

Resistance Gene Mapping for Witches' Broom Disease in *Theobroma cacao* L. in an F₂ Population using SSR Markers and Candidate Genes

J. Steven Brown and R.J. Schnell

Subtropical Horticulture Research Station, U.S. Department of Agriculture, Agricultural Research Service, 13601 Old Cutler Road, Miami, FL 33158

J.C. Motamayor

Masterfoods USA (Mars, Inc.), c/o U.S. Department of Agriculture, Agricultural Research Service, 13601 Old Cutler Road, Miami, FL 33158

Wilson Lopes

CEPEC/CEPLAC, Ilheus, Bahia, Brazil

David N. Kuhn

Department of Biological Sciences, Florida International University, Miami, FL 33199

James W. Borrone

Subtropical Horticulture Research Station, U.S. Department of Agriculture, Agricultural Research Service, 13601 Old Cutler Road, Miami, FL 33158

ADDITIONAL INDEX WORDS. QTL mapping, witches' broom disease, resistance mapping, genomic maps, *T. cacao*

ABSTRACT. A genetic linkage map was created from 146 cacao trees (*Theobroma cacao*), using an F₂ population produced by selfing an F₁ progeny of the cross Sca6 and ICS1. Simple sequence repeat (SSR) markers (170) were used principally for this map, with 12 candidate genes [eight resistance gene homologues (RGH) and four stress related WRKY genes], for a total of 182 markers. Joinmap software was used to create the map, and 10 linkage groups were clearly obtained, corresponding to the 10 known chromosomes of cacao. Our map encompassed 671.9 cM, approximately 100 cM less than most previously reported cacao maps, and 213.5 cM less than the one reported high-density map. Approximately 27% of the markers showed significant segregation distortion, mapping together in six genomic areas, four of which also showed distortion in other cacao maps. Two quantitative trait loci (QTL) for resistance to witches' broom disease were found, one producing a major effect and one a minor effect, both showing important dominance effects. One QTL for trunk diameter was found at a point 10.2 cM away from the stronger resistance gene. One RGH flanked the minor QTL for witches' broom resistance, implying possible association. QTLs mapped in F₂ populations produce estimates of additive and dominance effects, not obtainable in F₁ crosses. As dominance was clearly shown in the QTL found in this study, this population merits further study for evaluation of dominance effects for other traits. This F₂ cacao population constitutes a useful link for genomic studies between cacao and cotton, its only widely grown agronomic relative.

Cacao is a perennial tree crop grown in the tropics, native to South America. It is an economically important cash generator, especially for small landholders, and is well adapted to monoculture and to sustainable mixed tree cropping systems. The beans are used in the production of chocolate and cosmetics, and recent studies have found that chocolate may provide benefits to cardiovascular health through a variety of mechanisms (Steinberg et al., 2003). Interest has increased during recent decades in the production of specialty chocolates, produced by genotypes indigenous only to certain areas in South America (Fowler, 1994). While South America and Africa have been important production areas during the last two centuries, an outbreak of witches' broom disease in South America and frosty pod disease in Central America has shifted production such that Africa now supplies approximately 70% of world production, with approximately 20% coming from South America and 10% from Asia. Witches' broom has devastated production in the state of Bahia, Brazil, and has sig-

nificantly diminished production in Ecuador and other areas of South America (Yamada and Lopes, 1999).

Witches' broom disease is caused by the pathogen *Crinipellis pernicioso* (Stahel) Singer, which infects vegetative meristematic tissues, flower cushions, and young pods, and reduces yields by as much as 80%. Tree death can occur after successive cycles of infection when accompanied by abiotic stresses (Andebrhan, 1984). As chemical control and sanitation are limited in their efficacy in controlling this disease, resistance breeding is important. Sca6 has been used as a donor parent for resistance to witches' broom since the early 1950s (Johnson et al., 2003), and has been identified along with a sibling, Sca12, as being the most likely sources for apparent resistance in farmer selections within the state of Bahia (Yamada and Lopes, 1999). Field breeding experience has suggested that resistance to *C. pernicioso* coming from the Scavina sources is controlled by few genes with major effects (Ahnert and Pires, 2000), and the recent QTL mapping results for witches' broom resistance of Quieroz et al. (2003) support this. However, the extremely high LOD (log of the odds, referring to log likelihood) level required for significance in their work (34.96)

Received for publication 17 Aug. 2004. Accepted for publication 7 Nov. 2004.

does not preclude the possibility that genes with smaller effects could be involved. Resistance breeding for both witches' broom and frosty pod diseases, and searching for all possible sources of resistance are obvious priorities, among other objectives, for the cooperative cacao breeding program of the Subtropical Horticulture Research Station (SHRS) of the U.S. Dept. of Agriculture, Agricultural Research Service, in Miami, as described in an overall project description (Schnell et al., 2004).

Cacao has a low chromosome number ($n = x = 10$), and a small genome (0.4 pg/1 C) (Lanaud et al., 1992), about 2.8 times the size of that of *Arabidopsis thaliana* (L.) Heynh. (Couch et al., 1993). Lanaud et al. (1995) published the first linkage map of cacao resulting from an F_1 cross of UPA402, an Amazonian Forastero clone, produced from sib-mating two Ecuadorian Forastero clones, IMC60 and Na 34, crossed with a Trinitario selection, UF676, from Costa Rica. The initial map from 100 individual trees was developed from data on five isozyme loci, four functional genes, 55 RFLP produced from genomic DNA, and 28 RAPDs, and produced the first genomic map of cacao with 10 linkage groups, corresponding to the number of chromosome pairs in cacao. This cross was further saturated by Risterucci et al. (2000) with more types of markers and higher numbers of them (424 total markers), and 81 additional trees, to produce the first high-density linkage map of cacao. The recombination-based genome length from the map of Lanaud et al. (1995) was 759 cM, and the high-density map of Risterucci et al. (2000) covered 885.4 cM, with 10 linkage groups being clearly distinguishable in both maps, and marker alignment in good agreement. The first map was used also for mapping QTLs for resistance to black pod disease (Lanaud et al., 1999a). A more recent map has been reported from this population by Pugh et al. (2004), using 135 progeny and codominant markers from past maps, and including 201 new SSR markers and 16 resistance gene analog-based markers (Lanaud et al., 2004) for a total of 465 markers. Total genome length of this map was 782.8 cM, and the slight reduction in distance compared to the high-density map was attributed to the fact that AFLP markers were known to target specific AT-rich, heterochromatic sites located near the centromeres and telomeres (Pugh et al., 2004). Otherwise, colinearity was very high between the two maps, and genetic distances between loci common to both maps were generally of the same magnitude.

QTLs for yield and yield-related agronomic traits in cacao and resistance to black pod disease caused by *Phytophthora palmivora* (Butl.) Butl. have been reported by Crouzillat et al. (2000a, 2000b) and Flament et al. (2001). QTL were identified that accounted for moderate (17%) to relatively high levels (48%) of resistance to black pod disease. However, Flament et al. (2001) found no QTL in common across the three different measurements for resistance, and found results from artificial inoculation data to be poorly related to results based on field resistance, as did Lanaud et al. (1999a). Clement et al. (2003a, 2003b) used three test-cross populations of two Trinitario clones (DR1 and S52) and one upper Amazon clone (IMC78) crossed onto one homozygous lower Amazon parent (Catongo) for map creation and QTL colocalization for yield and yield-related traits and for black pod resistance. This effort was successful in mapping several QTLs for yield and yield-related traits with good reliability and commonality among correlated traits. Resistance QTLs to black pod were found in DR1 and IMC78 in a similar region of chromosome 4; however, possible interactions with morphological characteristics of the pods may have caused these apparent resistance QTL, as suggested by Flament et al. (2001).

The purpose of this investigation was to create a map of *T. cacao* containing codominant markers from an F_2 population segregating for resistance to witches' broom disease. As several years' data existed for the disease reaction, we hypothesized that QTLs for disease reaction could be located and quantified. Part of this population was mapped by Quieroz et al. (2003) using mostly dominant markers, and a major putative QTL for resistance to witches' broom was found on a map containing 25 linkage groups. With additional trees (146 compared to 82), and codominant, easily repeatable markers, it was our goal to locate QTLs on a map more closely corresponding to the cacao genome for use in breeding and future genetic studies, and for comparison with existing maps from F_1 populations discussed above.

Materials and Methods

PRODUCTION OF GENETIC MATERIAL AND PHENOTYPIC DATA.

The genetically segregating population used to construct this map was produced by selfing TSH516, a clone from a cross of Sca6 x ICS1, selected in Trinidad. Both Sca6 and TSH516 are self-incompatible. Self-incompatibility was overcome in TSH516 by using mentor pollen from the related genus *Herrania* Goudot (Pereira and Yamada, 1999). As stated above, Sca6 is considered to be resistant to *C. pernicioso*, and ICS1 to be susceptible. The initial F_2 population consisted of 151 trees, produced in two groups. The first group (Group 1) of 100 trees was produced in 1993, of which 79 individual tree samples were obtained, the remaining 21 trees having died. A second group (Group 2) of 80 trees was produced in 2000 to increase the population size, from which 72 DNA samples were extracted, giving a total of 151 unique tree samples for mapping. Groups 1 and 2 were used to construct the map, although phenotypic data was available only for Group 1. Two phenotypic traits were available, the number of brooms produced by witches' broom infection, and diameter of the main trunk. Number of brooms was recorded for each tree at the time of removal of brooms from 1997 to 2002, two or three times per year. Diameter of the main trunk was recorded in 2002 at the time of the last broom count.

DNA ISOLATION AND MARKER ASSAYS. DNA extraction was performed as described in Kuhn et al. (2003) using the Fast DNA Kit (Bio101, Carlsbad, Calif.) on 200-mg tissue samples taken from semi-adult leaves. Three different types of markers were used to construct this map (Table 1): 1) 174 microsatellite (SSR)

Table 1. Origin and number of microsatellite markers utilized for mapping the F_2 population, derived from selfing the cacao clone TSH516, derived from the cross, Sca6 x ICS1.

Marker origin	Marker nomenclature	No. of polymorphic markers
CIRAD, Montpellier, France	mTcCIR ^z	174
USDA-ARS, Miami	SHRSTc ^v	18
	RGH ^x	8
	WRKY ^w	4
Total		204

^zMicrosatellite markers; m = microsatellite, Tc = *Theobroma cacao*; CIR = CIRAD.

^vMicrosatellite markers; SHRS = Subtropical Horticultural Research Station; Tc = *T. cacao*.

^xResistance gene homologues.

^wWRKY gene, a transcriptional factor involved in the stress regulation.

markers developed at CIRAD, Montpellier, France (Lanaud et al., 1999b; Pugh et al., 2004; Risterucci et al., 2000) and 18 SSR markers from the SHRS, Miami (Kuhn et al., 2003); 2) eight resistance gene homologues (RGH) (Kuhn et al., 2003); and 3) four WRKY genes (Borrone et al., 2004); for a total of 204 markers. Our map shared 122 identical SSR markers with the map of Pugh et al. (2004). Markers based on RGH were described by Kuhn et al. (2003), and those based on WRKY genes by Borrone et al. (2004). RGH and WRKY markers were analyzed using single strand conformational polymorphism (SSCP) analysis, as described in detail in Kuhn et al. (2003). SSCP and analysis of SSR markers were conducted on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, Calif.) and analyzed with GeneScan Software (Applied Biosystems). All markers were first screened for polymorphism on the founding parents (Sca6 and ICS1) and the F₁ (TSH516).

MAPPING SOFTWARE AND MAPPING PROCEDURE. Genotype data output from Genotyper (Applied Biosystems) was imported into SAS V8.2 (SAS Institute, 2000), in which a program was written for the automated conversion of molecular weights of alleles from GeneScan output to the proper format for Joinmap, version 3 (Van Ooijen and Voorrips, 2001), a step in which many typographical errors can otherwise be introduced, reducing precision of the map (Hackett and Bradfoot, 2003). Joinmap was used to create the map, using the Kosambi mapping function. Data from five trees were eliminated due to excessive missing SSR data (two trees), apparent duplication (two trees), or poor DNA quality that repeatedly failed to produce PCR reactions (one tree). Data from eight additional trees were eliminated that had very distorted segregation, excessively skewed toward one parent or the other (>90%), and probabilistically unlikely in a population of this size. The final mapping population consisted of 138 individuals.

All markers were tested for conformity to the expected 1:2:1 segregation ratio by the chi-square statistic. Twenty-two markers with extremely skewed chi-square values [$P (>\chi^2) < 0.0001$] were eliminated from the data before constructing the final map. The criterion for eliminating these markers was having too few homozygous parental types of either Sca6 or ICS1 or their complete absence, while that parental type was present in the heterozygotes, causing suspicion as to whether these markers behaved as truly Mendelian loci in this population. Forty-nine additional markers, showing a lesser amount of segregation distortion (significance between $\alpha = 0.1$ and 0.0001) and comprising 27% of the remaining markers, were retained in the map, based on conclusions in a study by Hackett and Broadfoot (2003) that overall segregation distortion had little effect on map construction. The final dataset, with 182 total markers, had very few missing marker data points.

QTL MAPPING PROCEDURE. Both recorded characters, broom number and trunk diameter, were tested for normality using Proc Univariate of SAS (SAS Institute) and a Box-Cox macro also written in SAS. Trunk diameter data followed the normal distribution sufficiently well to be used without transformation. A square root transformation was found to best normalize broom number and to correct for any heterogeneity of variance. Phenotypic data were then merged with the marker data by tree for both Groups 1 and 2, leaving blank the missing values for Group 2. Retaining marker data from both tree groups was essential to perform QTL mapping with the best possible genomic map.

MapQTL version 4.0 (Van Ooijen and Maliépaard, 1996) was used to locate putative QTL for both phenotypic traits, using first simple interval mapping (SIM), followed by the multiple QTL

mapping (MQM) procedure (Jansen, 1993), and the restricted MQM (Res. MQM) procedure, both with cofactors. The population was fit with a model incorporating both additive and dominant effects. Automatic cofactor selection was done using backward elimination ($\alpha = 0.02$) by linkage group for both traits, beginning with approximately one cofactor every 10 to 20 cM (four to five markers per linkage group). For square root of broom number, a final set of six cofactors was retained, two on LG8, one on LG9, and three on LG6. For trunk diameter one cofactor on LG5 and one on LG9 were retained. The permutation method of Churchill and Doerge (1994) was used (1000 permutations) to obtain a LOD threshold for significance of QTL regions, resulting in a genome-wide LOD threshold at the 95th percentile of 3.7 for square root of broom number and 2.6 for trunk diameter.

It is known among cocoa researchers that a correlation exists between canopy circumference and trunk diameter. Resistance to witches' broom, evaluated by the number of brooms per tree, is affected by canopy size. An adjustment of broom count by analysis of covariance using trunk diameter as the covariate, similar to the adjustment of Quieroz et al. (2003) using canopy circumference, was unsuccessful; the coefficient for trunk diameter was nonsignificant, and its slope was negative, opposite of what would be expected. This population was apparently either of insufficient size (79 plants) or the trees had lost limbs with age and disease for this more general relationship to be apparent in the data. These two traits were therefore analyzed individually.

Results and Discussion

MAP CONSTRUCTION. The final linkage map is presented in Fig. 1. Ten linkage groups were formed, corresponding to the 10 chromosomes of cacao, and named following the conventions established by Lanaud et al. (1995). Linkage groups were clearly identifiable at minimal LOD scores of 2 to 3. Three markers, mTcCIR167, mTcCIR236, and mTcCIR44, had insufficient information to be mapped within their respective linkage groups (2, 3, and 8), and one marker, mTcCIR290, remained unlinked to any group, resulting in 178 effective markers. The map comprised a length of 671.9 cM, being 213.5 cM shorter than the high-density linkage map (885.4 cM) produced by Risterucci et al. (2000) and 110.9 cM shorter than the F₁ codominant map of Pugh et al. (2004). The high-density map with 424 markers (both codominant and dominant) had an average spacing of 2.1 cM between markers, and the recent F₁ codominant map had an average spacing of 1.7 cM between markers with 465 total markers. Our F₂ codominant map (178 markers) had an average distance of 3.8 cM between markers, a maximum distance of 22.3 cM and a median distance of 2.6 cM between markers. Linkage groups had a mean length of 67.3 cM, a range of 57.2 cM, the minimum and maximum lengths were 42.4 cM and 99.6 cM, respectively, and the median length was 58.7 cM. Remarkably good colinearity was observed between the F₁ and F₂ codominant maps, with only four pair-wise inversions of single marker positions, and distances between markers were also generally of the same magnitude, even for linkage group 10, which had only four markers in common and relatively few markers on either map. Disregarding a telomeric tail length of 11.9 cM on LG10, the two estimated linkage groups differed by only 4 cM (60.5 cM-F₁ vs. 56.5 cM-F₂). The mean of all Pearson correlation coefficients across linkage groups for markers in common on the two maps was 0.9927, with a maximum of 0.9981 and a minimum of 0.9805. The F₂ codominant map, with only three segregation types, and no need for constructing

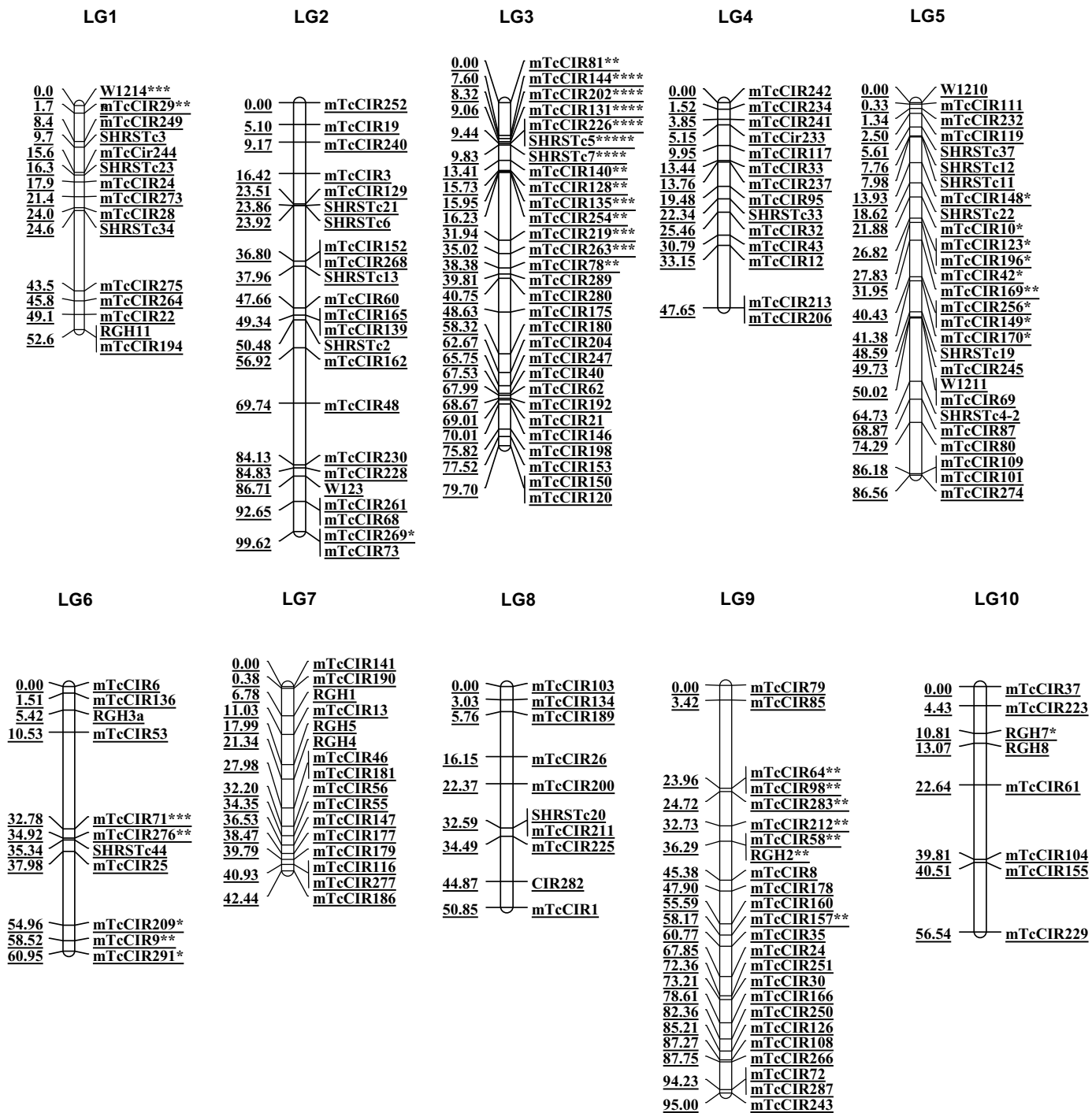


Fig. 1. Genetic linkage map of the genome of *Theobroma cacao* with 178 markers, based on an F₂ population derived from selfing the clone TSH516 (Sca6 x ICS1). The loci followed by asterisks are those whose segregation ratio did not fit the expected 1:2:1 ratio pattern when tested by χ^2 test [0.1 (*) and 0.05 (**), etc. probability]. The 10 linkage groups are presented in the order suggested by Lanaud et al. (1995). LG1 = linkage group 1; LG2 = linkage group 2, etc.

separate parental maps as with the dominant markers, would be expected to have good or better precision compared to F₁ maps, which would have a larger number of possible segregant types and therefore fewer trees per type for a given population size. These reasons seem to explain quite well why the F₂ codominant map has strong colinearity of marker order and similarity of distances among markers when compared to the previous maps of Risterucci et al. (2000) and Pugh et al. (2004), which had higher numbers of markers or progeny, or both.

SEGREGATION DISTORTION. The F₂ map had six regions of segregation distortion, covering 172.9 cM or approximately 25.7% of the genome, and involved 27% (49 of 182) markers. Risterucci et al. (2000) found distorted segregation ratios for 9.4% of all markers, located in four common areas of the genome. Segregation distortion in our map was seen to show a greater number of Sca6 alleles on LG1, LG3, and LG5 and a greater number of heterozygotes in groups LG6, LG9, and LG10 (one marker, RGH5, only). Approximately 9% of all markers in other published cacao

maps from F_1 populations have likewise demonstrated segregation distortion (Clement et al., 2003a, 2003b; Lanaud et al., 1995; Pugh et al., 2004). The distorted markers on our map resided in common areas on three linkage groups, LG5, LG6, and LG9, that showed a maximum distortion in the central region and decreased on both sides (see Fig. 1), suggesting a centralized cause. Distorted markers on LG1 and LG3 showed the same trend in a one-sided manner, the distorted markers being on one end of the linkage group. In the other codominant cacao map of Pugh et al. (2004), segregation distortion was found to be clustered on LG3 toward UF676 only, and also on LG5 and LG6 toward the heterozygote. Risterucci et al. (2000) also found distortion on LG5 and LG6, and on LG2 and LG9 as well. Therefore, distorted areas exist in common on LG5 and LG6 on three maps, with identical markers showing distortion on two maps, and in common on LG9 on our map and the high-density map. The fact that the maps involve different parents suggests that a genetic or chromosomal factor may cause the distortion. The argument of Risterucci et al. (2000) in using the Bonferroni correction for multiple independent tests, giving an individual threshold of 0.0001 for an experiment-wise type-I error of $P=0.05$, hinges on independence of the tests. Since the distorted markers on their map appear in linked groups, this argument is questionable [see also Lynch and Walsh, 1998 (p. 441)]. Chromosomal rearrangements (Tanksley, 1984; Tanksley et al., 1987) and gametic selection (Nakagahra, 1986) have both been proposed to explain segregation distortion in other species, though neither have been documented to exist in *T. cacao* (Risterucci et al., 2000). Lu et al. (2002), in a study of segregation distortion in four mapping populations of maize (*Zea mays* L.), found between 19% and 36% of all markers affected, similar to the amount we found. They found markers on a given chromosome to show unidirectional distortion, as we did, favoring the allele of the one parent or the heterozygous state, and also found common regions in different populations. They concluded that the commonality of regions across populations suggested the existence of segregation distorted regions (SDR), likely linked with three known gametophytic factors in maize. They also found that map distance was not affected when only one gametophytic factor was present per SDR. It seems likely that similar causal factors will be found in cacao, at least in these populations, to explain the blocks of linked, distorted markers on LG5, LG6, and LG9. The existence of these factors is also very important in association mapping based on linkage disequilibrium (Flint-Garcia et al., 2003), if, as in our map, up to one-quarter of the genome could be involved in them. Another causal factor for disequilibrium in this cross comes likely from the fact that Sca6 is homozygous for self-incompatibility, rendering its F_1 offspring with ICS1, TSH516, heterozygous for self-incompatibility. This incompatibility has a genetic basis and has been well described by Cope (1962a, 1962b) and by Knight and Rogers (1955). Segregation ratios described by these researchers include a self-compatible double homozygous segregant in the F_2 , which could contribute small amounts of segregation distortion toward Sca6. It is also conceivable that the use of mentor pollen is less than completely effective in overcoming the self-incompatibility of TSH516, and that some pollen genotypes could be selectively eliminated, resulting in segregation distortion of these genotypes. No reports of such genotype elimination are known, however, in using mentor pollen to overcome self-incompatibility. Finally, mapping populations with fewer than 200 individuals can quite possibly contain a few markers with distorted segregation ratios simply due to less-than-optimal population size for mapping.

QTL IDENTIFICATION. The results of SIM suggested two QTL regions for witches' broom resistance, one on LG9 with a very high LOD score (10.55) (Fig. 2A), explaining nearly 51% of the variance for the trait by SIM. We conclude that this region is identical to that described by Quieroz et al. (2003) as being linked to the RAPD marker, AV14.940 on their LG11, and is the important source of the simply inherited resistance reported by breeders. We also performed SIM on data from each year during which data were taken, and found this same region to appear consistently. Both witches' broom and trunk diameter had a cofactor and a QTL on LG9 (Fig. 2 A and B), most likely due to the (more generally observed) correlation between the two traits. The cofactors for each trait are also quite near the QTL peak for the opposite trait, less than 5 cM away. The MQM and Res. MQM procedures in MapQTL use cofactors retained after automatic cofactor selection for each trait. The MQM procedure creates a window of 5 cM around each region being tested for QTL presence by dropping any cofactor temporarily that would be a flanking marker in the interval in which a QTL is being fitted, as testing moves along the genome, leaving other cofactors in place. The Res. MQM procedure removes all cofactors temporarily from the entire linkage group on which a QTL test is being performed, as the procedure moves along the genome, but not from other linkage groups. Therefore the MQM mapping result is probably best viewed with caution, as non-flanking cofactors less than 5 cM away from a real QTL can detract misleadingly from its LOD (Jansen, 1993) (Fig. 2C). The position of the major QTL for number of brooms shifted slightly using MQM mapping vs. SIM (Fig. 2A) and Res. MQM mapping (Fig. 2D); the nearest marker changed from mTeCIR24 (LOD = 5.47) to mTeCIR157 (LOD = 8.15). The percentage of variance accounted for by the QTL also decreased from the original 51% with SIM, to 14.4% and 17.4%, respectively, using MQM and Res. MQM mapping. Theory would indicate that MQM mapping should be more precise in estimating the QTL position and effect; however, since only one-half of the population had phenotypic data (78 trees) and the cofactor resided less than 5 cM from both prospective peaks, there could be downward bias as suggested by Jansen (1993). Therefore caution is warranted when considering the MQM estimates of effects and variance; complete phenotypic data would allow more definitive resolution of the position and estimation of its effect. It seems quite likely that the actual location of the putative QTL lies within this region of approximately 5 cM between the MQM peak and the Res. MQM (and SIM) peaks, perhaps on the MQM peak. In using this information for selecting progeny from this cross for further crossing, we chose to use the Res. MQM peak (essentially equivalent to the SIM result), but suggest considering progeny favoring the other nearby peak, particularly if resistance data exist for witches' broom, and the phenotype is desirable. Another region was located on LG10, in the middle of a ~25 cM region with no markers present. This peak disappeared upon adding cofactors and using the MQM method, suggesting that it was an artifact of a long unmarked region, and it was therefore discarded. Using MQM with cofactors, the second putative QTL for witches' broom resistance was found on LG1, in an area well saturated with markers, one of which was RGH11. The peak had a LOD of 3.38, accounting for an estimated 6.7% of the variance, and was retained as a putative QTL. Both QTLs for witches' broom resistance have important dominance, the larger QTL toward Sca6, the resistant parent, and with the dominance of the minor QTL toward the susceptible parent, ICS1. Dominance effects in these two QTL alone could have been leading breeders

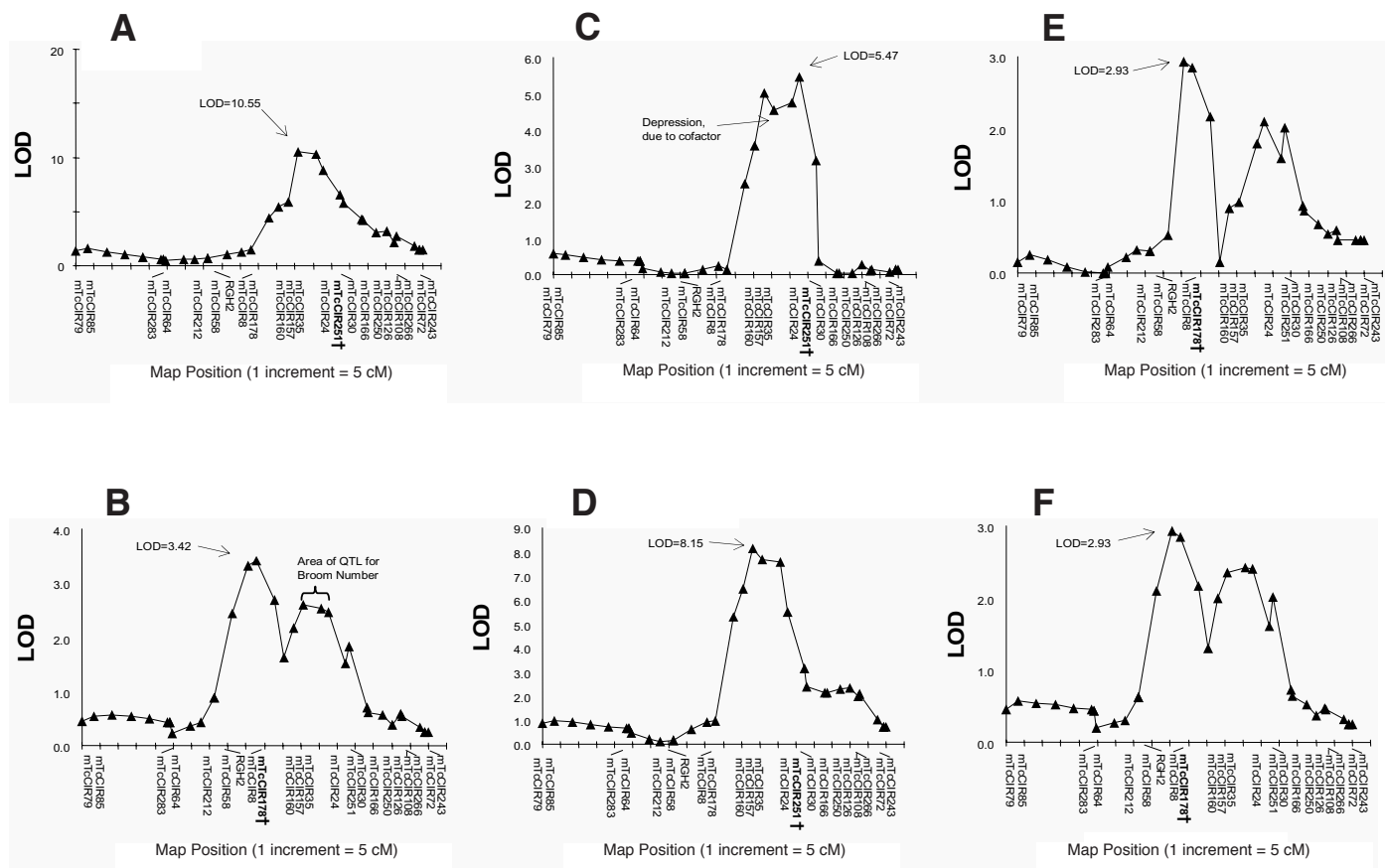


Fig. 2. Log of odds or log likelihood (LOD) curve plots for Linkage Group 9, with two quantitative trait loci (QTL) mapped on the genomic map. Markers in bold print with cross below are cofactors. The plots show: (A) the curve for broom number using simple interval mapping (SIM); (B) the curve for trunk diameter using SIM; (C) the curve for broom number using multiple QTL mapping (MQM); depression in curve is from cofactor; (D) the curve for broom number using restricted MQM (Res. MQM) mapping, dropping the cofactor; (E) MQM mapping for trunk diameter with deep dip in curve due to cofactor; and (F) Res. MQM mapping for trunk diameter without the cofactor effect in the curve, and a shallower dip in the curve.

to believe that “a few major genes” were involved in resistance rather than just one, the two gene ratios in progeny with dominance in opposite directions possibly resembling the segregation of “a few major QTL.” As this trait is quantitative, clear patterns of the segregation of two genes with opposing dominance effects are not likely to be observed by breeders when rating progeny for disease reaction.

For trunk diameter, LG9 shows two peaks using SIM, one with a LOD of 3.42 and the second with a LOD of 2.6, both significant (Fig. 2B). The second peak is nearer the peak found for broom number on that LG, however, and MQM mapping and Res. MQM mapping were used in addition to investigate the possibility of the second peak coming as a “ghost peak” from the indirect correlation of broom number with trunk diameter. The cofactor on LG9, just flanking the larger peak, clarifies the situation; the second peak is lowered to a LOD below the threshold of 2.6 by both MQM mapping (Fig. 2E) and Res. MQM mapping (Fig. 2F), leaving one well supported QTL for diameter, approximately 20 cM away from the QTL for broom number. Using a sequence of methods along with cofactors better supports the existence of both putative QTLs on the same linkage group, and seems to disprove the assertion that both peaks could be caused by one gene affecting the two indirectly correlated traits. A second potential QTL for trunk diameter on LG2 near

SHRSTc6, SHRSTc21, and mTcCIR129 was located by SIM; however, its LOD decreased to a point below the threshold using MQM and Res. MQM with cofactors, and it was discarded. The three final putative QTL regions found in this experiment are presented in Table 2, with their accompanying statistics and other information for breeders who might use these QTL for selection. Both SIM and Res. MQM results for the major resistance QTL are given, since a major resistance effect has long been known by breeders and it is highly likely to lie in this area. We consider the analysis of QTL position and effect by three methods to be important and complementary, especially since phenotypic data existed for only one-half of the trees. Comparing results of several methods is more likely to detect unpredictable multicollinearities among QTLs and cofactors in n-dimensional Euclidean space than analysis by one method alone, which could distort any of the associated estimates and cause questionable conclusions if only one method were used.

ASSOCIATIONS OF RESISTANCE WITH RGH. One of the candidate genes for disease resistance, RGH11, is a flanking marker for the minor resistance QTL on LG1, 3.5 cM to one side of the LOD peak. Another candidate gene, RGH2, maps at a distance of 24.6 cM away from the stronger QTL for resistance on LG9, probably too far away from the resistance QTL to provide support for RGH2 being related to witches’ broom. Lanaud et al. (2004)

Table 2. Genomic locations of quantitative trait loci (QTL) identified for both phenotypic traits in the F₂ population studied, flanking and nearest markers, log of odds or log likelihood (LOD) scores, percent variance explained, and estimates of additive and dominance effects. Three mapping methods were used; the method considered most appropriate for each QTL is given. The major QTL for resistance to witches' broom mapped by the restricted multiple QTL mapping (Res. MQM) method to a site 5.2 cM proximal to the site found by simple interval mapping (SIM). Given the importance of this QTL, results from both methods are given. The MQM method mapped it to the site identical to SIM, with effects and percent variance explained similar to Res. MQM.

Trait	Linkage group	Position (cM) from proximal end	Left flanking marker	Nearest marker	Right flanking marker	LOD	Mapping method	Variance explained (%)	Estimated additive effect ^x	Estimated dominance effect ^y
Broom no.	9	60.8	mTcCIR157	mTcCIR35	mTcCIR24	10.55	SIM	51.1	-0.62	-0.49
Broom no.	9	55.6	mTcCIR160	mTcCIR157	mTcCIR35	8.15	Res. MQM	12.7	-0.45	-0.33
Broom no.	1	49.1	mTcCIR264	mTcCIR22	RGH11	3.38	MQM	6.7	-0.11	0.38
Trunk diam	9	45.4	RGH2	mTcCIR8	mTcCIR178	2.93	MQM	15.6	-0.1	2.2

^{x,y}Estimated additive and dominance effects are given relative to Sca6. A negative number indicates a lower number of brooms or smaller diameter for the Sca6 allele for the additive effect. For the dominance effect, a negative number indicates that the Sca6 allele imposes dominance and ICS1 for a positive number.

found two clusters of candidate resistance genes containing the NBS domain clustered on LG7 and LG10. On LG10, RGH7 and RGH8 mapped approximately 3 cM apart, as did N171/F12 and NSCA6/A1 on the map of Lanaud et al. (2004). Sequence comparison suggests these RGH loci to be identical.

FINAL COMMENTS. As all previous cacao genomic maps (except that of Quiroz et al., 2003) have been made from heterozygous F₁ parents, the existence of at least one F₂ map with correspondence to the species chromosome number allows better comparisons of the cacao genome with many annual crops, in which maps are more often made from F₂ or recombinant inbred (RI) populations. *Theobroma cacao* was recently reclassified as a member of the taxonomic family Malvaceae (Bayer et al., 1998); therefore, the development of anchor markers with other agriculturally useful members of the genus *Theobroma* L., such as *T. grandiflorum* (Willd. Ex Spreng) Schum (de Sousa Silva et al., 2001), and with *Gossypium hirsutum* L. and other dicotyledons would be useful for genomic comparisons and for gene expression studies. Distances estimated among markers in this F₂ map are very comparable to those from F₁ maps, even with far fewer markers in the F₂ map. It is also possible to obtain estimates of separate additive and dominance effects with an F₂ population, which is not possible with dominant markers or with "pseudo-backcross" populations. The fact that rather strong dominance was found indicates that these resistance QTL can function well in the heterozygous state, though the second QTL for broom number has dominance toward the susceptible parent. This study importantly shows the existence of dominance effects in cacao on the genic level, as well, necessary to consider when breeders introgress resistance. Adding RFLP or AFLP markers to this population would serve to help relate the map to previous maps, and also to fill in certain gaps with different marker types (Risterucci et al., 2000). Currently, rooted cuttings from the entire population are being inoculated with *C. pernicioso* spores, which will enable reading of a disease response for the entire population. The future counting of brooms on the younger set of trees as they mature will be invaluable in raising precision of QTL mapped in this study and in locating additional modifier regions. The collection of other phenotypic traits (e.g., yield and yield components, morphological traits) would enable additional, valuable QTL mapping in this population. Very likely, the mapping of clusters of related resistance QTL and agronomic traits, as did Clement et al. (2003a, 2003b), would be possible in this F₂, as would be the estimation of dominance for such traits. This study has initiated a valuable start toward the comprehensive

mapping of resistance to *C. pernicioso* in cacao; clonal propagation and distribution of the population could provide future insight into host-pathogen interactions by exposing it to other pathotypes of *C. pernicioso*. Resistance to witches' broom in the pods is believed to be due to a different underlying set of genes, and this population merits evaluation for pod resistance. The shape of the LOD curve for resistance over the genome indicated that several other possible modifier genes could be present, not distinguished here due to the relatively high threshold required for significance. Finally, the investigation of this population with fluorescent in situ hybridization technique could be useful for locating regions containing chromosomal features known to inhibit or distort recombination, such as rearrangements or large amounts of heterochromatin, and the searching of LG5 and LG6 for gametophytic selection factors is likely to provide heretofore unknown information in the study of cacao.

Literature Cited

- Ahnert, D. and J.L. Pires. 2000. Use of the available genetic variability of cocoa in Brazil. Proc. Tech. Mtg.: State of the knowledge on mass production of genetically improved propagules of cocoa. 1:104-113.
- Andebran, T. 1984. Studies on the epidemiology and control of witches' broom disease of cocoa in the Brazilian Amazon. Proc. Intl. Cocoa Res. Conf. 9:395-402.
- Bayer, C., M.F. Fay, A.Y. de Bruijn, V. Savolainen, C.M. Morton, K. Kubitzki, W.S. Alverson, and M.W. Chase. 1998. Support for an expanded family concept of Malvaceae within a circumscribed order Malvales: A combined analysis of *atpB* and *rbcL* DNA sequences. Bot. J. Linnean Soc. 129:267-303.
- Borrone, J.W., D.N. Kuhn, and R.J. Schnell. 2004. Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. Theor. Appl. Genet. 109(3):495-507.
- Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.
- Clement, D., A.M. Risterucci, L. Grivet, J.C. Motamayor, J. N'Goran, and C. Lanaud. 2003a. Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. Genome 46:204-212.
- Clement, D., A.M. Risterucci, L. Grivet, J.C. Motamayor, J. N'Goran, and C. Lanaud. 2003b. Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. Genome 46:103-111.
- Cope, F.W. 1962a. The mechanism of pollen incompatibility in *Theobroma cacao* L. Heredity 17:157-182.
- Cope, F.W. 1962b. The effects of incompatibility and compatibility on genotype proportions in populations of *Theobroma cacao* L. Heredity 17:183-195.

- Couch J., H.A. Zintel, and P. Fritz. 1993. The genome of the tropical tree *Theobroma cacao* L. *Mol. Gen. Genet.* 237:123–128.
- Crouzillat, D., B. Menard, A. Mora, W. Phillips, and V. Petiard. 2000a. Quantitative trait analysis in *Theobroma cacao* using molecular markers. Yield QTL detection and stability over 15 years. *Euphytica* 114:13–23.
- Crouzillat, D., W. Phillips, P. Fritz, and V. Petiard. 2000b. Quantitative trait loci analysis in *Theobroma cacao* using molecular markers. Inheritance of polygenic resistance to *Phytophthora palmivora* in two related cacao populations. *Euphytica* 114: 23–36.
- de Souza Silva, C.R., A.V. de Oliveira Figueira, and C.A. Spaggiari Souza. Diversity in the genus *Theobroma* (in Portuguese), p. 40–80. In: L.A.S. Dias (ed.). Genetic improvement of cacao. FUNAPE, UFG, Viçosa, MG, Brazil.
- Flament, M.H., I. Kebe, D. Clement, I. Pieretti, A.M. Risterucci, J.A.K. N’Goran, C. Cilas, D. Despreaux, and C. Lanaud. 2001. Genetic mapping of resistance factors to *Phytophthora palmivora* in cacao. *Genome* 44:79–85.
- Flint-Garcia, S., J.M. Thornsberry, and E.S. Buckler, IV. 2003. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54:357–374.
- Fowler, M.S. 1994. Fine or flavour cocoas: Current positions and prospects. *Cocoa Growers’ Bul.* 48:17–23.
- Hackett, C.A., L. B. Broadfoot. 2003. Effect of genotyping errors, missing values, and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90:33–38.
- Jansen, R. 1993. Interval mapping of multiple quantitative trait loci. *Genetics* 135:205–211.
- Johnson, E.S., C.B. O’Connor, T.N. Sreenivasan, and R.J. Schnell. 2003. Population structure of the witches’ broom pathogen of cacao in Trinidad and Tobago. Abstract no. 126E. *Proc. Intl. Cocoa Res. Conf.* 14: (In press.)
- Knight, R. and H.H. Rogers. 1955. Incompatibility in *Theobroma cacao* L. *Heredity* 9:69–77.
- Kuhn, D.N., M. Heath, R.J. Wisser, A. Meerow, J.S. Brown, U. Lopes, and R.J. Schnell. 2003. Resistance gene homologues in *Theobroma cacao* as useful genetic markers. *Theor. Appl. Gen.* 107:191–202.
- Lanaud, C., P. Hamon, and C. Duperray. 1992. Estimation of the nuclear DNA content of *Theobroma cacao* L. by flow cytometry. *Café, Cacao, Thé* 36:3–8.
- Lanaud, C., A.M. Risterucci, J.A.K. N’Goran, D. Clement, M.H. Flament, V. Lauent, and M. Flaque. 1995. A genetic linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 91:987–993.
- Lanaud, C., I. Kebe, A.M. Risterucci, D. Clement, J.A.K. N’Goran, L. Grivet, M. Tahi, C. Cilas, I. Pieretti, A.B. Eskes, and D. Despreaux. 1999a. Mapping quantitative trait loci (QTL) for resistance to *Phytophthora palmivora* in *T. cacao* L. *Proc. Intl. Cocoa Res. Conf.* 12:99–105.
- Lanaud, C., A.M. Risterucci, I. Pieretti, M. Falque, A. Bouet, and P.J.L. Lagoda. 1999b. Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* 8:2141–2152.
- Lanaud, C., A.M. Risterucci, I. Pieretti, J.A.K. N’Goran, and D. Fargeas. 2004. Characterisation [sic] and genetic mapping of resistance and defence [sic] gene analogs in cocoa (*Theobroma cacao* L.). *Mol. Breeding* 13:211–227.
- Lu, H., J. Romero-Severson, and R. Bernardo. 2002. Chromosomal regions associated with segregation distortion in maize. *Theor. Appl. Genet.* 105:622–628.
- Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Assoc., Sunderland, Mass.
- Motilal, L.A., D.R. Butler, and V. Moolledhar. 2002. Verification in global cacao germplasm collections. *Ingenic Nwsl.* 7:4–8.
- Nakagahra, M. 1986. Geographic distribution of gametophyte genes in wide crosses of rice cultivars. *Proc. Intl. Rice Genet. Symp.* 73–82.
- Pereira. T.N.S. and M.M. Yamada. 1999. Hybridization in cacao (in Portuguese), p.153–174. In: A. Borem (ed.). Artificial hybridization of plants. Editora UFV, Viçosa, Brazil.
- Pugh, T., A.M. Fouet, P. Brottier, M. Abouladze, C. Deletrez, B. Courtois, D. Clement, P. Larmande, J.A.K. N’Goran, and C. Lanaud. (2004). A new cacao linkage map based on codominant markers: Development and integration of 201 new microsatellite markers. *Theor. Appl. Genet.* 108:1151–1161.
- Quiroz, V.T., C.T. Guimarães, D. Anher, I. Schuster, R.T. Daher, M.G. Pereira, V.R.M. Miranda, L.L. Loguercio, E.G. Barros, and M.A. Moreira. 2003. Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches’ broom disease. *Plant Breeding* 122:268–272.
- Risterucci, A.M., L. Grivet, J.A.K. N’Goran, I. Pieretti, M.H. Flament, and C. Lanaud. 2000. A high-density linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 101:948–855.
- SAS Institute Inc. 2000. SAS/procedures guide, version 8. SAS Inst., Cary, N.C.
- Schnell, R.J., J.S. Brown, J.C. Motamayor, and D.N. Kuhn. 2004. Development of DNA markers associated with disease resistance in cacao. 10 Nov. 2004. <<http://www.ars-grin.gov/ars/SoAtlantic/Miami/Pages/PlantSciences/project8.htm>>.
- Steinberg, F.M., M.M. Bearden, and C.L. Keen. 2003. Cocoa and chocolate flavonoids: Implications for cardiovascular health. *J. Amer. Medical Assn.* 103:215–23.
- Tanksley, S.D. 1984. Linkage relationships and chromosomal locations of enzyme-coding genes in pepper, *Capsicum annum*. *Chromosoma* 89:352–360.
- Tanksley, S.D., J. Miller, A.H. Paterson, and R. Bernatzky. 1987. Molecular mapping of plant chromosomes, p. 157–173. In: J.P. Gustafson and R.A. Apples (eds.). *Chromosome structure and function*. Plenum Press, New York.
- Van Ooijen, J.W. and C. Maliepaard. 1996. MapQTL version 3.0: Software for the calculation of QTL position on genetic maps. *Plant Res. Intl.*, Wageningen, The Netherlands.
- Van Ooijen, J.W. and R.E. Voorrips. 2001. JoinMap version 3.0, Software for the calculation of genetic linkage maps. *Plant Res. Intl.*, Wageningen, The Netherlands.
- Yamada, M. and U.V. Lopes. 1999. Paternity analysis of cacao trees selected for resistance to witches’ broom in plantations of Bahia, Brazil. *Agrotropica* 11(2):83–88.