates in base pairs per cM between pairs of markers. Physical distances were obtained from the *S. purpurea* v1.0 genome and genetic distances from the new linkage map of *S. viminalis*. The lowest recombination rate was found in the SD locus, that is, between markers XV-5_sa and R_33_sa_PIII with 15.93×10^6 base pairs per cM.

combination of G, A and A at the XV-5_sa, R_33_sa_pIII and R_51_sa markers. These three markers are referred to as the SD locus (Table 4).

Comparative mapping

When comparing the *S. viminalis* linkage map to the physical maps of *S. purpurea*, *S. suchowensis* and *P. trichocarpa* for chromosome XV, the synteny is overall shared, and no major rearrangement were detected (Figure 1). Individual marker sequences had a marginally better hit outside chromosome XV in *S. purpurea* and *P. trichocarpa*, that is, for *P. trichocarpa* the markers XV_f2 and XV_f82 had the best BLAST hit on chromosome XII, which is known to contain paralogous regions to chromosome XV. In *S. purpurea*, the markers XV_f51 and XV_f93 mapped best to two different scaffolds (scaffold 4340 and 0911

respectively). In all cases, the second best BLAST hit was of nearly identical quality and the position of that second hit was in syntheny. The three markers in the *S. viminalis* SD locus (XV-5_sa, R_33_sa_pIII and R_51) were present in all species, however, in *S. suchowensis*, XV-5_sa was on a different scaffold than the other two markers. SCAR 354 was located close to XV-5_sa on chromosome XV in the genomes of *S. suchowensis* and *P. trichocarpa* (Figure 1), but in the genome of *S. purpurea* it was located on scaffold 2584. SCAR AE08 did not map to chromosome XV in any of the species, but poor matches were found on chromosome V in *P. trichocarpa*, chromosome XVII in *S. purpurea* and to scaffold 238 in *S. suchowensis*. When comparing the *S. viminalis* linkage map to the physical maps of *P. trichocarpa* for chromosome XIX, again the overall synteny was shared and no chromosome XV marker mapped to XIX or vice versa (Figure 2). Marker XIX_f10, however, had a slightly better hit on scaffold 75, whereas marker XIX_f120 mapped marginally better to scaffold 53 (Figure 2). The SD locus, that is, between the markers XV-5_sa and R_51_sa represents a genomic region of about 2.5 Mbp in the *S. purpurea* genome, which harbours 48 protein-coding genes (Supplementary Table 2). When estimating the recombination rates between pairs of markers, the mean ratio for 171 analysed pairs was 138 kbp per cM (Figure 3). The most extreme ratio in the whole genome was found between the markers XV-5_sa and R_51_sa where the ratio was approximately 16 Mbp per cM (Figure 3), which is 115 times lower compared with the genome average.

DISCUSSION

In this study, we have for the first time identified and positioned one single SD locus on chromosome XV in the basket willow, *S. viminalis*. As the SCAR 354 marker that previously was demonstrated to be sex linked in other *S. viminalis* accessions (Alström-Rapaport *et al.*, 1998; Gunter *et al.*, 2003a) appear to be positioned near the XV-5_sa marker in the SD locus (based on BLAST searches in the genomes of *S. suchowensis* and *P. trichocarpa*), our and previous results together supports a general sex determination mechanism in *S. viminalis*

involving the SD locus in a so far unknown way. As the morphological trait 'sex' mapped near the XV-5_sa marker, this further strengthen the close association between the SD locus and sex as a morphological trait in this species. We further demonstrated that all female offspring had one maternally inherited haplotype, supporting a female heterogametic system. Sex determination could then either be under the control of a dominant W-like haplotype or alternatively, Z dosage determines sex. These two mechanisms are fundamentally different as in the first case the dominant W-like haplotype is assumed to harbour a master gene controlling female development, which is equivalent to the sex determination system in mammals (the SRY master male determining gene located on the Y chromosome). In the other case, sex would be determined by the dosage of one or more Z-linked genes and a higher dose would thus be required for male development, which would resemble the sex determination mechanism found in birds (Nanda *et al.*, 2008; Smith *et al.*, 2009). Further work is, however, needed in order to distinguish between these two mechanisms.

With the phase analysis, we reduced the SD locus to three markers that corresponded to a 2.5-Mbp genomic region in the *S. purpurea* genome, a region that harbours 48 protein-coding genes. This value should, however, be seen as an approximation and not as a definitive size of the region as *S. viminalis* and *S. purpurea* might have undergone lineage-specific changes since they diverged. In addition, if the sex determination regions are the same in *S. purpurea* and *S. viminalis*, the Z- and W-like genomic regions might have started to differentiate. As the *S. pupurea* genome assembly is based on sequencing data from a female individual the assembly process for this region might be a major challenge. There are reasons to believe that there are problems with the assembly as the SCAR 354 marker mapped to a scaffold (quite likely that it should be in the SD locus near XV-5_sa also in *S. purpurea*), and also that *S. purpurea* chromosome XV is much longer than chromosome XV in poplar. This can obviously be a problem when using the *S. purpurea* genome to get physical distances between the markers. If the region is misassembled and by that too long, the measure base pairs per cM will be overestimated in this region compared with the rest of the genome. As we found that the SD locus had an extremely low recombination rate compared with the rest of the genome this could then either be a consequence of misassembly or reflect suppression of recombination in the SD locus, or both. However, low recombination was also found when *P. trichocarpa* was used to estimate the physical distances between the markers (data not shown) indicating that the effect is not caused by potential problems in the *S. purpurea* genome assembly alone. In a female heterogametic system, recombination should be supressed only in females but not in males. However, in the present study, it was not possible to make this comparison, as only one of the three markers in the SD locus was polymorphic in the male parent (that is, XV-5_sa). Suppressed recombination in the SD locus could be a consequence of the position of the locus near the centromere, where recombination rates are known to be reduced. In other species, SD loci are located in pericentromeric regions, for example, in *P. tremuloides* (Kersten *et al.*, 2014), *P. alba* (Paolucci *et al.*, 2010), papaya (Yu *et al.*, 2007) and asparagus (Telgmann-Rauber *et al.*, 2007). However, as the recombination rate in the SD locus in the present study is the lowest found in the whole genome, other mechanisms that suppress recombination are probably operating in addition to the mere localization near the centromere on chromosome XV.

As both willows and poplars are dioecious, the most parsimonious explanation is that dioecy evolved before the two lineages diverged more than 45 million years ago. How come heteromorphic sex chromosomes have not evolved during this time? One likely explanation is that there has been high rates of sex chromosome turnover in the Salicaceae family, which previous studies support. For example, SD loci in poplars are located on different positions on chromosome XIX and both male and female heterogamety is found. In contrast, *S. viminalis* is female heterogametic and the SD locus is located on chromosome XV. Although the SD loci have different positions in poplar and willow genomes, they could still be homologous and harbouring homologous sex determination genes, but through translocations have swapped genomic positions. If we hypothesize that markers are expected to have the same position in the *S. viminalis* and the *P. trichocarpa* genomes, differences can be the result of lineage-specific rearrangements, such as translocations, inversions, fusions and fissions. Using this strategy, we previously demonstrated that a number of both large and small rearrangements have happened as the split between poplars and willows, for example, the fusion in the willow lineage between parts of poplar chromosome I and poplar chromosome XVI, and the fission of poplar chromosome I (Berlin *et al.*, 2010). The resolution of this method depends on the density of the marker in the genomic region of interest and the closer the distances between the markers the smaller rearrangements can be detected. By adopting this method here on the finemapped chromosome XV and XIX, we conclude that no chromosome XIX marker mapped to chromosome XV, or vice versa. Thus, the SD locus in *S. viminalis* is likely not related to the SD locus on chromosome XIX in *P. trichocarpa*, which suggests that different sex determination mechanisms exist in willows and poplars. Therefore, willows and poplars together represent a system with high sex chromosome turn over rates, where new master sex determination genes are replacing the old ones on new chromosomes. This rapid turnover has also led to repeated transitions between male and female heterogamety in this plant family. Why is the turnover rate high in the Salicaceae family? It has been demonstrated that sexually antagonistic selection can result in the spread of new sex determination genes when linked to another autosomal gene and this would be particularly likely in systems with undifferentiated sex chromosomes (van Doorn and Kirkpatrick, 2007). In another theoretical study, it was also shown that sexually antagonistic selection can cause transitions between male and female heterogamety (van Doorn and Kirkpatrick, 2010). More studies are needed to investigate the evolutionary processes responsible for the high rates of turnover in sex determination mechanisms in the Salicaceae plant family.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

to the three anonymous referees who gave insightful comments and suggestions that greatly improved the manuscript.

Alström-Rapaport C, Lascoux M, Gullberg U (1997). Sex determination and sex ratio in the dioecious shrub *Salix viminalis* L. *Theor Appl Genet* **94**: 493–497.

Alström-Rapaport C, Lascoux M, Wang YC, Roberts G, Tuskan GA (1998). Identification of a RAPD marker linked to sex determination in the basket willow (*Salix viminalis* L.). *J Hered* **89**: 44–49.

Bachtrog D (2008). The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**: 1513–1525.

Berlin S, Ellegren H (2004). Chicken W: a genetically uniform chromosome in a highly variable genome. *Proc Natl Acad Sci USA* **101**: 15967–15969.

Berlin S, Lagercrantz U, von Arnold S, Öst T, Rönnberg-Wästljung AC (2010). High-density linkage mapping and evolution of paralogs and orthologs in *Salix* and *Populus*. *BMC Genomics* **11**: 129.

Boucher LD, Manchester SR, Judd WS (2003). An extinct genus of Salicaceae based on twigs with attached flowers fruits, and foliage from the Eocene Green River Formation of Utah and Colorado, USA. *Am J Bot* **90**: 1389–1399.

Charlesworth B, Charlesworth D (2000). The degeneration of Y chromosomes. *Philos Trans R Soc Lond B* **355**: 1563–1572.

Charlesworth D (2002). Plant sex determination and sex chromosomes. *Heredity* **88**: 94–101.

Charlesworth D, Charlesworth B, Marais G (2005). Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**: 118–128.

Charlesworth D, Mank JE (2010). The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics* **186**: 9–31.

Dai X, Hu Q, Cai Q, Feng K, Ye N, Tuskan GA *et al.* (2014). The willow genome and divergent evolution from poplar after the common genome duplication. *Cell Res* **24**: 1274–1277.

Gaudet M, Jorge V, Paolucci I, Beritognolo I, Mugnozza GS, Sabatti M (2008). Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. *Tree Genet Genomes* **4**: 25–36.

Gunter LE, Roberts GT, Lee K, Larimer FW, Tuskan GA (2003a). The development of two flanking SCAR markers linked to a sex determination locus in *Salix viminalis* L. *J Hered* **94**: 185–189.

Gunter LE, Kopp RF, McCord RP, Tuskan GA (2003b). Analysis of sex-linked, sequence-characterized amplified region markers in *Salix eriocephala*. *Can J Forest Res* **33**: 1785–1790.

Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP *et al.* (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**: 695–716.

Höglund S, Larsson S, Wingsle G (2005). Both hypersensitive and non-hypersensitive responses are associated with resistance in *Salix viminalis* against the gall midge *Dasineura marginemtorquens*. *J Exp Bot* **56**: 3215–3222.

Kersten B, Pakull B, Fladung M (2012). Mapping of the sex trait and sequence analysis of two sex linked genomic regions in *Populus tremuloides*. *ScienceMed* **3**: 203–210.

Kersten B, Pakull B, Groppe K, Lueneburg J, Fladung M (2014). The sex-linked region in *Populus tremuloides* Turesson 141 corresponds to a pericentromeric region of about two million base pairs on *P. trichocarpa* chromosome 19. *Plant Biol* **16**: 411–418.

Lander ES, Botstein D (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.

Manchester SR, Judd WS, Handley B (2006). Foliage and fruits of early poplars (Salicaceae: *Populus*) from the eocene of Utah, Colorado, and Wyoming. *Int J Plant Sci* **167**: 897–908.

Mank JE, Avise JC (2009). Evolutionary diversity and turn-over of sex determination in teleost fishes. *Sex Dev* **3**: 60–67.

Miura I (2007). An evolutionary witness: the frog *Rana rugosa* underwent change of heterogametic sex from XY male to ZW female. *Sex Dev* **1**: 323–331.

Moghadam HK, Pointer MA, Wright AE, Berlin S, Mank JE (2012). W chromosome expression responds to female-specific selection. *Proc Natl Acad Sci USA* **109**: 8207–8211.

Nanda I, Schlegelmilch K, Haaf T, Schartl M, Schmid M (2008). Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. *Cytogenet Genome Res* **122**: 150–156.

Pakull B, Groppe K, Meyer M, Markussen T, Fladung M (2009). Genetic linkage mapping in aspen (*Populus tremula* L. and *Populus tremuloides* Michx.). *Tree Genet Genomes* **5**: 505–515.

Paolucci I, Gaudet M, Jorge V, Beritognolo I, Terzoli S, Kuzminsky E *et al.* (2010). Genetic linkage maps of *Populus alba* L. and comparative mapping analysis of sex determination across *Populus* species. *Tree Genet Genomes* **6**: 863–875.

Perrin N (2009). Sex reversal: a fountain of youth for sex chromosomes? *Evolution* **63**: 3043–3049.

Renner SS (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *Am J Bot* **101**: 1588–1596.

Semerikov V, Lagercrantz U, Tsarouhas V, Rönnberg-Wästljung A, Alström-Rapaport C, Lascoux M (2003). Genetic mapping of sex-linked markers in *Salix viminalis* L. *Heredity* **91**: 293–299.

Slancarova V, Zdanska J, Janousek B, Talianova M, Zschach C, Zluvova J *et al.* (2013). Evolution of sex determination systems with heterogametic males and females in *Silene*. *Evolution* **67**: 3669–3677.

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ *et al.* (2009). The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* **461**: 267–271.

Stephens M, Scheet P (2005). Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am J Hum Genet* **76**: 449–462.

Stephens M, Smith NJ, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**: 978–989.

Telgmann-Rauber A, Jamsari A, Kinney MS, Pires JC, Jung C (2007). Genetic and physical maps around the sex-determining M-locus of the dioecious plant asparagus. *Mol Genet Genomics* **278**: 221–234.

Temmel NA, Hardeep SR, Cronk QCB (2007). Sequence characterization of the putatively sex-linked *Ssu 72*-like locus in willow and its homologue in poplar. *Can J Bot* **85**: 1092–1097.

Tuskan GA, DiFazio S, Faivre-Rampant P, Gaudet M, Harfouche A, Jorge V *et al.* (2012). The obscure events contributing to the evolution of an incipient sex chromosome in *Populus*: a retrospective working hypothesis. *Tree Genet Genomes* **8**: 559–571.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U *et al.* (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

van Doorn GS, Kirkpatrick M (2007). Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**: 909–912.

van Doorn GS, Kirkpatrick M (2010). Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* **186**: 629–U294.

Van Ooijen JW (2006). JoinMap ® 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Kyazma BV (ed), Wageningen, Netherlands.

Van Ooijen JW. (2009). MapQTL ® 6, Software for the Mapping of Quantitative Trait Loci in Experimental Populations. Kyazma BV (ed), Wageningen, Netherlands.

Voorrips RE (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* **93**: 77–78.

Yin T, Difazio SP, Gunter LE, Zhang X, Sewell MM, Woolbright SA *et al.* (2008). Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Res* **18**: 422–430.

Yu Q, Hou S, Hobza R, Feltus FA, Wang X, Jin W *et al.* (2007). Chromosomal location and gene paucity of the male specific region on papaya Y chromosome. *Mol Genet Genomics* **278**: 177–185.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)