PROCEEDINGS

OF THE

26th ANNUAL MEETING

July 29 to August 4, 1990

Mayaguez, Puerto Rico

Published by:

Caribbean Food Crops Society

with the cooperation of the USDA-ARS-TARS

Mayaguez, Puerto Rico

THE AMERICAN COCOA RESEARCH INSTITUTE CACAO BIOTECHNOLOGY PROGRAM AT THE PENNSYLVANIA STATE UNIVERSITY

J. Couch, D. Chung, H. Dodo, W. Flynn, P. Fritz, D. Furtek, L. Glicenstein, G. Gora-Maslak, L. McHenry, M. Mirazon, J. Osei, T. Snyder, H. Tai, and H. Yeoh

Department of Food Science, 215 Borland Laboratory
The Pennsylvania State University
University Park, PA 16802, USA

ABSTRACT

Cacao molecular biology research at Penn State University is directed at six areas: 1) regenerating trees from tissue culture, 2) genetically transforming the tree, 3) cloning lipid biosynthetic genes, 4) characterizing the genomes, 5) isolating genes that may confer fungal disease resistance, and 6) creating an RFLP linkage map. An accomplishment with tissue culture studies is developing a method to readily induce "break" in excised lateral buds. These results will lead to the rapid propagation of selected cultivars. Because a long-term goal is to engineer the tree for a hard cocoa butter regardless of the climate, we have cloned a putative glycerol-3-phosphate acyltransferase gene, the first one in the cocoa butter biosynthetic pathway. Another project is to eventually engineer cacao to be more resistant to fungal diseases; thus cocoa genomic library is being screened for the gene encoding beta-1,3-glucanase, a hydrolase which degrades fungal cell walls. To more thoroughly understand the molecular biology of cacao, the nuclear genome is being characterized. Preliminary evidence suggests the cacao nuclear genome is smaller than that of most higher plants. Finally, efforts have begun to genetically transform the tree, and to develop a map of the cacao genome based on RFLPs.

INTRODUCTION

In 1986, the American Cocoa Research Institute (ACRI), in cooperation with Penn State University (PSU), established an endowed program in the College of Agriculture, Department of Food Science, to study the molecular biology of Theobroma cacao, the cocoa plant. Within the broad mandate of using the most recent advances in plant science to develop methods for producing improved cocoa trees, three areas of research emphasis were targetted for study: (1) micropropagation, (2) disease resistance, and (3) bean quality components. A better understanding of cocoa molecular genetics was seen as needed to provide a sound foundation for progress in all these areas. The present report summarizes published work, research in progress, and future plans.

Because of conflicting reports in the literature, a cytogenetic study was undertaken to establish the ploidy level of T. cacao (Glicenstein and Fritz, 1989 A). It was concluded that cocoa is a diploid taxon based on (1) strict bivalent pairing in over 400 meiotic cells in the diplotene, early diakinesis, and metaphase I stages, (2) only two nucleolus organizer region chromosomes per nucleus and a maximum of two nucleoli per root tip nucleus, (3) only two chromosomes associated with the nucleolus at the diplotene stage and early diakinesis, and (4) in situ hybridization studies with a lowcopy cocoa DNA probe. In the course of the cytological studies, a series of photomicrographs illustrating the stages of T. cacao meiosis during pollen development were compiled and published (Glicenstein and Fritz, 1989 B). The haploid number of chromosomes is 10.

Studies on nuclear (genomic) DNA revealed a lower than expected degree of methylation, and a melting temperature of 85°C, from which a G+C content of 38-39% was calculated. Reassociation kinetic studies to determine genome size were confounded by impurities in isolated DNA, so an original method for isolating DNA from cocoa leaves was developed (Couch and Fritz, 1990). The method was used successfully for several other plants having high levels of leaf polyphenolic compounds. Preliminary indications from flow cytometry, reassociation kinetics, and genomic library screening experiments suggest that the cocoa genome may be smaller and less repetitive than most higher plants, but the data have yet to be confirmed with the appropriate controls.

Isolation of pure chloroplast DNA from cocoa leaves in amounts sufficient for constructing restriction endonuclease maps was found to be a most difficult task, generally averaging about 60ng per g of leaf, almost three times lower than the lowest value reported for over 100 species of angiosperms, gymosperms, and ferns (Palmer, 1982). There appear to be several reasons for low yields. First, cocoa leaf cells are extremely small and contain only about three chloroplasts (Baker et al., 1975), hence the amount of DNA per cell is low. Second, the amount of nuclear DNA per leaf cell is several hundred times greater than the amount of cpDNA, and clean separation of nuclei from chloroplasts is troublesome. Finally, cocoa leaf cells contain large amounts of polyphenols whose physical properties make them extremely difficult to remove from DNA. In order to ensure an adequate amount of cpDNA for mapping and population diversity studies, we turned to molecular cloning. Close to 109 kilobase pairs of unique cocoa cpDNA have been cloned, 88.6 in a lambda bacteriophage vector, and 20.6 in plasmid pBR322. The figure of 109kbp is consistent with the size determined by transmission electron micrograph studies (Chung, 1988) and by the sum of fragments obtained by digestion with four different restriction endonucleases, suggesting that this is close to the full length of cocoa cpDNA. Thus, the cocoa chloroplast genome

appears to be among the smallest reported for any plant (Yeoh et al., 1990).

The First International Cocoa Genome Mapping Meeting was held at Penn State University in March, 1990. The purpose of the meeting was to explore ways to devise a strategy for constructing a cocoa genetic linkage map using isozyme and restriction fragment length polymorphism (RFLP) technology. It was agreed that the magnitude of the task was so great that in order to construct even a low-density map it was unrealistic to think that a single laboratory could complete the job in a reasonable length of time, so the concept of an international network was enthusiastically endorsed. Initially, the network will involve direct participation by six laboratories, one each in Costa Rica, Cote d'Ivoire, France, and Scotland, and two in the United States. A two map strategy is being planned because different segregating populations are available in Cote d'Ivoire and Costa Rica, and each map will serve as a backup for the other. Probe preparation from cocoa genomic libraries is currently underway in the Tropical Species Genome Analysis Laboratory (AGETROP), Montpellier, France, and in the PSU-ACRI lab. Cocoa RFLP studies have been underway at Penn State and at the Center for Tropical Agriculture Research and Teaching (CATIE) for several years (Mirazon, 1988; Mirazon et al., 1989; Febres, 1989).

Cocoa butter is the major storage product of mature cocoa seeds, making up more than 50% of the dry weight. It is at least 96% triacylglycerols (TG) consisting of 40% POS (palmiticoleic-stearic) TGs, 29% SOS, and 15% POP. TG biosynthesis, beginning with glycerol-3-phosphate, requires four enzymes, three acyltransferases and one phosphatase. This relatively simple metabolic pathway encouraged research in the PSU-ACRI cocoa molecular biology laboratory aimed at devising ways to use genetic engineering to make changes in the plant's genome designed to improve the quality and/or quantity of the cocoa butter. Studies on the first enzyme in the biosynthetic pathway have been underway for several years (Fritz, et al., 1986). The enzyme, glycerol-3-phosphate acyltransferase (G-3-PAT), was found to be present in significant amounts in the postmicrosomal supernatant fraction of mature cocoa seed extracts. Enrichments exceeding 1,000 fold were achieved and the purified enzyme was found to have a native molecular weight of 200,000 Daltons, and to be composed of 10, apparently identical, 20,000 Dalton subunits. The highly purified enzyme carries ten moles of fatty acid (palmitic and stearic) per mole of enzyme even after hydrophobic chromatography on octyl sepharose columns, or after polyacrylamide gel electrophoresis and isoelectric focusing. Thin layer chromatography of reaction products confirmed that the enzyme catalyzed transfer of fatty acids exclusively to the one position of glycerol-3-phosphate.

A near full-length putative G-3-PAT cDNA clone has been obtained from a cocoa seed lambda gt10 cDNA library using an oligonucleotide probe based on the N-terminal amino acid sequence (Tai, 1990). The cDNA clone is judged to be near full-length because it contains (1) 1006 bp excluding the poly A tail (approximately 60 As), (2) a 5' initiation codon beginning the only open reading frame capable of encoding a 20kD protein, (3) 75 bp encoding 25 amino acids, typically found in plant signal peptides, preceding the N-terminal amino acid of the mature protein, (4) an exact match of 20 amino acids between those deduced from the DNA sequence and those determined from N-terminal amino acid sequencing of the purified protein, and (5) 636 bp between the N-terminal amino acid codon and the translation termination codon, enough to code for 212 amino acids with a calculated molecular weight of 20,361 Daltons, consistent with the estimated 20,000 Daltons from SDS PAGE of the purified protein. Experiments to express the cocoa gene in transgenic tobacco plants are currently in progress. Also in progress are experiments to isolate and characterize the cocoa G-3-PAT gene from a genomic library.

Cocoa Black Pod disease, a fungal pod rot caused by Phytophtora species, occurs throughout the range of T. cacao and results in 10-30% annual crop loss worldwide. During the past several decades tropical research stations have been screening cacao for resistant phenotypes, a process that is slow and has often yielded conflicting information. Although some resistant clones are recognized, little information about the molecular basis of disease resistance is available. Amounts of B-1,3-glucanases and chitinases have been demonstrated in a number of plants to increase dramatically in response to fungal infection. These enzymes have fungitoxic, hydrolytic activity that directly degrades fungal cell walls and releases oligosaccharides that are elicitors of defense responses in plants. Resistant plants accumulate the hydrolytic enzymes, and other defense compounds such as phytoalexins, at a much faster rate than susceptible plants. One goal of the PSU-ACRI cocoa molecular biology laboratory is to obtain information on factors involved in the T. cacao: P. palmivora host:pathogen interaction for the purpose of understanding the regulation of such factors and in the hope that disease resistance genes can be identified that could be useful in genetically engineering cocoa plants resistant to the disease. The cocoa genomic library was screened with tobacco B-1,3-glucanase and chitinase genes (gifts from Dr. Fred Meins, Basel, Switzerland) using hybridization conditions adjusted for heterologous probes. Positive clones were obtained and DNA from these clones is being analyzed to locate the fragment encoding the gene for the glucanase.

A micropropagation system for cocoa is needed so that when superior trees are produced by conventional breeding, by genetic engineering, or by a combination of both, a method to produce large numbers of these desirable clones will be available. An in

vitro vegetative propagation system based on axillary bud growth has been developed (Flynn et al., 1990), but work to improve the system is proceeding. Explant response to current culture strategy now allows subdivisions of primary shoots to yield multiple explants each producing additional shoots. A method for obtaining large numbers of orthotropic buds suitable for clonal propagation has also been developed (Glicenstein et al.,1990).

To create transgenic cocoa plants containing foreign genes conferring valuable properties, such as disease resistance, altered lipid composition, reduced theobromine content, etc., an efficient genetic transformation and plant regeneration system must be devised. Thus, we will be attempting to genetically transform cocoa somatic embryos by either Agrobacterium tumefaciens or high-velocity microprojectiles (biolistics). Transgenic plants will then be regenerated as described by Duhem et al., (1989). Our first test experiments will be to create cocoa plants resistant to the antibiotic kanamycin, and expressing the bacterial gene for B-glucuronidase, a gene commonly used for transgenic plant studies.

REFERENCES

Baker, N.R., Hardwick, K., and Jones, P. 1975. Biochemical and physiological aspects of leaf development in cocoa (Theobroma cacao) II. Development of chloroplast ultrastructure and carotenoids. New Phytol. 75:513-518.

Chung, D. 1988. Theobroma cacao Chloroplast DNA: Isolation and characterization. Master of Science Thesis, Penn State University.

Couch, J.A., and Fritz, P.J. 1990, Isolation of DNA from plants high in polyphenolics. Plant Molec. Biol. Rep. 8:8-12.

Duhem, K., Lemercier, N., and Boxus, P. 1989. New data on induction and development of somatic embryos of Theobroma cacao Cafe' Cacao The' 33:9-14.

Febres, V.J. 1989. Marcadores genéticos en cacao para resistencia a la moniliasis, Anteproyecto de Tesis, CATIE.

Flynn, W.P., Glicenstein, L.J., and Fritz, P.J. (1990), Theobroma cacao L.: an axillary bud in vitro propagation procedure. Plant Cell, Tiss. and Org. Cult. 20:111-117.

Fritz, P.J., Kauffman, J.M., Robertson, C.A., and Wilson, M.R. (1986), Cocoa butter biosynthesis: Purification and characterization of a soluble sn-glycerol-3-phosphate acyltransferase from cocoa seeds. J. Biol. Chem. 261:194-199.

Glicenstein, L.J., Flynn, W.P., and Fritz, P.J. (1990), Clonal propagation of Theobroma cacao L. Cocoa Growers Bull. (in press).

Glicenstein, L.J., and Fritz, P.J. 1989A. Ploidy level in Theobroma cacao L. J. Hered. 80:464-467.

Glicenstein, L.J., and Fritz, P.J. 1989B. Meiosis in Theobroma cacao L. Turrialba 39:497-501.

Mirazon, M. 1988. Theobroma cacao: Genetic analysis using restriction fragment length polymorphisms. Master of Science Thesis, Penn State University.

Mirazon, M., Gora-Maslak, G., McHenry, L., and Fritz, P.J. 1989. Protocols for RFLP analysis. Turrialba 39:519-525.

Palmer, J.D. (1982), Physical and gene mapping of chloroplast DNA from Atriplex triangularis and Cucumis sativa. Nucl. Acids Res.190:1593-1605.

Tai, H. 1990. Theobroma cacao L.: Molecular cloning and sequencing of a seed cDNA putative glycerol-3-phosphate acyltransferase gene. Doctor of Philosophy Thesis, Penn State University.

Yeoh, H.H., Chung, D., and Fritz, P.J. 1990. Theobroma cacao Chloroplast DNA: Isolation, molecular cloning, and characterization. Cafe' Cacao The' (in press).