# 16

# Cocoa

V. Niral, M.K. Rajesh and V.A. Parthasarathy

## Introduction

Cocoa, *Theobroma cacao* L., is commercially exploited on a large scale for the production of cacao beans. Cocoa butter and powder, which are extracted from the fermented and dried cocoa beans, are the main ingredients used for the commercial manufacture of chocolate. Cocoa is a diploid, mainly cross-pollinated species (2n=20) and has a small nuclear genome. Lanaud *et al*. (1992) reported the size of the haploid cocoa genome by flow cytometry to be approximately $3.88 \times 10^8$ bp.

With recent support by molecular evidences by Motamayor *et al*. (2002), South America is presently accepted as the region of origin of cocoa. The evaluation of the genetic diversity of wild cacao plants based on morphology, agronomic characters and molecular markers supports the upper Amazon region as the putative center of diversity of cacao.

Establishing the genetic diversity and structure of cocoa natural populations is critical to define strategies for long-term conservation of genetic resources of this neotropical tree species, and to maintain the industry sustainability. Improved cocoa cropping requires the development of genetic materials that have a higher productivity and an increased resistance as well as cacao beans of good industrial quality. To obtain such characteristics, cocoa geneticists have made abundant use of molecular markers, especially microsatellites.

## Molecular Profiling

### *Isoenzymes*

Isozymes were the first molecular marker to be utilized in cacao. The first attempt to characterize cocoa germplasm using isoenzymes was made by Atkinson *et al*. (1986). Out of

the 24 enzyme systems tested, 17 showed activity, and in seven enzyme systems, reproducible variation in banding patterns was obtained.

Warren (1994) used isozyme polymorphism to estimate the genetic diversity in nine populations of cocoa from Peru, Ecuador, Grenada, Trinidad and Colombia maintained at the International Cocoa Genebank, Trinidad. The distribution of allelic variation was observed using four isoenzyme systems *viz*. acid phosphatase, malate dehydrogenase, isocitrate dehydrogenase and phosphoglucoisomerase. The low level of polymorphism and low genetic diversity indicated that all the populations analyzed from Peru were genetically similar and disproved the theory that Upper Amazon region of Peru was the centre of cacao genetic diversity. Although the available loci and the number of polymorphisms generated were low, this system contributed to assessment of genetic diversity, genotype identification and linkage mapping (Ronning and Schnell, 1994; Lachenaud *et al.,* 2004; Sounigo *et al.,* 2005). However, isoenzyme markers are presently outdated because of their low polymorphism and also the environmental effect on their phenotype.

### *DNA Markers*

Presently, DNA markers are used for molecular profiling of cocoa. The commonly used DNA markers are RFLP, RAPD, AFLP and SSRs.

#### *RFLP*

RFLP was the first DNA marker to be applied in cacao in early 1990s (Laurent *et al.*, 1994). The polymorphism of RFLP is moderately high in cacao and has been used for assessment of genetic diversity (N'Goran *et al*.,1994; Motamayor *et al.*, 2002). A high genetic diversity was revealed when 175 genotypes of cocoa belonging to different morphological groups and origins were analyzed using 27 RFLP/cDNA loci (N'Goran *et al*., 2000). Criollo genotypes appeared differentiated from Forastero genotypes as has been found by morphological traits. A significant, but rather low fixation index between Criollo and Trinitario genotypes pointed out the proximity of these two sub-groups and that Trinitario, as a hybrid group, had incorporated more genes from Criollo progenitors.

RFLP probes, however are invariably specific to a limited number of loci, and hence RFLP is not a very effective tool for cacao genotype identification. Another major drawback is that RFLP is not amenable to automation and data generation is both laborious and expensive.

#### *RAPD*

RAPD was the first PCR-based DNA fingerprinting method to be applied for the genetic characterization of cacao (Wilde *et al.,* 1992).This system is technically simple to perform but has low reproducibility. Russell *et al.* (1993) showed that a minimum of three RAPD primers could distinguidh 25 cocoa accessions based on their geographical origin. Various workers have used RAPDs for identification on accessions (Christopher *et al.,* 1999; Falerio *et al.*, 2002; Sounigo *et al.,* 2005) and for genetic diversity studies in cacao (Figueira *et al.,* 1994; N'Goran *et al.,* 1994; Lerceteau *et al.,* 1997; Whitkus *et al.,* 1998; Marita *et al.,* 2001; Sounigo *et al.,* 2005).

*AFLP*

AFLP combines the strength of RAPD and PCR. It has been utilized in varietal identification and Saunders *et al.* (2001) was able to distinguish closely related cocoa varieties through AFLP analysis. Queiroz *et al.* (2003) identified a marker associated with resistance to Witches' broom disease based on AFLP linkage mapping. However, the utility of AFLPs in cacao germplasm management has been rather limited.

*Microsatellites*

Since the development of simple sequence repeats (SSR) markers in cacao (Lanaud *et al.* 1999), SSR-based DNA fingerprinting has been increasingly applied in cacao germplasm characterization (Lanaud *et al.* 2001; Motamayor *et al.,* 2002; Saunders *et al.,* 2004; Schnell *et al.* 2005; Takrama *et al.,* 2005; Cryer *et al.*, 2006). Large numbers of cocoa microsatellite markers are now available (Kuhn *et al.,* 2003; Clement *et al.,* 2004; Pugh *et al.,* 2004) for use in fingerprinting studies. SSRs are typically codominant and multi-allelic allowing precise discrimination of individual clones based on the multi-locus banding profiles.

SSR analysis allows unambiguous verification of duplicates by a matching of multilocus DNA profiles between the potential duplicates (Takrama *et al.,* 2005; Cryer *et al.,* 2006; Zhang *et al.*, 2006). Zhang *et al.*, 2006, 2008 and Johnson *et al.* (2007) demonstrated the effectiveness of using SSR markers for the identification of synonymously mislabeled accessions and that the use of 15 SSR loci for cacao individual identification is statistically rigorous. This indicates the highly significant benefit of using molecular characterization for cacao genebank management.

SSRs have been applied to verify genetic identity in breeding programme (Takrama *et al.,* 2005) and to monitor the integrity of germplasm maintained *in vitro* (Rodriguez *et al.,* 2004). SSRs have also been used for determining the genealogical relationships and parentage analysis. Schnell *et al.* (2005) used SSR analysis to determine the parental population for a group of productive and unproductive seedlings. Seven SSR loci were found to have alleles that were associated with productive or unproductive seedlings. Motamayor *et al.* (2003), on the basis of SSR evidence, reported that Trinatario accao is the result of natural hybridization and subsequent introgressions between Lower Amazon Forastero and ancient Criollo varieties.

SSRs have lately been used to investigate the phylogeography and distribution of genetic diversity in cacao. Motamayor *et al.* (2002) assessed the allelic composition of cacao accessions from Central America (Guiana, Amazonia and Orinoco regions) using SSRs. Their results supported the hypothesis that cocoa originated in the upper Amazon and the likely dispersal route was from the Amazon to Central America and Mexico. Sereno *et al.* (2006) analyzed 94 cocoa accessions from the Forastero group, representing four populations from the Brazilian Amazon, using 11 microsatellite markers. The highest genetic diversity, the most polymorphic loci, the highest observed heterozygosity and the majority of rare alleles were detected from the population from the Brazilian upper Amazon, which suggested that this could be a part of the center of diversity of cocoa. Also, the population from Lower Amazon, which traditionally was considered to be highly homozygous, presented an unexpectedly high observed

heterozygosity and rare and distinct alleles. This population also had a large identity with the upper Amazon population. It was hypothesized that part of the lower Amazon population might have been derived from successive natural or intentional introduction of planting material from other provinces, mainly the upper Amazon.

Molecular marker studies have indicated that most of the gene diversity in cacao was found to occur within groups, rather than between groups (Ronning and Schnell, 1994; Lerceteau *et al.*, 1997; N'Goran *et al.*, 2000), being similar to other outcrossing woody perennial species (Ronning and Schnell, 1994). Sereno *et al.* (2006), therefore, suggested that *T. cacao* has a strong intra-population structure with small differentiation between populations, typical of species with high gene flow.

## Generation of Molecular Linkage Maps

Several linkage maps have been constructed for cocoa by different workers using different mapping populations. The first genomic map in cacao was produced by Lanaud *et al.* (1995) from a cross between an upper Amazon Forastero clone, 'UPA402', and a Trinitario clone produced by United Fruit in Costa Rica, 'UF676'. This map was made from 100 trees and contained several different kinds of markers from that era: RFLPs, RAPDs, isoenzymes, and phenotypic genes, for a total of 202 markers. The map had a total genomic length of 759 cM in 10 linkage groups, putatively corresponding to the known 10 chromosomes of cacao. This map has since been considered to be the "reference map" for the naming of linkage groups of cacao.

Risterucci *et al.* (2000) expanded this mapping population to contain 181 trees, and added markers based on AFLPs and SSRs. These markers were used for the full population of 181 trees, together with the data from the work of Lanaud *et al.* (1995). An integrated approach was used for producing this map, with a total of 424 mapped markers. Several different kinds of markers were used for this map, which provided more extensive genomic coverage, saturating certain areas that had not previously been mapped. The total genomic length was 885.4 cM, with an average distance of 2.09 cM between markers. Colinearity with the original map for markers in common was high. This map is considered as a good "reference map" for global genomic length in cacao, although several of the marker types in this map are seldom used any more.

Pugh *et al.* (2004) constructed another linkage map of cacao from the same population (UPA402 × UF 676) using only codominant markers. Two hundred one new SSR markers were added to the map, along with the previous isozyme data, RFLP data, and with 16 new candidate genes (resistance and defense gene analogs). Again, ten linkage groups were formed, with a total genomic length of 782.8 cM, and an average distance of 1.7 cM between markers. The length of the map established with only 268 SSRs was 769.6 cM, representing 94.8% of the total map of the new SSR markers, 26% were "bridge markers", heterozygous in both parents, allowing good integration of the map between parents. SSR markers were found to be distributed throughout the genome, however their distribution did not seem random; some genomic regions were more heavily saturated than others. No genomic gap was found to be

greater than 10 cM, indicating that a large portion of the genome was covered by SSR markers. Given the variety used in the complete marker set, some of which were known to be telomeric markers, there was a high degree of certainty that nearly all of the genome was covered. Separate maps were also made from both parents, which showed good colinearity to one another, with only a small number of markers showing inversions between closely linked markers. Good colinearity was also found when comparing common markers between the integrated map and the former two maps made from this population**.**

Crouzilla *et al.* (1996) developed a linkage map using a backcross population of 131 plants derived from a cross between a $F_1$ tree from 'Catango' × 'Pound 12' to a recurrent 'Catango'. The final map contained 140 markers (RFLPs, RAPDs) and included two morphologic loci (self compatibility and anthocyanin synthesis) and covered 944 cM (Crouzilla *et al.,* 2000b). An additional genetic map was developed for the related $F_1$ population from the cross Catongo × Pound 12 with 162 markers, covering 772 cM (Crouzillat *et al.,* 2000a).

Subsequently, another eight linkage maps were developed to search for genomic regions associated with resistance to various *Phytophthora* species and strains (Flament *et al.,* 2001, Motilal *et al.,* 2002; Clement *et al.,* 2003b). These maps were also simultaneously used for detection of QTLs associated with important agronomic traits in certain families.

Queiroz *et al.* (2003) constructed another cocoa genetic linkage map using an $F_2$ population derived from the cross between ICS-1 and Scavina-6 clones. Later, this map was saturated by Faleiro *et al.* (2006). Subseqently, Araújo *et al.* (2009) saturated the previous linkage map with AFLP, RAPD, and microsatellite markers, resulting in a total of 273 markers, distributed in 14 linkage groups (LGs). The total coverage of the genome was of 1,122 cM. This genome coverage was 31% (1,713 cM) smaller than that obtained by Queiroz *et al.* (2003) and 60% (670 cM) bigger than obtained by Faleiro *et al.* (2006), which mapped 193 loci (124 RAPD and 69 AFLP) along 25 LG and 342 loci (33 SSR, 77 AFLP, and 232 RAPD) along 16 LG, respectively, for the same population. The sizes of the 14 linkage groups ranged from 21 cM in LG14, to 138.3 cM in LG6, which clustered six and 41 markers, respectively. The marker-to-marker distance varied from 0.0 to 24.8 cM, with an average distance of 4.1 cM, with less than seven marker regions exceeding 20 cM. This average marker-to-marker distance is considerably shorter than observed for the map built by Queiroz *et al.* (2003), 8.8 cM. However, it is closer to the values determined in the maps developed by Faleiro *et al.* (2006), Risterucci *et al.* (2000), Crouzillat *et al.* (2000a, b), Flament *et al.* (2001), Risterucci *et al.* (2003), and Pugh *et al.* (2004) that were 2.7, 2.1, 4.8, 3.7, 3.2, and 1.7 cM, respectively.

Brown *et al.* (2008) constructed the first composite map of cacao from linkage data of one $F_2$ and two $F_1$ mapping populations with a high number of codominant markers in common. The combinatsion of linkage information from all three maps results in the currently most precise estimates of marker locations and distances between markers, especially in densely marked areas, quality assessment and mapping. Individual (sub-composite) maps and the composite map contained 10 major linkage groups, corresponding to the number of cacao chromosomes. Homogeneity of marker placement was very high among sub-composite maps, the composite map, and the designated "reference" map. The composite map places more

markers with higher precision than any individual map. This research clearly demonstrates for the first time a very high level of marker homogeneity among commercial cacao clones compared to other species. The observed homogeneity between different maps, including the composite one, is probably due to a narrow genetic base of commercial cacao clones. Markers which are linked to identified quantitative trait loci (QTLs) are more likely to retain linkage in other commercial clones, rendering the QTLs in cacao potentially more stable than in other species.

## QTL Mapping and MAS

QTL analysis has a greater potential for identification of markers for marker-assisted breeding in cocoa, especially for incorporation of disease resistance. Several genetic maps have been built for the cocoa tree from data generated by molecular markers, and most of them have been used to aid the identification of QTLs related to disease resistance and different traits of agronomic interest (Despreaux and Eskes, 1997; Crouzillat *et al.,* 1996, 2000a, 2000b, 2001; Ahnert, 2001; Motilal *et al.,* 2002; Clement *et al.,* 2003a, b; Queiroz *et al.,* 2003; Lanaud *et al.,* 2004; Risterucci *et al.*, 2004; Brown *et al.,* 2005). Recently, Araújo *et al.* (2009) identified QTLs related to cocoa butter content and hardness.

### *Disease Resistance*

An international project was initiated in 1995 to study the genetics of resistance to *P. palmivora* and *P. megakarya* in cocoa. This was implemented by IRAD (Cameroon), IDEFOR (Cote d'Ivoire), Cocoa Research Unit (Trinidad) and CIRAD (France), with an aim to characterize resistance factors, compare evaluation methods, identify quantitative trait loci (QTL) based on molecular markers. Using saturated linkage maps, the group has identified QTLs for field resistance in different cocoa populations (Despreaux and Eskes, 1997; Lanaud *et al.*, 2004). Further, Crouzillat *et al.* (2000b) identified QTLs conferring increased resistance in susceptible parents, suggesting the presence of transgressive traits and the possibility of selection in cacao. They also detected pleiotropic and epistatic effects for the QTLs associated with resistance of cacao to *Phytophthora*. Motilal *et al.* (2002) identified thee major QTLs on chromosomes 1, 9, and 3 or 8 that co-localized with QTLs detected in other studies. Risterucci *et al.* (2004) identified QTLs associated with resistance to *P. palmivora* using composite interval mapping. These QTLs were found located in six genomic regions. One of these was detected with five strains belonging to the three *Phytophthora* species. Two other regions were detected with two or three strains of two different species. Three additional QTLs were detected for only one species of *Phytophthora*. Each QTL explained between 8 to 12% of the phenotypic variation. For each strain, between 11.5% to 27.5% of the total phenotypic variation could be explained by the QTLs identified. The identification of multiple QTLs involved in resistance to *Phytophthora* offers the possibility to improve durability of resistance in cocoa by a possible accumulation of many different resistance genes located in different chromosome regions using marker-aided selection.

Lanaud *et al.* (2004) undertook characterization and genetic mapping of resistance and defense gene analogs in cocoa. Resistance gene homologues (RGHs) may make better molecular markers for disease resistance in cocoa because resistance genes are often physically clustered

in plants. By identifying putative members of the clusters, it may be possible to rapidly identify useful genetic markers which will show a stronger linkage to the disease resistance phenotype. Most of the resistance gene analogs (RGAs) could be mapped on the cocoa genome and three clusters of genes could be observed. Nucleotide binding site (NBS) like sequences (present in a number of plant resistance genes) clustered in two regions located on chromosomes 7 and 10. Pto-like sequences mapped in five genome regions of which one, located on chromosome 4, corresponded to a cluster of five different sequences. PR2-like sequences mapped in two regions located on chromosome 5 and 9, respectively. They observed co-localizations of RGAs and defense gene analogs (DGAs) and QTL for resistance to *Phytophthora* in several progenies, particularly on chromosome 4 where a cluster of Pto-like sequences and four QTLs for resistance to *Phytophthora* were observed. Many other serious diseases affect cocoa and the candidate genes, isolated in this study, could be of broader interest in cocoa disease management.

Witches' broom, *Crinipellis perniciosa* (Stahel) Singer, is a serious fungal disease of cacao in Latin in most of the cocoa growing, regions in South America and several Caribbean islands. Studies on QTL mapping for Witches broom resistance in Brazil by Ahnert, (2001) and Queiroz *et al.* (2003) using $F_2$ plants derived from the cross ICS 1 (susceptible parent) and SCA 6 (resistant parent) has resulted in the identification of a major QTL on the linkage group 11, explaining almost 35% of the resistance to witches' broom. Interval and composite interval mapping models identified one major genomic region associated with witches' broom resistance on the SCA 6 linkage group. These results are in accordance with the pattern of segregation observed in the field for SCA 6 progenies and may be used in a marker-assisted selection process for gene pyramidization. Brown *et al.* (2005) identified two QTLs for resistance to witches' broom disease, one producing a major effect and one a minor effect; with both showing important dominance effects. Further, 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. Subsequently, Faleiro *et al.* (2006) saturated the linkage map with additional markers and identified QTL of bigger effect that controls resistance to Witches' broom disease. To enable cloning of this resistance gene, a BAC library was constructed from 'Scavina 6' (Clement *et al.,* 2004) and a larger $F_2$ population is currently being developed in Brazil.

Frosty pod rot caused by *Moniliophora roreri* is another destructive disease causing considerable yield reduction in Central America. Brown *et al.* (2005) used a population of 'Pound 7' × 'UF273' for linkage mapping and evaluation for frosty pod resistance. They detected three major QTLs associated with resistance against internal and external lesions by *M. roreri* at the same region on chromosome 2, and 8 and an additional QTL for external resistance on chromosome 7.

### *Agronomic and Quality Traits*

Yield and yield component traits are of great economic importance. In cacao, pod yield is also influenced by the trunk diameter and canopy size. Crouzillat *et al.* (1996, 2000b) detected 2-4 QTLs for early flowering, trunk diameter, jorquette height and ovule number in an $F_1$ population of 'Cantaongo × Pound 2'. Subsequently, 10 yield QTLs were detected on eight

linkage groups. Some of these QTLs were frequently detected overs 15 years of production, while others were specific for a given year. Two major QTLs (E and I) explained ~20% of the total variance of the average yield over 15 years. Crouzillat *et al.* (2001) detected a major QTL for general agronomic value, explaining 27.1% of the variation, co-localized with QTL for early flowering and trunk diameter, indicating the possibility of pleiotropic and epistatic effects. They also observed that the QTL for weight of 100 cocoa beans was found on a different chromosomal location to the QTL for agronomic traits suggesting, that these traits could be selected independently.

Clement *et al*. (2003b) undertook QTL mapping for agronomic traits in cacao. Three heterozygous clones, one upper Amazon Forastero (IMC78) and two Trinitario (DR1 and S52), were crossed with the same male parent, a lower Amazon Forastero (Catongo), known to be highly homozygous. Observations were made on the progenies over nine consecutive years (1990-1998). One to three QTLs related to yield were detected in each of the three populations, located on chromosomes 1, 2, 4, 5, 9, and 10. They explained between 8.1 and 19.3% of the phenotypic variation and showed various levels of repeatability. In IMC78, the QTL detected on chromosome 5 was the most repeatable over the years. The QTL for the average individual pod weight on chromosome 4 was the most significant with an LOD of 17.3 and an $R^2$ of 43.7. QTL related to these traits were identified in the same region of the genome in clones of different genetic groups. Clement *et al.* (2003a) undertook QTL mapping for bean traits and number of ovules per ovary and identified QTL clusters on several chromosomes, particularly chromosome 4. Further, the QTLs governing bean traits were detected in the same/related genomic regions in diverse clones of different parentage (two with Trinitario parents and one with upper Amazon Forastero parent). A QTL for pod index was also identified on chromosome 4 of 'Pound 12' (Crouzillat *et al.,* 2000b). Clement *et al.* (2001) reported co-localization of QTLs for pod weight on chromosome 1 for Trinitario genotypes. This suggests that molecular markers can be used to improve cocoa varieties.

Araújo *et al.* (2009) used an $F_2$ population derived from the cross between the ICS-1 and Scavina-6 cocoa clones for molecular mapping and identified one quantitative trait locus associated to butter content at linkage group 9 (LG9) and two QTLs for butter hardness at linkage groups 9 and 7 (LG9 and LG7). The two QTLs mapped at the LG9 explained 51.0% and 28.8% of the phenotypic variation for butter content and hardness, respectively. These QTLs were concentrated in the same map region, suggesting a close genetic linkage or pleiotropic effect. The QTLs identified may be useful in further marker-assisted selection breeding programs aimed at cocoa butter quality improvement.

The prospect of progress from identification of QTLs to the isolation of gene(s) by map-based cloning has not yet been applied in cacao. Such an approach requires a very large population to accumulate a suitable number of recombinant individuals (around 1,000), which is not only expensive in terms of time and resources but is also extremely difficult in the case of cacao (Wilkinson, 2000). The bacterial artificial chromosome (BAC) library developed by Clement *et al.* (2004) can be a valuable resource for cloning genes corresponding to some important QTLs for agronomic traits such as resistance genes to major cocoa pathogens like *Phytophthora* spp. (*palmivora* and *megakarya*), *Crinipellis perniciosa* and *Moniliophthora*

*roreri*. Further, multiple QTLs identified for resistance to diseases and for different yield and yield component traits can be made use of for a possible accumulation of these favourable QTLs located in different chromosome regions using marker-aided selection.

## Future Perspectives

*Theobroma cacao* is mainly cultivated by small growers, especially in the developing world and is of great economic importance to the growers and the global food industry. The sustainability of cocoa cultivation will be improved if farmers have access to new planting material with improved agronomic traits such as yield, vigour, and pest and disease resistance. Progress in breeding programmes to accumulate favourable alleles for these traits can be accelerated using molecular marker techniques which allow more direct access to the genome.

## References

Ahnert, D., 2001. Use of QTLs for witches' broom resistance in cocoa breeding. In: Proceedings of the international workshop on new technologies and cocoa breeding, Kota Kinabalu, Sabah, Malaysia, 16-17 October 2000 (eds. Bekele,-F; End,-M; Eskes,-A), London, UK: INGENIC pp. 116-119.

Araújo, I.S., de Souza Filho, G.A., Pereira, M.G., Faleiro, F.G., de Queiroz, V.T., Guimarães, C.T., Moreira, M.A.,. de Barros, E.G., Machado, R.C.R.,. Pires, J.L., Schenell, R. and Lopes, U.V., 2009. Mapping of Quantitative Trait Loci for Butter Content and Hardness in Cocoa Beans (*Theobroma cacao* L.). Plant Mol. Biol. Rep., 27:177-183.

Atkinson, M.D., Withers, L.A. and Simpson, M.J.A., 1986. Characterisation of cacao germplasm using isoenzyme markers. 1. A Preliminary survey of diversity using starch gel electrophoresis and standardisation of the procedure. *Euphytica*, 35:741-750.

Brown, J.S.,Schnell, R.J., Motamayor, J.C., Lopes, U., Kuhn, D.N., Borrone, J.W., 2005. Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an $F^2$ population using SSR markers and candidate genes. *Journal of the American Society for Horticultural Science,* 130(3):366-373.

Brown, J.S., Sautter, R.T., Olano, C.T., Borrone, J.W., Kuhn, D.N., Motamayor, J.C. and Schnell, R.J., 2008. A Composite Linkage Map from three crosses between commercial clones of cacao, *Theobroma cacao* L. *Tropical Plant Biol.* 1:120-130. DOI 10.1007/s12042-008-9011-4.

Christopher, Y., Mooleedhar, V., Bekele, F., Hosein, F., 1999. Verification of accession in the ICG, T using botanical descriptors and RAPD analysis. In: Annual report 1998. The University of the West Indies, St. Augustine, Trinidad and Tobago, pp. 15-18.

Clément, D., Risterucci, A.M., Motamayor, J.C., Goran, J. and Lanaud, Co. 2003b. Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. *Genome*, 46:103-111.

Clement, D., Lanaud, C., Sabau, X., Fouet, O., Cunff, L., le, Ruiz, E., Risterucci, A.M., Glaszmann, J.C., Piffanelli, P., 2004. Creation of BAC genomic resources for cocoa (*Theobroma cacao* L.) for physical mapping of RGA containing BAC clones. *Theoretical and Applied Genetics*, 108(8):1627-1634.

Clement, D., Risterucci, A.M., Lanaud, C., 2001. Analysis of QTL studies related to yield and vigour traits carried out with different cocoa genotypes.In: Proceedings of the international workshop on new technologies and cocoa breeding, Kota Kinabalu, Sabah, Malaysia, 16-17 October 2000, London, UK: INGENIC. pp. 127-134.

Clement, D., Risterucci, A.M., Motamayor, J.C., Goran, N., J. Lanaud, C., 2003a. Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. *Genome*, 46(2):204-212.

Crouzillat, D., Rigoreau, M., Cabigliera, M., Alvarez, M., Bucheli, P., Pétiard, V., 2001. QTL studies carried out for agronomic, technological and quality traits of cocoa in Ecuador. In: Proceeding of the International Workshop-on new technology and Cocoa breeding, Kota Kimabalu, Sabah, Malaysia, 16-17, October, 2000, Londan, UIC INGENIC, pp. 120-126.

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. *Theoretical and Applied Genetics*, 93(1/2):205-214.

Crouzillat, D., Menard, B., Mora, A., Phillips, W., Petiard, V., 2000b. Quantitative trait analysis in *Theobroma cacao* using molecular markers. Yield QTL detection and stability over 15 years. *Euphytica*, 114(1):13-23.

Crouzillat, D., Phillips, W., Fritz, P.J., Petiard, V., 2000a. Quantitative trait loci analysis in *Theobroma cacao* using molecular markers. Inheritance of polygenic resistance to *Phytophthora palmivora* in two related cacao populations. Polygenic resistance to Phytophthora palmivora in cacao. *Euphytica*, 114(1):25-36.

Cryer, N.C., Fenn, M.G.E., Turnbull, C.J., Wilkinson, M.J., 2006. Allelic size standards and reference genotypes to unify international cocoa (*Theobroma cacao* L.) microsatellite data. *Genet. Resour. Crop Evol.*, 53:1643-1652.

Despreaux, D., Eskes, A., 1997. An international project on the genetic bases of resistance of cacao to *Phytophthora*. INGENIC Newsletter, 3:9

Faleiro, F.G., Queiroz, V.T., Lopes, U.V., Guimarães, C.T., Pires, J.L., Yamada, M.M., Araújo, I.S., Pereira, M.G., Schnell, R., De Souza Filho, G.A., Ferreira, C.F., Barros, E.G., Moreira, M.A., 2006. Mapping QTLs for witches' broom (*Crinipellis perniciosa*) resistance in cacao. *Euphytica*, 149:227-235.

Faleiro, F.G., Santos, I.S., Bahia, R.C.S., Santos, R.F., Yamada, M.M. and Anhert, D., 2002. Otimiza¸c˜ao da extra¸c˜ao e amplifica¸c˜ao de DNA de *Theobroma cacao* L. visando a obten¸c˜ao de marcadores RAPD. *Agrotr´opica*, 14:31-34.

Figueira, A., Janick, J., Morris, L. and Goldsbrough, P., 1994. Reexamining the classification of *Theobroma cacao* L. using molecular markers. *J. Amer. Soc. Hort. Sci.,* 119(5):1073-1082.

Flament, M.H., Kebe, I., Clement, D., Pieretti, I., Risterucci, A.M., N' Goran, J.A.K., Cilas, C., Despreaux, D., Lanaud, C., 2001. Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. *Genome*, 44(1):79-85.

Johnson, S.E., Mora, A., Schnell, R.J., 2007. Field guide efficacy in the identification of reallocated clonally propagated accessions of cacao. *Genet. Resour. Crop. Evol.*, 54:1301-1313.

Kuhn, D.N., Heath, M., Wisser, R.J., Meerow, A., Brown, J.S., Lopes, U. and Schnell R.J. 2003. Resistance gene homologues in *Theobroma cacao* as useful genetic markers. *Theor. Appl. Genet.*, 107:191-202.

Lachenaud, P., Sounigo, O., Oliver, G., 2004. Genetic structure of Guianan wild cocoa (*Theobroma cacao* L.) described using isozyme electrophoresis. *Plant Genetic Res Newsl.*, 139:24-30.

Lanaud, C., Hamon, P., Duperray, C., 1992. Estimation of the nuclear DNA content of *Theobroma cacao* L. by flow cytometry. *Cafe, Cacao, The*, 36:3-8.

Lanaud, C., Motamayor, J.C., Risterucci, A.M., 2001. Implications of new insight into the genetic structure of *Theobroma cacao* L. for breeding strategies. In: Proceedings of the International Workshop on New Technologies for Cocoa Breeding, Kota Kinabalu, Malaysia, London: INGENIC Press, 89-107.

Lanaud, C., Risterucci, A.M., N' Goran, J.A.K., Clement, D., Flament, M.H., Laurent ,V. and Falque, M., 1995. A genetic linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.*, 91:987–993.

Lanaud, C., Risterucci, A.M., Pieretti, I., Falque, M., Bouet, A., Lagoda, P.J.L., 1999. Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.*, 8:2141-2152.

Lanaud, C., Risterucci, A.M., Pieretti, I., N'goran, J.A.K. and Fargeas, D., 2004. Characterisation and genetic mapping of resistance and defence gene analogs in cocoa (*Theobroma cacao* L.). *Mol. Breeding*, 13(3):211-227.

Laurent, V., Risterucci, A.M., Lanaud, C.,1994. Genetic diversity in cocoa revealed by cDNA probes. *Theor. Appl. Genet.*, 88:193-198.

Lerceteau, E., Robert, T., Pe´ tiard, V., Crouzillat, D., 1997. Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theor. Appl. Genet.*, 95:10-19.

Marita, J.M., Nienhuis, J., Pires, J.L., Aitken, W.M., 2001. Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches' broom disease resistance. *Crop Sci.*, 41:1305-1316.

Motamayor, J.C., Risterucci, A.M., Heath, M., Lanaud, C., 2003. Cacao domestication. II. Progenitor germplasm of the Trinitario cacao cultivar. *Heredity*, 91:322-330.

Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., Lanaud, C. 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity*, 89:380-386.

Motilal, L.A., Sounigo, O., Thévenin J.M., Howell, M.H., Pieretti, I., Risterucci, A.M., Noyer, J.L. and Lanaud, C., 2000. *TheobromaTheobroma cacao* L.: genome map and QTLs for *Phytophthora palmivora* resistance. In: Proc Intl Cocoa Res Conf 13 Kota Kinabalu, Malaysia, 2000. Cocoa Producer's Alliance, Lagos, Nigeria, pp. 111-118.

N'Goran J.A.K., Laurent, V., Risterucci, A.M. and Lanaud, C. 2000. The genetic structure of cocoa populations (*Theobroma*

*cacao* L.) revealed by RFLP analysis. *Euphytica*, 115:83-90.

N'Goran, J.A.K., Laurent, V., Risterucci, A.M. and Lanaud, C., 1994. Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. *Heredity*, 73:589-597.

Pugh, T., Fouet, O., Risterucci, A.M., Brottier, P., Abouladze, M., Deletrez, C., Courtois, B., Clement, D., Larmande, P., N'Goran, J.A., 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.

Queiroz, V.T. Guimaraes, C.T., Anhert, D., Schuster, I., Daher, R.T., Pereira, M.G., Miranda, V.R.M., Loguercio, L.L., Barros, E.G., Moreira, M.A., 2003 Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. *Plant Breeding,* 122(3):268-272

Risterucci, A.M., Grivet, L., N,Goan, J.A.K., Pierct, I., Flament, M.H. and Lanaud, C. 2000. A high-density linkage map of *Theobroma cocoa* L. *Theor. Appl. Genet*., 101: 948-955.

Risterucci, A.M*.,* Paulin, D*.,* Ducamp, M*.,* N' Goran, J.A.K. and Lanaud, C., 2004. Identification of QTLs related to cocoa resistance to three species of *Phytophthora. Theor. Appl. Genet.,* 108(1):168-174.

Rodriguez, L.C.M.,Wetten, A.C. andWilkinson, M.J., 2004. Detection and quantification of in vitro-culture induced chimerism using simple sequence repeat (SSR) analysis in *Theobroma cacao* (L.). *Theor. Appl. Genet*., 110:157-166.

Ronning, C.M. and Schnell, R.J., 1994. Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *J. Heredity* 85(4): 291-295.

Russell, J.R., Hosein, F., Johnson, E., Waugh, R. and Powell, W., 1993. Genetic differentiation of cocoa (*Theobroma cacao* L.) populations revealed by RAPD analysis. *Mol. Ecol*., 2:89-97.

Saunders, J.A., Hemeida, A.A. and Mischke, S., 2001. USDA DNA fingerprinting programme for identification of *Theobroma cacao* accessions. In: F Bekele et al (ed) Proc Intl Workshop New Technol Cocoa Breeding 2000. INGENIC Press, London. pp. 108-114.

Saunders, J.A., Mischke, S., Leamy, E.A. and Hemeida, A.A., 2004. Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theor. Appl. Genet*., 110:41-47.

Schnell, R.J., Olano, C.T., Brown, J.S., Meerow, A.W., Cervantes-Martinez, C., Nagai, C. and Motamayor, J.C., 2005. Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. *J. Am. Soc. Hort. Sci*., 130:181-190.

Sereno, M.L., Albuquerque, P.S.B., Vencovsky, R. and Figueira, A., 2006. Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. *Conserv. Genet*., 7:13-24.

Sounigo, O., Umaharan, R., Christopher, Y., Sankar, A. and Ramdahin, S., 2005. Assessing the genetic diversity in the International Cocoa Genebank, Trinidad (ICG,T) using isozyme electrophoresis and RAPD. *Genetic Res. Crop Evol*., 52:1111-1120.

Takrama, J.F., Cervantes-Martinez, C.T., Phillips-Mora, W., Brown, J.S., Motamayor, J.C. and Schnell, R.J., 2005. Determination of off-types in a cacao breeding program using microsatellites. INGENIC *Newsl*., 10:2-8.

Warren, J.M., 1994. Isozyme variation in a number of populations of *Theobroma cacao* L. obtained through various sampling regimes. *Euphytica*, 72:121-126.

Whitkus, R., de la Cruz, M., Mota-Bravo L. and Go´ mez-Pompa, A., 1998. Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theor. Appl. Genet*., 96:621-627.

Wilde, J., Waugh, R. and Powell, W., 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. *Theor. Appl. Genet*., 83:871-877.

Zhang, D., Mischke, S., Johnson, E.S., Phillips-Mora, W. and Meinhardt, L., 2008. Molecular characterization of an international cacao collection using microsatellite markers. *Tree Genetics & Genomes,* 5:1-10.