BIOTECHNOLOGY FOR COCOA POD BORER RESISTANCE IN COCOA

Tetty Chaidamsari

Promotor:

Prof. dr. ir. E. Jacobsen
Hoogleraar in de Plantenveredeling,
in het bijzonder de genetische variatie en reproductie
Wageningen Universiteit

Co-promotor:

Dr. R. A. de Maagd
Senior Onderzoeker,
Business Unit Bioscience, Plant Research International

Promotiecommissie:

Prof. dr. W. J. Stiekema
Wageningen Universiteit

Prof. dr. ir. G. C. Angenent
Plant Research International/Radboud Universiteit Nijmegen

Prof. dr. C. Mariani
Radboud Universiteit Nijmegen

Prof. dr. R. G. F. Visser
Wageningen Universiteit

Dit onderzoek is uitgevoerd binnen de onderzoekschool
Experimentele Plantenwetenschappen (EPW)

BIOTECHNOLOGY FOR COCOA POD BORER RESISTANCE IN COCOA

Tetty Chaidamsari

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
Prof. dr. ir. L. Speelman
in het openbaar te verdedigen
op dinsdag 28 juni 2005
des namiddags te vier uur in de Aula

Chaidamsari, T. 2005
Biotechnology for cocoa pod borer resistance in cocoa
Ph.D. thesis Wageningen University, Wageningen, The Netherlands- with summaries in English, Dutch and Indonesian.
Plant Research International, P.O. Box 16, 6700 AA Wageningen,
The Netherlands.

ISBN: 90-8504-235-6

Contents

Chapter 1	Introduction	1
Chapter 2	Activity of *Bacillus thuringiensis* toxins against cocoa pod borer larvae	17
Chapter 3	Modification of the *Bacillus thuringiensis* toxin *SN19* gene for improved expression in plants	27
Chapter 4	Gene expression in the cocoa pod wall and pulp	45
Chapter 5	Characterization of a cocoa pod wall-specific gene, *TcLea5*, a comparison with its tomato homologue and characterization of its promoter in heterologous systems	67
Chapter 6	Development of cocoa transformation	85
Chapter 7	Isolation and characterization of the MADS-box genes *AGAMOUS* and *APETALA1* homologues from cocoa	103
Chapter 8	General discussion	119
Summary		133
Samenvatting		137
Ringkasan		141
Acknowledgements		145
Curriculum Vitae		149
Training and Supervision		150
Appendix		151

Chapter 1

Introduction

History and geography of cacao

The cocoa tree *(Theobroma cacao* L.) is a small evergreen tree native to the lower eastern slopes of the Andes in South America. It needs a warm and humid climate, regular rainfall, and grows in the shade, where the temperature is not much lower than 20°C or higher than 28°C. The headwaters of the Amazon basin have been said to be the origin of the cocoa tree. Natural cocoa populations are also present in the lower Amazon basin, Venezuela and the Guayanas.

Chocolate was enjoyed chiefly as a beverage for many centuries. Its popularity began in the Americas where the tree grew wild. Maya and Aztec people discovered the tasty secret of the cocoa tree more than a thousand years ago. They were cultivating cocoa beans before Columbus arrived in the Americas in 1492 and they mixed ground cocoa seeds with various seasonings to make a spicy, frothy drink, which the Aztecs called “chocolatl”. The Mayas and Aztecs also used cocoa beans in many social and religious events (11).

Europe’s first contacts with chocolate came during the conquest of Mexico in 1521 when Hernando Cortez recognized a potential commercial value in the beans of cocoa. Like the Mayas and Aztecs, Europeans also developed their own special protocol for drinking chocolate. Chocolate was a symbol of wealth and power.

In the sixteenth century after the conquest of Mexico, cocoa cultivation spread to the Caribbean Islands. Venezuela was also one of the first countries growing cocoa in the sixteenth century. Jamaica enjoyed a cocoa “boom” in 1670. Since 1678 Trinidad, Martinique, and Haiti were growing cocoa trees, which were introduced from Venezuela. Apart from movement within the Caribbean area, cocoa was taken across the Pacific to the Philippines around 1678. From there it spread later to Celebes and Java.

There is evidence of an introduction of cocoa to India and Sri Lanka from Ambon in the Moluccas in 1798. In the seventeenth century cocoa began to be grown in the lower Amazon basin. In 1746 a French planter brought seeds of cocoa to the state of Bahia. Following independence of Brazil, cocoa was introduced to Sao Tomé in 1822 and to Bioko (Fernando Po) in 1855 and later to Ghana and Nigeria, the basis for cocoa growing in West Africa (23).

Botany and types of cocoa

Theobroma cacao was the name given to the cocoa tree by Linnaeus in the first edition of his Species Plantarum published in 1753. In 1882 Morris of Jamaica distinguished two great classes of *T. cacao,* Criollo and Forastero. In 1964 Cuatrecasas divided the genus *Theobroma* into six sections containing twenty-two species (23). *Theobroma cacao* is the only species, which is cultivated widely, the other better known species in the genus being *T. bicolor* and *T. grandiflorum. T. grandiflorum,* known as "cupuaçu" in Brazil is well liked locally for the delicate flavour of the pulp around the beans. The genus *Theobroma,* together with the genera *Herrania, Guazuma,* and *Cola* belongs to the family *Sterculiaceae. T. cacao* is a diploid species with a chromosome number of 20, and was subdivided, also by Cuatrecasas, into two subspecies, *T. cacao* spp. *cacao* and *T. cacao* spp. *sphaerocarpum.* The subspecies *T. cacao* spp. *cacao* consists of the Criollo populations of Central and South America and *T. cacao* ssp. *sphaerocarpum* includes all other populations (17, 23).

Criollo cocoa is the variety that was cultivated since prehistoric times by the Mayas and other people in Central America according to archaeological evidence. There are two types of Criollo, Ancient and Modern Criollo (17). Criollo dominated cocoa production in Latin America during the pre-Columbian period until Forastero cocoa came into cultivation on a large scale in Brazil and Ecuador in historic times. Criollo cocoa had a higher quality than the Forastero type, but trees had low vigour and yield, and were extremely susceptible to diseases and pests.

The Forastero group is composed of very diverse populations with different geographic origins. They are Forastero of the lower Amazon basin, upper Amazon basin, Orinoco and Guyanas. Amelonado is the population of Forastero from the lower Amazon basin. Almost all Amelonado cocoa populations are uniform types, large trees with high yield, vigorous, precocious and disease-resistant. The pod is light green and has a smooth surface with shallow ridges.

Trinitario, a natural hybrid between Criollo and Amelonado from the Guyanas or the Lower Amazon regions, replaced the earlier Criollo cocoa in most American plantations after 1800 (23). Amelonado are the most extensively planted since 1950 in Brazil and West Africa (16). Trinitario is more productive and resistant to diseases compared to Criollo and Amelonado.

Early in the 20th Century cocoa breeding in Indonesia achieved progress in plant vigor and yield, in combination with good quality of the bean. In Central and West Java more than 10.000 ha of a Djati Roenggo (DR) clone, a hybrid of an introduced Venezuelan Trinitario produced in a seed garden of the plantation after which it was named, were planted. The current DR clones originate from this early population and are classified as Indonesian Trinitario (23). Trees of Forastero, Criollo Java-Red and Imperial College Selection (ICS) clones have been planted in Sumatra, Java, Sulawesi, Irian Jaya and the Moluccas.

Flowering, fertilization and pod development

Cocoa trees produce a large number of flowers at certain times of the year, depending on local conditions and cultivar. Only 1-5 percent of the flowers are successfully pollinated to produce a pod. The flowers of cocoa are formed on the trunk and branches, a habit referred to as cauliflorous or truncate (23). They are about 15 mm in diameter, are borne on long pedicels and have five prominent, pink sepals, five smaller yellowish petals, ten stamens and an ovary of five united carpels (Fig. 1). The petals have a curious shape, they are very narrow at the base but expand into a cup-shaped pouch and end in a broad tip or ligule. The ten stamen, which form the male part of the flower, are in two whorls, the outer whorl consisting of five long non-fertile staminodes, and the inner whorl with five fertile stamen, each bearing up to four anthers. The stamen are curled, and the anthers develop inside the petal pouches (Fig. 1).

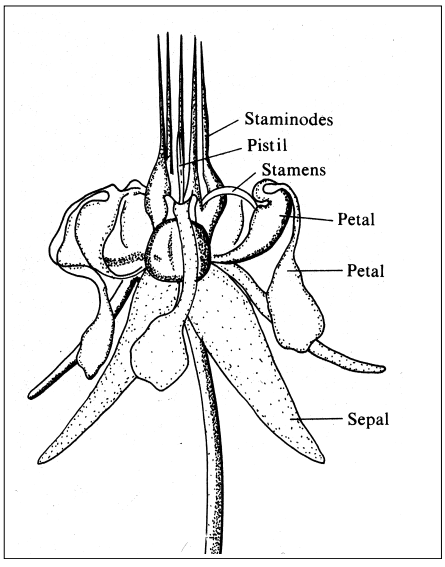


Fig. 1. Cocoa flower (23)

Inflorescence primordia can form in the axil of the first bract of an axillary bud, and can also appear adventitious in certain locations where a leaf never seems to have been formed. It takes about thirty days from initiation for the bud to emerge through the bark and mature. Small inflorescences are occasionally formed from subordinate buds lateral to the principal axillary buds of unabscised leaves. Leaf axils, which have produced flowers and pods for several years, become thickened and are then called flower cushions. When a bud matures, the sepals split during the afternoon and continue to open during the night. Early in the following morning the flowers

are fully opened and the anthers release their pollen. The style (female sporophyte) matures a bit later. Pollination, the process of depositing pollen on the style, is performed by various small insects. Pollination by flying insects results in approximately 25-50% cross-pollination on self-compatible trees. In Brazil on self-compatible cocoa trees, air is blown over the flowers to enhance pollination and increase the yield.

Self-incompatibly is a common phenomenon in tropical species as an efficient system to avoid inbreeding. *T. cacao* presents a unique incompatibility system, in which the incompatibility reaction does not occur at the style and stigma or by inhibition of pollen tube growth, but by failure of gamete nuclei fusion in the embryo sac. This system exhibits sporophytic and gamethopic features. In an incompatible pollination the proportion of non-fusion between ovules and male gametes in the ovary can vary form 25, 50 to nearly, if not quite 100%. In this case the ovary will fail to develop and the flower falls off three or four days after pollination. The degree of incompatibility varies between different cultivars. All the cultivars in the Amazon region are self-incompatible, but generally cross-compatible. A high proportion of Trinitario cultivars is self-incompatible, and not cross-compatible with other self-incompatible trees, but requiring pollen from self-compatible trees for successful pollination. The Amelonado population is self-compatible.

After fertilization the pod grows slowly for about 40 days. Between 40 to 50 days after fertilization, the zygote, a single cell in an ovule, makes its first division to become the embryo, from which the seedling develops. Eighty-five days after fertilization, the pod and ovule growth slow down at the expense of embryo growth. The ovule is filled with a jelly-like endosperm, which is completely consumed by the embryo about 140 days after fertilization. The seed or bean consists of two convoluted cotyledons and a small germ, together enclosed in the testa. The development of the pods takes 5-6 months from pollination to full ripeness. The number of seeds per pod is usually between 30 and 40. The seeds of cocoa are attached to a placenta and each bean is surrounded by a sweet pulp (mucilage). The fruit is botanically a non-dehiscent drupe and cocoa seeds display no dormancy period: the process of germination will start immediately after the seeds are removed from the pod. Natural dissemination depends on monkeys, rats or squirrels making a hole in the pod to extract the seeds and suck off the surrounding pulp, dropping the seeds on the ground.

Bean processing and chocolate production

Cocoa beans from the tree have to be processed before they can be marketed as chocolate. Unlike many contemporary crops, cocoa cannot be harvested by machine. Each pod must be plucked by hand and opened as soon as it is plucked. The pods are carefully cut with a machete, and the beans plus surrounding pulp are scooped out by hand to avoid damage to the beans. The beans are collected and placed in wooden boxes, where natural fermentation develops. The viability of the seeds is destroyed during fermentation and the mucilage will heat up, activating enzymes and creating compounds that give the beans their chocolate flavour, aroma and colour. Chocolate flavour develops in two stages, curing and roasting. Curing involves fermentation followed by drying. Fermentation is complete when the beans turn a rich deep brown colour. After drying, the beans are bagged and transported in containers. Cured beans can be safely stored for 2-3 years.

The manufacture of chocolate involves a series of processes by which flavour is further developed. The first process is to clean beans to remove any foreign matter using rocking and vibratory sieves. The second process is roasting, the most important stage in the development of chocolate flavour. In the roasting process the beans are heated to 100-120°C, and the process causes the beans to release their rich aroma and delicious taste. The roasted beans are next winnowed to remove the broken pieces of shells.

The nib (cotyledon) is ground into “cocoa liquor” or mass. The cocoa liquor is pressed to extract the cocoa butter leaving a solid mass called press cake. The cake is then broken down in a mill, resulting in cacao powder. Most cacao powders are made from mass or nib, which has been treated with alkali to improve the flavour. Cacao butter can be produced from whole beans, and mixtures of fine nib dusts, small nibs, and immature beans. Plain chocolate is made by mixing mass or nibs with sugar and sufficient cocoa butter. Other ingredients are incorporated and a fine texture is achieved. At present already 30% of all cacao is processed into liquor, butter and powder before export to a consumer country.

Cacao is more than a source for calories and confections. The chemicals and substances in cacao can be extracted and incorporated into cosmetics and medicine. Cacao is one of the most flavonoid-rich foods, containing catechin and epicatechin, which are responsible for some of its cardio protective properties (6, 13, 22).

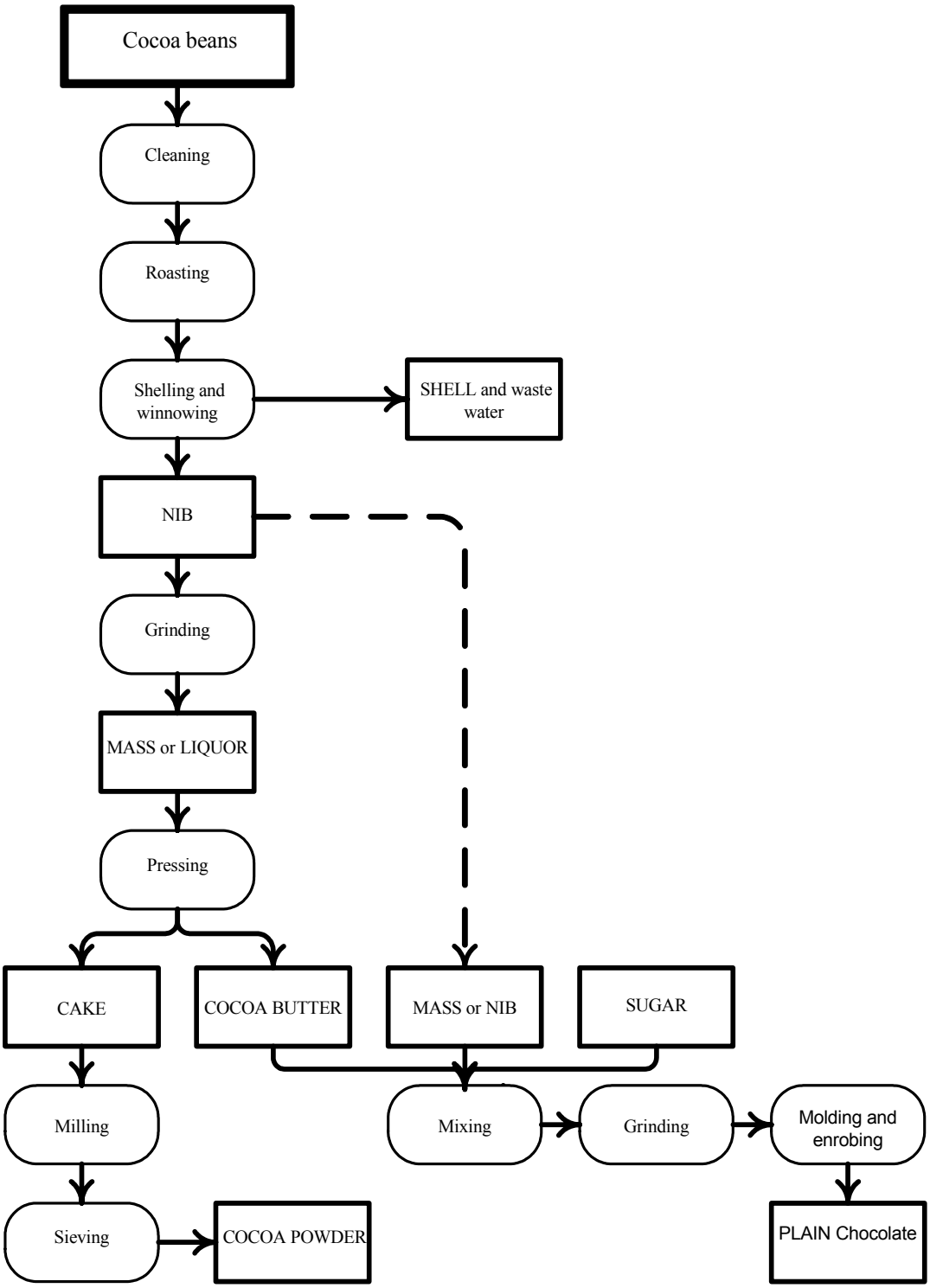


Fig. 2. Flow diagram for chocolate processing (26)

Pests and diseases of cocoa

Cacao is a vital export crop for many countries, particularly in Southeast Asia and West Africa. Up to 80% of the world cacao production is lost due to diseases and pests such as black pod, witches' broom, bark cancer, cocoa pod borer, frosty pod rot, *Fusarium roseum,* mealy pod rot, and cocoa swollen shoot virus (7, 14).

The black pod disease, or *Phytophtora* pod rot, has been the primary fungal disease affecting cocoa production since 1920. It is generally caused by *Phytophtora palmivora,* which is present worldwide, and more recently by *Phytophtora megakarya,* restricted to some African countries. *Phytophtora palmivora* black pod results in yield losses of 5–19%, but when the more virulent *Phytophtora megakarya* is involved, yield loss may range from 60 to 100% (14). The first symptoms of the disease are brown to black spots on the fruit, which increase in size. Several days after infection, a white mycelium, with sporangia, develops on the surface of the diseased pod, and the fruit become mummified. Humid conditions are favourable for the development of *Phytophtora* in the fruit. The inoculum develops from the soil or other parts of the tree such as infected roots, stems, leaves, and trunk bark. Pods can be infected at any age. The most significant loss arises from the infection of pods in the two months before they reach ripeness. The fungus can easily pass from the pod husk to the seed coat of the bean in a developing green pod as the husk and bean are in physical contact.

Witches' broom is a disease first described in 1785 by Fereira, who illustrated and named the malformation he observed "lagartão". The causal fungus (a basidiomycete) of Witches' broom of cocoa, first named *Marasmius perniciosus*, was renamed *Crinipellis perniciosa* (Stahel) Singer. *C. perniciosa* is indigenous to the Amazon Basin, and now found in Bolivia, Brazil, Columbia, Guayana, Ecuador and Trinidad (14, 19). The host range of *C. perniciosa* includes *Theobroma* spp., *Herrania* spp*., Solanum* spp*.,* and *Bixa orellana.* Symptoms induced in cocoa by *C. perniciosa* can occur in the vegetative parts as well as the pods (14, 19, 20). The symptoms on the pods vary and depend on their ages and size. Developing pods can be infected until about 12 weeks of age, causing severe damage to the pod, seed and failure to mature. Pods of 13-15 weeks of age can be infected, but the seeds within are little damaged. Pods older than 15 weeks apparently cannot be infected. Destruction of seed because of early infection of pods by *C. perniciosa* precludes formation of cocoa beans. Vegetative symptoms depend on the cultivars, the type of tissues involved and the stage of development. The most obvious symptoms, which give rise to its name, are

the characteristic shoots or brooms, caused by hypertropic growth of an infected bud. Vegetative shoots or brooms develop following infection of terminal and axillary buds. The most frequent brooms are fan brooms produced on the fan branches, when an axillary bud is infected. Leaves that are produced on fan brooms appear normal. An infected flower cushion produces vegetative shoots and abnormal flowers known as star blooms. The vegetative shoots are often 30 cm long. Vegetative brooms and cushion brooms may remain green for several weeks, but they finally change to a necrotic status. The physiological changes that induce hypoplasia and hypertophy in infected cocoa tissues are not known.

The frosthy pod rot pathogen, the fungus *Moniliophthora theobromae* is also called *Monilia roreri*. The symptoms of infection by *Moniliophthora* only develop in green pod tissues. The fungus only colonizes actively growing pod tissues, subsequently growing between the parenchyma cells of the cortex and producing conidia both within and on the surface of the host tissue. Infected pods never develop beans, the tissues being replaced by a disorganized gelatinous substance. In later infected pods the beans stick to each other and to the pod wall, making them difficult to remove. (10).

Trachysphaera fructigena the fungus causing mealy pod rot, is only recorded in West and Central Africa. The symptoms produced in infected pods are very similar to those of *Phytophthora* pod rot.

The virus diseases in cocoa are cocoa-swollen shoot (CSSV), cocoa necrosis (CNV), and cocoa yellow mosaic (CYMV). CSSV was first reported in 1936 in the Eastern Region of Ghana. There are many strains of CSSV, which differ in the symptoms they produce. The virulent strains predominate and cause various types of leaf chlorosis, root necrosis, root and stem swellings and dieback in Amelonado in Ghana. The trees are killed in two or three years from infection.

The primary insect pests in cocoa are groups of mirids (capsids). All regions have their specific mirids, causing severe damage to twigs, branches and young pods. *Distantiella* and *Sahlberghella* are the most important pests in West Africa, while the genus *Helopeltis* is the major mirid In Southeast Asia.

Cocoa pod borer

The cocoa pod borer, *Conopomorpha cramerella* (Snellen) is a member of the lepidopteran family *Gracillariidae.* Other members of this family of Microlepidoptera are primarily known

as leaf miners such as the Citrus leafminer, *Phyllocnistis citrella* Stainton. *C. cramerella* is also found in kola *(Sterculia acuminata),* nam-nam *(Cynometra cauliflora,)*, rambutan *(Nephelium lappaceuom),* and litchi *(Litchi chinensis* Sonn). Other members of the same genus*, C. sinensis* (Bradley), *C. oceanica* (Bradley) and *C. litchiella* (Bradley) also attack litchi and longan fruits in Taiwan and China. (18, 27). The cocoa pod borer is the most serious insect pest of cocoa in Southeast Asia (Indonesia, Philippines, Malaysia, and New Guinea). Crop losses due to this insect are up to 80% in some geographical regions (15, 24).

The strongly flattened orange-brown eggs are laid singly on the pods, mainly in the furrows, and are very difficult to see. Incubation lasts 3-6 days and the larva immediately after hatching bore into the pod and makes long frass-filled galleries in the pulp. Beans are not eaten but callus formation takes place, affecting bean development. When fully grown, after 15-18 days the larvae are 10-12 mm long. When in the pod they are off-white, but they become pale green as they eat their way out through the green husk. A cocoon is formed in which the larvae pupates, sometimes on the underside of leaves or on the pod surface but possibly more frequently on dead leaves on the ground. The adult emerges after 5-8 days, thus the total life cycle length is about 4 weeks. The moths fly feebly at sunset and can possibly be carried for some distance by wind. The larvae spend almost their whole life inside the pods of the cocoa fruits. This habit gives the larvae protection against natural enemies and insecticides, and allows them to disperse passively over long distances in the fruit.

Several essentially cultural methods for control of the cocoa pod borer are known. “Rampasan” consists of the total stripping of pods of more than 5 cm in length from farms or plantations once or twice a year, at periods of low fruiting to break the breeding sequence of the moth. Another method is the covering of developing pods with transparent plastic sleeves, open at the bottom, when the pods are 2 – 3 months old, before they are likely to be infested with eggs. Biological control, using black ants or parasitic wasps is also undertaken. Predation on pupae provides some limitation of numbers, but regulation of the cocoa pod borer by natural enemies is weak and the prospects for biological control are unpromising (7).

Bacillus thuringiensis

Bacillus thuringiensis (Bt), is a gram-positive bacteria that produces crystalline insecticidal proteins during sporulation. This common soil bacterium can be found almost everywhere in the world, in all types of terrain, including beaches, desert, and tundra habitats. *Bt* was first

isolated in 1901 by S. Ishawata, a Japanese biologist, as the causal agent of "soto disease" (sudden collapse bacillus) that was killing large populations of silkworms (1). Berliner described *Bacillus thuringiensis* for the first time in 1911 when he isolated a *Bacillus* from the flour moth *Anagasta kuehniella* from the province of Thuringia, Germany (1). In 1915, he reported on a parasporal crystal in sporulating *Bt* cells, but the activity of this crystal was not discovered until much later. Hannay and Fitz-James in 1956 reported that the main insecticidal activity against Lepidoptera (larvae of moths and butterflies) was due to the parasporal crystal and Angus demonstrated that the parasporal crystal was the cause of toxicity to the silk worm (8).

The crystal (Cry) proteins form a large, and still growing family of homologous proteins: more than 140 different crystal protein are known to date (2), with each protein being active on insect larvae of only one or a few species in a specific order. The Cry proteins generally form crystalline inclusions in the mother cell compartment. Depending on their protoxin composition, the crystal have various forms: bipyramidal (Cry1), cuboidal (Cry2), flat rectangular (Cry3A), irregular (Cry3B), spherical (Cry4A and Cry4B), and rhomboidal (Cry11A). Höfte and Whitely in 1989 published the first attempt to organize the genetic nomenclature of toxin genes based on the insecticidal activities of crystal proteins (also called δ-endotoxins). These *cry* genes were classified into four major classes and several sub-classes based on activity and nucleotide similarities. The *cryI* genes encoded proteins toxic to Lepidoptera (moths and butterflies), *cryII* to both Lepidoptera and Diptera (flies and mosquitoes), *cryIII* to Coleoptera (beetles), and c*ryIV* genes encoded proteins specific for Diptera (9). This system was revised by Crickmore et al. in 1998, and then entirely based on the degree of amino acid identity. Roman numerals were exchanged for Arabic numerals in the primary rank (eg CryIA became Cry1A) (2). More recent reports describe *Bt* strains with activity against Hemiptera, Mallophaga, Nematoda and Protozoa (21).

B. thuringiensis first became available as a commercial insecticide in France in 1938 and in the 1950s entered commercial use in the United States. *Bt* toxins are more pest-specific and more environmentally safe than many conventional pesticides. For these reasons, use of commercial spray formulations of *Bt* is allowed for organic farmers (3).

Mode of action of δ-endotoxins

After ingestion by insects, the crystal dissolves in the gut juice. Subsequently proteases activate the protoxins to produce an active toxin with a molecular weight of approximately 60 kDa. The activated toxin binds to receptors on the epithelial cell membranes and inserts to form pores, probably as oligomers. The binding affinity of these toxin fragments is often directly related to the toxicity, though binding does not assure toxicity. The pores disrupt the osmotic balance of these cells, causing them to lyse (4, 25). The insect soon stops feeding, and most susceptible insects will die within a few hours of ingestion.

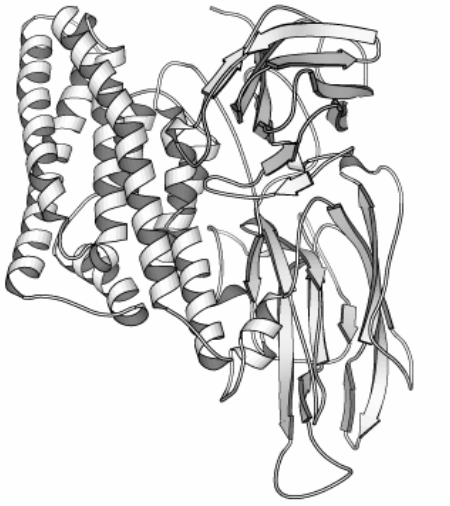


Fig. 3. Structure of a typical 3-domain Cry protein from *B. thuringiensis*

The structure of most activated Cry proteins consists of three domains, each with a specific function (Fig. 3). Domain I consists of 7 α-helices and is responsible for insertion into the gut membrane, and forming the pore. Domain II consists of three antiparallel ß-sheets, and is responsible for binding to receptors in the midgut. Domain III consists of two antiparallel ß-sheets, and is involved in ion channel function, as well as in receptor binding. Specificity is to large extent determined by the toxin receptor interaction, although solubility of the crystal and protease activation also plays a role.

Transgenic plants

Genes encoding Bt toxins have been incorporated into and expressed in transgenic crop plants since 1987 (3). Crops that have been genetically engineered for insect resistance can dramatically lower production costs and have provided farmers with new insect control options.

Genetically engineered crops have quickly become a significant component of agriculture. The production of Cry protein in planta can offer several benefits. In plants, the toxin is produced continuously and protected against degradation, and it is expressed in the appropriate tissues. Several *cry* genes have been introduced into plants starting with tobacco, potato, tomato, cotton, rice and maize. In 1997, Bt-cotton, maize, and potatoes covered nearly

10 million acres of land in the United States alone. The global GM crops area has grown 40-fold, from 1.7 million hectares in 1996 in the US only to 81 million hectares in 2004 in 17 countries. The US is the largest grower of GM crops, followed by Argentina, Canada , Brazil, China, Paraguay, India and South Africa (12). In 2004 the transgenic Bt crops in commercial culture were: transgenic maize containing the *cry1Ab* or *cry1Fa* gene, for resistance to European corn borer and Mediterranean corn borer, and *cry3Bb* for corn rootworm control. Cotton expressing *cry1Ac* or *cry1Ac* + *cry2Ab* to control cotton bollworm, pink bollworm, and tobacco bollworm, and *cry3Aa* in potato to control Colorado potato beetle.

Scope of the thesis

Indonesia is the second largest producer of cacao in the world. Cocoa Pod Borer (CPB) has been the single most important limiting factor for cacao production in Indonesia. At the beginning of the 20th century CPB largely destroyed the early cocoa industries in North Sulawesi and Java. In 1997 CPB was confirmed to have infected 10% of all cocoa plantations in Indonesia, and in 1998 this number had doubled. Numerous outbreaks of CPB have occurred, and now more than 50% of the total area of cocoa plantations in Indonesia has been infested (5). Larvae of CPB attack the economically important cocoa pod, which contains the beans used for chocolate production. So far, there has been no single cost effective and environmentally safe way to control this pest. Residing inside the pod, the larvae are out of reach for chemical sprays. Considering that CPB is a lepidopteran insect, cocoa resistant to CPB might be developed through genetic engineering, using a *cry* gene from *Bacillus thuringiensis*, encoding a toxin effective against CPB larvae.

In order to develop transgenic cocoa trees expressing an effective *cry* gene product, this thesis describes five major activities. First, twelve Cry proteins from the Cry1 class were tested for activity against the larvae of the cocoa pod borer in bioassays, on cocoa plantations in Indonesia, to determine whether this strategy might work for CPB. Subsequently an active toxin-encoding gene was modified for better expression in cocoa plants. Third, we studied gene expression in the pod wall and pulp and isolated and characterized a pod-specific promoter for controlling *cry* gene expression in cocoa. Fourth we optimized the procedures for transformation and regeneration of transgenic cocoa. Lastly, we identified and characterized two transcription factor encoding genes involved in determination of flowering time and floral organ identity.

Literature cited

1. **Beegle, C. C., and T. Yamamoto.** 1992. History of *Bacillus thuringiensis* berliner research and development. Can. Ent. **124:**587-616.
2. **Crickmore, N.** 2003. *Bacillus thuringiensis* toxin nomenclature. http://www.biols.susx.ac.uk/home/Neil_Crickmore/Bt/index.html.
3. **de Maagd, R. A., D. Bosch, and W. J. Stiekema.** 1999. *Bacillus thuringiensis* toxin mediated insect resistance in plants. Trends Plant Sci. **4:**9-13.
4. **de Maagd, R. A., A. Bravo, and N. Crickmore.** 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet. **17**: 193-199
5. **Depparaba, F.** 2002. Cacao moth (*Conopomorpha cramerella* Snellen) and its control measures. Jurnal Litbang Pertanian. **21:**69-74.
6. **Dillinger, T. L., P. Barriga, S. Escarcega, M. Jimenez, D. S. Lowe, and L. E. Grivetti.** 2000. Food of the Gods: cure for humanity? A cultural history of the medicinal and ritual use of chocolate. J. Nutr. **130:**2057S-2072S.
7. **Entwistle, P. F.** 1987. Insects and cocoa. p. 366-443. *In* G. A. R. Wood, and R. A. Lass (eds), Cocoa. Longman Scientific and Technical, Harlow.
8. **Glare, T. R., and M. O'Callaghan.** 2000. *Bacillus thuringiensis:* biology, ecology and safety. John Wiley & Sons, LTD, New York.
9. **Höfte, H., and H. R. Whiteley.** 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol Rev. **53:**242–255.
10. **Hoopen, G. M., Rees, R., Aisa, P., Stirrup, T., and Krauss, U.** 2003. Population dynamics of epiphytic mycoparasites of the genera *Clonostachys* and *Fusarium* for the biocontrol of black pod (*Phytophthora palmivora*) and moniliasis (*Moniliophthora roreri*) on cocoa (*Theobroma cacao*). Mycol. Res. **107:**587-96.
11. **Hurst, W. J., Tarka, S.M. Jr., Powis, T.G., Valdez.F Jr., and Hester, T.R.** 2002. Cacao usage by the earliest Maya civilization. Nature **418:**289-290.
12. **ISAAA.** 2004. Global status of commercialized transgenic crops: 2004. http://www.agbiotechnet.com.
13. **Keen, C. L.** 2001. Chocolate: Food as Medicine/Medicine as Food. J Am Coll Nutr. **20:**436S-439.
14. **Lass, R. A.** 1987. Disease. p. 265-365. *In* G. A. R. Wood, and R. A. Lass (eds), Cocoa. Longman Scientific and Technical, Harlow.
15. **Lim, G. T.** 1992. Biology, ecology and control of cocoa podborer *Conopomorpha cramerella* (Snellen). p. 85-100. *In* P. J. Keane, and C. A. J. Putter (eds), Cocoa pest and disease management in Southeast Asia and Australasia FAO Plant Production and Protection Paper No. 112. FAO, Rome.
16. **Motamayor, J. C., Risterucci, A.M., Heath, M., and Lanaud, C.** 2003. Cacao domestication II: progenitor germplasm of the Trinitario cacao cultivar. Heredity **91:**322-330.
17. **Motamayor, J. C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., and Lanaud, C.** 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity **89:**380-386.

18. **Nguyen Thanh, V., Anh Hai, Dao., Lachance, M. A., and Nguyen Thanh, Vu.** 2003. *Issatchenkia hanoiensis*, a new yeast species isolated from frass of the litchi fruit borer *Conopomorpha cramerella* Snellen. FEMS Yeast Res. **4:**113-117.

19. **Purdy, L., and R. Schmidt.** 1996. Status of cacao witches' broom. Biology, epidemiology and management. Ann. Rev. Phytopathol. **34:**573-594.

20. **Rudgard, S. A.** 1989. International Witches' Broom Project: Detailed descriptions of symptoms of witches' broom of cocoa caused by *Crinipellis perniciosa.* Cocoa Grow. Bull. **41:**1-32.

21. **Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler, and D. H. Dean.** 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 62: 775-806

22. **Schramm, D. D., J. F. Wang, R. R. Holt, J. L. Ensunsa, J. L. Gonsalves, S. A. Lazarus, H. H. Schmitz, J. B. German, and C. L. Keen.** 2001. Chocolate procyanidins decrease the leukotriene-prostacyclin ratio in humans and human aortic endothelial cells. Am. J. Clin. Nutr. **73:**36-40.

23. **Toxopeus, H.** 1987. Botany, types and populations. p. 11-37. *In* G. A. Wood, and R. A. Lass (eds), Cocoa, 4 ed. Longman Scientific and Technical, Harlow.

24. **Wardojo, S.** 1992. Major pests and diseases of cocoa in Indonesia. p. 63-67. *In* P. J. Keane, and C. A. J. Putter (eds), Cocoa pest and disease management in Southeast Asia and Australasia FAO Plant Production and Protection Paper No. 112. FAO, Rome.

25. **Whalon, M. E., and Wingerd, B.A.** 2003. Bt: Mode of action and use. Arch. Insect. Biochem. Physiol. **54:**200-211.

26. **Wood, G. A. R.** 1987. Consumtion and manufacture., p. 587-597. *In* G. A. R. Wood, and R. A. Lass (eds), Cocoa. Longman Scientific and Technical, Harlow.

27. **Yao, Z. W., and S. K. Liu.** 1990. Two gracillariid insect pests attacking litchi and longan. Acta Entomol. Sin. **33**: 207-212.

Chapter 2

Activity of *Bacillus thuringiensis* toxins against cocoa pod borer larvae

Djoko Santoso,[1†] Tetty Chaidamsari,[2†] Soekadar Wiryadiputra,[3] and Ruud A. de Maagd[2]

Pulbished in *Pest Manag Sci* 60 (2004):735–738 (online: DOI: 10.1002/ps.927)

[1]Biotechnology Research Institute for Estate Crops, Jl Taman Kencana 1, Bogor 16151, Indonesia

[2]Business Unit Bioscience, Plant Research International B.V., P.O. Box 16, 6700 AA Wageningen, The Netherlands

[3]Indonesian Coffee and Cacao Research Institute, Jl Jenderal Sudirman 90, Jember 68118, Indonesia

† These authors have contributed equally to this work

Summary

Twelve Cry proteins from *Bacillus thuringiensis* Berliner were tested in bioassays on cacao plantations in Indonesia for activity against the larvae of cocoa pod borer (*Conopomorpha cramerella* (Snellen)), an insect pest of the cacao tree. Through the damage caused by their feeding, the larvae of cocoa pod borer cause the pods of the cocoa tree to ripen prematurely. They are difficult to control with conventional measures. Preliminary assays identified five toxins that were more active than others. In two subsequent bioassays the activity of selected toxins was determined more accurately. Three Cry1 proteins with relatively little homology were all found to be toxic, opening perspectives for controlling cocoa pod borer by expression of Cry proteins in transgenic plants.

Introduction

The major insect pest threatening cocoa production in Indonesia, as well as in the rest of Southeast Asia, is cocoa pod borer (CPB), *Conopomorpha cramerella* (Snellen) (Lepidoptera: Gracillaridae). Whereas other members of this family of Microlepidoptera are primarily known as leaf miners (such as the Citrus leafminer, *Phyllocnistis citrella* Stainton), CPB larvae attack the economically important cocoa pod, which contains the beans used for chocolate production. Members of the same genus*, C. sinensis* (Bradley) and *C. litchiella* (Bradley) attack litchi and longan fruits in Taiwan and China [20]. The adult CPB is a mosquito–like, approximately 1 cm long moth. The female moth lays up to 200 eggs on the furrowed surface of the cocoa pod. After a few days the eggs hatch, larvae emerge and directly burrow into the pod, eating the pulp tissue (mesoderm) and placenta, disrupting the nutrient flow for developing beans. The pod dries up after the larva have fed on the pulp surrounding the beans and entry holes allow infections to rot the pod, which can cause the pod to turn yellow prematurely, as well as causing the clumping of deformed beans. Approximately two weeks after hatching, the larvae leave the pod and pupate on the pod or leaves [14, 19]. Crop losses due to cocoa pod borer can be as high as 80% [3]. So far, there has been no a single cost effective and environmentally safe way to control this pest. Residing inside the pod the larvae are out of reach for chemical sprays.

Bacillus thuringiensis (Bt) strains produce crystals containing proteins (Cry proteins) that are toxic to many important insect pests. The insecticidal activities of the various toxins differ, but they are considered harmless to most higher organisms, including people [8, 18]. Genes encoding Bt toxins (*cry* genes) have been incorporated into many crop plants to render them insect resistant. Since 1995 such plants have been commercialized for example cotton, maize and potato, resulting in efficient control of several insect pests and in large reductions in chemical pesticide use [6].

Considering that CPB is a lepidopteran insect, cacao resistant to CPB might be developed through genetic engineering using a *cry* gene from *Bacillus thuringiensis*, encoding a toxin effective against CPB larvae. For this purpose, we have tested twelve different Bt toxins for activity against the cocoa pod borer.

Materials and Methods

Expression vectors.

Bacillus thuringiensis toxins were prepared from twelve different bacterial clones. All Cry protein expression vectors, with exception of pSB204, are based on pBD10, a derivative of pKK233-2 [1]. These were pPB08 (*cry1Aa*), pBD140 (*cry1Ab*), pB03 (*cry1Ac*), pMH19 (*cry1Ba*), pBD150 (*cry1Ca*), pMH15 (*cry1Da*), pBD160 (*cry1Ea*), pMH21 (*cry1Fa*), pBD172 (*cry1Ia*), H04 and SN19 [1, 4, 5, 7, 15]. H04 and SN19 encode hybrid proteins with the domain composition (1Ab/1Ab/1Ca), and (1Ba/1Ia/1Ba), respectively [4, 15]. Cry1Cb protein was produced from *Escherichia coli* containing plasmid pSB204 with the *cry1Cb* gene from *Bacillus thuringiensis* subsp. g*alleriae* HD29 [13], provided by T Yamamoto (Sandoz, Palo Alto, Calif.,USA).

Protein isolation and bioassays.

Protoxins were expressed in *E. coli* strain XL-1, extracted, solubilized and activated with trypsin as described previously [1]. Solubilized protoxins and activated toxins were dialyzed overnight against 25 mM $NaHCO_3$/100mM NaCl, pH10. Protein concentration of the toxins was estimated using SDS-PAGE. Toxins were diluted in phosphate-buffered saline pH7.4 just prior to use in bioassays. Cry1Ba, Cry1Ia and SN19 were tested at first only as protoxins.

Bioassays with the toxins on CPB larvae were performed using an artificial diet containing cacao pulp, in the cacao plantations of Hasfarm Product Jahab Inc. located in Kabupaten Kutai, East Kalimantan, and of Inang Sari Inc. located in Lubuk Basung, West Sumatra, Indonesia. For 1 liter of artificial diet 25 g agar, 35 g casein, 100 g cellulose powder, 10 g Wesson salts, and 30 g wheat germs (premixed) were added to 150 g of a 2:1 mixture of cocoa pulp and water, and mixed in a blender. Next, to this 0.05 g cholesterol and 0.25 ml linoleic acid, dissolved in 50 ml dichloromethane, and 650 ml water were added and mixed, allowing the dichloromethane to evaporate. The pH was adjusted to 6.5 with 4M KOH and subsequently this mixture was autoclaved at 115°C for 20 minutes. After cooling to 60°C, 2.0 g Vanderzant vitamins, 3.0 g sucrose, and 0.5 g glucose (premixed) were added, and to this 15 mg streptomycinsulphate and 15 mg benzyl penicillin sodium dissolved in 7 ml water were added and mixed. Finally to this 15 ml of 90 mg/ml methyl-p-hydroxybenzoate and 120 mg/ml sorbic acid dissolved in 95% ethanol was added and mixed. For bioassays 20 ml of the

artificial diet was poured into 60 ml glass vials with an outer diameter of 5 cm and allowed to solidify.

For preliminary bioassays 100 ng/cm^2 of each toxin was overlaid onto freshly prepared artificial diet. After the fluid was absorbed into the medium, 5 to 10 3rd-4th instar larvae, collected from pods freshly picked on the plantation were placed in each vial. Larval mortality was scored every day up to 3 days. Fifty percent lethal concentrations (LC_{50}) at day three and 95% fiducial limits were determined by Probit analysis of results from five or six independent experiments using the PoloPC computer program [17].

Results and Discussion

Until this study, no attempts had been reported to asses activity of *Bacillus thuringiensis* Cry proteins against the cocoa pod borer. Such studies are difficult since there is no continuous laboratory colony of this insect available, and quarantine regulations forbid the movement of CPB larvae or eggs to our laboratories on the island of Java, which so far is free from CPB. As an alternative, we have performed bioassays on-site in infested plantations elsewhere, using an artificial diet and 3rd to 4th instar larvae isolated from freshly picked, infested pods. Bioassays were performed in three separate field trips, where available of time and manpower limited the number of bioassays that could be performed each time.

The artificial diet for CPB reported here allowed growth of larvae caught as 3rd or 4th instar for 5 days or more, or to pupation, whichever was earlier and with little or no control mortality. Bioassays of neonate or 1st instar larvae would have been preferable, because younger instars are generally more sensitive to Cry proteins. However, such larvae of CPB were invariably much weaker then the older specimens, giving 50% or higher control mortality. This was deemed unpractical for further studies.

The toxins we chose for the bioassays are all of the Cry1 class, which is known for its activity against lepidopteran larvae, though each toxin has a limited host range even within that insect order. Included were two hybrid proteins, H04 and SN19, which were developed in our laboratory [4, 15]. Toxins were presolubilized and activated by trypsin treatment to negate possible limitations to solubility and gut protease activation in a new insect host. However, it is known that in some toxins such as Cry1Ba, Cry1Ia, and thus possible also their hybrid SN19, trypsin will recognize and cleave a site within the first structural domain, which may render the protein inactive [7, 12]. For this reason these three proteins were tested initially

Table 1 Toxicity of *Bacillus thuringiensis* toxins against CPB larvae determined in three field trips to cacao plantations in Indonesia. Toxicity is indicated as % mortality at a single dose of 100 ng/cm^2 or LC_{50} (in ng/cm^2). Where appropriate, 95% fiducial limits are given between parenthesis

Toxin	% mortality		LC_{50} (95% fiducial limits)			
	Bioassay 1 Experiment		Bioassay 2		Bioassay 3	
	1	2				
Control	0	10				
Cry1Aa	53	60				
Cry1Ab	60	70			19.9	(14.3-25.8)
Cry1Ac	60					
Cry1Ba[a]	67	85	8.90	(0.20-24.9)		
Cry1Ca	60					
Cry1Cb	40					
Cry1Da	67	80				
Cry1Ea	67	85	16.2	(6.70-25.2)	49.3	(35.8-65.2)
Cry1Fa	47					
Cry1Ia[a]	73	90	3.90	(-)		
SN19[a]	80	70	4.70	(0.30-11.0)	37.8	(27.3-49.6)
SN19	ND	ND			24.3	(18.0-31.3)
H04	40					

[a]Tested as protoxins

only as solubilized protoxins. In the first assay, in order to make a first estimation of the activity of many proteins with a limited number of larvae, we tested twelve different (pro) toxins against CPB larvae in a cacao plantation located in Kabupaten Kutai, East-Kalimantan. Results of the bioassays are shown in Table 1. At the single dose of 100 nanograms toxin per square cm artificial diet, most of the Cry proteins had some activity. Eight out of twelve toxins tested could kill 50% or more CPB larvae in the 3rd-4th instar, while control mortality was zero.

The second experiment during the first field trip confirmed the activity of the 7 most toxic proteins using the same single dose. In general similar toxicity was found in both experiments (Table 1). For bioassays during a second field trip to the same location, we selected four of the seven more active toxins to test with different toxin concentrations, in an attempt to estimate a reliable 50% mortality dose. All four proteins, SN19, Cry1Ea, Cry1Ba, and Cry1Ia, again showed measurable activity against CPB. As far as could be determined with relatively broad or lacking (Cry1Ia) fiducial limits there was no significant difference in activity between these four toxins (Table 1).

The third and last set of bioassays was performed on a cocoa plantation located in Lubuk Basung, West-Sumatra. Since Cry1Ba, Cry1Ia, and their hybrid SN19 had similar activities and are relatively homology at amino acid level, we decided to proceed only with

SN19, and Cry1Ea, while including Cry1Ab for the first time. In this bioassay SN19 was tested as protoxin as well as trypsin-activated toxin to asses the effect of the earlier mentioned domain I processing on its activity. All four proteins had considerable activity and LC_{50}'s were determined with a high level of reliability, reflected in the narrower 95% fiducial limits (Table 1) and g-values of 0.07 to 0.09 as calculated by Polo PC. Trypsin-activated SN19 had an almost two times higher activity per weight then the equivalent protoxin, as might be expected for any activated Cry1 toxin compared to its full length protoxin, which has an approximately two times higher molecular weight. This indicated that trypsin-activation had no negative effect on SN19's activity against CPB. Comparing the activities of activated toxins, Cry1Ab and SN19 had similar activity and were approximately twice as toxic as Cry1Ea. For the toxins that were more accurately assayed in two locations (Cry1Ea and SN19 protoxin), there was approximately a 3 to 8-fold variation in LC_{50}'s. These differences may reflect variations in the experimental procedure or in the toxin sensitivity of the different CPB populations. However, it should be noted that the estimations from the third set of bioassays are the most accurate.

Conclusions

This is the first report of activity of *Bacillus thuringiensis* Cry proteins against cacao pod borer or for any member of the family Gracillaridae. Cry1Ab, Cry1Ea and SN19 (or its parents, Cry1Ba and Cry1Ia) occupy separate subbranches on the Cry1 branch of the Cry protein phylogenetic tree [2]. Since such relatively none-homologous Cry1 proteins in many target insects recognize different, non-overlapping receptors [9-11], this observation points to opportunities for resistance management by deploying two of these proteins simultaneously [16]. However, this idea awaits testing of receptor specificity for these toxins in CPB. These results indicate that cacao pod borer may be controlled effectively and in an environmentally safe way by expression of one or more Cry proteins in the pod walls of transgenic cocoa trees.

Acknowledgements

Part of this work was supported by the BIORIN Programme (Scientific Programme Indonesia-Netherlands) of the Royal Netherlands Academy of Arts and sciences (KNAW), and by an Indonesian International Joint Research Grant (RUTI) of the Ministry of Research and Technology of the Republic of Indonesia. We would like to thank Hansfarm Product

Jahab Inc. and Inang Sari Inc. for making these bioassays possible and Salvador Herrero for critically reading the manuscript.

Literature cited

1. **Bosch D, Schipper B, van der Kleij H, de Maagd RA, Stiekema WJ** (1994) Recombinant *Bacillus thuringiensis* crystal proteins with new properties: possibilities for resistance management. Bio/Technology 12: 915-918
2. **Crickmore N** (2003) *Bacillus thuringiensis* Toxin Nomenclature. http://www.biols.susx.ac.uk/home/Neil_Crickmore/Bt/index.html
3. **Day R** (1989) Effect of cocoa pod borer, *Conopomorpha cramerella*, on cocoa yield and quality in Sabah, Malaysia. Crop Protection 8: 332-339
4. **de Maagd RA, Kwa MSG, van der Klei H, Yamamoto T, Schipper B, Vlak JM, Stiekema WJ, Bosch D** (1996) Domain III substitution in *Bacillus thuringiensis* delta-endotoxin CryIA(b) results in superior toxicity for *Spodoptera exigua* and altered membrane protein recognition. Appl Environ Microbiol 62: 1537-1543
5. **de Maagd RA, Bakker PL, Masson L, Adang MJ, Sangadala S, Stiekema W, Bosch D** (1999) Domain III of the *Bacillus thuringiensis* delta-endotoxin Cry1Ac is involved in binding to *Manduca sexta* brush border membranes and to its purified aminopeptidase N. Mol Microbiol 31: 463-471
6. **de Maagd RA, Bosch D, Stiekema WJ** (1999) *Bacillus thuringiensis* toxin mediated insect resistance in plants. Trends Plant Sci 4: 9-13
7. **de Maagd RA, Weemen-Hendriks M, Stiekema W, Bosch D** (2000) *Bacillus thuringiensis* delta-endotoxin Cry1C domain III can function as a specificity determinant for *Spodoptera exigua* in different, but not all, Cry1-Cry1C hybrids. Appl Environ Microbiol 66: 1559-1563
8. **de Maagd RA, Bravo A, Crickmore N** (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet 17: 193-199
9. **Denolf P, Jansens S, Peferoen M, Degheele D, Van Rie J** (1993) Two different *Bacillus thuringiensis* delta-endotoxin receptors in the midgut brush border membrane of the European corn borer, *Ostrinia nubilalis* (Hübner) (Lepidoptera: Pyralidae). Appl Environ Microbiol 59: 1828-1837
10. **Escriche B, Martinez-Ramirez AC, Real MD, Silva FJ, Ferré J** (1994) Occurrence of three different binding sites for *Bacillus thuringiensis* delta- endotoxins in the midgut brush border membrane of the potato tuber moth, *Phthorimaea operculella* (Zeller). Arch Insect Biochem Physiol 26: 315-327
11. **Herrero S, Borja M, Ferré J** (2002) Extent of variation of the *Bacillus thuringiensis* toxin reservoir: The case of the geranium bronze, *Cacyreus marshalli* Butler (Lepidoptera : Lycaenidae). Appl Environ Microbiol 68: 4090-4094
12. H**öfte H, van Rie J, Jansens S, Houtven AV, Vanderbruggen H, Vaeck M** (1988) Monoclonal antibody analysis and insecticidal spectrum of three types of lepidopteran-specific insecticidal crystal proteins of *Bacillus thuringiensis.* Appl Environ Microbiol 54: 2010-2017

13. **Kalman S, Kiehne KL, Libs JL, Yamamoto T** (1993) Cloning of a novel CryIC-type gene from a strain of *Bacillus thuringiensis* subsp. *galleriae*. Appl Environ Microbiol 59: 1131-1137
14. **Lim GT** (1992) Biology, ecology and control of cocoa podborer *Conopomorpha cramerella* (Snellen). In: Keane PJ and Putter CAJ (eds) Cocoa Pest and Disease Management in Southeast Asia and Australasia FAO Plant Production and Protection, paper No 112. Rome: FAO, pp 85-100
15. **Naimov S, Weemen-Hendriks M, Dukiandjiev S, de Maagd RA** (2001) *Bacillus thuringiensis* delta-endotoxin Cry1 hybrid proteins with increased activity against the Colorado potato beetle. Appl Environ Microbiol 67: 5328-5330
16. **Roush RT** (1998) Two-toxin strategies for management of insecticidal transgenic crops: can pyramiding succeed where pesticide mixtures have not? Philos Trans R Soc Lond B Biol Sci 353: 1777-1786
17. **Russel RM, Robertson JL, Savin NE** (1977) POLO: a new computer program for Probit analysis. ESA Bulletin 23: 209-213
18. **Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH** (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev 62: 775-806
19. **Wardojo S** (1992) Major pests and diseases of cocoa in Indonesia. In: Keane PJ and Putter CAJ (eds) Cocoa Pest and Disease Management in Southeast Asia and Australasia FAO Plant Production and Protection, paper No 112 Rome: FAO, pp 63-67
20. **Yao ZW, Liu SK** (1990) Two gracillariid insect pests attacking litchi and longan. Acta Entomol Sin 33: 207-212

Chapter 3

Modification of the *Bacillus thuringiensis* toxin *SN19* gene for improved expression in plants

Tetty Chaidamsari,[1] Samir Naimov,[2] and Ruud A. de Maagd[1]

[1]Business Unit Bioscience, Plant Research International B.V., P.O. Box 16, 6700 AA Wageningen, The Netherlands

[2]Department of Plant Physiology and Molecular Biology, University of Plodiv, Plovdiv, Bulgaria

Summary

Many foreign proteins are not expressed at a sufficient level in transgenic plants. One of the causes of the low expression of *Bacillus thuringiensis cry* genes is their bacterial origin. We have made, by a combination of mutagenic PCR and PCR by overlap extension, a number of single nucleotide changes in fragments encoding domain I or domain III of the hybrid *cry* gene SN19, in order to improve expression in transgenic dicotyledonous plants. The efficacy of these changes was tested by testing each modified fragment by itself combined with an existing synthetic fragment or with the two modified fragments combined. Constructs were tested and compared in transgenic *Arabidopsis thaliana* plants, with insect bioassays, after determining that SN19 protein is active against the two targets (*Plutella xylostella* and *Pieris rapae*), and with quantitative immunoassays to determine the protein levels. Modification in both fragments achieved a marked improvement in protein expression levels and frequency of insect resistant transgenic plants.

Introduction

Cocoa Pod Borer (CPB) *(Conopomorpha cramerella)* is the most important insect pest threatening cacao production in Indonesia. We are studying the possibility to develop insect-resistant transgenic *Theobroma cacao* trees producing a Cry toxin from *Bacillus thuringiensis (Bt),* effective against CPB larvae.

The crystal proteins or δ-endotoxins of *Bt* are highly specific insecticidal proteins against the larvae of certain lepidopteran, coleopteran or dipteran insects, and exhibit no activity against humans, other vertebrates and beneficial insects (3). In an earlier study twelve different Cry toxins of *Bt* were tested for activity against CPB larvae. Results of the bioassays showed that SN19 (and its parents Cry1Ba and Cry1Ia), Cry1Ab, and Cry1Ea have considerable activity against this insect (19).

Cry toxins have been expressed in a large number of plant species, among which tomato, potato, tobacco, cotton and maize. Transgenic plants expressing insecticidal proteins from *Bt* were first commercialized in 1996 (4). Successful integration of a transgene into the plant genome does not automatically result in expression. Many cases of poor expression of foreign genes, in including *cry* genes, in plants have been investigated. Two approaches were used to increase the level of Cry toxins in genetically modified plants. The first approach is selectively removing DNA sequences predicted to inhibit efficient plant gene expression at the translational, transcriptional or mRNA level such as Cry1Ac in tomato and Cry1Ab in tobacco (17, 22) or Cry9Aa2 in potato (8). Secondary structure of mRNA, improper polyadenylation, improper nuclear transport, and mRNA instability motifs may result in accumulation of only a low level of mRNA, and thus a small amount of protein. The second approach goes one step further and requires the use of fully synthetic genes. In such an approach the sequence is optimized in order to match codon usage in plants (16-18, 20).

SN19 is a hybrid gene encoding a protein consisting of domains I and III of Cry1Ba and domain II of Cry1Ia (14). In an earlier study, 28 single nucleotides in the domain II encoding sequence of Cry1Ia were changed, for eliminating putative polyadenylation sites, mRNA instability sequences and consecutive C+G and A+T stretches. This modified domain II encoding sequence of Cry1Ia in combination with synthetic (propriety) Cry1Ba domain I and III encoding sequences has been used in transformation of potato plants, resulting in resistance against Colorado potato beetle (a coleopteran), potato tuber moth larvae, and European corn borer (both lepidopterans) (13).

Our goal is to express the *SN19* gene in various dicotyledonous plants, among which cocoa. Domain II of the hybrid *SN19* gene was modified previously, (14) and in this study we modified domain I and III of Cry1Ba, in order to construct our own optimized *SN19* gene. In our first approach we eliminated sequences that might cause mRNA destabilization and putative polyadenylation sites, by site directed mutagenesis. The purpose of redesigning domain I and III of the *SN19* gene was to create a modified gene that would be highly expressed in plants. We describe the construction of the modified domain I and III and the effect of these modifications on protein expression levels and insect resistance in the model plant *Arabidopsis thaliana.*

Material and methods

Modification of domains I and III of *cry1Ba*

For the modification of domains I and III of Cry1Ba, site directed mutagenesis by PCR with mutagenic primers and fragment assembly by recombination of overlapping fragments via PCR were performed (10, 13). Ten and five different primer pairs (Eurogentec), listed in Table 1 and 2, containing the desired mutations were used to modify domains I and domain III of *cry1Ba,* respectively. Four different DNA fragments of domain III were assembled in the first combinatorial PCR, as shown in Fig. 4. The parameters used during the first round PCR reactions were as follows: 95ºC, 1 min and 30 cycles: 95ºC, 30 s; 55ºC, 1 min; 68ºC, 2 min, and 68°C, 2 min. PCR fragments were analyzed by gel electrophoresis followed by purification of the fragments from the gel. For assembly and amplification of the overlapping fragments the following PCR program was used: 95ºC, 1 min and 35 cycles: 95ºC, 30 s; 68ºC, 1 min; 72ºC, 2 min and 72ºC, 4 min with the primers 18F and 5R. The amplified product from the first combinatorial PCR was used as a template for a second round of primary PCRs, followed by purification of fragments and a second combinatorial PCR with the primers 1BCSF and 25R. The PCR program used during the second round of mutagenic PCR was as described for the first round, except annealing and extension temperature were 62ºC, 1 min; and 68ºC, 2 min. The modification of domain I consisted of three rounds of mutagenic PCR. The PCR parameters of the first two rounds were the same as for the modification of domain III. The parameters used during the third round PCR reactions were as follows: 95ºC, 1 min and 35 cycles: 95ºC, 30 s; 65ºC, 1 min; 72ºC, 2 min, and 72ºC, 4 min. All PCR products were separated on 0.8% agarose gel and then purified using a QIAEX II agarose gel extraction kit

Table 1. Primers used for modification of the domain I encoding fragment of *cry1Ba*

ID	Nucleotide sequences
Cry1Ba-1For	5'-TTAACATAGCTGGTAGGATTCTAGGCGTATTGG-3'
Cry1Ba-1Rev	5'-CCAATACGCCTAGAATCCTACCAGCTATGTTAA-3'
Cry1Ba-2For	5'-CTTGTTGGTGAACTGTGGCCAAGGGGAAGAGATCAGTGG-3'
Cry1Ba-2Rev	5'-CCACTGATCTCTTCCCCTTGGCCACAGTTCACCAACAAG-3'
Cry1Ba-3For	5'-CGAACAACTTATCAACCAGCAGATTACAGAAAATGC-3'
Cry1Ba-3Rev	5'-GCATTTTCTGTAATCTGCTGGTTGATAAGTTGTTCG-3'
Cry1Ba-4 For	5'-GTGTTCTTTATACCCAGTACATCGCTTTAGAACTTGATTTCC-3'
Cry1Ba-4Rev	5'-GGAAATCAAGTTCTAAAGCGATGTACTGGGTATAAAGAACAC-3'
Cry1Ba-5For	5'-GCTTTAGAACTTGATTTCCTGAACGCTATGCCGCTTTTCGC-3'
Cry1Ba-5Rev	5'-GCGAAAAGCGGCATAGCGTTCAGGAAATCAAGTTCTAAAGC-3'
Cry1Ba-6For	5'-CGCTTTTCGCAATCAGGAATCAGGAAGTTCCATTATTGATGG-3'
Cry1Ba-6Rev	5'-CCATCAATAATGGAACTTCCTGATTCCTGATTGCGAAAAGCG-3'
Cry1Ba-7For	5'-TGCTCAAGCTGCCAACCTACACCTGCTATTATTGAGAGATGC-3'
Cry1Ba-7Rev	5'-GCATCTCTCAATAATAGCAGGTGTAGGTTGGCAGCTTGAGCA-3'
Cry1Ba-8For	5'-AGAATGGTATAACACAGGTCTAAATAGC-3'
Cry1Ba-8Rev	5'-GCTATTTAGACCTGTGTTATACCATTCT-3'
Cry1Ba-9For	5'-GCAAGTTGGGTAAGGTACAATCAGTTCCGTAGAGATC-3'
Cry1Ba-9Rev	5'-GATCTCTACGGAACTGATTGTACCTTACCCAACTTGC-3'
Cry1Ba-10For	5'-TCGCACTTATCCCATCAACACTAGTGCTCAGTTAAC-3'
Cry1Ba-10Rev	5'-GTTAACTGAGCACTAGTGTTGATGGGATAAGTGCGA-3'
Cry1Ba- 11Rev	5'-CCACATATTCATATACTGAGTGTTACTCCATCG-3'

(Qiagen). For all PCR steps, Pfu-Turbo DNA polymerase (Stratagene) was used. The modified fragments were ligated into pGEM-T Easy resulting in pTC6 and pTC1 for modified domain I and III respectively, and were sequenced in both directions.

Construction of plant transformation vectors

pTC7 containing the modified domain I of *cry1Ba*, the previously modified *cry1Ia* domain II and the synthetic *cry1Ba* fragment encoding domain III made available from Syngenta (13), was constructed by ligating a *Nco*I–*Mun*I fragment of pTC6 containing the modified domain I

Table 2. Primers used for modification of the domain III encoding fragment of *cry1Ba*

ID	Nucleotide sequences
5R	5'GGCCAGCCCTCAACTCTTG-3'
18F	5'CACATGTGAGAGCTTCGGTATATTC-3'
19R	5'-GGAAGTTCGGAAGCTCTTACCATTGGGATTTGGGTGATGCGGTTGGGTCCAATCG-3'
19F	5'-CGATTGGACCCAACCGCATCACCCAAATCCCAATGGTAAGAGCTTCCGAACTTCC-3'
2122R	5'-ACAAAATTTCCGTACTTCAGTTCGTCTCCACTGTTCATGGTACGTAGGAATCTG AAGTTGTTCAC AGTAGTACCTC-3'
2122F	5'-GAGGTACTACTGTGAACAACTTCAGATTCCTACGTACCATGAACAGTGGAGA CGAACTGAAGTAC GGAAATTTTGTG-3'
24R	5'-CTGGAATAATCTCGATCTTGTCTATATACAC-3'
24F	5'-GTGTATATAGACAAGATCGAGATTATTCCAG-3'
1B-CspF	5'-GGACGCATCGTAGTGCGGACCGTACGAATACGATTGG-3'
20R	5'-GCCCACCAGTGAAGCCTGGTCCTC-3'
20F	5'-GAGGACCAGGCTTCACTGGTGGGG-3'
23R	5'-GACGTTCGAATGATGTCCTGGATTTGTGTAAAAG-3'
23F	5'-CTTTTACACAAATCCAGGACATCATTCGAACTGC3'
25R	5'-AGAGCATTCACCGCGGATCCCTATCTTTCTAGGTCGTATTCTGCTTCG-3'

of *cry1Ba* (836bp), and a *Mun*I-*Bgl*II fragment of SN45 (13) containing the previously modified *cry1Ia* domain II and the synthetic *cry1Ba*-fragment encoding domain III into *Nco*I-*Bgl*II digested pBD1400. The entire gene was cut from pTC7 by *Nco*I and *Bgl*II digestion and ligated into the *Nco*I-*Bgl*II sites of pUCRBC1 (15), between the Chrysanthemum ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit promoter and terminator resulting in the expression cassette plasmid pTC8.

The fragment containing the modified domain III of *cry1Ba* (476 bp) was isolated by *Rsr*II and *Bam*HI digestion from pTC1 and used to replace the corresponding part of SN45, containing the synthetic *cry1Ba*-fragment encoding domain III, resulting in pTC2. The entire gene was cut from pTC2 by *Bsp*HI and *Bam*HI digestion and ligated into the *Nco*I-*Bgl*II sites of pUCRBC1, as described above, resulting in expression cassette plasmid pTC3.

An *Rsr*II-*Bam*HI fragment containing the modified domain III-encoding fragment of*cry1Ba* from pTC1 was cloned between the *Rsr*II and *Bgl*II-sites of pTC8 resulting in the expression cassette plasmid pTC11.

A binary vector containing unmodified *SN19*, as a negative control, was constructed by cutting of the Cry1Ba domain I and Cry1Ia domain II encoding fragment from SN19 using

*Nco*I and *Rsr*II. Primers 18F and 25R (Table 1) were used to amplify the Cry1Ba domain III-encoding fragment of *SN19*. The PCR program was as followed: 95ºC, 1 min and 35 cycles: 95ºC, 30 s; 55ºC, 1 min; 68ºC, 2 min, and 68ºC, 4 min. PCR fragments were analyzed by gel electrophoresis, followed by purification of the fragments from the gel. For the fragment containing the Cry1Ba domain III, the PCR product was digested with *Rsr*II and *Bam*HI. The fragment encoding the Cry1Ba domain I and Cry1Ia domain II, together with the Cry1Ba domain III-encoding PCR-fragment was ligated into the *Nco*I-*Bam*HI sites of pUCRBC1, as described above, resulting in expression cassette plasmid pTC13. All the expression cassettes were cloned in pBINPLUS using the *Hin*dIII and *Eco*RI sites flanking the cassettes, resulting in the binary vectors pTC4, pTC9, pTC12 and pTC13 (Fig. 5)

Plant transformation

All constructs were used for transformation of *Agrobacterium tumefaciens* strain Agl0 by electroporation, and introduced into *Arabidopsis thaliana* ecotype WS plants by the floral dip method (2). Resulting *Arabidopsis* seeds were surface sterilized by soaking 40 mg seeds in 70% ethanol for 1.5 min, and rinsing once with sterile water. The seeds were incubated in 4% bleach + Tween-20 (0.05%) for 10 min, followed by centrifugation at 1,000 g for 1 min. Subsequently the seeds were rinsed with sterile water four times, followed by centrifugation at 12,000 g for 30 sec in between. Next, the seeds were resuspended in sterile 0.1% agarose and plated on 50% MS medium containing 50 µg/ml kanamycin and 200 µg/ml cefotaxim, for selection of transformed seedlings. After 3 days at 4ºC, the plates were transferred to a growth chamber with long day condition (16 hrs light, 24ºC). Kanamycin-resistant plants were transferred to soil and the plants were grown in the greenhouse for bioassays.

Insect bioassays

Sensitivity of larvae of two insect species, *Pieris rapae* and *Plutella xylostella*, to SN19 protein was tested with a leaf dip assay. Untransformed leaves from *A. thaliana* ecotype WS were immersed in a test solution of Triton X-100 (0.05%), containing different concentrations of trypsin-activated SN19 protein for 10 sec and then allowed to dry with the adaxial leaf surface upwards on a Whatman filter paper for 15-30 min at room temperature. Control leaves were immersed in the test solution only. The leaves were then transferred to individual plastic Petri dishes (5 cm diameter) containing water agar (0.6% agar in water). Five larvae of *P.*

rapae or *P. xylostella* were placed on each leaf and each treatment was repeated three times on different days. Mortality was assessed after 4 days.

P. rapae and *P. xylostella* larvae were also used for bioassays with transgenic *Arabidopsis* leaves. All bioassays were performed on fully-grown detached leaves stuck in water agar. Eight or ten neonate larvae of *P. rapae* or *P. xylostella*, respectively, were placed on the upper leaf surface. Plants transformed with an unmodified *SN19* gene (pTC13) were used as negative controls. All bioassays were performed three times on different dates and mortality was assessed after 4 days.

Protein quantification

The expression levels of soluble SN19 protein in transgenic *A. thaliana* leaves were assayed using an antiserum against purified SN19 protein raised in rabbits. Leaf tissue (0.2 g) was ground with 400 µl of extraction buffer (50 mM NaOH, 20 mM NaS_2O_5, 5 mM EDTA and 10% Polyvinylpolypyrrolidone), subsequently neutralized with 80 µl 1 M Tris-HCl, pH 5.5, and centrifuged at 16,000 g for 10 min. The supernatant was transferred into a new eppendorf tube and additionally centrifuged at 16,000 g for 10 min. Protein concentrations in the supernatant were determined by the Bradford method (Bio-Rad Laboratories). The amount of SN19 protein was estimated by dot-blot analysis as follows. Equal amounts of soluble leaf protein (20 µg) were transferred to a nitrocellulose membrane using an S&S Manifold Dot blotter (Schleicher & Schuell). Immunological detection of SN19 protein was performed by treating the membrane with blocking solution containing Tris-buffered-saline (TBS: 10 mM Tris-HCl, pH 7.6, 150 mM NaCl), 5% (w/v) non-fat dry milk, and 3% (w/v) Bovine serum albumin for 1 h. The membrane was subsequently washed three times with TBST buffer (TBS buffer, with 0.2% Tween-20). 1:1000 Anti-SN19 serum diluted in TBST was applied and the membrane was incubated for 1 h at room temperature. After three times washing 5 min each with TBST, horseradish peroxidase conjugated anti-rabbit Ig (Amersham) was added (1:3000) and incubated for 1 h. The membranes were washed three times for 5 min each with TBST buffer, and once with carbonate buffer (0.1 M $NaHCO_3$, 1.0 mM $MgCl_2$, pH 9.8). The membranes were subsequently developed with lumi-light western blotting substrate (Roche). The membrane was covered with transparent plastic and the results were visualized with a chemiluminescent assay for 30 sec to 2 min. Serial dilutions of trypsin-activated SN19 in phosphate buffed saline (10 mM Na_2HPO4/KH_2PO_4, 0.8% NaCl) added to untransformed

Fig. 1. Modifications made in the domain I-encoding part of the wild-type *cry1Ba* gene. The sequence of the Cry1Ba domain I-encoding fragment and (below) the primers used for mutagenesis are shown. Only the nucleotides to be altered are shown in the primer sequences.

control plant extracts were used for comparison and estimation of SN19 content in leaf tissues.

Results

Modification of domains I and III of the *SN19* gene

SN19 is a hybrid toxin consisting of domains I and III of Cry1Ba, and domain II of Cry1Ia. In our previous study we combined modifications in domain II with domains I and III of a propriety synthetic *cry1Ba* gene (pSN48, see Fig. 5) (13). In this manuscript we describe our study on the effects of a limited number of modifications in wild type Cry1Ba domains I and III encoding sequences on SN19 protein expression and on insect resistance in the model plant *Arabidopsis thaliana.* In order to independently assess effects of modifications in domain I and in domain III, respectively, we choose to test the modified domains, individually and combined, in the background of the original SN48 construct, i.e. combined with the modified domain II or with modified domain II as well as a synthetic domain I or

Fig. 2. Modifications made in the domain III-encoding part of the wild-type *cry1Ba* gene The sequence of the Cry1Ba domain III-encoding fragment and (below) the primers used for mutagenesis are shown. Only the nucleotides to be altered are shown in the primer sequences.

domain III. All combinations used for transformation are shown in Fig. 5.

By a combination of PCR with mutagenic primers and annealing of fragments by overlap extension, 37 and 22 single nucleotides were changed to eliminate putative polyadenylation sites (ATTTA and AACCAA), mRNA instability sequences and consecutive C+G and A+T streches from domain I (Fig. 1) and domain III (Fig. 2) of SN19. The strategies for PCR assembly of the modified fragments is shown in Fig. 3 (domain I) and in Fig. 4 (domain III). Three different versions of a modified *SN19* gene were made and cloned between a promoter and terminator fragment derived from chrysanthemum. This resulted in transformation vectors pTC9, pTC4, and pTC12, respectively (Fig. 5). These constructs were introduced in *Arabidopsis thaliana* ecotype WS by *Agrobacterium tumefaciens*-mediated transformation. pSN48 and pTC13 were used as a positive and negative control respectively. Twenty-three, 21, 24, 22, and 17 transgenic plants per construct from pTC9, pTC4, pTC12, pSN48 and pTC13, respectively, were obtained and transferred to the greenhouse.

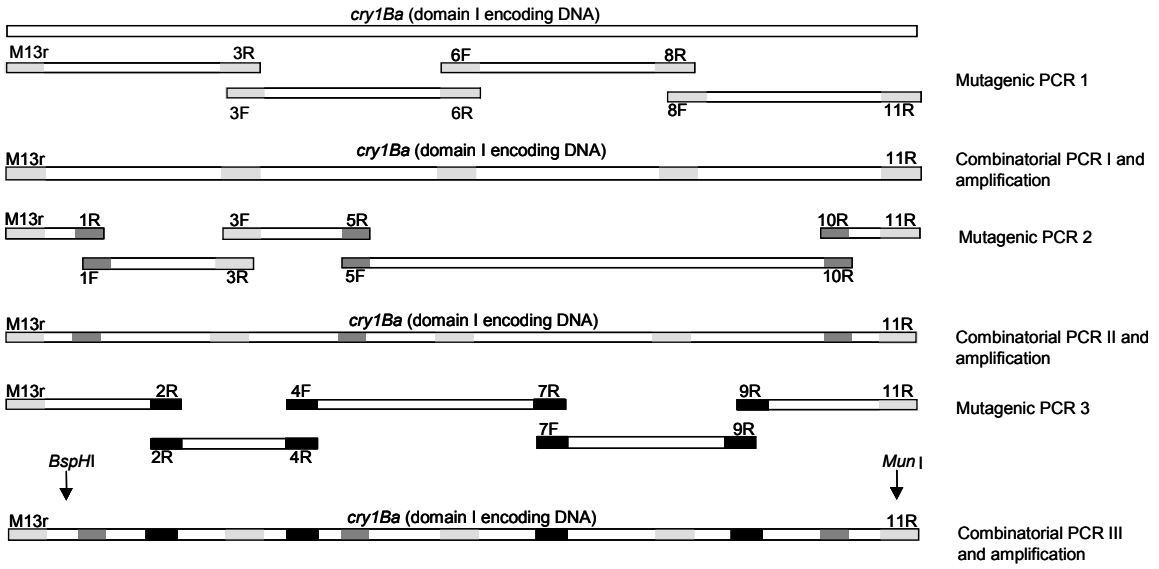


Fig. 3. Mutagenesis and recombination via PCR strategy for the domain I-encoding fragment of the *cry1Ba* gene. In separate amplifications, four fragments of the domain I fragment of the c*ry1Ba* gene were amplified using mutagenic primers. Denatured fragments annealing in the overlapping primer region were extended. This resulted in a full-length product, which was amplified using the two flanking primers (M13rev, Cry1Ba11Rev), and subsequently used as a template for the second round of mutagenic PCR reactions.

Insect Bioassays

First, we performed a leaf dip assay to determine sensitivity of *Plutella xylostella* (Diamondback moth) and *Pieris rapae* (Small White) to the SN19 protein. Five and seven larvae, of *P xylostella* and *P. rapae*, respectively, were tested with five different concentrations of SN19 toxin ranging from 0.005 to 50 µg/ml and the percentage of mortality was assessed each day up to four days. For both insects 0.05 µg/ml of SN19 toxin was the lowest concentration that was sufficient to give 100% mortality after 3 days. This is the first study showing that SN19 has activity against those two insect species. The results of the leaf dip assay suggested that both insects are also suitable test subjects for detection and estimation of protein levels in transgenic *Arabidopsis.*

Single leaves of transgenic *Arabidopsis* plants were used for insect bioassays with *P. xylostella* or *P. rapae*. All plants were classified, according to the percentage of mortality they caused, in one four classes (0, 0-50, 50-75 and 100% mortality). Results of the bioassays were equal for both insect species and are presented in Fig. 6. Positive control leaves (pSN48) resulted in 19 out of 22 plants with 100% mortality after 2 days. No sign of tunnels or visible damage were recorded from all of the plants with 100% mortality. All of the negative control

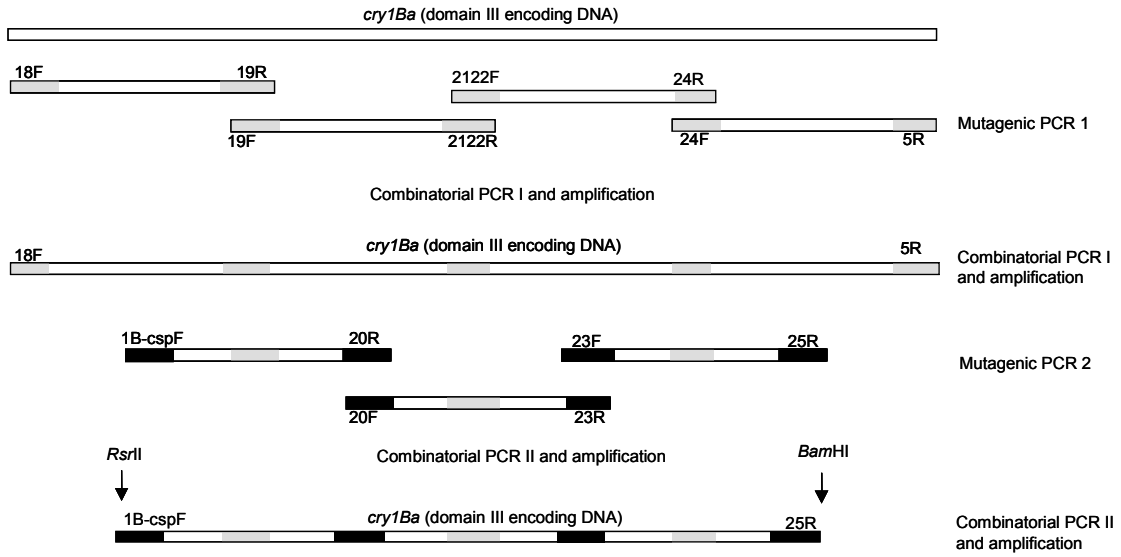


Fig. 4. Mutagenesis and recombination via PCR strategy for the domain III-encoding fragment of the *cry1Ba* gene. In separate amplifications, four fragments of the domain III fragment of c*ry1Ba* gene were amplified using mutagenic primers. Denatured fragments, annealing in the overlapping primer region, were extended by PCR. This results in a full-length product, which was amplified using the two flanking primers (18F and 5R), and subsequently used as a template for the second round of mutagenic PCR reactions.

leaves (pTC13) gave 0% mortality. Twelve plants out of 23 of the construct with modified coding sequence for domain I of Cry1Ba (pTC9) gave 100% mortality in all three replicate experiments. Six out of 24 transgenic plants transformed with pTC4 (modified domain III coding sequence) caused 100% mortality in all three experiments for both insects after 2 days, and 9 out of 24 plants transformed with pTC12 (both domain I and III coding sequences modified) caused 100% mortality.

Since both insect species were quite sensitive to the SN19 protein, transgenic plants giving 100% mortality may still contain a wide range of active protein concentrations. Therefore, leaves of transgenic plants with 100% mortality from all constructs were analyzed immunochemically for estimating the expressed amounts of the SN19 protein. For comparison, expression levels were again subdivided in four classes, and results are shown in Fig. 7. From the results it is clear that both insect species are indeed sensitive enough to be killed by levels of SN19, which were too low to be detected in our immunoassay. None of the tested 6 and 12 plants from pTC4 and pTC9, respectively, had detectable Cry protein levels

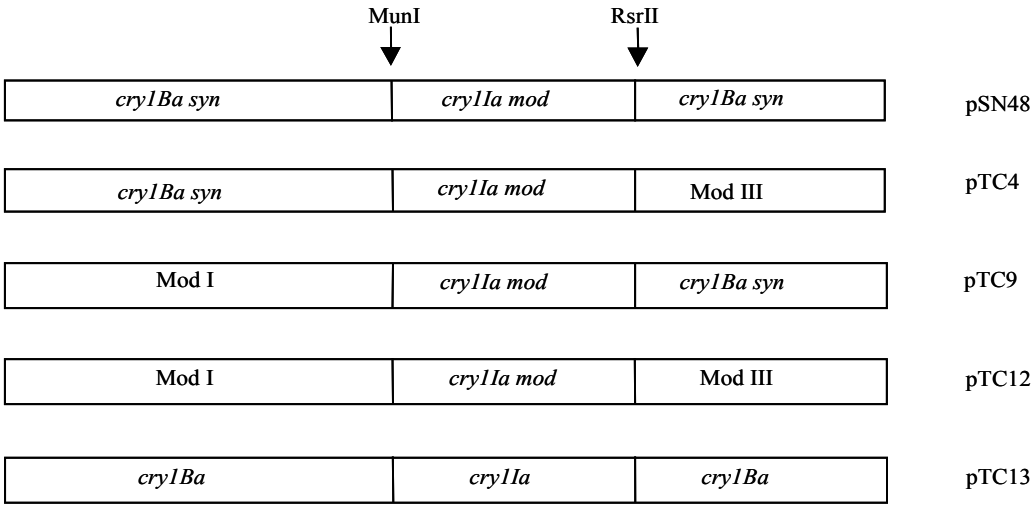


Fig. 5. Constructs used for transformation of *A. thaliana*. Included are the positive control (pSN48), containing the synthetic domain I and domain III-encoding fragments, and the negative control (pTC13) containing the wild type domains.

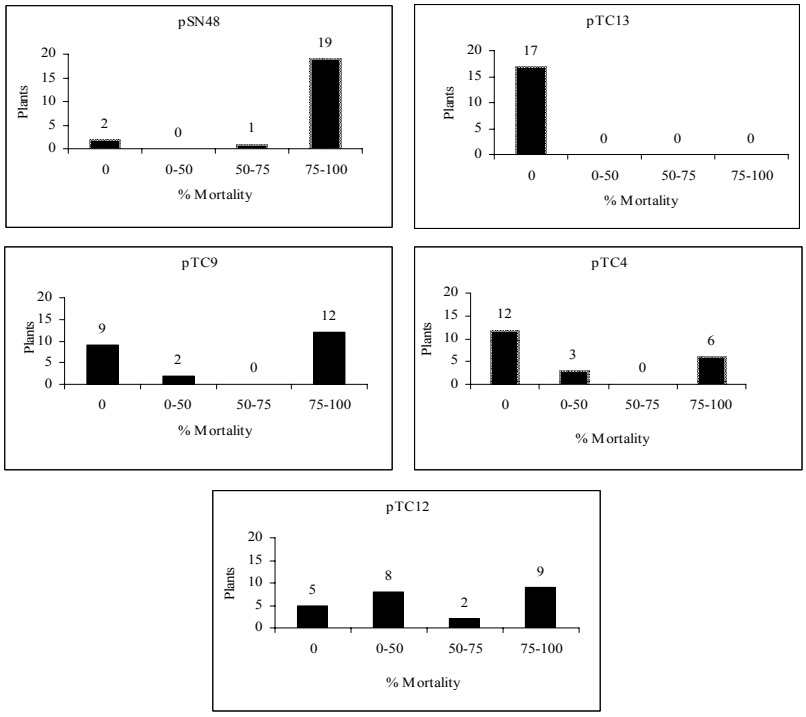


Fig. 6. Mortality (%) of *P. xylostella* and *P. rapae* larvae in bioassays on transgenic *Arabidopsis* leaves containing the various combinations of SN19 domain encoding fragments (see Fig. 5). Results were identical for the two insects

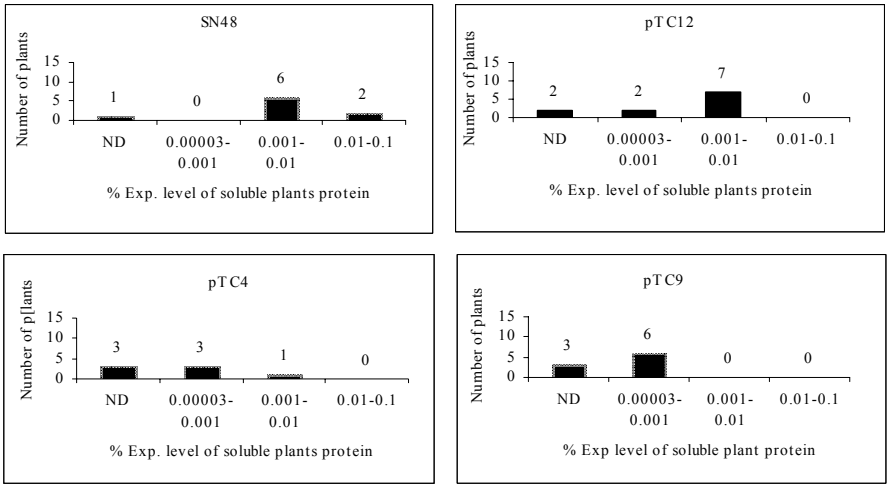


Fig. 7. Quantification of SN19 protein in soluble leaf proteins of transgenic *Arabidopsis*. Only plants giving 100% mortality (see Fig. 5) were tested.

above 0.01% of soluble leaf protein, yet some pTC12 plants, with both modified domains I and III showed levels of Cry protein between 0.001-0.01% of soluble protein. In contrast, positive control plants, containing both synthetic domains I and III encoding fragments (pSN48) had Cry protein levels up to between 0.01-0,1% of soluble protein.

Discussion

Gene transfer into plants has become a relatively simple and routine process in many crops, but integration of a transgene into a plant genome does not always result in a high level of expression of that gene. In some cases, poor expression of foreign genes in transgenic plants has been investigated (1, 7, 9, 11, 12). The problem can occur at one or more steps of gene expression in the transgenic plant. There are several methods for increasing expression of a particular gene in a plants host. Addition of an intron to the transcript and use of a proper untranslated leader for a particular host plant, removing RNA instability signals, optimizing codons used. Codon usage in native δ-endotoxin genes is clearly different from that found in typical plant genes and native genes have very low G+C content, around 37% (17, 21). mRNA accumulation may be limited at the level of transcription, but the problem may occur post transcriptionally as a result of splicing or degradation of the transcripts (7). Improper splicing, improper polyadenylation, improper nuclear transport, or instability of mRNA can all result in accumulation of only a low level of both mRNA and protein.

Transgenic tomato or tobacco plants containing a native *cry* gene, under control of the CaMV35S promoter were first reported in 1987 (1, 7, 21). Those studies reported the protein level to be too low for a quantitative estimation by immunoassay, although plants were reported to be resistant to the tobacco hornworm in laboratory and field tests. Insects that were not as sensitive to Bt toxins as tobacco hornworm would require considerably higher toxin levels in order to be controlled, which initiated a number of studies in the causes of low *cry* gene expression, and of methods to improve this. The poor expression of *cry* genes in plants can be caused by a very low steady state level of the transcript or premature polyadenylation of transcripts of the *cry* gene (5, 6, 12).

We have produced transgenic *Arabidopsis thaliana* plants carrying modified *SN19* genes under control of the Chrysanthemum ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit promoter. Although a version of *SN19* had been used before successfully to obtain transgenic insect-resistant potato plants (13), we had two reasons to attempt novel or further improvement on a wild type SN19 gene: i) To obtain freedom to operate with a version of SN19 modified by ourselves. ii) As the synthetic *cry1Ba* gene was tuned for high expression in monocotyledonous plants and codon usage is different between monocots and dicots further improvement of the expression could be possible.

Even though our modified gene (construct pTC12) is not quite as good as the earlier version (*SN48*), we did reach considerable improvement of expression over the wild type version, sufficient to render *Arabidopsis* resistant to two insect species' larvae. Whereas this may be sufficient for many applications, further improvements may still be obtained by studying in detail the causes for suboptimal expression at either transcriptional or translational level and testing directed changes to fix these. Alternatively, domain I and III encoding fragments could be synthesized de novo and optimized for codon usage in dicots, while trying to prevent sequences causing low expression for the reasons discussed above. We are currently taking this approach.

Literature cited

1. **Barton, K. A., H. R. Whiteley, and N.-S. Yang.** 1987. *Bacillus thuringiensis* delta-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to lepidopteran insects. Plant Physiol. **85:**1103-1109.
2. **Clough, S. J., and Bent, A.F.** 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:**735-43.

3. **de Maagd, R. A., A. Bravo, and N. Crickmore.** 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet. **17:**193-199.

4. **de Maagd, R. A., D. Bosch, and W. J. Stiekema.** 1999. *Bacillus thuringiensis* toxin mediated insect resistance in plants. Trends Plant Sci. **4:**9-13.

5. **De Rocher, E. J., T. C. Vargo Gogola, S. H. Diehn, P. J. Green, and E. J. de Rocher.** 1998. Direct evidence for rapid degradation of *Bacillus thuringiensis* toxin mRNA as a cause of poor expression in plants. Plant Physiol. **117:**1445-1461.

6. **Diehn, S. H., W. Chiu, E. J. d. Rocher, P. J. Green, E. J. de Rocher, and W. L. Chiu.** 1998. Premature polyadenylation at multiple sites within a *Bacillus thuringiensis* toxin gene-coding region. Plant Physiol. **117:**1433-1443.

7. **Fischoff, D. A., K. S. Bowdish, F. J. Perlak, P. G. Marrone, S. McCormick, J. G. Niedermeyer, D. A. Dean, K. Kusano-Kretzmer, E. J. Mayer, D. E. Rochester, S. G. Rogers, and R. T. Fraley.** 1987. Insect tolerant transgenic tomato plants. Bio-Technol. **5:**807-813.

8. **Gleave, A. P., D. S. Mitra, N. P. Markwick, B. A. M. Morris, and L. L. Beuning.** 1998. Enhanced expression of the *Bacillus thuringiensis* cry9Aa2 gene in transgenic plants by nucleotide sequence modification confers resistance to potato tuber moth. Mol. Breeding **4:**459-472.

9. **Haffani, Y. Z., S. Overney, S. Yelle, G. Bellemare, and F. J. Belzile.** 2000. Premature polyadenylation contributes to the poor expression of the *Bacillus thuringiensis cry*3Ca1 gene in transgenic potato plants. Mol. Gen. Genet. **264:**82-8.

10. **Ho, S., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989. Site-directed mutagenesis by overlap extension using polymerase chain reaction. Gene **77:**51-59.

11. **Koziel, M. G., N. B. Carozzi, and N. Desai.** 1996. Optimizing expression of transgenes with an emphasis on post-transcriptional events. Plant Mol. Biol. **32:**393-405.

12. **Murray, E. E., T. Rocheleau, M. Eberle, C. Stock, V. Sekar, and M. Adang.** 1991. Analysis of unstable RNA transcripts of insecticidal crystal protein genes of *Bacillus thuringiensis* in transgenic plants and electroporated protoplasts. Plant Mol. Biol. **16:**1035-1050.

13. **Naimov, S., S. Dukiandjiev, and R. A. de Maagd.** 2003. A hybrid *Bacillus thuringiensis* delta-endotoxin gene gives resistance against a coleopteran and a lepidopteran pest in transgenic potato. Plant Biotechnol. J. **1:**51-57.

14. **Naimov, S., M. Weemen-Hendriks, S. Dukiandjiev, and R. A. de Maagd.** 2001. *Bacillus thuringiensis* delta-endotoxin Cry1 hybrid proteins with increased activity against the Colorado potato beetle. Appl. Environ. Microbiol. **67:**5328-5330.

15. **Outchkourov, N. S., J. Peters, J. de Jong, W. Rademakers, and M. A. Jongsma.** 2003. Novel rubisco small subunit promoter from chrysanthemum (*Dendranthema grandiflora* L) yields very high foreign gene expression levels in plants. Planta **21:**1003-1012.

16. **Perlak, F. J., R. W. Deaton, T. A. Armstrong, R. L. Fuchs, S. R. Sims, J. T. Greenplate, and D. A. Fischoff.** 1990. Insect resistant cotton plants. Bio-Technol. **8:**939-943.

17. **Perlak, F. J., R. L. Fuchs, D. A. Dean, S. L. McPherson, and D. A. Fischoff.** 1991. Modification of the coding sequence enhances plant expression of insect control protein genes. Proc. Natl. Acad. Sci. USA **88:**3324-3328.

18. **Perlak, F. J., T. B. Stone, Y. M. Muskopf, L. J. Petersen, G. B. Parker, S. A. McPherson, J. Wyman, S. Love, and G. Reed.** 1993. Genetically improved potatoes: Protection from damage by Colorado potato beetles. Plant Mol. Biol. **22:**313-321.

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

20. **Strizhov, N., M. Keller, J. Mathur, Z. Konczkalman, D. Bosch, E. Prudovsky, J. Schell, B. Sneh, C. Koncz, and A. Zilberstein.** 1996. A synthetic *cryIC* gene, encoding a *Bacillus thuringiensis* delta-endotoxin, confers *Spodoptera* resistance in Alfalfa and Tobacco. Proc. Natl. Acad. Sci. **U**SA **93:**15012-15017.

21. **Vaeck, M., A. Reynaersts, H. Höfte, S. Jansens, M. D. Beuckeleen, C. Dean, M. Zabeau, M. V. Montagu, and J. Leemans.** 1987. Transgenic plants protected from insect attack. Nature (London) **328:**33-37.

22. **Van der Salm, T., D. Bosch, G. Honee, L. Feng, E. Munsterman, P. Bakker, W. J. Stiekema, and B. Visser.** 1994. Insect resistance of transgenic plants that express modified *Bacillus thuringiensis cryIA(b)* and *cryIC* genes: a resistance managment strategy. Plant Mol. Biol. **26:**51-59.

Chapter 4

Gene expression in the cocoa fruit pod wall and pulp

Tetty Chaidamsari[1], Mieke Weemen[1], Petra Bakker[1], Djoko Santoso[2] and Ruud A. de Maagd[1]

[1]Business Unit Bioscience, Plant Research International B.V, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[2]Biotechnology Research Institute for Estate Crops, Jl Taman Kencana 1, Bogor 16151, Indonesia

Summary

In this paper we describe a first venture into characterization of gene expression in various tissues of the fruit (pod) of the cocoa tree, *Theobroma cacao* L. Genes, highly expressed in these tissues, may yield promoters that can be used to control expression of transgenes aimed at protection against pathogens or pests attacking the pod. Two approaches were used to identify genes that are specifically expressed in the fruit pod wall or pulp, or both, in contrast to expression in other tissues such as the beans and leaves. The first approach was sequencing of randomly selected clones of cDNA libraries from the inner pod wall or pulp, which revealed many not previously identified expressed sequence tags (ESTs). Genes highly expressed in the pod wall or pulp were selected by reverse Northern dot-blotting. Further investigation of a selection of these showed that they are often highly expressed in the pod wall as well as in the pulp, and sometimes also in the leaf and bean. In order to obtain genes that are more specifically expressed in the pod wall, in the second approach we made a library of fragments from pod wall cDNA subtracted with mixed leaf and bean cDNA by Suppression Subtractive Hybridization (SSH). This yielded an EST library for which sequencing of randomly selected clones showed little overlap with the previous two libraries. Analysis of expression by Northern blotting revealed that the most frequent member of this library, homologous to the cotton *Lea5-D* gene, is almost exclusively expressed in the inner pod wall throughout development of the fruit, and therefore a good candidate for promoter isolation.

Introduction

Cacao is a vital export crop for many countries, particularly in Southeast Asia, West Africa, and South America. The fruit (pod) of the cocoa tree (*Theobroma cacao* L) holds the beans that are used for cacao production. Seeds develop in the pod and after harvest they are removed from the pod for fermentation, producing the characteristic cacao flavors. The pod is also the entry point for some of the most serious pathogens and insect pests which, depending on the geographical location, inflict serious economical damage (4). Species of the oomycete *Phytophthora* (*P. palmivora, P. megakarya, P. capsici, P. citrophthora,* depending on the region) cause black pod disease, also called Phytophthora pod rot (9). The basidiomycete *Crinipellis perniciosa* (Stahel) Singer is the causal agent of witches' broom disease, which occurs only in South America (14). Although as the name implies, the disease affects vegetative growth, infection of the pod also occurs and can lead to serious damage. Frosty pod rot or Monilia pod rot is caused by the basidiomycete *Moniliophthora roreri* Cif., which in the field more specifically infects pods, and occurs primarily in South America (12).

The cocoa pod borer (CPB), *Conopomorpha cramerella* (Snellen) is the most serious insect pest of the cocoa tree in Southeast Asia. The larvae of CPB directly bore the pod wall after hatching, making it difficult to control by chemicals or natural enemies. Inside the pod, the larvae feed themselves on the pulp surrounding the beans, causing deformation and clumping of the beans and premature ripening (7, 18).

Genetic modification of crop plants is a promising way of offering protection against pests and diseases, particularly if natural resistance is not available or difficult to apply. We have previously shown that some *Bacillus thuringiensis* (Bt)-derived toxins have activity against CPB larvae (15). Our long-term goal is to develop transgenic cocoa plants expressing one or more active Bt toxins in the pod wall for protection against this insect.

In order to successfully develop transgenic approaches for protection of the cocoa pod against diseases and pests, a number of problems will have to be resolved. In order to express a transgene or its indirect products in the pod to a sufficient level, a plant promoter with activity in the pod wall and/or pulp will be required. Moreover, to decrease the risk of the possible health or environmental effects (or more indirectly, to quench public anxiety) of expressing transgenes in tissues other then targeted by the pest organism (particularly the bean), such a promoter should preferably be highly specific for the target tissue. Characterization of gene expression processes in the pod during development or during

response to pathogens may also help to better understand and, where possible, to improve natural resistance of the pod wall to pests and pathogens.

In this manuscript, we describe the first steps towards characterizing gene expression in the developing cacao pod wall and in the pulp surrounding the beans. Two approaches were used to isolate (fragments of) pod wall and/or pulp-specific cDNAs. Using Northern blotting we studied the tissue specific expression of a selection of the isolated genes.

Material and Methods

Plant material

The plant material was obtained from the cacao tree clones RCC72 and Sca6, which were grown in fields of the Indonesian Coffee and Cacao Research Institute, Jember Indonesia. Fruits in different stages of development were shipped to Wageningen and there sliced, immediately frozen in liquid nitrogen, and stored at –80^{o}C until further use. Leaves were harvested from plants grown in the laboratory.

RNA extraction

The method used for RNA extraction from cocoa tissues is a modification of a method described earlier for banana tissues (2). One gram of tissue, mixed with 1.5% polyvinylpirrolidone was ground by hand using a pestle and mortar in the presence of liquid nitrogen. The powder was transferred to a 50 ml tube, resuspended in 15 ml of preheated (65ºC) extraction buffer (100 mM Tris-HCl pH 8.2, 1.4 M NaCl, 20 mM EDTA, 2 % (w/v) CTAB, and 75 µl of β-mercaptoethanol) and homogenized in a warring blender at maximum speed for two minutes. Next, the homogenate was incubated at 65°C for 1h, while mixing by gentle vortexing every 15 min. The tubes were kept at room temperature for 5 min and then an equal volume of chloroform: isoamyl alcohol (24:1) was added. The mixture was shaken vigorously until an emulsion was formed, and then centrifuged at 12,000 g for 15 min at room temperature. The upper aqueous phase containing nucleic acids and polysaccharides was carefully transferred to a new tube, and extracted subsequently with phenol, phenol: chloroform: isoamyl alcohol (25:24:1), and chloroform: isoamyl alcohol (24:1), each time followed by centrifugation at 12,000 g for 15 min. The aqueous phases were transferred to new tubes and 10 M LiCl was added to a final concentration of 2 M. The RNA was allowed to precipitate overnight at 4°C and then recovered by centrifugation at 17,000 g at 4°C for 30

min. The pellet was dissolved in milliQ grade water and was extracted with phenol, phenol: chloroform: isoamyl alcohol (25:24:1), and chloroform: isoamylalcohol (24:1), respectively each time followed by centrifugation 12,000 g for 15 min at 4ºC. The supernatant was transferred to a new tube and RNA was precipitated with 0.1 volumes of 3 M sodium acetate pH 5.8 and 3 volumes of 100 % ethanol at –70°C for 4 hrs to overnight. The RNA was recovered by centrifugation at 17,000 g in a microfuge at 4°C for 30 min. The pellet was washed with an equal volume of 70% cold ethanol, the ethanol was allowed to evaporate at room temperature for 15 min, and the purified RNA pellet was then resuspended in an appropriate volume of nuclease-free water. The purity and concentration of the purified RNA was determined spectrophotometrically at 230, 260, and 280 nm (2, 16).

cDNA library construction of the inner pod wall and pulp

The synthesis of cDNA from total RNA was carried out using the Smart PCR cDNA Construction kit (CLONTECH, Palo Alto, CA) according to the kit instructions. For the first strand cDNA synthesis 1 µg of total RNA was used. Double stranded cDNA was produced from 2 µl of first strand cDNA mixed with 10 mM dNTPs, Advantage 2 Polymerase, 5' PCR primer, and CDS III/3' PCR primer in a PE Biosystems DNA thermal Cycles 2400. The following PCR program was used: 95ºC for 20 s followed by 29 cycles: 95ºC, 5 sec; 68ºC, 6 min. To inactivate the polymerase activity, 2 µl of proteinase K (20 µg/µl) was added and incubated for 20 min at 45ºC, and then was followed by extraction with phenol: chloroform: isoamyl alcohol (25:24:1). The supernatant was transferred into a new tube and cDNA was precipitated with 0.1 volume of 3M sodium acetate, 1.3 µl of glycogen (20 µg/µl) and 2.5 volume of room-temperature 95% ethanol, followed by centrifugation for 20 min at 14.000 rpm. The pellet was washed with an equal volume of 80% ethanol, the ethanol was allowed to evaporate at room temperature for 10 min, and the purified cDNA pellet was then resuspended in 79 µl deionized H_2O. The purified ds cDNA was digested with *Sfi*I (16 units) at 50ºC for 2 hr and cDNA was fractionated by centrifugation at 400 g for 2 min through a "400 Spun Column" (Sepharose CL-40, Amersham). The eluate, containing the ds cDNA fraction larger then approximately 400 bp was ligated into a *Sfi*1-site previously created in the pGEM-T Easy vector, at 4ºC overnight. The ligated DNA was finally transformed into *E. coli* JM109 (*rec*A1, Promega) super competent cells.

In order to introduce a *Sfi*1-site into pGEM-T Easy, a double stranded *Sfi*1 linker was made. The double stranded oligo was made from *Sfi*1 and *Sfi*2 primers (Table 1). 5 µl (10

Table 1. Oligonucleotide primers used in this study

Name	Nucleotide sequence
Smart IV	AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG
CDS III/3'	ATTCTAGAGGCCGAGGCGGCCGACATG-d$(T)_{30}$ N-1N-
5' PCR Primer	AAGCAGTGGTATCAACGCAGAGT
*Sfi*1	GGCCATTATGGCCTGCAGGATCCGGCCGCCTCGGCCGA
*Sfi*2	CGGCCGAGGCGGCCGGATCCTGCAGGCCATAATGGCCA
cDNA synthesis primer	TTTTGTACAAGCTT$_{30}$ N_1N-
Adaptor1 SSH	CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT
Nested PCR SSH 1	TCGAGCGGCCGCCCGGGCAGGT
Adaptor 2R	CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT
Nested PCR SSH primer 2R	AGCGTGGTCGCGGCCGAGGT
M13 forward	GTAAAACGACGACGGCCAGT
M13 reverse	GGA AAC AGCTATGACCATG
T7	ATACGACTCACTATAGGGCG
SP6	ATTTAGGTGACACTATAGAATAC

nmol) of *Sfi*1 and *Sfi*2 were denatured by heating at 100ºC for 5 min in a block heater, and allowed to anneal by cooling to room temperature in the block. One ng of double strand oligo was ligated into pGEM-T Easy using T4 DNA ligase and the ligation mix was used to transform *E. coli* XL-1 blue competent cells.

Preparation of cDNA dot blots

The cDNA's from the inner pod wall and pulp libraries (96 colonies each) were amplified by colony PCR using T7 and SP6 primers (Table 1). The colony PCR was performed by picking up cells from white colonies on the X-gal indicator plates and suspending them in 10 µl milliQ grade water, followed by addition of 5 µl PCR reaction mix (1.5 µl 10x SuperTaq buffer, 0.3 µl dNTP's, 0.15 µl each of T7 and SP6 primers, 0.15 µl SuperTaq polymerase and 2.75 µl milliQ grade water). The parameters used during PCR reactions were as follows: 94ºC, 4 min. and 35 cycles: 94ºC, 30s; 50ºC, 30s; 72ºC, 1 min. In addition, 8 min at 72ºC for the last step was added. 1 µl of colony PCR product was diluted in 20 µl of DNA-dilution buffer [50 µg/ml herring sperm DNA in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0)]. The mixtures were heated for 10 min at 100°C, followed by incubation for 2 min on ice. 0.5

µl of diluted cDNA's was spotted on a nylon membrane (Roche Diagnostics, Almere, NL). The cDNA was then fixed by UV cross-linking (Crosslinker- Hoefer, San Fransisco, CA) for 30 sec.

Direct Labeling of cDNA probes with DIG-Chem-Link and dot blot hybridization

2.5 µg of total RNA from the inner pod wall or pulp was used for RT-PCR using the SUPERSCRIPT^{TM} II RNase H-Reverse Transcriptase kit and oligo-(dT) system (Invitrogen) in a single tube reaction protocol. The reverse transcriptase reaction was carried out for 2 hrs at 37ºC, followed by heat inactivating the AMV reverse transcriptase at 37ºC for 15 min. The cDNA was purified with a QIAquick (QIAGEN, Westburg BV, Leusden, NL) PCR purification kit. 27 µl of purified cDNA was then used to produce digoxigenin (DIG) labeled cDNA probes with DIG-Chem-Link kit (Roche) according to the manufacturer's instructions.

The inner pod wall and pulp cDNA dot blots were hybridized overnight with pod wall or pulp derived cDNA probes at 42ºC in Dig-easy Hyb buffer. The concentration of cDNA probe was 50-100 ng/ml hybridization buffer. The dot blots were then washed two times at room temperature for 10 min in 0.5 x SSC, 0.1 % SDS and 30 min at 68ºC in 0.5 x SSC, 0.1 % SDS. Hybridization was visualized by an enzyme-linked immunoassay, using alkaline phosphatase-conjugated anti-digoxigenin antibody and CDP *Star* (Roche) as a substrate. The light emitted from the blots was detected by CCD-camera in a Lumi-Imager set-up (Roche).

Dig-labeling of RNA and Northern blot analysis

The templates for in vitro transcription to produce RNA probes were made by PCR amplification of the complete inserts from 250 ng of the appropriate cDNA clones using the reverse and forward M13 primers (Table1). The amplified fragments were then used as templates for SP6/T7 RNA polymerase, with the Dig RNA *in vitro* transcription labeling kit (Roche) for 1 hr at 42ºC according to the manufacturer's instructions.

For Northern blots, 10 µg of denatured total RNA from the outer pod wall, inner pod wall, pulp, bean and leaf of cacao were separated on a 1% agarose gel containing 6% formaldehyde and MOPS buffer (200 mM MOPS, 50 mM sodium acetate, 20 mM EDTA, pH 7.0) for 3 hrs at 60 V. After electrophoresis, the separated RNA was blotted onto positively charged nylon membranes (Roche) by capillary transfer and hybridized with RNA probes overnight at 68ºC in Dig-easy Hyb buffer (Roche). After hybridization, the membranes were

washed two times at room temperature for 10 min in 0.5 x SSC, 0.1% SDS and 30 min at 68ºC in 0.5 x SSC, 0.1% SDS. Hybridization was visualized as described above for dot blots.

Isolation of pod wall-specific cDNAs by Suppression Subtractive Hybridization (SSH)

mRNAs from the inner pod wall, bean and leaf were isolated from total RNA using oligo $(dT)_{25}$ Dynabeads (GenoVision, Philadelphia, PA) as described in the manufacture's manual. Pod wall-specific cDNA fragments were isolated using the PCR-Select cDNA subtraction kit (CLONTECH) as described in the manufacturer's instructions. Starting material consisted of 1 µg inner pod wall mRNA as tester and 1 µg of leaf and bean mRNA (mixed equal amounts) as a driver. All SSH products were ligated into the pGEM-T Easy vector and transformed into JM109 super competent cells. Inserts from 96 cDNA clones were amplified by colony PCR using M13 forward and reverse primers (Table 1, see above) and were sequenced with the same primers.

Sequence analysis

The sequencing reaction was performed on PCR-amplified fragments using Bigdye reaction mix (Applied Biosystems) and was analyzed in a 3100 Genetic Analyzer (ABI Prisma). The parameters used during PCR reactions were as follows: 96ºC, 3 min; 32 cycles: 96ºC, 30s; 50ºC, 15s; 60ºC, 4 min. Sequences were analyzed and assembled into contigs using the Lasergene SeqMan II software (DNASTAR Inc., Madison, Wi.). Protein homology searches in the NCBI Genbank protein database were done using the BlastX algorithm (1), as were BlastX searches in the Solanaceae Genome Network tomato Unigene database (http://www.sgn.cornell.edu/cgi-bin/tools/blast/simple.pl). Nucleotide homology searches of the dbEST database and the TIGR cocoa gene index (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=cocoa) were done with the BlastN algorithm.

Results

Optimization of RNA Isolation

Cocoa fruits are pods, which mature in 5 to 7 months during which time many wilt and drop off in a natural thinning process. The pod grows to a length of 15 to 35 cm and a diameter of 10-15 cm. A pod normally contains 30 to 40 seeds (beans) surrounded by a mucilaginous pulp when the pod is ripe. The fruit does not soften during ripening and is non-dehiscent. The

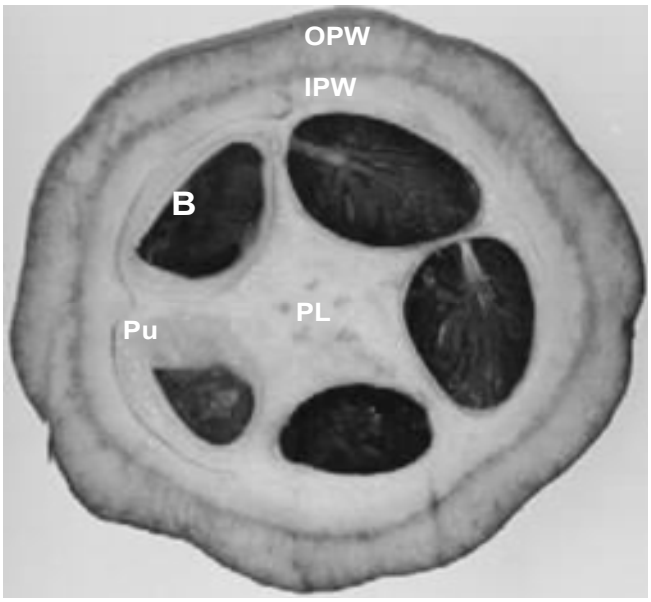


Fig. 1. Section through a ripe cocoa pod. Indicated are the outer pod wall (OPW), the inner pod wall (IPW), the pulp (PU), beans (B) and placenta (PL)

cacao seeds display no dormancy period: the process of germination will start immediately after the seeds are removed from the pod. Natural dissemination depends on animals feeding of the pulp and dropping seeds in the process. Fig. 1 shows a section of a ripe cocoa pod. Discernable are the thick cuticle, which can range in color from green, yellow to dark red, and the pod wall. The pod wall has two discernable layers (here called outer and inner pod wall), separated by a hard, sclerotized layer containing vascular tissue. Inside the pod lie the beans, connected to a central placenta and surrounded by mucilaginous pulp. In our current study we have isolated RNA from fruit tissues in order to study gene expression in the cocoa pod and to attempt isolating a pod wall-specific promoter.

We have compared several methods for RNA extraction from plants on various cacao tissues. Several commercial one-step extraction procedures: TriPure isolation reagent (Roche), RNeasy kit (Qiagen), SV Total RNA isolation kit (Promega, Leiden, NL), were tested. Invariably, the high carbohydrate (mucilage) concentration of the cocoa tissues resulted in gelling of the aqueous phase, so that no further purification or processing of RNA trapped in that phase was possible**.** A published method from Schultz *et al* (16) yielded good quality RNA except from bean and leaf, again because of a high carbohydrate content, but the quantity of RNA was very low (Fig 2A). Moreover, the chemical triisopropanylnaphthalene sulphonic acid, required for this protocol, became unavailable worldwide during the time of

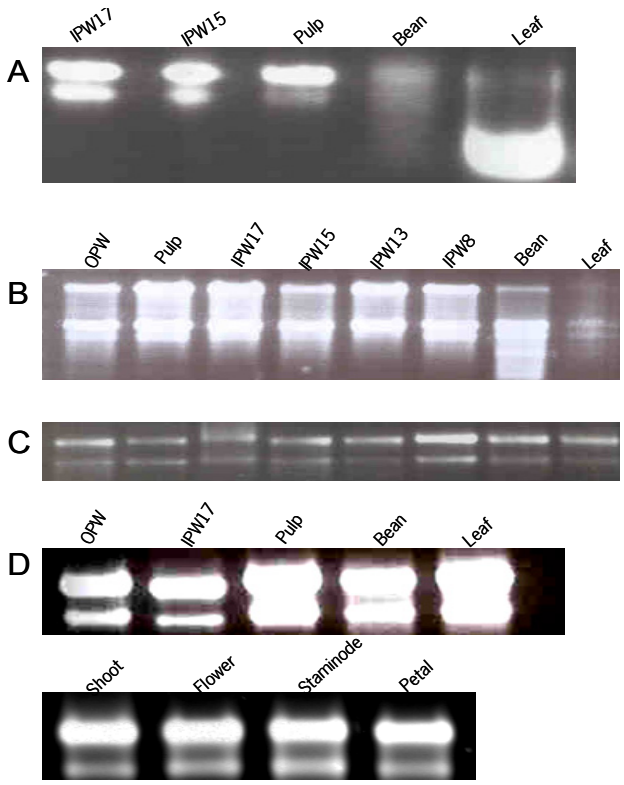


Fig. 2. Total RNA extracted from various cocoa tissues by different methods. **A.** Schultz *et al.* (16). **B**. Chang *et al.* (5). **C.** Asif *et al.* (2). **D.** Our modification of Asif *et al.* (see M&M section). OPW=outer pod wall; IPW=inner pod wall (number indicates length of the pod in cm).

quality RNA except from bean and leaf, again because of a high carbohydrate content, but the quantity of RNA was very low (Fig. 2A). Moreover, the chemical triisopropanylnaphthalene sulphonic acid, required for this protocol, became unavailable worldwide during the time of our research, and therefore this method was abandoned. Subsequently we tested and slightly modified the procedure from Chang *et al.* (5), by using PVPP instead of PVP during grinding of the tissue. The result shown was good but the yield was still low, and the RNA from bean and leaf was also contaminated with a colored substance and the RNA was not stable upon storing at -80°C for longer periods (Fig. 2B). In our hands, the method described for banana fruit from Asif *et al* (2), proved to be the most successful for RNA isolation from all cacao tissues, although the yield was still low, and especially the RNA from bean and leaf was still contaminated with carbohydrate and not very stable upon storage. This method uses

Table 2. Sequence analysis of 96 randomly selected clones from cocoa fruit pod wall and pulp cDNA libraries

			BlastN to dbEST	BlastX to Genbank results		
Contig	**No.seqs**[1]	**Tissue**	**EST id.**[2]	**E-Value** [3]	**GB identifier**	**Name**
1	6+1	PU, IPW	-	2e-69	gi|33325123|gb|AAQ08194.1|	eukaryotic translation initiation factor 5A isoform IV [Hevea brasiliensis]
2	6	IPW	-	1e-34	gi|3759184|dbj|BAA33810.1|	phi-1 [Nicotiana tabacum]
3	3+2	PU	CA796227	1e-11	gi|15149875|emb|CAC39481.2|	metallothionein-like protein [Quercus suber]
4	3	PU	-	8e-67	gi|602590|emb|CAA58230.1|	triosephosphate isomerase [Petunia x hybrida].
5	2	PU	CA794270	5e-62	gi|20426|emb|CAA79159.1|	trg-31 [Pisum sativum]
6	2	PU, IPW	-	8e-76	gi|22136876|gb|AAM91782.1|	putative sorbitol dehydrogenase
7	2	PU	-	3e-40	gi|32400804|gb|AAP80634.1|	geranylgeranylated protein ATGP1
8	1	PU	-	2e-12	gi|21593842|gb|AAM65809.1|	unknown [Arabidopsis thaliana]
9	1	PU	-	0.005	gi|18643339|gb|AAL76271.1	transcription factor EIL2 [Vigna radiata]
10	1	IPW	-	1e-50	gi|6730705|gb|AAF27100.1|	Putative phospatase 2A inhibitor [Arabidopsis thaliana]
11	2	P	-	3e-05	gi|46981317|gb|AAT07635.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
12	1	IPW	-	6e-52	gi|7327824|emb|CAB82281.1|	light-inducible protein ATLS1 [Arabidopsis thaliana]
13	2	PU, IPW	CA798562	2e-23	gi|33414046|gb|AAP03085.1|	class IV chitinase [Galega orientalis]
14	1	IPW	CA798423	8e-56	gi|4538910|emb|CAB39647.1|	S18.A ribosomal protein [Arabidopsis thaliana]
15	1	PU	CA795583	1e-33	gi|50939929|ref|XP_479492.1|	putative cytoplasmic ribosomal protein L 18 [Oryza sativa (japonica cultivar)]
16	1	IPW	-	2e-84	gi|5881239|gb|AAD55090.1|	thaumatin [Vitis riparia]
17	1	IPW	-	7e-31	gi|35187445|gb|AAQ84314.1|	fiber protein Fb25 [Gossypium barbadense]
18	1	IPW	CF974241	8e-74	gi|20147231|gb|AAM10330.1|	At1g68570/F24J5_7 [Arabidopsis thaliana].
19	1	PU	-	2e-10	gi|7211773|gb|AAF40430.1|	protein kinase MK5 [Mesembryanthemum crystallinum].
20	1	PU	-	2e-86	gi|1800281|gb|AAB68045.1|	polyubiquitin [Fragaria x ananassa]
21	1	IPW	CA797442	9e-7	gi|3043428|emb|CAA06491.1|	40S ribosomal protein S5 [Cicer arietinum]
22	1	PU	CA794992	4e-24	gi|34913680|ref|NP_918187.1|	putative DNA-damage-repair/toleration protein [Oryza sativa (japonica cultivar)]
23	1	IPW	-	4e-34	gi|50932749|ref|XP_475902.1|	putative ubiquitin-conjugating
24	1	PU	CA794487	7e-53	gi|18046|emb|CAA45866.1|	3-oxoacyl-[acyl-carrier protein] [Cuphea lanceolata]
25	1	IPW	-	1e-14	gi|6056417|gb|AAF02881.1|	alternative splicing factor SF2a [Arabidopsis thaliana]
26	2	PU, IPW	-	2e-38	gi|34146830|gb|AAQ62423.1|	At3g54420 [Arabidopsis thaliana]
27	1	IPW	-	3e-14	gi|8980432|emb|CAA65254.1|	potassium channel [Lycopersicon
28	1	PU	-	8e-63	gi|16904543|emb|CAD10740.1|	lipoxygenase [Corylus avellana]
29	1	PU	-	8e-93	gi|14715462|dbj|BAB62040.1	CjMDR1 [Coptis japonica]
30	2	PU, IPW	-	4e-35	gi|20465585|gb|AAM20275.1|	putative hydroxyproline-rich glycoprotein [Arabidopsis thaliana]
31	1	IPW	-	3e-23	gi|22135860|gb|AAM91512.1|	AP2 domain containing protein, putative [Arabidopsis thaliana]
32	1	PU	-	1e-30	gi|50937387|ref|XP_478221.1|	unknown protein [Oryza sativa (japonica cultivar-group)]
33	1	IPW	-	5e-46	gi|37534390|ref|NP_921497.1|	putative histone H2A [Oryza sativa (japonica cultivar-group)]
34	1	IPW	-	-	-	No significant hit
35	3	PU	CF974145	9e-63	gi|38325815|gb|AAR17080.1|	heat shock protein 70-3 [Nicotiana tabacum]
36	1	IPW	-	4e-80	gi|15485718|emb|CAC67492.1|	selenium binding protein [Lotus cornicuslatus var.japonicus]
37	1	PU	-	7e-63	gi|1619958|gb|AAB67996.1|	nucleoside diphosphate kinase
38	1	IPW		2e-07	gi|23198026|gb|AAN15540.1|	unknown protein [Arabidopsis thaliana]
39	1	IPW	CF972869	2e-42	gi|50906401|ref|XP_464689.1|	putative elongation factor 1-gamma [Oryza sativa (japonica cultivar-group)]
40	1	IPW	-	-	-	No significant hit
41	1	IPW	-	3e-43	gi|21842314|gb|AAM77753.1|	translation initiation factor B04 [Helianthus annuus]

Table 2. Sequence analysis of 96 randomly selected clones from cocoa fruit pod wall and pulp cDNA libraries

			BlastN to dbEST	BlastX to Genbank results		
Contig	**No.seqs**[1]	**Tissue**	**EST id.**[2]	**E-Value** [3]	**GB identifier**	**Name**
42	1	IPW	-	1e-08	gi|1208497|dbj|BAA07323.1|	ethylene-responsive element binding protein [Nicotiana tabacum]
43	1	IPW	CA798468	2e-27	gi|48209911|gb|AAT40505.1|	putative elongation factor [Solanum demissum]
44	3+7	PU, IPW	CP974081	1e-11	gi|14582465|gb|AAK69513.1	putative transcription factor [Vitis vinifera]
45	1+1	PU	CA794461	4e-75	gi|34304717|gb|AAQ63462.1|	calmodulin 8 [Daucus carota]
46	1	IPW	CF973824	6e-65	gi|1362041|pir||S56673	ribosomal protein S23.e, cytosolic (clone RJ3)-garden strawberry
47	1	IPW	-	5e-25	gi|4097553|gb|AAD09510.1|	ATFP6 [Arabidopsis thaliana]
48	1	IPW	-	4e-16	gi|26450598|dbj|BAC42411.1|	unknown protein [Arabidopsis thaliana]
49	1	PU	-	3e-15	gi|7329671|emb|CAB82665.1| .	putative protein [Arabidopsis thaliana]
50	1	IPW	-	-	-	No significant hit
51	1	PU	CF974266	8e-04	gi|20147259|gb|AAM10343.1|	At1g32920/F9L11_25 [Arabidopsis thaliana]
52	1	PU	CA794998	7e-38	gi|1332579|emb|CAA66667.1|	polyubiquitin [Pinus sylvestris]
53	1	IPW	CA795308	7e-30	gi|28193048|emb|CAD58628.1|	SUI1 protein [Coffea arabica]
54	1	PU	-	-	-	No significant hit
55	1	PU	-	1e-11	gi|2346974|dbj|BAA21921.1|	ZPT2-12 [Petunia x hybrida]
56	1	IPW	-	2e-39	gi|14423556|gb|AAK62460.1|	putative protein [Arabidopsis thaliana]
57	1	PU	-	9e-66	gi|20530131|dbj|BAB92019.1|	mitochondrial aldehyde dehydrogenase [Sorghum bicolor]
58	1	IPW	-	-	-	No significant hit
59	1	IPW	-	6e-35	gi|47027024|gb|AAT08727.1	| 60S ribosomal protein L11 [Hyacinthus orientalis]
60	1	IPW	-	6e-51	gi|37535224|ref|NP_921914.1|	cap-binding protein p28 [Oryza sativa (japonica cultivar-group)]
61	1	PU	-	8e-09	gi|21554633|gb|AAM63640.1|	unknown [Arabidopsis thaliana]
62	1	PU	-	7e-43	gi|15487258|emb|CAC69074.1|	STP13 protein [Arabidopsis thaliana].
63	1	PU	-	2e-48	gi|20148499|gb|AAM10140.1|	putative protein [Arabidopsis thaliana]
64	1	IPW	CA797373	2e-18	gi|5381255|dbj|BAA82307.1|	peroxidase [Nicotiana tabacum]
65	1	IPW	-	4e-26	gi|47026993|gb|AAT08714.1|	ribosomal protein L18A [Hyacinthus orientalis]
66	1	PU	CF974317	2e-14	gi|7488882|pir||T10900	late-embryogenesis protein homolog - mung bean
67	1	IPW	-	2e-48	gi|13194766|gb|AAK15545.1|	putative plasma membrane intrins
68	1	PU	-	1e-20	gi|9187622|emb|CAB97004.1|	WRKY DNA binding protein [Solanum tuberosum]
69	1	IPW	-	5e-13	gi|13171103|gb|AAK13589.1|	rRNA intron-encoded homing endonuclease [Oryza sativa]
P27	1		-	2e-91	gi|1332579|emb|CAA66667.1|	polyubiquitin [Pinus sylvestris]
PW6	1		-	1e-21	gi|21537118|gb|AAM61459.1|	unknown [Arabidopsis thaliana]
P3	1		-	4e-41	gi|30984524|gb|AAP42725.1|	At4g24570 [Arabidopsis thaliana]
PW17	1		-	5e-43	gi|3759184|dbj|BAA33810.1|	phi-1 [Nicotiana tabacum]

[1]Number of sequences used to assemble the contig. Additional sequences (15 in total) obtained from reverse-Northern-positive clones are added (11 in total) to the homologous contigs. The remaining 4 unique sequences are added to the bottom of the table.
[2]Only matches with >98 nucleotide identity are listed
[3]Hits with an expectation value E<e-2 were considered not significant

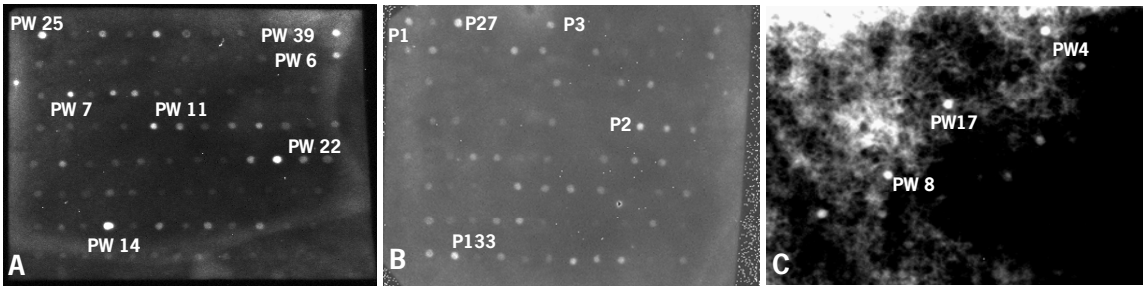


Fig. 3. Reverse Northern dot blots or macroarrays of 96 cDNA's of the inner pod wall (**A, C**) and of pulp (**B**), hybridized with labeled total cDNA probes from pulp (**A**, **B**) or the inner pod wall (**C**). Spots of cDNA's used in later experiments are labeled.

chloroform/isoamyl extraction prior to RNA precipitation with LiCl. In case of cocoa extracts addition of LiCl caused immediate formation of a cloudy precipitate of carbohydrates and possibly other contaminants. Addition of an extra series of phenol, phenol: chloroform : isoamyl alcohol (25:24:1) and chloroform : isoamyl alcohol (24:1) extraction, prior to LiCl precipitation (see Material and Methods section), resulted in a big improvement in quality, quantity and stability of RNA (Fig. 2C). The modified procedure could subsequently be used reliably to isolate high-quality RNA from all stages of fruit development, as well as from flower organs and leaves.

cDNA libraries from the cocoa pod wall and pulp

In order to identify pod specific gene expression in the cocoa fruit two cDNA libraries, from the pod wall and from the pulp surrounding the seeds, respectively, were made. To determine the quality of the library and the redundancy of sequences, 48 clones each of the pod wall and of the pulp libraries were sequenced and assembled into contigs where possible, and analyzed for homology with sequences in GenBank, as shown in Table 2. The 87 good quality sequences that resulted could be reduced to 69 different sequences by contig assembly: 13 contigs of 2 - 7 clones each and 56 unique sequences.

In two previous studies, 4455 expressed sequence tags (ESTs) from untreated leaves and beans (11) and 1256 ESTs from leaves treated with various elicitors (17) were reported. All contig sequences identified in this study were compared to cocoa ESTs available in the dbEST database by the BlastN algorithm. Of the 69 contig sequences, 20 were identical to a previously identified EST (Table 2). Thus, the cDNA libraries produced had relatively low redundancy and contained a considerable number of not previously characterized ESTs.

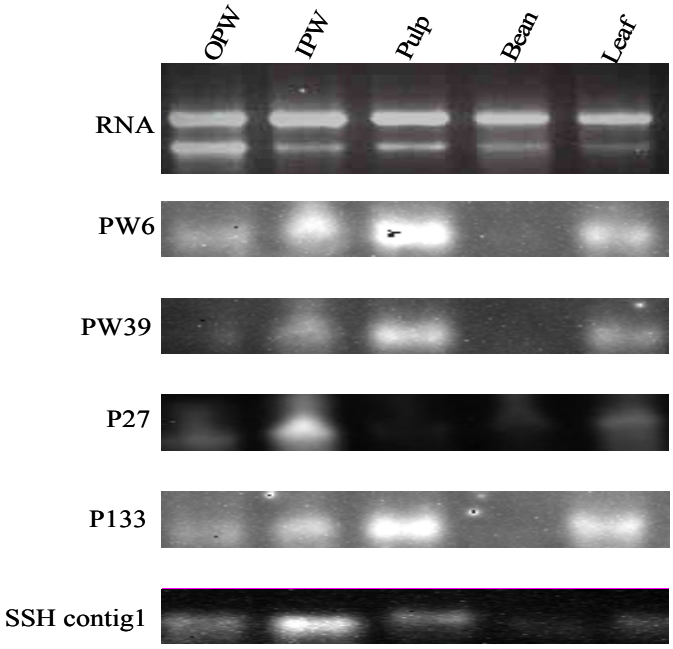


Fig. 4. Total RNA (top panel) and Northern blots of various cocoa pod tissues as well as of bean and leaf hybridized with 6 different probes, identified by reverse Northern blots (see Fig. 3; panels 2-5) or by SSH (panel 6).

To determine which of the isolated cDNA clones represented highly expressed genes in either pod wall, pulp or both, we performed a reverse Northern dot blot analysis. Blots spotted with PCR-amplified inserts of 96 random cDNA clones each of the pod wall or pulp cDNA library, respectively, were prepared and hybridized using DIG-labeled total cDNA probes prepared from the pod wall or the pulp mRNA. The results of 3 reverse northern blots are shown in Fig. 3. Labeled pulp cDNA probes on pod wall cDNA (Fig. 3A) or pulp cDNA spots (Fig. 3B), as well as pod wall cDNA probe on pod wall cDNA spots (Fig. 3C) clearly distinguished between cDNA clones of highly expressed genes and genes expressed at lower or non-detectable levels. From the fact that a pulp cDNA probe detects a large number of random pod wall cDNA spots (Fig. 3A), we concluded that there is considerable overlap in gene expression between the two tissues. cDNA's of 15 spots hybridizing with a high intensity on one of the three blots were sequenced. The obtained sequences were compared to the already obtained 69 contigs of random sequences as listed in Table 2. Eleven out of the 15 sequences were found to align with, and thus belong to, 4 of the already identified 69 contigs, while the remaining 4 represented new and unique sequences. Curiously, 7 out of 15 cDNA's were overlapping with one contig (44, see Table 2), showing homology to a putative

Table 3. Sequence analysis of randomly selected clones from the Suppression Subtractive Hybridization library obtained from the cocoa fruit pod wall cDNA fragments subtracted from leaf and bean cDNA fragments

		BlastN to dbEST	**BlastX to Genbank results**		
Contig[1]	No.seqs[2]	EST id.[3]	E-Value [4]	GB identifier	Name
1	14	CA794695	1e-05	gi|7484729|pir||T09877	late embryogenesis-abundant protein Lea5-D - upland cotton
2	10	-	1e-09	gi|21585695|gb|AAM55492.1|	allergenic-related protein Pt2L4 [Manihot esculenta]
3 (54)	2	-	0.002	gi|24021287|gb|AAN40994.1|	coat protein [Cucumber mosaic virus]
4	2	-	6e-23	gi|7262695|gb|AAF43953.1|	Strong similarity to an unknown protein from Arabidopsis thaliana
5	1	-	3e-26	gi|37901055|gb|AAP46154.1|	putative C3HC4-type RING zinc finger protein [Hevea brasiliensis]
6	1	-	-	-	No significant hit
7	3	-	4e-16	gi|37901055|gb|AAP46154.1|	putative C3HC4-type RING zinc finger protein [Hevea brasiliensis]
8	1		1e-23	gi|7677046|gb|AAF67003.1|	putative Hs1pro-1 homolog [Pisum sativum]
9	1	-	2e-21	gi|51535190|dbj|BAD38163.1|	NHL repeat-containing protein-like [Oryza sativa (japonica cultivar-group)]
10	1	-	0.004	gi|50978567|dbj|BAD34997.1|	hypothetical protein [uncultured bacterium]
11	2	-	8e-09	gi|21554872|gb|AAM63715.1|	unknown [Arabidopsis thaliana]
12 (16)	3	-	4e-08	gi|5881239|gb|AAD55090.1|	thaumatin [Vitis riparia]
13	1	-	9e-27	gi|23197792|gb|AAN15423.1|	putative protein [Arabidopsis thaliana]
14	1	-	-	-	No significant hit
15	1	-	1e-48	gi|21387087|gb|AAM47947.1|	proteasome subunit [Arabidopsis thaliana]
16	1	-	2e-50	gi|6522625|emb|CAB62037.1|	ubiquitin conjugating enzyme E2 (UBC13) [Arabidopsis thaliana]
17	7	CF973244	8e-47	gi|57283538|emb|CAG27609.1|	monosaccharide transporter [Populus tremula x Populus tremuloides]
18	1	-	7e-10	gi|21593806|gb|AAM65773.1|	putative RING zinc finger protein [Arabidopsis thaliana]
19	1	-	5e-26	gi|53793514|dbj|BAD54675.1|	putative iron inhibited ABC transporter 2 [Oryza sativa (japonica cultivar)]
20	1	-	8e-13	gi|38347592|emb|CAE04883.2|	OSJNBa0042I15.5 [Oryza sativa (japonica cultivar-group)]
21	1	CA796626	5e-56	gi|21234|emb|CAA40356.1|	acetohydroxy acid reductoisomerase [Spinacia oleracea]
22 (30)	3	-	-	-	No significant hit
23	1	-	-	-	No significant hit
24	4	-	2e-65	gi|2981131|gb|AAC06237.1|	AGAMOUS homolog [Populus balsamifera subsp. trichocarpa]
25	1	CF972672	9e-42	gi|50944095|ref|XP_481575.1|	hypothetical protein [Oryza sativa (japonica cultivar-group)]
26 (60)	1	-	5e-61	gi|22136272|gb|AAM91214.1|	putative WD-40 repeat protein [Arabidopsis thaliana]
27	1	-	5e-92	gi|6166483|gb|AAF04851.1|	putative alcohol dehydrogenase [Hibiscus syriacus]
28	1	-	-	-	No significant hit

[1]In parantheses the contig of the random libraries (Table 2) that overlap with this contig
[2]Number of sequences used to assemble the contig.
[3]Only matches with >98% nucleotide identity are listed
[4]Hits with an expectation value E<e-2 were considered not significant

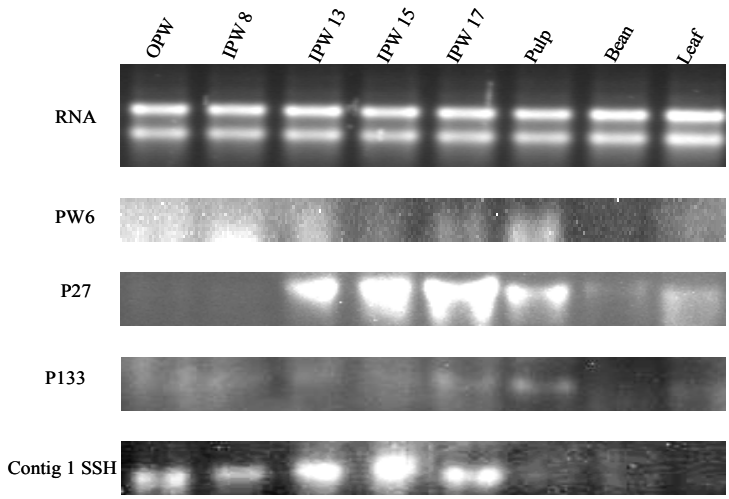


Fig. 5. Northern blots of various cocoa pod tissues, extended with samples of the inner pod wall from different developmental stages (indicated as pod length in cm) as well as of bean and leaf hybridized with 3 selected probes

transcription factor from grapevine.

To further investigate the expression level and tissue specificity of the genes encoding the cDNA's identified by reverse Northern blotting, Northern blots were made of the cocoa outer pod wall, inner pod wall, pulp, bean and leaf mRNA and DIG-labeled cDNA inserts from Table 2 used as probes. Eight probes were produced from pod wall cDNA's and three from pulp cDNA's. The results indicated that not all of the cDNA clones with a strong signal on the reverse Northern blot gave specific expression on Northern blots (Fig. 4). Only 4 (PW6, PW39, P27 and P133) of the 8 probes derived from the pod wall cDNA blots were found to have significant expression in the outer pod wall, inner pod wall, pulp, and (to a lesser extent) leaf but not in the bean. Probes PW39 and P133 both aligned with contig 44 of Table 2, and thus not surprisingly revealed the same expression pattern, although they originated from different libraries. In all cases, except with probe P27, expression in the pulp was higher than in the inner pod wall, where in turn expression was higher then in the outer pod wall. Probes of clones that initially appeared to be more specific for pod wall expression (those that hybridized in Fig. 3C, but not in Fig. 3A) gave no signal on the Northern blot. In contrast, probe P27 appeared to be hybridizing relatively specifically with the inner pod wall RNA and to a lower extent with the pulp RNA.

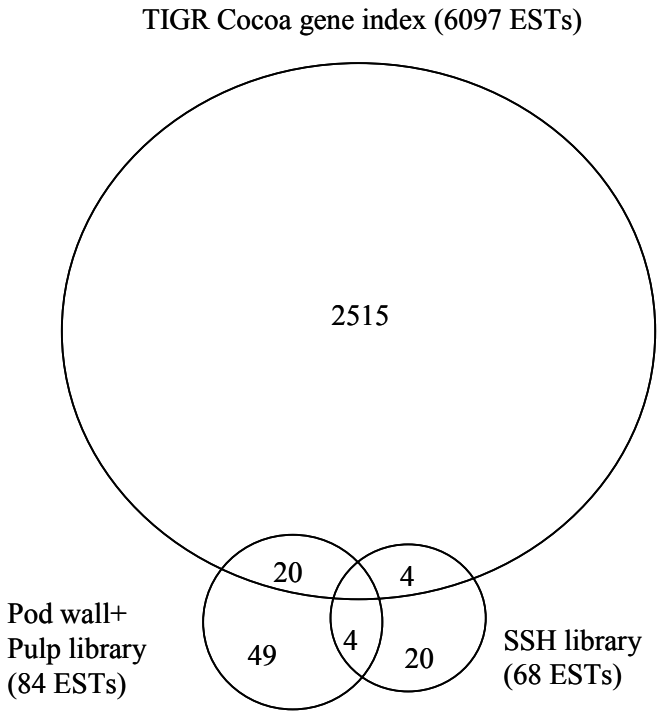


Fig. 6. Analysis of sequence content and overlap in a Venn diagram for the TIGR cocoa gene index, the pod wall + pulp cDNA library, and the SSH pod wall library. Numbers in the circles are unique sequences. Total numbers of EST sequences from which the unique sequences are derived are listed with the database/library name.

Isolation of pod wall-specific cDNA fragments by Suppression Subtractive Hybridization (SSH)

A second strategy that we used to identify pod wall specific cDNA's was by producing a Suppression Subtractive Hybridization (SSH) library resulting from a subtraction of pod wall cDNA as tester and mixed leaf + bean cDNA as driver. Subtracted cDNA was ligated into pGEM-T Easy and subjected to colony PCR, revealing inserts corresponding to cDNA fragments of 250 to 1000 bp. In total 96 random cDNA inserts were picked up, sequenced and analyzed. Sequence analysis using the SeqMan software package on 68 readable sequences showed that these amounted to 28 different sequences. These sequences consisted of 10 contigs of 2 - 14 sequences each and 18 single sequences (Table 3). Out of the 28 sequences from the SSH library, only 4 were overlapping with a

contig sequence (indicated in parentheses in the first column of Table 3) from the random libraries described in the previous section, suggesting that SSH samples mainly contained different cDNA's as compared to random cDNA cloning. Only 4 out of the 28 sequences were identical to previously identified cocoa ESTs deposited in GenBank. Whereas in the literature SSH is claimed to not only select cDNA's from one of the two compared pools, but also to normalize, i.e. equally represent highly and lowly expressed specific genes (6). In our study as in previous studies (results not shown) we found that there is an imbalance in the number of contig members when comparing all 28 contigs. This suggests that, as in earlier studies, the over-represented sequences (contigs 1, 2 and to a lesser extent contig 17) correspond to tissue-specific and highly expressed genes. For this reason we decided to further study the expression and regulation of the gene represented by SSH-contig 1 (from hereon called the cocoa *Lea5D*-homologue). Hybridization of a Northern blot with a probe made from the SSH contig 1 fragment (Fig. 4, bottom panel), showed that the gene represented by this contig is highly expressed in the inner pod wall, and to a lesser extent in the outer pod wall and pulp and further less in the leaf and virtually undetectable in beans. Next we extended the Northern blots of Fig. 4 with a series of inner pod wall RNA extracted from pods in various stages of development (8 -17 cm) and compared hybridization of SSH-contig 1 with that of the probes PW6 and P27 from the random cDNA libraries. As can be seen in Fig. 5, these three probes revealed overlapping but distinct expression patterns. PW 6 was expressed in all stages of development of the inner pod wall, as well as in the outer pod wall and in pulp. P27 was increasingly expressed in the inner pod wall as the pod develops, but not expressed in the outer pod wall or in the young stages of the inner pod wall, while expression in the pulp was lower, but discernable. Finally, cocoa *Lea5D* expression was more specific for the pod wall, both outer as well as inner, where it increased during development but in contrast to P27 it was also detectable in the youngest stage. Expression in the pulp was low. All three genes had little or no expression in the bean, and for PW6 and P27, some expression in the leaf.

Discussion

In this manuscript we describe the first attempt at characterizing gene expression in cocoa specifically in the tissues of the pod, the fruit of the cocoa tree. The ultimate goal of our investigation was the isolation of a cocoa gene promoter specifically driving high

expression in the pod wall and/or pulp, in order to use it to achieve organ-specific and high level. expression of transgenes. For example, expression in the pod wall or pulp of a gene encoding a cocoa pod borer (CPB)-active Bt toxin might protect transgenic cocoa pods against this pest

We used and compared two approaches to identify cDNA's that were highly expressed in the pod wall or pulp and preferably not in the beans and leaves. Both the random cDNA library as well as the SSH pod wall library contained considerable diversity in sequence and little overlap with each other or with the public dbEST and TIGR cocoa gene index databases (Fig. 6). This may not be surprising since the TIGR cocoa gene index still has a limited size (a total of 6097 deposited sequences, constituting 2539 unique sequences; compared to the tomato gene index: 164,108 and 31,838 respectively). However the ESTs in the database are almost all derived from two libraries: a mixed library of leaf and bean cDNA's (11) and one from elicitor-treated leaves (17). Thus, our results indicate that the libraries in our studies may reveal many new, more pod-specific cDNA sequences.

Reverse Northern blotting (alternatively called a macroarray) revealed a number of cDNA's from either of the two random libraries that seemed to be highly expressed at least in the inner pod wall. Most of these were also and often higher expressed in the pulp, suggesting that there is considerable overlap in gene expression between inner pod wall and pulp. Most of these also showed some expression in the leaves, indicating a low level of tissue-specificity. Exception was clone P27, which although originating from the pulp library, showed considerably higher expression in the inner pod wall and no expression in bean or leaf. However, more in-depth analysis of expression in various developmental stages of the pod showed that this gene is indeed highly expressed in the ripe inner pod wall, but much less in the earlier stages of development.

Suppression subtractive hybridization is for many systems an attractive method for selecting cDNA fragments, which are specific for one pool as opposed to another pool of cDNA's. We used this technique in order to identify genes specifically expressed in the pod wall as opposed to expression in leaf or bean. The most frequent cDNA fragment in the set of sequenced fragments from this library was a homologue of the cotton Lea5D gene (8). Northern blotting revealed highest expression in the inner pod wall and very low expression in beans and leaf. Moreover, expression of this gene was detectable in the outer pod wall as well as, in contrast to clone P27, in early developmental stages of the inner pod wall. This expression pattern would be better suited for a Bt toxin encoding

gene aimed at protection against CPB, as CPB moths are known to lay their eggs preferably on pods 3-4 weeks before ripening, but also earlier (3), and thus early expression of the toxin would be necessary for optimal protection against the boring larvae. The choice for isolation of the promoter of the cocoa *Lea5* homologous gene was further supported by earlier studies of the expression pattern of a tomato homologue of this gene. TblastN searching of the TIGR tomato gene index identified a homologous tentative contig (*TC153879*) consisting of 156 ESTs, of which 86 were isolated from tomato fruit in different developmental stages, with the only other significant occurrence in root cDNA libraries. The GenBank BlastX search also revealed homology to a potato (*Solanum demissum*) *Lea5* homologue (CAA66948) expressed during tuber formation (10). The tomato homologue of this protein was highly expressed in developing and mature green tomato fruit pericarp and locular tissue (10, 13). Based on these data we decided for the subsequent work to clone and characterize the promoter of the cocoa *Lea5D* gene and compare its regulation with that of the tomato *Lea5D* homologue in the tomato fruit.

Literature cited

1. **Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25:**3389-3402.
2. **Asif, M. H., P. Dhawan, and P. Nath.** 2000. A simple procedure for the isolation of high quality RNA from banana fruit tissue. Plant Mol. Biol. Rep. **18:**1 -7.
3. **Azhar, I., and G. E. Long.** 1996. Effect of cocoa and pod age on egg distribution and egg parasitism of the cocoa pod borer in Malaysia. Entomol. Exp. Appl. **81:**81-89.
4. **Bowers, J. H., B. A. Bailey, P. K. Hebbar, S. Sanogo, and R. D. Lumsden.** 2001. The impact of plant diseases on world chocolate production. Plant Health Progress Online. doi:10.1094/PHP-2001-0709-01-RV.
5. **Chang, S., Puryear, J., and Cairney, J.** 1993. A simple and efficient methods for isolating RNA from pine trees. Plant Mol. Biol. Rep. **16:**9-18.
6. **Diatchenko, L., Y. F. C. Lau, A. P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E. D. Sverdlov, and P. D. Siebert.** 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA **93:**6025-6030.
7. **Entwistle, P.** 1987. Insects and cocoa, p. 366-443. *In In* G. A. R. Wood and R.A. Lass (ed.), Cacao, vol. 4. Longman Scientific and Technical, Harlow.
8. **Galau, G. A., H. Y. Wang, and D. W. Hughes.** 1993. Cotton *Lea5* and *Lea14* encode atypical late embryogenesis-abundant proteins. Plant Physiol. **101:**695-6.

9. **Gregory, P. H.** 1974. Prospect. In 'Phytophthora Disease of Cocoa' (Ed. PH Gregory): 301-308. Longman, London.

10. **Jackson, S., J. Gascon, E. Carrera, E. Monte, and S. Prat.** 1997. Cloning and expression analysis of a gene that shows developmental regulation upon tuberization in potato. Plant Mol. Biol. **33:**169-74.

11. **Jones, P. G., D. Allaway, D. M. Gilmour, C. Harris, D. Rankin, E. R. Retzel, and C. A. Jones.** 2002. Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta **216:**255-264.

12. **Lass, R. A.** 1987. Diseases, p. 265-365. *In* G. A. R. Wood and R.A. Lass (ed.), Cacao, vol. 4. Longman Scientific and Technical, Harlow.

13. **Morgens, P., J. Pyle, and A. Callahan.** 1987. Searching for molecular mechanisms involved in fruit ripening., p. 157-166. *In* J. E. Fox and M. Jacobs (ed.), Molecular biology of plant growth control. Alan R. Liss, New York.

14. **Purdy, L. H., and R. A. Schmidt.** 1996. Status of cacao withches' broom: biology, epidemiology, and management. Annu. Rev. Phytopathol. **34:**573-594.

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

16. **Schultz, D. J., R. Craig, D. L. Cox-Foster, R.O. Mumma, and J. L. Medford.** 1994. RNA isolation from recalcitrant plant tissue. Plant Mol. Biol. Rep. **12:**310-316.

17. **Verica, J. A., S. N. Maximova, M. D. Strem, J. E. Carlson, B. A. Bailey, and M. J. Guiltinan.** 2004. Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. Plant Cell Rep. **23:**404-413.

18. **Wardojo, S.** 1992. Major pests and diseases of cocoa in Indonesia., p. 63-67. *In* P. J. Keane and C. A. J. Putter (ed.), Cocoa pest and disease management in Southeast Asia and Australasia FAO Plant Production and Protection Paper No. 112. FAO, Rome.

Chapter 5

Characterization of a cocoa pod wall-specific gene, *TcLea5*, a comparison with its tomato homologue and characterization of its promoter in heterologous systems

Tetty Chaidamsari[1], Mieke Weemen[1], Djoko Santoso[2] and Ruud A. de Maagd[1]

[1]Business Unit Bioscience, Plant Research International B.V, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[2]Biotechnology Research Institute for Estate Crops, Jl Taman Kencana 1, Bogor 16151, Indonesia

Summary

In order to isolate and characterize a pod wall-specific promoter from cocoa (*Theobroma cacao* L.), we have cloned a 3314 bp genomic fragment encompassing a gene homologous to a cotton *Lea5* gene. This gene, which is almost exclusively expressed in the pod wall of cocoa, was named *TcLea5*. To identify the promoter of this gene a series of constructs, comprising different lengths of the 5' upstream sequence fused to the *uidA* (GUS) gene, was tested for transient expression in tomato fruits as well as in transgenic *Arabidopsis thaliana* plants. A minimal promoter fragment of 451 bp (relative to the start codon) that drives expression in *A. thaliana* was identified. However, unexpectedly, expression in *A. thaliana* also occurred in leaves. LEA5 proteins form an atypical family in the larger collection of LEA families, proteins that are usually expressed in late developmental stage embryos as well as often in desiccating vegetative tissues. In contrast, *TcLea5*, and its homologue in tomato are almost exclusively expressed in developing fruit pericarp (as well as locular tissue in tomato), where they belong to the most abundant transcripts although their function is unknown.

Introduction

Genetic modification of crop plants is a promising way of offering protection against pests and diseases, particularly if natural resistance is not available or difficult to apply. We have previously shown that some *Bacillus thuringiensis* (Bt)-derived toxins have activity against the cocoa pod borer (CPB) (14). Our long-term goal is to develop transgenic cocoa trees expressing one or more active Bt toxins in the pod wall for protection against CPB.

In order to successfully develop transgenic approaches for protection of the cocoa pod against diseases and pests, a number of problems will have to be resolved. In order to express a transgene or its indirect products in the pod to a sufficient level, a plant promoter with activity in the pod wall and/or pulp is required. Moreover, to decrease the risk of possible consumer health or environmental effects (or more indirectly, to quench public anxiety) of expressing transgenes in tissues other then targeted by the pest (particularly the bean), such a promoter should preferably be highly specific for the target tissue.

In Chapter 4 of this thesis, we described the expression pattern of an apparently pod wall-specific cDNA, a fragment from an SSH-library encoding a peptide with homology to Lea5 proteins. The gene was primarily expressed in the inner pod wall and somewhat in the outer pod wall and not in pulp, beans, and leaves, thus meeting the requirements for the promoter driving transgene expression, as described above.

In this study, we describe how the corresponding genomic region for the cocoa Lea5 homologue was cloned and the putative promoter characterized. Due to the long generation time it is difficult to study promoter activity in transgenic cocoa fruits. Therefore we studied the promoter's activity in heterologous systems and investigated in parallel the expression of a tomato homologue in tomato fruit, in order to attempt to link processes in the cocoa pod wall development with those of the development of a model species' fruit. This should improve and make easier future studies of pod wall gene expression in cocoa.

Material and Methods

Plant material

The plant material was obtained from the cacao tree clones RCC72 and Sca6, which were grown in fields of the Indonesian Coffee and Cacao Research Institute, Jember Indonesia. Cocoa fruits in different stages of development were shipped to Wageningen and upon arrival were sliced, immediately frozen in liquid nitrogen, and stored at –80°C until further use.

Cocoa leaves were harvested from plants grown in the laboratory in Wageningen. Tomato leaves and fruits were obtained from the greenhouse at PRI. *Arabidopsis thaliana* ecotype WS plants for transformation were grown in the growth chamber at Plant Research International (16 h light, 21°C).

RNA extraction

The method used for RNA extraction from cocoa tissues is a modification of a method described earlier for banana tissues (2), and is described in detail in chapter 4 of this thesis.

The hot phenol method was used for RNA extraction from tomato fruits and leaves. 4 gram of tissue was ground by hand using a pestle and mortar in the presence of liquid nitrogen. The powder was transferred to a 50 ml tube, resuspended in 12 ml preheated (80^{o}C) phenol and 12 ml extraction buffer (100 mM NaAc pH 4.8, 100 mM LiCl, 10 mM EDTA, and 1% SDS). Next, the homogenate was mixed by vortexing for 30 sec. Subsequently 12 ml chloroform: isoamyl alcohol was added, mixed by vortexing for 30 sec, and then centrifuged at 10,000 g for 15 min at 4°C. The upper aqueous phase containing nucleic acids and polysaccharides was carefully transferred to a new tube, and 4 M LiCl was added to a final concentration of 2 M. The RNA was allowed to precipitate 1 h at -80°C and then recovered by centrifugation at 17,000 g at 4°C for 25 min. The pellet was washed with an equal volume of 70% cold ethanol, the ethanol was allowed to evaporate at room temperature for 15 min, and the purified RNA pellet was then resuspended in an appropriate volume of nuclease-free water. The purity and concentration of the purified RNA was determined spectrophotometrically at 230, 260, and 280 nm.

The RNA isolation from *Arabidopsis thaliana* tissues was carried out using the TriPure Isolation Reagent kit (Roche) according to the kit‘s instructions.

Dig-labelling of RNA and Northern blot analysis

The templates for in vitro transcription to produce RNA probes were made by PCR amplification of the complete inserts from 250 ng of the appropriate cDNA clones using the reverse and forward M13 primers (Table 1). A probe for the tomato Lea5 homologue was made from the amplified insert (700 bp) of plasmid TUS15A17 (obtained from the Boyce Thompson Institute, Center for Gene Expression Profiling, Cornell, NY). Labeling by in vitro transcription, Northern blotting, hybridization and detection were described in Chapter 4 of this thesis.

Table 1. Oligonucleotide primers used in this study

Name	Nucleotide sequence (5' to 3')
M13 forward	GTAAAACGACGACGGCCAGT
Adaptor genome walking	GTAATACGACTCACTATAGGGC
Nested adaptor genome walking	ACTATAGGGCACGCGTGGT
GSP1 contig1	CCATGGGAAAGCTAGTAGGGGTTACAATAC
GSP2 contig1	GTGGTGGTCATTAACATTAGAAATGGCC
promBspHIrev	GAGCGAGTCATGATTTCTATTTGC
promA for	GGAATTCTGGATATCCATTTAAACATAAAC
promB for	GGAATTCCTTAATCCCTGCC
promC for	GGAATTCAAGATAAGATCCAGCTAAG
promD for	GGAATTCCCCAAGGCTAAGACTG
linkerEcoRINcoI	ATTCCTCATGAC
linkerNcoIEcoRI	GTACCAGTACTC
promBspHIrev	GAGCGAGTCATGATTTCTATTTGC

Cocoa leaf DNA extraction

For extraction of DNA from cocoa leaf cell nuclei, an earlier described method (5, 11) was used with small modifications. One gram young pale green leaf was ground into a coarse powder in liquid nitrogen by hand using a pestle and mortar. Once the liquid nitrogen had boiled off, 10 ml extraction buffer (0.1 M sodium citrate, 0.35 M glucose, 5 mM EDTA, 0.5 % (v/v) ß-mercaptoethanol, 2% (w/v) polyvinylpyrrolidone 10.000, 1% (w/v) BSA, pH 6,0) was added and the mixture homogenized in a warring blender at maximum speed for two min. The homogenate was filtered through four layers of cheesecloth or nylon mesh, transferred to clean 30 ml tubes and then centrifuged at 2,000 g for 10 min at 4°C. The supernatant was discarded and the cell pellet was resuspended in lysis buffer (100 mM TrisHCl pH 8.0, 25 mM EDTA, 1.5 M NaCl, 2.5% CTAB, 0.2% (v/v) β-mercaptoethanol, 1% (w/v) polyvinypirrolidone and was mixed by inversion for 5 min, until a gelatinous homogenate was formed. The mixture was incubated for 1 hour at 60°C while mixing by inversion several times. An equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed by inversion for 10 min and the mixture was centrifuged at 12,000 g for 10 min at 4°C. The upper clear aqueous layer was transferred to a new tube and 0.5 volume 5 M NaCl was added. This mixture was mixed by inversion, and 0.6 volume of isopropanol was added, and then incubated at 4°C for 1 hour, followed by centrifugation at 12,000 g for 10 min. The pellet was

washed twice with cold 80% (v/v) ethanol, and dried under vacuum for 10 min. The pellet was resuspended in 200 µl of milliQ grade water.

Genome Walking

Amplification of cocoa genomic DNA fragments was performed with the Universal Genome Walker kit (CLONTECH, Palo Alto, USA) according to the manufacturer's instructions. DNA was digested for 16 h with four blunt-end cutting restriction enzymes: *Dra* I, *Eco* RV, *Pvu* II, and *Stu* I. After digestion the DNA was purified by phenol/chloroform extraction and ethanol precipitation. Genome walker adaptors were ligated to each batch of genomic fragments, which were then used for PCR amplification of specific fragments. For each genomic fragment batch, two rounds of PCR were performed using the Advantage 2 Polymerase mix (CLONTECH) in a PE Biosystems DNA thermal Cycler 480. The primary PCR was performed using an adaptor-specific primer and a gene-specific primer. This was followed by a second PCR with a nested adaptor-specific primer and a nested gene-specific primer (Table 1). The parameters used during primary PCR reactions were as follows: 7 cycles: 94°C, 25s; 72°C, 3 min; 32 cycles: 94°C, 25s; 67°C, 3 min. An additional 7 min at 67°C for the last step was added. For the nested PCR the same program as for primary PCR was used, with the exception that the number of cycles was different, i.e. 5 cycles and 24 cycles respectively. The PCR fragments were analyzed on 1.5% agarose/EtBr gels, and the band was extracted from the gel, and subsequently cloned into pGEM-T Easy (Promega, Madison USA). The ligated DNA was finally transformed into *E. coli* XL-1 Blue competent cells.

Sequence analysis

DNA sequencing reactions were performed on PCR-amplified fragments using Bigdye reaction mix (Applied Biosystems) and analyzed in a 3100 Genetic Analyzer (ABI Prisma). The parameters used during PCR reactions were as follows: 96°C, 3 min; 32 cycles: 96°C, 30s; 50°C, 15s; 60°C, 4 min. Sequences were analyzed and assembled into contigs using the DNAstar SeqMan program (Lasergene Biocomputing Software for Windows). Protein homology searches in the NCBI Genbank protein database were done using the BlastX algorithm (1), as were BlastX searches in the Solanaceae Genome Network tomato Unigene database (http://www.sgn.cornell.edu/cgi-bin/tools/blast/simple.pl).

Promoter-*gus* gene fusion constructs

To test genome walking products for promoter activity, 4 different constructs were made based on a genomic *Pvu*II fragment bordering TcLea5. All the putative promoter fragments were obtained by PCR using *Pfu* polymerase with the primers, promA, promB, promC and promD (Table 1; all containing a 5' tail with an *EcoR*I-recognition site) for amplification of 1739, 813, 451, and 289 bp genomic fragments respectively (Fig. 1). In all cases primer prom*Bsp*HI (Table 1; containing a *Bsp*HI-recognition site) was the other primer. The PCR-fragments were digested with *EcoR*I and *Bsp*HI and ligated into the *EcoR*I and *Nco*I sites, replacing the CaMV 35S promoter, of the vector pCAMBIA 1303-kan (Chapter 6), preceding the modified *uidA* (GUS) gene encoding β-glucuronidase, yielding plasmids pMH60, pMH61, pMH62, and pMH63, respectively. The CaMV 35S-promoter in pCAMBIA1303kan, (Chapter 6 in this thesis) was used as a positive control. For a negative control, the ends of *EcoR*I/*Nco*I-digested pCAMBIA1303-kan were joined by ligation with an *EcoR*I/*Nco*I-linker, giving plasmid pMH64. The constructs were introduced into *Agrobacterium tumefaciens* strain AGL0 by electroporation and transformants were selected on kanamycin (50 µg/ml).

Qualitative promoter activity testing in tomato fruit

To determine promoter activity by transient gene expression in tomato fruit, we used the method of injection of an *Agrobacterium* suspension into the tomato fruit (15). All the strains with plasmids containing the putative promoter fragments, or the controls, were grown at 28°C in 10 ml YEB medium (5 g/L beef extract, 1 g/L yeast extract, 5 g/L sucrose and 2 mM $MgSO_4$). When the culture reached an OD_{600} of approximately 0.8 it was centrifuged and the bacterial pellet was resuspended in 1 ml Murashige and Skoog (MS) medium. The *Agrobacterium* suspension (500 µl) was injected in one place in the fruit pericarp using a sterile 1 ml syringe with a 27Gx3/4" needle. The outside of the injected fruits was dried and the fruits were placed for 2 d at 22°C with a 16h photoperiod. After this incubation period, fruits were cut in half through the injection site, and immersed in GUS staining solution [1 mM X-Gluc (5-bromo-4-chloro-3-indolyl ß-D-glucuronide), 100 mM phosphate buffer pH 7.2, 0.1% Triton X-100, 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 10 mM EDTA, and 20% methanol]. After a vacuum treatment for 5 min, to facilitate the penetration of the substrate solution, the immersed tissues were incubated overnight in the dark at 37°C. Destaining was done with 70% (v/v) ethanol.

Arabidopsis transformation

All plasmids containing the putative promoter fragments, including positive and negative controls were used for transformation of *Arabidopsis thaliana* ecotype WS plants by the floral dip method (4). After harvesting, seeds were surface-sterilized by soaking 40 mg seeds in 70% ethanol for 1.5 min followed by rinsing once with sterile water and incubation in 4% bleach, 0.05% Tween-20 for 10 min. Seeds were subsequently rinsed four times with sterile water, resuspended in sterile 0.1% water-agarose, followed by plating on 50% MS medium containing 50 µg/ml kanamycin and 200 µg/ml cefotaxim. After 3 days at 4°C, the plates were transferred to a growth chamber with long day conditions (16 hrs light, 24°C). Kanamycin-resistant plants were transferred to soil and the plants were grown in the greenhouse until they were used for GUS assays (as described above).

Desiccation experiments

In order to test the effects of desiccation on gene expression in leaf tissue, the tomato and cacao leaves were harvested and were put on room temperature for 1 h for the tomato leaf and 6 hrs for the cacao leaf. After 1 and 6 hrs, respectively the leaves were put in a plastic bag for 24 hour after which metabolism was stopped by immersion in liquid nitrogen. The frozen leaf samples were ground in liquid N_2 and divided into portions that were subsequently used for RNA extraction.

Results

Cloning of the *TcLea5* promoter by genome walking

In Chapter 4 of this thesis, a cocoa homologue of a cotton *Lea5-D* gene (SSH-contig1, hence called *TcLea5*) seemed to be the most pod wall-specific of all studied cDNAs. We have studied the expression of this gene and its regulation in more detail. The 5' upstream genomic sequence of *TcLea5* was amplified and cloned by genome walking using a series of gene specific primers derived from the SSH-contig 1. Two rounds of genome walking yielded a 3314 bp genomic sequence as shown in Fig. 1. The sequence is at the 3' end bounded by the original contig 1 fragment.

The 3314 bp sequence composed of the genome walking products and the original contig 1 sequence was analyzed by BlastX comparison with the Genbank database. Besides the extended homology with the cotton Lea5 protein encoding DNA, we located at the 5'end

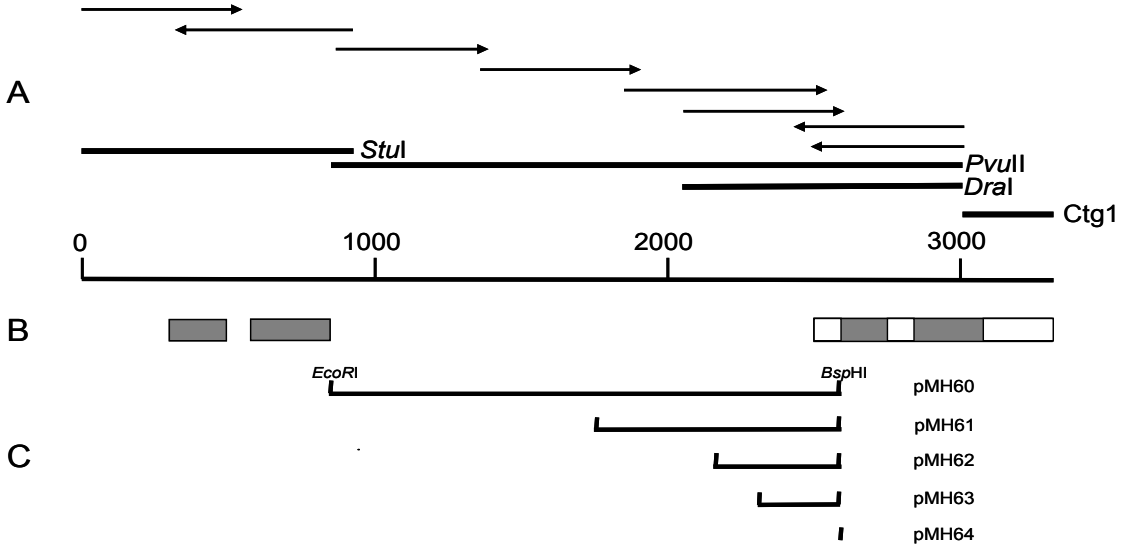


Fig. 1. A. Genome walking and sequencing strategy of the *TcLea5* genomic DNA and its 5'-flanking region. B. Position of (putative) ORF's (gray boxes), and for the *Lea5* gene the entire transcribed region is indicated as an open box C. Promoter constructs used in this study.

two open readings frames (base pairs 308-496 and 573-856) encoding peptides with homology to the product of gene At3g16660 (with unknown function) from *Arabidopsis thaliana.* The position of these ORF's is indicated in Fig. 1**.** At the 3'end of the sequence we found two open reading frames (base pairs 2666-2739 and 2845-3085) with homology to the cotton Lea5-D protein (6). The two ORF's are separated by a single putative intron at the same location as the single intron found in the cotton Lea5-D gene. The location of this intron was furthermore confirmed by alignment with a cocoa EST (GenBank: CA794695) sequence (10) published during our study. The 5' sequence of this EST suggests that the transcription start of the *TcLea5* gene is 5' of position 2509 of the genomic fragment in Fig. 1. Although the published EST sequence contains only 500 bp (to position 3009), our contig 1 sequence extends to 3314. This suggests that the *TcLea5* mRNA has a 5' UTR of 157 nucleotides and a 3' UTR of at least 305 nucleotides. The sequence of the *TcLea5* gene with exons and UTR's indicated, as well as its predicted amino acid product are shown in Fig. 2.

The *TcLea5* gene encodes a protein of 104 amino acids which shares 68%, and 66% amino acid identity with the predicted LEA5-D and LEA5-A amino acid sequences of cotton, respectively (6). Further homology was found with a LEA5 protein from *Citrus sinensis* (59%) (13), an embryonic abundant Lea5 from cotton (61%) and to a lesser extent, a drought-

Fig. 2. Detail of the 3' end of the *TcLea5* genomic fragment indicating the proposed location of the TATA box (shaded) and the 5'end of the mRNA. Amino acid translation of the two exons is given under the nucleotide sequence.

induced cDNA from *Arabidopsis thaliana* (Di21, 50%) (7), and cDNAs from tobacco (49%), tomato (48%) (12), *Solanum demissum* (49%) (9), soy bean (47%) (3) and mung bean (ARG2, 43%) (18). An amino acid alignment of these proteins is shown in Fig. 3. The conserved N-terminal amino acids may constitute a signal sequence for localization in the endoplasmatic reticulum, and TcLea5 also contains the C-terminal sequence motif WAPDPVTGYYRP typical for this family of proteins and not found in other classes of LEA proteins.

The sequence 5' to the *TcLea5* first exon was analyzed for the presence of promoters using the Softberry TSSP software package (http://www.softberry.com) for predicting plant RNA polymerase II promoters. The most probable TATA-box was localized around 216 bp (position 2449; see Fig. 2) upstream of the predicted start codon, close to the 5' end of the EST CA794695 cDNA.

Promoter analysis by transient expression in the tomato fruit

In order to study the role of the putative promoter region of *TcLea5* in regulation and tissue specific expression, we produced a number of promoter subfragment-GUS fusion constructs for use in heterologous systems. Fragments representing the full length of the region between the identified open reading frames of the At3g16660 homologue and of *TcLea5*, as well as three 5' deletions thereof (starting at positions -1793, -813, -451, and -289 relative to the start of the open reading frame), all including the 5' untranslated leader encoding part of *TcLea5* at

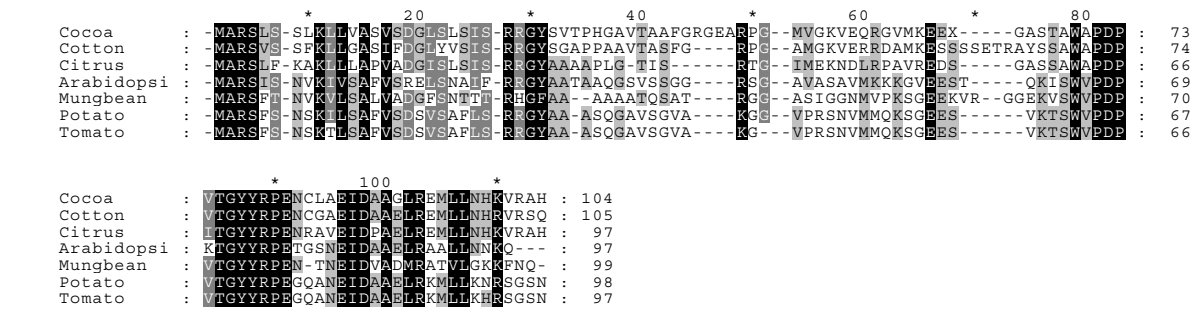


Fig. 3. Amino acid sequence alignment of the cocoa Lea5-homologous gene and it's closest relatives in cotton, citrus, *Arabidopsis*, Mungbean, tomato and potato.

the 3' end, were produced by PCR. These were subsequently fused to the open reading frame of the *uidA* gene of the plasmid pCAMBIA1302-kan, giving plasmids pMH60-63 (Fig. 1).

All fragments of the putative pod specific promoter were tested with transient gene expression assays in the tomato fruit. The tomato fruits were injected with a suspension of *A. tumefaciens* containing the respective plasmids. pMH64, containing no promoter, was used as a negative control, and the CaMV35S promoter in pCAMBIA1303-kan was used as positive control. As can be seen in Fig. 4, plasmids pMH60 and 61 showed the strongest activity, whereas staining was significantly less with pMH62 and 63, while absent for the negative control, pMH64.

Regulation of the *LcLea5* promoter in transgenic *Arabidopsis* plants

As the promoter-GUS fusions constructs used for transient expression contained an intron-less *uidA* (GUS) gene, it is difficult to distinguish the transient expression in the tomato fruit cells from expression in the bacterial cells, which at least for the 35S promoter is known to occur. Therefore, we also produced transgenic *Arabidopsis* plants with the same constructs. Seeds from *Arabidopsis* plants transformed by the floral dip method were grown on soil until flowering and fruit formation. Seven to ten transgenic plants per construct were tested for GUS expression by staining overnight in substrate solution, followed by destaining. Results of three plants each of every construct are shown in Fig. 5. Leaves, flowers, siliques of various ages, and roots were tested for all plants. As can be seen in Fig. 5 for constructs pMH60, 61 and 62 both leaves as well as siliques of various developmental stages, but excluding fully

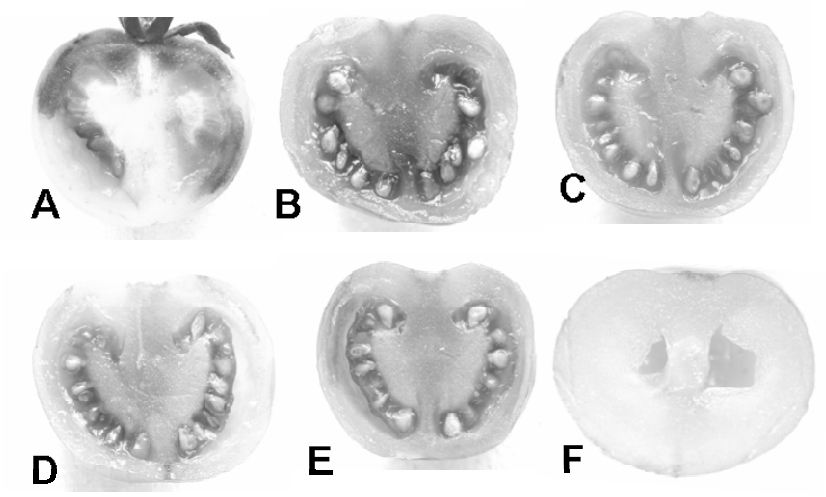


Fig. 4. Transient expression of the cocoa promoter-GUS fusions in tomato fruit. Tomato fruits were injected with *A. tumefaciens* cells containing the respective constructs and after incubation halved and stained with GUS substrate. **A.** pCAMBIA1303 (CaMV 35S promoter)**.**

grown and drying siliques, were showing GUS activity at approximately equal levels. All parts of plants containing the constructs pMH63 and 64 were negative for GUS activity. Observations at intermediate times during incubation in substrate solution did not reveal obvious differences in GUS activity between leaf and siliques for the same construct, or between the same tissues for different constructs. Thus the promoter activity in transgenic *Arabidopsis* requires at least part of the sequence between -451, and -289 relative to the start codon of *TcLea5*.

Expression of the tomato *Lea5* homologue during tomato fruit development

As mentioned earlier, a BlastX homology search of GenBank sequences identified proteins homologous to TcLea5 in tomato and potato. BlastX search of the TIGR tomato gene index and the SGN tomato Unigene database identified a major group of tomato ESTs (TIGR TC123969, 154 ESTs; SGN U212930, 147 ESTs) with homology to the tomato entry in GenBank. We have compared the developmental regulation of expression of this tomato gene in the tomato fruit wall with that of the *TcLea5* gene, as presented in Chapter 4. Northern blots of RNA from tomato fruits (separated in pericarp and locular tissue, with seeds removed) in various stages of development, as well as from mature seeds and leaves were probed with a labeled RNA derived from a tomato Lea5-like clone (TUS15A17). Results are

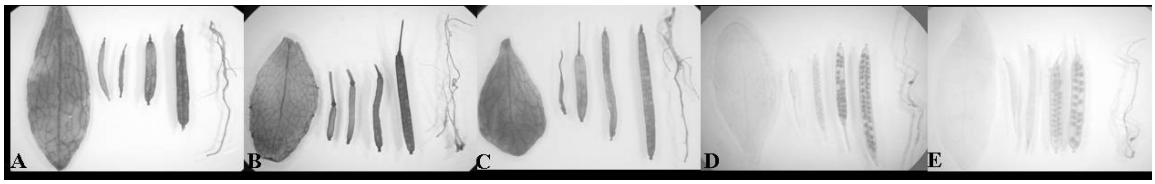


Fig. 5. Expression of the cocoa promoter-GUS fusions in transgenic *Arabidopsis* tissues. Root, siliques in various stages and leaf of one plant for each construct were stained for GUS activity. **A.** pMH60; **B.** pMH61; **C.** pMH62; **D.** pMH63; **E.** pMH64

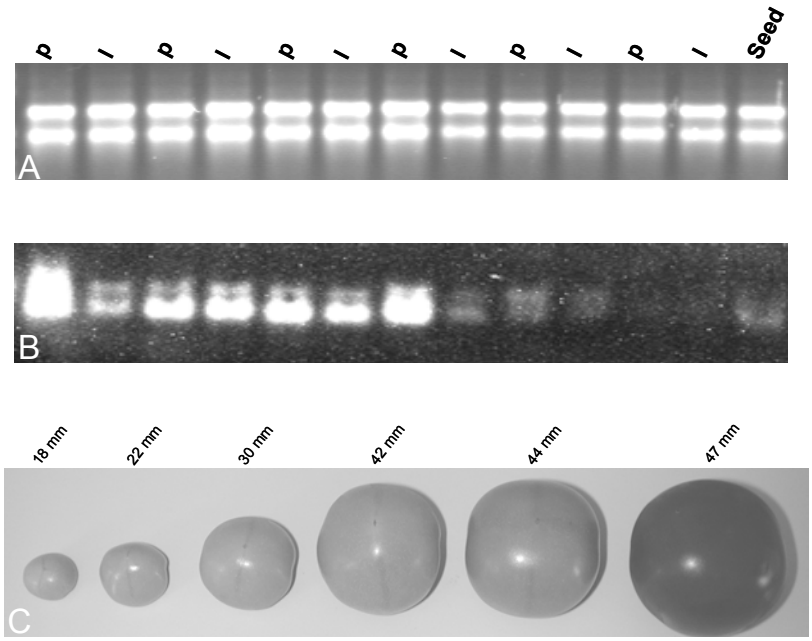


Fig. 6. Northern blot analysis of expression of the tomato *Lea5*-homologue in tomato fruit tissues. Pericarp (P) and locule or jelly tissue (L) as well as mature seed RNA were used. **A.** Total RNA. **B.** Northern blot probed with labelled TUS15A17 RNA; **C.** tomato fruits from which RNA in panels A and B was extracted.

shown in Fig. 6.

As can be seen in Fig. 6, the tomato *Lea5* gene is highly expressed in both pericarp and locular tissue of the developing tomato fruit up to and including the breaker stage, whereas expression is almost undetectable in ripe red fruit tissues Expression in tomato seeds is also virtually undetectable. In fruit tissues as well as in leaves (see below), there is weak cross-hybridization with a slightly larger transcript.

Induction of Lea5 expression in tomato and cocoa leaves by desiccation

Since for several members of the Lea5 family (presented in Fig. 3) was reported that the expression is induced by desiccation (and in some cases, other forms of abiotic stress), we have tested the effect of desiccation on expression of the respective Lea5 homologues in detached leaves of cocoa and tomato, incubated for several hours with and without water. Northern blotting and hybridization with the appropriate probes (Fig. 7) showed that under the conditions used here, neither the tomato *Lea5* homologue nor the cocoa *Lea5* homologue expression was induced by desiccation as applied in this experiment.

Discussion

The *TcLea5* promoter

In this study we have sequenced and partially characterized the promoter region of a cocoa Lea5-homologous gene (*TcLea5*), for which we had shown previously (Chapter 4 of this thesis) that it is almost exclusively expressed in the developing pod wall. Our preliminary characterization of the promoter activity of the 1800 base pairs between the open reading frames of *TcLea5* and a probable gene preceding it (homologous to *Arabidopsis* At3g16660) indicated that the sequence up to 451 bases upstream of the start codon are required and sufficient to show the full range of expression phenotypes, as far as tested and observed, in *Arabidopsis*. This is partially supported by the transient expression data in tomato fruits, although in that case the data have to be considered with some reserve because we cannot exclude (due to the use of an intron-less GUS gene) that promoter activity in the injected bacteria may also have lead to substrate conversion in those experiments.

The promoter activity in transgenic *Arabidopsis* plants appeared to deviate from the expression pattern in cocoa trees. In cocoa, expression in leaves was low compared to that in the pod wall. In contrast to *Arabidopsis* the promoter activity was equally high in leaves as in,

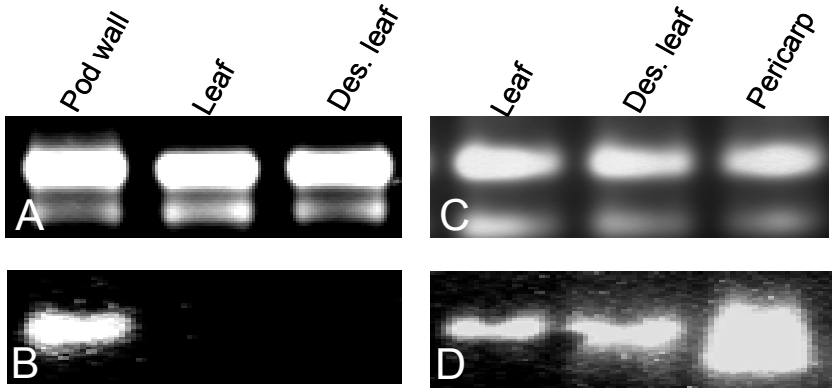


Fig. 7. Northern blot analysis of Lea5 expression in cocoa (A) and tomato (B) leaves. **A.** Total RNA of the cocoa pod wall (positive control) and leaves of cocoa with and without desiccation. **B.** Northern blot of A. probed with labelled*TcLea5*. **C.** Total RNA from leaves of tomato with and without desiccation, of fruit pericarp (positive control). **D.** Northern blot of C., probed with labeled TUS15A17.

the pod wall. In contrast to *Arabidopsis* the promoter activity was equally high in leaves as in, what might be considered an equivalent of the pod, the siliques. Possible explanation for this difference may be that the promoter functions in a fundamentally different way in *Arabidopsis* as compared to cocoa. Alternatively the cloned fragments of *TcLea5* (5' upstream sequence and 5' untranslated leader) may not contain all cis-acting sequences involved in expression regulation, in which case intron or 3' untranslated leader sequences may be required for more correct expression regulation.

What is the function of *Lea5* expression in tomato and cocoa fruits?

The classification of the cocoa gene studied here as *TcLea5* is solely based on the high homology of its product with the proteins encoded by the cotton *Lea5* genes. LEA (Late Embryogenesis Associated) proteins are a large group of probable desiccation protectants that in many plants are expressed during the post-abscission stage in the embryo, when desiccation occurs, as well as in many cases in vegetative tissues during desiccation (8). The cotton homologue was initially cloned from seeds, as well as was shown to be expressed in leaves of water-stressed plants or water-stressed detached leaves (6). It was also recognized in this study that *Lea5* is an atypical LEA protein because it has primarily a hydrophobic character. Most typical LEA proteins are hydrophilic and thought to act as desiccation protectants by

binding or replacing water (8). More recent bioinformatics studies suggest that LEA proteins may form filament-like structures, but conclude again that Lea5 proteins are not typical LEA proteins in this respect either (16, 17).

Since the initial finding of expression of a *Lea5* homologue in cocoa fruits as opposed to seeds (containing the embryo) was puzzling, we looked for similar situations in other species. Curiously, a potato *Lea5* homologue expressed in leaves during tubererization had a tomato homolog expressed in immature fruits, whereas expression diminished in red ripe fruit, although no distinction was made between different parts of the fruit (9). This gene could be traced further back to a study of genes specifically expressed in green and ripe tomato fruits (12). There, it was shown that the tomato *Lea5* homologue (clone 204) is expressed in both pericarp as well as locular tissue of immature and mature green fruit and expression decreased in the breaker stage to undetectable in the red ripe stage. In contrast, our analysis of online available microarray data for the homologous Unigene sequence in the Solanaceae Genomics Network database (SGN-U212930) suggested that expression was increased during the first two weeks of fruit development and then remained constant, including in the ripening stage (data not shown). Our own Northern blot analysis with a tomato cDNA clone identical to the clone 204 confirms that expression is high in the early developmental stages of the tomato fruit in both pericarp as well as in locular tissue, and that expression drops to almost undetectable between the breaker and red ripe stages. The fact that expression of *TcLea5* does not drop off in late developmental stages may well reflect the fact that cocoa pods do not have a climacteric ripening phase accompanied by softening of the fruit. Both in tomato and cocoa *Lea5* expression is also atypical in that there is no expression in the seeds and that expression in the leaves is not induced by desiccation. However, due to the limited scope of our experiments the possibility that other kinds of stress or other application methods may induce gene expression, cannot be excluded yet. Thus, both cocoa as well tomato fruits show high and almost exclusive expression of a *Lea5* homologue, for which the function is unknown but which seems increasingly unlikely to be a typical *Lea* gene both in expression pattern as well as in proposed function. It will be interesting to study in the future the function of the *Lea5* genes in fruit development.

Is the *Lea5* promoter suitable for pod wall specific gene expression?

The initial goal of our study was to identify and clone a cocoa promoter that may be used to direct high levels of transgene expression in the cocoa pod wall with the least possible

expression in other tissues, especially in the beans used for consumption. Although we have tentatively identified a minimum promoter fragment functioning in transgenic *Arabidopsis*, the different expression pattern in those plants (high expression in leaves) raises some questions about the validity of this testing system. Furthermore the fact that expression of several members of the *Lea5* family can be induced by desiccation raises the question whether this may happen in cocoa as well. Although our initial experiments with desiccated leaves suggest that this is not the case for tomato and cocoa, these experiments probably need to be repeated under several different stress conditions. On the other hand, we found very low expression of *TcLea5* in ripe beans, where expression of a transgene is least desirable. The difficulty of testing cocoa promoters in the homologous background may remain an obstacle for this work if a suitable heterologous model system is not found. Possibly expression analysis in transgenic tomato plants, where we have identified a clear *Lea5* homologue, and which in contrast to *Arabidopsis* do produce fleshy fruits, may give clearer results.

Literature cited

1. **Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25:**3389-3402.
2. **Asif, M. H., P. Dhawan, and P. Nath.** 2000. A simple procedure for the isolation of high quality RNA from banana fruit tissue. Plant Mol. Biol. Rep **18:**1 -7.
3. **Burns, W., N. Maitra, and J. Cushman.** 1996. Isolation and characterization of a cDNA encoding a LEA5-like protein from Soybean. Plant Physiol. **112:**1398.
4. **Clough, S. J., and Bent, A.F.** 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:**735-43.
5. **Couch, J. A., and P.J. Fritz.** 1990. Isolation of DNA from plant high in polyphenolics. Plant Mol. Biol. Rep. **8:**8-12.
6. **Galau, G. A., H. Y. Wang, and D. W. Hughes.** 1993. Cotton Lea5 and Lea14 encode atypical late embryogenesis-abundant proteins. Plant Physiol. **101:**695-6.
7. **Gosti, F., N. Bertauche, N. Vartanian, and J. Giraudat.** 1995. Abscisic acid-dependent and -independent regulation of gene expression by progressive drought in *Arabidopsis thaliana*. Mol. General Genet. **246:**10-8.
8. **Ingram, J., and D. Bartels.** 1996. The molecular basis of dehydration tolerance In Plants. Annu. Rev. Plant. Phys. **47:**377-403.
9. **Jackson, S., J. Gascon, E. Carrera, E. Monte, and S. Prat.** 1997. Cloning and expression analysis of a gene that shows developmental regulation upon tuberization in potato. Plant Mol. Biol. **33:**169-74.

10. **Jones, P. G., D. Allaway, D. M. Gilmour, C. Harris, D. Rankin, E. R. Retzel, and C. A. Jones.** 2002. Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta **216:**255-264.

11. **Khanjua, S. P. S., Shasany, A. K., Darokar, M. P. and Kumar, S.** 1999. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Mol. Biol. Rep. **17:**1-7.

12. **Morgens, P., J. Pyle, and A. Callahan.** 1987. Searching for molecular mechanisms involved in fruit ripening., p. 157-166. *In* J. E. Fox and M. Jacobs (ed.), Molecular biology of plant growth control. Alan R. Liss, New York.

13. **Naot, D., G. Ben-Hayyim, Y. Eshdat, and D. Holland.** 1995. Drought, heat and salt stress induce the expression of a citrus homologue of an atypical late-embryogenesis *Lea5* gene. Plant Mol. Biol. **27:**619-22.

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

15. **Spolaore, S., L. Trainotti, and G. Casadoro.** 2001. A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*. J. Exp. Bot. **52:**845-850.

16. **Wise, M. J.** 2003. LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. BMC Bioinformatics **4:**52.

17. **Wise, M. J., and A. Tunnacliffe.** 2004. POPP the question: what do LEA proteins do? Trends Plant Sci. **9:**13-7.

18. **Yamamoto, K. T.** 1994. Further characterization of auxin-regulated mRNAs in hypocotyl sections of mung bean [*Vigna radiata* (L.) Wilczek]: sequence homology to genes for fatty-acid desaturases and atypical late-embryogenesis-abundant protein, and the mode of expression of the mRNAs. Planta **192:**359-64.

Chapter 6

Development of cocoa transformation

Tetty Chaidamsari[1], Herti Sugiarti [2] , Sri Winarsih[3], Djoko Santoso[2], Evert Jacobsen[4], and Ruud A. de Maagd[1]

[1]Business Unit Bioscience, Plant Research International B.V, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[2]Biotechnology Research Institute for Estate Crops, Jl Taman Kencana 1, Bogor 16151, Indonesia

[3]Indonesian Coffee and Cacao Research Institute, Jl Jenderal Sudirman 60, Jember 68118, Indonesia

[4]Laboratory of Plant Breeding, Department of Plant Sciences, Wageningen University, P.O. Box 386, 6700 AJ Wageningen, The Netherlands

Summary

Cocoa is a recalcitrant crop in tissue culture and for micro-propagation by somatic embryos. Regeneration of somatic embryos can be obtained by using zygotic embryos or staminodes as explants. It has been shown that all 5 selected genotypes could regenerate somatic embryos after using zygotic embryos as explant. There was a clear effect of medium composition in all steps of the protocol and of the genotypes used. The primary somatic embryos could be used as explant for secondary embryogenesis. After maturation of the somatic embryos, they could be germinated into plantlets. The frequency of normal looking embryos was high. Staminodes of 2 genotypes were used successfully to develop a somatic embryogenesis protocol. The regeneration protocol with zygotic embryos was successful and a basis for the development of a transformation protocol by using the *gfp* reporter gene as a mean of selection from the beginning. After treatment of primary somatic embryos as explant for transformation with *Agrobacterium tumefaciens* it was possible to select solid looking, *gfp*-expressing transformants. It is expected that the staminode regeneration protocol also can be used for transformation. In this way, both seed and vegetatively propagated varieties can be used for crop improvement by transformation.

.

Introduction

Genetic transformation provides a very powerful plant breeding tool, and has been utilized for many plant species. Crops that have been genetically engineered for insect resistance can dramatically lower production costs and have provided farmers with new insect control options. Genes encoding Bt toxins have been incorporated into and expressed in a large number of plant species since 1987, among which tomato, tobacco, cotton and maize (3, 14, 17).

Genetic modification offers a new opportunity to speed up the development of genetically improved genotypes of crops like the cocoa tree, *Theobroma cacao* L., which is propagated by seed or vegetatively. It includes amongst others tissue culture, introduction of novel genes into the plant and rapid propagation. *Theobroma cacao* L, like most tropical trees is recalcitrant to tissue culture. Early attempts to develop a somatic embryogenesis-based system for cocoa propagation were focused on direct embryogenesis from immature cocoa zygotic embryos. Later, Pence *et al.* (18) and Wen *et al*. (13) improved the process using zygotic embryos in a Murashige and Skoog medium containing 8 μM naphthalene acetic acid (NAA) and 10% (v/v) of coconut water. Somatic embryos can be obtained in many ways. Litz *et al.* (8) was able to induce embryogenesis from cocoa leaf explants, but a complete root system and leaves were not obtained. Lopez-Baez *et al.* (9) have reported the induction of somatic embryogenesis from flower parts including staminodes and petals of cocoa and Sondhal *et al.* (16) produced somatic embryos from nucellar tissue. On recent efforts in tissue culture, Li *et al.* (7) reported the induction of somatic embryogenesis using thidiazuron (TDZ) and 2,4 dichlorophenoxyacetic acid (2,4-D) and Maximova *et al*. (11) further increased the efficiency of this method by developing a technique for secondary embryogenesis using cotyledon explants from primary somatic embryos.

The aim of the research presented here was to optimize and increase the efficiency of *Agrobacterium tumefaciens* mediated transformation of cocoa embryos using somatic embryogenesis, kanamycin resistance as a selectable marker and *gfp* as a reporter gene under control of the CaMV 35S promoter. We have produced germinating somatic embryos derived both from zygotic embryos as well as from staminodes of several cocoa clones. Transformation experiments using some of these embryo cultures produced gfp-expressing secondary and tertiary embryos.

Material and Methods

Somatic embryogenesis

Induction of somatic embryos from zygotic embryos

Two month-old fruits of cocoa from clones Sca6, RCC72, TSH858, ICS60 and DR2 were washed with tap water and surface sterilized with 70% (v/v) ethanol for 5 min. The fruits were opened and subsequently immersed in 70% (v/v) ethanol for 5 min. Zygotic embryos were aseptically excised from the fruits and placed on culture medium. All the chemicals for medium preparation were purchased from Duchefa Biochemie BV (Haarlem, The Netherlands) or Sigma Chemical Co. (St Louis, MO). All media were adjusted to pH 5.8 with 0.1M KOH and autoclaved at 121°C for 20 min. Zygotic embryos were first cultured for 3-4 weeks in the dark at 25°C in somatic embryo induction medium with two different auxins. Induction media consisted of Murashige and Skoog (MS) medium including MS vitamins; one was supplemented with naphthalene acetic acid (NAA) and the other with indole-3-butyric acid (IBA). The concentration of NAA was 8 μM and of IBA were 5, 10, 15, and 20 μM, respectively. Both media were complemented with 10 mg/L coconut water, 15 g/L sucrose, and 2 g/L phytagel. After 3 weeks on induction medium, emerging primary somatic embryos were subcultured on multiplication medium, and maintained for 2-4 weeks under the culture conditions described above. Secondary somatic embryos were induced by subculturing the primary somatic embryos on multiplication medium every 2-4 weeks. Multiplication medium consisted of MS medium including MS vitamins supplemented with 0.05 μM NAA, 1.5 μM 2-isopentenyl adenine (2iP), 40 g/L glucose, and 2 g/L phytagel. Mature secondary somatic embryos with fully developed embryonic axis, hypocotyl and two cotyledons were subsequently subcultured with 2 treatments. (A) The secondary embryos were subcultured on multiplication medium in the dark at 25°C until a next generation of secondary somatic embryos developed. (B) Cotyledon explants excised from secondary somatic embryos were cut with a scalpel into approximately 4X4 mm pieces. This was followed by reintroduction in induction medium with 20 μM IBA and incubation in the dark at 25°C for three weeks, until development of a second generation of secondary somatic embryos emerged. Obtained mature second generation secondary embryos were then subcultured in regeneration medium containing MS medium (including MS vitamins), supplemented with 10 g/L glucose, 0.3 g/L charcoal, and 2 g/L phytagel. The cultures were cultivated under a 12-hr light regime (60 umol/sec/m^2) at 25°C for one month. Later, the

embryos obtained were germinated in a half-strength MS medium containing 0.75 µM gibberellic acid (GA) and 3.8 µM abscisic acid (ABA).

Induction of somatic embryos from staminodes

Floral buds (4-6 mm long), from clone Sca6 and the UAH hybrid clone were picked prior to opening and washed in tap water followed by immersion in 70% (v/v) ethanol for 5 min followed by washing 3 times in sterile distilled water. Staminodes were aseptically excised from buds and placed on culture medium. Cultures were maintained in the dark at 25°C for 14 days. Staminodes were first cultured for 2 weeks on primary callus growth (PCG) medium that consisted of DKW basal medium (Driver and Kuniyuki 1984, Duchefa DO247) (4), supplemented with 20 g/L glucose, 1.7 mM glutamine, 0.5 mM *myo*-inositol, 9 µM 2,4-D, 0.02 µM TDZ, and 2.0 g/L Phytagel. After 14 days on PCG medium, the explants were transferred to secondary callus growth-1 (SCG-1) medium and maintained for another 14 days under the culture conditions described above. SCG-1 consisted of basal salts of the low salt McCown's woody plant medium (Lioyd and McCown, 1980, Sigma M-6774), supplemented with Gamborg's vitamins solution (Gamborg, 1966 Sigma G-1019), 20 g/L glucose, 50 ml/L coconut water, 9 µM 2,4 D, 1.4 µM kinetin and 2.2 g/L phytagel. The explants were subcultured every two weeks in SCG-1 until embryogenic callus occurred. Primary somatic embryos were induced by transfer of floral tissue-derived calli to embryo development (ED) medium and subcultured every 14d, at 25°C in the dark. ED medium consisted of DKW basal salts medium, 1 g/L glucose, 20 g/L sucrose, and 2.2 g/L phytagel. Healthy, mature primary embryos with developed cotyledons were selected. The cotyledons were excised and cut into approximately 4 X 4 mm pieces and subsequently subcultured on SCG-2 medium until development of secondary somatic embryos occurred (7, 11). SCG-2 consisted of McCown's salts supplemented with Gamborg's vitamins solution, 20 g/L glucose, 9 µM 2,4 D, 0.2 µM 6-benzylaminopurine (BA) and 2.2 g/L phytagel.

Transformation of embryos using *Agrobacterium tumefaciens*

Binary plasmid pCAMBIA1303-kan was constructed by replacing the hygromycin resistance gene with the kanamycin resistance gene from pCAMBIA2301. The kanamycin fragment *EcoR*I – *Sac*I from pCAMBIA2301 was ligated into the *EcoR*I – *Sac*I site of pCAMBIA1303. The expression vector pCAMBIA1303-kan has GUS/GFP with the intron under the control of

CaMV 35S promoter. Introduction of pCAMBIA1303-kan in the disarmed *Agrobacterium tumefaciens* strain AGL-0 (6) was performed by electroporation.

The recombinant *A. tumefaciens* strain was grown in 100 ml LB medium supplemented with 25 mg/L kanamycin and 50 mg/L rifampicin at 28°C on a gyratory shaker at 200 rpm, until the cultures reached an optical density (OD) of 1, measured at 600 nm. Bacterial cells were pelleted by centrifugation at 750 g at 25°C for 15 min. The pellet was resuspended in liquid multiplication medium supplemented with 250 µM acetosyringone and the OD was adjusted to 0.5 at 420 nm. Virulence was induced for 5 h at 25°C and 100 rpm agitation. After induction, between 10 and 50 cocoa somatic embryos were added to the *A. tumefaciens* suspension. The mixture was sonicated for 30 s in a Bransonic water bath, followed by 10 min of infection at 25°C and 50 rpm agitation and continued with vacuum infiltration in a desiccator for 10 min. After infection, the *A. tumefaciens* suspension was vacuum-aspirated and the somatic embryos were transferred to fresh liquid multiplication medium overlaid with one sheet of Whatman filter paper and incubated for 48 h at 25°C in the dark for co-cultivation of embryos with *A. tumefaciens* (10).

Detection of GFP expression and selection for transformants

After co-cultivation for 48 hrs the inoculated embryos were transferred into a 50-ml centrifuge tube containing 40 ml of sterile water with 400 mg/L claforan, followed by inversion for several times to wash out attached bacteria from the surface of the embryos. The washing process was repeated and subsequently the embryos were placed briefly on the surface of a sterile paper towel to remove excessive water. The embryos were incubated on solid multiplication medium with 400 mg/L claforan at 25°C in the dark for 4 days. During this period the presence of green fluorescence derived from the expression of the *gfp* gene on the surface of inoculated embryos was monitored on a daily basis using a dissecting microscope equipped with an MVI fluorescence (The Leica MZ FLIII fluorescent binocular) attachment. Embryos with GFP-positive fluorescence were transferred to liquid multiplication medium with 200 mg/L kanamycin and 400 mg/L cefotaxim. Cultures were maintained on a gyratory shaker at 100 rpm under dim light. Embryos were transferred to fresh medium every 10 days. GFP-positive embryos about 3 mm in size were transferred and cultured on solid multiplication medium with 100 mg/L cefotaxim and 200 mg/L kanamycin. Secondary embryos were produced by transferring GFP-positive primary somatic embryos subsequently to multiplication medium with kanamycin and cefotaxim. The concentration of kanamycin

was reduced to 50 mg/ml on regeneration medium and subsequently transferred to regeneration medium without kanamycin. *gfp* expressing plantlets were recovered by following the culture steps for somatic embryo maturation and plant regeneration as described above.

Expression of SN19 protein in transgenic cocoa callus

For expression in cocoa callus the SN19 gene versions from pSN48 (containing synthetic domains I and III) and from pTC11 (containing modifications as described in Chapter 3) were cloned as *BspH*I-BglII or *Nco*I-*Bgl*II fragments in the *Nco*I/*Bgl*II-sites of pCAMBIA1302-kan between the CaMV 35S promoter and NOS-terminator. This resulted in plasmids pMH58 and pMH59, respectively. These were transferred to *A. tumefaciens* Agl0 and used for transformation of somatic embryos derived from zygotic embryos of clone Sca6, as described above. After formation of secondary, transgenic embryos in selection medium, secondary embryos were transferred to PCG medium (see above) to allow callus formation. Callus was maintained on the same medium for 5 months before harvesting. Callus was ground in extraction buffer and SN19 protein was detected by a dot-blot immunoassay as described in Chapter 3.

Results

Somatic embryos derived from zygotic embryos

To evaluate the effect of genotype on the frequency of somatic embryogenesis, zygotic embryos (Fig. 1A) from seeds of 5 different clones were cultured on embryo induction medium. Somatic embryogenesis of cocoa can be induced in MS medium with a high concentration of auxin and a low concentration of cytokinin (12). Two different auxins, NAA and IBA were used to induce the formation of primary somatic embryos. The concentration of NAA used was 8 μM and four different concentrations of IBA 5, 10, 15 and 20 μM were tested. All cultivated zygotic embryos on the different induction media showed somatic embryo formation after 2 to 8 weeks. Fifty zygotic embryos were used from each genotype. The number of somatic embryos formed was found to be between 1- 40 per zygotic embryo when using NAA 8 μM and between 8 - 59 per zygotic embryo after using IBA 20 μM (Table 1) after 7 weeks, depending on the clone. The total number of primary somatic embryos somatic embryos produced reached a maximum between 4 and 7 weeks after primary culture

Table1. Effect of NAA and IBA on primary somatic embryo initiation from zygotic embryos on induction medium after 7 weeks.

Clone	% of zygotic embryos showing somatic embryos formation						
	NAA (μM)*		or	IBA (μM)			
Conc.	0	8	0	5	10	15	20
Sca6	0	33	0	5	7	25	59
RCC72	0	40	0	12	3	30	40
UAH	0	27	0	3	3	20	36
ICS60	0	15	0	10	4	8	18
TSH858	0	1	0	10	12	8	8

*Embryogenesis without germination

initiation (data not shown). In general, the highest concentration of IBA was better than the media with only NAA or IBA at a concentration lower than 20 μM. Among the different clones Sca6, RCC72 and UAH showed the highest embryogenic potential. In 4-7 weeks after start of culture, primary somatic embryos with globular, heart, or torpedo shape were obtained (Fig. 1B, C).

In order to achieve secondary somatic embryogenesis, somatic embryos in globular, heart, or torpedo shape were subcultured on multiplication medium. Multiplication of primary

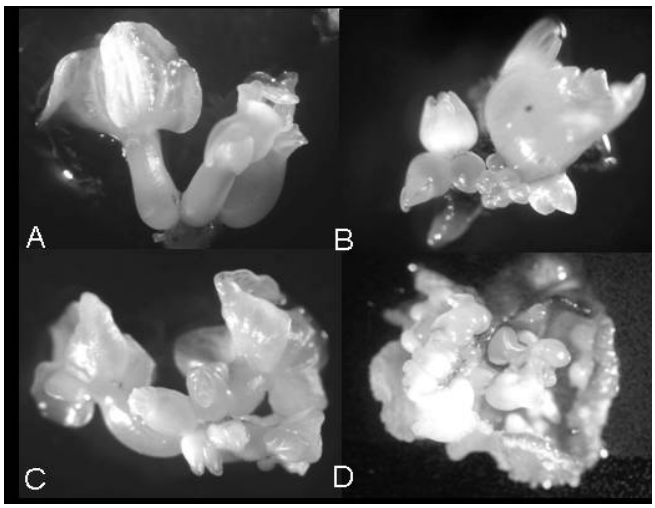


Fig. 1. Direct somatic embryogenesis on zygotic embryos of cocoa. **A.** Zygotic embryo as explant on embryo induction medium. **B.** Primary somatic embryo formation. **C.** Mature somatic embryos on multiplication medium. **D.** Secondary somatic embryo formation on the primary somatic embryo after 10 days on induction medium.

Table 2. Primary and secondary somatic embryos per zygotic embryo after 7 weeks in induction medium with IBA 20 (μM)

Clones	Primary somatic embryo		Secondary somatic embryo	
	% of zygotic embryos producing primary somatic embryos	Average number of embryos per zygotic embryo	% somatic embryo producing sec.embryos	Number of embryos per primary somatic embryo
Sca6	48	8.48	45	2-20
RCC72	48	7.5	47	2-18
UAH	42	7.14	40	1-15
TSH858	10	4.47	7	1-10
ICS60	30	5.5	10	1-15

somatic embryos is defined as the activity to develop secondary somatic embryos. After 4 weeks of subculture on multiplication medium, primary somatic embryos produced secondary somatic embryos (Fig. 1D). The mean number of secondary somatic embryos formed per primary somatic embryo explant was between 2 to 20 (Table 2). The zygotic embryos produced both normal as well as abnormal looking embryos. Each well-developed somatic embryo appeared as a bipolar structure with one embryo axis and two cotyledons (Fig. 2A, B). The abnormally developing somatic embryos appeared with one embryo axis and one cotyledon (Fig. 2C) or one embryo axis without any cotyledons (Fig. 2D, E). In our case, less than 0.01 % of primary somatic embryos produced abnormal somatic embryos (data not shown).

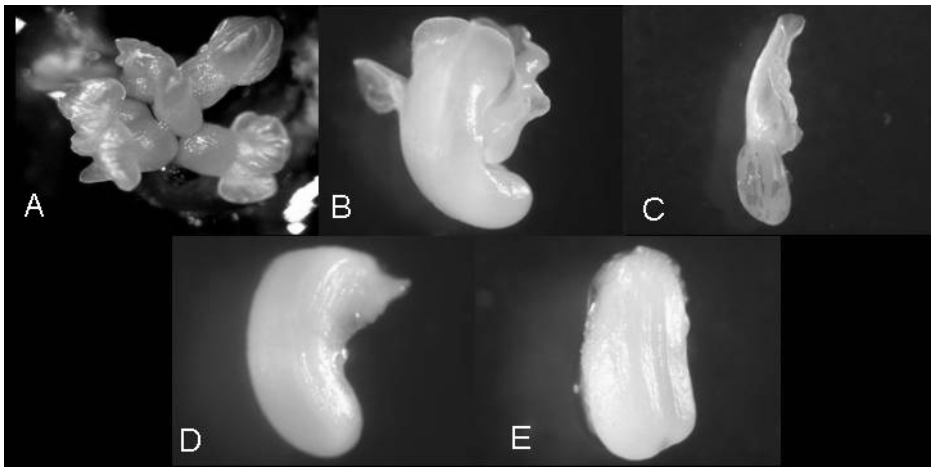


Fig. 2. Embryogenic structures observed during primary and secondary embryogenesis on zygotic embryos of cocoa. **A-B.** Normal primary somatic embryos. **C**. Primary somatic embryo with one cotyledon. **D-E.** Primary somatic embryos without cotyledons.

Table 3. Plantlets produced from secondary somatic embryos (which were derived from primary somatic embryos on zygotic embryos).

Clone	Number of sec. somatic embryos	Number of mature embryos (%)	Number of plantlets (%)
Sca6	56	45 (80)	21 (46)
RCC72	82	69 (84)	46 (54)
ICS60	42	36 (85)	21 (58)
TSH858	33	22 (66)	5 (0.7)

The embryos produced from pieces of somatic embryo-derived cotyledon explants were transferred on induction medium for secondary somatic embryo initiation and the embryo formation was counted every week. The amount of embryos produced was 2-8 per piece in 4-6 weeks after transfer to induction medium. The cotyledon explants on induction medium produced somatic embryos faster compared with the cotyledons transferred to multiplication medium (data not shown). The numbers of somatic embryos produced on both media were similar.

For further development the somatic embryos were transferred to regeneration medium. For maturation, somatic embryos at the torpedo shape were transferred to 45 ml of liquid regeneration medium. The somatic embryos were cultured until small roots appeared on the embryos, this occurred within two weeks. At this developmental stage the germinating

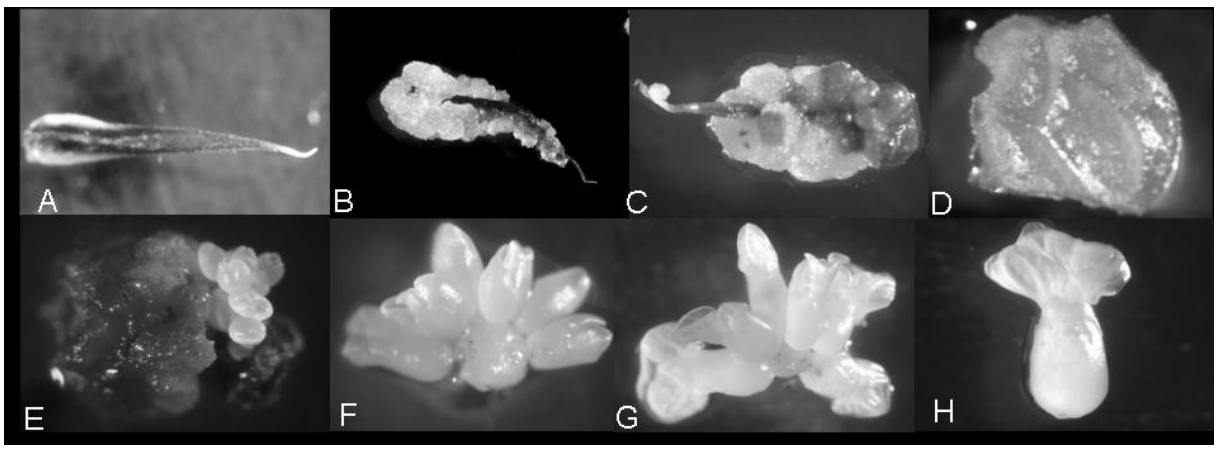


Fig. 3. Somatic embryogenesis of cocoa derived from callus on staminode explants. **A.** Staminode explant after 1 week on primary callus growth (PCG) medium. **B.** Staminode with callus on PCG medium. **C.** Staminode on secondary callus growth medium (SCG-1). **D.** A piece of cotyledon on SCG-2. **E.** Secondary somatic embryos on SCG-2. **F.** Secondary somatic embryos on embryo development (ED) medium. **G.** Somatic embryos on ED medium after 1 months. **H.** Maturing of somatic embryos on ED medium.

Table 4. Total number of *gfp* expressing embryos from the second round of secondary somatic embryos on multiplication medium with kanamycin. Zygotic embryos were used as the original explant.

Genotype	Total number of embryos infected	Total of inoculated embryos expressing gfp	Number of *gfp* expressing secondary embryos	Number of *gfp* expressing tertiary embryos per sec. embryo
Sca6	80	24	27(from 24)	1 - 14
RCC72	75	14	28(from 14)	1 - 9
UAH	60	10	8 (from 10)	1 - 5

embryo was transferred to a solid medium until small true leaves appeared in two months. Based on the initial number of embryos produced, the rate of mature embryo formation was transferred to a solid medium until small true leaves appeared in two months. Based on the initial number of embryos produced, the rate of mature embryo formation was 37 to 80% depending on the clone, whereas the conversion rate of mature somatic embryos to germination was 33-66 % (Table 3).

Somatic embryos derived from staminodes

Primary somatic embryos derived from staminodes were produced from Sca6 and UAH genotypes. On PCG medium after 1 week (Fig. 3A) the staminode explant size increased two to three times compared to the original explant size. Following this period, compact callus developed over the entire length of the staminodes after two weeks in PCG medium (Fig. 3B). This compact callus subsequently was transferred to SCG-1 medium. After 2 weeks or more on SCG-1 medium globular callus clusters (Fig. 3C) were obtained. Subsequently, the globular callus was transferred to ED medium to develop globular shaped embryos. Three or four weeks after culture on ED medium the upper surface of the staminode explants were covered with globular and heart shape somatic embryos. 1% of the staminodes cultured produced mature primary somatic embryos directly. The cotyledons of mature primary somatic embryos were cut into 4 mm^2 pieces and incubated on SCG-2 medium until secondary somatic embryos developed (Fig. 3D,F). After 2 weeks on SCG-2 medium each individual cotyledon produced 1 to 10 secondary somatic embryos (Fig. 3. E,H). Secondary somatic embryos were transferred to ED medium for maturation and subsequently mature somatic embryos were successfully transferred to plant germination medium for plant regeneration (Fig. 4).

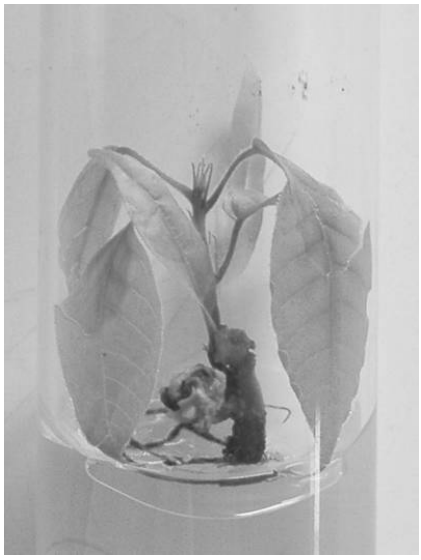


Fig. 4. Regenerated plant from a germinated secondary embryo derived from a staminode.

Transgenic embryo formation and transformation efficiency

All the genotypes studied here produced somatic embryos derived from zygotic embryos. Only the somatic embryos of RCC72, Sca6 and UAH obtained from zygotic embryos were used for transformation.

Somatic embryos in the globular phase were inoculated with *A. tumefaciens* containing pCAMBIA 1303-kan with *gus* and *gfp* as a reporter gene and kanamycin as a selectable marker. In this study, because of lack of sufficient source material, only the *gfp* gene driven by the CaMV 35S promoter and terminator sequences was used as reporter gene. The *nptII* gene driven by the NOS promoter and terminator sequences was used as the selectable marker gene. Green fluorescence in the infected explants due to GFP expression was detected 2 days after infection (Fig 5A). Between 50-90% of the explants from the 3 different genotypes of cocoa had large numbers of fluorescent spots, after infection. The control with pCAMBIA2301-AGL-0 without the *gfp* gene were also performed, and no fluorescence was observed after infection. During the following 22 days, some fluorescent areas enlarged and their fluorescence intensities increased. Somatic embryos emerged on selection medium with kanamycin after 30 days (Figure 5E). The formation of individual gfp expressing secondary somatic embryos was observed between 60- 90 days after infection (Fig. 5G). The formation of transgenic secondary somatic embryos was observed in the three different genotypes after 90 days. 27 of the second generation of gfp expressing secondary

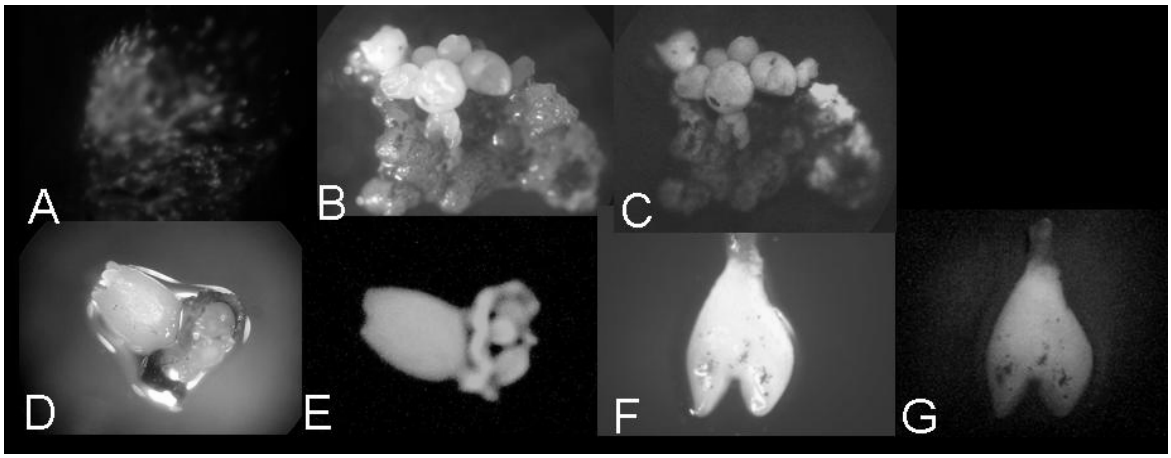


Fig. 5. Transformation of somatic embryos with pCAMBIA1303kan. **A, C, E, G,** fluorescent images. **B, D, F,** visible light images. **A.** Green fluorescent protein expression in cocoa somatic embryos formed on zygotic embryos 2 days after inoculation with *A. tumefaciens*. **B, C.** Transgene expressing somatic embryos. **D, E.** GFP-positive somatic embryos on multiplication medium after 30 days. **F, G.** Mature GFP-positive secondary somatic embryos on multiplication medium after 60-90 days.

somatic embryos were produced from 80 somatic embryos of Sca6, 28 from 75 somatic embryos of RCC72, and 8 from 60 somatic embryos of UAH. The range of second generation somatic embryos formed per embryo was 1 to 14 per somatic embryo (Table 4). The *gfp* expressing secondary somatic embryos were cultured on multiplication medium to produce a new generation of secondary embryos expressing *gfp*, which where regenerated on regeneration medium. Secondary somatic embryos proved positive for the NPTII and *uidA* genes in PCR reactions with specific primers (results not shown). Transgenic plantlets produced *in vitro* on regeneration medium proved to be fully positive for GFP (Fig. 6).

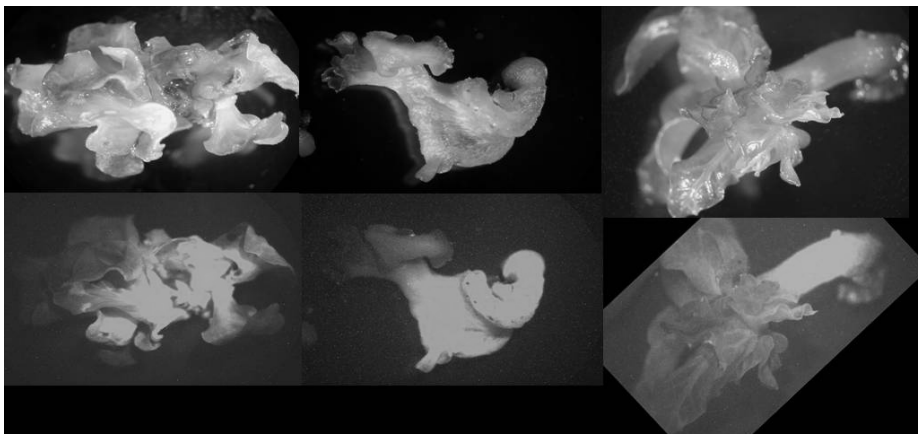


Fig. 6. Regenerating transgenic plantlets of transgenic (*gfp*) secondary embryos derived from zygotic embryos. Top panels: visible light images. Bottom panel: fluorescent images.

Expression of SN19 protein in transgenic callus

In order to test the expression of *Bacillus thuringiensis* toxin gene *SN19* (see Chapters 2 and 3) with the transformation system described here, we constructed transformation vectors pMH58 and pMH59 containing, under control of the CaMV 35S promoter, *SN19* with synthetic domains I and III-encoding fragments (comparable with pSN48 in Chapter 3) and with our own modified fragments (compare with pTC12 in Chapter 3), respectively. Somatic embryos derived from staminodes of clone Sca6 were transformed with these constructs. The secondary somatic embryos arising after transformation in selection medium were transferred to callus growth medium and callus was allowed to grow and maintained for 5 months.

Fifteen and ten calli, for pMH58 and pMH59, respectively, were tested for SN19 protein levels in an immunoassay. Preliminary data show that as expected, calli from pMH58 overall showed the highest expression levels, ranging from 0 to 1.4% of total soluble protein, while pMH59 ranged from 0 to 0.2%.

Discussion

Somatic embryogenesis from cocoa tissue might be a seasonal phenomenon, affected by the physiological stage of the explant tissues (18). The limitation in primary embryogenesis and secondary embryogenesis was that the efficiency was genotype dependent (2, 7, 11, 13, 16). However our results showed that all of the clones could produce somatic embryos. The clones RCC72, Sca6 and UAH produced more embryos per explant compared to the other two clones ICS60 and TSH858. TSH858 produced very low numbers of embryos. A difference was also found in different zygotic embryos from different pods of the same clone, because. the development stage of cocoa pods played an important role in the somatic embryogenesis (11, 18). Wen *et al.* (18) found that MS medium supplemented with 32 μM NAA showed a somatic embryogenesis ratio of 25% when using Amelonado as variety. In contrast Esan *et al* found with Mexican varieties that MS medium supplemented with 8 μM NAA showed a somatic embryogenesis ratio of 27%. According to previous results (data not shown) and Esan *et al.* (5) we decided to use in our experiment 8 μM of NAA, because of the origin of our clones

Our results showed that IBA 20 μM in induction medium gave the highest percentage of somatic embryos derived from zygotic embryos (Table1) and more than 20 % of mature somatic embryos could be regenerated into plantlets. 8 μM NAA (Table 2) in induction

medium produced more than 10% somatic embryos, but most of these somatic embryos failed to develop into normal plantlets after germination.

Recent advances in somatic embryogenesis of cocoa have enabled the production of large numbers of aseptically grown cocoa plants from staminode and petal cultures. Our results showed that primary somatic embryos derived from zygotic embryos developed in shorter time periods compared to primary somatic embryos derived from staminodes. The production of somatic embryos from staminodes takes 2 to 6 months from callus initiation to embryo production (1) and from zygotic embryos as explant it required 2 to 7 weeks. As expected, embryogenesis rates were also different among staminode explants from different clones. In our case the embryogenesis rate of staminodes was lower compared to that of somatic embryogenesis from zygotic embryos.

Our results showed that 1% of the staminodes directly produced primary somatic embryos after culturing on secondary callus growth (SCG) medium. The secondary somatic embryos from staminodes were also more uniform and developed in shorter time periods compared to the primary somatic embryos (Fig 3F-H). After 30 days on SCG medium the staminodes produced large numbers of cell clusters consisting of round cells which later developed into somatic embryos (11).

Less than 0.01 % (data not shown) of the primary somatic embryos derived from zygotic embryos were abnormal. Abnormal embryo formation could be the result of variation in the number of cells that participated in the formation of each individual embryo (11). The staminodes produced normally shaped somatic embryos including cotyledons and hypocotyls.

The use of the *gfp* reporter gene enabled us to follow the early events in *Agrobacterium* – mediated cocoa transformation. The intron-*gfp*-reporter gene system is useful for monitoring the transfer of the *gfp* gene from bacterial cells into the plant cell in the early stages of regeneration. Interruptions of a reporter gene by a plant intron prevented expression of the reporter gene at the bacterial level and allowed the expression of this gene only at the plant cell level. Our results showed that 25% of the explants produced one or more *gfp*-expressing somatic embryos. Until now, only a few publications about successful transformation systems in cocoa do exist (10, 15). In addition, plant regeneration via somatic embryogenesis provides an alternative approach for clonal propagation of cocoa. GFP expressing somatic embryos were produced from Sca6, RCC 72 and UAH after 30 days of inoculation. Ninety days after inoculation individual gfp-expressing secondary somatic embryos were released. The *gfp*-expressing secondary somatic embryos looked fully

transgenic because they were fluorescing throughout the whole embryo. Naturally this requires more molecular proof once the regenerated plants are bigger or more preferably, through confirmation of transmission of the transgene through seeds. Plants derived from somatic embryos are in essence genetically identical to their parental donor cells (zygote or staminode). The mean number of total *gfp*-expressing secondary somatic embryos, obtained from three different clones, was similar. We were also able to multiply the gfp-expressing somatic embryos in multiplication medium. Mature gfp expressing somatic embryos will be multiplied on multiplication medium and healthy embryos will be cut into 4 X 4 mm pieces and transferred to induction medium to produce a new generation of *gfp*-expressing somatic embryos. These embryos did germinate, as thus we have strong indications that we produced transgenic regenerants from zygotic embryos. It is expected that the same results can be obtained after using staminodes as explants, which could be used to transform varieties or clones in a direct way.

Our preliminary experiments with transformation of Bt toxin SN19 expression constructs have shown that the gene is expressed in transgenic cocoa calli as it was before in *Arabidopsis* (Chapter 3). Thus regeneration of plantlets from these calli may yield transgenic cocoa trees expressing *SN19* under control of the CaMV 35S promoter.

Acknowledgements

We would like to thank Dr. Frans Krens (Plant Research International) for critically reading the manuscript.

Literature cited

1. **Abdoulaye, T., S.N. Maximova, and M. J. Guiltinan.** 2003. Micropopagation of *Theobroma cacao* L using somatic embryo-derived plants. In Vitro Cell. Dev. Biol. Plant **39:** 332-337.
2. **Alemanno L., B. M., Michaux-Ferrière N.** 1996. Histology of somatic embryogenesis from floral tissues of cocoa. Plant Cell Tiss. Org. **46:**187-194.
3. **de Maagd, R. A., D. Bosch, and W. J. Stiekema.** 1999. *Bacillus thuringiensis* toxin mediated insect resistance in plants. Trends Plant Sci. **4:**9-13.
4. **Driver, J. A., Kuniyuki, A. H.** 1984. In vitro propagation of Paradox walnut rootstock. HortScience **19:**507–509.
5. **Esan, E. B.** 1977. In vitro studies on morphogenesis and plant regeneration in tissue cultures of cacao and *Herrenia* species Presented at the International Cacao Res. Conf., Ibadan, Nigeria.
6. **Gerard. R. Lazo., Pascal. A. Stein., and R. A. Ludwig.** 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. Bio-Technol. **9:**963-967.

7. **Li, Z., Traore, A., Maximova, S., Guiltinan, M. J.** 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.

8. **Litz, R. E.** 1986. Tissue culture studies with *Theobroma cacao.*Presented at the Cacao Biothecnology Symposium, Pennsylvania State University, university Park, PA.

9. **Lopez-Baez, O., H. Bollon, A. Eskes, and V. Petiard.** 1993. Embriogenese somatique de cacaoyer Theobroma cacao La portir de pieces florales. Comptes rendus de l'academie des sciences Serie III - Sciences de la Vie **316:**579-584.

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

11. **Maximova, S. N., Alemanno, L., 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.

12. **Ndoumou, D. O., Ndzomo, G.T., Niemenak, N.** 1997. Phenol content, acidic peroxidase and IAA-oxidase during somatic embryogenesis in *Theobroma cacao* L. Biol. Plantarum **39:**337-347.

13. **Pence, V. C., Hasegawa, P.M., and Janick, J.** 1979. Asexual embryogenesis in *Theobroma cacao* L. J. Am. Soc. Hortic. **104:**145-148.

14. **Perlak, F. J., R. L. Fuchs, D. A. Dean, S. L. McPherson, and D. A. Fischoff.** 1991. Modification of the coding sequence enhances plant expression of insect control protein genes. Proc. Natl. Acad. Sci. USA **88:**3324-3328.

15. **Sain, S. L., Oduro, K.K.and Furtek, D.B.** 1994. Genetic transformation of cocoa leaf cells using *Agrobacterium tumefaciens*. Plant Cell Tiss. Org. **37:**243-251.

16. **Sondahl, M. R., Liu, S., Bellato, C.M. and Bragin, A.** 1993. Cacao somatic embryo-genesis. Acta Hortic. **336:**245-248.

17. **Vaeck, M., A. Reynaersts, H. Höfte, S. Jansens, M. D. Beuckeleen, C. Dean, M. Zabeau, M. V. Montagu, and J. Leemans.** 1987. Transgenic plants protected from insect attack. Nature (London) **328:**33-37.

18. **Wen. M.C, and J. E. Kinsela.** 1991. Somatic embryogenesis and plantlet regeneration of *Theobroma cacao* L. Food Biotechnol. **5:**119-137.

Chapter 7

Isolation and characterization of the MADS-box genes *AGAMOUS* and *APETALA1* homologues from cocoa

Tetty Chaidamsari[1], Samanhudi[2], Herti Sugiarti[3], Djoko Santoso[3], Gerco C. Angenent[1], and Ruud A. de Maagd[1]

[1]Business Unit Bioscience, Plant Research International BV - PO Box 16, 6700 AA Wageningen, The Netherlands

[2]Faculty of Agriculture, Sebelas Maret University - Jl. Ir Sutami 36A, Surakarta 57126, Indonesia

[3]Indonesian Biotechnology Research Institute for Estate Crops - Jl. Taman Kencana 1 Bogor 16151, Indonesia

Summary

The ABC model for floral development proposes that the coordinate expression and function of A, B, and C-type transcription factor genes determines the identity of the flower organs (sepals, petals, stamens, pistil). The MADS-box transcription factor gene *AGAMOUS (AG)* from *Arabidopsis thaliana* is the C-type gene and is responsible for carpel and stamen development as well floral determinacy. Here we report the cloning of a cDNA encoding *TcAG*, an *AG* homologue from cocoa (*Theobroma cacao* L.), which was expressed primarily in stamens and ovaries, comparable to AG in Arabidopsis. We found that *TcAG* is also expressed in the fruit (pod) wall during its entire development.

The MADS-box transcription factor gene *APETALA1(AP1)* from *A. thaliana* is one of two A-type genes and is co-responsible for sepal and petal development. Additionally, it functions in earlier developmental stages in determination of the floral meristem identity and formation of flowers. We have cloned a cDNA encoding *TcAP1,* an *AP1* homologue from cocoa, which is primarily expressed in sepals and petals, but for which expression can also be detected in flowering cushions, the meristematic stem tissue from which flowers develop.

Introduction

Cocoa trees produce a large number of flowers at certain times of the year from the age of 3 years, under good growing conditions. Flowers emerge from floral cushions (Fig. 1A), a meristematic tissue in the bark of the stem and branches at the location of old leaf axils (16), a habit called cauliflorous or truncate. The flowers are small, about 15 mm in diameter and form long pedicels (Fig. 1B). The flowers have 5 unfused sepals and petals, 10 stamens and an ovary of 5 united carpels (Fig. 1B, C). The petals are very narrow at the base but expand into a cup-shaped pouch and end in a broad tip or ligule. The ten stamens are in two whorls, the outer whorl is consisting of 5 long non-fertile staminodes and the inner whorl of five fertile stamens (Fig. 1C). Cocoa trees produce large numbers of flowers in certain times of the year, when flowering progresses in waves from the lower stem flowering cushions to the branches. Flowers are open pollinated by small insects and only 1-5 % of the flowers are successfully pollinated and proceed to produce a pod (24).

In most higher eudicotyledonous flowering plants the floral organs are arranged in four different whorls, containing sepals, petals, stamens and carpels, respectively. The specification of floral organ identity is explained by the ABC model (9, 23) and extensions thereof (13). Expression of class A genes specifies sepal formation in whorl 1. The combination of class A and B genes specifies the formation of petals in whorl 2. Class B and C genes specify stamen formation in whorl 3, and expression of the class C gene alone determines the formation of carpels in whorl 4. Class A, B, and C genes have been isolated from several eudicotyledonous model plants such as *Arabidopsis*, *Antirrhinum* and *Petunia*, and most of them belong to the family of MADS-box genes, encoding transcription factors

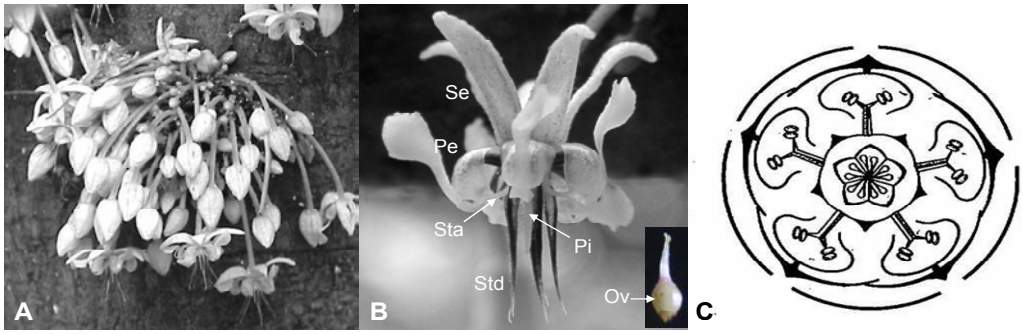


Fig. 1. A. Cocoa flowering cushion. **B.** A cocoa flower with organs indicated. Se, sepal; Pe, petal; Sta, stamen; Std, staminode; Pi, pistil; Ov, ovary; The insert shows a dissected ovary and pistel. **C.** Floral diagram (source: (24))

containing a conserved 56 amino acid motif within their DNA-binding domains, named MADS after the four original members (for a review: see (13)). In *Arabidopsis thaliana* A function is performed by the *AP* (*APETALA*) 1 and 2 genes, where only *AP1* is a MADS-box gene. In *Arabidopsis, AP1* also functions together with the non-MADS transcription factor gene *LEAFY* (*LFY*) in determination of the floral meristem identity and the concomitant formation of the flowers (17). In addition, in *Arabidopsis AP1* acts in a redundant manner with another, closely related MADS-box gene, *FRUITFULL* (*FUL*) to control inflorescence architecture and floral meristem identity (12). The *Arabidopsis* C-function, establishing identity of the reproductive organs as well as determinacy of the floral meristem, is performed by the MADS-box gene *AGAMOUS* (*AG*) (28). In *Antirrhinum majus*, two closely related genes, *PLE* and *FAR* are partially redundant in carrying out C-function (10). Moreover, *AG* functions in a redundant manner with the closely related genes *SHATTER PROOF* (*SHP*) *1/2* in the control of carpel development and with *SHP1/2* and *SEED STICK* (*STK*) in determining ovule identity (21).

In this manuscript we describe the cloning and characterization of two cDNAs from flowers and fruits of the cocoa tree (*Theobroma cacao* L.), homologous to the A-type (*AP1*) and C-type (*AG*) MADS-box genes, respectively. Results from expression analysis as well as from the phenotypes caused by their ectopic expression in transgenic *Arabidopsis* plants suggest that the genes represented by these cDNAs have similar functions in regulating cocoa flowering induction and flower organ specification as their homologues in *A. thaliana*.

Materials and methods

Plant material

The plant material was obtained from cocoa trees, which were grown in the fields of the Biotechnology Research Institute for Estate Crops, Bogor Indonesia. Flower organs were separated in Indonesia in liquid nitrogen before shipping. Fruits in different stages of development were shipped to Wageningen and after arrival sliced, immediately frozen in liquid nitrogen, and stored at -80ºC until further use.

RNA extraction

Total RNA from developing pods and flower organs was isolated as described in chapter 4 using our modification of the method described by Asif *et al.* (2).

Table 1. Oligonucleotide primers used in this study

Name	Nucleotide sequence (5' to 3')
TcAP1F	ATG GGA AGA GGT AGG GTT CA
TcAG-R	GGTGACCGTAGCACTTACTCCACCAGA
TcAG-F	CCATGGAGTACCAAAGTGAATCC
TcAG-F396	CGCATTGCCTATGAAGGATC
AP1FRT-PCR	ATGGAAGAGGTAGGGTTCA
AP1RRT-PCR	CTTCTCCCATGTAGTGCCTG
M13 forward	GTAAAACGACGACGGCCAGT
M13 reverse	GGAAACAGCTATGACCATG

Amplification of cDNAs from cocoa homologous to *APETALA1* and *AGAMOUS*

To amplify a cDNA of an *AGAMOUS* homologue from cocoa, a reverse (3') primer complementary to the sequence beyond the stop codon of SSH-contig 24 (see Chapter 4) was designed (TcAG-R, Table 1). The forward (5') primer (TcAG-F, Table 1) was designed on basis of the consensus around the start codon of the sequences of *AG* homologues of 6 different plants (*Populus balsamifera subsp. trichocarpa PTAG1* (Accession nr. AF052570) and *PTAG2* (AF052571), *Gossypium hirsutum* MADS-3 (AY083173), *Cucumis sativus CUM1* (AF035438), *Arabidopsis thaliana AGAMOUS* (X53579), *Malus domestica* C-type MADS box protein MADS14 (AJ251117), and *Petunia hybrida* PMADS3 (X72912). Two µl of the first strand cDNA from the cocoa flower bud were used as a template for PCR using the kit HSRT-20 (Sigma Bioscience). The parameters used during the PCR reaction were as follows: 95°C, 2 min; 36 cycles: 95°C, 25s; 50°C, 30s; 70°C, 2 min. An additional 5 min at 70°C was added in the last step.

In order to amplify a full length *AP1* homologous cDNA from cocoa, a 5' (forward) primer (TcAP1f, Table 1) was designed based on a consensus of the conserved region including the start codon ATG of the sequences of *Apetala1* genes from 14 different species: *Citrus sinensis* (AY338974), *Populus balsamifera* (AY615966), *Sinapsis alba* (X81480), *Arabidopsis thaliana* (NM10558, *Oryza sativa* (AF109153), *Betula pendula* (X99655) *Vitis vinifera* (AY538746), *Chloranthus spicatus* (AY316311), *Malus domestica* (AY071921), *Brassica oleracea* (AJ505845), *Pisum sativum* (AJ279089), *Eucalyptus globulus* (AF305076), *Dendrobium grex* (AF198174), and *Nicotiana tabacum* (AF009126). First strand cDNA was synthesized from total RNA of cocoa flowers, using Superscript Reverse Transcriptase

(InVitrogen) with oligo(dT)$_{18}$-anchor primers according to the manufacturer's instructions. Two µl of the first strand cDNA were used as a template for PCR with the gene specific primer TcAP1f and oligo(dT)$_{18}$ with superTaq DNA polymerase (5U/µl). The parameters for the PCR reaction were as follows: 95°C, 4 min; 40 cycles: 95°C, 30s; 49°C, 35s; 72°C, 2 min. In an additional 5 min at 72°C was added in the last step.

The obtained PCR-products from both *AP1* and *AGAMOUS* reactions were gel purified using a QiaexII gel extraction kit (Qiagen, Hilden, Germany) and subsequently cloned in a pGEM-T Easy cloning vector (Promega, Madison, WI, USA). The ligated DNA was finally transformed into *E. coli* XL-1 Blue competent cells.

Sequence analysis

Sequencing reactions were performed on PCR amplified fragments using the DYnamic ET Terminator Cycle Sequencing DETT reaction mix (Amersham Pharmacia Biotech, U.S.A.) and analyzed in a 3100 Genetic Analyzer (ABI Prisma). The parameters used for PCR reactions were as follows: 96°C, 3 min; 29 cycles: 96°C, 30s; 50°C, 15s; 60°C, 4 min. Sequences were analyzed and assembled into contigs using the LASERGENE SEQMAN II software (DNASTAR, Madison, Wi). Protein homology searches in the NCBI Genbank protein database were done using the BlastX algorithm (1). Multiple protein alignments were produced with the ClustalW program of the DNAstar Megalign software package and visualized using the TreeView program (19).

Expression analysis by reverse transcription polymerase chain reaction (RT-PCR)

First stand cDNA was synthesized from total RNA from stem cushions with vegetative shoots, passive cushions, actively flowering cushions, flower buds, sepals, petals, staminodes, stamens, ovaries, developing pods (17, 15 and 11 cm length), pulp, beans and leaves, using oligo(dT)$_{23}$ primer and Superscript Reverse Transcriptase (InVitrogen). 2 µl each of the first strand cDNAs were used as a template for PCR with gene specific primers for AP1 or AG. The primers for *AGAMOUS* were TcAG-F396 and TcAG-R (Table 1). The primers for *AP1* were AP1FRT-PCR and AP1RRT-PCR (Table 1). The parameters used during PCR reactions were as follows: 94°C, 4 min; 39 cycles: 95°C, 30s; 50°C, 30s; 70°C, 1 min. In an additional 5 min at 70°C was added in the last step.

Plant vectors and transformation

The fragment containing the full *TcAP1* cDNA (824 bp) was excised by *Nco*I and *Spe*I digestion from pGEM-T Easy and ligated into the *Nco*I-*Nhe*II sites of pCAMBIA2301, between the CaMV 35S promoter and nos terminator, giving plasmid pSM2. The fragment containing the *TcAg* cDNA was isolated by *Nco*I and *BstE*II digestion from pGEM-T Easy and ligated into the *Nco*I- *BstE*II sites of pCAMBIA2301, giving plasmid pTC22. The ligation reactions were incubated overnight at 16°C. The ligated DNA was finally transformed into *E. coli* XL1-Blue competent cells.

The plasmids pSM2 and pTC22 were used for transformation into *Agrobacterium tumefaciens* strain Agl0 by electroporation, and then introduced into *Arabidopsis thaliana* ecotype Columbia by the floral dip method (8). Resulting *Arabidopsis* seeds were surface sterilized by soaking 40 mg seeds in 70% ethanol for 1.5 min, and rinsed once with sterile water. The seeds were incubated in 4% bleach + Tween 20 (0.05%) for 10 min, followed by centrifugation at 1,000 g for 1 min. Subsequently the seeds were rinsed with sterile water four times, with centrifugation at 12,000 g for 30 sec in between. Finally the seeds were resuspended in sterile 0.1% agarose, followed by plating on 50% MS medium containing 50 mg/L kanamycin and 200 mg/L cefotaxim, for selection of transformed seedlings. After 3 days at 4°C the plates were transferred to a growth chamber with long day conditions (16 hrs light, 24°C). Kanamycin-resistant plants were transferred to soil and the plants were grown in the growth chamber (16 hrs light, 21°C) for analysis of the flowering time.

Results

Isolation and sequence analysis of *TcAG* cDNA from cocoa

During our earlier work (Chapter 4) we found several copies of a fragment of a cocoa cDNA homologous to *AGAMOUS*-like genes from several plant species in a Suppression Subtractive Hybridization (SSH) library putatively enriched for pod wall cDNAs as compared to leaf and bean cDNAs. This sequence of approximately 780 bp corresponded to the part of the gene of its closest homologue, *Populus balsamifera* subsp. *trichocarpa AG1* (Accession number: AAC06237) encoding the C-terminal 2/3 of the protein. In order to amplify a cDNA covering the entire open reading frame of the cocoa *AG* homologue, we used a 3' primer complementary to the area including and following the stop codon, as well as a 5' primer covering the start codon, based on an alignment of the six most homologous AG proteins

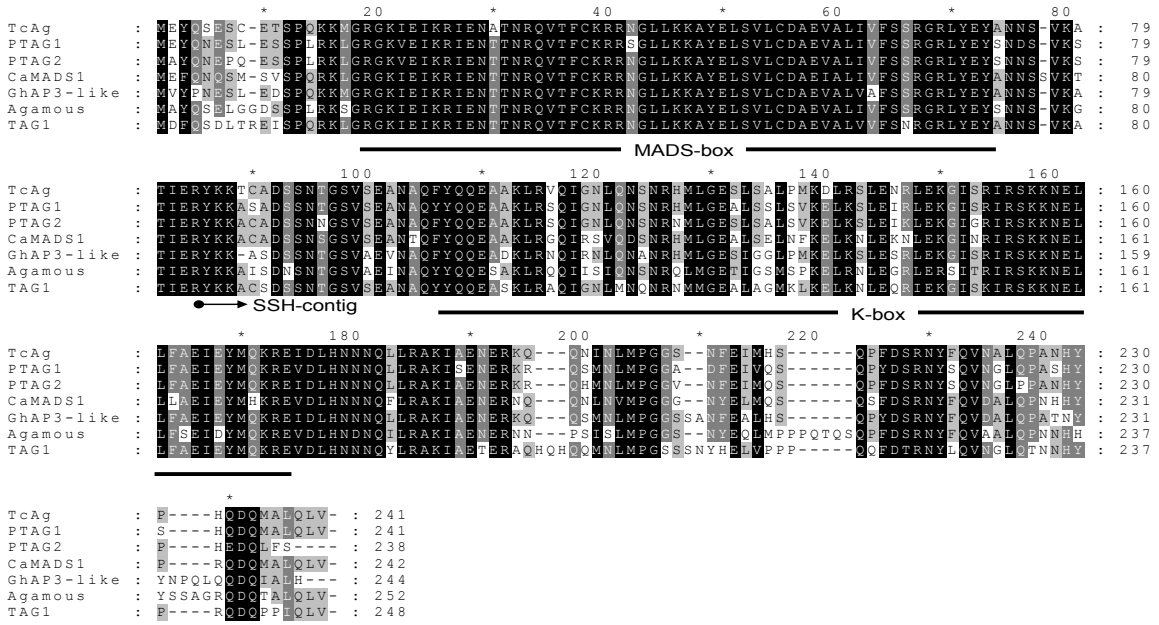


Fig. 2. Amino acid alignment of TcAG and its closest homologues (for accession numbers, see the text). Gh, *Gossypium hirsutum* (cotton); Pt, *Populus balsamifa* subsp. *trichocarpa* (black cottonwood); TAG1, tomato AG. Alignments were made with the ClustalW program and edited with the GeneDoc program (http://www.psc.edu/biomed/genedoc). The conserved MADS-box and K-box are underlined, and the start of the original SSH contig sequence is indicated

from other species. The primers were applied in an RT-PCR reaction on RNA from flowers and resulted in a cDNA product of approximately 1000 bp, which was cloned and sequenced. The sequence coded for a protein, hence called TcAG of 241 amino acids, of which the terminal 158 amino acids were identical to those encoded by the earlier identified SSH-fragment. The resulting protein has as its closest homologues in Genbank (85-86% amino acid identity) the two black cottonwood, *Populus balsamifera subsp. trichocarpa AG* proteins, followed by hazel nut, *Corylus avellana* CaMADS1 (AAD03486), cotton, *Gossypium hirsutum* AP3-like protein (AAL92522), peach, *Prunus persicae* PpMADS4 (AAU29513) and European white birch, *Betula pendula* MADS4 (CAB95649). An amino acid alignment of TcAG with it most closely related relatives as well as with the more distant *Arabidopsis* AG (P17839) and tomato TAG1 (AAA34197) proteins is shown in Fig. 2. Clearly recognizable and indicated are the strongly conserved MADS-box and somewhat lesser conserved K-box, while the C-terminus is the least conserved portion of the AG homologues. TcAG also contains the extension N-terminal to the MADS-box that is typical for AGAMOUS and closely related proteins. *A. thaliana AG*, together with *STK* and the SHP1/2 genes is part of a subfamily of MIKC-type genes, which likely represent a monophyletic clade, and whose

members have partially redundant functions (4). In order to more accurately position TcAG in this subfamily, we created a phylogenetic tree of a selection of AG, STK, and SHP protein homologues from selected species (Fig. 3). As can be seen in this figure, TcAG is firmly embedded in the AG subclade, closest to a cotton protein, curiously called AP3-like, and to the two AG proteins from poplar, while clearly separated from the SHP and STK subclades.

TcAG expression pattern

To gain insight into the developmental role of *TcAG,* its expression pattern was analyzed by RT-PCR, using primers designed to amplify the less-conserved 3'region, on various vegetative and reproductive tissues of cocoa trees: leaf, shoot producing stem cushions, passive and actively flowering cushion tissue, flower buds, floral organs (sepals, petals, staminodes, stamens and ovaries) and developing (young, middle and mature) cocoa pod wall, mature pod pulp, and beans. The results (Fig. 4) revealed that the transcript was not detectable in vegetative tissues (leaf, vegetative shoot producing stem cushion), nor in passive or actively flowering cushion tissue, but only first detectable in flower buds. Within open flowers, transcript levels were highest in the organs of the inner two whorls (stamens and ovary) while less in staminodes and barely detectable in sepals and petals. *TcAG* transcript was also clearly detectable in all stages of the developing pod wall. This expression pattern

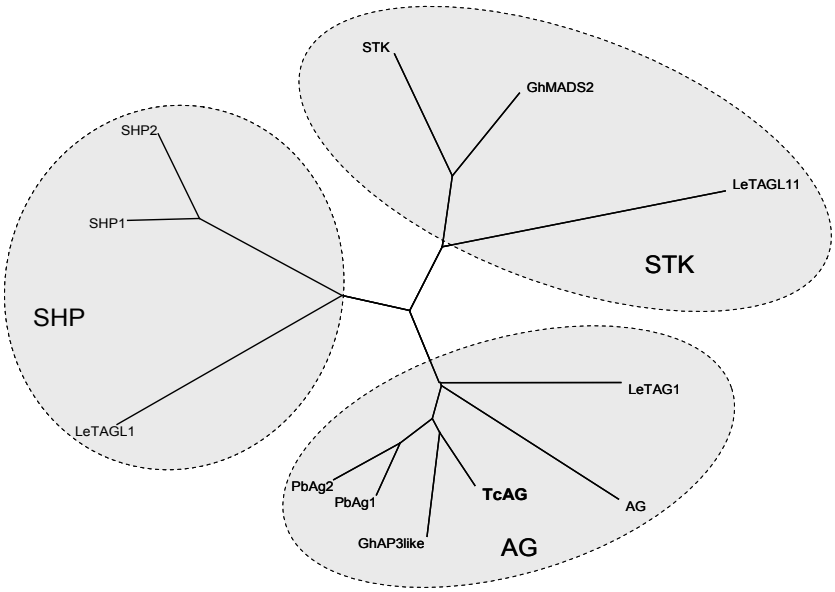


Fig. 3. Phylogenetic tree of TcAG and related MADS-box proteins, indicating the subclades represented by AG, STK and SHP from *Arabidopsis*. Le: *Lycopersicon esculentum* (tomato); Gh: *Gossypium hirsutum* (cotton); Pb: *Populus balsimefera* subsp. *trichocarpa*.

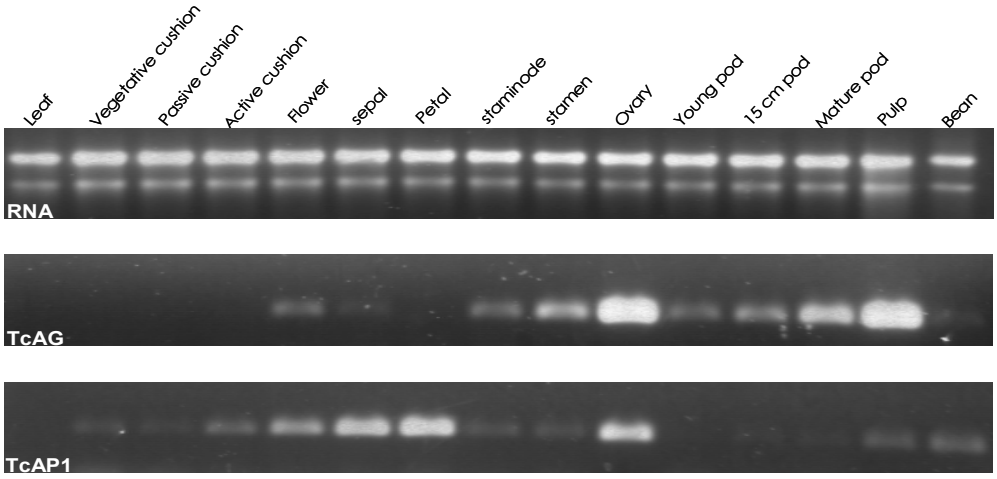


Fig. 4. RT-PCR products from RNA of various vegetative tissues (leaf, and vegetative shoot producing stem cushion), passive flowering cushions, actively flowering cushions, flower buds, floral organs (sepals, petals, staminodes, ovaries) and three stages of the developing inner pod wall. Upper panel shows total RNA used for the reverse transcription. Middle panel*: TcAG*; Lower panel: *TcAP1*.

suggested that this gene is involved in the development of reproductive floral organs since it is expressed in carpels and stamens, while expression is almost undetectable in sepals and petals of the cocoa flower.

Isolation and sequence analysis of *TcAP1* cDNA from cocoa

In order to isolate the full length *AP1* homologous cDNA from cocoa, a forward degenerate primer was designed based on the conserved N-terminus found in an alignment of AP1 proteins from 14 different species. Together with an oligo(dT)$_{18}$ as 3' primer these were used in an RT-PCR on first strand cDNA derived from cocoa flower bud mRNA. PCR products between 750 and 900 bp were cloned in the pGEM-T Easy vector and 4 products were sequenced and analyzed for homology with sequences in Genbank. The resulting protein had as its closest homologue in Genbank (85 % amino acid identity) the grapevine, *Vitis vinifera* AP1-like protein (AAT07447), followed by *Populus balsamifera* subsp. *trichocarpa* PTAP1-1 (AY615964), Quaking aspen, *Populus tremuloides* PTM2 (AAF12700). *Betula pendula* MADS3 protein (CAA67967). An amino acid alignment of TcAP1 protein with its most closely related relatives as well as with the more distant *Arabidopsis* AP1 (CAA78909) and tomato LeMADS-MC (AAM15774) proteins is shown in Fig. 5. Also here are clearly recognizable and indicated the strongly conserved MADS-box and somewhat lesser conserved K-box, while the C-terminus is the least conserved portion of the AP1 homologues.

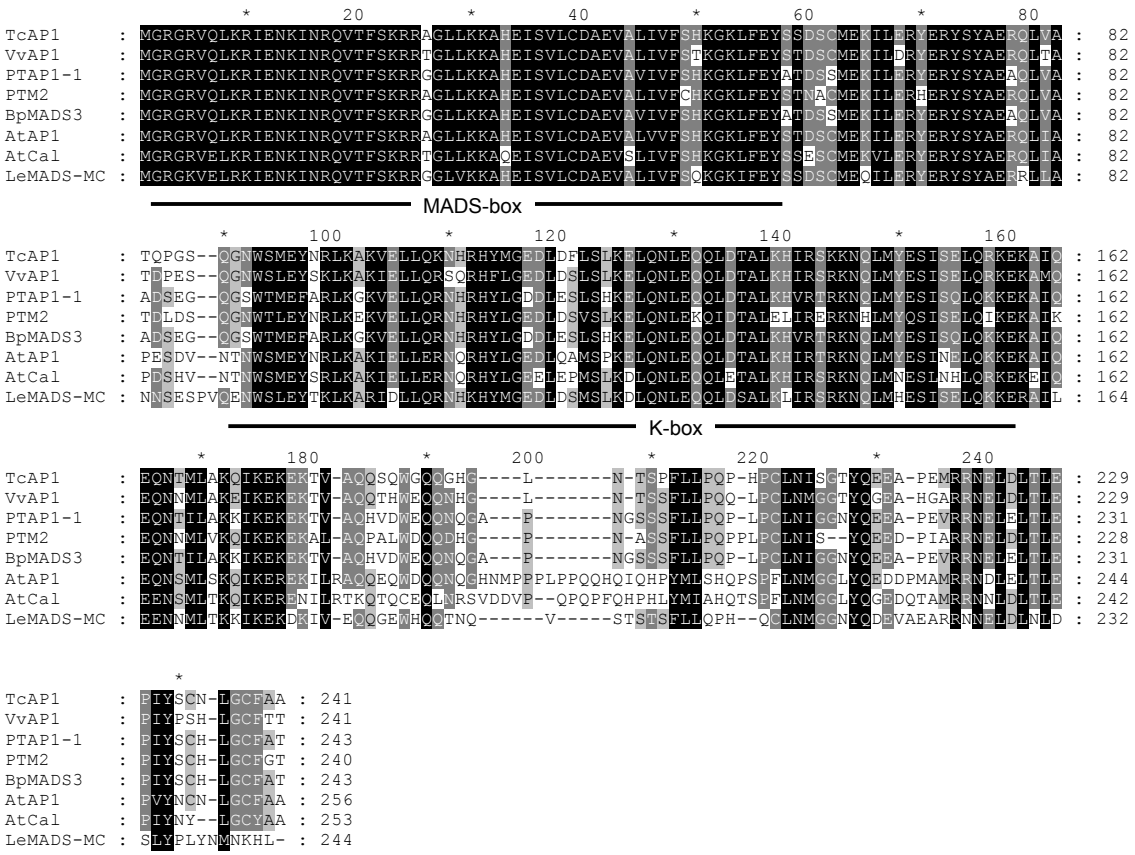


Fig. 5. Amino acid alignment of TcAP1 and its closest homologues (for accession numbers, see text). The conserved MADS-box and K-box are underlined. Vv, *Vitis vinefera* (grapevine); Ptm, *Populus tremoloidus*; (quaking aspen); Bp, *Betula pendula* (European white birch); Le, *Lycopersicon esculentum* (tomato):

FBP29 protein (for which no *Arabidopsis* ortholog has been identified so far) is also part of a subfamily of the MADS-box proteins, which likely represent a monophyletic clade, and whose members have partially redundant functions (4). In order to more accurately position *TcAP1* in this subfamily, we created a phylogenetic tree of a selection of AP1/CAL, FUL and FBP29 protein homologues from selected species (Fig. 6). As can be seen in this figure, *TcAP1* is firmly embedded in the AP1/CAL-subclade, closest to a birch gene (BpMADS3), while clearly more remote from the FUL and FBP29 subclades.

TcAP1 expression pattern

To clarify the expression pattern of the *TcAP1* gene, gene specific RT-PCR was performed using RNA prepared from various tissues as described above for *TcAG.* The results of the RT-PCR (Fig. 4, lower panel) revealed that *TcAP1* is particularly expressed at high levels in the.

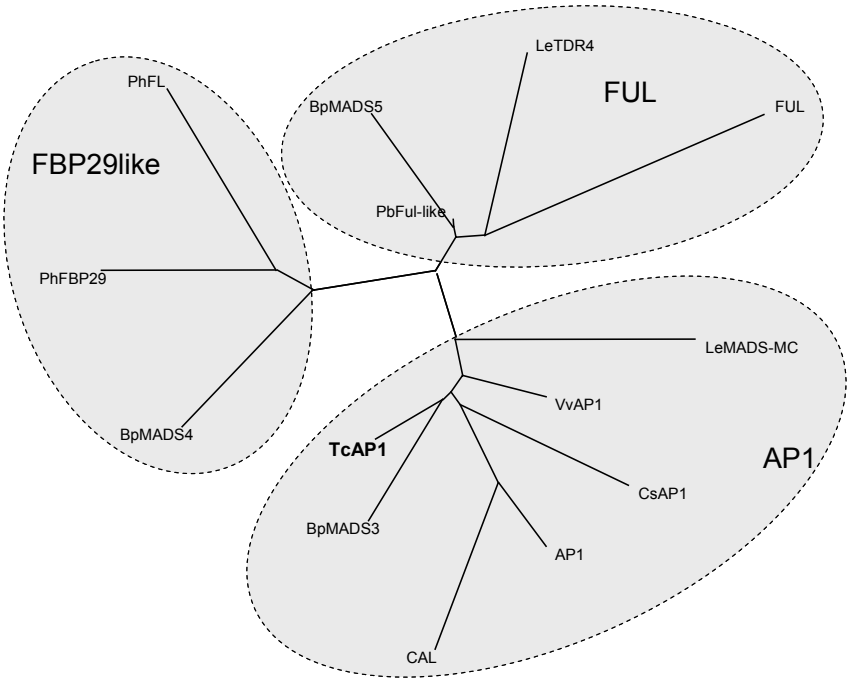


Fig. 6. Phylogenetic tree of TcAP1 and related MADS-box proteins, indicating the subclades represented by AP1, and FUL from *Arabidopsis*, and by FBP29 from *Petunia hybrida* (Ph); Cs, *Cucumis sativus* (cucumber)

staminodes, stamens and ovaries, an expression pattern which, at least in the flowers, is more or less complementary to that of *TcAG*. Moreover, in contrast to *TcAG*, expression was low but detectable in vegetative shoot producing and passive flowering cushions, and increased in actively flowering cushions. Expression in pod tissues was low, but detectable in some samples

Discussion

In this manuscript, we describe for the first time the cloning and characterization of two MADS-box type transcription factor genes involved in flower development of a tropical tree species, cocoa (*Theobroma cocoa* L.). In several aspects regulation of the onset of flowering as well as of the floral architecture in cocoa are substantially different from other plant species, in which the role of MADS-box genes during flowering was studied so far. Therefore, the study of the role of MADS-box genes during flowering of cocoa may yield important insights in the mechanisms leading to the evolutionary diversity of floral architecture and the regulation of the onset of flowering.

Based on the sequence of a cDNA-fragment from a library of subtracted pod wall-specific cDNA's, we amplified and cloned a full-length cDNA representing an *AGAMOUS* homologue (*TcAG*) from flowers of cocoa. Phylogenetic analysis of this and other AG protein sequences (Fig. 3) placed TcAG clearly in the AGAMOUS clade rather then in one of the two closely related clades represented by the *Arabidopsis* STK and SHP proteins, respectively. Identification of *TcAG* as a true ortholog of *AG* was further supported by its expression pattern in the flower. *TcAG* was typically expressed in the so-called C-domain of the flower: the inner whorls containing the staminodes, stamens, and the ovary. The occurrence of highly specialized staminodes is an interesting aspect of cocoa flower architecture. Staminodes are defined as stamen having lost the function of pollen production in evolution and are, although relatively uncommon, widely distributed taxonomically (26). Staminodes may occur simply as reduced stamen or stamen remnants, but often they have acquired during evolution new functions, such as in attracting pollinators or in regulating pollination (26). In cocoa they are likely to function in the prevention of self-pollination of the flower by typical dipteran pollinators (small midges). After opening the flower has a male phase, in which the staminodes surround the style and protect it against contact with the pollinators. During the following female phase the staminodes bend outwards and the midges can land on them, contacting the style (29). The staminodes of the cocoa flower are likely to have developed from the outer of two whorls of stamens. The expression of *TcAG* in the staminodes as well as in the stamens supports their common ontology.

Expression of *TcAG* was also found in the pod wall throughout its development, as well as, to a high level, in the pulp surrounding the beans. The pulp is probably derived from the seed testa, although the published evidence for this is scarce (18). AG in *Arabidopsis* as well as TAG1 in tomato are expressed in ovule integuments (7, 22), but little is known of their expression in later stages. Expression of *TAG1* in developing tomato fruit occurs particularly in the turning and red ripe stages (3), but no differentiation was made between seed and pericarp, whereas in the initial stages after fertilization *TAG1* expression in carpels diminished (25). *AG* in *Arabidopsis* is expressed during the early stages of silique development but drops off in the last stages (11). Thus it is likely that AG orthologues are expressed in the pericarp during some stages of fruit development in *Arabidopsis* and tomato as well as in cocoa, and its precise tissue localization of expression and its role in fruit development will be interesting to study. In grapevine, an AG homologue is also expressed in the skin of the berry during development up to the ripening stage, when it drops off (6).

Using primers based on conserved regions of several *AP1* orthologs, we amplified a cocoa homologue (*TcAP1*) of the Arabidopsis *APETALA1* gene. Phylogenetic analysis of this and other AP1 protein sequences (Fig. 3) placed *TcAP1* clearly in the *AP1* clade rather then in one of the two closely related clades represented by the *Arabidopsis* FUL and *Petunia* FBP29 proteins, respectively (Fig. 6). The *TcAP1* identity as a true *AP1* orthologue is further supported by its expression pattern in the flower, where it is expressed in the typical A-domain, the two outer whorls of sepals and petals, and much less so in the inner whorls, although expression was detected in the ovary. As for *TcAG*, further support for the identity of *TcAP1* as a true orthologue of its *Arabidopsis* counterpart may come from ectopic expression of the genes in *Arabidopsis*. The importance of the MADS-box genes in determining floral organ identity is reflected by their strong evolutionary conservation of function and of interactions with other transcription factors of the MADS family (14) and ectopic expression of orthologues from other species in *Arabidopsis* will often result in producing a similar phenotype as ectopic expression of the *Arabidopsis* gene will (5, 15, 20). In the case of *AP1*, this results in early flowering, reflecting the role of *AP1* in co-determining flowering time (17).

Cocoa is unique in its habit of flowering among the plants studied so far with respect to the molecular mechanisms regulating flowering. Annual plants like *Arabidopsis*, where the vegetative meristems are irreversibly converted to inflorescence meristems at the onset of flowering, are probably a poor model to study flowering in perennial plants. It would be interesting to determine to what extent cocoa conforms to the flowering pattern of other perennials such as the more studied *Populus* (30). In cocoa, flower buds initiate from meristems located in the axils of already shed leaves on the stem and older branches (16). These meristems can produce flowers for several years, eventually forming flowering cushions, but often have intermittent flowering activity regulated by environmental factors and possibly plant hormones. It is not known at what developmental stage flowering is arrested in resting flowering cushions: before formation of inflorescence meristems, at floral meristem formation or in the early flower bud stage. Here we report expression of *TcAP1* in all tested flower cushions, including those producing vegetative shoots, but not in leaves. These results do suggest that *TcAP1* may be involved in co-regulating flowering in cocoa, at a very early stage of inflorescence development. More detailed study of developmental processes within flowering cushions and characterization of other regulators of flowering time

such as an orthologue of *Arabidopsis LEAFY* (27) from cocoa may shed more light on the regulation of flowering in cocoa in the future.

Literature cited

1. **Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25:**3389-3402.
2. **Asif, M. H., P. Dhawan, and P. Nath.** 2000. A simple procedure for the isolation of high quality RNA from banana fruit tissue. Plant Mol. Biol. Rep. **18:**1 -7.
3. **Bartley, G. E., and B. K. Ishida.** 2002. Digital fruit ripening: Data mining in the TIGR tomato gene index. Plant Mol. Biol. Rep. **20:**115-130.
4. **Becker, A., and G. Theissen.** 2003. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol. Phylogen. Evol. **29:**464-489.
5. **Benedito, V. A., P. B. Visser, J. M. van Tuyl, G. C. Angenent, S. C. de Vries, and F. A. Krens.** 2004. Ectopic expression of *LLAG1*, an *AGAMOUS* homologue from lily (*Lilium longiflorum* Thunb.) causes floral homeotic modifications in *Arabidopsis.* J. Exp. Bot. **55:**1391-9.
6. **Boss, P. K., M. Vivier, S. Matsumoto, I. B. Dry, and M. R. Thomas.** 2001. A cDNA from grapevine (*Vitis vinifera* L.), which shows homology to *AGAMOUS* and *SHATTERPROOF*, is not only expressed in flowers but also throughout berry development. Plant Mol. Biol. **45:**541-53.
7. **Bowman, J. L., G. N. Drews, and E. M. Meyerowitz.** 1991. Expression of the *Arabidopsis* floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development. Plant Cell. **3:**749-58.
8. **Clough, S. J., and Bent, A.F.** 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:**735-43.
9. **Coen, E. S., and E. M. Meyerowitz.** 1991. The war of the whorls: genetic interactions controlling flower development. Nature **353:**31-7.
10. **Davies, B., P. Motte, E. Keck, H. Saedler, H. Sommer, and Z. Schwarz-Sommer.** 1999. *PLENA* and *FARINELLI*: redundancy and regulatory interactions between two *Antirrhinum* MADS-box factors controlling flower development. EMBO J. **18:**4023-34.
11. **de Folter, S., J. Busscher, L. Colombo, A. Losa, and G. C. Angenent.** 2004. Transcript profiling of transcription factor genes during silique development in *Arabidopsis*. Plant Mol. Biol. **56:**351-66.
12. **Ferrandiz, C., Q. Gu, R. Martienssen, and M. F. Yanofsky.** 2000. Redundant regulation of meristem identity and plant architecture by *FRUITFULL, APETALA1* and *CAULIFLOWER*. Development **127:**725-734.
13. **Ferrario, S., R. G. Immink, and G. C. Angenent.** 2004. Conservation and diversity in flower land. Curr. Opin. Plant Biol. **7:**84-91.
14. **Immink, R. G. H., and G. C. Angenent.** 2002. Transcription factors do it together: The hows and whys of studying protein-protein interactions. Trends Plant Sci. **7:**531-534.

15. **Kotoda, N., M. Wada, S. Kusaba, Y. Kano Murakami, T. Masuda, and J. Soejima.** 2002. Overexpression of *MdMADS5*, an *APETALA1*-like gene of apple, causes early flowering in transgenic *Arabidopsis*. Plant Sci. **162:**679-687.

16. **Lent, R.** 1966. The origin of cauliflorous inflorescence of *Theobroma cacao*. Turrialba. **16:**352-358.

17. **Mandel, M. A., and M. F. Yanofsky.** 1995. A gene triggering flower formation in *Arabidopsis*. Nature **377:**522-524.

18. **Nichols, R.** 1964. Studies of fruit development of cacao (*Theobroma cacao*) in relation to cherelle wilt. I. Development of pericarp. Ann. Bot-London. **28:**619-634.

19. **Page, R. D.** 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl. Biosci. **12:**357-8.

20. **Pena, L., M. Martin Trillo, J. Juarez, J. A. Pina, L. Navarro, and J. M. Martinez Zapater.** 2001. Constitutive expression of Arabidopsis *LEAFY* or *APETALA1* genes in citrus reduces their generation time. Nat. Biotechnol. **19:**263-267.

21. **Pinyopich, A., G. S. Ditta, B. Savidge, S. J. Liljegren, E. Baumann, E. Wisman, and M. F. Yanofsky.** 2003. Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature **424:**85-88.

22. **Pnueli, L., D. Hareven, D. Rounsley Steven, F. Yanofsky Martin, and E. Lifschitz.** 1994. Isolation of the tomato AGAMOUS gene TAG1 and analysis of its homeotic role in transgenic plants. Plant Cell. **6:**163-173.

23. **Schwarz-Sommer, Z., P. Huijser, W. Nacken, H. Saedler, and H. Sommer.** 1990. Genetic control of flower development: homeotic genes of *Antirrhinum majus*. Science. **250:**931–936.

24. **Toxopeus, H.** 1987. Botany, types and populations. p. 11-37. *In* G.A.R. Wood, and R.A.Lass (eds), Cacao, vol. 4. Longman Scientific and Technical, Harlow.

25. **Victoria, B. M., B. Claudia, D. Cecilia, H. C. Mauricio, B. B. Silvana, V. M. Estela, and Z. Eduardo.** 2003. MADS-box genes expressed during tomato seed and fruit development. Plant Mol. Biol. **52:**801-815.

26. **Walker-Larsen, J., and L. D. Harder.** 2000. The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention. Am J Bot. **87:**1367-1384.

27. **Weigel, D., J. Alvarez, D. R. Smyth, M. F. Yanofsky, and E. M. Meyerowitz.** 1992. LEAFY controls floral meristem identity in *Arabidopsis*. Cell. **69:**843-59.

28. **Yanofsky, M. F., H. Ma, J. L. Bowman, G. N. Drews, K. A. Feldmann, and E. M. Meyerowitz.** 1990. The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. Nature **346:**35-9.

29. **Young, A. M., E. H. Erickson Jr., M. A. Strand, and B. J. Erickson.** 1987. Pollination biology of *Theobroma* and *Herrania* (Sterculiaceae). I. Floral biology. Insect science and its application. **8:**151-164.

30. **Yuceer, C., and D. S. Luthe.** 2003. Use of *Populus* as a model for perennial flowering. Flowering newsletter **36:**35-42.

Chapter 8

General Discussion

Introduction

In this thesis studies on a number of aspects of cocoa biotechnology are presented that should contribute to the advance of the general knowledge of the molecular biology of cocoa, more particularly in fruit and flower development, and to the production of disease or pest-resistant transgenic varieties.

The results of this thesis are briefly discussed and placed in context of the larger world research efforts in biotechnology and, in particular, recent developments in cocoa research.

Cocoa breeding

Cocoa breeding may be done for the purpose of increasing yield potential, resistance to diseases or pests, uniformity and stability, lower input, or improved quality (16). Despite its long history of cultivation, breeding of *T. cacao* has received relatively little attention and the production of new clones from any breeding technique has been declining in the past few decades (25). This may be partially due to the lack of good breeding and selection programs. Breeding programs for cocoa have to deal with the same problems as other fruit crops: long generation times, long juvenile phase without fruit production, and requirement for large planting areas. With interbreeding tree crops like cocoa, clonal selection may be the best way to exploit genetic variation. However, selection of clones for yield has proven to be rather unsuccessful because of the low predictive value of the parental clone yield for the yield of its progeny (25). Currently there are two major efforts to revive cocoa breeding by using molecular markers, one at CIRAD/BIOTROP in Montpellier (France), and one in the USDA-ARS Subtropical Horticulture Research Station in Miami (USA).

Molecular breeding of cocoa

Cocoa is a diploid crop (2n=20) and has a haploid genome size of approximately 0.43 pg (415,000 kbp), only three times as much as the small genome of *Arabidopsis thaliana* (18). A first genetic map was produced with RFLPs, isozyme markers and RAPDs (22) and later refined with resistance gene homologues (RGA's) (21), WRKY transcription factor homologues (5), and microsatellite markers (28). Recently a BAC (Bacterial Artificial Chromosome) library of cocoa was produced, which was used for physical mapping of RGA's (6). The molecular markers were used to map genome regions involved in agronomic traits (yield, vigor) or in *Phytophthora* resistance (7-11, 30). The latter studies have confirmed

that *Phytophthora* resistance is of a polygenic nature, as is resistance to witches' broom disease (29). Although resistance gene homologues are being studied, their role in defense against the major pathogens in cocoa is unclear (21, 23). Genetic resistance to cocoa pod borer is largely unknown. The thickness of the sclerotic mesocarp of the pod, functioning as a physical barrier to the boring larvae may play a role in quantitative resistance and this is being used as a criterion in clonal selection (3, 4).

Propagation of cocoa trees

Cocoa trees are currently still mainly propagated by seedlings, from seeds selected by farmers from their own material. Selected hybrid varieties are cultivated on less then 30% of the global area (16). The seeds used for propagation are selected to be uniform, of good quality, and from seed-derived clones. Propagation by seeds results in high levels of heterogeneity of the crop, genetic variation and low yielding trees. Generally this variation is considered acceptable for genetically uniform types, such as the self-compatible Amelonado, because of its cheapness and for lack of better alternatives (36).

For genetically more heterogeneous types such as the self-incompatible Trinitario, vegetative propagation is essential to produce true to type trees. Budding and the use of rooted cuttings are common practice throughout the cocoa-growing regions. In Indonesia, the budding system has been used for many years to produce a series of DR (Djati Roenggo) clones in Java. Seedlings of Forestero cocoa are grown as rootstock and then grafted in the field with bud wood of DR1, DR2, and DR38 clones. Propagation of cocoa trees by rooted cuttings uses orthotropic and plagiotropic shoots. Orthotropic (Chupon) shoots produce cocoa clones with the same morphology as the seed-derived trees, but only a small quantity can be produced. The plagiotropic (Fan) branches produce morphologies different from that of natural seedlings, with more fibrous root systems, exhibiting bushy growth, not performing well in the field and requiring extensive pruning.

Embryogenesis for in vitro propagation and for transformation

Traditional vegetative propagation techniques are considered rather inefficient and costly for many developing countries. To produce a large number of cocoa plants and rapidly generating asexually propagated uniform plants, *in vitro* propagation techniques form a promising alternative. *In vitro* propagation offers a system for the clonal production of orthotropic plants with normal dimorphic architecture and taproot formation. Production and testing of disease-

free material and possibly germplasm conservation as embryo could also contribute to further improvement, distribution and preservation of germplasm. Somatic embryogenesis as means for in vitro propagation of cocoa was first reported by Esan (15) working with immature zygotic embryo tissue explants, a method also deployed by us as described in **Chapter 6**. These methods, however, do not produce clonal progeny and genetic gain is lost. Subsequently a number of groups have worked on somatic embryogenesis (SE) from vegetative tissues with mixed results. A more efficient method, capable of inducing primary SE in floral explants of a wide variety of genotypes used thidiazuron and 2,4-dichlorophenoxyacetic acid in the medium as phytohormones (24). With the optimization of regeneration, this method proved to be capable of producing large numbers of offspring that reached maturity (32). In vitro propagation methods contain the risk of producing mutant regenerants (somaclonal variation) to the extent, for example by delaying flowering, that the method becomes no longer commercially viable. Whether this is the case for cocoa will probably be determined in the next few years. A further improvement came through the production of secondary somatic embryos from explants of primary somatic embryo cotyledons. As we observed in **Chapter 6** for Indonesian clones, secondary embryogenesis is easier to achieve and thus multiplies the number of progeny possible from one primary somatic embryo. Moreover, whereas primary somatic embryos arise from groups of cells, secondary embryos were shown to arise predominantly from a single cell (26). This will be important for the development of a reliable transformation system: secondary embryos formed in selective medium after infection with *A. tumefaciens* will be likely to have arisen from a single transformed cell and thus eventually yield fully transgenic plants as opposed to chimeras growing from a mixture of transformed and non-transformed cells.

Cocoa genomics and molecular biology

International efforts in Cocoa breeding are coordinated by the International Group for the Genetic Improvement of Cocoa (INGENIC). At a meeting in Miami, 2002, researchers from 10 countries, including research institutes and companies agreed to form the “International Working Group on Cocoa Genome Studies” with the main objective to identify and share molecular genetic information to improve cocoa varieties, of which results should be in the public domain. Main activities planned were the creation of EST libraries, micro-array

studies, increase of SSR markers, BAC mapping, genotype identification studies, creation and management of segregating populations and database management (1).

As with other molecular studies of cocoa, the knowledge of the transcriptome of cocoa, in the form of characterized EST (Expressed Sequence Tags) libraries and expression studies is still relatively small. Two EST libraries have been reported: a mixed library of leaf and bean cDNAs (20) and one from elicitor-treated leaves (34). Together with sporadic earlier characterized sequences, this brings the publicly accessible TIGR cocoa gene index to a size of 6097 deposited sequences in total, constituting 2539 unique sequences. Compared to other gene indices in the TIGR database (for example the tomato gene index: 164,108 and 31,838 respectively) this is still very small and probably represents only a few percent of all possible cocoa transcripts. There is so far only one example of gene expression studies using a microarray containing 1380 unigenes from the leaf + bean library (20). It is to be expected that a microarray based on also the second EST library (elicitor-induced transcripts) will be quite valuable for the study of cocoa-pathogen interactions. Furthermore some additional useful data may arise from molecular studies and gene discovery in cotton (currently there are 40,348 unique sequences in the TIGR cotton gene index). Of all (model) crops or plant species that have been characterized to some useful extent by EST sequencing, cotton (*Gossypium hirsutum*) is phylogenetically closest to *T. cacao*, both belonging to the order Malvaceae sensu latu (35). This is supported by our findings that for cocoa sequences found and characterized in this work (**Chapters 4, 5, 7**), of all found homologues the cotton homologue sequence (if available) is phylogenetically closest to the cocoa sequence.

One of the action points at the Miami meeting was the production of EST libraries from different tissues. Cocoa pods were not part of this list and thus our construction of a specific pod wall cDNA library (**Chapter 4** of this thesis) may well complement this work. Although with the current small number of EST sequences available any set of cocoa sequences will contain some novel ones, the use of cDNA libraries of yet unused organs like the pod (excluding the beans) will probably substantially increase the rate of gene discovery in cocoa.

Fruit and flower development

Productivity of cocoa can be limited by many factors, among others by the irregular timing of flowering and by the competition between vegetative growth, flowering and fruit

development. In most places flowers can be found all year round but flowering peaks occur, in which flowering progresses as a wave from the lower stem flower cushions upwards. Pod development likely influences the development of new flowers. In the phenomenon called cherelle wilt, probably similar to fruit drop in other species, the cherelles (small fruit) turn yellow, wilt, and turn a brown color yet stay on the tree (31). In periods of increased vegetative growth (flushing) cherelle wilt is increased, indicating that it may be caused by sink competition for photoassimilates between fruit and vegetative parts (33). Thus an increased understanding of the regulation of flowering and of fruit development may help to understand and alleviate some of these problems. Since due to the long generation time of cocoa, experimental verification of hypotheses will be difficult and time consuming, in **Chapters 4 and 7** of this thesis a first attempt was made to link molecular processes in pod and flower development, respectively, with the equivalent processes in the model species *Arabidopsis* or tomato. Although limited in its scope, in **Chapter 7** it was shown that MADS box genes, which we isolated, and which in other species are known to (co-) regulate inflorescence formation (*AP1*) and/or to determine flower organ identity (*AP1 and AG*) are expressed in a pattern similar to that in flowers of other plants such as *Arabidopsis*. These genes may well function similarly in cocoa and specifically the role of *TcAP1* in flowering induction may lead to interesting insights in regulation of flowering in cocoa. Interestingly, we found continued high expression of *TcAG* in pod walls throughout development, highlighting the fact that the role of *AG* in the stages after flower organ formation is still unclear. We found in both developing pod walls as well as in tomato pericarp high expression of a *Lea5* gene homologue (**Chapter 5**), for which the function is still unknown. Although this gene family and its (sometime) induction of expression by desiccation shares some aspects with other *Lea* gene families, the reason for its high expression in pericarp tissue and its function in fruit development is an intriguing mystery.

Genetic modification of cocoa

Transformation and regeneration of cocoa has only been achieved relatively recently (27) and thus genetic modification of cocoa is still in its infancy. Relatively few groups in the world are taking this approach, and progress is probably hampered by the current unpopularity of GM approaches to crop improvement. In **Chapter 6** of this thesis we describe the development of transformation and regeneration protocols, for a number of cocoa clones

including some typical Indonesian clones, which are partially based on the work of the Guiltinan lab. In our studies we were most successful in transforming and regenerating secondary embryos derived from zygotic embryos. When cross-pollination occurs in the field, the use of zygotic embryos carries the risk of heterozygosity of the embryo. Thus, some of the favorable and selected properties of the parental clone chosen for enhancement by genetic modification may be lost. For this reason we have also included in our studies embryogenesis from petals and staminodes of agriculturally well performing clones. The latter explant tissues had been proven to be promising starting points for embryogenesis and regeneration in earlier studies (24). Somatic embryogenesis from vegetative parts of plants *after* transformation will also be a valuable strategy to maintain the transgene in the same genetic background during propagation of solid, non-chimaeric transgenic plants and prevent segregation of the transgene as would occur with propagation by seed. The Guiltinan lab at Pennsylvania State University has developed a transformation and regeneration protocol using explants from primary somatic embryos derived from staminode explants, and has successfully transformed plants with an enhanced green fluorescent protein gene and with a cocoa chitinase encoding gene, under the control of an enhanced CAMV 35S promoter. In that study mature plants expressing green fluorescent protein were obtained and characterized (27). Plants containing the chitinase gene were produced with the purpose to obtain or increase resistance to fungal pathogens. Enzymatically active chitinase was expressed but the resistance phenotype of these plants has not been reported so far (2).

Transgenic crop plants

Since the initial relatively unnoticed introduction of a transgenic food product in 1995 (FlavrSavr tomato), the introduction of transgenic crops has been a road with obstacles, except in Northern America and Argentina. Despite the initial (commercial) failure of transgenic tomatoes with longer shelf life, since 1996 the global area of GM crops has grown steadily, to 81 million hectares in 2004 (19). The majority of these were herbicide resistant maize, canola, soybean and cotton (72%), with almost all the rest planted with Bt maize and cotton (19%) or cotton and maize with stacked herbicide and insect resistance genes. Major countries in the adoption of GM crops are the United States, Argentina, Canada, Brazil, China and Paraguay.

For a number of reasons, mostly as a result of consumer concerns and protest actions by non-governmental groups, the adoption of GM crops in Europe (in particular in the European Union) has been slow and came to a complete standstill during a de facto, and later European Commission imposed moratorium from 1998 to 2001. Following the revision of the Deliberate Release Directive regulating the release of GMOs into the environment (Directive 2001/18/EC, repealing Directive 90/220/EEC) the moratorium was officially lifted. However, several European countries still maintain the moratorium until the issues of traceability, labeling and co-existence with conventional and organic crops are resolved. Spain was in 2004 still the only EU country with a significant area of GM crops (Bt-maize).

Whereas in some countries the introduction of GM crops has carried on, a number of countries, especially in the developing world, have followed the example of the EU countries and taken a cautionary approach. A notable exception is China, where not only Bt-cotton has virtually replaced conventional cotton in many areas, but which is also developing many of its own varieties and strategies for GM crops, such as GM-rice.

Bt-crops

Crop plants expressing a *Bacillus thuringiensis* toxin for resistance to pest insects (Bt-crops) have been among the first to be commercialized since 1996. Although many plant species have been transformed with a Bt toxin encoding gene, commercialization has been limited to three species: cotton, potato, and maize. Whereas Bt-potato use has been discontinued, Bt-cotton and Bt-maize have been commercialized successfully in different regions of the world (19).

Concerns about the use of Bt-crops in particular have focused on two issues: non-target effects and resistance development in the insect. Although Bt toxins have a rather high level of specificity, there is concern that non-target organisms, especially insects and more particularly beneficial insects such as predators and parasitoids of pest insects, may be negatively effected by direct or indirect (through their prey) exposure to the toxin in the transgenic crops. Research on this issue has been reviewed elsewhere (12, 13). Although at present there is little evidence for undesirable non-target effects of Bt-crops in the field, obviously every new introduction, such as that of a Bt-cocoa, will have to be preceded by careful risk assessment and accompanied by post-release monitoring for any effects. In this thesis (**Chapter 5**), we explored a method to decrease the risk of non-target (including

chocolate consumers) effects by aiming at expression of Bt toxin specifically in the pod wall. This would be one strategy to reduce the risk of non-target effects, namely by reducing the chances of exposure to the toxin for non-target organisms. Although proposed for this purpose, tissue-specific expression has not been deployed in any of the commercialized Bt-crops so far.

A second concern about the use of Bt-crops is for the rapid development of resistance to a Bt-toxin in the targeted pest population. As for any insecticide, resistance alleles for Bt toxins are present in insect populations and can be selected for. Resistance to various Bt toxins and in different insect species has been selected, mostly in the laboratory, but in some cases in the field (with Bt sprays) (17). The fear is that through the increased deployment of Bt crops increased selection pressure will lead to rapid development of resistant populations. For the currently commercial Bt-crops a number of measures have been deployed to try to prevent or delay the development of resistance (reviewed in (12, 13)). Although the role of these measures is difficult to asses, since the introduction in 1996 there have not yet been any reports of insect resistance to Bt-crops. One proposed measure is to delay resistance development by the deployment of two or more resistance genes with different modes of action, where the likelihood of two resistance alleles occurring in the same insect is very small. In this thesis (**Chapter 2**) we not only identified Bt proteins that are active against Cocoa pod borer and thus may be used in Bt-cocoa for CPB-resistance, but we also identified different Cry proteins with activity, some of which are likely to recognize different receptors in the CPB midgut and thus have a lower chance of inducing cross-resistance in the insect. Although the latter fact awaits experimental confirmation, this suggests that stacking these genes, like for other *cry* genes in the newest generation of Bt-cotton (BollGardII) is a useful option for resistance management. Other options would be the installment of refugia, areas with non-Bt cocoa trees to act as a source for unselected insects to breed with any arising resistant insects and to so dilute and prevent accumulation of resistance alleles in the insect population. Yet another possibility would be combination of resistance through a *cry* gene with some kind of natural (possibly quantitative) resistance, such as structural resistance of the pod wall to boring in varieties with a thicker sclerotic pod wall layer (see above). Thus, genetic modification for insect resistance does not preclude the use of more traditional breeding and selection techniques, but rather can add another layer of resistance to that selected by breeding.

Almost all successful reports of genetic modification of plants for insect resistance using a Bt toxin encoding gene have made use of genes with modified coding regions in order to adapt the originally bacterial gene for higher expression in plants (14). In **Chapter 3** of this thesis we describe our attempts to modify the open reading frame of the selected SN19 gene for higher expression in plants, and characterization of the expression of the modified gene versions in transgenic *Arabidopsis thaliana.* A comparison was made with an earlier version of a modified SN19 gene containing synthetic gene fragments obtained from a propriety synthetic *cry1Ba* gene. Although our own modifications substantially improved expression and resistance over a non-modified version, expression levels did not yet match the level obtained with the earlier version. Since the synthetic parts of the earlier version were optimized for expression in a monocotyledonous plant (maize) and since codon bias in general differs between monocot and dicotyledonous plants, it is expected that creation of a synthetic gene optimized for dicot plants, which is currently being undertaken, will raise expression in *Arabidopsis* and eventually in cocoa above the level of the earlier versions of SN19. In addition we have shown that expression of SN19 also confers resistance to *Plutella xylostella* (diamondback moth) and *Pieris rapae* (large cabbage white) larvae in *Arabidopsis*.

Outlook for transgenic cocoa in Indonesia

From a socio-economic point of view, cocoa has been an important commodity to Indonesia, the second largest cocoa bean producing country in the world. Data reported for 2003 showed that Indonesia exported 450,000 ton of cocoa beans with a value of between US$1.410 - US$1.420 per ton. More than 70% of the total area of cocoa plantations in Indonesia is owned and managed by small planters. In cocoa plantations managed by big companies, each hectare requires usually 1 employee to do routine work on the crop. Thus, with a total area of 0.6 million hectares, the Indonesian cocoa plantations hire at least 600,000 people. However, the total area of cocoa plantations is slowly diminishing. Several companies in some provinces of Indonesia converted cocoa plantations to plantations of other estate crops such as oil palm, and some are waiting to do so. The reason is that the cocoa plantations are becoming less productive, and therefore are generating less profit. The major factor contributing to this production loss is the cocoa pod borer (CPB). CPB is a serious pest and so far there is no single cost effective and environmentally safe way to control this pest. Residing inside the pod during most of the life cycle, the larvae are untouchable by chemical sprays. To solve this

problem, planting materials genetically tolerant to CPB are preferable. These may be obtained by classical breeding as well as by genetic modification, and a combination of resistance mechanisms is preferable for resistance management (see above). Genetic engineering of cocoa for resistance to CPB may have two clear benefits: the offering of an effective solution for the CPB problem and the building of Indonesia's capacity in modern agricultural biotechnology.

The last determining factor in the successful development of marketable transgenic crops obviously is public acceptance. Some transgenic plants, cotton, soybean and maize have been planted in Indonesia, of which only Bt-cotton on a commercial scale. In field trials in South Sulawesi, this cotton was claimed to increase productivity between 3 to 4 fold compared to local cotton. Complying with the Indonesian Biosafety Guideline, this field trial was preceded by thorough technical as well as regulatory examinations by the National Commission for Biosafety. Nevertheless, public acceptance in Indonesia is still low. Some people believe that public acceptance to transgenic crops particularly in Indonesia is a matter of public relations instead of technical aspects. One aspect of this may be to have risk analysis be accompanied by an analysis of the potential benefits of introduction of the crop. Whereas for many European consumers the benefits of transgenic crops may be less obvious or are ignored, transgenic CPB-resistant Indonesian cocoa may one day make a big difference for many Indonesians.

Literature cited

1. **Anonymous.** 2002. Meeting Summary. USDA-ARS Cacao Research Meeting, Miami, Fl.
2. **Antúnez de Mayolo, G.** 2003. PhD Thesis. Pennsylvania State University.
3. **Azhar, I.** 1990. Studies on the sclerotic layer hardness of cocoa pods. MARDI Res. J. **18:**63-70.
4. **Azhar, I., G. E. Long, and M. J. Musa.** 1995. Qualitative and multivariate analyses of clonal cocoa resistance to cocoa pod borer (Lepidoptera: Gracillariidae). Planter **71:**307-321.
5. **Borrone, J. W., D. N. Kuhn, and R. J. Schnell.** 2004. Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. Theor. Appl. Genet. **109:**495-507.
6. **Clement, D., C. Lanaud, X. Sabau, O. Fouet, L. Le Cunff, E. Ruiz, A. M. Risterucci, J. C. Glaszmann, and P. Piffanelli.** 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.
7. **Clement, D., A. M. Risterucci, J. C. Motamayor, N. G. J, and C. Lanaud.** 2003. Mapping QTL for yield components, vigor, and resistance to *Phytophthora palmivora* in *Theobroma cacao* L. Genome **46:**204-212.

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

9. **Crouzillat, D., E. Lerceteau, V. Petiard, J. Morera, H. Rodriguez, D. Walker, W. Phillips, C. Ronning, R. Schnell, J. Osei, and P. Fritz.** 1996. *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor. Appl. Genet. **93:**1-2.

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

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

12. **de Maagd, R. A.** 2004. Biotechnology meets ecology. p. 117-131. *In* J.-P. Nap, A. Atanassov, and W. J. Stiekema (eds), Genomics for Biosafety in Plant Biotechnology, vol. 359. IOS Press, Amsterdam.

13. **de Maagd, R. A.** 2002. Environmental impact of insect-resistant crop plants expressing a Bt-toxin. II. Non-target effects (updated) and Bt-resistance in target insects. Plant Research International Note 218.

14. **de Maagd, R. A., D. Bosch, and W. J. Stiekema.** 1999. *Bacillus thuringiensis* toxin mediated insect resistance in plants. Trends Plant Sci. **4:**9-13.

15. **Esan, E.** 1977. Tissue culture studies on cacao (*Theobroma cacao* L.). A supplementation of current research. 5th Int. Cacao Res. Conf. Cacao Res. Inst. Nigeria, Ibadan.

16. **Eskes, B.** 2004. Synthesis and conclusions of the INGENIC Workshop on "Cocoa breeding for improved production systems". INGENIC Newslett. **9:**2-7.

17. **Ferré, J., and J. Van Rie.** 2002. Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. Ann. Rev. Entomol. **47:**501-533.

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

19. **James, C.** 2004. Global status of commercialized biotech/GM crops: 2004. ISAAA.

20. **Jones, P. G., D. Allaway, D. M. Gilmour, C. Harris, D. Rankin, E. R. Retzel, and C. A. Jones.** 2002. Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta. **216:**255-264.

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

22. **Lanaud, C., A. M. Risterucci, N. G. AKJ, D. Clement, M. H. Flament, V. Laurent, and M. Falque.** 1995. A genetic linkage map of *Theobroma cacao* L. Theor. Appl. Genet. **91:**6-7.

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

24. **Li, Z. J., A. Traore, S. Maximova, and M. J. Guiltinan.** 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.

25. **Lockwood, R.** 2003. Who needs clothing? INGENIC Newslett. **8:**2-5.

26. **Maximova, S., L. Alemanno, A. Young, N. Ferriere, A. Traore, and M. Guiltinan.** 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.

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

28. **Pugh, T., O. Fouet, A. M. Risterucci, P. Brottier, M. Abouladze, C. Deletrez, B. Courtois, D. Clement, P. Larmande, N. G. Jak, and C. Lanaud.** 2004. A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. Theor. Appl. Genet. **108:**1151-1161.

29. **Queiroz, V., C. Guimaràes, D. Anhert, I. Schuster, R. Daher, M. Pereira, V. Miranda, L. Loguercia, E. Barros, and M. Moreira.** 2003. Idenfication of a major QTL in cocoa (*Theobroma cacao* L.) associated with resistance to witches broom disease. Plant Breeding. **122:**268-272.

30. **Risterucci, A. M., D. Paulin, M. Ducamp, J. A. K. N'Goran, and C. Lanaud.** 2003. Identification of QTLs related to cocoa resistance to three species of *Phytophthora*. Theor. Appl. Genet. **108:**168-174.

31. **Toxopeus, H.** 1987. Botany, types and populations. p. 11-37. *In* G.A.R. Wood, and R.A. Lass (eds), Cacao, vol. 4. Longman Scientific and Technical, Harlow.

32. **Traore, A., S. N. Maximova, and M. J. Guiltinan.** 2003. Micropropagation of *Theobroma cacao* L. using somatic embryo-derived plants. In Vitro Cell. Dev. Biol.-Plant. **39:**332-337.

33. **Valle, R., A. De Almeida, and R. De Leite.** 1990. Energy costs of flowering, fruiting, and cherelle wilt in cacao. Tree Physiol. **6:**329-336.

34. **Verica, J. A., S. N. Maximova, M. D. Strem, J. E. Carlson, B. A. Bailey, and M. J. Guiltinan.** 2004. Isolation of ESTs from cacao (*Theobroma cacao* L.) leaves treated with inducers of the defense response. Plant Cell Rep. **23:**404-413.

35. **Whitlock, B., C. Bayer, and D. Baum.** 2001. Phylogenetic relationship and floral evolution of the Byttnerioideae ("Sterculiaceae" or "Malvaceae") based on sequence of the chloroplast *ndhF*. Sys. Botan. **26:**420-437.

36. **Wood, G.** 1987. Propagation. p. 93-118. *In* G.A.R. Wood, and R.A. Lass (eds), Cacao, vol. 4. Longman Scientific and Technical, Harlow.

Summary

The cocoa tree (*Theobroma cacao* L.) produces the beans that are the source of cacao, the basis for chocolate production, and it is an important commodity crop for many tropical countries in South America, West Africa, and Southeast Asia. Up to 80% of the world cacao production may be lost due to fungal and viral diseases or insect pests. In Indonesia, the second largest cacao producing country in the world, as well as in the rest of Southeast Asia, the major insect pest is the Cocoa pod borer (CPB, *Conopomorpha cramerella*). The larvae of this small moth bore through the wall of the fruits (pods), feed on the inside, and interfere with bean development, leading to crop losses of up to 80%. There is no single effective method for the control of CPB and progressing infestation of plantations is endangering the future of the cocoa culture in Indonesia.

This thesis describes the first steps in a biotechnological approach to the control of CPB through the production of transgenic, CPB-resistant cocoa trees that produce an insecticidal toxin from *Bacillus thuringiensis* in their pod wall. **Chapter 1**, the **Introduction** gives an overview of the origins and botany of cocoa, of chocolate production and of the main diseases and pests threatening cacao production all over the world. Additionally, it describes the nature and long-time safe agricultural use of the insecticidal activity of *Bacillus thuringiensis* (Bt) and its toxins, how the toxins work, and how these toxins are currently being applied in transgenic crops.

Bt toxins are a large family of proteins, each protein being only active against a few insect species. Thus, for CPB-resistant cocoa, active toxins had to be selected. **Chapter 2** describes the selection of toxins with activity against CPB from a repertoire of Bt toxins. Of 12 Bt proteins (Cry proteins) tested, 5 appeared more active then others. Toxicity of three of these toxins was determined more accurately. Since all three were active and yet relatively different in amino acid sequence, the result gave a good perspective for sustainable CPB control. One active toxin, SN19, was selected for further studies.

The genes encoding Bt toxins are of bacterial origin and, therefore, usually require modification of the coding sequence to acquire sufficiently high expression in a transgenic plant. **Chapter 3** describes the modification of the gene encoding SN19 for improved expression. The effects on expression of a number of modifications was tested and compared to the expression of an earlier version of the gene containing synthetic parts. Due to the long

generation time of cocoa, transgenic *Arabidopsis thaliana* was used as a model plant to test new gene constructs in combination with a green tissue-specific promoter of the Rubisco small subunit gene from Chrysanthemum. The possible target spectrum of SN19 was further extended with *Pieris rapae* and *Plutella xylostella*, pests of cruciferous crops. The modifications to SN19 resulted in a marked improvement in expression level and insect resistance, but did not quite match the levels of the earlier used version.

Tissue- or organ-specific expression of a transgene requires the use of a proper promoter that has tissue- or organ-specific activity, to control the expression. In order to isolate a pod wall-specific promoter to drive Bt toxin expression, in **Chapter 4** two approaches were used to isolate pod wall or pulp-specific gene transcripts. Random cDNA libraries of the pod wall and of the pulp were screened in a reverse Northern blot to identify the more highly expressed genes. Subsequently Northern blotting with RNA of various tissues was used to assess tissue-specificity. Although a number of genes, highly expressed in pod wall sand/or pulp, were identified, many were also expressed in leaves or even beans. The second approach, the production of a subtractive library of pod wall cDNA-fragments subtracted with cDNA from beans and leaves yielded a number of interesting, more pod wall-specific cDNAs. One of these, representing a gene homologous to a cotton *Lea5* gene, was expressed in the pod wall starting from an early developmental stage and was used for further study and for the isolation of its promoter, as described in **Chapter 5**. Cloning by genome walking was used to isolate a 3411 base pair genomic fragment of cocoa encompassing the coding region and the 5' flanking region, containing the promoter. The extent of the promoter and its regulation were tested by transforming *Arabidopsis* with a construct containing different lengths of the 5' flanking region fused to the GUS reporter gene. In this way, the minimal fragment (spanning 451 bp relative to the start codon) of the promoter necessary for activity in *Arabidopsis* was determined. However, this promoter in *Arabidopsis* was also active in leaves and roots, indicating that either not all regulatory elements for fruit-specific expression were present or active in this plant and more data are needed. Interestingly, the *TcLea5* product belongs to a family of proteins (Late embryogenesis associated) that is usually highly expressed in desiccating embryos and in some cases desiccating vegetative tissues. However, as shown in Chapter 5, in cocoa as well as in tomato the *Lea5* homologue is highly expressed in the fruit wall, where its function is yet unclear.

Producing transgenic insect-resistant cocoa trees requires an effective transformation and regeneration protocol for cocoa. In the study described in **Chapter 6** a number of

protocols and different tissues was tested for the formation of somatic embryos, the starting material for transformation and regeneration. Somatic embryos originating from either zygotic embryos or from staminodes could be transformed by *Agrobacterium tumefaciens* carrying a CaMV 35S promoter-*gfp* (green fluorescent protein) construct, followed by secondary embryogenesis and plant regeneration. Solid *gfp*-expressing plantlets were obtained from zygotic embryo-derived somatic embryos and are expected from staminode-derived somatic embryos. Preliminary results with *SN19* expression under control of the CaMV 35S promoter in transgenic cocoa callus indicate that our *SN19* gene constructs are capable of expressing this gene in cocoa tissues.

In **Chapter 7** a preliminary characterization of two cocoa genes encoding transcription factors of the MADS-box class, probably involved in regulation of flowering time and in determining flower architecture, is presented. A cocoa homologue of the *Arabidopsis* gene *APETALA1* (*TcAP1*) may well be involved in co-regulation of inflorescence initiation, as well as in determining the identity of the two outer whorls of floral organs, i.e. sepals and petals. As expected for its proposed role and similar to the situation in other flowering plant species, *TcAP1* was only expressed in inflorescences and within the flower predominantly expressed in sepals and petals, though also at a lower level in other floral organs. A cocoa homologue of the *Arabidopsis* gene *AGAMOUS* (*TcAG*) may well be involved in determining the identity of the inner whorls of floral organs, staminodes, stamen and ovaries. As expected for its proposed role and similar to the situation in other flowering plant species, *TcAG* was only expressed in flowers and within the flower predominantly in staminodes, stamen and ovaries. We also found continuing expression of *TcAG* in the pod wall throughout its development.

Finally, **Chapter 8** –the **General Discussion**- discusses the results presented in this thesis and places them in the context of an overview of recent developments in biotechnology and genomics of cocoa. Although progress in cocoa breeding has been slow and the adoption of transgenic crops in many countries is lagging, but because of the economic importance of cocoa it is expected that one day, based on a positive outcome of a risk/benefit-analysis, CPB-resistant transgenic cocoa trees will be accepted and will make a big difference for many Indonesians.

Samenvatting

De cacaoboom (*Theobroma cacao* L.) maakt de bonen waaruit cacao, de basis voor chocoladeproductie, wordt gemaakt en is een belangrijk gewas voor veel tropische landen in Zuid-Amerika, West-Afrika en Zuidoost Azië. Tot 80% van de wereld cacaoproductie kan verloren gaan door ziekten, veroorzaakt door schimmels en virussen, of door insectenplagen. In Indonesië, de tweede cacaoproducent van de wereld, maar ook in de rest van Zuidoost Azië, is de cacao podboorder (CPB, *Conopomorpha cramerella*) de belangrijkste insectenplaag. De larven van deze kleine mot boren door de wand van de vrucht (pod), voeden zich met de inhoud en verstoren de ontwikkeling van de bonen, leidend tot wel 80% productieverlies. Er is geen enkelvoudige effectieve manier om CPB te bestrijden en de voortschrijdende invasie van plantages brengt de toekomst van de productie van cacao in Indonesië in gevaar.

Dit proefschrift beschrijft de eerste stappen op het pad van een biotechnologische benadering van CPB bestrijding door de productie van transgene, resistente cacaobomen, die een insectendodend eiwit van *Bacillus thuringiensis* in de vruchtwand aanmaken. **Hoofdstuk 1**, de **Inleiding** geeft een overzicht van de herkomst en botanische eigenschappen van cacao, van de chocoladeproductie en van de vele ziekten en plagen die de cacaoproductie wereldwijd bedreigen. Tevens beschrijft het de aard en langdurig veilig gebruik in de landbouw van de insectendodende activiteit van *Bacillus thuringiensis* (Bt) en zijn eiwitten, hoe die eiwitten werken, en hoe ze momenteel in transgene gewassen worden toegepast.

Bt toxinen vormen een grote familie van eiwitten, waarbij elk eiwit slechts actief is tegen een of enkele insectensoorten. Om die reden moesten eerst eiwitten met activiteit tegen CPB larven worden geselecteerd. **Hoofdstuk 2** beschrijft de selectie van eiwitten met activiteit tegen CPB larven uit een verzameling van 12 Bt toxines. Van de 12 Bt toxinen (Cry eiwitten) die getest werden, bleken er 5 actiever dan de anderen. De activiteit van drie van deze vijf werd meer nauwkeurig bepaald. Daar alle drie hoge activiteit hadden én ze aanzienlijk in aminozuurvolgorde verschilden, is er goede hoop voor duurzame bestrijding van CPB met behulp van genen die voor deze toxinen coderen. Een van de actieve eiwitten, SN19, werd gebruikt voor het verdere onderzoek.

De genen die coderen voor Bt toxinen zijn van bacteriële oorsprong en moeten daarom gewoonlijk veranderingen in het coderende gedeelte ondergaan om voldoende hoog tot

expressie te komen in transgene planten. **Hoofdstuk 3** beschrijft het aanbrengen van zulke veranderingen in het gen voor SN19 teneinde de expressie te verhogen. De effecten van een aantal veranderingen werden getest en vergeleken met een oudere versie van het gen, welke volledig synthetische delen bevat. Omdat het zo lang duurt om vruchten van transgene cacao te krijgen, werden transgene *Arabidopsis thaliana* (zandraket) planten gebruikt als modelplant voor het testen van nieuwe genconstructen, in een combinatie met een groene delen-specifieke promoter uit chrysant. Het aantal mogelijke doelwitplagen van SN19 werd verder uitgebreid met *Pieris rapae* (klein koolwitje) en *Plutella xylostella* (koolmotje), plagen van kruisbloemige gewassen. De veranderingen in het gen van SN19 leverden een aanmerkelijke verbetering in de expressie en insectenresistentie op, maar haalden nog niet het niveau van de eerder gebruikte versie van het SN19 gen.

Weefsel- of orgaanspecifieke expressie van een transgen is alleen mogelijk indien gereguleerd door een promoter die weefsel- of orgaanspecifieke activiteit vertoont. Teneinde een dergelijke vruchtwandspecifieke promoter voor expressie van een Bt toxinegen te verkrijgen, worden in **Hoofdstuk 4** twee benaderingen voor het isoleren van vruchtwand- of pulpspecifieke transcripten beschreven. “Random” cDNA banken van de vruchtwand en van de pulp werden getest in een omgekeerde Northern blot om zo de hoogst tot expressie komende genen te identificeren. Vervolgens werd Northern blotting met RNA van verschillende weefsels en organen gebruikt om de specificiteit van die expressie te bepalen. Hoewel een aantal genen met hoge expressie in de vruchtwand of pulp werd geïdentificeerd, vertoonden vele daarvan ook expressie in bladeren en bonen. De tweede benadering, het maken van een subtractiebank van vruchtwand cDNA-fragmenten met daaruit verwijderd het cDNA dat ook in bonen en bladeren voorkwam, leverde een aantal interessante, meer vruchtwandspecifieke cDNA’s op. Een van deze, homoloog met een *Lea5* gen van katoen, kwam hoog tot expressie in de vruchtwand vanaf een vroeg ontwikkelingsstadium en werd daarom voor verder onderzoek en voor isolatie van zijn promoter geselecteerd. **Hoofdstuk 5** beschrijft hoe door “genome walking” een fragment van het cacaogenoom ter grootte van 3411 baseparen, met daarop het volledig cacao *Lea5* gen inclusief promoter, werd gekloneerd. De minimale actieve lengte van die promoter, en de mate van regulatie werden bepaald door het transformeren van *Arabidopsis* met constructen waarin stukken van de promoter, met verschillende lengten, waren gefuseerd met het “reporter”gen coderend voor GUS. Op deze manier werd bepaald dat de minimale lengte van de promoter, benodigd voor activiteit in *Arabidopsis* 451 baseparen (gerekend vanaf het startcodon) was. Echter, in *Arabidopsis* was

de promoter ook actief in bladeren en wortels, wat erop wijst dat niet alle benodigde elementen voor vruchtwandspecifieke expressie aanwezig of actief waren in de *Arabidopsis* planten, en dat meer onderzoek nodig is. Het product van het cacaogen, *TcLea5*, maakt deel uit van een familie van eiwitten (Late embryogenesis associated – geassocieerd met de laatste stadia van de embryogenese) welke gewoonlijk hoog tot expressie komen in uitdrogende embryo's, en soms in uitdrogende vegetatieve delen van de plant. Echter, zoals in hoofdstuk 5 werd aangetoond, zowel in cacao als in tomaat komt het *Lea5* gen tot hoge expressie in de vruchtwand, waar de functie niet duidelijk is.

Voor de productie van transgene, insectenresistente cacaobomen is een effectief transformatie- en regeneratieprotocol nodig. **Hoofdstuk 6** beschrijft de ontwikkeling, het testen en de vergelijking van een aantal protocollen, met verschillende weefsels, voor de productie van somatische embryo's, het uitgangsmateriaal voor transformatie en regeneratie. Somatische embryo's voortkomend uit zygotische embryo's of uit staminoden konden worden getransformeerd door *Agrobacterium tumefaciens* met een CaMV 35S promoter-*gfp* ("green fluorescent protein" – groen fluorescerend eiwit) construct, gevolgd door secundaire embryogenese en plantregeneratie. Volledig GFP-positieve plantjes werden verkregen uit somatische embryo's, die ontstaan waren uit zygotische embryo's. Dezelfde GFP-positieve plantjes zijn in ontwikkeling uit staminodeëxplantaten-geproduceerde somatische embryo's. De eerste resultaten van expressie van *SN19* onder controle van de CaMV 35S promoter in transgeen cacao callus laten zien dat onze SN19 constructen in staat zijn het gen in cacaoweefsels tot expressie te brengen.

In **Hoofdstuk 7** wordt een eerste karakterisatie beschreven van twee cacaogenen die coderen voor transcriptiefactoren van de "MADS-box" klasse, en die waarschijnlijk betrokken zijn bij de regulatie van de bloeitijd en bij het bepalen van de bloembouw. Een cacao homoloog van het *Arabidopsis* gen *APETALA1* (*TcAP1*) kan betrokken zijn bij het meebepalen van het moment van vorming van de bloeiwijze, evenals bij het bepalen van de identiteit van de twee buitenste ringen van bloemorganen, de kelkbladen en de bloembladen. Als verwacht vanwege zijn voorspelde functie, en vergelijkbaar met de situatie in andere plantensoorten, kwam *TcAP1* alleen tot expressie in bloeiwijzen, en in de bloem vooral in kelk- en bloembladen, hoewel er ook expressie op een lager niveau in de andere bloemorganen werd gevonden. Een cacao homoloog van het *Arabidopsis* gen *AGAMOUS* (*TcAG*) is waarschijnlijk betrokken bij het bepalen van de identiteit van de binnenste ringen van bloemorganen, de staminoden, de meeldraden en het vruchtbeginsel. Als verwacht

vanwege zijn voorspelde functie en vergelijkbaar met de situatie in andere plantensoorten kwam *TcAG* alleen tot expressie in bloemen en daarin voornamelijk in staminoden, meeldraden en vruchtbeginsels. Ook werd aanhoudende expressie van *TcAG* in de vruchtwand gedurende de hele vruchtontwikkeling gevonden.

Tenslotte bespreekt **Hoofdstuk 8** –de Algemene Discussie- de resultaten in dit proefschrift en plaatst ze in de context van recente ontwikkelingen in de biotechnologie en genomica van cacao. Ondanks dat de voortgang in de cacaoveredeling langzaam is en de in gebruikneming van transgene gewassen in veel landen achter blijft, wordt vanwege het economische belang van cacao verwacht dat, volgend op een positieve balans van de geschatte risico's en verwachte voordelen, CPB-resistente transgene cacaobomen zullen worden geaccepteerd en grote invloed op het welzijn van veel Indonesiërs zullen hebben.

Ringkasan

Kakao (*Theobroma cacao*L) menghasilkan biji sebagai bahan dasar coklat dan merupakan komoditi penting bagi Amerika Selatan, Afrika Barat, dan Asia Tenggara. Lebih dari 80% produksi kakao dunia kehilangan akibat fungal, virus atau hama penyakit. Indonesia sebagai penghasil coklat nomer 2 di dunia, juga mempunyai hama utama seperti negara Asia tenggara lainnya adalah penggerek buah kakao (PBK), *Conophomorpha cramerella.* Larva dari PBK ini, menggerek dinding buah, dan tinggal di dalam buah serta menghambat perkembangan biji kakao, menyebabkan tanaman kakao kehilangan produksi lebih dari 80%. PBK sebagai ancaman yang berbahaya bagi tanaman kakao di Indonesia hingga kini belum ditemukan teknik yang effektif untuk mengontrol dan menghambat meluasnya ke wilayah Indonesia lainnya.

Thesis ini menguraikan tahap awal pendekatan secara bioteknologi untuk mengontrol PBK dengan menghasilkan tanaman transgenik kakao tahan PBK yang menghasilkan toksin dari *Bacillus thuringiensis* pada dinding buahnya. **Bab 1** menguraikan tentang pendahuluan yang memberikan gambaran tentang asal dan biologi tanaman kakao, produksi coklat, hama dan penyakit utama yang menyerang tanaman kakao dunia. Juga diuraikan tentang penggunaan insektisida *Bacillus thuringiensis (Bt)* dan toksin yang dihasilkannya, serta mekanisme kerjanya yang saat ini banyak digunakan pada tanaman transgenik.

Bt toksin adalah keluarga protein, dimana setiap protein hanya aktif pada beberapa insek spesies tertentu saja. Untuk mengetahui Bt tahan terhadap PBK, dilakukan seleksi dengan menguji dilapangan beberapa Bt toksin terhadap PBK. **Bab 2** menguraikan seleksi beberapa Bt toksin yang berasal dari group Bt yang aktif terhadap PBK . Dari 12 protein Bt (Cry protein) yang di gunakan, 5 diantaranya lebih aktif, dan 3 diantara toksin yang aktif tersebut, mempunyai sekuen asam amino yang tidak berbeda satu dengan lainnya, menunjukkan hasil yang positif untuk mengontrol PBK. SN19 sebagai salah satu toksin yang aktif dipilih untuk digunakan lebih lanjut.

Gen sebagai penanda toksin Bt berasal dari bakteri, oleh sebab itu biasanya diperlukan modifikasi dari “coding sequence” untuk menghasilkan ekspresi yang diharapkan pada tanaman transgenik. **Bab 3** menguraikan modifikasi yang dilakukan pada gen *SN19* untuk mendapatkan ekspresi yang diharapkan pada tanaman. Effek dari modifikasi di test, dan dilakukan perbandingan dengan modifikasi yang telah dilakukan

sebelumnya. Berhubung regenerasi tanaman kakao memerlukan waktu yang cukup panjang, maka *Arabidopsis thaliana* digunakan sebagai tanaman model, untuk menguji konstruk gen yang dihasilkan dengan menggunakan promoter "green tissue" dari Rubisco.small dari Chrysanthemum. *Pierris rapae* dan *Plutella xylostella* pest dari tanaman Cruciferous digunakan sebagai target insek.. Hasil menunjukkan bahwa ekspresi dari SN19 yang telah dimodifikasi, tidak lebih baik dari modifikasi yang telah dilakukan sebelumnya.

Untuk mengontrol ekspresi toksin pada jaringan atau organ spesifik diperlukan promoter yang mempunyai aktivitas spesifik pada jaringan atau organ spesifik. Untuk mengisolasi promoter spesifik dari dinding buah kakao yang mengekspresikan Bt toksin, pada **Bab 4,** dua jenis pendekatan digunakan untuk isolasi pod wall atau pulp yang diharapkan mempunyai transkrip gen spesifik. Random pustaka cDNA dari pod wall dan pulp di "screen" menggunakan reverse Northern blot untuk mengidentifikasi gen-gen yang memiliki ekspresi yang tinggi pada pod wall atau pulp. Selanjutnya Northern blot dengan RNA dari berbagai jaringan digunakan untuk mengetahui gen-gen yang mempunyai ekspresi spesifik. Diantara gen-gen yang walaupun diketahui mempunyai ekspresi yang kuat pada pod wall dan atau pulp, diantaranya ada juga yang mempunyai ekspresi spesifik pada daun dan juga biji. Pendekatan yang kedua adalah pustaka subtractive (SSH) cDNA dari dinding buah yang di substrak terhadap cDNA biji dan cDNA daun. Pendekatan ini menghasilkan beberapa hal yang menarik, diantaranya menghasilkan gen yang mempunyai homologous terhadap *Lea5* gen yang berasal dari kapas, dan mempunyai ekspresi pada pod wall mulai pada awal perkembangan buahan digunakan selanjutnya untuk mengisolasi promoter yang diuraikan pada **Bab 5.** Cloning menggunakan "genome walking" digunakan untuk mengisolasi 3411 bp genomic fragment dari kakao termasuk coding region dan 5' flanking region dan juga termasuk di dalamnya promoter. Untuk mengetahui ukuran minimum dan regulasi promoter , promoter dengan ukuran berbeda dari 5' flanking region di fusion dengan GUS sebagai reporter gen ditransformasi ke tanaman *Arabidopsis*. Dalam hal ini minimal fragmen (451 bp dari start codon) dari promoter yang diperlukan untuk dapat diukur aktivitasnya pada tanaman arabidopsis. Walaupun promoter ini pada arabidopsis juga aktif pada daun dan akar, ini menandakan belum semua regulasi elemen yang diperlukan atau yang aktif pada tanaman ini untuk spesifik ekspresi pada buah diketahui, oleh sebab itu masih memelukan data lebih banyak. Yang menarik dari *TcLea 5* yang termasuk kelompok protein (Late embryogenesis

association) mempunyai ekspresi yang kuat ebrio yang mengalami kekeringan dan dalam beberapa kasus dijumpai pada jaringan vegetatif yang mengalami kekeringan. Bagaimanapun seperti diuraikan pada bab 5 pada kakao dan juga pada tomat *Lea 5* homologus mempunyai ekspresi yang kuat pada dinding buah, dimana fungsi sebenarnya belum diketahui dengan jelas.

Untuk menghasilkan tanaman kakao transgenik kakao tahan terhadap insek diperlukan prosedur transformasi dan regenerasi yang effektif. Pada **Bab 6** beberapa prosedur dengan menggunakan berbagai jaringan yang berbeda, telah digunakan untuk menghasilkan embrio somatik sebagai bahan untuk transformasi dan regenerasi tanaman kakao. Somatik embrio yang dihasilkan dari embrio zigotik atau staminode dapat ditransformasi dengan menggunakan *Agrobacterium tumefaciens* yang membawa konstruk gfp (green fluorescent protein) menggunakan CaMV 35S promoter, yang menghasilkan secondary embriogenesis dan regenerasi tanaman. Tanaman yang menghasilkan ekspresi gfp dihasilkan dari somatik embrio yang berasal dari embrio zigotik, hal yang sama diharapkan dari somatik embrio yang berasal dari staminode.

Pada **Bab 7,** karakteristik awal dari 2 jenis gen kakao yang mengkode faktor transkripsi dari kelas MADS-box , yang diperkirakan mempengaruhi regulasi waktu pembungaan dan pembentukan bunga. *Apetala 1 (TcAP1)* Kakao gen homolog dari *Arabidopsis* berpengaruh dengan baik pada koregulasi pada inisiasi pembungaan, juga pada determinasi identitas dari 2 kelopak bunga bagian luar dari organ pembungaan seperti sepal dan petal. Keadaan yang sama pada pembungaan species tanaman yang lain, *TcAP1* hanya diekspresikan pada pembungaan dan dalam bunga predominant terekspressi pada sepal dan petal dan pada level yang rendah pada organ pembungaan yang lain. Kakao homolog dari *Agamous Arabidopsis* gene (*TcAG)* berpengaruh pada pembentukan kelopak bagian dalam dari organ pembungaan, staminode, stamen dan ovari. Seperti diketahui pada tanaman yang lain, *TcAG* hanya terekspresi pada pembungaan dan predominant pada staminode, stamen dan ovari. Kami menemukan *TcAG* juga terekspresi pada dinding buah kakao pada berbagai tingkat perkembangannya.

Akhirnya pada **Bab 8** diskusi umum tentang hasil keseluruhan yang diperoleh dalam thesis ini, dan menyajikan dalam pandangan yang ditinjau dari sudut pandang bioteknologi dan genomic kakao. Walaupun kemajuan dari pemuliaan tanaman kakao relatif lambat, dan adopsi dari tanaman transgenik di beberapa negara sedikit lambat, namun karena dipandang dari segi ekonomi kakao mempunyai arti yang cukup penting,

dan berdasarkan hasil yang positif dari segi keuntungan, transgenik kakao tahan PBK pada masa yang akan datang, akan diterima dan membuat arti yang besar bagi kebanyakan penduduk Indonesia.

Acknowledgements

First I would like to thank Ir. Basuki MSc and Dr. Darmono Taniwiryono, my former and present directors at the Biotechnology Research Institute for Estate Crops in Bogor, for giving me permission and support to leave temporarily my office for doing a PhD in Wageningen.

The research for this thesis was a part of the BIORIN (Biotechnology Research Indonesia-Netherlands) program with the financial aid from the Royal Netherlands Academy of Arts and Sciences (KNAW).

After more than 4 years, finally I came to the end of my PhD here in Plant Research International, Wageningen, the City of Life Science. In the beginning I was upset, feeling lonely and homesick. I never imagined that I could stay in Holland so long. However, this became possible because I had nice environment, helping people around and a very nice place to work and learn things, which helped me to survive and made me feel like at home. For this, I would like to express my sincere gratitude to all of the people who were involved during this period with me at scientific and personal level.

Dr Ruud de Maagd, my supervisor, contributed the most from the first day of my arrival to today and guided me during the whole period of my PhD. Without your sincere guidance and valuable advice, and critical comments, I wouldn't have seen this day. I learned a lot from you at every step of my PhD period, as your ability to understand and solve the problems is par excellence. I can never forget your knowledge, enthusiasm, and above all your patience. You taught me how to be brave, tough, and never give up, and your critical comments were very important for my experience as a PhD student. You are not only a good supervisor but also a good friend; you organized many dinners and parties with others friends from many countries. I will miss your leadership and enthusiasm. Many thanks for numero uno.

I wish to express my gratitude to my promoter Prof. Evert Jacobsen, with his support and his stories of his own experiences about research and life, who helped me to be strong and encouraged me constantly, especially in difficult times. He was always available for help whenever I needed.

My grateful thanks are due to Dr. Djoko Santoso, my Indonesian supervisor. He visited PRI several times and joined with me from the beginning until the last moment of my

PhD period. He contributed significantly throughout my research, with helpful discussions and advice until the last moment. He always supported me with constant motivation.

From my Bt group, Petra, who helped me in the lab in the beginning, helping me finding stuff in the institute, operating different equipment and so on. She was always ready to help, whenever needed. I learned a lot from her practical advice. We spent almost 2 years together, but you were always in my mind, even when you were not in PRI anymore. Thanks again. Jos, who always called me Tetty double U. After Petra, we had Mieke and Bert. Mieke, with her lovely children Roel and Nina, we shared a lot of stories together about work and life and you always said "Tetty do you need help, just tell me". She joins with me until the end of my stay in PRI. Bert helped me with purified toxin and immunological assays. Samir, who always has answers when we had a problem, especially with the computer, and who taught me about modification of *Bt* genes. Nice to have you in the group, you visited us several times, and I hope to meet you again in the future. Rumyana, good luck with your PhD and your lovely daughter Sofia. She is a very good Dutch teacher for you; soon you can speak Dutch very well. Tzanko, always serious, thank you too. Yien, I liked having you in our group, and after you finished we had Ellen, Tim and Liora, and from Colombia Jairo and his wife Sonia. During my stay, almost every year we had some Indonesian joining with the Bt group and often helping me doing a part of experiments: Eddy, Wiwin and Herti, who helped me with the tissue culture of cocoa, and finally Samanhudi, we had a very nice time. I felt like in my own lab in Bogor.

Among the colleagues from the Bt group my special thanks are due to Salvador Herrero. Salva, whom I always call my Fa-fa, I would like to thank for the care and support you gave me especially when I had a lot of problems about life and research. You always came to me and we shared feelings about many things from happiness to disappointments. I sincerely appreciate your constant motivation, and as you always said "who is the number 1?" We talked about a lot of stories, about research, about life; you always supported me from close by or from far away. I'm very happy I had the chance to have you not only as a friend but also as a brother. Having you around was very nice.

Maarten, my previous cluster leader, I would like to thank you for the care and support, you and your family gave me when my sister and brother passed away; thanks. From your group: Ralph (12) I thank for the care you always gave to me; Limei, Nikolay, Ashok, and Tania, also made a good environment for me.

Harro, who agrees with the functions of my magic oil, and his group: Cinzia and Ale, Per, Iris Altug, Iris Kappers, Sun, and Francel, always happy and smiling, even though your number is still a mystery. Ludo, always ready with his camera when I needed. Spinach Teun, also doing PhD in PRI. Kumkum, whom I always call Kumikumi, joined PRI two years back. I would like to thank you for sharing with me many things and spending a good time together. Last but not the least I have Rada from Harro's group, with whom I shared almost every day until 10 o'clock in PRI. We also enjoyed so many things together, during my stay in PRI. Rada we managed to life PRI almost in the same time and thank you for helping me take pictures of my lovely *Arabidopsis*.

Jules, for kindness, we were in the same lab for a couple of years. You are not only a friend for discussion about molecular biology, but also a good Dutch teacher; I would like to thank you for the care and supporting me with my work and for your advice. There were a lot of people in PRI who helped me and were kind to me: Jeroen always confusing between cocoa and banana, thanks a lot for your help with the bombardment, and Robert who helped me a lot with the tissue culture and took care of the pictures of my cultures. Arnoud Bovy, who always had the tomato fruit available, even in the winter time. Froukje, Adele, Hanny, Hetty and others ....

I would like to thank Ans, thank for you help, you always happy, smiling make a good atmosphere from the beginning.

I would like to thank Henk, Threes, and Yvonne, who always offer to help me when I needed chemicals.

Prof. Gerco Angenent, my present cluster leader, gave advice with the flowering system. Stefan helped me taking pictures and we discussed a lot, especially about *Arabidopsis*. Ralph thanks for your RNA of the green tomatoes. Jan Kodde and Michiel gave me the RNA of tomato seeds, and Marco, thanks for your protocols.

Andi and his group: Shital, Aarti, Fang and Wang who took a lot of pictures for me and Gert, thanks for helping me with tomato transformation. Raffa (26) and Asaphino, who also support me from close or far away. Nayelli, who helped me with designing my thesis cover, even though you are still busy with you lovely son, Olin.

I also share nice moments with the international friends from PRI, they have been visiting the institute for periods of a few weeks to several years, and with whom I shared times together, making life more easy: Xinying-Xinying we sharing with many things and had a good time together, Mazen, Vagner, Gao, Red life, Qi, Marina, Anita, Cahyo, Edwige,

Hong, Ziufu, Sarah, Rahim, Michia, Yuri, Fabricio, Alex, Kim, Ri, Isabella, Stefano, Agata, Remziye, Liliana, and many others, thanks for good friendships.

My great thanks also to Huub Löffler (BIORIN project leader) for his contribution and leadership to arrange and support the project making our BIORIN project running well. Together with Ence Darmo Jaya Supena, Reni Chaerani, Vivi Anggaini, Rudi Trijatmiko, and Sigit Purwantomo, my PhD student group, we spent a lot of time together, discussing about experiments and sharing many things: dinner, travelling, and so on. I really enjoyed your company. Sjaak van Heusden always helped me with the tomato fruit, and sometimes he brought tomatoes from the Plant Breeding's green house.

Without cocoa fruit most of this work couldn't have been performed. I would like to thank greatly my Indonesian friends Vivi, Reni, Putu, Wiratno, Drupadi Dillon, Murniati, Titi, Rani, Retno, Yurdi, Hadi, Dian, Cahyo, Ruanda and Muli, who were always most helpful in bringing fruits and plant material in time when I needed them. No one came, especially from Bogor, without cocoa fruit or flower or plant tissue in their luggage.

I would like to thank the Indonesian students, friends and their families for their warm friendship and kindness. My great appreciation goes to Sugardiman, Mulyanto, Budi,Yurdi, Marco, Peter, Hadi and their families, and semua yang tidak dapat disebutkan satu persatu.

My very special thanks go to my husband Yoharmus Syamsu, who always guaranteed that one day I will finish my PhD. Thanks for taking care of our lovely children Fauzan, Azi and Hafiz. Thanks for your continuous support, understanding, and your patience throughout my whole PhD period. My deepest gratitude goes to Yati, who always supported and helped my family with her love and patience during hard times, especially after my sister passed away. I would like to thank to my father and mother, for you it was always very difficult to accept my living alone so far from home. Thanks for your support, I was always happy because I knew you were there and prayed for me to have a good life. I never forget the support from my big families in Indonesia.

At the end of all of the praises, I would like to thank deeply Allah, the Lord of the Alameen, the Most Beneficent, the Most Merciful, for his blessing. Time to go back home to my family, and hope to see you again in the future.....

Tetty,

Wageningen, May 2005

Curriculum Vitae

Tetty Chaidamsari was born in Bukittinggi, in the West Sumatra province of the Indonesia on 24th of September 1957. She grew up in Makassar, the capital city of South Sulawesi. In 1976 after finishing high school, she started studying in the Department of Chemistry of Hasanuddin University for BSc degree. From 1980 till 1984 she continued her studies at the Bandung Institute of Technology (ITB) in Bandung. From 1985 till 1993 she was employed by the Rubber Research Institute in Medan. From 1993 till today she was reseacher at the Biotechnology Research Institute for Estate Crops in Bogor. In 1998 she received her MSc degree from Bogor Agriculture University (IPB). In February 2001 she came to the Netherlands, to work in a PhD project within the BIORIN program, funded by KNAW.

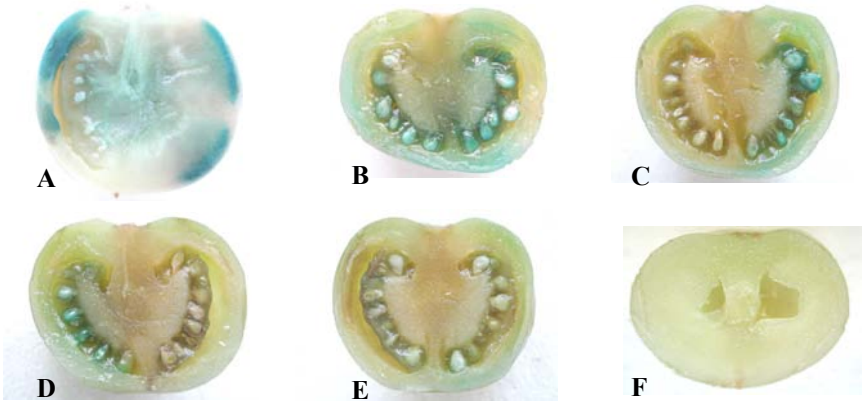
After finishing her PhD, her plan is to continue her work as a researcher at the Biotechnology Research Institute for Estate Crops in Bogor, Indonesia.

Training and Supervision Plan

The Graduate School of Experimental Plant Science

1) Start-up phase		*date*
►	**First presentation of your project**	19-11-2001
►	**Preliminary literature study of own research project**	2001/2002
►	**Research discussions**	2001 to 2005
	Subtotal Start-up phase	10.5
2) Scientific Exposure		*date*
►	**EPS PhD student days (2001 and 2004)**	
	Poster presentation, Utrecht	20-12-2001
	Poster presentation, Amsterdam	3-6-2004
►	**EPS theme symposia**	
	Theme 2, "Interactions between Plants and Biotic Agents"	12-12-2003
	Theme 1, "Developmental biology of plants"	17-2-2004
	Theme 4, "Genome Plasticity"	9-dec-04
►	**Seminars (series), workshops and symposia**	
	Seminar series "Biodiversity and Plant Breeding" at PRI	2001 to 2005
	Lecturer by M. Estelle. Prof. Molecular Genetic	22-10-2002
	Flying seminars 2002, 2003, 2004 (selected topics)	2002 to 2004
	Seminar series "Plant Cell Biology"(selected topics)	2002 to 2003
	Lecture by Prof. J Talbot Investigating the biology of plant infection by fungus	`29-4-2003
	Thematic meeting Business unit + Oral presentation	17-11-2003
	Symposium on System Biology in Honor of Prof.dr. Pierre de Wit	4-11-2004
►	**International conferences**	
	Society for Intervertebrate Pathology, Finland (Poster presentation)	1-6 Aug. 20'04
	Workshop Biotech.Res.Indonesia-Netherl.(BIORIN), Bogor + Oral pres.	Feb. 2002
	Workshop Biotech.Res.Indonesia-Netherl.(BIORIN), Wag. + Oral pres.	25-2-2003
	Workshop Biotech.Res.Indonesia-Netherl.(BIORIN), Wag. + Oral pres.	16-6-2004
►	**National meetings**	
	Lunteren Meeting(ALW)	7-8 April 2003
Subtotal Scientific Exposure		12.5
3) In-Depth Studies		*date*
►	**EPS courses or other PhD course**	
	Resistance mechanisms genomics disease resistance management	14-16 Oct 2002
►	**Journal club**	
	Cluster meeting discussions PRI	2001 to 2005
	Subtotal In-Depth Studies	3.0
4) Personal development		*date*
►	**Skill training courses**	
	English course	2003
	Techniques for writing and presenting a scientific paper	2004
	Presentation skills	2004
	Subtotal Personal Development	5.0
TOTAL NUMBER OF CREDIT POINTS		31

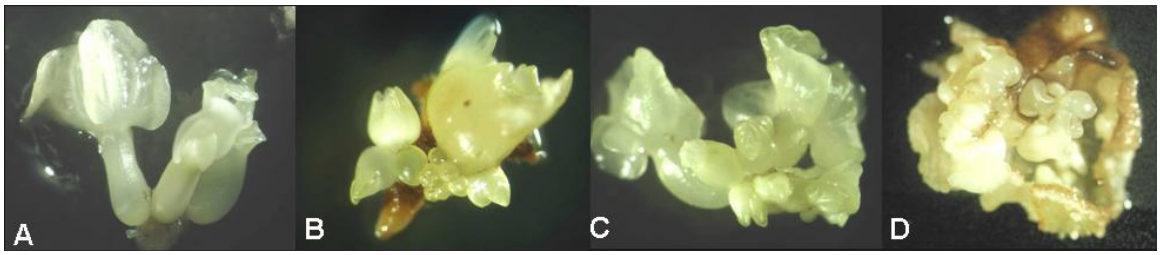
COVER AND INVITATION DESIGN BY NAYELLI MARSCH MARTINEZ



Chapter 5 Fig. 4 page 78

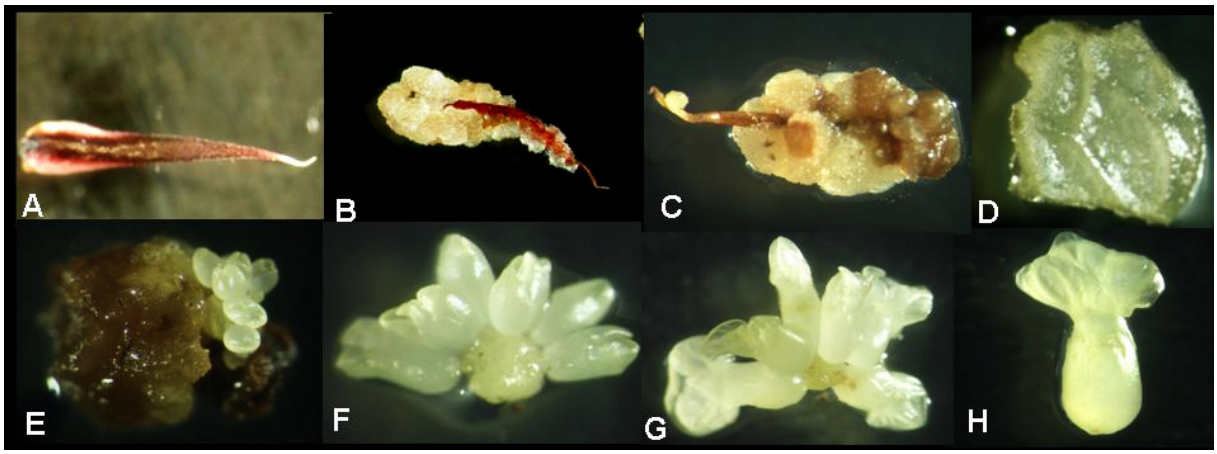
Chapter 5 Fig. 5 page 79

Chapter 5 Fig. 6C page 79



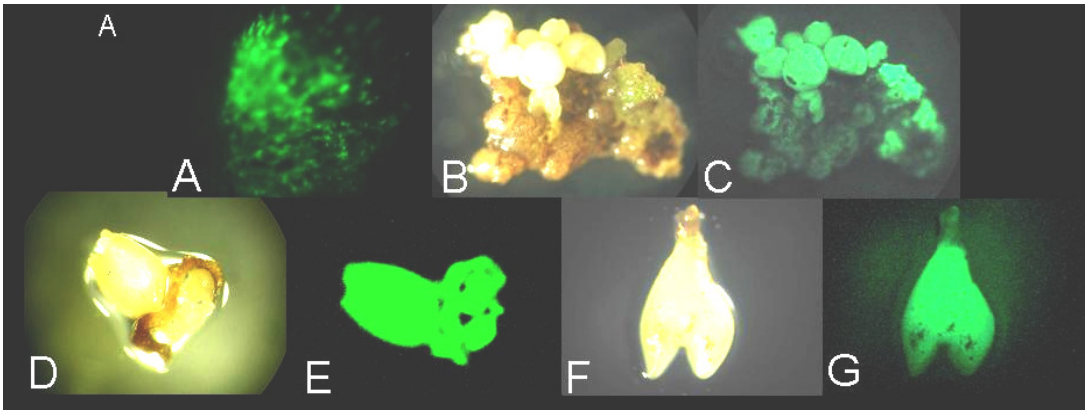
Chapter 6 Fig. 1 page 92

Chapter 6 Fig. 2 page 93

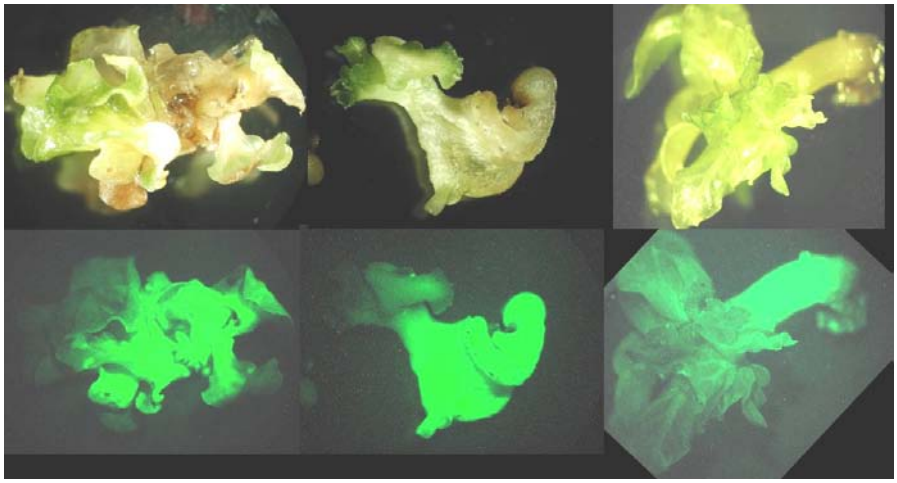


Chapter 6 Fig. 3 page 94

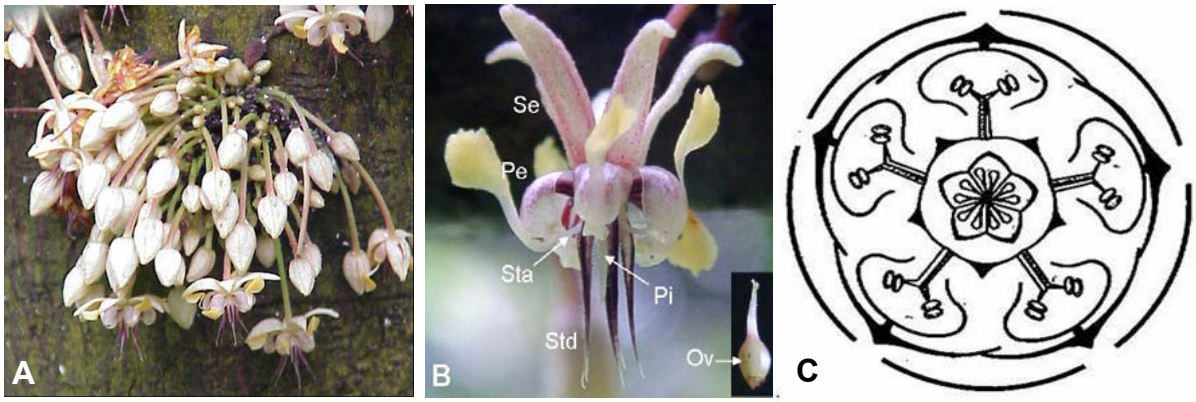
Chapter 6 Fig. 4 page 94



Chapter 6 Fig. 5 page 97



Chapter 6 Fig. 6 page 97



Chapter 7 Fig. 1 page 105