

GENETICS

Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.110.116772/DC1>

A Genetic Model for the Female Sterility Barrier Between Asian and African Cultivated Rice Species

Andrea Garavito, Romain Guyot, Jaime Lozano, Frédéric Gavory, Sylvie Samain, Olivier Panaud, Joe Tohme, Alain Ghesquière and Mathias Lorieux

Copyright © 2010 by the Genetics Society of America
DOI: 10.1534/genetics.110.116772

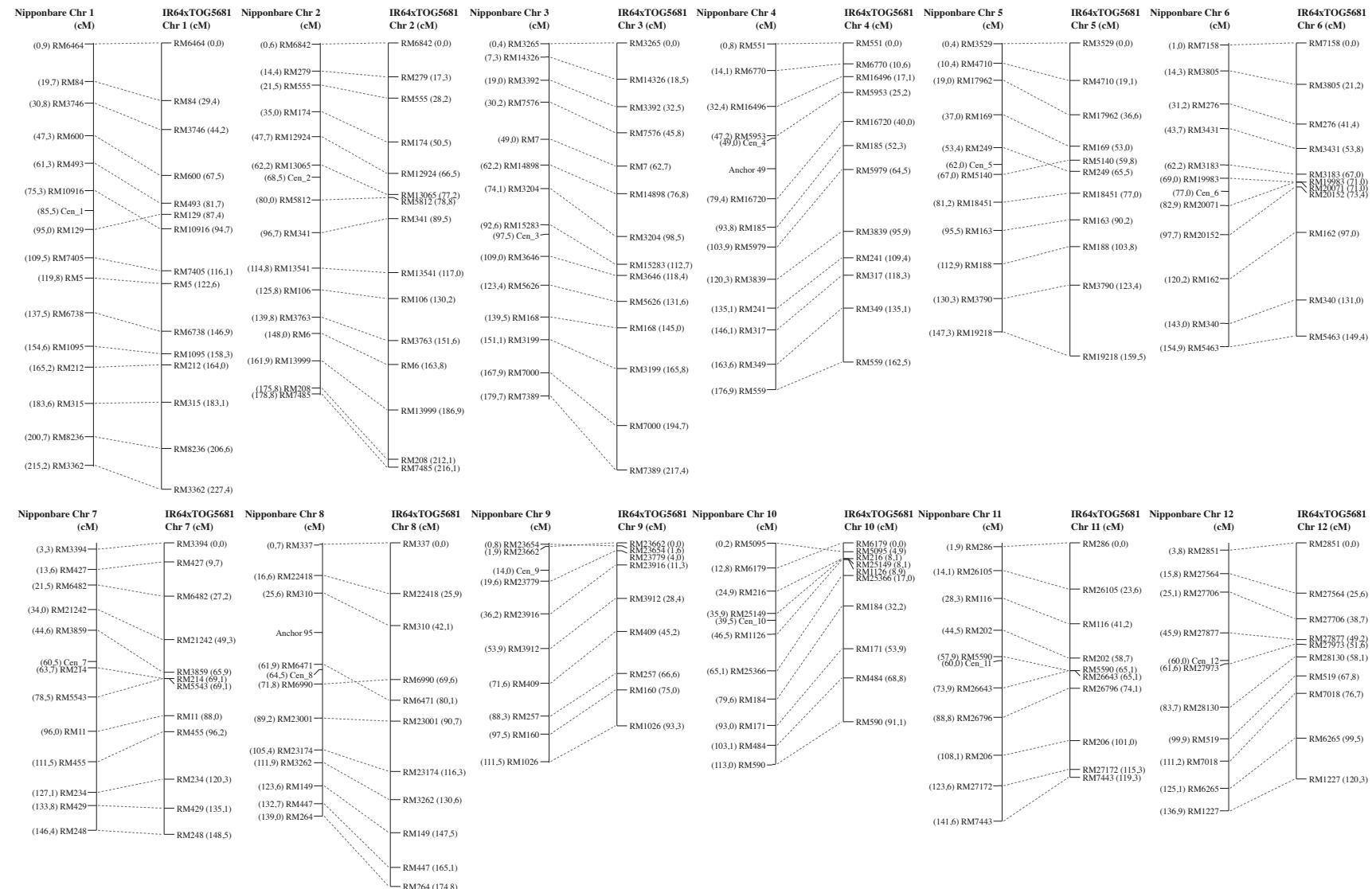


FIGURE S1.— Comparison of the *O. sativa* x *O. glaberrima* genetic map with the *O. sativa* cv. Nipponbare physical map. Comparison of the obtained genetic map for the IR64/TOG5681/IR64 cross, with the physical localization of SSR markers in the *O. sativa* cv. Nipponbare genome, based in the TIGR V.6 available data. **Cen:** Centromeric region.

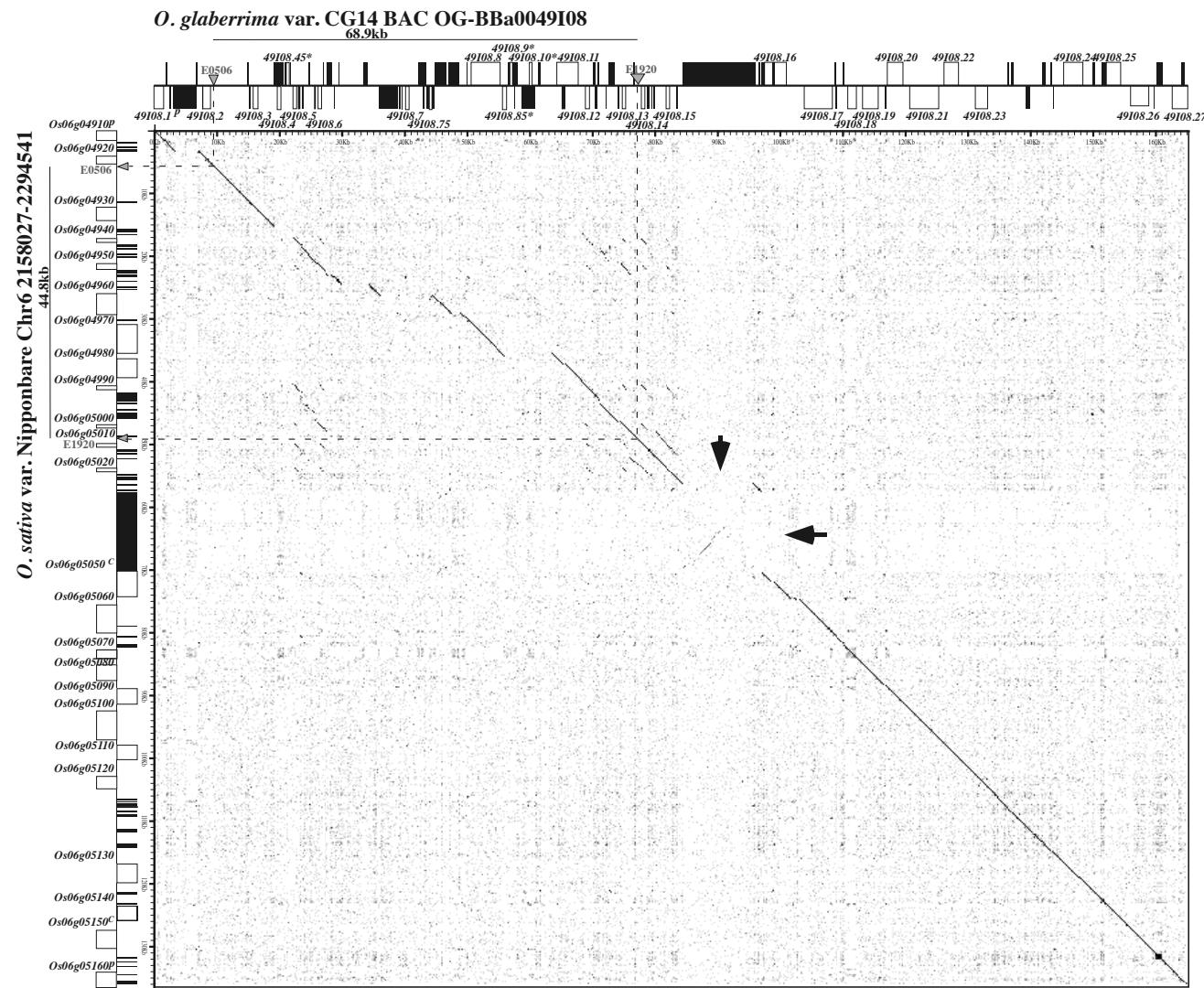


FIGURE S2.—Graphical comparison between *O. glaberrima* cv. CG14 BAC OG-BBa0049I08 and the orthologous *O. sativa* cv. Nipponbare genomic segment. Pairwise alignments of the *O. glaberrima* cv. CG14 BAC OG-BBa0049I08 and the respective *O. sativa* cv. Nipponbare orthologous region. Predicted genes (empty boxes) and TEs (dark bars) are shown for each of the sequences. Gray triangles mark the locus *S₁*, as shown in Figure 4. **C:** Nipponbare genes manually corrected. **P:** Partial gene.

#: Pseudogene. The insertion of two LTR retrotransposons are indicated with arrows.

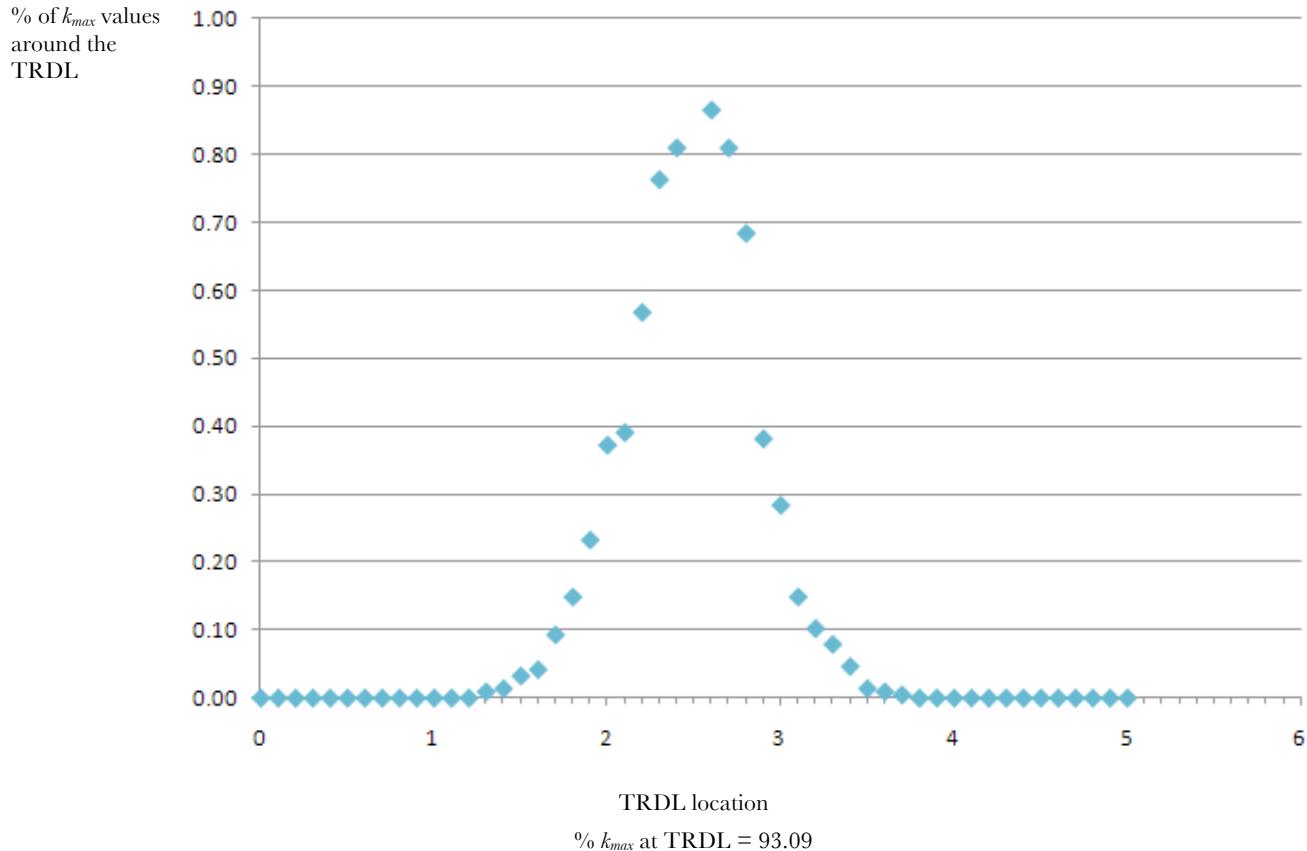
FILE S1**Genetic models for the allelic female gamete elimination caused by the *SI* locus**

File S1 is available for download as an Excel file (.xlsx) at <http://www.genetics.org/cgi/content/full/genetics.110.116772/DC1>.

FILE S2

Estimation of the precision of the mapping of a transmission ratio distorter locus (TRDL) using the k_{max} measurement by simulation

In order to estimate the precision of the mapping of a transmission ratio distorter locus (TRDL) using the k_{max} measurement, we simulated 10,000 BC1F1 populations of 734 individuals each, bearing one chromosome segment of 5 cM long spanned by 50 markers (10 markers per cM). A TRDL was located in the middle of the segment at the position 2.5 cM, and induced a differential viability rate of 0.95. For each population generated, the k statistic was computed for each marker. We then determined the percentage of cases where the maximum value of k (k_{max}) was observed on each marker surrounding the TRDL and plotted this percentage value against the position on the chromosome segment.



The simulation results showed a percentage of k_{max} of 93.09 % at the TRDL position, and showed that 95.58 % of the k_{max} fell in an interval of 0.3 cM around the TRDL, meaning that there is a 95 % probability of observing a k_{max} in an interval of slightly less than 0.3 cM around the true TRDL position. Converting this genetic distance to a physical one using the local rate of 80 kbp/cM (derived from our data, see Figure 3) leads to an interval of slightly less than 24 kb.

FILE S3

**Equations for calculating the expected gametic and allelic frequencies
under Model 1**

The expected frequencies F_c of the different 1n-2c configurations can be derived from the recombination frequencies between the three S_I genes and between the S_I complex locus and the centromere (see schematic representation 1 below), and are given in Table S3-1.

The expected frequencies of each gamete that belong to a specific 1n-2c configuration are equal to $F_c/4$, where F_c is the expected frequency of the specific 1n-2c configuration.

Let D be the sum of the frequencies of the viable gametes. Under model 1, D simplifies to:

$$D = \left(1 - 2r_5\right) \left(\frac{r_2 + 1}{2}\right) + 2r_5 \left[\frac{1}{2} + \frac{r_1}{2} - r_1 r_2\right]$$

The expected frequencies of *viable* gametes are obtained by dividing their corresponding expected frequency by D .

Thus, the final frequency of gametes bearing the S_I^s allele, $f(S_I^s)$, is obtained by summing the individual frequencies of viable gamete that bear this allele. Under Model 1, this simplifies to:

$$f(S_I^s) = \left[\frac{r_2}{2} + r_5(r_1 - r_2 - 2r_1 r_2) \right] / D$$

Similar equations can be derived for Models 2 to 4, with numerical examples based on observed data given in File “S1 - Genetic models.xlsx”.

Under all models, the estimated recombination fractions r_1 and r_2 are naturally given by dividing the observed recombination fractions, $r_{1\max}$ and $r_{2\max}$, by $2D$.

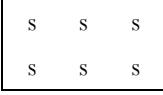
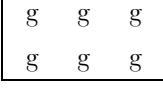
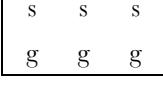
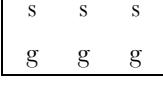
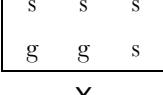
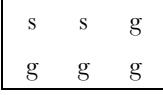
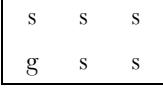
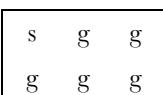
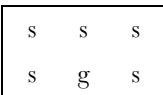
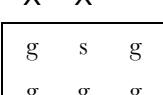
$$\hat{r}_1 = \frac{r_{1\max}}{2D}; \quad \hat{r}_2 = \frac{r_{2\max}}{2D}.$$

In the case of the presence of the additional factor $S_I C$ (see schematic representation 2 below), the effect of $S_I C$ is applied after the initial selection due to the $S_I A^- S_I^- S_I B$ locus action.

As $S_I C$ only affects the S_I^s gametes when a recombination event occurs between $S_I B$ and $S_I C$, the expected frequencies of these gametes are obtained by multiplying them by $1 - r_3$, where r_3 is the recombination fraction between $S_I B$ and $S_I C$, and dividing them by the relative sum of all gamete frequencies, $1 - r_3 f(S_I^s)$. The frequencies of other viable gametes are obtained by dividing them by the relative sum of all gamete frequencies, $1 - r_3 f(S_I^s)$.

TABLE S3-1

Expected frequencies F_c of the different 1n-2c configuration expressed in function of the recombination frequencies between the three S_I genes, and associated gametic frequencies under no selection and Model 1 selection

1n-2c configuration	<i>Expected configuration frequency without selection (F_c)</i>	<i>Expected gamete frequency without selection (F_g)</i>	<i>Survival under Model 1</i>	<i>Expected gamete frequency under selection (Model 1)</i>
$S_{IA} \quad S_I \quad S_{IB}$				
	$F_c^1 = (1 - 2r_5)[1 - 2(r_1 + r_2 - 2r_1r_2)]$	$F_c^1/4$	Aborted	
		$F_c^1/4$	Aborted	
	$F_c^2 = 2r_5[1 - 2(r_1 + r_2 - 2r_1r_2)]$	$F_c^2/4$	Viable	F_g/D
		$F_c^2/4$	Viable	F_g/D
	$F_c^3 = 2r_2(1 - 2r_1)(1 - 2r_5)$	$F_c^3/4$	Viable	F_g/D
		$F_c^3/4$	Viable	F_g/D
	$F_c^4 = 2r_1(1 - 2r_2)(1 - 2r_5)$	$F_c^4/4$	Aborted	0
		$F_c^4/4$	Aborted	0
	$F_c^5 = 4r_1r_2(1 - 2r_5)$	$F_c^5/4$	Viable	F_g/D
		$F_c^5/4$	Viable	F_g/D

<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>s</td><td>s</td><td>s</td></tr> <tr><td>s</td><td>s</td><td>g</td></tr> <tr><td colspan="3">X</td></tr> <tr><td>g</td><td>g</td><td>s</td></tr> <tr><td>g</td><td>g</td><td>g</td></tr> </table>	s	s	s	s	s	g	X			g	g	s	g	g	g	$F_c^6 = 4r_2r_5(1 - 2r_1)$	$F_c^6/4$ $F_c^6/4$ $F_c^6/4$ $F_c^6/4$ $F_c^6/4$ $F_c^6/4$	Aborted Aborted Viable Viable Viable Viable	0 0 F_g/D F_g/D F_g/D F_g/D
s	s	s																	
s	s	g																	
X																			
g	g	s																	
g	g	g																	
<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>s</td><td>s</td><td>s</td></tr> <tr><td>s</td><td>g</td><td>g</td></tr> <tr><td colspan="3">X</td></tr> <tr><td>g</td><td>s</td><td>s</td></tr> <tr><td>g</td><td>g</td><td>g</td></tr> </table>	s	s	s	s	g	g	X			g	s	s	g	g	g	$F_c^7 = 4r_1r_5(1 - 2r_2)$	$F_c^7/4$ $F_c^7/4$ $F_c^7/4$ $F_c^7/4$	Aborted Viable	0 F_g/D
s	s	s																	
s	g	g																	
X																			
g	s	s																	
g	g	g																	
<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>s</td><td>s</td><td>s</td></tr> <tr><td>g</td><td>s</td><td>g</td></tr> <tr><td colspan="3">X X</td></tr> <tr><td>s</td><td>g</td><td>s</td></tr> <tr><td>g</td><td>g</td><td>g</td></tr> </table>	s	s	s	g	s	g	X X			s	g	s	g	g	g	$F_c^8 = 8r_1r_2r_5$	$F_c^8/4$ $F_c^8/4$ $F_c^8/4$ $F_c^8/4$	Aborted Aborted Viable Viable	0 0 F_g/D F_g/D
s	s	s																	
g	s	g																	
X X																			
s	g	s																	
g	g	g																	

The sum of the frequencies of the viable gametes thus simplifies to:

$$D' = D / \left(1 - f(S_1^s) \right) r_3,$$

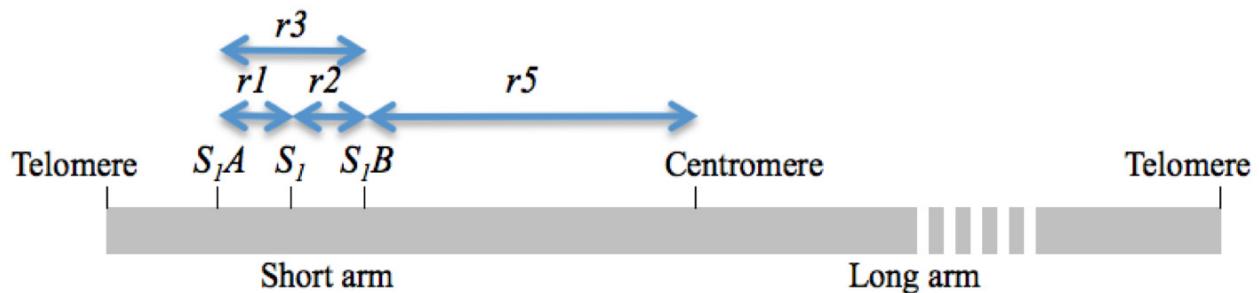
while the final frequency of gametes bearing the S_1^s allele converts to:

$$f(S_1^s) = (1 - r_3) f(S_1^s) / \left(1 - f(S_1^s) r_3 \right),$$

and the estimates for r_1 and r_2 convert to:

$$\hat{r}_1 = \frac{r_1 \max}{2D'}, \quad \hat{r}_2 = \frac{r_2 \max}{2D'}.$$

Schematic representation 1: positions of the three S_I genes on rice chromosome 6, expressed as recombination fractions between each other and between the S_I locus and the centromere.



Schematic representation 2: positions of the three S_I genes and the $S_I C$ gene on rice chromosome 6, expressed as recombination fractions between each other and between the S_I locus and the centromere.

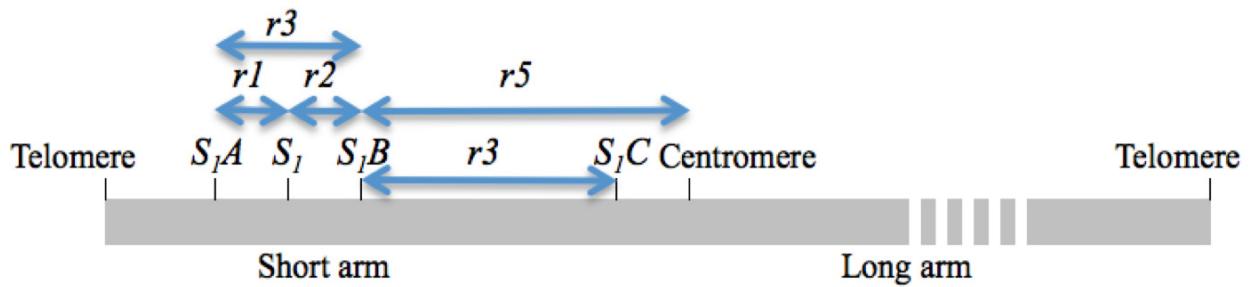


TABLE S1**SSR markers and PCR conditions used for saturation of chromosome 6**

<i>Molecular marker</i>	Position in Nipponbare Chromosome 6 (TIGR V.6)	Annealing Temperature °C	MgCl ₂ final concentration (mM)
RM7158	216 292	50	2.5
RM435	537 327	55	1.5
RM589	1 380 866	55	1.5
RM190	1 764 638	55	1.5
RM19350	1 967 120	55	1.8
RM19353	2 084 117	57	1.5
RM19357	2 175 040	55	1.8
RM19359	2 205 367	59	2.0
RM19360	2 215 120	55	2.0
RM19361	2 245 211	55	1.8
RM5199	2 270 164	55	1.8
RM19363	2 288 872	55	1.8
RM19367	2 335 233	55	1.5
RM19369	2 349 474	55	2.0
RM19377	2 440 456	55	2.0
RM19387	2 659 693	55	1.5
RM19391	2 714 833	55	1.5
RM3805	2 853 068	55	1.5
RM19414	2 941 180	59	1.5
RM19420	3 076 926	61	1.5
RM204	3 168 377	55	1.5
RM19496	4 351 871	55	1.8
RM276	6 230 046	55	1.8
RM3431	8 744 955	50	1.8
RM3183	12 447 059	50	1.5
RM19983	13 808 188	55	1.5
RM20071	16 542 271	55	1.8
RM20152	19 542 217	55	1.8
RM6036	21 548 402	59	1.5
RM20313	23 161 984	55	1.5
RM162	24 035 501	61	2.0
RM5371	25 825 382	55	1.5
RM340	28 599 182	55	1.8
RM5463	30 985 008	55	1.8

TABLE S2
New SSR and InDel markers in the *S₁* locus

Molecular marker	Position in Nipponbare Chromosome 6 (TIGR V.6)	Forward Primer	Reverse Primer	Annealing Temperature °C	MgCl ₂ concentration mM
RMC6_21678	2 167 869	AGGGAGCTCCAGTGTGAGA	TTGTGCCCAAGCTTAAGGTA	59	1.5
C6_21774	2 177 472	CAGTTTGGGCATTTTCG	ATAGGCGATCCCTAGAGCAA	61	1.5
C6_21788	2 178 838	GCTGCGATGATAAGGCAATA	CCCCTAGCTAGTTGGTGTGA	61	1.5
C6_21804	2 180 401	CCTTTACTTCTTTGGAAGTTGG	GCACCACAGTTTCAGGCTCT	61	1.5
C6_21824	2 182 413	CTGCTACCATCCCTCTGCTC	GGAGCAATTGAGAAGCTG	61	1.5
C6_21837	2 183 714	TGGATAGGGCTATAGGAGCA	CTTCCCTCTCTGGATCGAC	61	1.5
RMC6_21851	2 185 106	AAGGGCCTTCCTCTAGCC	AGGGCGATTGACGCACAT	59	1.5
RMC6_21942	2 194 257	TCCAATCCCCCTAAATTTCAG	TTTACCATGGGAACTTGGA	59	1.5
RMC6_21989	2 198 927	CCTGATGAGCTCTGCTGATG	GTGCGTGCTTGTGAATCTG	59	1.5
RMC6_22028	2 202 836	CACCATCCTCCTCCATGTTT	CCCTTTCATTTGCCGTCTA	59	1.5
RMC6_22046	2 204 666	AAGAAGGTGAAGACATCAACGA	GCTGCTGAGTTCATGCGTAA	59	1.5
RMC6_22105	2 210 541	GATTITGCCGTCAAGTTGG	CCATACACGCAGCTCACAG	59	1.5
C6_22305	2 230 535	CAAGTGAAGCTGGTCATTGG	TCCGTTGTTGGAATGGAAT	61	1.5
C6_22386	2 238 675	TGTATGAGAGGGTAGGATCAGTCA	TGCTCATGGAGAGCGATAAA	61	1.5
RMC6_22500	2 250 065	CTGCATCCAAAGTTCTGAAGG	ATCTGTTCATGGGGCTGTA	59	1.5
RMC6_22639	2 263 917	TTGCCATTCATGGAAATTAG	GACTCTTAGCCACCAAGGAAA	59	1.5
C6_22773	2 277 347	GACCTTGATGTCACCAGCA	TCGAAATTGGAGCACTCTGA	61	1.5
RMC6_22854	2 285 471	GGCTAGAGCGGTGATTTC	TGGTAGGCACAAGGATAGGG	59	1.5
RMC6_25568 *	2 556 896	ACGTCAGGGAGAGGTACTTT	CCCTAGGAAGGAAGAGAGAA	55	1.5

SSR and InDels based molecular markers, designed by comparison of the *O. glaberrima* cv. CG14 BAC OG-BBa0049I08 with the orthologous segment from *O. sativa* cv. Nipponbare. PCR conditions used for each marker are shown. *: SSR marker designed based only in the TIGR V6 data (putative_ss_r_21715).

TABLE S3**List of identified putative TEs in the *O. glaberrima* BAC OG-BBA0049I08**

Class	Repeat type	Group	Number	Size (bp)
I	LTR retrotransposon	LTR retrotransposon	7	18534
	Non-LTR retrotransposon	LINE	1	161
		SINE	2	401
II	Transposons	DNA Transposon	29	16887
		MITE	34	8279
	Unclassified	Unclassified	1	905
Total			74	45167

TABLE S4

Sequence annotation and comparison between orthologous coding sequences in the *O. glaberrima* cv. CG14 BAC OG-BBa0049I08 and *O. sativa* cv. Nipponbare

<i>O. glaberrima</i> Gene Name	Best BLASTN homology	Best BLASTN EST homology	Best BLASTX homology (swiss prot)	Protein domain	Putative Function	Putative <i>O. sativa</i> (Nipponbare) orthologous gene	% of nucleotide identity	% of protein identity	% of similarity	Length (bp)	Ka (dN)	Ks (dS)	Ka/Ks
OG-BBa0049I08.1*	NM_001063292 <i>O. sativa</i> Os06g0141100 (0.0)	CB618108 <i>O. sativa</i> (0.0)	Q94A82 <i>A. thaliana</i> NADH pyrophosphatase NUDT19 (8e-69) <i>O. sativa</i> flcDNA (0.0)	pfam09296, NUDIX-like; cd03429, NADH pyrophosphatase	Putative NADH pyrophosphatase protein	LOC_Os06g04910 p	99	98.4	98.4	1224	0.0025	0.0000	99
OG-BBa0049I08.2	CT830197 <i>O. sativa</i> flcDNA (0.0)	CT850081 <i>O. sativa</i> (0.0)	O13801 <i>Schizosaccharomyces pombe</i> RNA-binding protein (2e-9)	pfam00641, Zn-finger	Putative Zinc finger protein	LOC_Os06g04920	99.3	99.3	100	438	0.0028	0.0266	0.1063
OG-BBa0049I08.3	AK121184 <i>O. sativa</i> flcDNA (2e-131)	EE592322 <i>O. sativa</i> (5e-169)	Q02921 <i>G. max</i> Early nodulin 93 (2e-11)	pfam03386, Early nodulin 93 ENOD93 protein	Putative ENOD93 protein	LOC_Os06g04930	61.3	60.9	62.2	/	/	/	/
OG-BBa0049I08.4	AC136219 <i>O. sativa</i> BAC (0.0)	CT843513 <i>O. sativa</i> (1e-92)	P73714 Synechocystis Zinc metalloprotease (1e-9)	cd06163 zinc metalloproteases	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.45*	/	/	/	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.5	AK122162 <i>O. sativa</i> flcDNA (2e-148)	EL586675 <i>O. sativa</i> (2e-162)	Q02921 <i>G. max</i> Early nodulin 93 (8e-20)	pfam03386, Early nodulin 93 ENOD93 protein	Putative ENOD93 protein	LOC_Os06g04940	65	54.1	59.5	/	/	/	/
OG-BBa0049I08.6	AK121791 <i>O. sativa</i> flcDNA (1e-157)	CI240527 <i>O. sativa</i> (1e-158)	Q02921 <i>G. max</i> Early nodulin 93 (4e-14)	pfam03386, Early nodulin 93 ENOD93 protein	Putative ENOD93 protein	LOC_Os06g04950	89.1	76.8	78.4	/	/	/	/
OG-BBa0049I08.7	/	/	/	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.75	/	/	/	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.8	NM_001063299 <i>O. sativa</i> Os06g0142000 (0.0)	CB622551 <i>O. sativa</i> (0.0)	Q9SH88 <i>A. thaliana</i> Ribosome biogenesis regulatory protein (3e-48)	/	Putative protein	LOC_Os06g04970	98.6	97.7	98.4	2073	0.0112	0.0237	0.4728
OG-BBa0049I08.85*	AP002838 <i>O. sativa</i> BAC (2e-156)	CT858525 <i>O. sativa</i> (8e-67)	/	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.9*	AP004090 <i>O. sativa</i> BAC (8e-69)	CT844843 <i>O. sativa</i> (3e-68)	/	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/
OG-BBa0049I08.10*	AC121363 <i>O. sativa</i> BAC (3e-128)	CB649669 <i>O. sativa</i> (1e-57)	P47927 <i>A. thaliana</i> Floral homeotic protein APETALA 2 (4e-8)	/	Hypothetical protein	no orthologous gene	/	/	/	/	/	/	/

OG-BBa0049I08.11	AK073027 <i>O. sativa</i> flcDNA (0.0)	CB658549 <i>O. sativa</i> (0.0)	Q9FJT2 <i>A. thaliana</i> F-box/FBD/LRR-repeat (5e-4) <i>O. sativa</i> (0.0)	/	Putative F-box protein	LOC_Os06g04980	98.1	96.5	97.7	1542	0.0173	0.0198	0.8715
OG-BBa0049I08.12	AK122162 <i>O. sativa</i> flcDNA (2e-175)	CI252819 <i>O. sativa</i> (2e-175)	Q02921 <i>G. max</i> Early nodulin 93 (1e-17) <i>O. sativa</i> (4e-177)	pfam03386, Early nodulin 93 ENOD93 protein (1e-14)	Putative ENOD93 protein	LOC_Os06g04990	99.7	100	100	348	0.0000	0.0142	0.001
OG-BBa0049I08.13	CT837557 <i>O. sativa</i> flcDNA (4e-177)	CT856080 <i>O. sativa</i> (4e-177)	Q02921 <i>G. max</i> Early nodulin 93 (1e-14) <i>O. sativa</i> (5e-176)	pfam03386, Early nodulin 93 ENOD93 protein (1e-17)	Putative ENOD93 protein	LOC_Os06g05000	100	100	100	348	0.0000	0.0000	99
OG-BBa0049I08.14	CT827995 <i>O. sativa</i> flcDNA (1e-158)	EE592204 <i>O. sativa</i> (5e-176)	Q02921 <i>G. max</i> Early nodulin 93 (1e-17) <i>O. sativa</i> (1e-176)	pfam03386, Early nodulin 93 ENOD93 protein (1e-17)	Putative ENOD93 protein	LOC_Os06g05010	99.4	100	100	351	0.00002	0.0260	0.001
OG-BBa0049I08.15	CT827995 <i>O. sativa</i> flcDNA (9e-173)	CI050172 <i>O. sativa</i> (1e-176)	Q02921 <i>G. max</i> Early nodulin 93 (1e-17) <i>O. sativa</i> (1e-176)	pfam03386, Early nodulin 93 ENOD93 protein (1e-17)	Putative ENOD93 protein	LOC_Os06g05020	99.1	98.3	99.1	348	0.0074	0.0143	0.5175
OG-BBa0049I08.16	NM_001063304 <i>O. sativa</i> Os06g0142500 (0.0)	CB617832 <i>O. sativa</i> (0.0)	Q9LMN7 <i>A. thaliana</i> Wall-associated receptor kinase 5 (5e-130)	cd00180, Serine/Threonine protein kinases	Putative Serine/Threonine protein kinases	LOC_Os06g05050	92	87.8	92.1	2291	0.0447	0.0571	0.7828
OG-BBa0049I08.17	NM_001063305 <i>O. sativa</i> Os06g0142600 (0.0)	CB635675 <i>O. sativa</i> (0.0)	O82804 <i>A. thaliana</i> Protein Early flowering 3 (1e-55)	/	Putative Early flowering protein	LOC_Os06g05060	99.2	98.8	99.1	2283	0.0057	0.0155	0.3669
OG-BBa0049I08.18	AP000399	CT860562 <i>O. sativa</i> BAC (0.0)	O23081 <i>A. thaliana</i> Cysteine-rich receptor-like protein kinase 41 (2e-57)	cd00180, Serine/Threonine protein kinases	Putative Serine/Threonine protein kinases	LOC_Os06g05070	97.6	97.6	98.5	1386	0.0081	0.1700	0.0465
OG-BBa0049I08.19	CT834467 <i>O. sativa</i> flcDNA (0.0)	CT848007 <i>O. sativa</i> (0.0)	P00428 <i>Bos taurus</i> Cytochrome c oxidase subunit 5B, mitochondrial (6e-14)	cd00924, Cytochrome c oxidase subunit Vb	Putative cytochrome c oxidase	LOC_Os06g05080	99.8	99.4	99.4	465	0.0032	0.0000	99
OG-BBa0049I08.20	AK103971 <i>O. sativa</i> flcDNA (0.0)	CA998930 <i>O. sativa</i> (0.0)	Q9SNQ2 <i>O. sativa</i> protein arginine N-methyltransferase (0.0)	cd02440, S-adenosylmethionine-dependent methyltransferases	Putative S-adenosylmethionine-dependent methyltransferases	LOC_Os06g05090	98.8	98.4	99	1143	0.0071	0.0299	0.2366
OG-BBa0049I08.21	AK121920 <i>O. sativa</i> flcDNA (0.0)	EC366141 <i>O. sativa</i> (0.0)	Q38854 <i>A. thaliana</i> 1-deoxy-D-xylulose-5-phosphate synthase, chloroplastic (0.0)	cd02007, Thiamine pyrophosphate (TPP) family, DXS subfamily, TPP-binding module	Putative 1-deoxy-D-xylulose-5-phosphate synthase	LOC_Os06g05100	99.9	99.6	99.7	2169	0.0018	0.0000	99
OG-BBa0049I08.22	AK071301 <i>O. sativa</i> flcDNA (0.0)	CB672512 <i>O. sativa</i> (0.0)	P22302 <i>Nicotiana plumbaginifolia</i> Superoxide dismutase chloroplastic (5e-60)	PRK10543, superoxide dismutase; pfam02777, Iron/manganese superoxide dismutases	Putative Superoxide dismutase	LOC_Os06g05110	99.9	99.6	100	768	0.0016	0.0000	99
OG-BBa0049I08.23	AK061597 <i>O. sativa</i> flcDNA (9e-80)	CI020918 <i>O. sativa</i> (9e-80)	/	/	Putative protein	LOC_Os06g05120	/	/	/	/	/	/	/

OG-BBa0049I08.24	AK120946 <i>O. sativa</i> flcDNA (0.0)	CX101001 <i>O. sativa</i> (0.0)	Q9SQI3 <i>Gossypium hirsutum</i> Myristoyl-acyl carrier protein thioesterase, chloroplastic (6e-146)	cd00586, 4-hydroxybenzoyl- CoA thioesterase (4HBT)	Putative Myristoyl- acyl carrier protein thioesterase	LOC_Os06g05130	97.4	99	99	1284	0.0044	0.0300	0.1457
OG-BBa0049I08.25	AP003487 <i>O. sativa</i> BAC (0.0)	CA083999 <i>Saccharum officinarum</i>	O49287 <i>A. thaliana</i> Putative pentatricopeptide repeat- (1e-99) containing protein (2e-93)	pfam01535, PPR repeat	Putative PPR protein	LOC_Os06g05140	99.8	99.6	99.6	2292	0.0018	0.0037	0.4789
OG-BBa0049I08.26	AP003487 <i>O. sativa</i> BAC (0.0)	CF326731 <i>O. sativa</i> (0.0)	Q80SY5 <i>M. musculus</i> Pre-mRNA- splicing factor 38B (2e-40)	pfam03371, PRP38 family	Putative protein	LOC_Os06g05150	98	97.2	97.5	1296	0.0070	0.0048	1.471
OG-BBa0049I08.27 ^p	AK067270 <i>O. sativa</i> flcDNA (0.0)	FL869738 <i>Panicum</i> <i>virgatum</i> (0.0)	Q9LW86 <i>A. thaliana</i> Probable sulfate transporter (1e-141)	pfam00916, Sulfate transporter family	Putative sulfate transporter	LOC_Os06g05160	/	/	/	/	/	/	/

p: Partial gene; *****: Pseudogene; **C:** Nipponbare genes manually corrected. Genes in grey are located in the *S*₇ locus.