

Identification of marker-trait associations for self-compatibility in a segregating mapping population of *Theobroma cacao* L.

Stefan Royaert · Wilbert Phillips-Mora · Adriana M. Arciniegas Leal · Kathleen Cariaga · James Steven Brown · David N. Kuhn · Raymond J. Schnell · Juan Carlos Motamayor

Received: 8 September 2010 / Revised: 6 June 2011 / Accepted: 16 June 2011
© Springer-Verlag 2011

Abstract Increasing yield, quality, and disease resistance are important objectives for cacao breeding programs. Yield reduction in material improved for other traits is one of the main constraints caused by self-incompatibility (SI). Genes regulating SI in cacao have not been identified; therefore, knowledge of the location of genetic markers for and the effects of alleles determining SI will be useful for selecting uniformly self-compatible cultivars with higher yields. In a mapping population originating from a cross between a self-incompatible clone, Pound-7, and a self-compatible clone, UF-273, we observed important differences in flower retention at 15, 21, and 28 days after pollination. Our

results suggest that the best time to measure flower retention is 15 days after pollination or later and that selecting self-compatibility (SC) thresholds is genotype specific. Using marker-trait association analysis we identified one marker, mTcCIR222, strongly associated with SC, as well as three surrounding markers (mTcCIR168, mTcCIR115 and mTcCIR158), all located near the proximal end of linkage group 4. These markers are currently being tested in our marker-assisted selection program.

Keywords *Theobroma cacao* · Self-incompatibility · Pollination · Genetic map · Marker-trait association

Communicated by D. Grattapaglia

Electronic supplementary material The online version of this article (doi:10.1007/s11295-011-0403-5) contains supplementary material, which is available to authorized users.

S. Royaert · K. Cariaga · J. S. Brown · D. N. Kuhn · R. J. Schnell · J. C. Motamayor
USDA-ARS, Subtropical Horticulture Research Station,
13601 Old Cutler Road,
Miami, FL 33158, USA

W. Phillips-Mora · A. M. Arciniegas Leal
Tropical Agricultural Research and Higher Education
Center (CATIE),
Turrialba, Costa Rica

J. C. Motamayor
MARS Inc.,
Hackettstown, NJ 07840, USA

J. C. Motamayor (✉)
MARS Inc. c/o USDA-ARS, Subtropical Horticulture
Research Station,
13601 Old Cutler Road,
Miami, FL 33158, USA
e-mail: juan.motamayor@effem.com

Introduction

Theobroma cacao L. (cacao) is a tree known for being the source of cacao beans and its main derived product, chocolate. Cacao ($2n=2x=20$, Dantas and Guerra 2010) belongs to the Malvaceae sensu lato (Alverson et al. 1999), and its genome size is estimated to be 430 Mbps (Argout et al. 2011). Native to the Amazonian basin (Motamayor et al. 2008; Motamayor et al. 2002), cacao trees are now grown in the Americas, West and Central Africa, South East Asia, and Papua New Guinea. In the main production areas, yields are reduced by diseases and pests and these differ in each of the production areas. Yields are also low due to the heterogeneity of planting material; up to 80% of the yield is often obtained from 20% of the trees (Hunter 1990). Self-incompatibility (SI) plays a role in this heterogeneity and contributes to inconsistent and poor bean yields.

Various mechanisms of SI exist in angiosperms to prevent self-fertilization (de Nettancourt 1997) and promote outcrossing to increase genetic variability. These mecha-

nisms operate: in recognition and rejection of pollen, prevention and interruption of pollen germination, pollen tube growth, syngamy, or, as an example of late-acting self-incompatibility (LSI), by arresting embryo development (Sage et al. 1994). In most angiosperms, SI is controlled by a single multi-allelic S-locus encoding a pollen-specific S protein and a corresponding pistil-specific S protein (Takayama and Isogai 2005). In addition to LSI, the two other main mechanisms are gametophytic SI (GSI) and sporophytic SI (SSI). GSI is determined by the male gametophyte (haploid pollen genotype), whereas SSI is controlled by the female sporophyte (diploid maternal genotype). GSI is not only mainly known in Solanaceae, Rosaceae, and Plantaginaceae but also in Papaveraceae (for reviews, see McClure and Franklin-Tong 2006; Takayama and Isogai 2005; see also Wheeler et al. 2009), whereas SSI has been studied mainly in Brassicaceae (reviewed by Takayama and Isogai 2005).

E.J. Pound (1932) and J. Marshall (1933) were first to describe SI in cacao trees (Cope 1962; Knight and Rogers 1955). Trees were designated self-incompatible when less than 10% of the hand-pollinated flowers set fruits (Pound 1933). Cytology has often been used to study SI in cacao (Bouharmont 1960; Cheesman 1927; Cope 1939, 1940, 1958, 1962; Knight and Rogers 1955; Posnette 1940). Shortly after pollination, pollen grains start to germinate and pollen tubes grow through the hollow style towards the ovules (Cheesman 1927). The embryo sacs in the ovules are penetrated via the micropyle and via one of the degenerating synergids, and the two sperm cells are released. One sperm cell moves towards the egg cell and egg nucleus, the other sperm cell moves towards the central cell and the two polar nuclei (Cheesman 1927). Until this point, there are no visible differences between a self-compatible and self-incompatible reaction. In the case of self-compatibility, at approximately 24 to 48 h after pollination, the plasma membranes of one of the sperm nuclei and the egg nucleus fuse and syngamy occurs while the other sperm nucleus fuses with one of the polar nuclei (Bouharmont 1960). This gives rise to a zygote and the triploid primary endosperm. At approximately 72 h after pollination, the triploid endosperm starts its first division. In the case of SI, neither fusion of the plasma membranes nor syngamy occurs and this condition, without fusion, persists until 72 h after pollination after which abscission occurs. In compatibly pollinated flowers, the endosperm continues to divide (Bouharmont 1960; Cheesman 1927; Cope 1939, 1940, 1958, 1962; Knight and Rogers 1955; Posnette 1940). According to Cheesman (1927), in self-compatible reactions the zygote undergoes its first round of division approximately 50 days after pollination; the embryo then grows very slowly, to about 0.5 mm by approximately 90 days post-pollination and to 3.5 mm after 120 days, and

reaches maturity about 130–150 days after pollination. Pollination to mature pods takes, on average, about 6 months (Bouharmont 1960; Falque et al. 1995).

Based on cross-pollination experiments among progenies of three self-incompatible cultivars, Knight and Rogers (1953, 1955) concluded that SI in cacao is genetically under sporophytic control. They proposed that SI is controlled by a single S-locus at which five different alleles, with dominance relationships among them wherein $S_1 > S_2 = S_3 > S_4 > S_5$, could reside. Cytological observations performed by Cope (1958) demonstrated that the SI mechanism is based on the non-fusion of gametes and results in three different categories of non-fusion ovule content: 25%, 50%, or near 100%. Further investigations led to the discovery of a sixth allele which is recessive to the S_5 allele (Cope 1962). Cope (1962) and Bartley and Cope (1973) concluded that the genetic system of SI in cacao has aspects of both sporophytic and gametophytic control. Cope (1962) also concluded that, in order to explain every segregation pattern observed, two additional loci (A and B) exist that encode products that act before meiosis and as accessories to the product(s) of the S-locus.

Abscission of unpollinated cacao flowers 24–48 h after anthesis or of incompatibly pollinated flowers 72–96 h after pollination has also been linked to the physiological effects of changes in phytohormone concentration levels (Baker et al. 1997). Abscisic acid (ABA), ethylene, and the auxin indole-3-acetic acid (IAA) have all been implicated in this process (Baker et al. 1997). These results suggested that the SI reaction or ‘recognition’ happens in two steps; the first occurs soon after pollination and is primarily a response to ABA and the second occurs after contact of the pollen tube with the ovule and is in response to IAA and ethylene (Baker et al. 1997).

The SI mechanism in cacao is not absolute as there are often fusions between gametes containing self-incompatible alleles and resulting fertilized ovules develop into normal seeds. SI can also be manipulated through the use of mentor pollen (Glendinning 1960; Lanaud et al. 1987; Opeke and Jacob 1967; Posnette 1940), or the use of *Herrania* pollen followed by pollination with cacao self-pollen (Bartley 1969, 2005; Opeke and Jacob 1967). Mixing irradiated compatible pollen and normal self-incompatible pollen also results in viable seeds (Adu-Ampomah et al. 1991). Another way to overcome SI is to apply higher CO₂ concentrations around the stigmatic surface right after pollination (Aneja et al. 1994).

Increasing yield, quality, and disease resistance are important objectives for cacao breeding programs; however, seed yield is one of the major constraints caused by SI. Genes regulating SI in cacao have not been previously identified, although a region associated with SI was identified on linkage group 4 by Crouzillat et al. (1996;

2000). Our objective for this study was to use an existing mapping population (Brown et al. 2007) to identify markers linked to SI phenotypes, markers that could be useful for selecting uniformly self-compatible cultivars. In this report, we propose a modified assay for SI and describe its use to identify molecular markers associated with a locus in the cacao genome responsible for the regulation of self-compatibility in cacao.

Material and methods

Segregating mapping population

A segregating mapping population developed by the Cacao Breeding Program at the Tropical Agricultural Research and Higher Education Center (CATIE) in Turrialba, Costa Rica, which consists of 256 offspring from a cross between Pound-7 and UF-273, was used for this study. This population was used previously to construct a genetic map using 177 simple sequence repeats, two resistance gene homologues and one WRKY stress-related marker and QTLs for resistance to frosty pod and black pod and for other horticultural traits were identified (Brown et al. 2007). Briefly, trees were planted in 1998 at CATIE's La Montaña Farm. Five cloned trees of UF-273 were used as male parents, one of which had differing alleles for 22 markers out of the 180 used for mapping and is referred to as UF-273 Type II. The other four male parental trees are referred to as UF-273 Type I. The final genetic map contains 170 markers in ten linkage groups with a total genome length of 884.8 cM (Brown et al. 2007). The population also segregated for self-incompatibility with Pound-7 as the self-incompatible parent and UF-273 as the self-compatible parent.

Pollination procedure and determination of compatibility status

Self-pollinations were carried out between 2008 and 2010. Each tree in the mapping population was evaluated at least four times and was pollinated no less than 20 and at most 40 times. No further pollinations, after a minimum of 20, were performed in those trees where at least 95% of the self-pollinations did not result in flower retention. Three trees did not produce any flowers. Twelve trees produced only 10 to 16 flowers; however, these had a clear flower drop versus retention reaction. Mature flower buds were selected for use as male or female parents, in the early afternoon (from 11:30 a.m. to 1:30 p.m.) on the day preceding pollination, based on their size and initiation of sepal separation. In order to prevent unwanted pollinations (via wind or insects) selected flowers were isolated by

enclosing them in glass tubes covered with gauze at one side and fixed to the stem or branch using modeling clay and a rubber band. The next morning pollinations were carried out between 6 and 11 AM. First, freshly opened flowers were collected as pollen donors and the petals were removed to liberate stamens and anthers. Pollen quality was checked visually; good pollen has a pearly white color. Tubes were then removed from the flowers serving as female parents. Three to five staminodes surrounding the pistil were removed to better access the pistil and one to three stamens of freshly collected pollen donors were rubbed onto the stigma. Afterwards, the tube with the modeling clay and rubber band were replaced and the pollinated flower was labeled with a tag indicating the female parent × male parent and the pollination date. The tube was removed 15 days after pollination and flower drop (the self-incompatible reaction) or retention (the self-compatible reaction) was monitored.

Characterization of the self-compatibility threshold

Traditionally in CATIE and other institutions, trees with flower retention (FR) of 10% or less after 15 days have been classified as self-incompatible, whereas trees with 30% or more flowers after the same span of time were considered self-compatible; for trees with FR between 10% and 30%, pollinations were repeated until the reaction could be clearly defined as self-incompatible or self-compatible. We tested fruit set potential of the parents (Pound-7 and UF-273) after self-pollination by performing 40 self-pollinations and monitoring flower drop/retention every day for 34 days. Based on the fruit set potential of the self-incompatible parent and on the frequency distribution of the trait in the mapping population at 15 days after pollination (DAP), we used 30% FR as the lower threshold for self-compatibility (SC); trees with values lower than 30% were considered to be self-incompatible. Pearson's chi-square test was applied; the null hypothesis was that the distribution did not differ from a 1:1 segregation ratio in the two observed classes.

Identification of marker-trait associations and parental alleles favorable for self-compatibility at 15 days after pollination

SC was treated as a categorical variable with 1=self-compatible and 0=self-incompatible. Because of this binary nature and the fact that the trait, when considered quantitative using the proportion of successful self-pollinations on each tree, was not normally distributed, the Kruskal–Wallis non-parametric method, using MapQTL 5.0 software, was carried out (van Ooijen 2004). Kruskal–Wallis was performed with a SC threshold of 30% but also

with SC thresholds of 1%, 15%, or 50% in order to investigate the potential of the threshold to influence the identification of and confidence levels of the markers that were identified. After identification of significant markers (same markers as used to create the genetic map by Brown et al. 2007) associated with SC/SI, contingency tables were created for each locus to test the frequency of self-incompatible trees versus self-compatible trees for heterogeneity among the parental alleles of each locus. Fisher's exact test was used to calculate the associated chi-square statistics.

Identification of marker-trait associations at 28 days after pollination

Flower retention at 15, 21, and 28 DAP was studied in 144 trees, about half of the mapping population. The number of pollinations per tree across the time course experiment varied from 3 to 30 with a mean of 12. Kruskal–Wallis was performed with a SC threshold of 30% and 50% and identification of and confidence levels of the markers identified were compared within the experiment. Contingency tables were created for each marker locus to test the frequencies of self-incompatible trees versus self-compatible trees for heterogeneity among the parental alleles of each locus. Fisher's exact test was used to calculate the associated chi-square statistics.

Results

SC/SI thresholds in the parents

Flower drop as end result of SI in the self-incompatible parent, Pound-7, only became obvious 14 to 15 DAP and even more evident at 21 DAP. Percentages of FR dropped from 93% at 96 h after pollination to 27% at 15 DAP and 13% at 21 DAP. After 21 days, the percentage of FR stabilized around 6–7%. However, during the first 7 DAP, Pound-7 resembled a self-compatible tree, with more than 40% FR (Fig. 1). In the self-compatible parent, UF-273, FR went from 97.5% at 4 DAP to 65% at 14 DAP, 62.5% at 22 DAP, and stabilized at 57.5% at 29 DAP (Fig. 1).

Determination of self-compatibility thresholds in the mapping population

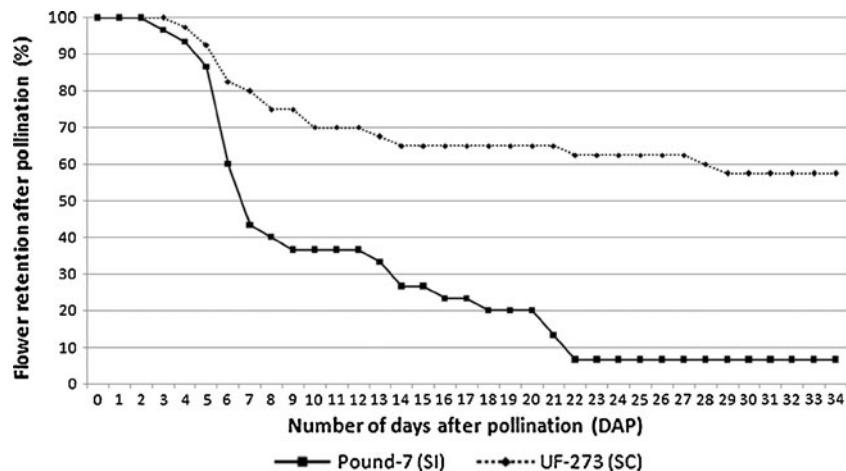
SC/SI was measured in 251 of the 256 trees in the mapping population at 15 DAP. Of the original 256 trees, three trees did not produce flowers, one died, and one tree was identified as an off-type. While FR at 15 DAP in the self-incompatible parent was approximately 27% (Fig. 1) and the frequency distribution of the trait demonstrated a clear

delineation around 30% (Fig. 2), the level of SC/SI varied greatly in the segregating population and the frequency distribution of FR at 15 DAP clearly showed a bimodal distribution (Fig. 2). Three different groups were discerned. The first group contained 102 trees (40.6%) and had less than 10% FR at 15 DAP. In this group, 67 trees had no FR at all. The second group, with FR between 10% and 29%, contained 15 trees (6%). The third group, with FR between 30% and 80%, contained 134 trees (53.4%) (Fig. 2). Therefore, at the 30% threshold of FR, 134 trees were SC and 117 trees were SI (Fig. 2), which does not significantly differ from a 1:1 segregation (two-tailed $P=0.3437$, Pearson's chi-square test).

Identification of marker-trait associations for self-incompatibility and favorable allele composition

Results of the Kruskal–Wallis non-parametric method are presented in Table 1 which also includes the K^* (Chi-square) and P values (significance levels) for those marker loci that are significantly associated with SC/SI for different percentages of SC thresholds. Figure 3 shows six of the ten linkage groups (LG) of the cacao genetic map and the markers associated with SC. Microsatellite markers mTcCIR168, mTcCIR222, mTcCIR115, and mTcCIR158, located near the proximal end of LG4, were highly significantly associated with self-compatibility for SC thresholds of 15% to 50%. The highest association was for mTcCIR222 and, therefore, that marker is likely to be closest to the actual gene; it is located at 6.7 centimorgans (cM) near the proximal end of LG4 and has a K^* value ranging from 28.9 (for a 1% SC threshold) to 56.1 (for a 50% SC threshold) and a significance level of $P<0.0001$ for each SC threshold. Thirty-two per cent of the trees containing this marker retained no self-pollinated flowers. The other three markers had significance levels ranging between $P<0.1$ (1% SC threshold) and $P<0.0001$ (50% SC threshold). All four markers were located within 8.5 cM (Table 1; Fig. 3). Four other markers (mTcCIR18, mTcCIR107, mTcCIR213, and mTcCIR206), also located on LG4 (between 34.41 and 44.817 cM), were significantly associated with SC ($P<0.005$) but only at the 50% SC threshold. An additional marker, mTcCIR17, was also significantly associated with SC ($P<0.005$) but only at the 50% SC threshold. Marker mTcCIR17 could not be assigned to any linkage group in this mapping population but was positioned in the same interval on LG4 on other genetic maps (Brown et al. 2008). Two other markers, mTcCIR146 (LG3) and mTcCIR124 (LG9), were associated with SC at the 15% ($P<0.005$) and 50% SC thresholds ($P<0.01$), respectively (Table 1). Marker mTcCIR222 had the heterozygous 218/220 allelic configuration in the self-incompatible parent and the heterozy-

Fig. 1 Time course of self-pollination of parents, presenting the percentage of flower retention observed during the 34 days following pollination



gous 216/220 allelic configuration in the self-compatible parent (Table 2). Fisher's exact test, used to calculate the associated chi-square statistics, indicated that the heterozygous 216/218 combination was the most favorable allelic configuration of marker mTcCIR222 for self-compatible trees (two-tailed $P < 0.0001$) and was therefore positively associated with SC (Table 2). Of the 63 trees with the 216/218 combination in the mapping population, only seven trees (11%) were self-incompatible, the other 56 trees were self-compatible (Table 2). Of those seven trees, two trees had an FR of 20%, four trees had an FR of less than 10%, and one tree retained no single flowers at all. Fifty-three trees (84%) had an FR of 50% or more and three trees (5%) had an FR between 30% and 50% (Electronic supplementary material (ESM 1)). The other three allelic configurations for mTcCIR222 were not significantly associated with SI. A summary of the other marker loci that were associated with SI, the allelic configurations of those markers in the parents of the mapping population, their possible allelic configurations in the mapping population, and the

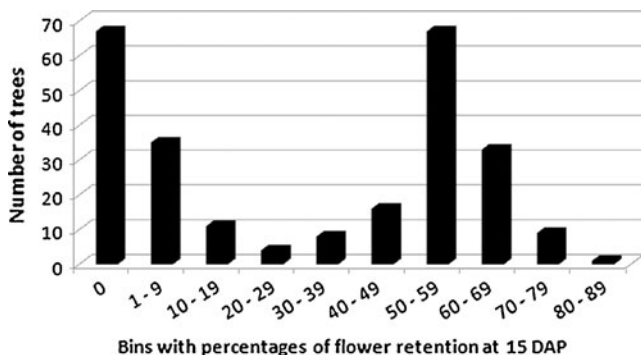


Fig. 2 Frequency distribution of the number of trees against bins representing ranges of percentage of flower retention on trees of the whole mapping population at 15 days after pollination

favorable allele combinations and their significance at the 30% SC threshold is presented in Table 2.

Identification of marker-trait associations at 28 days after pollination

To determine whether the SC threshold remained unchanged over time, different measurements of the trait were taken over a time course and compared with the same time course measurements of the self-incompatible parent. The observation period was extended from 15 DAP to 21 and 28 DAP. At 15 DAP, 43.1% of the trees in the mapping population were self-incompatible, whereas the percentage of self-incompatible trees increased to 47.2% at 21 DAP and to 51.4% at 28 DAP (ESM 1). Despite the fact that only about half of the mapping population was used for this time course experiment and there were a limited number of pollinations per tree (from 3 to 30, with a mean of 12), the percentage of self-incompatible trees in this experiment was similar to the percentage identified in our non-time course experiment, 43% versus 46%, respectively.

The Kruskal–Wallis test was performed on our time course results using SC thresholds of 30% and 50% (Table 3). The most significant markers identified were again mTcCIR168, mTcCIR222, mTcCIR115, and mTcCIR158, located near the proximal end of LG4 but only mTcCIR222 remained highly significant ($P < 0.0001$) at the 30% SC threshold at 28 DAP and mTcCIR222 became much less significant ($P < 0.05$) at the 50% SC threshold at 28 DAP. Marker mTcCIR249, which mapped to LG1, and markers SHRS13 and mTcCIR176, which mapped to LG2, only became significantly associated with SC ($P < 0.01$) at the 50% SC threshold at 28 DAP. Marker mTcCIR229, mapped to LG10, was significant ($P < 0.01$) at 15 DAP but not at 28 DAP (Table 3). The heterozygous 216/218 combination was still the most favorable allelic configuration of marker

Table 1 Marker loci significantly associated with self-compatibility for different self-compatibility threshold percentages for the whole mapping population at 15 days after pollination

LG	Position	Locus	1% SC		15% SC		30% SC		50% SC	
			<i>K</i> *	Significance	<i>K</i> *	Significance	<i>K</i> *	Significance	<i>K</i> *	Significance
3	68.533	mTcCIR146	3.952	–	12.966	****	8.596	**	2.499	–
4	0	<i>mTcCIR168</i>	3.382	*	11.885	*****	12.86	*****	21.16	*****
4	6.683	<i>mTcCIR222</i>	28.919	*****	43.016	*****	42.514	*****	56.128	*****
4	6.931	<i>mTcCIR115</i>	3.788	*	13.429	*****	13.112	*****	16.71	*****
4	8.512	<i>mTcCIR158</i>	3.788	*	13.429	*****	13.112	*****	18.851	*****
4	34.41	mTcCIR18	0.919	–	3.427	*	5.478	**	9.627	****
4	36.4	mTcCIR107	1.071	–	2.915	*	4.865	**	8.949	****
4	44.817	mTcCIR213	3.84	–	6.083	–	7.005	*	15.147	****
4	44.817	mTcCIR206	3.84	–	6.083	–	7.005	*	15.147	****
9	54.579	mTcCIR124	1.797	–	4.341	–	4.951	–	11.353	***
U	ND	mTcCIR17	1.192	–	2.659	–	4.433	**	9.437	****

Loci which are highly significantly associated with self-compatibility are set in italics

(–) not significant, *U* unassigned to any LG, *ND* not determined

*0.1; **0.05; ***0.01; ****0.005; *****0.001; *****0.0005; *****0.0001

mTcCIR222 for self-compatible trees (two-tailed $P < 0.0001$) using the 30% and 50% SC thresholds at 15 DAP and the 30% SC threshold at 28 DAP, but became less significant using the 50% SC threshold at 28 DAP (two-tailed $P < 0.01$); there was no significance associated with the other three allelic configurations (data not shown).

Discussion

Levels of self-compatibility/incompatibility in parents

We determined the levels of SC/SI in the parents and in the mapping population developed at CATIE via hand pollina-

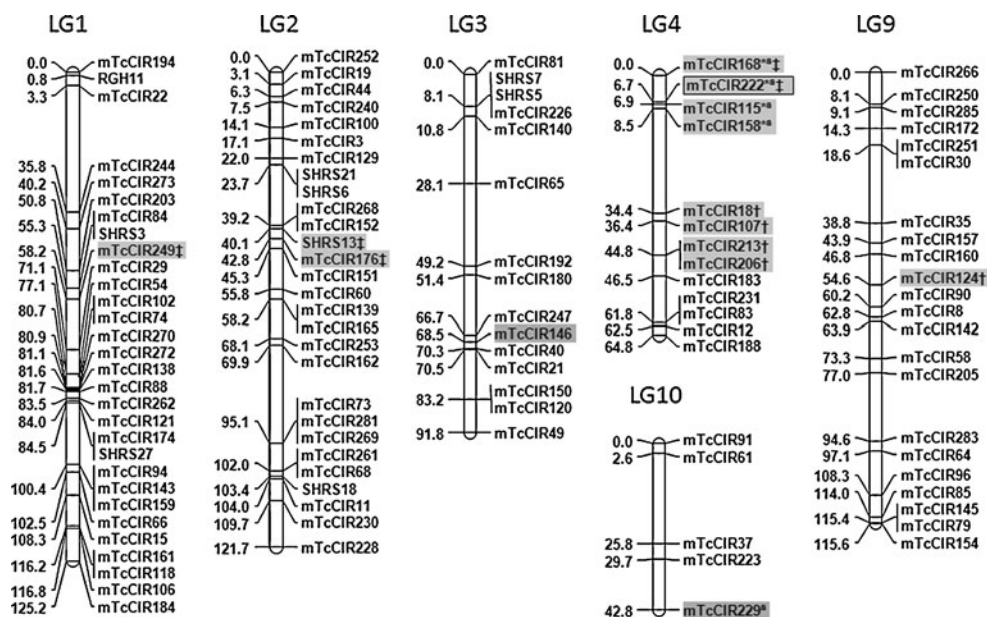


Fig. 3 Genetic linkage groups and indication of markers associated with self-compatibility, with mTcCIR222 (boxed) being the most consistent and most significantly associated marker. Note: Asterisk, highly significant association with self-compatibility at 15 DAP for the 30% and 50% self-compatibility threshold for the whole mapping population; dagger, significant associations at 15 DAP, only for the

50% self-compatibility threshold for the whole population; super-scripted letter *a*, significant associations only at 15 DAP for the 30% threshold for the subset of 144 trees; double dagger, significant associations only at 28 DAP for the 50% threshold for the subset of 144 trees. Marker mTcCIR146 is only significant at 15 DAP for the 15% self-compatibility threshold

Table 2 Contingency table at 30% self-compatibility threshold with identification of favorable allele combinations for different marker loci associated with self-compatibility in the whole mapping population at 15 days after pollination

Locus	LG	Parental alleles (bp)		Allelic combinations offspring		Contingency table at 30% SC threshold				
						# SI trees		# SC trees		2-tailed <i>P</i> value
						Comb 1	Comb 2	Comb 1	Comb 2	
		Pound-7	UF-273	Comb 1 vs. Comb 2						
mTcCIR168	4	176/176	176/180	176/176	176/180	39	78	75	59	<i>0.00037</i>
mTcCIR222	4	218/220	216/220	216/218	216/220	7	36	56	24	<i>1.46E-08</i>
				216/218	218/220	7	38	56	28	<i>1.81E-08</i>
				216/218	220/220	7	36	56	26	<i>2.29E-08</i>
				216/220	218/220	36	38	24	28	0.85689
				216/220	220/220	36	36	24	26	0.85567
				218/220	220/220	38	36	28	26	1.00000
mTcCIR115	4	192/192	190/192	190/192	192/192	74	43	54	80	<i>0.00038</i>
mTcCIR158	4	214/214	214/218	214/214	214/218	43	74	80	54	<i>0.00038</i>

Favorable alleles are indicated in bold. The *P* values of the allelic combinations significantly associated with self-compatibility are set in italics

tion of approximately 40 flowers per tree and measurements of the SC/SI reaction were performed at 15 DAP. Contrary to the general belief that the SI reaction in cacao is expressed 96 h after pollination, the self-incompatible reaction developed more slowly in the Pound-7 parent in this study. For Pound-7 at that time point, 93% of the flowers were still attached to the tree. Percentages of FR dropped to 43% at 7 DAP, 27% at 14 DAP, and 13% at 21 DAP, after which it stabilized at 6–7% (Fig. 1). Pollination experiments performed at the Subtropical Horticulture Research Station (SHRS) in Miami and at

the Mars Center for Cocoa Science (MCCS) in Brazil with other self-incompatible trees also showed the quick drop in FR at 96 h after pollination (0% FR for cultivars EET-59, Gainesville-10, and IMC-105 after 4 to 7 days, unpublished data). In the particular case of cultivar TSH-1188, 96 h after pollination the percentage of FR dropped to 14%, to 2% at 15 DAP, and 0% at 16 DAP (unpublished data). FR in the self-compatible parent, UF-273, went from approximately 80% at 7 DAP to 65% at 14 DAP, then dropped to 62.5% at 22 DAP and finally stabilized at 57.5% at 29 DAP (Fig. 1).

Table 3 Time course *K** and *P* values for marker loci significantly associated with self-compatibility for different self-compatibility threshold percentages

LG	Position	Locus	15 DAP–30% SC		15 DAP–50% SC		28 DAP–30% SC		28 DAP–50% SC	
			<i>K*</i>	Significance	<i>K*</i>	Significance	<i>K*</i>	Significance	<i>K*</i>	Significance
1	54.904	mTcCIR249	5.353	–	4.508	–	5.764	–	11.464	***
2	40.061	SHRS13	3.049	–	5.236	–	4.808	–	11.364	***
2	42.81	mTcCIR176	3.189	–	5.661	–	5.46	–	12.688	***
4	0	<i>mTcCIR168^a</i>	13.282	*****	14.571	*****	13.439	*****	6.859	***
4	6.683	<i>mTcCIR222^b</i>	27.117	*****	26.971	*****	25.189	*****	15.872	****
4	6.931	<i>mTcCIR115^c</i>	10.405	****	12.155	*****	11.306	*****	5.897	**
4	8.512	<i>mTcCIR158^d</i>	9.471	****	11.013	*****	10.161	****	6.556	**
10	42.819	mTcCIR229	7.733	***	6.273	**	5.342	**	4.282	**

Loci which are highly significantly associated with self-compatibility are set in italics

(–) not significant

0.05; *0.01; ****0.005; *****0.001; *****0.0005; *****0.0001

^aAJ566497 (GenBank accession number)

^bAJ566543 (GenBank accession number)

^cAJ566457 (GenBank accession number)

^dAJ566489 (GenBank accession number)

Determination of self-compatibility threshold in the mapping population

Trees with 30% FR or more at 15 DAP were considered self-compatible, based on the level of SI in Pound-7 at 15 DAP, i.e., 27% (Fig. 1), and on the frequency distribution of FR in the mapping population (Fig. 2). However, historically, various approaches with differing numbers of pollinated flowers, SC thresholds and number of days after pollination have been taken to determine levels of self-incompatibility (Knight and Rogers 1955; Posnette 1945; Pound 1932, 1933; Voelcker 1936, 1937). More recently, Lopes and Yamada (2000) tried to identify the optimal number of flowers to be pollinated and the criterion to distinguish between self-compatible and self-incompatible clones. They used two different methods. The first one required self-pollination of at least 26 flowers and cross-pollination of at least 26 flowers of the same tree with compatible pollen; using this method, trees were considered self-compatible if FR was between 5% and 30% at 15 DAP. For the second method, only self-pollinations were performed and trees with 15–17% FR were classified as self-compatible.

We found that as the level of FR in the parents decreased over time (Fig. 1), the timing of the observations and the threshold for detecting SC in the mapping population members had to be reconsidered. The observation period was extended from 15 DAP to 21 and 28 DAP. At 15 DAP, 43.1% of the trees had FR of less than 30% and therefore were considered self-incompatible. At 21 DAP, 47.2% of the trees were self-incompatible whereas at 28 DAP, 51.4% of the trees were self-incompatible (ESM 1). This study suggests that 15 days after pollination or later is a reasonable time to measure FR in the parents and offspring of this mapping population as there are important differences between the percentages of flowers retained between 96 h and 15, 21, and 28 DAP. Our results also suggest, in agreement with previous studies (Lopes and Yamada 2000 and references therein), that selecting thresholds for self-compatibility is genotype specific. We also note that the amount of variation in FR that occurs after self-incompatible pollination has not been characterized well. For all of these reasons, we conclude that more data on the various factors affecting FR in cacao, including cytological, hormonal and other developmental information, especially between 96 h and 50 days after pollination and with emphasis on between 15 and 28 DAP, are necessary before a reasonable hypothesis for the mechanism(s) underlying our observations can be established. Studies of the 15 trees in which we found FR ranged from 10% to 29% might be particularly enlightening. In the meantime, we recommend that SC thresholds should be examined and set genotype specifically in cacao.

Identification of marker-trait associations for self-compatibility

We detected a particularly significant association between one molecular marker and the SC/SI trait using the genetic map of Brown et al. (2007). The results obtained using different SC thresholds allowed us to identify a single microsatellite marker, mTcCIR222, associated with SC ($P < 0.0001$) and located 6.7 cM from the proximal end of LG4 (Table 1; Fig. 3). Three other microsatellite markers, mTcCIR168, mTcCIR115, and mTcCIR158, near mTcCIR222, also became significantly associated with SC. Thus, we have identified, in agreement with the 1:1 segregation ratio we observed for SC, a single locus that is strongly associated with this trait. We note, however, that at the 50% SC threshold the segregation ratio started to deviate significantly from 1:1 ($P < 0.0001$) (Fig. 2, Table 1). This deviation could indicate that mechanisms other than SC/SI (such as cherelle wilt, see below) might be involved in flower drop/retention and that some of these additional markers may not be involved in SC. This discrepancy should be investigated further.

Warren et al. (1995), using isozyme markers, identified two markers, *Acp* and *Idh*, which were associated with SC in a segregating population (72 trees) of an ICS cacao clone. That population also displayed a 1:1 segregation for SC. Cruzillat et al. (1996) generated a map based on the analysis of 131 BC₁ trees and another map based on 55 F₁ trees (Cruzillat et al. 2000), all derived from a cross between Catongo (self-compatible) and Pound-12 (self-incompatible). Two of their markers, i.e., self-compatibility (*Autoc*) and anthocyanidin synthesis (*Anth*), were linked with SC and positioned at what was referred to as the distal end of linkage group 5, which is the same as the proximal end of linkage group 4 in the consensus map we are using (Brown et al. 2007; Fritz et al. 1995; Pugh et al. 2004). Lanaud et al. (2009) performed a meta-QTL analysis using 16 different QTL and association studies and positioned the *Autoc* locus at 1.9 cM from the *Idh* locus and 4.6 cM from the mTcCIR222 locus. These studies all corroborate our findings that mTcCIR222 is tightly linked to the SC/SI trait.

Favorable allele combination for self-compatibility

Marker mTcCIR222 was strongly associated with SC at 15 DAP. Fifty-six of the 134 self-compatible trees (42%) at the 30% threshold had the heterozygous 216/218 allelic genotype (Table 2). Over the whole population (251 trees), the average percentage of FR for this allelic configuration was 54%, whereas for the other combinations the percentages were 23%, 24%, and 23% for the 216/220, 218/220, and 220/220 allelic configurations, respectively (data not

shown). For the time course experiment (144 trees), more than 45% of the self-compatible trees had the 216/218 allelic configuration (ESM 1). Marker mTcCIR222 and its 216/218 allelic configuration, together with the favorable allelic configuration of 176/176 for mTcCIR168, 192/192 for mTcCIR115, and 214/214 for mTcCIR158 (all of these alleles are positively associated with SC) are currently being tested as candidates for selecting self-compatibility in our marker-assisted selection program (Schnell et al. 2007).

The fact that other markers appeared to become significantly associated with SC at 28 DAP, and more specifically at the 50% SC threshold when the segregation was significantly different from 1:1 (two-tailed $P < 0.05$), suggests to us that the genes linked to these markers may be involved in other mechanisms such as cherelle wilt. It is postulated that the young pods, or cherelles, wilt because of a lack of hormones produced by the endosperm, which causes a decrease in water and food uptake. This phenomenon of cherelle wilt may begin at fertilization and last until approximately 80 days after pollination (McKelvie 1956).

Concluding remarks

An important step towards the selection of uniformly self-compatible cultivars is the identification of genomic regions involved in SC/SI. We have identified a single genomic region and a marker that is strongly associated with the SC/SI trait. Further research, as part of the cacao genome sequencing project (www.cacaogenomedb.org), will involve the fine mapping of the trait in another existing mapping population, additional and combined genetic and physical mapping, and gene expression studies in our effort to identify those genes that are the key players in the SI process. Comparisons between the sequence data available and known genes involved in the major SI mechanisms in other plant species (for reviews, see McClure and Franklin-Tong (2006) and Takayama and Isogai (2005)), however, have not yet led to the identification of significant homology, which may indicate the uniqueness of the SI system in cacao. Additional cacao genome sequencing information should allow the cacao scientific community to make more rapid progress in breeding efforts to produce cultivars that are self-compatible and have increased yields, improved quality, and disease resistance as well.

Acknowledgments We would like to thank Tony Alfaro and Allan Mata Quiros at CATIE. Thanks are due to Cecile Tondo and Belinda Martineau for helpful comments on the manuscript. We also would like to thank Mars, Inc. and their funding of Trust Agreement # 58-6631-6-123 that in part supported this research.

References

- Adu-Ampomah Y, Novak FJ, Klu GYP, Lamptey TVO (1991) Use of irradiated pollen as mentor pollen to induce self-fertilization of two self-incompatible Upper Amazon cacao clones. *Euphytica* 51:219–225
- Alverson WS, Whitlock BA, Nyffeler R, Bayer C, Baum DA (1999) Phylogeny of the core Malvales: evidence from *ndhF* sequence data. *Am J Bot* 86:1474–1486
- Aneja M, Gianfagna T, Ng E, Badilla I (1994) Carbon dioxide treatment partially overcomes self-incompatibility in a cacao genotype. *Hortscience* 29:15–17
- Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto JF, Sabot F, Kudrna D, Ammiraju JSS, Schuster SC, Carlson JE, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Berard A, Viot C, Boccara M, Risterucci AM, Guignon V, Sabau X, Axtell MJ, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahí M, Akaza JM, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie WR, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C (2011) The genome of *Theobroma cacao*. *Nat Genet* 43:101–108
- Baker RP, Hasenstein KH, Zavada MS (1997) Hormonal changes after compatible and incompatible pollination in *Theobroma cacao* L. *Hortscience* 32:1231–1234
- Bartley BGD (1969) Selfing of self-incompatible trees. Annual Report on Cacao Research (1968) Trinidad, pp 22–23
- Bartley BGD (2005) The genetic diversity of cacao and its utilization. CABI Publishing, Wallingford, p 341
- Bartley BGD, Cope FW (1973) Practical aspects of self-incompatibility in *Theobroma cacao* L. In: Moav R (ed) Agricultural genetics—selected topics. Wiley, New York, pp 109–134
- Bouharmont J (1960) Recherches cytologiques sur la fructification et l'incompatibilité chez *Theobroma cacao* L. Publication de l'Institut National pour l'Etude Agronomique du Congo. Série scientifique no 89
- Brown JS, Phillips-Mora W, Power EJ, Krol C, Cervantes-Martinez C, Motamayor JC, Schnell RJ (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in *Theobroma cacao* L. *Crop Sci* 47:1851–1858
- Brown J, Sautter R, Olano C, Borrone J, Kuhn D, Motamayor J, Schnell R (2008) A composite linkage Map from three crosses between commercial clones of cacao, *Theobroma cacao* L. *Trop Plant Biol* 1:120–130
- Cheesman EE (1927) Fertilization and embryogeny in *Theobroma cacao* L. *Ann Bot* 41:107–126
- Cope FW (1939) Studies in the mechanism of self-incompatibility in cacao I. 8th Annual Report on Cacao Research (1938), Trinidad, pp 20–21
- Cope FW (1940) Studies in the mechanism of self-incompatibility in cacao II. 9th Annual Report on Cacao Research (1939) Trinidad, pp 19–23
- Cope FW (1958) Incompatibility in *Theobroma cacao*. *Nature* 181:279–279
- Cope FW (1962) The mechanism of pollen incompatibility in *Theobroma cacao* L. *Heredity* 17:157–182
- Crouzillat D, Lerceteau E, Petiard V, Morera J, Rodriguez H, Walker D, Phillips W, Ronning C, Schnell R, Osei J, Fritz P (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. *Theor Appl Genet* 93:205–214
- Crouzillat D, Ménard B, Mora A, Phillips W, Pétiard V (2000) Quantitative trait analysis in *Theobroma cacao* using molecular markers. *Euphytica* 114:13–23

- Dantas LG, Guerra M (2010) Chromatin differentiation between *Theobroma cacao* L. and *T. grandiflorum* Schum. *Genet Mol Biol* 33:94–98
- de Nettancourt D (1997) Incompatibility in angiosperms. *Sex Plant Reprod* 10:185–199
- Falque M, Vincent A, Vaissiere B, Eskes A (1995) Effect of pollination intensity on fruit and seed set in cacao (*Theobroma cacao* L.). *Sex Plant Reprod* 8:354–360
- Fritz PJ, Phillips-Mora W, Rodriguez H (1995) New tools for 21st century plant breeding. DNA markers: theory and applications. Technical Series Technical Bulletin/CATIE No 251:122 p
- Glendinning DR (1960) Selfing of self-incompatible cocoa. *Nature* 187:170–170
- Hunter J (1990) The status of cacao (*Theobroma cacao*, Sterculiaceae) in the western hemisphere. *Econ Bot* 44:425–439
- Knight R, Rogers HH (1953) Sterility in *Theobroma cacao* L. *Nature* 172:164–164
- Knight R, Rogers HH (1955) Incompatibility in *Theobroma cacao*. *Heredity* 9:69–77
- Lanaud C, Sounigo O, Amefia YK, Paulin D, Lachenaud P, Clément D (1987) Nouvelles données sur le fonctionnement du système d'incompatibilité du cacaoyer et ses conséquences pour la sélection. *Café, Cacao, Thé* 31:267–277
- Lanaud C, Fouet O, Clément D, Boccara M, Risterucci A, Surujdeo-Maharaj S, Legavre T, Argout X (2009) A meta-QTL analysis of disease resistance traits of *Theobroma cacao* L. *Mol Breed* 24:361–374
- Lopes UV, Yamada MM (2000) Sexual incompatibility in cacao (*Theobroma cacao* L.). I. Sample size and approaches to characterize germplasm. In: 13th International Cocoa Research Conference, Kota Kinabalu, Malaysia
- Marshall J (1933) Fertility in cacao. 3rd Annual Report on Cacao Research (Trinidad), p 34
- McClure B, Franklin-Tong V (2006) Gametophytic self-incompatibility: understanding the cellular mechanisms involved in “self” pollen tube inhibition. *Planta* 224:233–245
- McKelvie AD (1956) Cherville wilt of cacao: I. Pod development and its relation to wilt. *J Exp Bot* 7:252–263
- Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, Lanaud C (2002) Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* 89:380–386
- Motamayor JC, Lachenaud P, da Silva e Mota JW, Loor R, Kuhn DN, Brown JS, Schnell RJ (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L.). *PLoS ONE* 3:e3311
- Opeke LK, Jacob VJ (1967) Studies on methods of overcoming self-incompatibility in *Theobroma cacao* Linn. 2e Conférence Internationale sur les Recherches Cacaoyères, pp 356–359
- Posnette AF (1940) Self-incompatibility in cocoa (*Theobroma* spp.). *Trop Agric* 17:67–71
- Posnette AF (1945) Incompatibility in Amazon cacao. *Trop Agric* 22:184–187
- Pound FJ (1932) Studies on fruitfulness in cacao, I and II. First Annual Report on Cacao Research, 1931, Trinidad, pp 24–28
- Pound FJ (1933) Studies on fruitfulness in cacao III. Second Annual Report on Cacao Research (1932), Trinidad
- Pugh T, Fouet O, Risterucci AM, Brottier P, Abouladze M, Deletrez C, Courtois B, Clément D, Larmande P, N’Goran JAK, Lanaud C (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theor Appl Genet* 108:1151–1161
- Sage TL, Bertin RI, Williams EG (1994) Ovarian and other late-acting self-incompatibility systems. In: Williams EG, Clarke AE, Knox RB (eds) Genetic control of self-incompatibility and reproductive development in flowering plants. Kluwer, Norwell
- Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W, Amores FM, Motamayor JC (2007) Development of a marker assisted selection program for cacao. *Phytopathology* 97:1664–1669
- Takayama S, Isogai A (2005) Self-incompatibility in plants. *Annu Rev Plant Biol* 56:467–489
- van Ooijen JW (2004) MapQTL®5. Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V, Wageningen
- Vœlcker OJ (1936) Self-incompatibility in cacao. 6th Annual Report on Cacao Research (Trinidad), pp 2–5
- Vœlcker OJ (1937) Self-compatibility in cacao II. 7th Annual Report on Cacao Research (Trinidad), pp 2–5
- Warren J, Misir S, Kalai (1995) Isozyme markers for self-compatibility and yield in *Theobroma cacao* (cacao). *Heredity* 74:354–356
- Wheeler MJ, de Graaf BHH, Hadjosif N, Perry RM, Poulter NS, Osman K, Vátovec S, Harper A, Franklin FCH, Franklin-Tong VE (2009) Identification of the pollen self-incompatibility determinant in *Papaver rhoeas*. *Nature* 459:992–995