Supplementary Table 1. Origins of male individuals for the *SlX9/SlY9* diversity study, the geographic region, and the intron type of the X-linked copy ("long" or "short", see text).

ID	Country	Population	Geographic region [1]	Intron 2 length
1	Germany	Bissendorf	1	short
2				short
3	Sweden	Oland, Grasgard	1	short
4				long
5	France	Relais des Chenes	3	Not determined
6				short
7	Greece	Ioanninon	3	short
8				long
9	Poland	Krakow, near Conference Centre	2	long
10				short
11	Denmark	Aarhus Botanic Garden [2]	1	short
12				short
13	Austria	Vienna, Heiligenstadt	2	short
14				short

---

[2] Population of origin unknown.

Supplementary Table 1, continued

15	Norway	Lardalsoyri	1	short
16				short
17	Greece	Evrou	3	short
18				short
19	Italy	near Tarquinia, Lazio	3	Not determined
20				long
21	Estonia	Vinupea	2	long
22				Not determined
23	Austria	Dietmans	2	Not determined
24				long
25	France	Canche	1	short
26	France			short
27	Portugal	Serre de Nogere	4	long
28				long
29	Netherlands	River Kraal	1	long
30				long
31	France	Vitry en Artois	1	short
32				short
33	Italy	Piedmont, Ceres	3	Not determined
34				short
35	UK	Dalkeith	1	long
36				long

Supplementary Table 1, continued

37	UK	Somerset, Burnham-on-Sea	1	short
38				short
41	Portugal	Segier-Chavez	4	short
42				Not determined
43	Spain	Madrid, El Escorial	4	long
44				long
45	Ukraine	Ukraine/Belarus border	2	Not determined
46				Not determined
47	Germany	near Glaubitz	1	long
48				short

Supplementary Table 2. *S. dioica* individuals used in this study.

Plant ID	Location of origin	Sex
Family 99K24.1, plant 1	France	female
Family A2009.1, plant 1	Finland [3]	female
Family A2009.1, plants 7 and 9	Finland	male
Family A2009.2, plant 1	Finland	female
Family 99M24.2 plant	Unknown	male
Family 99M9.1 plant	Unknown	male
FS01- FS09	Pentland Hills, Scotland	male
FS10- FS14	Pentland Hills, Scotland	female

---

[3] Tvärminne Zoological Station, Southern Finland (map reference: N59.844 E23.249)

Supplementary Table 3. Primers used in this study, and their uses (see also main text).

Name	Sequence	Notes
RB18_F	CTTGTGGAACTTCTGGTGGAAG	Establishing sex-linkage and deletion mapping of *SlY9*
RB18_R	GTCCAATCACATTCAAGTCTCTCC	
RB18_male_intron_2_F	TCTTTCACACCCAATTTGATCC	Amplifying *SlY9* with RB18_E3_rev
RB18_male_intron_2_R	GTACAGGGAAGAGCAAAGCAC	Amplifying *SlY9*
RB18_exon_1_F	AGCTAGCAGTTTTGCAGCATC	
RB18_Y_E3-R	GACATTCAAGTCTCTCCTCAGCCAA	Amplifying cDNA to get Y copy
RB18-Intron2-male-F-2	AAGGACAACAATTCAATGGGATG	*SlY9*-diversity with RB18-3' UTR
RB18-Intron2-male-F-3	GGGATGGAGGGAGTATGTTATTATTG	
RB18-3' UTR	GATGAATCTAAAATCAAACAGTGAAAC	
RB18_exon_1_F	AGCTAGCAGTTTTGCAGCATC	*SlX9*-diversity

Supplementary Table 3, continued

RB18_exon_4_R	TCTTCAGTCCTTCCTTTGAAGC
RB18-male-exon1-F	ACTCTCTCTCGCTCTTACTCC
RB18_male_exon3-R	GACATTCAAGTCTCTCCTCAGCCAA
RB18-E3-R-beg	TCCTCAGCCAGTCTCTTTGAA
RB18-E2-F-beg	GTTGCTGCAGTGAACCCTCT
RB18-EXON2-R-Male	TGCCTCAGAGGGTTTACTGC

Supplementary Table 4. Categories of the 94 variants found in the *S. latifolia SlX9* sequences. Variants in the two intron 2 types, long (L) and short (S) and listed separately. There were no sites where all S sequences differ from those in *S. dioica*.

Category of variants	Total numbers of polymorphic sites	Numbers of sites in categories: Variants in L intron 2 type	Numbers of sites in categories: Variants in S intron 2 type	Numbers of sites in categories: Variants shared by L and S haplotypes
I. Most sequences same as *S. dioica* state [4]	69			
Variants exclusive to *S. latifolia*	55 (36)	15 + 4	18 + 6	3 + 9
Shared polymorphisms with *S. dioica*	14	0	9 + 0	2 + 3
II. All L sequences entirely or largely different from *S. dioica* [5]	25			
Fixed difference between L and S sequences	3 + 1	0	0	—
Polymorphisms exclusive to S sequences	12	—	12 [6]	—
Polymorphisms exclusive to L sequences	1	1	—	—
Polymorphisms in S sequences shared with *S. dioica*	8	0	8	0
All S sequences different from *S. dioica*	0	—	—	—

---

[4] Excluding an intron 2 region that is present only in the S sequence type within *S. latifolia* (and also in *S. dioica*). The 17 polymorphic sites in this region are included in category II. After position 2024 we have sequences for only two *S. dioica* individuals, limiting the ability to find polymorphisms shared in both species; for category I sites, the values shown are the numbers before and after this point.

[5] In this category, sequences with the S intron 2 type all, or almost all, have the same state as *S. dioica*. Includes the intron 2 region present only in the S sequence type (see footnote 1).

[6] 9 of these are in regions present only in the S sequence type. At the 3 polymorphic sites present in both the L and S sequences, 2 polymorphisms in the S sequences are for the state found in the L sequences.

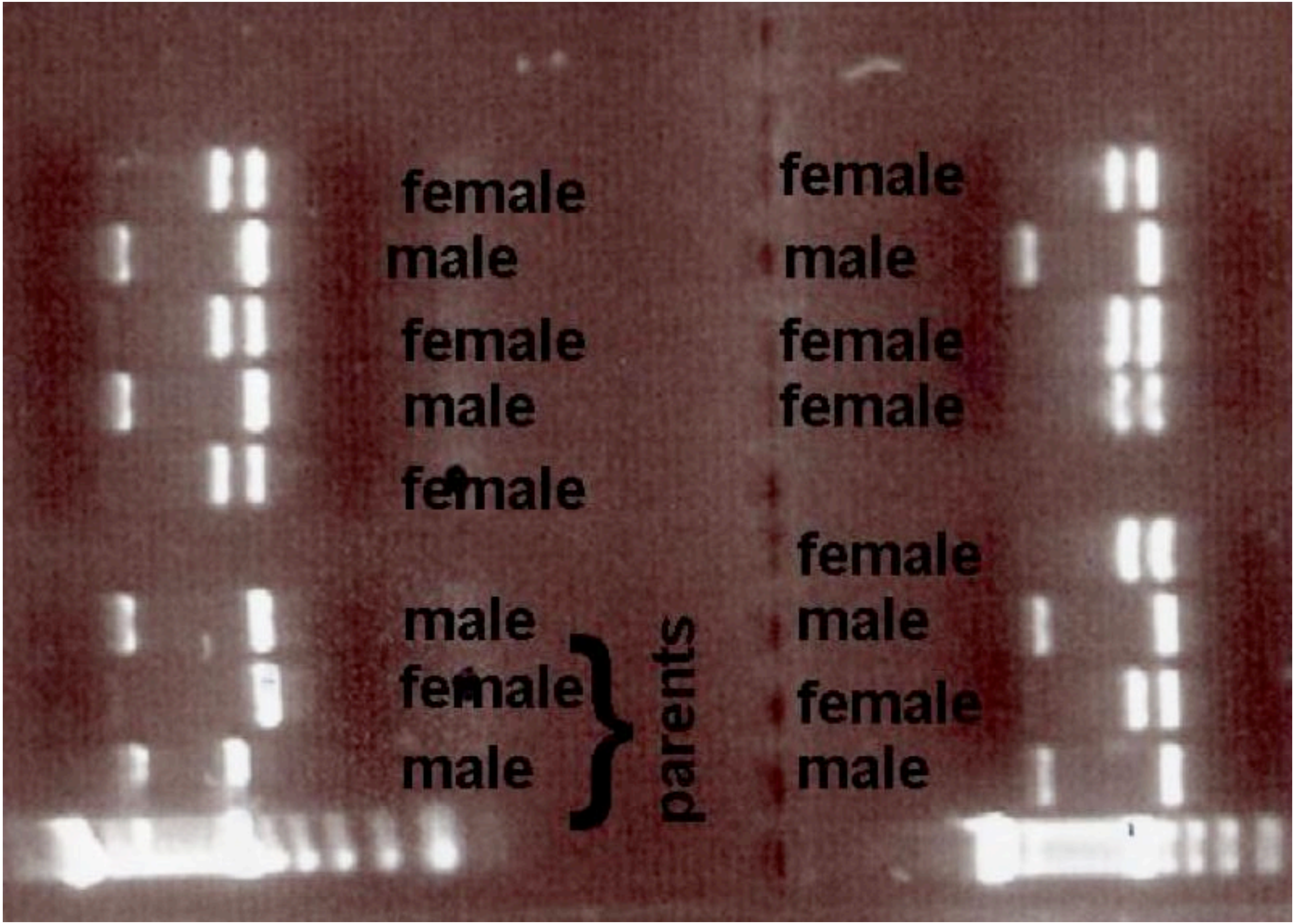
Supplementary Table 5. Sites with significant linkage disequilibrium in the *S. latifolia SlX9* diversity sample.

Position 1 (bp)	Position 2 (bp)	Distance (bp)
734	2268	602
794	875	69
2057	2138	79
2057	2141	82
2108	2114	6
2126	2138	12
2126	2141	15
2138	2141	3
2195	2198	3

Supplementary Figures

Figure S1. PCR products of *SlX9* in family H2005-1, showing the segregation of the informative intron size variants in the family H2005-1 (see Methods).

The female parent carries only one variant (480 bp*,* *SlX9*). The father carries two differently sized variants, (a long Y-linked genomic DNA fragment, of ~ 1200 bp, inherited by all male offspring, and one of intermediate size, 600 bp, which segregates as X-linked*;* all female offspring inherited the maternal 480 bp band, plus the paternal 600 bp one).


female
male
female
male
female
male
female
male
parents
female
male
female
female
female
male
female
male

Supplementary Figure 2. Polymorphisms is the SlX9 and SlY9 sequences.

See Excel file.