Cacao

MARK J. GUILTINAN

The Department of Horticulture, 422 Life Sciences Building, The Pennsylvania State University,

University Park, PA, 16802

Corresponding author: M. J. Guiltinan

Fax: 814 863-6139

Email: mjg9@psu.edu

Guiltinan M (2007) Cacao. In: Pua EC, Davey MR (eds) Biotechnology in Agriculture and Forestry - Transgenic Crops VI. Springer-Verlag, Berlin Heidelbelg (in press)

1 Introduction

Theobroma cacao L. (cacao) is a small under-story tree endemic to the lowland rainforests of the Amazon basin (Wood and Lass 1985; Bartley 2005). Cacao was domesticated in pre-Columbian times by the Olmec and Maya civilizations, the latter of which used the seeds ("cocoa beans") to produce beverages for royalty and religious ceremonies, and as currency (Coe and Coe 1996; Motamayor et al. 2002; Emch 2003). Today, cacao is grown throughout the humid tropics, often in agroforestry-ecosystems with other fruit and commodity crops.

1.1 Worldwide Market

The ability to store dried cocoa beans for long periods of time and the strong demand for the production of chocolate make it one of the world's most valuable commodity crops. The annual world production of cocoa is approximately 3 million tons, with 2/3 being processed into cocoa powder and cocoa butter, and the remaining 1/3 used for cocoa liquor (the flavor and color component of chocolate) (Wood and Lass 1985). Cocoa is the major export commodity of several countries in West Africa (68% of world production), providing major economic resources to Ivory Coast, Cameroon, Nigeria and Ghana. Other major cocoa exporters include Ecuador, Venezuela, Brazil, Panama, Costa Rica, Malaysia and Indonesia. More recently, cacao has become a priority crop in regions of Vietnam and India. It is also an important replacement cash crop for illegal narcotics in Colombia and Peru, offering farmers a legal alternative for their livelihoods. Worldwide, approximately 5 to 6 million smallholder farmers grow 95% of the worlds' production on more than seven million hectares, providing an important source of income for their families. World cocoa export commerce is $5 to $6 billion/year and the use of cocoa and cocoa butter in chocolate manufacturing, cosmetics and other products drive an approximately seventy billion dollar market providing over 60,000 jobs in the US alone (Morais 2005). US chocolate production also uses large amounts of sugar, nuts and milk, valued at approx. $3 billion/year in receipts to American farmers and similar benefits to many other countries.

1.2 Sustainability and Ecological Benefits of Cacao

Cacao-growing regions are largely centered in important biodiversity hotspots, impacting 13 of the world's most biologically diverse regions (Piasentin and Klare-Repnik 2004). For example, Southern Bahia, Brazil, a major cacao-producing region, contains some of the last remaining vestiges of the Atlantic Rainforest ("Mata Atlântica"). Ivory Coast and Ghana, world leaders in cocoa production, are both located in the Upper Guinea Forest, which contains more than half of the mammalian species in Africa, the highest mammalian diversity density on earth, many of which are endangered. Because cacao is a shade grown perennial tree crop, with a cropping cycle of more than 50 years, cultivation creates environmental benefits such as enhancement of biodiversity in avian migratory corridors, soil and watershed conservation, and providing buffer zones near endangered rainforest habitats (Rice and Greenberg 2003; Ruf and Zadi 2003). Development agencies such as USAID, Conservation International (CI), The World Wildlife Federation (WWF) and the World Cocoa Foundation, are increasingly aware of the role of cacao in stabilizing local economies and environments and have stepped up their involvement with cacao farmers in these regions (Guyton et al. 2003).

1.3 Economic/Social Impact of Cacao Diseases

Cacao pathogens reduce the potential crop by an estimated 810,000 tons annually (30% of world production) and individual farm losses can approach 100% (Keane 1992; Bowers et al. 2001). For example, in Southern Bahia, Brazil, Witches' broom disease (caused by the fungus *Crinipellis perniciosa*) resulted in a decrease of production from 300,000 tons in 1989 to 130,000 tons 10 years later, for an estimated loss of $220 million each year (Pereira et al. 1990). This devastating disease has resulted in an economic and ecological disaster, resulting in widespread homelessness and a major social crisis. Widely dispersed cacao pathogens include several species of *Phytophthora* that cause multiple diseases of economic importance, including pod rot, trunk canker and leaf-blight (Appiah et al. 2003, 2004; Chowdappa et al. 2003). *Phytophthora megakarya*, the most aggressive and damaging species, has been reported to have entered Ivory Coast, the world's leading cocoa producer (Nyasse et al. 2002; Opoku et al. 2002; Appiah et al. 2003; Risterucci et al. 2003; Efombagn et al. 2004). Other important diseases and pests include frosty pod in central America (Evans et al. 2003), the cocoa pod borer in Asia

(Day 1984; Santoso et al. 2004) and cocoa swollen shoot virus, which is particularly devastating in Ghana, Africa and is also spreading into the Ivory Coast (Hanna 1954; Hervé et al. 1991; Muller and Sackey 2005).

It is estimated that only 30% of the varieties grown today are the result of selection programs, and the remaining 70% are essentially fortuitous introductions of plants with low disease resistance and yield potential, originally derived from the Amazon basin and propagated by farmers through seeds passed down through the generations (Bennett 2003). While large genetic variation has been identified in wild populations throughout the Amazon, this diversity has not been widely utilized in terms of practical incorporation into cultivated varieties (Bartley 2005). This dramatically illustrates the primitive state of domestication of cacao, but it also highlights the opportunities for future enhancement through the application of modern molecular/genetic breeding approaches. It is widely accepted that although agronomic and agrochemical technologies can play important roles in integrated pest management strategies for cacao, breeding has the advantage of delivering genetically encoded traits such as disease resistance to the farmer, so that less inputs are required once the trees have been planted. This is especially important since cacao is often grown by poor farmers who cannot afford expensive agricultural inputs. Disease resistance is currently the primary trait targeted by cacao breeders. Other important traits for cacao include yield efficiency, flavor characteristics, cocoa butter content (% seed lipid content) and quality (fatty acid saturation), tolerance to abiotic stress, and various horticultural traits such as precocity, rootstock/scion interactions, plant height and stature.

1.4 Cacao Molecular Genetics Research Community

Today's cacao genetics research community is well organized, highly collaborative, and poised to make use of new genomics resources, which has been reviewed by Bennett (2003). In order to formally foster collaboration and communication between cacao breeders and geneticists, the International Group for Genetic Improvement of Cocoa (INGENIC) was formed in 1994. It now includes over 300 members, representing 35 developing and developed countries around the world. The INGENIC Study Group for Molecular Biology (INGENIC-MOL-BIOL) was formally chartered in October of 2003 (Johnson 2003) to coordinate the activities of the INGENIC members interested in molecular approaches. An

international research symposium is held by INGENIC in a developing country every third year. More information about INGENIC is available at the website (http://ingenic.cas.psu.edu).

In addition to cacao research centers in developed countries such as the US, France and the UK, most of the cocoa producing countries have research facilities funded by international and national organizations that support agricultural research. The sophistication of these facilities varies from elementary to excellent. The main strengths of these organizations are the many scientists with strong experiences in cacao agriculture, as well as their extensive field sites and breeding programs. It is essential that researchers in developed and developing countries establish and maintain strong working relationships and collaborative research and training programs to maximize the potential impact of our research on the cacao farmers and the environment.

2 Recent Advances: A Recent Literature Review

2.1 Molecular Markers

2.1.1 Germplasm Evaluation via DNA Fingerprinting

One of the central applications of molecular biology to plant genetics is the use of molecular markers in the analysis of relationships between accessions using phylogenetic analysis. For cacao, this has been approached with isozymes, Random Amplified Polymorphic DNA (RADP), Restriction Fragment Length Polymorphisms (RFLP) and other types of genomic DNA markers. Most recently, microsatellite markers (SSR) have gained acceptance as the most accurate and reliable method (Atkinson et al. 1986; Fritz 1991; Wilde et al. 1991; Figueira et al. 1994; Ronning et al. 1995; Lerceteau et al. 1997; Crouzillat et al. 2000; Gomes et al. 2000; Niella et al. 2000; Dias 2001; Motamayor et al. 2002; Kuhn et al. 2003; Borrone et al. 2004; Faleiro et al. 2004; Pugh et al. 2004; Brown et al. 2005; Sereno et al. 2005). Using these approaches, our understanding of the evolutionary relationships between different cacao accessions and of the relationships between members of germplasm collections have been advancing rapidly (Figueira et al. 1994; Whitkus et al. 1998; Marita et al. 2001; Motamayor et al. 2002). These investigations have ranged from characterizing germplasm collections, to elucidating the origin of different cacao lineages spanning back to the ancient Mayas. For example, molecular markers have been used to shed light on a

long standing controversy, the origins of a particular lineage of cacao, the *Criollo*, believed to have been cultivated by the Mayas over 1500 years ago (Motamayor et al. 2002, 2003). This analysis suggests that contrary to the hypothesis put forward by others, *Criollo cacao* does not represent a separate subspecies, but may have derived from a few individuals in South America and then later spread by man into Central America. As discussed in the epilogue of his recent book, Bartley (2005) points out that this indicates that the Criollo group represents an insignificant proportion of the total diversity in the species. Understanding these lineages has important implications to breeders who wish to incorporate diverse sources of resistance and other traits into breeding programs.

In a broader sense, the same methodology can be used to investigate the evolutionary relationships of cacao with other species. In this way, using molecular phylogenetic analysis of chloroplast DNA sequences, *Theobroma cacao* and its related species have been reclassified from the family Sterculiaceae to the Malvaceae (Alverson et al. 1999; Bayer et al. 1999; Whitlock et al. 2001). This is important to cacao researchers who wish to extrapolate information from other related species, such as cotton, which is in the same family as cacao, and *Arabidopsis thaliana*, a model plant for molecular-genetic research worldwide that is also closely related to cacao.

Cacao Germplasm collections are found in several sites throughout the world, and the diversity maintained in these collections is an important resource for breeders. These accessions are systematically being evaluated for various traits of interest, such as disease resistance and quality (Iwaro et al. 2003). One of the clear messages from the germplasm characterization work has been the realization that many of the genotypes residing in collections are mis-labeled and that DNA markers can be used to fingerprint accessions and rectify some of these discrepancies (Charters and Wilkinson 2000). International standards for cacao fingerprinting have been agreed upon and are being adjusted and modified by the cacao research community through mutual cooperation and testing (Saunders et al. 2004; Cryer et al. 2005). With these new standards, germplasm collections are now being systematically characterized, and mislabeling problems corrected. Cryer et al. (2005) described two methods of standardization of microsatellite allele profiles between different laboratories. This was performed on a total of 429 cacao accessions. They fingerprinted these accessions with 15 microsatellite markers that have been agreed upon as common set for worldwide standardization. Importantly, the data was deposited in the International Cocoa Germplasm Database, which is an easily accessible and searchable

database for cacao molecular genetic data, and another new important resource for the cacao research community (http://www.icgd.reading.ac.uk/netscape_index.htm).

An excellent example of how molecular markers can be used to study the genetic diversity and population structure of cacao collections was published recently (Sereno et al. 2005). Four populations from the upper and lower Amazon were analyzed using microsatellite makers. Interestingly, the accessions from the upper Amazon contained the largest diversity, and thus the authors suggest that this specific sub-region might be the center of diversity for cacao. Another example of using fingerprinting was published by Faleiro et al. (2004), who studied genetic variability of 19 accessions from around the Amazon basin. RAPD and microsatellite markers were used to demonstrate the high genetic variability of this collection, which is not surprising considering the Amazon is the center of origin for *Theobroma cacao* (Cruz et al. 1995; Motamayor et al. 2002, 2003).

Individual gene sequences from different accessions can also be used to study genetic diversity. Sousa Silva and Figueira (2005) published a study using the sequences of the Kunitz-like trypsin inhibitor gene from 11 *Theobroma* and three related *Herrania* species. The phylogenetic analysis carried out using these sequences was consistent with the phylogeny previously proposed based on morphological characters.

One aspect of developing molecular makers for germplasm characterization and marker-assisted selection is to generate highly reproducible and economical methods that can be applied in a large scale in the future. Kuhn et al. (2003, 2005) and Kuhn and Schnell (2005) reported on the development of new genetic markers for cacao using a technique, capillary array electrophoresis-single strand conformation polymorphism (CAE-SSCP), which has never before been applied to cacao. This method is capable of detecting single base pair differences between two alleles of a gene, and multiple markers can be multiplexed in single reaction tubes. Using automated sample handling and high throughput automated DNA fragment detection systems, a very high throughput system was developed. The authors showed that CAE-SSCP was capable of detecting single nucleotide polymorphisms in cacao, and specified parameters to increase the reliability of the method. If this method can be adopted internationally, it could accelerate mapping and fingerprinting efforts.

Molecular markers can also be used to study mutation rates in cells, as was recently reported (Rodriguez Lopez et al. 2004). In this study, sequence repeat (SSR) analysis was used to detect

mutations in *in vitro* grown plantlets derived from cacao somatic embryos. Analysis of 233 regenerated plants revealed that 31% of the plants contained putative chimeric mutations, that is to say, portions of the plants appeared to have mutations. The plants tested were not regenerated and acclimated into glasshouse conditions, a process that is known to select against undesirable mutations from populations of embryos. It will be interesting to see if this result is specific for the population of plants tested in this study, or is representative of *in vitro* cacao plants regenerated in other labs or with different protocols.

2.1.2 Molecular Mapping

The recent advances in the field of cacao genetic mapping and breeding have been reviewed recently (Bennett 2003; Figueira and Alemanno 2005). Much of this work is performed in collaborative projects between laboratories in developed and in producing countries. Recent progress has involved the development of useful markers for mapping, the production of a reference molecular-genetic map and the application of these markers in quantitative trait loci (QTL) mapping studies. One landmark study was the creation of a reference molecular-genetic map of cacao, which was developed by the Lanaud group in France (Risterucci et al. 2000; Pugh et al. 2004). This map currently consists of >250 SSRs and >400 RFLP, RAPD, AFLP and isozyme markers, covering ~900 cM of the 10 cacao chromosomes with an average marker distance of ~2 cM. This framework map is important to serve as a scaffold for genetics and genomics research in the future, and will allow integrating QTL maps from different genotypes for various traits. Another landmark has been the establishment of a large number of microsatellite markers for mapping in cacao (Lanaud et al. 1999; Pugh et al. 2004). These authors created a specific genomic DNA sub-library enriched for short repetitive DNA elements (microsatellite DNA), which are very useful as molecular markers because the length of the repeat elements are highly variable between genotypes and are easy to measure using conventional PCR based analysis.

Quantitative Trait Loci (QTL) mapping is a technique central to developing markers to assist breeders in molecular based selection schemes. This approach determines regions of the genome that contribute to complex, multi-genic traits, and allows the development of molecular probes that can be used to screen progeny in breeding populations for the desired allelic combinations of these genomic regions. QTLs for various traits in cacao have been identified, including resistance to fungal diseases,

and various yield and morphological traits of interest such as fruit size and seed size (Lanaud et al. 1996; Motilal et al. 2000; Flament et al. 2001; Clement et al. 2003a, 2003b; Queiroz et al. 2003; Risterucci et al. 2003). An example of the power of these methods in identification of important traits is the QTL mapping of a major genomic locus conditioning resistance to the severe pathogen of cacao, *Crinipellis perniciosa* (Queiroz et al. 2003). In this study, a F2 population derived from a cross between the resistant Scavina-6 genotype and highly susceptible ICS-1 genotype were used. Mapping of 193 markers (AFLP and RAPD) resulted in identification of several QTLs, including one major QTL responsible for 35% of the resistance. Using a related population of plants with an increased number of individuals (146 trees) and using co-dominant markers, Brown et al. (2005) created a linkage map using SSR markers and 12 candidate resistance genes. A much better genetic linkage map was produced using this approach, and two strong QTLs for resistance to witches' broom disease showing dominance were detected. In addition, a QTL for trunk diameter close to one of the resistance QTLs was also found. Similar QTL maps have been produced for many different traits, as reviewed previously (Figueira and Alemanno 2005). Much of the currently available resources are now going into the generation of more markers, increasing marker density on maps and defining higher resolution to the major QTLs that have been identified for major disease resistances. Although not yet accomplished, it is likely that the coming years will see the first isolation of a cacao gene responsible for a QTL using a positional cloning approach, and the major QTL for resistance to witches' broom is a likely first target.

Schnell et al. (2005) published a study on the use of association genetic methods as applied to cacao. This method has been shown to be useful in human genetics and with other species where controlled populations and specific genetic designs are not possible. The application of this method to perennial tree crops could potentially circumvent some of the difficulties faced with the development of cacao populations for QTL mapping studies. In this study, microsatellite markers were used to identify loci associated with productivity in a selection of mature plants. The majority of these markers co-localized to QTLs previously identified for productivity by conventional QTL mapping (Clement et al. 2003a, 2003b), demonstrating the utility of this approach. This approach can potentially speed up the development of molecular markers useful for enhanced cacao breeding programs.

2.2 Gene Discovery

2.2.1 BAC Library Resources

Two cacao Bacterial Artificial Chromosome libraries (BAC libraries) have been constructed recently. A BAC library is useful for gene discovery in that it contains large fragments of genomic DNA each cloned into a bacterial strain that is arrayed in microtiter plates, making it simple to screen and to isolate specific genomic regions. Clement et al. (2004) reported on the construction of a BAC library from the genotype Scavina-6, is one of the most well-known and utilized genotypes of cacao. This genotype was collected in 1938 near the Ucayali River in Peru on the eastern side of the Andes in the center of origin of cacao, the Amazon basin (Pound 1940). Seeds from a tree were used to produce a series of clones that still exist in Trinidad and from which cuttings have been disseminated around the world. Although poor yielding due to its small pods and seeds, the Scavina 6 genotype contains genes conditioning tolerance to several different fungal pathogens, and is used widely in breeding programs. This library contains approximately ten genome equivalents with an average insert size of 120 KB. A second BAC library was created in collaboration between the USDA Miami Subtropical Horticultural Research Station and the Clemson BAC Resource Center, which also distributes this resource (http://www.genome.clemson.edu). This library was made using the genotype LCT-EEN 37 and represents approximately 11 genome equivalents with an average insert size of 120 KB. These resources will facilitate genome mapping and gene discovery efforts in the future.

2.2.2 EST Resources for Cacao

Prior to 2002, only a handful of cacao DNA sequences were deposited in the NCBI Genbank database. Since then, the number has increased to the thousands. Expressed Sequence Tags (ESTs) are short stretches of DNA sequences determined from collections of cDNAs synthesized from RNA, and thus represent a snapshot of the genes expressed in a plant. Jones et al. (2002) published the first study using an EST approach to generate large numbers of expressed sequence tags for cacao. In this study, leaf and seed cDNA libraries were sequenced and a unigene set of 1380 sequences were assembled. Additionally, these sequences were used to create a microarray, which was used to demonstrate the specificity of tissue

specific expression of a number of genes. In a similar approach, Verica et al. (2004) used subtracted-normalized cDNA libraries to sequence a unigene set of 1256 gene fragments, some of which were shown to be up-regulated by a known inducer of the plant defense response Nep 1. The authors used suppressive-subtractive hybridization (SSH) and macroarray analysis to identify cacao ESTs representing genes induced by methyl jasmonate, ethylene, the salicylic acid analog benzothiadiazole (BTH), and Nep1. A total of 475 clones that showed an increase of expression level of at least two-fold were designated as up-regulated clones and sequenced. Additionally, 1639 randomly chosen cDNAs were sequenced. After contig assembly (joining of overlapping sequences), 1256 unique non-overlapping sequences (unigenes) (367 contigs and 889 singletons) were obtained, including 330 representing up-regulated genes, and 865 unigenes were assigned to functional classes using BLAST, while 8% of the sequences up-regulated by the defense inducers were similar to defense proteins in other plants. The cDNAs were isolated which are similar to known defense-related genes from other plants included heat shock proteins, NPR1 (a transcriptional regulator that mediates the expression of SA- and JA-responsive genes), and several PR genes, including chitinases, which have been shown to enhance resistance against fungal pathogens. An additional 8% of the up-regulated genes are predicted to play roles in signaling, although their roles in defense are unclear.

A major activity of INGENIC-MOL-BIOL is the establishment of a large EST database for cacao. This project, led by Claire Lanaud of CIRAD and implemented by the French CNS (Centre National de Sequençage), will sequence both ends of 100,000 clones from 28 cDNA libraries that have been contributed in a cooperative effort by the international cacao molecular biology community. All DNA sequencing data and clones will be made publicly available. The project is scheduled to be completed by mid to late 2006.

2.2.3 Microarray Resources for Cacao

A microarray is a specialized microscope slide on which thousands of small droplets of DNA are deposited, creating a substrate which can be used to assay the expression of thousands of genes simultaneously. In a major INGENIC led collaboration, a unigene dataset of all known cacao DNA sequences including those described above was compiled in 2004 (Guiltinan, unpublished). From 6,659

sequences contributed, a 2,781 sequence unigene set was assembled. From this set of sequences, a set of 50-mer oligonucleotides unique for each sequence were designed and synthesized and used to create a spotted microarray (Guiltinan, unpublished). Examples of cacao genes represented on the microarray involved in plant defense response include basic endochitinase, phenylalanine ammonia-lyase, disease resistance protein (CC-NBS-LRR class, quercetin 3-O-methyltransferase 1/flavonol 3-*O*-methyltransferase 1/caffeic acid/5-hydroxyferulic acid *O*-methyltransferase (OMT1), lipoxygenase (LOX2, pathogenesis-related protein, PR-1 protein, caffeoyl-CoA 3-*O*-methyltransferase, isoflavone reductase, chalcone synthase/naringenin-chalcone synthase, ethylene response sensor/ethylene-responsive sensor (ERS), catalase 1, endo-1,4-β-glucanase, pathogen-responsive α-dioxygenase and flavonol synthase 1 (FLS1). These microarrays are currently being evaluated for reproducibility and sensitivity in the author's laboratory and preliminary results are encouraging. In the near future, this array can be used to test the effects of various conditions on cacao gene expression. For example, they can be used to ask questions such as, what genes are induced by endophytic microorganisms? Which genes are induced by pathogen infection and which genes are expressed differently in different genotypes?

2.2.4 Resistance Gene Analogs

Plant disease resistance genes have been identified in a number of model plant species such as rice, maize, tomato and potato, and the results of these studies are beginning to be applied to cacao. Using a degenerate primer-PCR based approach, Lanaud et al. (2004) isolated a set of defense gene analogs from cacao including several kinases similar to those shown in other species to be resistance genes, and several pathogenesis-related genes of the PR class 2 and 5 families. Interestingly, several of these genes appear to be clustered together, and reside on chromosomes in locations close to known QTLs for disease resistance. This is similar to results seen in other species, and may reflect and evolutionary process for the generation of multigene families for disease resistance. Ultimately, this approach could lead to map based cloning of genes implicated in major QTLs for disease resistance.

Using a similar PCR based strategy with degenerate primers based on highly conserved motifs within plant resistance genes, Kuhn et al. (2003) and Borrone et al. (2004) isolated a collection of cacao

genomic fragments containing candidate resistance genes. One key class of pathogen resistance genes isolated, the so called WRKY genes, are also found in many different plant species and have been shown to encode transcription factors involved in resistance to both biotic and abiotic stress. In both cases, multiple gene fragments were isolated and some of these were converted into polymorphic markers suitable for genetic mapping experiments, as discussed in Sections 2.1.1 and 2.1.2. These types of markers are useful in that they encode potentially important genes for resistance, and thus are considered candidate genes. The sequences can also be used in detailed studies to understand the regulation of defense responses in cacao, and the evolution of plant resistance to pathogens.

2.2.5 Floral Development Genes

Comparison of the developmental biology of cacao with *Arabidopsis* can provide insights into the conservation of the molecular mechanisms that determine cell fate, differentiation and signal transduction in cacao. One such system that has been studied is flower development, which is very well characterized in *Arabidopsis*. Research in the authors laboratory carried out by former graduate student J.D. Swanson addressed this comparison in a detailed study of cacao flower development from both the morphological and molecular levels of analysis (Swanson 2005). In addition to documenting the dynamics of cacao flower development, Swanson isolated a series of cacao floral specific genes using degenerate RT-PCR and used these to examine gene expression during early development using *in situ* hybridization. The results showed a remarkable conservation of expression pattern of these regulatory genes, each showing exquisite tissue and stage specificity, highlighting the high degree of conservation of the molecular mechanisms controlling flower development between these two related species. Major facts arriving from this research can be found at the Guiltinan laboratory website (http://guiltinanlab.cas.psu.edu/Research/Cocoa/flowers.htm).

2.3 Plant Tissue Culture and Cryopreservation

Somatic embryogenesis (SE) continues to be the primary method for *in vitro* propagation of cacao (Sondahl et al. 1993; Figueira and Janick 1995; Alemanno et al. 1997; Li et al. 1998; Traore 2000; Maximova et al. 2002). Although it has been applied to a wide variety of genotypes, a large variation of

efficiency has been observed and this appears to be dependant on genotype. One factor, which has been attributed to the difficulty sometimes encountered in cacao tissue culture, is the high concentrations of oxidized phenolic compounds that can accumulate. Alemanno et al. (2003) published a study that investigated the phenolic composition of cocoa floral explants and how they change during culture. Using this knowledge and the methods developed, it may be possible to modify the phenolic composition or change the oxidation status of cacao tissue cultures and thus improve the efficiency and/or genotype variation of cacao embryogenesis. In a more recent study of cacao somatic embryogenesis, Santos et al. (2005) investigated the somatic embryogenesis receptor kinase (SERK) of cacao, which is known to be highly expressed and essential during SE in other plants. These authors found that the cacao genome contains a single functional copy of the SERK gene and that this gene is expressed during SE. Since SERK gene expression is required for SE to proceed in other species, this raises the interesting possibility of using the SERK gene as a marker for SE in cacao, and as a predictor of the potential for SE capacity for different genotypes, although this idea has yet to be tested. A discussion of the use of molecular markers to investigate genetic changes which might occur during somatic embryogenesis in cacao (Rodriguez Lopez et al. 2004) is given in Section 2.1.1.

Cryopreservation of cacao has been long considered a promising application of biotechnology to cacao conservation Considering the loss of wild cacao germplasm due to deforestation, and the risk of loss of germplasm stored in living collections. After development of the SE systems in the mid to late 1990s, cryopreservation systems were developed for cacao (Florin et al. 2000; Fang et al. 2004). In the most recent of these publications, Fang et al. (2004) described a cryopreservation method using encapsulation and dehydration. ABA (abscisic acid) and sugars were used as cryoprotectants during a preculture step. Three genotypes were preserved with a 33% recovery rate, and recovered plants were phenotypically identical to non-cryopreserved SE-derived plants. Work performed at the Nestlé's research center in Tours, France has developed a cryopreservation system based on a different approach (Florin et al. 2000 and V Pétiard / B Florin, person. comm.). This system makes use of hormonally induced floral explants (staminodes) that are then pretreated and frozen prior to SE development. After thawing, the induced staminodes recover and proceed directly to SE, with excellent recovery rates. This process has the advantage of requiring less time prior to freezing and this would be a large advantage for a large-scale germplasm preservation projects.

2.4 Haploid Plants

Doubled haploids have long been considered as an important potential method to generate homozygous lines of cacao for the production of hybrid seed for germplasm propagation (Dublin 1973; Lanaud 1987a, 1987b, 1988a, 1988b, 1988c; Sounigo et al. 2003). However, progress is this area has been slow. A publication by Sounigo et al. (2003) documented the work carried out in this area and summarized the conclusions to date. In this study, twelve doubled haploids were used as parents in field trials in the Ivory Coast, West Africa. The results showed that some of the doubled haploids showed a significantly higher combining value than their parents, demonstrating the potential of this technology for rapid improvement of parents. The authors concluded that many more crosses between doubled haploids would need to be tested to identify potential crosses for improving cacao germplasm. This is a challenging but potentially very rewarding proposition that could lead to the development of seed production systems for cacao germplasm distribution and plantation improvement. Notably, tissue culture has yet to be applied to the development of doubled haploids of cacao.

2.5 Genetic Transformation

Genetic transformation offers a tool for performing basic research on gene structure and function, and also a means to potentially introduce genes from other organisms for crop improvement. In cacao, an initial report demonstrated the susceptibility of cacao cells to *Agrobacterium*, a commonly used bacterium capable of introducing DNA into plant cells (Purdy and Dickstein 1989), and another manuscript described transformation of cacao cells (Sain et al. 1994). Two reports of using particle bombardment (the use of high velocity gold particles to introduce DNA into cultured plant cells) have been reported (Perry et al. 2000; Santos et al. 2002). In both of these studies, the particle bombardment method was used to demonstrate that reporter genes could be introduced into cacao cells and visualized. The method was further refined through optimization of a pretreatment step with osmotic adjustment that increased transient transformation frequencies. However, none of these efforts resulted in the regeneration of transgenic cacao plants. While regeneration of cacao plants through somatic

embryogenesis was possible, as was the introduction of DNA into cacao tissue cultured cells, the ability to regenerate whole plants from the individual transformed cells and cell clusters remained elusive.

Using *A. tumefaciens* based transformation of cultured somatic embryos, a transformation system for cacao capable of producing whole plants was established recently in this laboratory (Maximova et al. 2003). Using the green fluorescent protein marker gene as a way to identify transgenic somatic embryos, the authors recovered a series of transgenic plants that were grown to maturity. The growth and development of the plants was shown to be the same as control, untransformed plants. Transgene insertion, gene expression and stability were all shown to be similar as in other transgenic plants of different species. Details of some of the experiments in which various parameters were optimized were also published separately (Traore 2000; Antunez de Mayolo et al. 2003).

Cacao plants were grown to maturity and the transgenes were shown to be stable through seed to the next generation (Guiltinan, unpublished). More recently, this system has been used to demonstrate the function of a cacao chitinase gene in plant defense against fungal pathogens (Guiltinan lab, in review). This has shown that the cacao transformation system can be useful in analysis of gene function. The utility of this system for crop improvement however remains to be seen, as continued opposition by the public against genetically manipulated crops continues to be debated. Because of this controversy, there are currently no experiments in progress or planned by the author involving a field release of these plants in any country.

2.6 Cacao Pests and Pathogens

2.6.1 Plant-Pathogen Interactions

Researchers are beginning to use newly discovered cacao genes as probes to investigate various aspects of cacao development and, in particular, their responses to interactions with pathogenic organisms. Bailey et al. (2005 a, b) have recently published two manuscripts in which the expression of genes in cacao leaves under different stress or induction conditions were evaluated. These studies have revealed interactions between gene expression changes during leaf development and in response to various inducers such as ethylene, wounding, and methyl jasmonate and pathogen infection. This is a starting point from which to begin dissecting these pathways for a better understanding of the metabolic changes

related to resistance to pathogens in cacao and the underlying differences between cacao genotypes. Such knowledge will provide a fundamental background of resistance mechanisms that will be helpful to assist the planning of breeding programs in the future.

Using a physiological and metabolic approach Aneja and Gianfagna (2001) demonstrated a large increase in caffeine levels in stems and leaves in response to wounding, pathogen infection and by treatment with salicylic acid and benzothiadiazole, inducers of the plant defense responses. Caffeine was shown to inhibit the growth of *Crinipellis perniciosa in vitro*. This finding is consistent with the later finding that the gene encoding caffeine synthase is induced in cacao leaves in response to various defense response inducers (Bailey et al. 2005a, b). In a more comprehensive metabolic study, Scarpari et al. (2005) investigated changes in a large number of metabolic compounds during the infection and development of witches' broom disease. Changes in sugars, amino acids, alkaloids tannins, chlorophyll, fatty acids, glycerol carotenoids, xanthophylls and notably, ethylene were seen during the infection process. Since ethylene is a well known plant hormone, the authors suggest that it may play a key role in the formation of brooms during disease development. This study sets a new precedence for the comprehensive analysis of metabolites in the investigation of cacao/pathogen interactions.

2.6.2 Associations with Beneficial Microorganisms and Biocontrol

Endophytic and epiphytic fungi and bacteria have been shown to live on the surfaces and inside of most plant species studied to date and, in cacao, this has been intensely studies in the past few years because of the potential beneficial use of these organisms as biocontrol agents for protection against cacao pathogens. A recent study characterized some of the naturally occurring endophytic fungi in leaves of cacao growing in Panama (Arnold et al. 2003). A wide range of fungal species was found in cacao leaves, up to 13 different taxa in a single leaf sample, and 344 different morphotaxa were identified overall. The authors also showed that the presence of the endophytes significantly decreased leaf necrosis and mortality when plants were challenged with *Phytophthora* sp., and suggested that this protection may be due to changes in leaf chemistry, a hypothesis consistent with the observations described above with caffeine level changes in cacao leaves.

There are several devastating insect pests of cacao and the worst of which include the myrids of West Africa and the Cocoa Pod Borer (CPB) of Indonesia. In an effort to find potential insecticidal proteins active against the devastating CPB, Santoso et al. (2004) screened 12 Cry proteins from *Bacillus thruringiensis* for activity against CPB larvae. Five of the toxins were shown to be more active, opening the possibility of using these proteins or the genes encoding them as insecticides.

2.6.3 Molecular Analysis of Pathogen Diversity

Molecular biology and biotechnology is also becoming increasingly applied in attempts to understand more about the genetic structure of cacao pests and diseases, and to find ways to fight them. For example, PCR and DNA sequencing have been applied to the analysis of the genetic variability of the cocoa swollen shoot virus, a major pathogen in parts of West Africa (Muller et al. 2001; Muller and Sackey 2005). In the most recent of these studies, the genomes of five isolates were sequenced and compared to the reference sequence. Up to 29.4% sequence variability was detected and these differences could be related to the geographic distributions of the isolates. Interestingly, one of the proposed open reading frames of the CSSV genome, ORF X, varied greatly in size between several of the isolates. The authors suggest that this may indicate that this open reading frame may actually be non-functional.

Using similar approaches, researchers have studied the genetic structures and evolutionary relatedness of a number of cacao fungal pathogens (Niella et al. 2000; Appiah et al. 2003; Chowdappa et al. 2003; de Arruda et al. 2003a, 2003b; Evans et al. 2003; Ploetz et al. 2005). These studies are beginning to shed light on the evolution, geographic distribution and movements of the important cacao fungal pathogens. In an analysis of the fungal pathogen responsible for frosty pod rot (*Moniliophthora roreri*), Phillips (2003) demonstrated the very close relationship of this species to another cacao pathogen, *Crinipellis perniciosa*. Phylogenetic analysis of ribosomal RNA sequences from both nuclear and mitochondrial genes were compared to sequences from other fungi, clearly supporting the re-assignment of *M. roreri* into the Basidiomycete and its close relatedness to *C. perniciosa*. The rapidly advancing understanding of relationships between cacao pathogens and their population structures will contribute to accelerated plant breeding for durable, horizontal resistance.

3 Future Prospects

3.1 Translational Genomics

Theobroma cacao is a simple diploid with ten chromosomes ($2n=2x=20$) and for plant species, a small genome. Published genome size estimates vary from 390 Mb to 415 Mb (Figueira et al. 1992; Couch et al. 1993). *Theobroma cacao* is a member of the order Malvales that includes the important crop plant cotton. Both are members of the Eurosids II group of plants that contains the Brassicales, including *Arabidopsis* (Soltis et al. 2002). The close evolutionary relatedness of these three species suggests that cotton and cacao are excellent crop plants for translational research. What can be learned from the model plant species cotton and *Arabidopsis* and how can this be used to speed up cacao improvement? These questions are a central focus of the author's current research objectives. The recent advances in methodologies and strategies for whole genome sequencing, and the dramatic reduction in costs, makes it highly likely that the cacao genome will be completely sequenced in the next five years. With the full genome sequence, translational genomics will be greatly accelerated, and cacao molecular biologists would be relieved of tedious marker development and gene discovery efforts, and be able to focus more directly on gene function and trait mapping for crop improvement. Bioinformatics will become increasingly important to the future of cacao genetics research.

3.2 Marker Assisted Selection Based Breeding

While the accomplishments reviewed here represent an encouraging beginning towards the applications of genomics to cacao breeding, in order to realize the practical benefits of marker-assisted selection, a much more extensive set of genomics resources are needed. In the future, with such resources available, it can be envisaged that plant breeders will use marker-assisted selection to vastly speed up cacao breeding programs. In addition to genetic markers, metabolic markers, such as caffeine concentrations, could also be used for breeding purposes, once the specific mechanisms important to disease resistance are better understood. In the future, high-throughput metabolic analysis could be used for screening germplasm collections and progeny of breeding trials.

In the near future, local accessions, well adapted to regional environmental and soil conditions, will be crossed with internationally tested genotypes with disease resistance, high yield, and other quality traits. Molecular markers will be used to screen segregating progeny for desired traits while retaining locally desired adaptive phenotypes. Preemptive breeding for resistance to diseases not yet spread to local areas can be included in selection schemes. Flavor and other quality traits will also need to be maintained or improved simultaneously. Gene pyramiding will be used to enhance resistance durability. Selected progenies will be vegetatively propagated through a combination of tissue culture, grafting and rooted cuttings, for distribution to farmers. The genetic diversity of the wild cacao germplasm will be safeguarded in large cryopreservation storehouses.

3.3 Increased Scientific Capacity

To reach these goals, and to have a lasting impact, it is also essential that we increase the scientific capacity of cocoa-producing countries. This can be accomplished through graduate and postgraduate education, exchange visits and training workshops, community building and user-oriented bioinformatics resources. An important complementary component will be the training of scientists in the developed countries in international agriculture in addition to their training in plant molecular genetics, genomics, and bioinformatics, to form a future cadre of well educated plant scientists versed and experienced in international agricultural issues. While cocoa does not fight hunger directly as it is an export crop, it is an important cash crop for farmers who otherwise can grow most of their own food for sustenance, but need money to improve their quality of life or to purchase food they cannot grow. Our investments in research and education in the basic plant sciences, genetics and molecular biology of cacao now, will help to ensure the well being of the cacao farmer in the future.

Acknowledgements

The author would like to thank many colleagues who responded to requests for reprints and suggestions for articles which may otherwise have been overlooked and to Drs. Jose R. Peralta-Videa, Antonio Figueira, Lizz Johnson and Basil Bartley for critical reading and comments on this manuscript.

References

Alemanno L, Berthouly M, Michaux-Ferriere N (1997) A comparison between *Theobroma cacao* L. zygotic embryogenesis and somatic embryogenesis from floral explants. In vitro Cell Dev Biol-Plant 33: 163-172

Alemanno L, Ramos T, Gargadenec A, Andary C, Ferriere N (2003) Localization and identification of phenolic compounds in *Theobroma cacao* L. somatic embryogenesis. Ann Bot 92: 613-623

Alverson WS, Whitlock B, A., 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 (2001) Induction and accumulation of caffeine in young, actively growing leaves of cocoa (*Theobroma cacao* L.) by wounding or infection with *Crinipellis perniciosa*. Physiol Mol Plant Pathol 59: 13-16

Antunez de Mayolo G, Maximova SN, Pishak S, Guiltinan MJ (2003) Moxalactam as a counter-selection antibiotic for *Agrobacterium*-mediated transformation and its positive effects on *Theobroma cacao* somatic embryogenesis. Plant Sci 164: 607-615

Appiah AA, Flood J, Archer SA, Bridge PD (2004) Molecular analysis of the major *Phytophthora* species on cocoa. Plant Pathol 53: 209-219

Appiah AA, Flood J, Bridge PD, Archer SA (2003) Inter- and intraspecific morphometric variation and characterization of *Phytophthora* isolates from cocoa. Plant Pathol 52: 168-180

Arnold AE, Mejia LC, Kyllo D, Rojas EI, Maynard Z, Robbins N, Herre EA (2003) Fungal endophytes limit pathogen damage in a tropical tree. Proc Natl Acad Sci USA 100: 15649-15654

Atkinson MD, Withers LA, Simpson MJA (1986) Characterization of cacao germplasm using isoenzyme markers.1. A preliminary survey of diversity using starch gel electrophoresis and standardisation of the procedure. Euphytica 35: 74-750

Bailey B, Bae H, Strem MD, Antunez de Mayolo G, Guiltinan MJ (2005a) Gene expression in leaves of *Theobroma cacao* in response to mechanical wounding, ethylene, or methyl jasmonate. Plant Sci 128: 1247-1258

Bailey B, Bae H, Strem MD, Mayolo GA, Guiltinan M, Verica JA, Maximova S, Bowers JH (2005b) Developmental expression of stress response genes in *Theobroma cacao* leaves and their response to Nep1 treatment and a compatible infection by *Phytophthora megakarya*. Plant Physiol Biochem 43: 611-622

Bartley BGD (2005) The genetic diversity of cacao and its utilization. CAB International, Wallingford, UK

Bayer C, Fay MF, De Bruijn AY, Savolainen V, Morton CM, Kubitzki K, Alverson WS, Chase MW (1999) Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: A combined analysis of plastid atpB and rbcL DNA sequences. Bot J Linn Soc 129: 267 - 303

Bennett AB (2003) Out of the Amazon: *Theobroma cacao* enters the genomic era. Trends Plant Sci 8: 561-563

Borrone JW, Kuhn DN, Schnell RJ (2004) Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. Theor Appl Genet 109: 495-507

Bowers JH, Bailey BA, Hebbar PK, Sanogo S, Lumsden RD (2001) The impact of plant diseases on worldwide chocolate production (http://www.plantmanagementnetwork.org/pub/php/review/cacao/)

Brown JS, Kuhn DN, Lopez U, Schnell RJ (2005) Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an F2 population using SSR markers and candidate genes. J Am Soc Hortic Sci 130: 366-373

Charters YM, Wilkinson MJ (2000) The use of self-pollinated progenies as "in-groups" for the genetic characterization of cocoa germplasm. Theor Appl Genet 100: 160-166

Chowdappa P, Brayford D, Smith J, Flood J (2003) Molecular discrimination of *Phytophthora* isolates on cocoa and their relationship with coconut, black pepper and bell pepper isolates based on rDNA repeat and AFLP fingerprints. Curr Sci 84: 1235-1237

Clement D, Lanaud C, Sabau X, Fouet O, Le Cunff L, Ruiz E, Risterucci AM, Glaszmann JC, Piffanelli P (2004) Creation of BAC genomic resources for cocoa (*Theobroma cacao* L.) for physical mapping of RGA containing BAC clones. Theor Appl Genet 108: 1627-1634

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

Clement D, Risterucci AM, Motamayor JC, Ngoran J, Lanaud C (2003b) Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L. Genome 46: 103-111

Coe SD, Coe MD (1996) The true history of chocolate. Thames and Hudson, New York, USA

Couch JA, Zintel HA, Fritz PJ (1993) The genome of the tropical tree *Theobroma cacao* L. Mol Gen Genet 237: 123-128

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

Cruz M, Whitkus R, Gomez-Pompa A, Monta-Bravo L (1995) Origins of cacao cultivation. Nature 375: 542-543

Cryer NC, Fenn MGE, Turnbull CJ, Wilkinson MJ (2005) Allelic size standards and reference genotypes to unify international cocoa (*Theobroma cacao* L.) microsatellite data. Genetic Resour Crop Evol (in Press)

Day RK (1984) Population dynamics of cocoa pod borer *Acrocercops cramerella*: Importance of host plant cropping cycle. In: International Conference on Cocoa and Coconuts, Kuala Lumpur, Malaysia, pp 1-9

de Arruda MC, Ferreira MA, Miller RN, Resende ML, Felipe MS (2003a) Nuclear and mitochondrial rDNA variability in *Crinipellis perniciosa* from different geographic origins and hosts. Mycol Res 107: 25-37

de Arruda MC, Miller RN, Ferreira MASV, Felipe MS (2003b) Comparison of *Crinipellis perniciosa* isolates from Brazil by ERIC repetitive element sequence-based PCR genomic fingerprinting. Plant Pathol 52: 236-244

Dias LAS (2001) Genetic iImprovement of cacao. (http://ecoport.org/ep?SearchType=earticleView&earticleId=197)

Dublin P (1973) Phytogenetique - Haploidie chez *Theobroma cacao*. Les Comptes Rendus de l'Académie des Sci 276: 757-759

Efombagn MIB, Marelli JP, Ducamp M, Cilas C, Nyasse S, Vefonge D (2004) Effects of fruiting traits on the field resistance of cocoa (*Theobroma cacao* L.) clones to *Phytophthora megakarya*. Phytopathology 152: 557-562

Emch M (2003) The human ecology of Mayan cacao farming in Belize. Human Ecol 31: 111-132

Evans HC, Holmes KA, Reid AP (2003) Phylogeny of the frosty pod rot pathogen of cocoa. Plant Pathol 52: 476-485

Faleiro FG, Pires JL, Monteiro WR, Lopes UV, Yamada MM, Piedra AG, Moura AD, Arevalo-Gardini EA, Marques J, Gramacho K, Faleiro A, Santos M (2004) Variability in cacao accessions from the Brazilian, Ecuadorian, and Peruvian Amazons based on molecular markers. Crop Breed Appl Biotechnol 4: 227-233

Fang JY, Wetten A, Hadley P (2004) Cryopreservation of cocoa (*Theobroma cacao* L.) somatic embryos for long-term germplasm storage. Plant Sci 166: 669-675

Figueira A, Alemanno L (2005) *Theobroma cacao* (Cacao). In: Litz RE (ed) Biotechnology of fruit and nut crops. CAB International Biosciences, Wallingford, UK, pp 639-669

Figueira A, Janick J (1995) Somatic embryogenesis in cacao (*Theobroma cacao* L.). In: Newton R (ed) Somatic embryogenesis in woody plants, vol 2. Klumer Academic Publishers, Dordrecht, The Netherlands, pp 291-310

Figueira A, Janick J, Goldsbrough P (1992) Genome size and DNA polymorphism in *Theobroma cacao*. J Am Hortic Sci 117: 673-677

Figueira A, Janick J, Levy M, Goldsbrough P (1994) Reexamining the classification of *Theobroma cacao* L. using molecular markers. J Am Hortic Sci 119: 1073-1082

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

Florin B, Brulard E, Pétiard V (2000) In vitro cryopreservation of cacao genetic resources. In: Engelman F; Tagaki H (eds) Cryopreservation of tropical plant germplasm. Japanese International Research Centre for Agricultural Sciences and IPGRI, Rome, Italy, pp 344-347

Fritz PJ, Osei JK, Goodin MM and Furtek DB (1991) Molecular and biochemical markers for genetic analysis of *Theobroma cacao*. Proc. Inter. Cocoa Res. Conf.-Kuala Lumpur, Malaysia: 391-400

Gomes LMC, Melo GRP, Faleiro FG, Silva SDM, Araujo RC, Bahia RC, Morales MG, Ahnert D (2000) Genetic diversity of *Crinipellis perniciosa* from the south region of Bahia-Brazil using RAPD molecular markers. Fitopatologia Brasileira 25: 377

Guyton B, Lumsden R, Matlick BK (2003) Strategic plan for sustainable cocoa production. Manufact Confect 83: 55-60

Hanna AD (1954) Application of a systemic insecticide by trunk implantation to control a mealybug vector of the cacao swollen shoot virus. Nature 173: 730-731

Hervé L, Djiekpor E, Jacquemond M (1991) Characterization of the genome of cacao swollen shoot virus. J Gen Virol 72: 1735-1739

Iwaro AD, Bekele FL, Butler DR (2003) Evaluation and utilisation of cacao (*Theobroma cacao* L.) germplasm at the International Cocoa Genebank, Trinidad. Euphytica 130: 207-221

Johnson L (2003) INGENIC Workshop, Cocoa genomics group. Gro Cocoa 4: 4-5 (http://www.cabicommodities.org/Acc/ACCrc/PDFFiles/GROC/GROC.htm)

Jones PG, Allaway D, Gilmour DM, Harris C, Rankin D, Retzel ER, Jones CA (2002) Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta 216: 255-264

Keane PJ (1992) Diseases and pests of cocoa: An overview. Cocoa pest and disease management in Southeast Asia and Australasia, FAO Plant Production and Protection Paper 112: 1-12

Kuhn DN, Borrone J, Meerow AW, Motamayor JC, Brown JS, Schnell RJ (2005) Single-strand conformation polymorphism analysis of candidate genes for reliable identification of alleles by capillary array electrophoresis. Electrophoresis 26: 112-125

Kuhn DN, Heath M, Wisser RJ, Meerow A, Brown JS, Lopes U, Schnell RJ (2003) Resistance gene homologues in *Theobroma cacao* as useful genetic markers. Theor Appl Genet 107: 191-202

Kuhn DN, Schnell RJ (2005) Use of capillary array electrophoresis single strand conformational polymorphism analysis to estimate genetic diversity of candidate genes in germplasm collections. Meth Enzymol 395: 238-258

Lanaud C (1987a) Doubled haploids of cocoa (*Theobroma cacao* L.) I. Observations of fertility. Plant Breed 99: 187-195

Lanaud C (1987b) Doubled haploids of cocoa (*Theobroma cacao* L.) II. Observations of monogenic and polygenic characters. Plant Breed 99: 196-202

Lanaud C (1988a) Behavior in crossing of the haploids of cacao-trees (*Theobroma cocoa*). Can J Bot 66: 1986-1992

Lanaud C (1988b) Comportement en croisement des haploides de cacaoyers (*Theobroma cacao*). Can J Bot 66: 1986-1992

Lanaud C (1988c) Origin of haploids and semigamy in *Theobroma cacao* L. Euphytica 38: 221-228

Lanaud C, Kebe I, Risterucci AM, Clement D, N'Goran JKA, Grivet L, Tahi M, Cilas C, Pieretti I, Eskes AB, Despreaux D (1996) Mapping quantitative trait loci (QTL) for resistance to *Phytophthora palmivora* in *T. cacao*. 12th International Cocoa Research Conference Salvador, Bahia, Brazil, pp 99-105

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

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

Lerceteau E, Robert T, Petiard 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

Li Z, Traore A, Maximova S, Guiltinan MJ (1998) Somatic embryogenesis and plant regeneration from floral explants of cacao (*Theobroma cacao L.*) using thidiazuron. In vitro Cell Dev Biol-Plant 34: 293-299

Marita JM, Nienhuuis J, Pires JL, Aitken WM (2001) Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches broom disease resistance. Crop Sci 41: 1305-1316

Maximova S, Alemanno, Young A, Ferriere N, Traore A, Guiltinan M (2002) Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L. In vitro Cell Dev Biol-Plant 38: 252-259

Maximova S, Miller C, Antunez de Mayolo G, Pishak S, Young A, Guiltinan MJ (2003) Stable transformation of *Theobroma cacao* L. and influence of matrix attachment regions on GFP expression. Plant Cell Rep 21: 872-883

Morais R (2005) The Gnomes of cocoa. In: Forbes Magazine, March 14th, New Your, USA, pp 110-112

Motamayor JC, Risterucci AM, Heath M, Lanaud C (2003) Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. Heredity 91: 322-330

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

Motilal LA, Sounigo O, Thevenin JM, Howell MH, Pieretti I, Risterucci AM, Noyer JL, Lanaud C (2000) *Theobroma cacao* L. : Genome map and QTLs for *Phytophthora palmivora* resistance. In: 13th Annual Cocoa Research Conference, Kota Kinabalu, Malaysia

Muller E, Jacquot E, Yot P (2001) Early detection of cacao swollen shoot virus using the polymerase chain reaction. J Virol Meth 93: 15-22

Muller E, Sackey S (2005) Molecular variability analysis of five new complete cacao swollen shoot virus genomic sequences. Arch Virol 150: 53-66

Niella GR, Resende MLV, De Castro HA, Figueira AR, Silva SDM, Araujo IS, Gomes LMC, Faleiro FG (2000) Genetic diversity of the monosporic isolates of *Crinipellis perniciosa* from different Brazilean States, using RAPD markers. Fitopatol Brasil 25: 400

Nyasse S, Despreaux D, Cilas C (2002) Validity of a leaf inoculation test to assess the resistance to *Phytophthora megakarya* in a cocoa (*Theobroma cacao* L.) diallel mating design. Euphytica 123: 395-399

Opoku IY, Akrofi AY, Appiah AA (2002) Shade trees are alternative hosts of the cocoa pathogen *Phytophthora megakarya*. Crop Prot 21: 629-634

Pereira JL, Ram A, Defigueiredo JM, Dealmeida LCC (1990) First occurrence of witches' broom disease in the principal cocoa-growing region of Brazil. Trop Agric 67: 188-189

Perry M, Power J, Lowe K, MR D (2000) Biolistic transformation of cacao (*Theobroma cacao* L.). Trop Agric 77: 64-66

Phillips W (2003) Origin, biogeography, genetic diversity and taxonomic affinities of the cacao fungus *Moniliophthora roreri* as determined using molecular, phytopathological and morpho-physiologycal evidence. Ph.D. thesis. University of Reading, Reading, UK

Piasentin F, Klare-Repnik L (2004) Biodiversity conservation and cocoa agroforests. Gro Cocoa 5: 7-8 (http://www.cabi-commodities.org/Acc/ACCrc/PDFFiles/GROC/GROC.htm)

Ploetz RC, Schnell RJ, Ying Z, Zheng Q, Olano CT, Johnson E, Motamayor JC (2005) Molecular diversity in *Crinipellis perniciosa* with AFLPs. Euro J Plant Pathol 111: 317-326

Pound FJ (1940) Witches' broom resistance in cacao. Trop Agric 17: 6 - 8

Pugh T, Fouet O, Risterucci AM, Brottier P, Abouladze M, Deletrez C, Courtois B, Clement D, Larmande P, Ngoran JAK (2004) A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. Theor Appl Genet 108: 1151-1161

Purdy LH, Dickstein ER (1989) *Theobroma cacao*, a host for *Agrobacterium tumefaciens*. Plant Dis 73: 638-639

Queiroz VT, Guimaraes CT, Anhert D, Schuster I, Daher RT, Pereira MG, Miranda VRM, Loguercio LL, Barros EG, Moreira MA (2003) Identification of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches' broom disease. Plant Breed 122: 268-272

Rice RA, Greenberg R (2003) The Chocolate Tree: Growing cacao in the forest. Nat His 112: 36-43

Risterucci AM, Grivet L, Ngoran JA, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of *Theobroma cacao* L. Theor Appl Genet 101: 948-955

Risterucci AM, Paulin D, Ducamp M, N'Goran JAK, Lanaud C (2003) Identification of QTLs related to cocoa resistance to three species of *Phytophthora*. Theor Appl Genet 108: 168-174

Rodriguez Lopez CM, Wetten AC, Wilkinson MJ (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 CM, Schnell RJ, Kuhn DN (1995) Inheritance of random amplified polymorphic DNA (RAPD) markers in *Theobroma cacao* L. J Am Soc Hortic Sci 120: 681-686

Ruf F, Zadi H (2003) Cocoa: From deforestation to reforestation. (http://nationalzoo.si.edu/ConservationAndScience/MigratoryBirds/Research/Cacao/ruf.cfm).

Sain SL, Oduro KK, Furtek DB (1994) Genetic transformation of cocoa leaf cells using *Agrobacterium tumefaciens*. Plant Cell Tissue Organ Cult 37: 342-351

Santos M, Albuquerque de Barros E, Tinoco M, Brasileiro A, Aragão F (2002) Repetitive somatic embryogenesis in cacao and optimisation of gene expression by particle bombardment. J Plant Biotechnol 4: 71-76

Santos MO, Romano E, Yotoko K, Tinoco M, Dias B, Aragao F (2005) Characterisation of the cacao *somatic embryogenesis receptor-like kinase* (SERK) gene expressed during somatic embryogenesis. Plant Sci 168: 723-729

Santoso D, Chaidamsari T, Wiryadiputra S, de Maagd RA (2004) Activity of *Bacillus thuringiensis* toxins against cocoa pod borer larvae. Pest Manag Sci 60: 735-738

Saunders JA, Mischke S, Leamy EA, Hemeida AA (2004) Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. Theor Appl Genet 110: 41-47

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AWV, Schiavinato MA, Cascardo JCM, Pereira GAG (2005) Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa*. J Exp Bot 56: 865-877

Schnell RJ, Olano CT, Brown JS, Meerow AW, Cervantes-Martinez C, Nagai C, Motamayor JC (2005) Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. J Am Soc Hortic Sci 130: 181-190

Sereno ML, Albuquerque PSB, Vencovsky RAF (2005) Genetic diversity and natural population structure of cacao (*Theobroma cacao* L.) from the Brazilian Amazon evaluated by microsatellite markers. Conser Genet (in press)

Soltis DE, Soltis PS, Albert VA, Oppenheimer DG, DePamphilis CW, Ma H, Frohlich MW, Theissen G (2002) Missing links: the genetic architecture of flower and floral diversification. Trends Plant Sci 7: 22-31

Sondahl MR, Liu S, Bellato C, Bragin A (1993) Cacao somatic embryogenesis. Acta Hort 336: 245-248

Sounigo O, Lachenaud P, Bastide P, Cilas C, N'Goran J, Lanaud C (2003) Assessment of the value of doubled haploids as progenitors in cocoa (*Theobroma cacao* L.) breeding. J Appl Genet 44: 339-353

Sousa Silva CRS, Figueira A (2005) Phylogenetic analysis of *Theobroma* (Sterculiaceae) based on Kunitz-like trypsin inhibitor sequences. Plant Syst Evol 250: 93-104

Swanson JD (2005) Flower development in *Theobroma cacao* L.: An assessment of morphological and molecular conservation of floral development between *Arabidopsis thaliana* and *Theobroma cacao*. PhD Dissertation. Pennsylvania State University, University Park, Pennsylvania, USA

Traore A (2000) Somatic embryogenesis, embryo conversion, micropropagation and factors affecting genetic transformation of *Theobroma* cacao L. PhD thesis. Pennsylvania State University, University Park, Pennsylvania, USA

Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ (2004) Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. Plant Cell Rep 23: 404-413

Whitkus R, Delacruz M, Motabravo L, Gomezpompa A (1998) Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. Theor Appl Genet 96: 621-627

Whitlock B, A., Bayer C, Baum DA (2001) Phylogenetic relationships and floral evolution of the Byttnerioideae ("Sterculiaceae" or Malvaceae s.l.) based on sequences of the chloroplast gene, *ndhF*. Syst Bot 26: 420-437

Wilde J, Waugh R, Powell W (1991) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor Appl Genet 83: 871-877

Wood GAR, Lass RA (1985) Cocoa, 4th Edn. Longman Scientific & Technical, New York, USA